ALGÈBRE LINÉAIRE AVANCÉE I, SECTION DE PHYSIQUE DE L'EPFL

DONNA M. TESTERMAN

Date: December 7, 2024.

1.1. **Groupes, définitions et exemples.**

Définition 1.1.1. Un *groupe* est un ensemble muni d'une loi de composition $* : G \times G \to G$, qui envoie $(a, b) \mapsto a * b$, tel que les conditions suivantes sont vérifiées:

- (associativité de la loi de composition) $a * (b * c) = (a * b) * c$ pour tous $a, b, c \in G$.
- (existence de l'élément neutre à gauche) Il existe $e \in G$ tel que $e * g = g$ pour tout $g \in G$.
- (existence des inverses à gauche) Pour tout $g \in G$, il existe $g^* \in G$ tel que $g^* * g = e$.

Conséquences directes de la définition: Soit $(G, *)$ un groupe avec $e \in G$ l'élément neutre à gauche.

(1) Grâce à l'associativité, on écrira $a * b * c$ sans parenthèses sans que cela entraîne une ambiguïté quelconque.

(2) Soit $g \in G$. On montre que g^*, l'inverse à gauche de g, est aussi un inverse à droite: on a

$$(g^*)^* * g^* * g * g^* = (g^*)^* * e * g^* = (g^*)^* * g^* = e.$$

Aussi,

$$(g^*)^* * g^* * g * g^* = e * g * g^* = g * g^*.$$

On a donc l'égalité $g * g^* = e$, et g^* est un inverse à droite de g.

(3) L'élément e est également un élément neutre à droite: En effet, pour tout $g \in G$, $g * e = g * g^* * g = e * g = g$.

(4) L'élément e est unique: Si $f \in G$ satisfait aussi $f * g = g$ pour tout $g \in G$, on a $f * e = e$ car f est un élément neutre à gauche, et $f * e = f$ car e est un élément neutre à droite (cf point (3)). Donc $e = f$.

(5) L'élément g^* est unique. (à faire en exercice) On utilisera la notation g^{-1} pour désigner cet élément (l'inverse de g).

(6) Simplification, à gauche et à droite: pour tous $a, b, c \in G$, on a $a * b = a * c \Longrightarrow b = c$. En effet:

$$a * b = a * c \Longrightarrow a^{-1} * a * b = a^{-1} * a * c \Longrightarrow b = c.$$

De même,

$$a * b = c * b \Longrightarrow a = c.$$

(7) Pour tous $a, b \in G$, $(a * b)^{-1} = b^{-1} * a^{-1}$ et $(a^{-1})^{-1} = a$. (Exercice.)

Définition 1.1.2. 1. Un groupe $(G, *)$ dans lequel la loi de composition est commutative, c'est-à-dire que pour tous $a, b \in G$ on a $a * b = b * a$, est appelé *groupe abélien*. Si G n'est pas abélien, on dit que G est *non abélien*.

2. Soit G un groupe. On appelle *l'ordre de* G le cardinal de G, noté, comme pour les ensembles, $|G|$, ou $\mathrm{Card}(G)$.

Notation 1.1.3. Soit G un groupe.

1. On écrira souvent ab pour $a * b$ dans un groupe où la loi n'est pas précisée.
2. Si G est abélien, on utilisera parfois la notation $+$ pour désigner la loi de composition, 0 pour l'élément neutre et $-g$ pour l'inverse de $g \in G$.
3. Soient $g \in G$ et $m \in \mathbb{Z}$. On écrit g^m pour désigner

$$\begin{cases} g \cdots g, (m \text{ copies de } g), & \text{si } m > 0 \\ g^{-1} \cdots g^{-1}, (|m| \text{ copies de } g^{-1}), & \text{si } m < 0 \\ e & \text{si } m = 0 \end{cases}$$

On vérifie que pour tous $\ell, n \in \mathbb{Z}$, on a $g^\ell g^n = g^{\ell+n}$ et $(g^\ell)^n = g^{\ell n}$. Si le groupe est abélien et qu'on utilise la notation additive, on écrit mg et $-mg$ à la place de g^m et g^{-m}.

Exemples 1.1.4. Voici quelques exemples de groupes.

1. $(\mathbb{Z}, +)$
2. $(\mathbb{R} \setminus \{0\}, \cdot)$
3. Soit X un ensemble non vide. On note $\mathrm{Bij}(X) = \{f : X \to X, f \text{ est bijective}\}$, l'ensemble des applications bijectives de X dans X. On munit $\mathrm{Bij}(X)$ d'une loi de

composition $\circ : \mathrm{Bij}(X) \times \mathrm{Bij}(X) \to \mathrm{Bij}(X)$, qui est la composition d'applications et on vérifie que $(\mathrm{Bij}(X), \circ)$ est un groupe.

4. Cas particulier: prenons $X = \{1, 2 \ldots, n\}$ dans l'exemple précédent. On appelle $\mathrm{Bij}(X)$ le *groupe des permutations de l'ensemble* X, ou le *groupe symétrique de degré* n, et ce groupe est souvent noté $\mathfrak{S}_n$, Sym_n, ou simplement S_n. C'est un groupe d'ordre $n!$. On indique ici une des façons de représenter les éléments de S_n. On donne un tableau dans la première ligne duquel on trouve les entiers $1, 2, \ldots, n$ dans l'ordre croissant, et dont la deuxième ligne donne les images, dans l'ordre, de ces éléments par la permutation. Donc, pour $\sigma \in S_n$, on écrit

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix}.$$

Par exemple, si $n = 3$, le groupe S_3 comprend les 6 permutations :

$$\text{l'élément neutre } e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix},\ \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},\ \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$

$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},\ \sigma_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},\ \text{et } \sigma_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

On note que $\sigma_1\sigma_2 = \sigma_4 \neq \sigma_5 = \sigma_2\sigma_1$, et donc S_3 est un groupe non abélien.

5. Le groupe dit "trivial" qui consiste en un seul élément, $G = \{e\}$.

6. Les *entiers modulo* n:

Fixons un nombre naturel $n \in \mathbb{N}$, $n \neq 0$. On définit une relation sur $\mathbb{Z}$:

pour tous $a, b \in \mathbb{Z}$, on dit que $a \sim b$ si et seulement si n divise $b - a$, c'est-à-dire qu'il existe $m \in \mathbb{Z}$ tel que $b - a = nm$.

Assertion: $\sim$ est une relation d'équivalence sur $\mathbb{Z}$:

- (réflexive) Pour tout $a \in \mathbb{Z}$, $a \sim a$ puisque $a - a = 0 = n \cdot 0$.
- (symétrique) Pour $a, b \in \mathbb{Z}$, si $a \sim b$, on a $b - a = nm$ pour un certain $m \in \mathbb{Z}$, et donc $a - b = n(-m)$ et par conséquent $b \sim a$.

- (transitive) Pour $a, b, c \in \mathbb{Z}$, si $a \sim b$ et $b \sim c$, alors il existe $m, \ell \in \mathbb{Z}$ tel que $b - a = mn$ et $c - b = \ell n$, et donc $c - a = c - b + b - a = \ell n + mn = (\ell + m)n$ et $a \sim c$.

On note $\overline{a}$ la classe d'équivalence de a, c'est-à-dire $\overline{a} = \{b \in \mathbb{Z} \mid a \sim b\}$. Pour $b \in \mathbb{Z}$ tel que $a \sim b$, on écrit $a \equiv b \mod n$, on dit que a *est congru à* b *modulo* n, et on appelle la classe d'équivalence de a *la classe de congruence de* a *modulo* n. On écrit $\mathbb{Z}/n\mathbb{Z}$ pour désigner l'ensemble des classes d'équivalence de $\mathbb{Z}$ par rapport à cette relation.

On remarque que la notation $\overline{a}$ a un défaut, dans le sens où le n n'y apparaît pas. Si on souhaite travailler avec des classes de congruence modulo différents entiers, on utilisera la notation $[a]_n$ pour indiquer la classe de a modulo n.

Proposition 1.1.5. $\mathrm{Card}(\mathbb{Z}/n\mathbb{Z}) = n$

Preuve. Pour chaque entier $a \in \mathbb{Z}$, la division euclidienne de a par n donne l'existence de $q, r \in \mathbb{Z}$ avec $0 \leq r < n$ et $a = qn + r$. Donc n divise $a - r = qn$ et $a \sim r$. On déduit que $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$, et pour tous $0 \leq r < s < n$, on a $0 < s - r < n$, et donc $\overline{r} \neq \overline{s}$. □

Structure de groupe de $\mathbb{Z}/n\mathbb{Z}$: On munit $\mathbb{Z}/n\mathbb{Z}$ d'une loi de composition qui lui donnera la structure de groupe abélien.

Pour $a, b \in \mathbb{Z}$, on pose $\overline{a} + \overline{b} := \overline{a+b}$. Il faut montrer que l'association $(\overline{a}, \overline{b}) \mapsto \overline{a} + \overline{b} = \overline{a+b}$ définit une application de $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. Il est clair que $\overline{a+b} \in \mathbb{Z}/n\mathbb{Z}$.

Supposons que $\overline{a} = \overline{a'}$ et $\overline{b} = \overline{b'}$ pour $a, a', b, b' \in \mathbb{Z}$. Donc il existe $\ell, m \in \mathbb{Z}$ tels que $a - a' = \ell n$ et $b - b' = mn$. Alors $(a+b)-(a'+b') = (a-a')+(b-b') = \ell n + mn = (\ell+m)n$ et donc, $\overline{a+b} = \overline{a'+b'}$. L'application + est bien définie. Nous avons donc une loi de composition sur $\mathbb{Z}/n\mathbb{Z}$. On vérifie aisément l'associativité de la loi (qui suit directement de l'associativité de l'addition dans $\mathbb{Z}$). L'élément neutre est $\overline{0}$, et l'inverse de $\overline{a}$, pour $a \in \mathbb{Z}$, est l'élément $\overline{-a} = \overline{n-a}$.

1.2. **Sous-groupes.**

Définition 1.2.1. Soit G un groupe. Une partie H de G est un *sous-groupe de G* si la loi de composition restreinte à H munit H d'une structure de groupe. Si H est un sous-groupe de G, on écrit $H \leq G$.

Proposition 1.2.2. *Soit G un groupe et $H \subseteq G$ une partie de G. Alors H est un sous-groupe de G si et seulement si les conditions suivantes sont vérifiées:*

(i) *H est non vide.*

(ii) *Pour tous $h, k \in H$, on a $hk \in H$.*

(iii) *Pour tout $h \in H$, on a $h^{-1} \in H$.*

Preuve. Tout d'abord, on suppose que H est un sous-groupe de G. Comme H possède un élément neutre, H est non vide. Aussi, comme la loi de composition de G se restreint à une loi de composition sur H, pour tous $h, k \in H$, l'image de la loi $(h, k) \mapsto hk$ appartient à H. Maintenant, on montre que l'élément neutre de H, e_H, est égal à e, l'élément neutre de G. En effet, nous avons $e_H e = e_H = e_H e_H$ et par simplification à gauche on obtient que $e = e_H$. Maintenant, pour $h \in H$, soit $h^* \in H$ tel que $h^* h = e_H = e$. On a aussi $h^{-1}h = e$ et par simplification à droite, on obtient que $h^* = h^{-1}$ et $h^{-1} \in H$.

Maintenant, supposons que les conditions (i), (ii) et (iii) soient vérifiées. On montre que la restriction de la loi de composition sur G au sous-ensemble H munit H d'une structure de groupe. On a par (ii) que pour tous $h, k \in H$, $hk \in H$. Cela veut dire que l'image de la loi de composition $G \times G \to G$ de tout couple $(a, b) \in H \times H$ appartient à H. Donc la restriction définit bien une loi de composition sur H. La loi est associative car c'est déjà le cas dans G. Par (i), il existe $h \in H$. Par (iii), $h^{-1} \in H$ et par (ii), $hh^{-1} = e \in H$. Donc H possède un élément neutre, notamment e, l'élément neutre de G, et par (iii), H possède les inverses. □

Exemples 1.2.3. Voici quelques exemples de sous-groupes.

1. Soit G un groupe quelconque. Alors $H = \{e\}$ et $H = G$ sont des sous-groupes de G.
2. $\{1, -1\} \leq (\mathbb{R} \setminus \{0\}, \cdot)$.
3. $H = \{\text{entiers pairs}\} \leq (\mathbb{Z}, +)$.
4. Soit $G = \mathbb{Z}/6\mathbb{Z}$. Alors $H = \{\overline{0}, \overline{2}, \overline{4}\}$ est un sous-groupe. (Y en a-t-il d'autres?)

5. Soit $G = S_3$, le groupe symétrique de degré 3. Alors $H = \{e, \sigma_4, \sigma_5\}$ est un sous-groupe de G (voir la notation introduite dans l'exemple 1.1.4(4)).
6. Soit G un groupe. On pose $Z(G) := \{z \in G \mid za = az \text{ pour tout } a \in G\}$, le *centre de* G. On montre que $Z(G)$ est un sous-groupe de G. (exercice)

1.3. **Morphismes de groupes.**

Définition 1.3.1. (1) Soient $(G_1, \cdot), (G_2, *)$ des groupes. Un *homomorphisme* (ou simplement *morphisme*) *de groupes de* G_1 *dans* G_2 est une application $\phi : G_1 \to G_2$, telle que $\phi(x \cdot y) = \phi(x) * \phi(y)$ pour tous $x, y \in G_1$.

(2) Un *endomorphisme* d'un groupe G est un homomorphisme de groupes $\phi : G \to G$.

(3) Un *isomorphisme* entre deux groupes G_1 et G_2 est un homomorphisme de groupes bijectif.

(4) Un *automorphisme* d'un groupe G est un endomorphisme bijectif de G.

(5) S'il existe un isomorphisme de groupes $\phi : G_1 \to G_2$, on dit que G_1 *est isomorphe à* G_2 et on écrit $G_1 \cong G_2$.

Exemples 1.3.2. Voici quelques exemples de morphismes de groupes.

1. Soit G un groupe. Alors l'application identité $\mathrm{id} : G \to G$ est un morphisme de groupes, tout comme l'application $\phi : G \to \{e\}$.
2. Soit $H = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \right\} \leq S_3$. On définit $\phi : \mathbb{Z}/3\mathbb{Z} \to H$ par $\phi(\overline{0}) = e$, $\phi(\overline{1}) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ et $\phi(\overline{2}) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$. On vérifie que ϕ est un isomorphisme de groupes.
3. $\phi : \mathbb{Z} \to (\mathbb{R} \setminus \{0\}, \cdot)$ et $\phi(m) = (\sqrt{2})^m$.
4. $\phi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ et $\phi(m) = \overline{m}$.

Lemme 1.3.3. *Soit $\phi : G_1 \to G_2$ un morphisme de groupes de $(G_1, \cdot)$ dans $(G_2, *)$, et soit $a \in G_1$. Alors*

(i) $\phi(e_{G_1}) = e_{G_2}$*, où* e_{G_i} *est l'élément neutre de* G_i *pour* $i = 1, 2$*,*

(ii) $\phi(a^{-1}) = \phi(a)^{-1}$

(iii) Pour $n \in \mathbb{Z}$, $\phi(a^n) = \phi(a)^n$.

Preuve. (i) Pour tout $a \in G_1$, $e_{G_2} * \phi(a) = \phi(a) = \phi(e_{G_1} \cdot a) = \phi(e_{G_1}) * \phi(a)$. La simplification à droite donne le résultat. Les énoncés (ii) et (iii) sont laissés en exercice. □

1.3.1. *Morphismes et sous-groupes.*

Définition 1.3.4. Le *noyau* d'un homomorphisme de groupes $\phi : G_1 \to G_2$ est l'ensemble

$$\ker(\phi) = \{x \in G_1 \mid \phi(x) = e_{G_2}\}.$$

On désigne l'image de ϕ par $\mathrm{im}\,(\phi) = \{\phi(g) \mid g \in G_1\} \subseteq G_2$.

Proposition 1.3.5. *Soit $\phi : G_1 \to G_2$ un homomorphisme de groupes. On a*

(i) $\ker(\phi) \leq G_1$, *et*

(ii) $\mathrm{im}\,(\phi) \leq G_2$.

Preuve. (i) exercice.

(ii) Comme G_1 est non vide et ϕ est une application de G_1 dans G_2, $\mathrm{im}\,(\phi)$ est non vide. Soient maintenant $a, b \in \mathrm{im}\,(\phi)$, et $x, y \in G_1$ tels que $\phi(x) = a$ et $\phi(y) = b$. Alors $ab = \phi(x)\phi(y) = \phi(xy)$. Comme G_1 est un groupe, $xy \in G_1$ et $\phi(xy) \in \mathrm{im}\,(\phi)$. De même, on a que $x^{-1} \in G_1$ et, de ce fait, $a^{-1} = \phi(x)^{-1} = \phi(x^{-1}) \in \mathrm{im}\,(G_1)$, où la dernière égalité provient du Lemme 1.3.3. Par conséquent, $\mathrm{im}\,(\phi) \leq G_2$. □

Exemple 1.3.6.

Soit $n \in \mathbb{N}, n \geq 1$ et soit $\phi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ le morphisme défini par $\phi(a) = \overline{a}$. Alors $\ker(\phi) = \{a \in \mathbb{Z} \mid \overline{a} = \overline{0}\} = \{a \in \mathbb{Z} \mid n \text{ divise } a\} = n\mathbb{Z}$.

1.4. **Anneaux.**

Définition 1.4.1. Un *anneau unitaire* $(A, +, \cdot)$ est un ensemble muni de deux lois de composition $+$ et $\cdot$:

$$(\dagger) \qquad \begin{array}{rccc} + : & A \times A & \to & A \\ & (a,b) & \mapsto & a+b \end{array} \qquad \begin{array}{rccc} \cdot : & A \times A & \to & A \\ & (a,b) & \mapsto & a \cdot b \end{array}$$

satisfaisant les axiomes suivants :

(1) $(A, +)$ est un groupe abélien.

(2) $a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \forall a, b, c \in A.$ (associativité de $\cdot$)

(3) il existe $1_A \in A$ tel que $1_A \cdot a = a \cdot 1_A = a \forall a \in A.$ (élément neutre pour $\cdot$)

(4) $(a + b) \cdot c = (a \cdot c) + (b \cdot c) \quad \forall a, b, c \in A.$ (distributivité I)

(5) $a \cdot (b + c) = (a \cdot b) + (a \cdot c) \quad \forall a, b, c \in A.$ (distributivité II)

Cachée dans (†), la stabilité des lois de composition est une propriété essentielle d'un anneau.

Remarque 1.4.2. Soit $(A, +, \cdot)$ un anneau unitaire.

(1) On écrira souvent ab à la place de $a \cdot b$, et on parlera de la "multiplication" dans A.

(2) Un ensemble A muni de deux lois de composition $+$ et $\cdot$, qui satisfont les axiomes (1), (2), (4) et (5) s'appelle un *anneau*. Nous ne considérerons que les anneaux unitaires ici, et nous nous permettons de parler simplement des anneaux quand nous voulons dire anneaux unitaires.

(3) On écrit 0 pour l'élément neutre par rapport à $+$, et $-a$ pour l'inverse de a par rapport à $+$, pour tout $a \in A$. Pour l'opération $+$, on parlera de "l'addition" dans A.

(4) Les inverses multiplicatifs n'existent pas nécessairement. Pour $a \in A$, s'il existe $b \in A$ tel que $ab = 1_A = ba$, on dit que a est *inversible* et que b est *l'inverse de* a. Et souvent, on écrit $b = a^{-1}$.

(5) Si pour tous $a, b \in A$, on a $ab = ba$, alors on dit que A est un anneau *commutatif*.

(6) **Convention.** On suppose que $1_A \neq 0$, et donc que $A \neq \{0\}$.

Exemples 1.4.3.

- $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, où $+$ et $\cdot$ sont les opérations usuelles, sont des anneaux commutatifs.
- Soit $n\mathbb{Z} = \{nz : z \in \mathbb{Z}\}$ pour $n \in \mathbb{N}$ et $n \geq 1$ (avec la multiplication et l'addition usuelles de nombres réels). Alors, $(n\mathbb{Z}, +, \cdot)$ n'est pas un anneau pour $n \geq 2$, car il ne possède aucun élément neutre pour la multiplication.
- Soit E un ensemble non vide et $(A, +, \cdot)$ un anneau. On définit

$$\mathrm{App}(E, A) := \big\{ f \mid f \text{ est une application de } E \text{ vers } A \big\}$$

et les opérations

$$(f+g)(x) := f(x)+g(x), \qquad (f\cdot g)(x) := f(x)g(x).$$

Alors, $(\mathrm{App}(E,A),+,\cdot)$ est un anneau (non commutatif, si A est non commutatif).

- Soit $n \in \mathbb{N}$ et $n \geq 2$. On munit l'ensemble $\mathbb{Z}/n\mathbb{Z}$ des entiers modulo n d'une deuxième loi de composition: on associe à la paire $(\overline{a},\overline{b})$ l'élement $\overline{ab}$. Il faut vérifier que cette association définit bien une application de $\mathbb{Z}/n\mathbb{Z}\times\mathbb{Z}/n\mathbb{Z}$ dans $\mathbb{Z}/n\mathbb{Z}$, donnée par $(\overline{a},\overline{b})\mapsto\overline{ab}$. Il est clair que $\overline{ab}\in\mathbb{Z}/n\mathbb{Z}$. Maintenant, on suppose que pour $a,a',b,b'\in\mathbb{Z}$, on a $\overline{a}=\overline{a'}$ et $\overline{b}=\overline{b'}$, c'est-à-dire, qu'il existe $k,\ell\in\mathbb{Z}$ tels que $a'=a+nk$ et $b'=b+n\ell$. On considère $a'b'-ab=(a+nk)(b+n\ell)-ab=ab+n\ell a+nkb+n^2k\ell-ab=n(\ell a+kb+nk\ell)$. On déduit que n divise $a'b'-ab$, et par la définition de $\mathbb{Z}/n\mathbb{Z}$, on a $\overline{ab}=\overline{a'b'}$, ce qui montre que nous avons une application bien définie $\cdot:\mathbb{Z}/n\mathbb{Z}\times\mathbb{Z}/n\mathbb{Z}\to\mathbb{Z}/n\mathbb{Z}$ donnée par $(\overline{a},\overline{b})\mapsto\overline{a}\cdot\overline{b}:=\overline{ab}$. Les autres axiomes de la définition 1.4.1 sont faciles à vérifier. $(\mathbb{Z}/n\mathbb{Z},+,\cdot)$ est donc un anneau commutatif.

Lemme 1.4.4. *Soit $(A,+,\cdot)$ un anneau. Alors,*

(i) $0\cdot a=a\cdot 0=0$ *pour tout* $a\in A$,

(ii) $(-a)b=a(-b)=-(ab)$ *pour tous* $a,b\in A$,

(iii) $(-a)(-b)=ab$ *pour tous* $a,b\in A$.

Preuve.

(i) Par l'axiome de distributivité I,

$$0\cdot a=0\cdot a+0=0\cdot a+0\cdot a+(-(0\cdot a))=(0+0)\cdot a+(-(0\cdot a))=0\cdot a+(-(0\cdot a))=0.$$

De la même manière, on montre que $a\cdot 0=0$.

(ii) Par les axiomes de distributivité I+II et la partie (i) du lemme,

$$ab+(-a)b=(a+(-a))b=0\cdot b=0$$

et

$$ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0.$$

Maintenant, par l'unicité des inverses additifs, on a que $(-a)b = -(ab)$ et $a(-b) = -(ab)$.

(iii) Ici, on applique (ii) deux fois et on utilise le fait que $-(-a) = a$ pour obtenir $(-a)(-b) = -(a(-b)) = -(-(ab)) = ab$.

■

Le lemme 1.4.4 permet d'écrire $-ab$ sans ambiguïté.

Définition 1.4.5. Soient $(A, +, \cdot)$ et $(B, \oplus, \odot)$ deux anneaux, avec éléments neutres $1_A, 1_B$, respectivement, par rapport à $\cdot$, et $\odot$. Un *morphisme d'anneaux* est une application $f : A \to B$ telle que

$$f(a + b) = f(a) \oplus f(b), \qquad f(a \cdot b) = f(a) \odot f(b), \qquad \forall a, b \in A,$$

et

$$f(1_A) = 1_B.$$

Si de plus f est bijective, on dit que f est un *isomorphisme d'anneaux* et que les anneaux $(A, +, \cdot)$ et $(B, \oplus, \odot)$ sont isomorphes. On note $(A, +, \cdot) \cong (B, \oplus, \odot)$, ou plus simplement $A \cong B$ (en tant qu'anneaux).

Définition 1.4.6. Soit $(A, +, \cdot)$ un anneau et $U \subseteq A$. On dit que $(U, +, \cdot)$ est un *sous-anneau* de A si

(i) $(U, +)$ est un sous-groupe de $(A, +)$,

(ii) si $a, b \in U$ alors $a \cdot b \in U$,

(iii) L'élément neutre multiplicatif 1_A de A appartient à U.

Lemme 1.4.7. *Soient $(A, +, \cdot)$ un anneau et $U \subseteq A$. Alors, les assertions suivantes sont équivalentes:*

(i) $(U, +, \cdot)$ est un sous-anneau de $(A, +, \cdot)$

(ii) $1_A \in U$, et pour tous $a, b \in U$, on a $a - b \in U$ et $a \cdot b \in U$.

Preuve. (i) $\Rightarrow$ (ii) découle de la définition d'un sous-anneau.

(ii) $\Rightarrow$ (i) Comme $1_A \in U$, U est non vide et comme $1_A - 1_A = 0$, on a que $0 \in U$. Si $b \in U$ alors $-b = 0 - b \in U$. Si $a, b \in U$, alors $a + b = a - (-b) \in U$. Donc $(U, +)$ est un sous-groupe de $(A, +)$. Si $a, b \in U$, alors $a \cdot b \in U$ d'après (ii) et donc $(U, +, \cdot)$ est bien un sous-anneau de $(A, +, \cdot)$. ∎

1.5. **Corps, corps finis.** Un corps est un anneau unitaire commutatif dans lequel tout élément non nul est inversible par rapport à la loi de composition $\cdot$.

Définition 1.5.1. Un *corps* $(K, +, \cdot)$ est un anneau unitaire (avec l'élément neutre multiplicatif 1) avec $K \neq \{0\}$ tel que:

(i) K est commutatif, c'est-à-dire que pour tous $a, b \in K$, on a $a \cdot b = b \cdot a$.

(ii) pour tout $a \in K \setminus \{0\}$, il existe $a^{-1} \in K$ tel que $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

On remarque que $(K, +, \cdot)$ est un corps si et seulement si $(K, +)$ et $(K \setminus \{0\}, \cdot)$ sont des groupes abéliens et $(a + b) \cdot c = a \cdot c + b \cdot c$ pour tous $a, b, c \in K$.

Une liste de tous les axiomes d'un corps $(K, +, \cdot)$:

(1) $a + b \in K, \quad a \cdot b \in K \quad \forall a, b \in K.$ (stabilité)

(2) $a + b = b + a, \quad \forall a, b \in K.$ (commutativité+)

(3) $a + (b + c) = (a + b) + c, \quad \forall a, b, c \in K.$ (associativité+)

(4) Il existe $0 \in K$ tel que $0 + a = a$ pour tout $a \in K$. (élément neutre+)

(5) Pour tout $a \in K$, il existe $-a \in K$ tel que $a + (-a) = 0$. (inverse+)

(6) $a \cdot b = b \cdot a, \quad \forall a, b \in K.$ (commutativité$\cdot$)

(7) $a \cdot (b \cdot c) = (a \cdot b) \cdot c, \quad \forall a, b, c \in K.$ (associativité$\cdot$)

(8) Il existe $1 \in K$ tel que $1 \cdot a = a$ pour tout $a \in K$. (élément neutre$\cdot$)

(9) Pour tout $a \in K \setminus \{0\}$, il existe $a^{-1} \in K$ tel que $a \cdot a^{-1} = 1$. (inverse$\cdot$)

(10) $(a + b) \cdot c = a \cdot c + b \cdot c, \quad \forall a, b, c \in K.$ (distributivité I)

(11) $a \cdot (b + c) = a \cdot b + a \cdot c, \quad \forall a, b, c \in K.$ (distributivité II)

En fait, la commutativité de $\cdot$ implique que les deux lois de distributivité I et II sont équivalentes.

Exemples :

- $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, où $+$ et $\cdot$ sont les opérations usuelles, sont des corps.
- $(\mathbb{Z}, +, \cdot)$ n'est pas un corps parce qu'il n'y a pas d'inverse multiplicatif en général.

Un morphisme (isomorphisme) de corps est simplement un morphisme (isomorphisme) des anneaux sous-jacents.

Proposition 1.5.2. *Soit $n \in \mathbb{N}$, $n \geq 2$. Si n est premier, alors tout élément de $\mathbb{Z}/n\mathbb{Z}$ différent de $\overline{0}$ est inversible.*

Preuve. On suppose $n = p$ un nombre premier. Soit maintenant $a \in \mathbb{Z}$ tel que $\overline{a} \neq \overline{0}$. Alors p ne divise pas a et donc $\text{pgcd}(a, p) = 1$. Par l'identité de Bézout, il existe $c, d \in \mathbb{Z}$ tels que $ac + pd = 1$. Donc $ac - 1 = -pd$, p divise $ac - 1$, et on déduit que $ac \equiv 1 \mod p$. Par conséquent, $\overline{ac} = \overline{1}$ et $\overline{a}$ est inversible. ∎

Corollaire 1.5.3. *Soit $p \in \mathbb{N}$ un nombre premier. Alors l'anneau $\mathbb{Z}/p\mathbb{Z}$ est un corps.*

Notation 1.5.4. On écrit $\mathbb{F}_p$ pour désigner le corps fini $\mathbb{Z}/p\mathbb{Z}$.

1.6. **Le corps des nombres complexes.** Un *nombre complexe* est une paire ordonnée (couple) (x, y) où $x, y \in \mathbb{R}$. En définissant l'*unité imaginaire* i, on écrit

$$x + \mathrm{i}y$$

au lieu de (x, y). L'ensemble des nombres complexes se note

$$\mathbb{C} = \{x + \mathrm{i}y : \ x, y \in \mathbb{R}\}.$$

Quelques conventions : Soit $z = x + \mathrm{i}y \in \mathbb{C}$.

- On note $x = \mathsf{Re}(z)$ et on dit que x est la *partie réelle* de z.
- On note $y = \mathsf{Im}(z)$ et on dit que y est la *partie imaginaire* de z.
- Si $y = 0$, il est usuel d'identifier le nombre complexe z avec le nombre réel x, et on dit que z est réel. De plus, on n'écrit pas le terme i0.
- Si $x = 0$, on dit que z est *imaginaire pur* ou *totalement imaginaire*, et on n'écrit pas le terme 0 dans l'expression $0 + \mathrm{i}y$.
- Si $y = \pm 1$, on écrit $x \pm \mathrm{i}$ (au lieu de $x + \mathrm{i}(\pm 1)$).

On définit une loi + (addition des nombres complexes) et une loi $\cdot$ (multiplication des nombres complexes) sur $\mathbb{C}$:

$$\begin{aligned} + :&\quad \mathbb{C}\times\mathbb{C}\to\mathbb{C},\quad (x_1+\mathrm{i}y_1)+(x_2+\mathrm{i}y_2) := (x_1+x_2)+\mathrm{i}(y_1+y_2),\\ \cdot :&\quad \mathbb{C}\times\mathbb{C}\to\mathbb{C},\quad (x_1+\mathrm{i}y_1)\cdot(x_2+\mathrm{i}y_2) := (x_1x_2-y_1y_2)+\mathrm{i}(x_1y_2+y_1x_2). \end{aligned}$$

Exemple 1.6.1. On a

$$(1+\mathrm{i})+(-2+\mathrm{i}) = -1+2\mathrm{i}$$

et

$$(1+\mathrm{i})(-2+\mathrm{i}) = -3-\mathrm{i}.$$

Comme $(\mathbb{C},+)$ hérite des propriétés de $(\mathbb{R},+)$, on voit que $(\mathbb{C},+)$ est un groupe abélien. L'élément neutre est $0 = 0+0\mathrm{i}$, et l'inverse additif de $z = x+\mathrm{i}y \in \mathbb{C}$ est $-z := -x-\mathrm{i}y$. On définit la soustraction des nombres complexes par

$$z_1 - z_2 := z_1 + (-z_2) = (x_1-x_2)+\mathrm{i}(y_1-y_2).$$

C'est laborieux de vérifier directement les propriétés de la multiplication. Néanmoins, on peut montrer:

Théorème 1.6.2. *L'ensemble $\mathbb{C}$ muni des opérations + et $\cdot$ définies ci-dessus est un corps.*

On constate que $\mathrm{i}^2 = \mathrm{i}\cdot\mathrm{i} = -1$. Cela suffit pour retrouver la loi de multiplication :

$$\begin{aligned} (x_1+\mathrm{i}y_1)\cdot(x_2+\mathrm{i}y_2) &= x_1x_2+\mathrm{i}y_1x_2+\mathrm{i}x_1y_2+\mathrm{i}^2y_1y_2\\ &= x_1x_2+\mathrm{i}y_1x_2+\mathrm{i}x_1y_2-y_1y_2\\ &= (x_1x_2-y_1y_2)+\mathrm{i}(x_1y_2+y_1x_2). \end{aligned}$$

Remarque 1.6.3. Grâce au théorème précédent, on a que la multiplication est commutative, et on peut écrire aussi bien $x+\mathrm{i}y$ que $x+y\mathrm{i}$ car $\mathrm{i}y = (0+\mathrm{i})\cdot(y+\mathrm{i}0) = (y+\mathrm{i}0)\cdot(0+\mathrm{i})$

Définition 1.6.4. Le *conjugué* d'un nombre complexe $z = x+\mathrm{i}y$ est le nombre complexe $\overline{z}$ défini par $\overline{z} := x-\mathrm{i}y$.

Lemme 1.6.5. *Soient $z_1, z_2, z \in \mathbb{C}$. Alors,*

(i) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$

(ii) $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$

(iii) $\overline{\overline{z}} = z$

(iv) $\mathsf{Re}(z) = \frac{1}{2}(z + \overline{z})$

(v) $\mathsf{Im}(z) = \frac{1}{2\mathrm{i}}(z - \overline{z})$.

Preuve. Les preuves de ces propriétés sont des exercices faciles. ∎

Les parties (i)–(iii) du lemme 1.6.5 impliquent que la conjugaison est un isomorphisme du corps $(\mathbb{C}, +, \cdot)$ dans lui-même.[1]

Définition 1.6.6. Le *module* d'un nombre complexe $z = x + \mathrm{i}y$ est le nombre réel positif $|z|$ défini par $|z| := \sqrt{x^2 + y^2}$.

Lemme 1.6.7. *Soient $z_1, z_2, z \in \mathbb{C}$. Alors,*

(i) $z\overline{z} = |z|^2$

(ii) $z^{-1} = \frac{\overline{z}}{|z|^2} \quad (z \neq 0)$

(iii) $\overline{z^{-1}} = \overline{z}^{-1} \quad (z \neq 0)$

(iv) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$

(v) $|z_1 + z_2| \leq |z_1| + |z_2|$ *avec égalité si et seulement si il existe $\alpha \geq 0$ tel que $z_1 = \alpha z_2$ ou $z_2 = \alpha z_1$.*

Preuve. (i). $z\overline{z} = (x + \mathrm{i}y)(x - \mathrm{i}y) = x^2 + y^2 + \mathrm{i}(xy - yx) = x^2 + y^2 = |z|^2$.
(ii) découle de (i) et (iii) découle de (ii).
(iv). En utilisant le lemme 1.6.5 et la commutativité de la multiplication complexe, on obtient

$$|z_1 \cdot z_2|^2 = z_1 \cdot z_2 \cdot \overline{z_1 \cdot z_2} = z_1 \cdot z_2 \cdot \overline{z_1} \cdot \overline{z_2} = z_1 \cdot \overline{z_1} \cdot z_2 \cdot \overline{z_2} = |z_1|^2 \cdot |z_2|^2.$$

[1] Un isomorphisme d'une structure algébrique dans elle-même est dit *automorphisme*.

(v). L'inégalité découle de

$$\begin{aligned} |z_1+z_2|^2 &= (z_1+z_2)\overline{(z_1+z_2)} = (z_1+z_2)(\overline{z_1}+\overline{z_2}) \\ &= |z_1|^2 + z_2\overline{z_1} + z_1\overline{z_2} + |z_2|^2 = |z_1|^2 + 2\,\mathsf{Re}(z_1\overline{z_2}) + |z_2|^2 \\ &\leq |z_1|^2 + 2\,|z_1||\overline{z_2}| + |z_2|^2 = |z_1|^2 + 2\,|z_1||z_2| + |z_2|^2 = (|z_1|+|z_2|)^2. \end{aligned}$$

On a utilisé le fait que le module est toujours supérieur ou égal à la partie réelle. L'inégalité ci-dessus devient une égalité si $\mathsf{Re}(z_1\overline{z_2}) = |z_1\overline{z_2}|$, c-à-d si $\beta = z_1\overline{z_2}$ est réel positif. Si $z_2 = 0$, on a bien $z_2 = \alpha z_1$ avec $\alpha = 0$. Si $z_2 \neq 0$, alors

$$z_1\overline{z_2}z_2 = \beta z_2 \quad \Rightarrow \quad z_1 = \frac{\beta z_2}{|z_2|^2} = \alpha z_2 \text{ avec } \alpha = \frac{\beta}{|z_2|^2} \geq 0.$$

La réciproque est évidente. ∎

La division d'un nombre complexe z_1 par un nombre complexe $z_2 \neq 0$ est définie par $z_1/z_2 := z_1 z_2^{-1}$. D'après le lemme 1.6.7 (ii), on a

$$\frac{z_1}{z_2} = \frac{z_1\bar{z}_2}{|z_2|^2}.$$

Par exemple,

$$\frac{2+3\mathrm{i}}{1+\mathrm{i}} = \frac{(2+3\mathrm{i})(1-\mathrm{i})}{1+1} = \frac{5+\mathrm{i}}{2} = \frac{5}{2} + \frac{1}{2}\mathrm{i}.$$

1.6.1. *Plan complexe et forme polaire.* On note que nous n'utiliserons pas la matière de ce paragraphe ni du paragraphe §1.7 dans le cours d'algèbre linéaire ce semestre, mais ce sera repris dans les cours de physique et d'analyse.

Par définition, les nombres complexes $\mathbb{C}$ sont des couples de nombres réels. Pour cette raison, tout nombre complexe correspond uniquement à un vecteur dans le plan $\mathbb{R}^2$ (qu'on appelera ici *plan complexe*). La somme des nombres complexes correspond à la somme des vecteurs, et la conjugaison correspond à la réflexion par rapport à l'axe réel, voir figure 1.

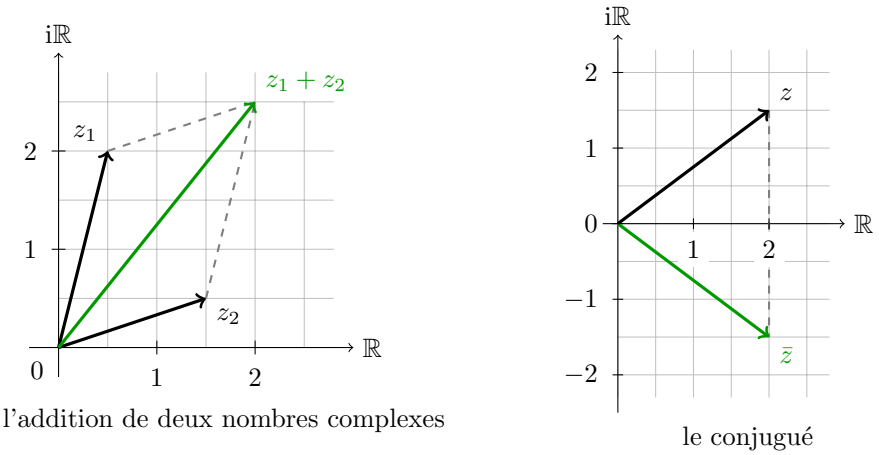

FIGURE 1. L'addition et le conjugué dans la plan complexe.

Soit $z = x + \mathrm{i}y \in \mathbb{C} \setminus \{0\}$. En notant $r = \sqrt{x^2 + y^2} > 0$ la longueur, et $\theta = \arctan \frac{y}{x} \in]-\pi, \pi]$ l'angle du vecteur (x, y) dans la plan complexe, on peut écrire

$$(x, y) = (r\cos\theta, r\sin\theta).$$

Ainsi, on a

$$z = x + \mathrm{i}y = r\cos\theta + \mathrm{i}r\sin\theta = r(\cos\theta + \mathrm{i}\sin\theta),$$

où θ est défini à $2k\pi$ près avec $k \in \mathbb{Z}$. On l'appelle la *forme polaire* de z. L'angle $\theta = \mathsf{arg}(z)$ est l'*argument* de z.

Par les identités trigonométriques, la forme polaire permet de multiplier facilement deux nombres complexes:

$$\begin{aligned} z_1 z_2 &= \rho_1(\cos\varphi_1 + \mathrm{i}\sin\varphi_1) \cdot \rho_2(\cos\varphi_2 + \mathrm{i}\sin\varphi_2) \\ &= \rho_1\rho_2\big(\cos(\varphi_1 + \varphi_2) + \mathrm{i}\sin(\varphi_1 + \varphi_2)\big). \end{aligned} \tag{1}$$

Alors, le produit $z_1 z_2$ représente géométriquement une multiplication de la longueur de z_1 par ρ_2 et une rotation anti-horaire de z_1 d'angle φ_2.

Lemme 1.6.8 (Formule de Moivre)**.** *Pour tous $r > 0$, $\theta \in \mathbb{R}$ et $n \in \mathbb{N}$, on a*

$$\big(r(\cos\theta + \mathrm{i}\sin\theta)\big)^n = r^n\big(\cos(n\theta) + \mathrm{i}\sin(n\theta)\big).$$

Preuve. Par récurrence utilisant (1). ∎

1.7. **La fonction exponentielle complexe.**

Définition 1.7.1. Pour $z = x + \mathrm{i}y \in \mathbb{C}$, on définit

$$e^z = \exp(z) := e^x(\cos y + \mathrm{i}\sin y) = e^{\mathsf{Re}z}(\cos(\mathsf{Im}z) + \mathrm{i}\sin(\mathsf{Im}z))$$

où e^x est la fonction exponentielle réelle usuelle.

Propriétés de l'exponentielle:

(1) $|e^z| = e^x = e^{\mathsf{Re}z}$

(2) $\mathsf{arg}(e^z) = \mathsf{Im}z$ (à $2k\pi$ près avec $k \in \mathbb{Z}$)

(3) si $\mathsf{Im}z = 0$, on a $e^z = e^{\mathsf{Re}z}$

(4) $e^{z+2k\pi\mathrm{i}} = e^x\big(\cos(y + 2k\pi) + \mathrm{i}\sin(y + 2k\pi)\big) = e^z$ pour tout $k \in \mathbb{Z}$

(5) $e^{w+z} = e^w \cdot e^z$ pour tous $w, z \in \mathbb{C}$

La *formule d'Euler* s'écrit, pour $\theta \in \mathbb{R}$,

$$e^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta.$$

En particulier pour $\theta = \pi$ on obtient *l'identité d'Euler*

$$e^{\mathrm{i}\pi} + 1 = 0.$$

1.8. **Anneaux de polynômes.** Soit $(A, +, \cdot)$ un anneau (nous utiliserons la juxtaposition pour indiquer la loi de composition $\cdot$ dans A).

Soit $A^{(\mathbb{N})}$ l'ensemble des suites ordonnées $(a_0, a_1, \ldots)$ d'éléments de A avec $a_i \neq 0$ pour un nombre fini de i, autrement dit

$$A^{(\mathbb{N})} := \{(a_0, a_1, \ldots) \mid a_i \in A \text{ pour tout } i \text{ et } \{i \geq 0 \mid a_i \neq 0\} \text{ est fini}\}.$$

On note que deux suites $(a_0, a_1, \ldots)$, $(b_0, b_1, \ldots)$ sont égales si et seulement si $a_i = b_i$ pour tout $i \geq 0$.

On définit deux lois de composition $\oplus$ et $*$ sur $A^{(\mathbb{N})}$:

$$(a_0, a_1, \ldots) \oplus (b_0, b_1, \ldots) := (a_0 + b_0, a_1 + b_1, \ldots);$$

$$(a_0, a_1, \ldots) * (b_0, b_1, \ldots) := (c_0, c_1, \ldots), \text{ où } c_m = \sum_{i+j=m} a_i b_j \text{ pour } m \geq 0.$$

Soient $M \in \mathbb{N}$ tel que $a_i = 0$ pour $i > M$, et $N \in \mathbb{N}$ tel que $b_j = 0$ pour $j > N$. Alors $a_m + b_m = 0$ pour $m > \max\{M, N\}$, et $c_m = 0$ pour $m > M + N$, car pour tous i, j avec $i + j = m > M + N$, soit on a $i > M$ soit on a $j > N$. On a donc soit $a_i = 0$, soit $b_j = 0$.

Par conséquent $(a_0, a_1, \ldots) \oplus (b_0, b_1, \ldots) \in A^{(\mathbb{N})}$ et $(a_0, a_1, \ldots) * (b_0, b_1, \ldots) \in A^{(\mathbb{N})}$.

On note que $(A^{(\mathbb{N})}, \oplus)$ est un groupe abélien avec élément neutre $0 = (0, 0, \ldots)$, l'inverse de $(a_0, a_1, \ldots)$ est $(-a_0, -a_1, \ldots)$, et l'associativité et la commutativité de $\oplus$ sont héritées de celles de A.

L'élément neutre pour $*$:

$$(1, 0, 0, \ldots) * (a_0, a_1, \ldots) = (c_0, c_1, \ldots) \text{ avec } c_m = \sum_{i+j=m} b_i a_j$$

où $b_0 = 1$ et $b_i = 0$ pour $i > 0$. Donc $c_m = b_0 a_m = a_m$ pour tout $m \geq 0$, d'où

$$(1, 0, 0, \ldots) * (a_0, a_1, \ldots) = (a_0, a_1, \ldots)$$

et de même pour $(a_0, a_1, \ldots) * (1, 0, 0, \ldots)$. L'élément neutre pour $*$ est alors $(1, 0, 0, \ldots)$.

On vérifie maintenant l'associativité de $*$:

Soient $a_i, b_i, d_i \in A$, et posons $c_\ell = \sum_{i+j=\ell} a_i b_j$ pour $\ell \geq 0$. Alors

$$((a_0, a_1, \ldots) * (b_0, b_1, \ldots)) * (d_0, d_1, \ldots) = (c_0, c_1, \ldots) * (d_0, d_1, \ldots) = (f_0, f_1, \ldots),$$

où

$$f_m = \sum_{\ell+k=m} c_\ell d_k = \sum_{\ell+k=m} \Big(\sum_{i+j=\ell} a_i b_j \Big) d_k = \sum_{i+j+k=m} (a_i b_j) d_k$$

$$= \sum_{i+j+k=m} a_i (b_j d_k) = \sum_{i+r=m} a_i \Big(\sum_{j+k=r} b_j d_k \Big).$$

On pose $(b_0, b_1, \ldots) * (d_0, d_1, \ldots) = (s_0, s_1, \ldots)$, alors $s_r = \sum_{j+k=r} b_j d_k$ et on déduit que

$$(a_0, a_1, \ldots) * ((b_0, b_1, \ldots) * (d_0, d_1, \ldots)) = (a_0, a_1, \ldots) * (s_0, s_1, \ldots) = (g_0, g_1, \ldots), \text{ où}$$

$g_m = \sum_{i+r=m} a_i s_r$ pour $m \geq 0$, c'est-à-dire $g_m = f_m$, ce qui établit l'associativité de $*$.

On laisse la vérification des propriétés de distributivité comme exercice.

Notation 1.8.1. On écrit

- $0 = (0, 0, \ldots)$
- $t := (0, 1, 0, 0, \ldots)$
- $1 = (1, 0, 0, \ldots)$
- pour $a \in A$, $a = (a, 0, 0, \ldots)$

On remplace $*$ par $\cdot$, ou simplement par la juxtaposition d'éléments, et $\oplus$ par $+$. Enfin, on désigne l'anneau $(A^{\mathbb{N}}, \oplus, *)$ par $A[t]$; on l'appelle *l'anneau des polynômes à coefficients dans* A (voir Proposition 1.8.2).

Conséquences de cette nouvelle notation:

(1) Pour $a \in A$, $(a, 0, 0, \ldots) = a \cdot 1 = a$.

(2) Pour $a \in A$, $(0, a, 0, 0, \ldots) = (a, 0, 0, \ldots) * (0, 1, 0, \ldots) = at$.

(3) Pour $m \in \mathbb{N}$, $m \geq 1$, $(0, 0, \ldots, 1, 0, 0, \ldots) = t^m$ où la valeur 1 est à la $(m+1)$-ième place.

(4) Pour $a_i \in A$, $i \in \mathbb{N}$ et $a_k = 0$ pour tout $k > m$, on a $(a_0, a_1, \ldots) = a_0 + a_1 t + a_2 t^2 + \cdots + a_m t^m$.

Proposition 1.8.2. *L'ensemble $A[t]$, avec $+$ et $\cdot$, est un anneau. Si A est commutatif, alors $A[t]$ est commutatif. L'application $\phi : A \to A[t]$ définie par $\phi(a) = (a, 0, 0, \ldots)$ est un morphisme d'anneaux injectif.*

Remarque 1.8.3. Etant donné le morphisme ϕ de la proposition précédente, on identifie A avec $\phi(A)$, et A devient un sous-anneau de l'anneau de polynômes $A[t]$.

Définition 1.8.4. Soit A un anneau.

(1) Soit $f = a_0 + a_1 t + \cdots + a_m t^m \in A[t]$ avec $a_m \neq 0$. On dit que f *est de degré* m, et on écrit $\deg(f) = m$. On pose $\deg(0) = -\infty$.

(2) Pour $f \in A[t]$, si $\deg(f) = m$ et $a_m = 1$, on dit que f est *unitaire*.

(3) Si $\deg(f) = 0$, ou si $f = 0$, on dit que f *est un polynôme constant*.

1.8.1. *Polynômes à coefficients dans un corps.* Si K est un corps, en particulier $(K, +, \cdot)$ est un anneau commutatif, on sait d'après la proposition 1.8.2 que l'ensemble des polynômes à coefficients dans K, muni de l'addition et de la multiplication des polynômes, est un anneau commutatif.

Définition 1.8.5. Soit K un corps, sous-anneau d'un anneau A. Soit $p \in K[t]$ avec $p(t) = a_0 + a_1 t + \cdots + a_n t^n$. *L'évaluation de* p en $s \in A$, notée $p(s)$, est l'élément de A suivant:

$$a_0 + a_1 \cdot s + \cdots + a_n \cdot s^n, \quad \text{où} \quad s^j := \underbrace{s \cdot s \cdots s}_{j \text{ fois}}.$$

On peut montrer (et ici on admettra) que l'application d'évaluation en c, $p \mapsto p(c)$ est un morphisme d'anneaux de $K[t]$ dans A.

Exemple 1.8.6.
Soient $K = \mathbb{R}$, $A = \mathbb{C}$, $p(t) = t^2 + 1 \in$
$\mathbb{R}[t]$.

$$\begin{aligned} p(\mathrm{i}) &= \mathrm{i}^2 + 1 = -1 + 1 = 0 \\ p(\mathrm{i}+1) &= (\mathrm{i}+1)^2 + 1 \\ &= \mathrm{i}^2 + 2\mathrm{i} + 1 + 1 = 2\mathrm{i} + 1 \end{aligned}$$

Définition 1.8.7. Soit K un corps, sous-anneau d'un anneau A. Un élément $c \in A$ s'appelle une *racine* de $p \in K[t]$ si $p(c) = 0$.

Théorème 1.8.8 (division euclidienne des polynômes)**.** *Soient $p, q \in K[t]$ avec $q \neq 0$. Alors, il existe un unique couple de polynômes $g, r \in K[t]$ tels que*

$$p = gq + r \qquad \textit{avec} \qquad \deg r < \deg q.$$

Preuve. (*Facultative*)

Existence. Par récurrence sur $n = \deg p$. Si $n < \deg q$, alors $g = 0$ et $r = p$ conviennent. (A noter que ceci inclut le cas $p = 0$.) Supposons le résultat montré pour tout polynôme de degré strictement inférieur à n et $n \geq m := \deg q$. On pose

$$p(t) = a_0 + a_1 t + \cdots + a_n t^n, \qquad q(t) = b_0 + b_1 t + \cdots b_m t^m.$$

Posons $f(t) = p(t) - a_n/b_m \cdot t^{n-m}q(t)$, alors $\deg f < n$. Par hypothèse de récurrence, on a

$$f = g_1 q + r \qquad \text{avec} \qquad \deg r < \deg q.$$

Alors,

$$p(t) = g_1(t)q(t) + r + \frac{a_n}{b_m} \cdot t^{n-m}q(t) = \big(\underbrace{g_1(t) + \frac{a_n}{b_m} \cdot t^{n-m}}_{=:g(t)} \big) q(t) + r,$$

où g, r possèdent les propriétés demandées.

Unicité. Si $p = g_1 q + r_1 = g_2 q + r_2$, alors

$$(g_1 - g_2)q = r_2 - r_1 \qquad \text{avec} \qquad \deg(r_2 - r_1) < \deg q.$$

Si $g_1 - g_2 \neq 0$,

$$\deg((g_1 - g_2)q) = \deg(g_1 - g_2) + \deg q \geq \deg q,$$

ce qui est absurde. Donc $g_1 = g_2$, et par suite $r_1 = r_2$. ∎

Corollaire 1.8.9. *Soient $p \in K[t]$ et $c \in K$. Alors c est une racine de p si et seulement si $t - c$ divise p (sans reste), c-à-d $p(t) = g(t)(t - c)$ pour un certain $g \in K[t]$.*

Preuve. L'assertion découle du théorème 1.8.8 en posant $q(t) = t - c$ et en utilisant le fait que l'évaluation en c est un morphisme d'anneaux.. ∎

Définition 1.8.10. On dit qu'un polynôme $p \in K[t]$ de degré $n \geq 1$ est *scindé* si

$$p(t) = \alpha(t - c_1)(t - c_2) \cdots (t - c_n), \qquad \alpha, c_1, \ldots, c_n \in K.$$

Théorème 1.8.11 (Théorème fondamental de l'algèbre)**.** *Tout polynôme à coefficients dans $\mathbb{C}$ est scindé.*

Preuve. Admis sans preuve. ∎

Remarque 1.8.12. La suite et la fin de ce chapitre seront utilisées au deuxième semestre, et par conséquent, ne sont pas couvertes au premier semestre.

Vocabulaire: Soient $p, q \in K[t]$ avec $q \neq 0$. On dit

- que q *divise* p,
- que q est un *diviseur* de p,
- que p est *divisible* par q, ou
- que p est un *multiple* de q,

si le reste de la division de p par q est nul.

Définition 1.8.13. Un polynôme $p \in K[t]$ est dit *irréductible* (sur K) si

(i) $\deg p \geq 1$

(ii) les seuls diviseurs de p sont les polynômes de degré 0 (les polynômes constants) et $c \cdot p(t)$ avec $c \in K \setminus \{0\}$.

Exemples :

(1) Tout polynôme de degré 1 est irréductible.

(2) Le polynôme $t^2 + 1 \in \mathbb{R}[t]$ est irréductible.

(3) Le polynôme $at^2 + bt + c \in \mathbb{R}[t]$ est irréductible si et seulement si $b^2 - 4ac < 0$.

Théorème 1.8.14. *Tout polynôme $p \in K[t]$ de degré ≥ 1 peut s'écrire de manière unique (à permutation des facteurs près)*

$$p = \alpha\, g_1 g_2 \cdots g_r \qquad \text{où} \qquad \alpha \in K \tag{2}$$

et g_i, $i = 1, \ldots, r$, sont des polynômes irréductibles unitaires.

Preuve. (*Facultative*)

Sans perte de généralité, on peut supposer que p soit unitaire.

Existence. Si p est irréductible, on obtient directement (2). Sinon, on peut écrire $p = p_1 p_2$, où p_1, p_2 sont des polynômes de degré strictement inférieur à $\deg p$. Ainsi, on obtient (2) par la récurrence.

Unicité. Soit $p = g_1 g_2 \cdots g_r = h_1 h_2 \cdots h_s$, où h_i, $i = 1, \ldots, s$, sont des polynômes irréductibles unitaires. Comme h_1 est irréductible, h_1 divise un des g_i. Mais, comme g_i est aussi irréductible, $h_1 = g_i$. Soit σ une permutation avec $\sigma(1) = i$. Alors,

$$\prod_{\substack{j=1 \\ j \neq \sigma(1)}}^{r} g_j = h_2 h_3 \cdots h_s.$$

En continuant de cette manière, on obtient $r = s$ et l'existence d'une permutation σ telle que $h_i = g_{\sigma(i)}$, $i = 1, \ldots, r$. ■

2.1. **Définitions, premières propriétés, exemples.**

Définition 2.1.1. Soit K un corps. Un *K-espace vectoriel* est un ensemble V muni de deux lois:

- une loi dite *interne* $V \times V \to V$,

  $(u, v) \mapsto u + v$,
- une loi dite *externe* $K \times V \to V$,

  $(\lambda, v) \mapsto \lambda \cdot v$ (ou simplement λv),

qui satisfont les conditions suivantes:

i) $(V, +)$ est un groupe abélien,

ii) pour tous $\lambda, \mu \in K$ et $v \in V$, $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$,

iii) pour tous $\lambda, \mu \in K$ et $v \in V$, $(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v)$,

iv) pour tous $\lambda \in K$ et $v, w \in V$, $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$, et

v) $1 \cdot v = v$ (ici $1 = 1_K$).

Les éléments de V s'appellent les *vecteurs* et les éléments de K s'appellent les *scalaires*. On parle de *l'addition* ou *la somme* pour $u+v$ et de la *multiplication par un scalaire* pour λv. On notera l'élément neutre du groupe $(V, +)$ par 0 (ou 0_V si nécessaire) et l'inverse de $v \in V$ par $-v$.

Exemples 2.1.2 (Quelques exemples)**.** 1. Soit V le produit cartésien $K \times \cdots \times K$ (n fois), qu'on notera $V = K^n$. On munit V d'une structure de groupe via

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n),$$

pour tous $x_i, y_i \in K$, $1 \leq i \leq n$. On définit une loi externe, la multiplication par un scalaire, comme suit:

pour tous $\lambda \in K$ et $(x_1, \ldots, x_n) \in V$, on a

$$\lambda \cdot (x_1, \ldots, x_n) := (\lambda x_1, \ldots, \lambda x_n).$$

On vérifie que V est un K-espace vectoriel.

2. Soit $K[t]$ l'anneau de polynômes à coefficients dans K. Alors $(K[t], +)$ est un groupe abélien et on définit une loi externe, la multiplication par un scalaire, comme suit :

   pour $\lambda \in K$ et $f \in K[t]$, $f(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$, on pose

$$(\lambda \cdot f)(t) := \lambda \cdot f(t) = \lambda a_n t^n + \lambda a_{n-1} t^{n-1} + \cdots + \lambda a_1 t + \lambda a_0$$

3. Soit X un ensemble et soit $\mathcal{F}(X, K)$ l'ensemble des applications de X dans K. Alors $\mathcal{F}(X, K)$ est un anneau. En particulier, $(\mathcal{F}(X, K), +)$ est un groupe abélien avec l'addition définie par $(f + g)(x) = f(x) + g(x)$, pour tous $f, g \in \mathcal{F}(X, K)$ et pour tout $x \in X$. On définit une loi externe, la multiplication par un scalaire, par :

   pour tout $\lambda \in K$ et pour tout $f \in \mathcal{F}(X, K)$, $(\lambda \cdot f)(x) = \lambda \cdot f(x)$, pour tout $x \in X$.

   On vérifie que $\mathcal{F}(X, K)$ est un K-espace vectoriel.

4. Une matrice $n \times m$ à coefficients dans K est un tableau à n lignes et m colonnes constitué d'éléments de K:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}.$$

   On appelle les a_{ij} les *composantes de la matrice* A ou les *coefficients de* A. Les indices (i, j) indiquent la ligne et la colonne à l'intersection desquelles se trouve la composante a_{ij}. On écrit $A = (a_{ij})$. On pose $M_{n\times m}(K)$ l'ensemble des matrices $n \times m$ à coefficients dans K et on définit une loi interne $+$ et une loi externe $\cdot$ comme suit: soient $A, B \in M_{n\times m}(K)$, avec $A = (a_{ij})$ et $B = (b_{ij})$ et soit encore $\lambda \in K$. On définit l'application

$$\begin{aligned} + : M_{n\times m}(K) \times M_{n\times m}(K) &\to M_{n\times m}(K) \\ (A, B) &\mapsto A + B, \end{aligned}$$

où $(A+B)_{ij} = a_{ij} + b_{ij}$, pour tous $1 \leq i \leq n$ et $1 \leq j \leq m$, et la multiplication par un scalaire:

$$\cdot : K \times M_{n\times m}(K) \to M_{n\times m}(K)$$

$$(\lambda, A) \mapsto \lambda \cdot A,$$

où $(\lambda \cdot A)_{ij} = \lambda \cdot a_{ij}$, pour tous $1 \leq i \leq n$ et $1 \leq j \leq m$.

Quelques conséquences directes de la Définition 2.1.1:

Proposition 2.1.3. *Soient $\lambda \in K$ et $v \in V$. On dénote par 0_V l'élément neutre de V et par 0 l'élément neutre par rapport à l'addition dans K, c'est-à-dire le scalaire $0 \in K$. On a*

a) $\lambda \cdot 0_V = 0_V$.

b) $0 \cdot v = 0_V$.

c) Si $\lambda \cdot v = 0_V$, alors soit $\lambda = 0$, soit $v = 0_V$.

d) $(-\lambda) \cdot v = \lambda \cdot (-v) = -(\lambda \cdot v)$.

Preuve. (a) $\lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V$. On simplifie à gauche (ce qui est possible dans un groupe) et on obtient $0_V = \lambda \cdot 0_V$.

(b) $0 \cdot v = (0+0) \cdot v = 0 \cdot v + 0 \cdot v$ et on conclut comme dans (a).

(c) Supposons $\lambda \cdot v = 0_V$ et que $\lambda \neq 0$. Alors il existe $\lambda^{-1} \in K$ (l'inverse multiplicatif) et on a

$$\lambda^{-1} \cdot (\lambda \cdot v) = (\lambda^{-1}\lambda) \cdot v = 1_K \cdot v = v.$$

Mais aussi $\lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot 0_V = 0_V$, par (a). Donc $v = 0_V$.

(d) On calcule $(-\lambda) \cdot v + \lambda \cdot v = (-\lambda + \lambda) \cdot v = 0 \cdot v = 0_V$, par (b). Par l'unicité des inverses dans un groupe, $(-\lambda) \cdot v = -(\lambda \cdot v)$. De façon similaire, on a $\lambda \cdot (-v) + \lambda \cdot v = \lambda \cdot (-v + v) = \lambda \cdot 0_V = 0_V$ et on conclut comme avant. □

2.2. **Sous-espaces vectoriels.** On fixe un corps K.

Définition 2.2.1. Soit V un K-espace vectoriel. Une partie W de V s'appelle un *sous-espace vectoriel* (ou simplement un *sous-espace*) *de* V si les restrictions des deux lois $+$ et $\cdot$ à W font de W un K-espace vectoriel.

Pour qu'une partie $W \subset V$ soit un sous-espace, il faut que

- W soit non vide,
- pour tous $w_1, w_2 \in W$ on a $w_1 + w_2 \in W$ et $-w_1 \in W$ (stable par $+$ et l'existence des inverses), et
- pour tous $\lambda \in K$ et $w \in W$, on a $\lambda \cdot w \in W$ (stable par la multiplication par un scalaire).

En fait, la proposition suivante démontre qu'on peut remplacer ces trois conditions par deux conditions et que ces deux conditions sont suffisantes pour assurer que W est un sous-espace.

Proposition 2.2.2. *Soit V un K-espace vectoriel, et soit W une partie de V. Alors W est un sous-espace de V si et seulement si*

1. *$W \neq \emptyset$, et*
2. *pour tout $\lambda \in K$ et pour tous $w_1, w_2 \in W$, on a $\lambda w_1 + w_2 \in W$.*

Preuve. La nécessité des conditions a été discutée ci-dessus.

Supposons que $W \subset V$ soit une partie non vide de V qui satisfait à la condition 2. Comme W est non vide, il existe $w \in W$. Prenons $\lambda = -1$ et appliquons la condition 2. On trouve $(-1) \cdot w + w = -w + w = 0_V \in W$. Donc W possède l'élément neutre. Aussi pour $w_1, w_2 \in W$, on a $1 \cdot w_1 + w_2 = w_1 + w_2 \in W$, et $-w_1 = (-1) \cdot w_1 + 0_V \in W$. Donc W est stable par l'addition et possède les inverses. La condition 2. implique que W est stable par multiplication par les scalaires (prendre $w_2 = 0_V$). Donc $(W, +)$ est un groupe abélien muni de la loi externe $K \times W \to W$, $(\lambda, w) \to \lambda \cdot w$. Les conditions ii), iii), iv) et v) dans la définition d'un K-espace vectoriel sont satisfaites, car elles le sont déjà dans V. □

Exemples 2.2.3. 1. Soit V un K-espace vectoriel, et soient $v_1, \ldots, v_r \in V$. Posons

$$\text{Vect}\,(v_1, \ldots, v_r) := \{\lambda_1 v_1 + \cdots + \lambda_r v_r \mid \lambda_i \in K \text{ pour } 1 \leq i \leq r\}.$$

On vérifie que $\text{Vect}\,(v_1, \ldots, v_r)$ est un sous-espace vectoriel de V, appelé *le sous-espace de V engendré par $v_1, \ldots, v_r$*.

2. Soit V un K-espace vectoriel. Soit $X \subseteq V$ une partie de V. On pose

$$\text{Vect}\,(X) = \bigcap_{W \in \mathcal{S}} W, \text{ où } \mathcal{S} = \{U \mid U \text{ un sous-espace de } V \text{ avec } X \subseteq U\},$$

l'intersection de tous les sous-espaces de V qui contiennent le sous-ensemble X. On l'appelle le *sous-espace engendré par X*. On vérifie que $\text{Vect}\,(X)$ est bien un sous-espace et dans les exercices vous montrerez que $\text{Vect}\,(X) = \{0\}$ si $X = \emptyset$ et si $X \neq \emptyset$ alors

$$\text{Vect}\,(X) := \{\lambda_1 v_1 + \cdots \lambda_r v_r \mid r \in \mathbb{N}, v_i \in X, \lambda_i \in K \text{ pour } 1 \leq i \leq r\}.$$

3. Soit $A = (a_{ij}) \in M_{m \times n}(K)$. On pose un système d'équations :

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\ \vdots & & & & \vdots & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & 0 \end{cases}$$

On dit que $(\alpha_1, \ldots, \alpha_n) \in K^n$ est *une solution du système* si pour tout $1 \leq t \leq m$ on a

$$a_{t1}\alpha_1 + \cdots + a_{tn}\alpha_n = 0.$$

L'ensemble des solutions du système forme un sous-espace vectoriel de K^n

D'autres exemples sont développés en cours.

2.3. **Comment former de nouveaux sous-espaces à partir d'autres sous-espaces.**

Définition 2.3.1. Soient W_1 et W_2 deux sous-espaces d'un K-espace vectoriel V. La *somme* $W_1 + W_2$ est l'ensemble $\{u + w \mid u \in W_1, w \in W_2\}$.

Lemme 2.3.2. *Soient W_1, W_2, V comme dans la définition précédente. Alors*

(a) *$W_1 + W_2$ est un sous-espace vectoriel de V.*

(b) *$W_1 \cap W_2$ est un sous-espace vectoriel de V.*

Preuve. Exercice. □

Définition 2.3.3. Soient W_1, W_2, V comme dans la définition précédente. On dit qu'un sous-espace U de V est la *somme directe de* W_1 *et* W_2 si

- $U = W_1 + W_2$, et
- $W_1 \cap W_2 = \{0\}$.

Dans le cas où U est la somme directe de W_1 et W_2, on écrit $U = W_1 \oplus W_2$.

On peut généraliser cette définition au cas de plus de deux sous-espaces:

Définition 2.3.4. Soient $W_1, \ldots, W_t$ des sous-espaces vectoriels d'un K-espace vectoriel V.

(a) On dénote par $W_1 + \cdots + W_t$ l'ensemble $\{x_1 + \cdots + x_t \mid x_i \in W_i,\ \forall 1 \leq i \leq t\}$. On écrit

$$\sum_{i=1}^{t} x_i = x_1 + \cdots + x_t.$$

(b) On dit qu'un sous-espace U de V est la *somme directe de* $W_1, \ldots, W_t$ si

- $U = W_1 + \cdots + W_t$, et
- $W_i \cap (\sum_{j \neq i} W_j) = \{0\}$ pour tout $1 \leq i \leq t$.

On démontre (par récurrence par exemple) que $W_1 + \cdots + W_t$ est un sous-espace vectoriel de V. Si U est la somme directe des W_i, $1 \leq i \leq t$, on écrit $U = W_1 \oplus \cdots \oplus W_t$.

Théorème 2.3.5. (Caractérisation des sommes directes) *Soient $W_1, \ldots, W_t, U$ des sous-espaces vectoriels d'un K-espace vectoriel V avec $W_i \subseteq U$ pour tout $1 \leq i \leq t$. Les conditions suivantes sont équivalentes:*

(1) $U = W_1 \oplus \cdots \oplus W_t$.

(2) *Chaque vecteur $u \in U$ s'écrit de façon unique comme $w_1 + \cdots + w_t$ avec $w_i \in W_i$ pour tout $1 \leq i \leq t$.*

Preuve. (1)$\Longrightarrow$(2): Chaque vecteur $u \in U$ s'écrit comme $u = w_1 + \cdots + w_t$ pour certains $w_i \in W_i$, $1 \leq i \leq t$, car $U = W_1 + \cdots + W_t$. Supposons que $u = w_1 + \cdots + w_t = y_1 + \cdots + y_t$, pour certains $w_i, y_i \in W_i$, $1 \leq i \leq t$. Alors pour chaque $1 \leq i \leq t$, on a $w_i - y_i = (y_1 - w_1) + \cdots + (y_{i-1} - w_{i-1}) + (y_{i+1} - w_{i+1}) + \cdots + (y_t - w_t)$. Ce vecteur

appartient à la fois à W_i et à la somme $\sum_{j\neq i} W_j$ et par la définition de la somme directe, l'intersection de ces deux sous-espaces est l'ensemble $\{0\}$. On déduit que $w_i = y_i$ pour tout $1 \leq i \leq t$ et par conséquent l'écriture est unique.

(2)$\Longrightarrow$(1): Comme $W_i \subseteq U$ pour tout i et que U est un sous-espace de V, on a $W_1 + \cdots + W_t \subseteq U$. Par hypothèse, pour tout $u \in U$, $u = w_1 + \cdots + w_t \in W_1 + \cdots + W_t$ et on déduit que $U = W_1 + \cdots + W_t$. Supposons que $x \in W_i \cap (\sum_{j\neq i} W_j)$. Alors d'une part, $x = 0 + 0 \cdots + x + 0 + \cdots + 0$, où le terme x est dans le sous-espace W_i, et d'autre part $x = w_1 + w_2 + \cdots + w_{i-1} + 0 + w_{i+1} + \cdots + w_t$ pour certains $w_j \in W_j$. Comme l'écriture est unique, on déduit que $x = 0$. □

On fixe un corps K et V un K-espace vectoriel.

3.1. **Dépendance et indépendance linéaire.**

Définition 3.1.1. a) Soient $v_1, \dots, v_r \in V$. Une *combinaison linéaire de* $v_1, \dots, v_r$ est un vecteur $v \in V$ de la forme $v = \lambda_1 v_1 + \cdots + \lambda_r v_r$, pour $\lambda_i \in K$. On dit que v *est une combinaison linéaire de* $v_1, \dots, v_r$ et que $\lambda_1, \dots, \lambda_r$ *sont les coefficients de la combinaison linéaire.*

b) Une partie X de V s'appelle un *système générateur*, ou une *partie génératrice* de V, si $V = \mathrm{Vect}\,(X)$. Si $X \neq \emptyset$, X est une partie génératrice de V si et seulement si tout $v \in V$ est une combinaison linéaire de vecteurs dans X.

Exemples 3.1.2. 1. Dans l'espace vectoriel $K[t]$, l'ensemble $\{1, t, t^2, \dots\}$ est un système générateur.

2. Dans l'espace vectoriel K^3, $\{(1,0,0),(0,1,0),(0,0,1)\}$ est un système générateur, où 1 désigne 1_K.

3. Dans l'espace vectoriel $\mathbb{R}^3$, $\{(1,1,0),(0,1,0),(0,0,2),(1,0,1)\}$ est un système générateur.

On souhaite préciser la notion d'une partie génératrice minimale.

Définition 3.1.3. (1) Soient $v_1, \dots, v_r \in V$. On dit que $v_1, \dots, v_r$ sont *liés*, ou *linéairement dépendants*, s'il existe $\lambda_1, \dots, \lambda_r \in K$, non tous nuls, avec $\lambda_1 v_1 + \cdots + \lambda_r v_r = 0$.

(2) Une partie X de V est dite *liée*, ou *linéairement dépendante*, s'il existe $v_1, \dots, v_t \in X$ distincts qui sont liés.

On a le critère utile suivant:

Proposition 3.1.4. *Soient $v_1, \dots, v_r \in V$, $r \geq 2$ vecteurs distincts. Alors $v_1, \dots, v_r$ sont linéairement dépendants si et seulement si l'un des v_i est une combinaison linéaire des autres.*

Preuve. Supposons que $v_1, \ldots, v_r$ sont linéairement dépendants. Alors il existe $\lambda_1, \ldots, \lambda_r \in K$, non tous nuls, tels que $\lambda_1 v_1 + \cdots + \lambda_r v_r = 0$. Supposons que $\lambda_i \neq 0$. On a

$$v_i = \lambda_i^{-1}(-\sum_{j \neq i} \lambda_j v_j).$$

En particulier v_i est une combinaison linéaire des v_j, $j \neq i$.

Réciproquement, si l'un des v_i est une combinaison linéaire des v_j, $j \neq i$, on a $v_i = \sum_{j \neq i} \alpha_j v_j$ pour des $\alpha_j \in K$ et on déduit que $v_i - \sum_{j \neq i} \alpha_j v_j = 0$, c'est-à-dire, les vecteurs $v_1, \ldots, v_r$ sont linéairement dépendants. □

Il y a une notion opposée à la dépendance linéaire:

Définition 3.1.5. (1) Soient $v_1, \ldots, v_r \in V$ distincts. On dit que $v_1, \ldots, v_r$ sont *libres*, ou *linéairement indépendants*, si $v_1, \ldots, v_r$ ne sont pas liés.

(2) Une partie X de V est dite *libre*, ou *linéairement indépendante*, si toute partie finie $\{v_1, \ldots, v_r\} \subseteq X$ de r vecteurs distincts est libre.

Remarque 3.1.6. Les r vecteurs distincts $v_1, \ldots, v_r$ sont libres si et seulement si toute égalité $\lambda_1 v_1 + \cdots + \lambda_r v_r = \mathbf{0}$ pour des scalaires $\lambda_i \in K$, implique que $\lambda_i = 0$ pour tout i.

Exemples 3.1.7. 1. L'ensemble vide $\emptyset \subseteq V$ est libre.

2. Dans l'espace vectoriel $K[t]$, la partie $\{1, t, t^2, \cdots\}$ est une partie libre.
3. Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$, l'ensemble $\{e^x, \sin x, x^2\}$ est un ensemble de vecteurs linéairement indépendants.
4. Soit F un corps avec $K \subseteq F$. Alors $(F, +)$ est un groupe abélien et aussi un K-espace vectoriel par rapport à la loi externe :

$$K \times F \to F,\ (\alpha, \beta) \mapsto \alpha\beta \in F.$$

Si on prend $K = \mathbb{R}$ et $F = \mathbb{C}$, alors les vecteurs $1, i \in \mathbb{C}$ sont linéairement indépendants.

3.2. **Base.**

Définition 3.2.1. Une partie X de V s'appelle une *base* si les deux conditions suivantes sont satisfaites:

- X est une partie génératrice de V, et
- X est une partie libre.

On démontre la proposition suivante.

Proposition 3.2.2. *Soit V un K-espace vectoriel, $V \neq \{\mathbf{0}\}$. Une partie X de V est une base de V si et seulement si pour tout $v \in V$, $v \neq \mathbf{0}$, il existe $x_1, \ldots, x_t \in X$ et $\lambda_1, \ldots, \lambda_t \in K \setminus \{0\}$, uniquement déterminés tels que $v = \lambda_1 x_1 + \cdots + \lambda_t x_t$.*

Preuve. Supposons que X est une base de V. Par définition, X est une partie génératrice et donc tout $v \in V$ est une combinaison linéaire de vecteurs dans X. On doit montrer l'unicité de l'expression. Supposons que pour $v \in V$, on a

$$v = \alpha_1 x_1 + \cdots \alpha_t x_t = \beta_1 y_1 + \cdots + \beta_r y_r,$$

où $\alpha_i, \beta_j \in K$ et $x_i, y_j \in X$. En rajoutant des termes à coefficients nuls, on peut supposer que

$$v = \gamma_1 u_1 + \cdots + \gamma_m u_m = \delta_1 u_1 + \cdots + \delta_m u_m,$$

où $\gamma_i, \delta_i \in K$ et $u_i \in X$ pour $1 \leq i \leq m$. On a donc que

$$\sum_{i=1}^{m} (\gamma_i - \delta_i) u_i = \mathbf{0}.$$

Mais comme X est libre cela implique que $\gamma_i = \delta_i$ pour tout i et donc les deux expressions sont identiques. En particulier, l'ensemble des u_i avec $\gamma_i \neq 0$ et l'ensemble des u_j avec $\delta_j \neq 0$ sont les mêmes. On déduit que l'ensemble $\{x_1, \ldots, x_t\}$ et les scalaires $\alpha_1, \ldots, \alpha_t$ sont uniquement déterminés.

Supposons maintenant que tout $v \in V$, $v \neq \mathbf{0}$ s'écrit de manière unique comme combinaison linéaire d'éléments de X. En particulier, X est un système générateur pour V et comme $V \neq \{\mathbf{0}\}$, X n'est pas vide et il existe $x \in X$, $x \neq \mathbf{0}$. On note aussi que le vecteur nul n'appartient pas à X car sinon pour $x \in X$, $x \neq \mathbf{0}$, on a $1 \cdot x = 1 \cdot x + 1 \cdot \mathbf{0}$, ce qui contredit l'unicité de l'expression.

Maintenant supposons que $\sum_{i=1}^{m} \alpha_i x_i = \mathbf{0}$ pour $\alpha_i \in K$ et $x_i \in X$. On suppose que $\alpha_i \neq 0$ pour tout i. On a donc

$$-\alpha_i x_i = \sum_{j=1, j\neq i}^{m} \alpha_j x_j$$

et par l'unicité de l'ecriture on déduit que ce vecteur est le vecteur nul. Mais cela veut dire que $\alpha_i x_i = \mathbf{0}$. Comme $\alpha_i \neq 0$ on déduit que $x_i = \mathbf{0}$, une contradiction. Donc tout α vaut 0 et X est libre. Cela montre que X est une base de V. □

Remarque 3.2.3. Le résultat précédent permet de compter les éléments d'un espace vectoriel sur un corps fini. Soit K un corps fini à p^a éléments (p un nombre premier, $a \in \mathbb{Z}, a \geq 1$). Si V est un K-espace vectoriel avec base $\{v_1, \ldots, v_m\}$ alors V possède exactement $(p^a)^m = p^{am}$ éléments.

Corollaire 3.2.4. *Supposons $\{f_1, \ldots, f_n\}$ est une base de V. Alors $V = \text{Vect}\,(f_1) \oplus \cdots \oplus \text{Vect}\,(f_n)$.*

Preuve. Exercice. □

Définition 3.2.5. Soit $B = \{v_1, \ldots, v_n\}$ une base de V. Les *composantes*, ou *coordonnées*, d'un vecteur $v \in V$ par rapport à la base B sont les $\lambda_i \in K$ tels que $v = \sum_{i=1}^{n} \lambda_i v_i$.

Définition 3.2.6. L'espace vectoriel V est dit *de dimension finie* si V possède un système générateur fini.

Théorème 3.2.7 (l'existence d'une base)**.** *Soit V un K-espace vectoriel de dimension finie, avec une partie génératrice finie $S \subseteq V$. Soit $L \subseteq S$ une partie libre. Alors il existe une base B de V avec $L \subseteq B \subseteq S$.*

Preuve. Si $V = \{\mathbf{0}\}$, alors $L = \emptyset$ et l'ensemble $B = \emptyset$ satisfait au résultat. On suppose maintenant que $V \neq \{\mathbf{0}\}$.

Soit B une partie libre incluse dans S, contenant L, et maximale sous les conditions d'être libre et de contenir L. Si $B = S$ alors B est libre et génératrice, donc une base de V.

Si $B \neq S$, soit $x \in S \setminus B$. Alors $B \cup \{x\}$ est plus grand que B et contient L. Par la maximalité de B, on a que $B \cup \{x\}$ n'est pas libre. Par conséquent, il existe $\lambda_x \in K$ et $\{\lambda_\gamma \mid \gamma \in B\} \subseteq K$ non tous nuls tels que

$$\lambda_x x + \sum_{\gamma \in B} \lambda_\gamma \gamma = 0.$$

Comme B est libre, $\lambda_x \neq 0$. On déduit que

$$x = -\lambda_x^{-1}(\sum_{\gamma \in B} \lambda_\gamma \gamma).$$

Par conséquent, tout élément dans $S \setminus B$ est une combinaison linéaire des éléments de B.

Soit maintenant $v \in V$. Comme S est une partie génératrice, $v = \sum_{s \in S} c_s s$ pour certains $c_s \in K$. Donc

$$v = \sum_{s \in B} c_s s + \sum_{s \in S \setminus B} c_s s.$$

Par l'argument qui précède, pour $s \in S \setminus B$, nous pouvons écrire $s = \sum_{b \in B} a_{sb} b$ et enfin, nous avons que

$$v = \sum_{s \in B} c_s s + \sum_{s \in S \setminus B} c_s (\sum_{b \in B} a_{sb} b),$$

une combinaison linéaire d'éléments de B. Donc B est une partie génératrice et libre de V, donc une base de V. □

Corollaire 3.2.8. *Soit V un K-espace vectoriel de dimension finie. Alors V possède une base finie.*

Preuve. Par Définition 3.2.6, il existe un système générateur fini, disons S. On prend $L = \emptyset$ dans le théorème précédent. □

Théorème 3.2.9. *Soit V un K-espace vectoriel de dimension infinie. Alors V possède une base.*

Nous admettrons ce résultat sans démonstration. La preuve demande la mise en place du lemme de Zorn, ce que nous ne ferons pas dans ce cours. Pour une preuve, vous pouvez consulter le polycopié de D. Kressner, §4.3.1, pages 61-63.

Théorème 3.2.10 (de la dépendance linéaire)**.** *Soit V un K-espace vectoriel de dimension finie. Soit $S = \{f_1, \ldots, f_n\}$ un système générateur (où $S = \emptyset$ si $n = 0$). Si $p > n$, tout ensemble de p vecteurs $\{v_1, \ldots, v_p\}$ est linéairement dépendant. Autrement dit, si $L = \{w_1, \ldots, w_q\}$ est une partie libre dans V, alors $q \leq n$*

Preuve. Soit p un entier avec $p > n$. On procède par récurrence sur n. Si $n = 0$, alors $V = \text{Vect}\,(\emptyset) = \{\mathbf{0}\}$ et tout ensemble non vide de vecteurs de V contient le vecteur nul et est donc linéairement dépendant. On traite aussi le cas $n = 1$: Ici, $V = \text{Vect}\,(f_1) = \{\lambda f_1 \mid \lambda \in K\}$, et $p > 1$. Prenons $v_1, v_2 \in V$, $v_i = \lambda_i f_1$, pour $\lambda_i \in K$, $i = 1, 2$. Si $v_1 = 0$, alors v_1 et v_2 sont linéairement dépendants. Si $v_1 \neq \mathbf{0}$, alors $\lambda_1 \neq \mathbf{0}$ et $\mathbf{0} = -\lambda_2\lambda_1^{-1}(\lambda_1 f_1) + \lambda_2 f_1 = -\lambda_2\lambda_1^{-1}v_1 + v_2$ et v_1 et v_2 sont linéairement dépendants.

Supposons maintenant que $n \geq 2$ et que le résultat est vérifié pour tout espace vectoriel W avec un système générateur de moins que n vecteurs.

Posons $W := \text{Vect}\,(f_1, \ldots, f_{n-1})$. Tout vecteur dans $v \in V$ s'écrit comme $v = w + \alpha f_n$, pour $w \in W$ et $\alpha \in K$. En particulier, nous avons

$$\begin{cases} v_1 &= w_1 + \alpha_1 f_n \\ v_2 &= w_2 + \alpha_2 f_n \\ \vdots & \vdots \\ v_p &= w_p + \alpha_p f_n, \end{cases}$$

pour $w_i \in W, \alpha_i \in K$, $1 \leq i \leq p$.

Si $\alpha_i = 0$ pour tout i, alors $v_i \in W$ pour tout i et par l'hypothèse de récurrence sur n, $\{v_1, \ldots, v_p\}$ est un ensemble de vecteurs linéairement dépendants de W et donc de V.

Supposons donc que $\alpha_1 \neq 0$ (sans perte de généralité). Dans ce cas, nous avons $f_n = \alpha_1^{-1}(v_1 - w_1)$. On trouve par substitution que $v_i = w_i + \alpha_i\alpha_1^{-1}(v_1 - w_1)$, pour $1 \leq i \leq p$, ce qui implique que

$$v_i - \alpha_i\alpha_1^{-1}v_1 = w_i - \alpha_i\alpha_1^{-1}w_1 \in W.$$

Nous avons les vecteurs

$$v_2 - \alpha_2\alpha_1^{-1}v_1, v_3 - \alpha_3\alpha_1^{-1}v_1, \ldots, v_p - \alpha_p\alpha_1^{-1}v_1,$$

qui appartiennent à $W = \text{Vect}\,(f_1, \ldots, f_{n-1})$. S'il y a des répétitions dans cette liste de vecteurs, c'est-à-dire si $v_j - \alpha_j\alpha_1^{-1}v_1 = v_k - \alpha_k\alpha_1^{-1}v_1$ pour $j \neq k$, alors on a $v_j - \alpha_j\alpha_1^{-1}v_1 - v_k + \alpha_3\alpha_1^{-1}v_1 = 0_V$ et les vecteurs $v_1, \ldots, v_p$ sont linéairement dépendants. Si ces vecteurs sont distincts, nous avons $p-1$ vecteurs dans W et $p-1 > n-1$. Par l'hypothèse de récurrence, ces vecteurs sont linéairement dépendants. Il existe $\lambda_2, \lambda_3, \ldots, \lambda_n \in K$, non tous nuls, tels que

$$\lambda_2(v_2 - \alpha_2\alpha_1^{-1}v_1) + \lambda_3(v_3 - \alpha_3\alpha_1^{-1}v_1) + \cdots + \lambda_p(v_p - \alpha_p\alpha_1^{-1}v_1) = \mathbf{0}.$$

Donc $-(\lambda_2\alpha_2\alpha_1^{-1} + \lambda_3\alpha_3\alpha_1^{-1} + \cdots + \lambda_p\alpha_p\alpha_1^{-1})v_1 + \sum_{j=2}^{p}\lambda_j v_j = \mathbf{0}$, une relation de dépendance. Donc $\{v_1, \ldots, v_p\}$ est un ensemble de vecteurs linéairement dépendants. □

3.3. **Dimension.**

Théorème 3.3.1 (de la dimension)**.** *Soit V un K-espace vectoriel de dimension finie. Alors toutes les bases de V sont finies et possèdent le même nombre d'éléments.*

Preuve. Soit S un système générateur fini (qui existe par la définition de dimension finie). Par le Thm. 3.2.7, il existe une base $B \subseteq S$ finie (prendre $L = \emptyset$ dans le théorème). Soit maintenant B' une autre base de V, donc un ensemble libre. Par le Thm. 3.2.10, $\text{Card}(B') \leq \text{Card}(B)$. En particulier, B' est finie. Pour compléter la preuve, on échange les rôles de B et B' pour obtenir que $\text{Card}(B) \leq \text{Card}(B')$. □

Définition 3.3.2. Le cardinal d'une base dans un K-espace vectoriel V de dimension finie s'appelle *la dimension de V* et se note $\dim(V)$.

Exemples 3.3.3. La dimension de l'espace $V = \{\mathbf{0}\}$ est 0. La dimension de K est égale à 1 et plus généralement, $\dim(K^n) = n$. L'espace vectoriel $K[t]$ n'est pas de dimension finie.

Proposition 3.3.4 (Critère de la dimension finie et infinie)**.** *Soit V un K-espace vectoriel.*

(a) *L'espace vectoriel V est de dimension infinie si et seulement si pour tout $n \in \mathbb{N}$, il existe une partie libre L de V de cardinal n.*

(b) *L'espace vectoriel V est de dimension finie si et seulement s'il existe $m \in \mathbb{N}$ tel que toute partie de V de cardinal m est liée.*

Preuve. Les deux affirmations étant équivalentes, il suffit de montrer (b).

Supposons que V est de dimension finie $\dim(V) = n$. Prenons $m = n + 1$. Par le Thm. 3.2.10, toute partie de m éléments est liée.

Maintenant supposons qu'il existe $m \in \mathbb{N}$ tel que toute partie de V de cardinal m est liée, et par conséquent, toute partie d'au moins m éléments est liée. Soit $L \subseteq V$ une partie libre maximale (c'est-à-dire qui n'est incluse dans aucune partie libre de V à part elle-même). On montre que $V = \text{Vect}\,(L)$. Par les remarques précédentes, $\text{card}(L) < m$. Soit $v \in L$, alors $v \in \text{Vect}\,(L)$. Prenons maintenant $v \in V \setminus L$. Posons $W = \text{Vect}\,(L \cup \{v\})$. L'espace W est un espace de dimension finie, car engendré par $L \cup \{v\}$. Par le Thm. 3.2.7, il existe une base B de W, avec $L \subseteq B \subseteq L \cup \{v\}$. Donc $W = \text{Vect}\,(B)$. Mais L maximal pour la propriété d'être libre implique $L = B$ et par conséquent $v \in \text{Vect}\,(L)$. Donc $V = \text{Vect}\,(L)$, comme voulu. □

Théorème 3.3.5 (complétion en une base)**.** *Soit V un K-espace vectoriel de dimension finie et soit L une partie libre de V.*

(a) *L'ensemble L est fini et* $\text{card}(L) \leq \dim V$.

(b) *L'ensemble L peut être complété en une base de V, c'est-à-dire, il existe une base B de V avec $L \subseteq B$.*

(c) *Si* $\text{Card}(L) = \dim V$, *alors L est une base de V.*

Preuve. Soit $\dim(V) = n$ et soit B une base de V.

(a) Par le Thm 3.2.10, si une partie libre L possède au moins m éléments distincts, alors $m \leq \text{Card}(B)$. Donc L est fini et $\text{Card}(L) \leq \text{Card}(B) = \dim V$.

(b) On a que $B \cup L$ est un système générateur de V, car B l'est. Aussi, $L \subseteq B \cup L$. Par le Thm. 3.2.7, il existe une base B' de V avec $L \subseteq B' \subseteq B \cup L$.

(c) Par (b), il existe une base B' avec $L \subseteq B'$. Mais $\text{Card}(B') = \dim V$ par le Thm. 3.3.1. Donc $\text{Card}(L) = \text{Card}(B')$ et $L = B'$ est une base de V. □

Théorème 3.3.6 (extraction d'une base)**.** *Soit V un K-espace vectoriel de dimension finie et $S \subseteq V$ un système générateur fini de V.*

(a) $\dim V \leq \text{card}(S)$.

(b) *On peut extraire de S un sous-ensemble qui est une base de V.*

(c) *Si* $\mathrm{Card}(S) = \dim V$*, alors* S *est une base de* V*.*

Preuve. Par le théorème de la dépendance linéaire, tout ensemble de vecteurs avec plus que $\mathrm{card}(S)$ vecteurs est lié. Donc une base possède au plus $\mathrm{card}(S)$ vecteurs, ce qui montre (a). L'affirmation de (b) se déduit du théorème de l'existence d'une base en prenant $L = \emptyset$. Enfin, supposons que $\mathrm{card}(S) = \dim V$. Par (b), il existe une base B de V avec $B \subseteq S$. Mais le théorème de la dimension montre que $\mathrm{card}(B) = \dim V$. Comme $\dim V = \mathrm{card}(S)$ par hypothèse, on a que $B = S$ et donc S est une base de V. □

Théorème 3.3.7 (des sous-espaces)**.** *Soit* V *un* K*-espace vectoriel de dimension finie, et soit* W *un sous-espace vectoriel de* V*.*

(a) *L'espace vectoriel* W *est de dimension finie.*

(b) $\dim W \leq \dim V$*.*

(c) *Toute base de* W *peut être complétée en une base de* V*.*

(d) *Si* $\dim W = \dim V$*, alors* $W = V$*.*

Preuve. (a) et (b) Posons $n = \dim V$. Alors par le théorème de la dépendance linéaire, toute partie de $n+1$ éléments de W est liée. Donc par le critère de la dimension finie, W est de dimension finie. Enfin, comme toute partie de $n+1$ éléments de W est liée, une base de W contient au plus n éléments et donc $\dim W \leq n$.

(c) Soit B' une base de W. On peut la compléter en une base B de V, $B' \subseteq B$, car B' est libre.

(d) Si $\dim W = \dim V$, on a $B' = B$ et on déduit que $\mathrm{Vect}\,(B) = \mathrm{Vect}\,(B')$ et $V = W$. □

Théorème 3.3.8 (du supplémentaire)**.** *Soit* V *un* K*-espace vectoriel de dimension finie et soit* W *un sous-espace vectoriel de* V*. Alors il existe un sous-espace* U *de* V *tel que* $V = W \oplus U$ *et* $\dim V = \dim W + \dim U$*.*

Preuve. Soit B' une base de W, qu'on complète en une base B de V, $B' = \{e_1, \ldots, e_m\}$ et $B = \{e_1, \ldots, e_m, e_{m+1}, \ldots, e_n\}$. Posons $U = \mathrm{Vect}\,(e_{m+1}, \ldots, e_n)$. Alors $U \cap W = \{\mathbf{0}\}$ et $U + W = V$. □

Définition 3.3.9. (1) Pour V, W et U comme ci-dessus, on appelle U un *supplémentaire de W dans V*.

(2) Pour V et W comme ci-dessus, si $\dim V = n$ et $\dim W = m$, alors on appelle $n - m$ *la codimension de W dans V* et on note $\mathrm{codim}_V(W) = n - m$.

Théorème 3.3.10 (formule des dimensions)**.** *Soient W_1 et W_2 deux sous-espaces vectoriels d'un K-espace vectoriel V. Supposons W_1 et W_2 de dimension finie. Alors*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Preuve. Soit $\{u_1, \ldots, u_t\}$ une base de $W_1 \cap W_2$. On la complète tout d'abord en une base $\{u_1, \ldots, u_t, w_1, \ldots, w_s\}$ de W_1 et aussi en une base $\{u_1, \ldots, u_t, v_1, \ldots, v_r\}$ de W_2. On montre que

$$S = \{u_1, \ldots, u_t, w_1, \ldots, w_s, v_1, \ldots, v_r\}$$

est une base de $W_1 + W_2$, et donc

$$\dim(W_1 + W_2) = t + s + r = (t + s) + (t + r) - t = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Système générateur: soit $v \in W_1 + W_2$, $v = x + y$, $x \in W_1$, $y \in W_2$. Alors $x = \sum \alpha_i u_i + \sum \beta_j w_j$ et $y = \sum \gamma_i u_i + \sum \delta_j v_j$ et $x + y \in \mathrm{Vect}\,(S)$.

Indépendance linéaire: Supposons que $\sum \alpha_i u_i + \sum \beta_j w_j + \sum \gamma_k v_k = \mathbf{0}$, pour certains $\alpha_i, \beta_j, \gamma_k \in K$. Posons $u = \sum \alpha_i u_i$, $w = \sum \beta_j w_j$ et $v = \sum \gamma_k v_k$, d'où $u + w + v = \mathbf{0}$. On a donc $u + w = -v$ et ce vecteur est à la fois un vecteur dans W_1 et aussi un vecteur dans W_2, donc $v \in W_1 \cap W_2$. On déduit que $v = \sum \delta_i u_i$. Mais cela entraîne que $\sum \delta_i u_i = \sum \gamma_i v_i$, et $\sum \delta_i u_i - \sum \gamma_i v_i = \mathbf{0}$. Par l'indépendance linéaire de $\{u_1, \ldots, u_t, v_1, \ldots, v_r\}$ on déduit que $\delta_i = 0 = \gamma_i$ pour tout i et donc $v = 0$.

Ensuite on a que $0 = v = u + w = \sum \alpha_i u_i + \sum \beta_i w_i$. Par l'indépendance linéaire de $\{u_1, \ldots, u_t, w_1, \ldots, w_s\}$, $\alpha_i = 0 = \beta_i$ pour tout i. Donc S est libre.

Comme nous avons montré que S est un système générateur, S est une base de $W_1 + W_2$. □

Corollaire 3.3.11. *Soient $W_1, \ldots, W_t$ des sous-espaces vectoriels d'un K-espace vectoriel V. Supposons W_i de dimension finie pour tout i et aussi $W_1 + \cdots + W_t = W_1 \oplus \cdots \oplus W_t$.*

Alors

$$\dim(W_1 + \cdots + W_t) = \sum_{i=1}^{t} \dim W_i.$$

On fixe un corps K. Soient V et W deux K-espaces vectoriels.

4.1. **Définitions, exemples.**

Définition 4.1.1. Une application $\varphi : V \to W$ est dite *K-linéaire* si

i) φ est un homomorphisme de groupes (pour l'addition), et

ii) $\varphi(\lambda v) = \lambda\varphi(v)$, pour tous $\lambda \in K$ et $v \in V$.

iii) Une application K-linéaire de V dans V est aussi appelée un *endomorphisme* de V.

iv) Une applicaiton K-linéaire de V dans W qui est bijective s'appelle un *isomorphisme de V dans W* ou bien *entre V et W*. Et s'il existe un isomorphisme, $\phi : V \to W$ on dit que V et W sont isomorphes.

S'il est clair que nous parlons du corps K, on dit simplement *une application linéaire*.

Donc $\varphi : V \to W$ est une application K-linéaire si et seulement si

i) $\varphi(u + v) = \varphi(u) + \varphi(v)$, pour tous $u, v \in V$, et

ii) $\varphi(\lambda v) = \lambda\varphi(v)$, pour tous $\lambda \in K$ et $v \in V$.

Ces deux conditions se résument en une seule:

(1) $\varphi(\lambda u + v) = \lambda\varphi(u) + \varphi(v)$, pour tous $\lambda \in K$ et $u, v \in V$.

On a les propriétés suivantes:

Proposition 4.1.2. *Soit $\varphi : V \to W$ une application K-linéaire. On note $\mathbf{0}_V$, respectivement $\mathbf{0}_W$, pour l'élément neutre de V, respectivement W.*

(1) On a $\varphi(\mathbf{0}_V) = \mathbf{0}_W$ (car φ est un morphisme de groupes).

(2) $\varphi(\lambda_1 v_1 + \cdots + \lambda_t v_t) = \lambda_1\varphi(v_1) + \cdots + \lambda_t\varphi(v_t)$, pour tous $\lambda_i \in K$ et $v_i \in V$ (par récurrence sur t).

Comme chaque élément de v s'écrit de manière unique comme une combinaison linéaire des éléments d'une base B de V, l'application linéaire φ est entièrement déterminée par ses images de la base B, et il suffit de connaître les valeurs de φ sur la base B pour connaître l'application φ.

On donne plusieurs exemples en cours. Mais un exemple qui en englobe plusieurs autres est le suivant:

Exemple 4.1.3. Soit $V = K^n$ et $W = K^m$. On fixe des scalaires a_{ij} pour $1 \leq i \leq m$ et $1 \leq j \leq n$. On définit une application $\varphi : V \to W$ par

$$\varphi(x_1, \ldots, x_n) = (\sum_{j=1}^{n} a_{1j}x_j, \ldots, \sum_{j=1}^{n} a_{mj}x_j),$$

pour $(x_1, \ldots, x_n) \in V$. C'est-à-dire, chaque coordonnée de $\varphi(v)$ est une combinaison linéaire des coordonnées de v. Vérifions que φ est bien K-linéaire.

Soient $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in V$ et $\lambda \in K$.

$$\begin{aligned}
\varphi(\lambda(x_1, \ldots, x_n) + (y_1, \ldots, y_n)) &= \varphi(\lambda x_1 + y_1, \ldots, \lambda x_n + y_n) \\
&= (\sum_{j=1}^{n} a_{1j}(\lambda x_j + y_j), \ldots, \sum_{j=1}^{n} a_{mj}(\lambda x_j + y_j)) \\
&= (\lambda \sum_{j=1}^{n} a_{1j}x_j + \sum_{j=1}^{n} a_{1j}y_j, \ldots, \lambda \sum_{j=1}^{n} a_{mj}x_j + \sum_{j=1}^{n} a_{mj}y_j) \\
&= \lambda(\sum_{j=1}^{n} a_{1j}x_j, \ldots, \sum_{j=1}^{n} a_{mj}x_j) + (\sum_{j=1}^{n} a_{1j}y_j, \ldots, \sum_{j=1}^{n} a_{mj}y_j) \\
&= \lambda\varphi(x_1, \ldots, x_n) + \varphi(y_1, \ldots, y_n).
\end{aligned}$$

Ceci montre que φ est bien K-linéaire.

4.2. **Opérations sur les applications K-linéaires.**

Définition 4.2.1. Soient V, W deux K-espaces vectoriels et $\varphi : V \to W$, $\psi : V \to W$ deux applications K-linéaires.

(1) On définit la *somme* de φ et ψ comme étant l'application $\varphi + \psi : V \to W$, $(\varphi + \psi)(v) = \varphi(v) + \psi(v)$. On vérifie que $\varphi + \psi$ est une application K-linéaire.

(2) Pour $\lambda \in K$, on définit l'application $\lambda \cdot \varphi : V \to W$ par $(\lambda \cdot \varphi)(v) = \lambda \cdot \varphi(v)$. On vérifie que $\lambda \cdot \varphi$ est une application K-linéaire. (Par la suite, on laissera tomber le 'point' entre le scalaire λ et l'application φ et on notera cette nouvelle application simplement par $\lambda\varphi$.)

(3) On dénote par $\mathcal{L}(V,W)$ l'ensemble des applications K-linéaire de V dans W.

(4) On utilisera aussi la notation $\mathrm{End}_K(V)$ pour désigner le K-espace vectoriel $\mathcal{L}(V,V)$.

On montre facilement:

Proposition 4.2.2. *Soient V et W deux K-espaces vectoriels. On munit $\mathcal{L}(V,W)$ de l'addition et la multiplication scalaire définies ci-dessus. Alors $\mathcal{L}(V,W)$ est un K-espace vectoriel.*

On peux aussi composer les applications linéaires:

Proposition 4.2.3. (1) *Soient $\varphi : V \to W$ et $\psi : W \to U$ deux applications K-linéaires de K-espaces vectoriels V, W, U. Alors l'application $\psi \circ \varphi : V \to U$ est une application K-linéaire.*

(2) *Si φ est bijective, alors l'application inverse, qu'on dénote par φ^{-1}, est une application K-linéaire $\varphi^{-1} : W \to V$.*

Preuve. Pour (1), on prend $\lambda \in K$ et $x, y \in V$. On a $(\psi \circ \varphi)(\lambda x + y) = \psi(\varphi(\lambda x + y)) = \psi(\lambda\varphi(x) + \varphi(y))$, car φ est K-linéaire. Et ensuite, $\psi(\lambda\varphi(x) + \varphi(y)) = \lambda\psi(\varphi(x)) + \psi(\varphi(y))$, car ψ est K-linéaire. Mais ce dernier est précisément $\lambda(\psi \circ \varphi)(x) + (\psi \circ \varphi)(y)$.

Pour (2), on se rappelle que l'application inverse φ^{-1} est définie comme suit: pour tout $w \in W$, il existe un unique $v \in V$ avec $\varphi(v) = w$. On pose $\varphi^{-1}(w) = v$. (C'est bien défini par la bijectivité de φ.)

Maintenant, soient $x, y \in W$ et $\lambda \in K$. Donc $x = \varphi(u)$ et $y = \varphi(v)$ pour certains $u, v \in V$. Alors

$$\varphi^{-1}(\lambda x + y) = \varphi^{-1}(\lambda\varphi(u) + \varphi(v)) = \varphi^{-1}(\varphi(\lambda u + v)),$$

car φ est K-linéaire. Ensuite, on a

$$\varphi^{-1}(\varphi(\lambda u + v)) = (\varphi^{-1} \circ \varphi)(\lambda u + v) = \lambda u + v = \lambda\varphi^{-1}(x) + \varphi^{-1}(y),$$

ce qui montre que φ^{-1} est K-linéaire. □

4.3. **Noyau, image et le théorème du rang.**

Définition 4.3.1. Soit $\varphi : V \to W$ une application K-linéaire. *L'image de* φ est l'ensemble

$$\mathrm{im}\,(\varphi) = \{w \in W \,|\, \exists v \in V \text{ avec } w = \varphi(v)\}.$$

On vérifie que $\mathrm{im}\,(\varphi)$ est un sous-espace vectoriel de W.

Définition 4.3.2. Si $\mathrm{im}\,(\varphi)$ est de dimension finie, alors $\dim(\mathrm{im}\,\varphi)$ s'appelle le *rang* de φ.

Définition 4.3.3 (Rappel)**.** Soit $\varphi : V \to W$ une application K-linéaire. Le *noyau de* φ est l'ensemble

$$\ker(\varphi) = \{v \in V \,|\, \varphi(v) = 0\}.$$

C'est exactement le même ensemble qu'on a défini pour un morphisme de groupes.

On vérifie aussi que $\ker(\varphi)$ est un sous-espace vectoriel de V (on sait déjà que c'est un sous-groupe du groupe $(V, +)$). La proposition suivante montre que le noyau de φ détermine si φ est injective.

Proposition 4.3.4 (Critère d'injectivité)**.** *Soit $\varphi : V \to W$ une application K-linéaire. Alors φ est injective si et seulement si* $\ker(\varphi) = \{0_V\}$.

Preuve. On suppose tout d'abord que φ est injective. Alors si $v \in \ker\varphi$, par définition, on a que $\varphi(v) = 0_W$. Mais comme $\varphi(0_V) = 0_W$ aussi, l'injectivité de φ implique que $v = 0_V$. On conclut que $\ker(\varphi) = \{0_V\}$.

Supposons maintenant que $\ker\varphi = \{0_V\}$. Soient $v, w \in V$ et supposons que $\varphi(v) = \varphi(w)$. On a

$$\varphi(v) - \varphi(w) = 0_W \text{ ce qui implique que } \varphi(v - w) = 0_W.$$

Par conséquent $v - w \in \ker\varphi$ et par hypothèse $v - w = 0_V$. On déduit que $v = w$ et φ est bien injective. □

On aimerait un critère aussi clair pour déterminer si une application K-linéaire est surjective. Comme $\mathrm{im}\,(\varphi)$ est un sous-espace vectoriel, si W est de dimension finie, φ est surjective si et seulement si $\dim(\mathrm{im}\,(\varphi)) = \mathrm{rang}(\varphi) = \dim W$. Le théorème suivant montre comment utiliser le noyau pour calculer le rang de φ.

Théorème 4.3.5 (Théorème du rang)**.** *Soit $\varphi : V \to W$ une application K-linéaire. Supposons V de dimension finie. Alors*

$$\dim(V) = \dim(\ker(\varphi)) + \dim(\mathrm{im}\,(\varphi)) = \dim(\ker(\varphi)) + \mathrm{rang}(\varphi).$$

Preuve. Comme V est de dimension finie, $\ker(\varphi)$ est aussi de dimension finie. On choisit une base $\{e_1, \ldots, e_m\}$ de $\ker(\varphi)$ et on la complète en une base B de V, $B = \{e_1, \ldots, e_m, e_{m+1}, \ldots, e_n\}$. On montre que $S = \{\varphi(e_{m+1}), \ldots, \varphi(e_n)\}$ est une base de $\mathrm{im}\,(\varphi)$, ce qui donne $\dim(\ker\varphi) + \dim(\mathrm{im}\,\varphi) = m + (n-m) = n = \dim V$.

Tout d'abord on montre que S est un ensemble libre: supposons que $\alpha_{m+1}\varphi(e_{m+1}) + \cdots + \alpha_n\varphi(e_n) = \mathbf{0}$, pour certains $\alpha_i \in K$. Par la linéarité de φ, on a que $\mathbf{0} = \varphi(\alpha_{m+1}e_{m+1} + \cdots + \alpha_n e_n)$. On déduit que $\alpha_{m+1}e_{m+1} + \cdots + \alpha_n e_n \in \ker(\varphi)$. On peut alors écrire $\alpha_{m+1}e_{m+1} + \cdots + \alpha_n e_n = \beta_1 e_1 + \cdots + \beta_m e_m$, pour certains $\beta_j \in K$. Mais ensuite on obtient que $\beta_1 e_1 + \cdots + \beta_m e_m - \alpha_{m+1}e_{m+1} - \cdots - \alpha_n e_n = \mathbf{0}$. Par l'indépendance linéaire de la base B, $\alpha_i = 0 = \beta_j$ pour tous i, j et donc S est libre.

On montre que S est un système générateur de $\mathrm{im}\,(\varphi)$. Soit $w \in \mathrm{im}\,(\varphi)$. Alors il existe $v \in V$ avec $\varphi(v) = w$. On écrit $v = \sum_{i=1}^n \lambda_i e_i$. Maintenant, on a $\varphi(v) = \varphi(\sum_{i=1}^n \lambda_i e_i) = \sum_{i=1}^n \lambda_i \varphi(e_i)$. Ce dernier est égal à $\sum_{i=m+1}^n \lambda_i \varphi(e_i)$, car $\varphi(e_j) = \mathbf{0}$ pour $j \leq m$. On a alors que

$$w = \varphi(v) = \sum_{i=m+1}^{n} \lambda_i \varphi(e_i) \in \mathrm{Vect}\,(\varphi(e_{m+1}), \ldots, \varphi(e_n)),$$

et $S = \{\varphi(e_{m+1}), \ldots, \varphi(e_n)\}$ est un système générateur de $\mathrm{im}\,(\varphi)$. □

Un corollaire direct de ce résultat est

Théorème 4.3.6 (Critère de bijectivité)**.** *Soit $\varphi : V \to W$ une application K-linéaire et supposons V de dimension finie.*

(1) Si φ est bijective, alors W est aussi de dimension finie et $\dim V = \dim W$.

(2) Si W est aussi de dimension finie et $\dim V = \dim W$, alors φ est bijective $\Leftrightarrow$ φ est injective $\Leftrightarrow$ φ est surjective.

Preuve. (Bon exercice à faire soi-même) □

On conclut avec un dernier exemple d'application linéaire qui nous sera utile par la suite:

Définition 4.3.7. Soit V un K-espace vectoriel avec sous-espaces U et W tels que $V = U \oplus W$. La *projection sur W le long de U* est l'application linéaire $\pi : V \to W$ définie par $\pi(u + w) = w$, pour tous $u \in U$ et $w \in W$.

Comme chaque $v \in V$ s'écrit de manière unique sous la forme $v = u + w$, cette application est bien définie et on vérifie que π est K-linéaire. De plus $\mathrm{im}\,(\pi) = W$ et $\ker(\pi) = U$.

Mais attention: comme il y a beaucoup de supplémentaires différents de W dans V, il y a beaucoup de projections différentes sur W, selon le supplémentaire choisi U.

4.4. **Le groupe linéaire général.** Posons $\mathrm{End}_K(V) = \mathcal{L}(V, V)$ (les endomorphismes de V). Alors $\mathrm{End}_K(V)$ est un sous-ensemble des applications de V dans V. Définissons également le groupe $\mathrm{Bij}(V)$, l'ensemble des applications bijectives de V dans V. On pose $\mathrm{GL}(V) = \mathrm{End}_K(V) \cap \mathrm{Bij}(V)$, les applications linéaires bijectives de V dans V. Alors la propositon 4.2.3 montre que $\mathrm{GL}(V)$ est un groupe avec comme opération binaire la composition d'applications. Ce groupe s'appelle le *groupe général linéaire sur V*. Nous considérons quelques sous-groupes dans les exemples suivants et dans les exercices.

Remarque 4.4.1. Il y a une confusion dans la littérature avec la notation $\mathrm{End}(V)$, car $(V, +)$ est un groupe abélien avec la loi de composition $+$. Dans ce contexte, $\mathrm{End}(V) = \{\phi : V \to V \mid \phi(u + v) = \phi(u) + \phi(v)\}$, l'ensemble des morphismes de groupe de V dans V. C'est pour cela que nous adopterons la notation $\mathrm{End}_K(V)$ pour l'ensemble $\mathcal{L}(V, V)$. Malheureusement, on voit aussi dans la littérature $\mathrm{End}(V)$ utilisé pour $\mathcal{L}(V, V)$.

Pour éviter cette confusion dans le cours, je vais utiliser la notation $\mathcal{L}(V, V)$ pour désigner l'ensemble des applications K-linéaires de V dans V.

On fixe un corps K.

5.1. **Algèbres des matrices.**

Définition 5.1.1. Soient $A \in M_{n\times m}(K)$ et $B \in M_{m\times \ell}(K)$. On définit le *produit* $A \cdot B$ (ou simplement AB) comme étant la matrice $C \in M_{n\times \ell}(K)$ telle que pour $i = 1, ..., n$ et $j = 1, ..., l$:

$$C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{im}B_{mj} = \sum_{k=1}^{m} A_{ik}B_{kj}.$$

(On multiplie successivement les coefficients le long de la i-ème ligne de A et la j-ème colonne de B et on additionne.)

Attention: Le produit AB n'est défini que si le nombre de colonnes de A est égal au nombre de lignes de B. En particulier, il arrive que AB soit défini et que BA ne soit pas défini.

Quelques premières propriétés découlent de cette définition:

(1) Associativité: Pour toutes $A \in M_{p\times q}(K)$, $B \in M_{q\times m}(K)$, et $C \in M_{m\times r}(K)$, on a $(AB)C = A(BC)$.

(2) Distributivité: Pour toutes $A, B, D \in M_{n\times m}(K)$ et $C, E, F \in M_{m\times \ell}(K)$, on a $(A + B)C = AC + BC$ et $D(E + F) = DE + DF$.

(3) Matrice identité: Soit $I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$, c'est-à-dire, I_n est la matrice $n \times n$ telle que

$$(I_n)_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon .} \end{cases}$$

Alors pour toute $A \in M_{n\times m}(K)$ et pour toute $B \in M_{m\times n}(K)$, on a $BI_n = B$ et $I_nA = A$.

Preuve. Pour (1), on compare les coefficients ij des matrices $A(BC)$ et $(AB)C$.

$$((AB)C)_{ij} = \sum_{k=1}^{m}(AB)_{ik}C_{kj} = \sum_{k=1}^{m}(\sum_{\ell=1}^{q} A_{i\ell}B_{\ell k})C_{kj} = \sum_{k=1}^{m}\sum_{\ell=1}^{q}(A_{i\ell}B_{\ell k})C_{kj}.$$

Et aussi

$$(A(BC))_{ij} = \sum_{\ell=1}^{q} A_{i\ell}(BC)_{\ell j} = \sum_{\ell=1}^{q} A_{i\ell}(\sum_{k=1}^{m} B_{\ell k}C_{kj}) = \sum_{\ell=1}^{q}\sum_{k=1}^{m} A_{i\ell}(B_{\ell k}C_{kj}).$$

On constate que les derniers termes des deux égalités sont identiques, ce qui établit (1). On raisonne de la même façon pour (2).

Pour (3), $(I_nB)_{ij} = \sum_{k=1}^{n}(I_n)_{ik}B_{kj} = (I_n)_{ii}B_{ij} = B_{ij}$ (la deuxième égalité découle du fait que $(I_n)_{ik} = 0$ si $k \neq i$). □

Notation 5.1.2. On écrit $M_n(K)$ pour l'ensemble des matrices $n \times n$ à coefficients dans K, c'est-à-dire, $M_n(K)$ désigne l'ensemble $M_{n\times n}(K)$.

On a déjà vu que $M_n(K)$ est un K-espace vectoriel, donc un groupe abélien pour l'addition. Avec les propriétés (1) à (3) ci-dessus, on a

Théorème 5.1.3. *Soit K un corps. Alors $M_n(K)$ est un anneau.*

Noter que cet anneau est non commutatif, avec des "diviseurs de zero"; c'est-à-dire, il existe $A, B \in M_n(K)$ non nulles avec $AB = 0$.

5.2. **Des matrices carrées particulières.**

Définition 5.2.1. (1) Une matrice $D \in M_n(K)$ est dite *diagonale* si $D_{ij} = 0$ à chaque fois que $i \neq j$, c'est-à-dire, D est une matrice carrée de la forme $\begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ 0 & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$,

où $\lambda_i \in K$ pour $1 \leq i \leq n$.

(2) Une matrice diagonale de la forme λI_n, pour $\lambda \in K$, s'appelle une *matrice scalaire*; c'est un multiple scalaire de la matrice identité.

(3) Une matrice $T \in M_n(K)$ telle que $T_{ij} = 0$ si $i > j$ s'appelle *une matrice triangulaire supérieure*. Une matrice $S \in M_n(K)$ telle que $S_{ij} = 0$ si $j > i$ s'appelle une matrice *triangulaire inférieure*.

(4) Une matrice $A \in M_n(K)$ est dite *inversible* s'il existe $B \in M_n(K)$ telle que $AB = I_n = BA$.

(5) On note $\mathrm{GL}_n(K) = \{A \in M_n(K) \mid A \text{ inversible}\}$ l'ensemble des matrices $n \times n$ inversibles.

Remarques: (1) Comme la multiplication de matrices est associative, si une matrice $A \in M_n(K)$ est inversible, elle possède un unique inverse qu'on note A^{-1}.

(2) Avec la multiplication de matrices, $\mathrm{GL}_n(K)$ est un groupe. En anglais, on dit 'the general linear group.'

5.3. **La matrice d'une application linéaire.**

Définition 5.3.1. Une *base ordonnée* d'un K-espace vectoriel V de dimension finie est un n-uplet $(f_1, \ldots, f_n)$ ordonné, c'est-à-dire un élément du produit cartésien $V \times \cdots \times V$ (n copies), tel que $\{f_1, \ldots, f_n\}$ soit une base de V.

Dans la suite de ce chapitre, tous nos espaces vectoriels seront de dimension finie et toutes nos bases seront des bases ordonnées.

Définition 5.3.2. Soit $\phi : V \to W$ une application K-linéaire. On fixe une base de V, $B_V = (e_1, \ldots, e_n)$, et une base de W, $B_W = (f_1, \ldots, f_m)$. On définit *la matrice de ϕ par rapport aux bases B_V et B_W*, notée $(\phi)_{B_V}^{B_W}$ comme suit:
On exprime $\phi(e_j)$ par rapport à la base B_W, $\phi(e_j) = a_{1j}f_1 + \cdots + a_{mj}f_m$. La matrice $(\phi)_{B_V}^{B_W}$ est la matrice $m \times n$ dont la j-ème colonne est $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$.
Si $V = W$ et on fixe une base B de V, on simplifie la notation $(\phi)_B^B$ en écrivant $(\phi)_B$.

Proposition 5.3.3. *Avec les notations de la Définition 5.3.2 et $\psi \in \mathcal{L}(V, W)$, on a*

(1) $(\phi + \psi)_{B_V}^{B_W} = (\phi)_{B_V}^{B_W} + (\psi)_{B_V}^{B_W}$.

(2) Pour tout $\lambda \in K$, $(\lambda\phi)_{B_V}^{B_W} = \lambda \cdot (\phi)_{B_V}^{B_W}$.

(3) L'application $\Theta : \mathcal{L}(V,W) \to M_{m\times n}(K)$ *définie par* $\Theta(\phi) = (\phi)_{B_V}^{B_W}$, *pour tout* $\phi \in \mathcal{L}(V,W)$, *est une application* K*-linéaire bijective.*

Preuve. Les propriétés (1) et (2) sont à montrer en exercices. Pour (3): L'application Θ est K-linéaire par les propriétés (1) et (2).

Pour la bijectivité, on détermine d'abord $\ker(\Theta)$. Pour $\phi \in \mathcal{L}(V,W)$, $\Theta(\phi) = \mathbf{0}_{m\times n}$ si et seulement si $\phi(e_i) = \mathbf{0}_W$ pour tout i, si et seulement si $\phi = 0$, l'application nulle. Donc $\ker(\Theta) = \{0\}$ et Θ est injective.

Pour la surjectivité, on prend $C \in M_{m\times n}(K)$, $C = (c_{ij})$, et on définit une application K-linéaire $\phi : V \to W$ comme suit

$$\phi(e_i) = \sum_{k=1}^{m} c_{ki} f_k,$$

pour tout i. On vérifie que ϕ est une application linéaire et $\Theta(\phi) = C$. □

Corollaire 5.3.4. $\dim \mathcal{L}(V,W) = \dim V \cdot \dim W$.

Définition 5.3.5. Soit B_V comme ci-dessus et $v \in V$. Si $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$, alors on pose

$$(v)_{B_V} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in M_{n\times 1}(K),$$

appelée la *matrice de* v *par rapport à la base* B_V. On dit aussi le *vecteur colonne de* v *par rapport à la base* B_V.

Remarque 5.3.6. On note que l'association $v \mapsto (v)_{B_V}$ définit une application K-linéaire bijective de V dans $M_{n\times 1}(K)$.

Théorème 5.3.7. *Soit* $\phi \in \mathcal{L}(V,W)$. *On fixe une base* B_V *de* V *et une base* B_W *de* W. *Alors pour tout* $v \in V$, *on a*

$$(\phi(v))_{B_W} = (\phi)_{B_V}^{B_W} \cdot (v)_{B_V} .$$

Noter que l'expression à droite de l'égalité est un produit matriciel.

Preuve. Fixons $B_V = (e_1, \ldots, e_n)$ et $B_W = (f_1, \ldots, f_m)$ et $A = (\phi)_{B_V}^{B_W}$. Pour $v \in V$, $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$, nous avons

$$\phi(v) = \phi(\lambda_1 e_1 + \cdots + \lambda_n e_n) = \sum_{i=1}^{n} \lambda_i \phi(e_i) = \sum_{i=1}^{n} \lambda_i (A_{1i} f_1 + \cdots + A_{mi} f_m)$$

$$= \sum_{i=1}^{n} \lambda_i (\sum_{j=1}^{m} A_{ji} f_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i A_{ji} f_j = \sum_{j=1}^{m} (\sum_{i=1}^{n} \lambda_i A_{ji}) f_j.$$

On déduit que le coefficient de f_j dans l'expression de $\phi(v)$ est égal à $\sum_{i=1}^{n} \lambda_i A_{ji} = \sum_{i=1}^{n} A_{ji} \lambda_i$. Mais ce dernier est précisément le j-ème terme du produit $(\phi)_{B_V}^{B_W} \cdot (v)_{B_V}$. Donc $(\phi)_{B_V}^{B_W} \cdot (v)_{B_V} = (\phi(v))_{B_W}$. □

Proposition 5.3.8. *Soient $A, B \in M_{m \times n}(K)$. Si $AX = BX$ pour tout $X \in M_{n \times 1}(K)$, alors $A = B$.*

Preuve. Par la Propriété 5.3.3(3), via la surjectivité de Θ, on sait que $A = (\phi)_{B_V}^{B_W}$ et $B = (\psi)_{B_V}^{B_W}$ pour certains $\phi, \psi \in \mathcal{L}(V, W)$, où V est un K-espace vectoriel de dimension n avec base B_V, et W est un K-espace vectoriel de dimension m avec base B_W. Soit $v \in V$ et posons $Y = (v)_{B_V}$. Par hypothèse, $AY = BY$. Donc

$$(\phi)_{B_V}^{B_W} \cdot (v)_{B_V} = (\psi)_{B_V}^{B_W} \cdot (v)_{B_V} \implies (\phi(v))_{B_W} = (\psi(v))_{B_W} \implies \phi(v) = \psi(v).$$

Comme ceci est vrai pour tout $v \in V$, on a que $\phi = \psi$ et par conséquent $A = \Theta(\phi) = \Theta(\psi) = B$. □

Théorème 5.3.9 (Matrice d'une composition)**.** *Soient $\phi : U \to V$ et $\psi : V \to W$ deux applications K-linéaires. Soient B_U, B_V et B_W des bases de U, V et W respectivement. Alors*

$$(\psi \circ \phi)_{B_U}^{B_W} = (\psi)_{B_V}^{B_W} \cdot (\phi)_{B_U}^{B_V} .$$

Preuve. Posons $A = (\phi)_{B_U}^{B_V}$, $B = (\psi)_{B_V}^{B_W}$ et $C = (\psi \circ \phi)_{B_U}^{B_W}$. Soit $X \in M_{n \times 1}(K)$ et posons $v \in U$ avec $(v)_{B_U} = X$.

On a $((\psi \circ \phi)(v))_{B_W} = (\psi \circ \phi)_{B_U}^{B_W} \cdot X = CX$.

Aussi $((\psi \circ \phi)(v))_{B_W} = (\psi(\phi(v)))_{B_W} = (\psi)_{B_V}^{B_W}(\phi(v))_{B_V} = (\psi)_{B_V}^{B_W} \cdot (\phi)_{B_U}^{B_V} \cdot (v)_{B_U} = BAX$. Donc $CX = BAX$ pour tout $X \in M_{n\times 1}(K)$. Par la Proposition 5.3.8, $C = BA$. □

Corollaire 5.3.10. *Soit V un K-espace vectoriel de dimension finie n. Soit $\Theta : \mathcal{L}(V,V) \to M_n(K)$ l'application bijective définie dans la Proposition 5.3.3. Alors Θ est un isomorphisme d'anneaux, c'est-à-dire, un homomorphisme d'anneaux bijectif.*

Preuve. Nous avons déjà montré que Θ est un morphisme de groupes abéliens pour l'addition et que Θ est bijective. Le Théorème 5.3.9 montre que $\Theta(\phi \circ \psi) = \Theta(\phi)\Theta(\psi)$. Enfin nous avons aussi que $\Theta(\mathrm{id}_V) = I_n$. Donc Θ est un isomorphisme d'anneaux. □

Corollaire 5.3.11 (Applications bijectives)**.** *Soit $\phi \in \mathcal{L}(V,W)$. Si ϕ est bijective, alors la matrice $(\phi)_{B_V}^{B_W}$ est inversible et*

$$(\phi^{-1})_{B_W}^{B_V} = ((\phi)_{B_V}^{B_W})^{-1} .$$

De plus, si $A \in M_n(K)$ est une matrice inversible, alors il existe $\psi \in \mathcal{L}(V,W)$ bijective telle que $A = (\psi)_{B_V}^{B_W}$.

Preuve. Exercice. □

Proposition 5.3.12. *Soit $\mathbb{F}_p$ le corps fini à p éléments. Alors le groupe $\mathrm{GL}_n(\mathbb{F}_p)$ est un groupe fini de cardinal $(p^n-1)(p^n-p)\cdots(p^n-p^{n-1})$.*

Preuve. Par le Corollaire 5.3.11, $\mathrm{GL}_n(\mathbb{F}_p)$ est en bijection avec l'ensemble des applications bijectives de $\mathbb{F}_p^n$ dans $\mathbb{F}_p^n$. Une application linéaire $\varphi : \mathbb{F}_p^n \to \mathbb{F}_p^n$ est une bijection si et seulement si l'image d'une base de $\mathbb{F}_p^n$ par φ est de nouveau une base de $\mathbb{F}_p^n$. De plus, φ est déterminée par l'image d'une base fixée. On fixe une base ordonnée $(e_1,\ldots,e_n)$ de $\mathbb{F}_p^n$ et on dénombre les images possibles, ce qui nous donnera le nombre d'applications linéaires bijectives distinctes.

L'image de e_1 par φ peut être n'importe quel vecteur dans $\mathbb{F}_p^n$ sauf le vecteur nul. Il y a (p^n-1) choix possibles pour $\varphi(e_1)$. Ensuite, comme $\varphi(e_2)$ doit être linéairement indépendant de $\varphi(e_1)$, on a le choix entre tous les vecteurs qui appartiennent à $\mathbb{F}_p^n$, sauf

les vecteurs de Vect $(\varphi(e_1))$; il y a $(p^n - p)$ choix possibles pour $\varphi(e_2)$. On continue ainsi jusqu'au choix de $\varphi(e_n)$, où on doit prendre un vecteur qui appartient à $\mathbb{F}_p^n$, mais qui n'appartient pas au sous-espace Vect $(\varphi(e_1), \varphi(e_2), \ldots, \varphi(e_{n-1}))$. Nous avons donc $(p^n - p^{n-1})$ choix possibles pour cette dernière image. Par le principe d'une suite de choix, on trouve que le nombre d'applications linéaires bijectives entre $\mathbb{F}_p^n$ et $\mathbb{F}_p^n$ est le produit indiqué. □

5.4. **Changement de base.**

Définition 5.4.1. Soient $B = (e_1, \ldots, e_n)$ et $B' = (f_1, \ldots, f_n)$ deux bases d'un K-espace vectoriel V. On exprime les f_j en termes de la base B:

$$f_j = p_{1j}e_1 + p_{2j}e_2 + \cdots p_{nj}e_n, \text{ pour } 1 \leq j \leq n \text{ et } p_{ij} \in K.$$

On définit la matrice $P \in M_n(K)$ par $P = (p_{ij})$; donc la j-ème colonne de P est le vecteur colonne $(f_j)_B$. La matrice $P \in M_n(K)$ s'appelle la *matrice de changement de base entre la base B' et la base B*. On dit aussi que P est la *matrice de passage entre la base B' et la base B*. On note que P est la matrice de l'application identité id $: V \to V$, par rapport aux bases B' et B; c'est-à-dire $P = (\mathrm{id})_{B'}^{B}$.

Proposition 5.4.2. *Soient V, B, B' et P comme dans la Définition* 5.4.1 *et soit $v \in V$. Alors $P \cdot (v)_{B'} = (v)_B$.*

Preuve. Par définition, $P \cdot (v)_{B'} = (\mathrm{id})_{B'}^{B} \cdot (v)_{B'}$. Par le Théorème 5.3.7, $(\mathrm{id})_{B'}^{B} \cdot (v)_{B'} = (\mathrm{id}(v))_B = (v)_B$. □

Proposition 5.4.3 (Matrice de passage inverse)**.** *Soit $P = (\mathrm{id})_{B'}^{B}$ la matrice de passage entre les bases B' et B. Alors P est inversible et son inverse P^{-1} est la matrice de passage $(\mathrm{id})_{B}^{B'}$ entre les bases B et B'.*

Preuve. Par le Corollaire 5.3.11, la matrice $P = (\mathrm{id})_{B'}^{B}$ est inversible et son inverse est la matrice de l'application id^{-1}, par rapport aux bases B et B' (dans cet ordre). Donc $P^{-1} = (\mathrm{id}^{-1})_{B}^{B'} = (\mathrm{id})_{B}^{B'}$. □

Théorème 5.4.4 (changement de base)**.** *Soit $\phi \in \mathcal{L}(V,W)$. Soient B et B' deux bases de V et soient C et C' deux bases de W. Posons $S = (\mathrm{id}_V)_{B'}^{B}$ et $T = (\mathrm{id}_W)_{C'}^{C}$ (deux matrices de passage), et $A = (\phi)_B^C$ et $B = (\phi)_{B'}^{C'}$. Alors*

$$B = T^{-1}AS, \text{ c'est-à-dire } (\phi)_{B'}^{C'} = (\mathrm{id}_W)_C^{C'} \cdot (\phi)_B^C \cdot (\mathrm{id}_V)_{B'}^{B} .$$

Preuve. Théorème 5.3.9 $\implies (\mathrm{id}_W)_C^{C'} \cdot (\phi)_B^C \cdot (\mathrm{id}_V)_{B'}^{B} = (\mathrm{id}_W \circ \phi \circ \mathrm{id}_V)_{B'}^{C'} = (\phi)_{B'}^{C'}$. □

Définition 5.4.5. Soient $A, B \in M_n(K)$. On dit que A et B sont *semblables* s'il existe une matrice inversible $P \in M_n(K)$ telle que $P^{-1}AP = B$. On vérifie que 'être semblable' est une relation d'équivalence sur $M_n(K)$. (Exercice.)

Exemple 5.4.6. Soit $\phi \in \mathcal{L}(V,V)$, où $\dim V = n$. Soient B et C deux bases de V. Alors les matrices $(\phi)_B$ et $(\phi)_C$ sont semblables, car

$$(\phi)_B = (\mathrm{id})_C^B \cdot (\phi)_C \cdot (\mathrm{id})_B^C = ((\mathrm{id})_B^C)^{-1} \cdot (\phi)_C \cdot (\mathrm{id})_B^C .$$

On fixe un corps K.

6.1. **Définitions et premières propriétés.**

Notation 6.1.1. a) Soit X un ensemble. Pour $a, b \in X$, le symbole de Kronecker δ_{ab} désigne le nombre réel tel que $\delta_{ab} = 0$ si $a \neq b$ et $\delta_{ab} = 1$ si $a = b$.

b) Pour $1 \leq r \leq n$, $1 \leq s \leq m$, on définit une matrice $E_{rs} \in M_{n\times m}(K)$ dont le coefficient $(E_{rs})_{ij}$ satisfait

$$(E_{rs})_{ij} = \delta_{ri}\delta_{sj},$$

c'est-à-dire, E_{rs} est la matrice $n \times m$ telle que le seul coefficient non nul est $(E_{rs})_{rs}$ et ce coefficient est égal à 1. On note que $\{E_{ij} \mid 1 \leq i \leq n, 1 \leq j \leq m\}$ est une base de $M_{n\times m}(K)$.

c) Pour $A \in M_{n\times m}(K)$, on note $A = (A_{ij})_{\substack{1\leq i\leq n \\ 1\leq j\leq m}} = (A_{ij}) = (a_{ij})$, et on note A_i la i-ème ligne de A.

Définition 6.1.2. On définit trois types d'opérations sur les lignes d'une matrice, appelées *opérations élémentaires*.

Type I Echanger deux lignes de la matrice.

Type II Multiplier une ligne de la matrice par un scalaire non nul $\lambda \in K$.

Type III Additionner à une ligne de la matrice un multiple scalaire d'une autre ligne de la matrice.

Remarque 6.1.3. On note que ces trois opérations sont 'réversibles' dans le sens qu'il existe une opération élémentaire qui renvoie à la matrice de départ.

i. Type I est son propre inverse.

ii. Type II: On multiplie la même ligne par $\frac{1}{\lambda}$.

iii. Type III: Si on additionne $\lambda \cdot A_i$ à A_j, alors l'opération inverse est d'additionner $-\lambda \cdot A_i$ à A_j.

Définition 6.1.4. On dit que $A, B \in M_{n\times m}(K)$ sont *lignes équivalentes* (ou *équivalentes par lignes*), si B peut être obtenue à partir de A en faisant une suite (finie) d'opérations élémentaires.

Remarque 6.1.5. Comme toutes ces opérations sont réversibles, 'être lignes équivalentes' est une relation d'équivalence sur $M_{n\times m}(K)$. En particulier, la relation est symétrique, c'est-à-dire, si on peut obtenir B à partir de A par une suite d'opérations élémentaires, on peut également obtenir A à partir de B par une suite d'opérations élémentaires.

Propriété. Chaque opération élémentaire sur les lignes de $A \in M_{n\times m}(K)$ correspond à la multiplication à gauche par une certaine matrice de $\mathrm{GL}_n(K)$.

Preuve. **Type I:** Echanger les lignes A_r et A_s de la matrice A.

On considère la permutation $\tau = (rs) \in S_n$ du groupe symétrique de degré n. Posons $T_{rs} \in M_n(K)$ la matrice dont le coefficient $(T_{rs})_{ij}$ satisfait:

$$(T_{rs})_{ij} = \delta_{\tau(i),j}.$$

On vérifie que

- $(T_{rs})_{ii} = 1$ si $i \notin \{r, s\}$,
- $(T_{rs})_{rs} = 1 = (T_{rs})_{sr}$ et
- tous les autres coefficients de T_{rs} sont nuls.

On montre que $T_{rs}A$ est la matrice obtenue à partir de A en échangeant les lignes A_r et A_s:

Preuve. $(T_{rs}A)_{ik} = \sum_{j=1}^{n}(T_{rs})_{ij}A_{jk}$. Comme $(T_{rs})_{ij} = \delta_{\tau(i),j} = 0$ sauf si $\tau(i) = j$, on a

$$(T_{rs}A)_{ik} = \sum_{j=1}^{n}(T_{rs})_{ij}A_{jk} = A_{\tau(i),k}.$$

Et ce dernier terme est égal à A_{ik} si $i \neq r, s$, A_{sk} si $i = r$, et A_{rk} si $i = s$. Cela veut dire que $T_{rs}A$ est bien la matrice obtenue en échangeant les lignes A_r et A_s.

On note que $T_{rs} = (T_{rs})^{-1}$.

Type II: Multiplier la ligne A_r par $\lambda \in K$, $\lambda \neq 0$.

Posons $D_r(\lambda) \in M_n(K)$ la matrice dont le coefficient $D_r(\lambda)_{ij}$ satisfait

$$D_r(\lambda)_{ij} = \begin{cases} 1, & \text{si } i = j \neq r \\ 0, & \text{si } i \neq j \\ \lambda, & \text{si } i = j = r \end{cases}$$

Alors $(D_r(\lambda)A)_{ik} = \sum_{j=1}^{n} D_r(\lambda)_{ij} A_{jk} = D_r(\lambda)_{ii} A_{ik} = \begin{cases} A_{ik}, & i \neq r \\ \lambda A_{rk} & i = r \end{cases}$.
Donc $D_r(\lambda)A$ est la matrice obtenue en multipliant la r-ème ligne de A par λ. De ce fait, on a

$$(D_r(\lambda))^{-1} = D_r\left(\frac{1}{\lambda}\right).$$

Type III: Additionner $\lambda \cdot A_s$ à A_r .

Posons $L_{rs}(\lambda) = I_n + \lambda E_{rs} \in M_n(K)$. Pour simplifier, on écrit $I = I_n$. On calcule

$$L_{rs}(\lambda)A = (I + \lambda E_{rs})A = A + \lambda E_{rs}A.$$

$$\begin{aligned} (L_{rs}(\lambda)A)_{ik} &= A_{ik} + \textstyle\sum_{j=1}^{n} (\lambda E_{rs})_{ij} A_{jk} \\ &= A_{ik} + \delta_{ri} \lambda A_{sk} \end{aligned}.$$

Ce dernier est égal à A_{ik} si $i \neq r$ et à $A_{rk} + \lambda A_{sk}$ si $i = r$. Donc c'est bien le résultat d'additionner $\lambda \cdot A_s$ à A_r. De ce fait, on a

$$(L_{rs}(\lambda))^{-1} = L_{rs}(-\lambda).$$

□

Définition 6.1.6. Les matrices T_{rs}, $D_r(\lambda)$ et $L_{rs}(\lambda)$ s'appellent les *matrices élémentaires.*

Remarque 6.1.7.

- $T_{rs}T_{rs} = I_n$ et par conséquent $T_{rs}^{-1} = T_{rs}$.
- $D_r(\lambda^{-1})D_r(\lambda) = I_n$ et par conséquent $(D_r(\lambda))^{-1} = D_r(\lambda^{-1})$.
- $L_{rs}(-\lambda)L_{rs}(\lambda) = I_n$ et par conséquent $(L_{rs}(\lambda))^{-1} = L_{rs}(-\lambda)$.

6.2. **Echelonnage et la méthode de Gauss.**

Définition 6.2.1. On dit qu'une matrice $A = (a_{ij}) \in M_{p\times n}(K)$ est *échelonnée* si soit $A = \mathbf{0}_{p\times n}$, soit il existe un entier $r \leq n$, $r \leq p$, et des entiers $j_1, j_2, \ldots, j_r$ entre 1 et n avec les propriétés suivantes:

1. $1 \leq j_1 < j_2 < \cdots < j_r \leq n$;
2. $a_{1j} = 0$ pour tout $j < j_1$ et $a_{1j_1} \neq 0$;

   $a_{2j} = 0$ pour tout $j < j_2$ et $a_{2j_2} \neq 0$;

   $\vdots$

   $a_{rj} = 0$ pour tout $j < j_r$ et $a_{rj_r} \neq 0$;
3. Si $r < p$ les lignes $A_{r+1}, \ldots A_p$ sont nulles.

Si $A \neq \mathbf{0}$, les entiers $j_1, \ldots, j_r$ s'appellent les *échelons* de la matrice A et les éléments a_{ij_i}, $i = 1, ..., r$, s'appellent les *pivots*.

Définition 6.2.2. Une matrice $A = (a_{ij}) \in M_{p\times n}(K)$ est dite *échelonnée réduite* si soit $A = \mathbf{0}$, soit A est échelonnée avec échelons $j_1 < j_2 < \cdots < j_r$ et si de plus on a les propriétés suivantes:

1. $a_{1j_1} = 1$, $a_{2j_2} = 1, \ldots, a_{rj_r} = 1$;
2. $a_{kj_i} = 0$ pour tout $k \neq i$ (c'est-à-dire, dans toute la colonne à l'échelon j_i, le seul coefficient non nul est a_{ij_i}).

Théorème 6.2.3 (l'échelonnage d'après la méthode de Gauss)**.** *Toute matrice est ligne équivalente à une matrice échelonnée réduite. Autrement dit, toute matrice peut être transformée en une matrice échelonnée réduite par une suite d'opérations élémentaires sur les lignes de la matrice.*

Preuve. (La preuve est constructive, dans le sens qu'elle donne un algorithme pour trouver une forme échelonnée réduite; cette méthode s'appelle la *méthode de Gauss* ou la *méthode d'élimination de Gauss*.)

Soit $A \in M_{p\times n}(K)$. On suppose $A \neq 0$, car sinon A est déjà échelonnée réduite.

Soit j_1 le plus petit indice colonne pour lequel un coefficient de A est non nul, disons $a_{ij_1} \neq 0$. Par une opération de type I (échanger les lignes 1 et i), on est ramené au cas ou $a_{1j_1} \neq 0$ (le pivot est en première ligne).

Par une opération de type II, multiplier la première ligne par $a_{1j_1}^{-1}$, on est ramené au cas où $a_{1j_1} = 1$ (le pivot en première ligne est égal à 1).

Par une suite d'opérations de type III, on annule tous les autres coefficients de la j_1-ème colonne (rajouter $-a_{kj_1} \times$ ligne 1 à la ligne k).

Ainsi, on aboutit à une matrice de la forme

$$A' = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{pmatrix}.$$

Posons B la matrice constituée des lignes 2 à p de A'. Si $B = 0$, alors A' est échelonnée réduite. Sinon soit j_2 le plus petit indice colonne pour lequel B possède un coefficient non nul. Alors $j_2 > j_1$. On applique à la matrice B la méthode utilisée ci-dessus, ce qui crée (en reportant les opérations sur la matrice A') une matrice

$$A'' = \begin{pmatrix} 0 & \cdots & 0 & 1 & * & * & * & * & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \cdots & \vdots & 0 & \cdots & 0 & 0 & * & \cdots & * \end{pmatrix},$$

où nous avons un pivot $a_{1j_1} = 1$ et $a_{2j_2} = 1$. Avec une opération de type III, on annule le coefficient a_{1j_2} à la ligne 1 (rajouter $-a_{1j_2} \times$ ligne 2 à la ligne 1) . Cette dernière opération ne modifie pas les éléments sur la première ligne de A'' précédant la colonne j_2, i.e. les $a''_{1\ell}$ avec $\ell < j_2$, car tous les coefficients correspondant sur la ligne 2, i.e. les $a''_{2\ell}$ avec $\ell < j_2$, sont nuls.

On répète ce procédé jusqu'à ce qu'on obtienne une matrice échelonnée réduite. □

Remarque 6.2.4. Soient $A, B \in M_{p\times n}(K)$ des matrices lignes équivalentes. Alors chaque ligne de B est une combinaison linéaire des lignes de A et chaque ligne de A est une combinaison linéaire des lignes de B.

Corollaire 6.2.5 (L'unicité de la forme échelonnée réduite)**.** *Soient $A, R, R' \in M_{p\times n}(K)$ lignes équivalentes. Si R et R' sont échelonnées réduites, alors $R = R'$.*

Esquisse de preuve. Par les remarques ci-dessus, les lignes de R sont des combinaisons linéaires des lignes de R' et vice versa.

Supposons que R ait des échelons aux colonnes $j_1 < j_2 < \cdots < j_r$ et R' aux colonnes $k_1 < k_2 < \cdots < k_s$. On déduit que $j_1 = k_1$, sinon soit R_1 n'est pas dans le sous-espace engendré par les lignes de R', soit R'_1 n'est pas dans le sous-espace engendré par les lignes de R.

Maintenant

$$R_1 = \begin{pmatrix} 0 & \cdots & 0 & 1 & a_{1j_1+1} & \cdots & a_{1j_2-1} & 0 & * & \cdots \end{pmatrix},$$

avec le pivot 1 à la place $(1, j_1)$ et

$$R_2 = \begin{pmatrix} 0 & \cdots & 0 & 1 & a_{2j_2+1} & * & \cdots \end{pmatrix},$$

avec le pivot 1 à la place $(2, j_2)$, et

$$R'_1 = \begin{pmatrix} 0 & \cdots & 0 & 1 & b_{1j_1+1} & \cdots & b_{1k_2-1} & 0 & * & \cdots \end{pmatrix},$$

avec le pivot 1 à la place $(1, k_1) = (1, j_1)$ et

$$R'_2 = \begin{pmatrix} 0 & \cdots & 0 & 1 & b_{2k_2+1} & * & \cdots \end{pmatrix},$$

avec le pivot 1 à la place $(2, k_2)$.

Alors R_2 est une combinaison linéaire des lignes R'_k, $k \geq 2$, de R' et R'_2 est une combinaison linéaire des lignes R_k, $k \geq 2$, de R (la première ligne étant toujours exclue, car $R_{2,j_1} = R'_{2,j_1} = 0$). Pour la même raison que précédemment, on déduit que $j_2 = k_2$ et donc $(a_{1j_1+1}\ a_{1j_1+2}\ \cdots\ a_{1j_2-1}) = (b_{1j_1+1}\ b_{1j_1+2}\ \cdots\ b_{1k_2-1})$.

La suite consiste à comparer les lignes R_i et R'_i de manière analogue. □

Corollaire 6.2.6. *Soit $A \in M_{p\times n}(K)$. Alors il existe une matrice inversible $P \in \mathrm{GL}_p(K)$ telle que PA est échelonnée réduite.*

Preuve. On applique la méthode de Gauss; chaque opération élémentaire correspond à multiplier à gauche par un élément de $\mathrm{GL}_p(K)$. Le produit d'éléments dans $\mathrm{GL}_p(K)$ est aussi dans $\mathrm{GL}_p(K)$. On a $P_N \cdots P_1 A$ échelonnée réduite et $P_N \cdots P_1 \in \mathrm{GL}_p(K)$. □

6.3. **Applications de la méthode de Gauss.**

6.3.1. *Systèmes de vecteurs; base d'un sous-espace défini par un système de générateurs.* Soit $(u_1, \ldots, u_p)$ un p-uplet ordonné de vecteurs dans K^n. On souhaite:

- Trouver une base aussi simple que possible de $\mathrm{Vect}\,(u_1, \ldots, u_p) = U$.
- Trouver $\dim U$.
- Compléter cette base en une base de K^n.

Méthode: On écrit les coordonnées des vecteurs $u_1, \ldots, u_p$ dans les lignes d'une matrice $A \in M_{p\times n}(K)$, c'est-à-dire on pose $A = (a_{ij}) \in M_{p\times n}(K)$, où $A_i = (a_{i1}, \ldots, a_{in}) = u_i$. On effectue les opérations élémentaires sur les lignes de A pour la transformer en une matrice échelonnée réduite R. Soit $w_i \in K^n$ le vecteur avec coordonnées données par la ligne R_i. Alors $\mathrm{Vect}\,(w_1, \ldots, w_p) = U$, car chaque opération élémentaire est inversible. Donc chaque w_i est une combinaison linéaire des $u_1, \ldots, u_p$ et chaque u_j est une combinaison linéaire des $w_1, \ldots, w_p$. Les w_i non nuls forment une base de U. Donc $\dim U$ est égale au nombre de lignes non nulles de la matrice R. Pour compléter cette base en une base de K^n, on prend les vecteurs $\{e_\ell \mid \ell \text{ n'est pas un échelon de la matrice } R\}$.

Définition 6.3.1.

(1) Un *système de vecteurs* dans un K-espace vectoriel V est un uplet ordonné de vecteurs $(v_1, \ldots, v_t)$ dans V.

(2) On dit qu'un système de vecteurs $(w_1, \ldots, w_t)$, avec $w_i \in K^n$, est *échelonné (réduit)* si la matrice ayant pour i-ème ligne les coordonnées de w_i est échelonnée (réduite).

(3) Pour un système de vecteurs $(u_1, \ldots, u_t)$ dans V, $\dim(\mathrm{Vect}\,(u_1, \ldots, u_t))$ s'appelle le *rang* du système.

On note que le rang du système $(u_1, \ldots, u_t)$ est le nombre maximal de vecteurs linéairement indépendants dans l'ensemble $\{u_1, \ldots, u_t\}$.

Définition 6.3.2. Soit $A \in M_{m\times n}(K)$. On regarde les lignes $A_1, \ldots, A_m$ de A comme vecteurs dans K^n. On dit que le *rang-ligne* de A est le rang du système $(A_1, \ldots, A_m)$.

On note que le rang-ligne de A est égal au rang-ligne de R, où R est la forme échelonnée réduite de A, et que le rang-ligne de R est le nombre d'échelons de R.

Plus généralement, soit V un K-espace vectoriel de dimension n avec base ordonnée $B = (f_1, \ldots, f_n)$. On a une application linéaire bijective $\varphi : K^n \to V$ donnée par $\varphi((a_1, \ldots, a_n)) = \sum_{i=1}^n a_i f_i$. Soit maintenant $(v_1, \ldots, v_p)$ un système de vecteurs dans V. Alors comme φ est bijective, $\dim \mathrm{Vect}\,(v_1, \ldots, v_p) = \dim \mathrm{Vect}\,(\varphi^{-1}(v_1), \ldots, \varphi^{-1}(v_p)) =$ le rang-ligne de la matrice avec lignes $\varphi^{-1}(v_1), \ldots, \varphi^{-1}(v_p)$. Ainsi, on peut utiliser la méthode développée ci-dessus pour déterminer $\dim \mathrm{Vect}\,(v_1, \ldots, v_p)$, une base de cet espace, et également pour compléter cette base en une base de V.

6.3.2. *L'image d'une application linéaire.*

Définition 6.3.3. Soit $B \in M_{m\times n}(K)$. On note par $c_i \in K^m$ le vecteur dont les coordonnées se trouvent dans la i-ème colonne de B. Le *rang-colonne* de B est $\dim(\mathrm{Vect}\,(c_1, \ldots, c_n))$ comme sous-espace de K^m.

Remarque 6.3.4 (Lien avec les applications linéaires)**.** Soit $\phi : V \to W$ une application K-linéaire entre espaces vectoriels de dimension finie, et soient B_V et B_W des bases ordonnées de V et W respectivement. On rappelle que le rang de ϕ est égal à la dimension de $\mathrm{im}\,(\phi)$. Posons $B = (\phi)_{B_V}^{B_W} \in M_{m\times n}(K)$ et notons les colonnes de B par $c_1, \ldots, c_n \subset K^m$. Alors $\dim \mathrm{im}\,(\phi) = \dim(\mathrm{Vect}\,(c_1, \ldots, c_n))$, ce qui est le rang-colonne de B.

Définition 6.3.5. Soit $A \in M_{p\times q}(K)$. La transposée de A, notée A^t, est la matrice $q \times p$ telle que $(A^t)_{ij} = A_{ji}$.

Remarque 6.3.6. Comme les colonnes de A^t sont les lignes de A et les lignes de A^t sont les colonnes de A, le rang-colonne de A est égal au rang-ligne de A^t. On utilise ce fait pour calculer le rang-colonne d'une matrice et le rang d'une application linéaire, et même pour trouver une base de $\mathrm{im}\,\phi$.

6.3.3. *Systèmes d'équations linéaires.*

Définition 6.3.7. Soit un système de p équations linéaires à n inconnues:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad \vdots \qquad &\vdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n &= b_p, \end{aligned}$$

avec $a_{ij}, b_k \in K$. Si $b_i = 0$ pour tout $1 \leq i \leq p$, on dit que le système est *homogène* et s'il existe i, $1 \leq i \leq p$, avec $b_i \neq 0$, on dit que le système est *inhomogène*.

En termes matriciels: Posons $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix}$, *la matrice des coefficients*, $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, *le vecteur colonne des inconnues*, et $b = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$, *le vecteur colonne des termes constants*. Alors le système d'équations est équivalent à l'équation matricielle $AX = b$.

Définition 6.3.8. Le *rang* du système est le rang-ligne de la matrice des coefficients A.

En termes d'applications linéaires: la matrice A représente une application linéaire $\phi : K^n \to K^p$ (par rapport aux bases canoniques des deux espaces). L'existence d'une solution du système veut dire qu'il existe $x = (x_1, \ldots, x_n) \in K^n$ tel que $\phi(x) = b$. C'est-à-dire b appartient à l'image de ϕ.

Si $b \notin \operatorname{im}(\phi)$, alors le système ne possède aucune solution. Si $b \in \operatorname{im}(\phi)$, le système possède au moins une solution.

Cas particulier: si le système est homogène, c'est-à-dire $b = 0$, alors il existe au moins une solution car $\phi(0) = 0$. Même, l'ensemble des solutions du système est égal à $\{v \in K^n \mid \phi(v) = 0\} = \ker(\phi)$, et par conséquent on a le résultat suivant:

Proposition 6.3.9. *L'ensemble des solutions d'un système homogène est un sous-espace vectoriel de K^n, car c'est le noyau d'une application K-linéaire.*

Pour résoudre un système linéaire par la méthode de Gauss, on réduit à un système échelonné réduit $A'X = b'$. Pour ce faire, on forme la matrice augmentée du système $(A \mid b)$ où on rajoute b comme une dernière colonne. Ensuite on réduit cette nouvelle matrice à sa forme échelonnée réduite, $(A' \mid b')$. Soit r le rang du système (égal au rang-ligne de la matrice A'), et soient $j_1, \ldots, j_r$ les échelons de la matrice A'.

Définition 6.3.10. Les inconnues qui apparaissent aux échelons du système échelonné réduit, $x_{j_1}, \ldots, x_{j_r}$ s'appellent *les inconnues principales* et les autres (s'il y en a) s'appellent *les inconnues libres* (qui sont $n - r$ en nombre).

Description des solutions:

Cas I: Si l'un des scalaires $b'_{r+1}, \ldots b'_p$ n'est pas nul, on a l'équation $0 = b'_i$, et le système ne possède aucune solution.

Cas II: Si $b'_{r+1} = ... = b'_p = 0$, ou si $r = p$, le système possède au moins une solution. Pour décrire les solutions, on donne des valeurs arbitraires aux inconnues libres et on détermine la valeur de chaque inconnue principale en termes des inconnues libres. Donc s'il existe des inconnues libres, il y a plus qu'une solution et s'il n'existe aucune inconnue libre ($\Leftrightarrow r = n$), le système possède une solution unique.

Proposition 6.3.11 (L'ensemble des solutions dans le cas homogène)**.** *Considérons un système homogène de p équations linéaires à n inconnues, et de rang r.*

(1) On a $r \leq p$ et $r \leq n$ (par définition du rang).

(2) Il existe toujours la solution dite triviale*, i.e. la solution $x \in K^n$ où $x_i = 0$ pour tout i.*

(3) L'ensemble des solutions est un sous-espace vectoriel de K^n de dimension $n - r$ (= le nombre d'inconnues libres).

(4) Si $n \leq p$ et $n = r$, il n'y a que la solution triviale.

(5) Si $n > p$, on a $n > p \geq r$ et donc $n - r > 0$, et il existe des solutions non triviales.

La méthode pour trouver une base du sous-espace des solutions d'un système est expliquée en cours.

Proposition 6.3.12 (L'ensemble de solutions dans le cas inhomogène)**.** *Considérons un système inhomogène de p équations linéaires à n inconnues, de rang r, et avec système échelonné réduit associé $A'X = b'$.*

(1) Le système ne possède aucune solution si et seulement s'il existe b'_i avec $i \geq r+1$ et $b'_i \neq 0$.

(2) Si $y = (y_1, \ldots, y_n) \in K^n$ est une solution du système $AX = b$, alors l'ensemble des solutions du système est $\{y + x \mid x \in K^n$ est une solution du système $AX = 0\}$.

(3) Si $n = p$ et $r = n$, le système possède une solution unique.

Preuve. L'affirmation de (1) est claire.

Pour (2): Posons $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ tel que $AY = b$. Prenons $(\alpha_1, \ldots, \alpha_n) \in K^n$ une solution du système homogène $AX = 0$. Posons $Z = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$, donc $AZ = 0$. Par conséquent $A(Y + Z) = AY + AZ = b + 0 = b$. De plus, si Y' est une autre solution de $AX = b$, alors $AY' = b = AY$, et donc $A(Y' - Y) = b - b = 0$ et $Y' = Y + (Y' - Y)$, où $Y' - Y$ est bien une solution du système homogène.

(3): Si $n = p$, alors le nombre d'équations est égal au nombre d'inconnues et $r = n$ veut dire que le nombre d'échelons est aussi égal au nombre d'équations et donc il n'y a pas de lignes de $(A' \mid b')$ de la forme $(0\ 0\ \cdots\ 0 \mid b'_i)$ avec $b'_i \neq 0$ et par (1), il existe une solution. S'il existe deux solutions Y et Y' à $AX = b$, alors le système $AX = 0$ possède une solution $Z = Y - Y'$ qui ne peut être que triviale (i.e. $Y = Y'$) par la partie (4) de la proposition précédente. □

6.4. **Le rang d'une matrice.** Soit $A \in M_{m\times n}(K)$. Nous avons défini le rang-ligne et le rang-colonne de la matrice A. Le rang-ligne est le nombre maximal de lignes de A qui

sont linéairement indépendantes vues comme des vecteurs dans K^n et le rang-colonne est la dimension de $\operatorname{im}(\varphi)$, où $\varphi : K^n \to K^m$ est l'application linéaire représentée par la matrice A par rapport aux bases canoniques de K^n et K^m (i.e. $(\varphi)_E^E = A$).

Théorème 6.4.1. *Le rang-ligne de A est égal au rang-colonne de A.*

Preuve. Posons $r =$ le rang-ligne de A. Alors $\ker\varphi$ est égal à l'ensemble des solutions de l'équation $AX = 0$ et, par la Proposition 6.3.11, est de dimension $n - r$. Par le théorème du rang, $\dim K^n = \dim\ker\varphi + \dim\operatorname{im}\varphi$.

On a donc $n = (n - r) + \dim\operatorname{im}\varphi$ et on déduit que $\dim\operatorname{im}\varphi = r$, c'est-à-dire, le rang-ligne est égal au rang-colonne. □

Définition 6.4.2. Le *rang d'une matrice A* est le rang-ligne (ou le rang-colonne) de A, noté $\operatorname{rang}(A)$ ou $\operatorname{rg}(A)$.

6.5. **Inversion des matrices carrées.**

Théorème 6.5.1 (d'inversibilité)**.** *Soit $A \in M_n(K)$. Les conditions suivantes sont équivalentes.*

(a) *A est inversible.*

(b) *Il existe $C \in M_n(K)$ telle que $AC = I_n$.*

(c) *Il existe $B \in M_n(K)$ telle que $BA = I_n$.*

(d) *Le système $AX = 0$ possède une solution unique, la solution triviale.*

(e) $\operatorname{rang}(A) = n$.

(f) *La matrice échelonnée réduite qui est ligne équivalente à A est la matrice identité I_n.*

Preuve. D'abord on montre que (a), (b) et (c) sont équivalents. Il est clair que (a)$\Longrightarrow$(b) et (a)$\Longrightarrow$(c).

Supposons maintenant qu'il existe $C \in M_n(K)$ telle que $AC = I_n$. Soient $\varphi : K^n \to K^n$ tel que $(\varphi)_E^E = A$ et $\psi : K^n \to K^n$ telle que $(\psi)_E^E = C$, où E est la base canonique de K^n. Alors $\varphi\circ\psi = \mathrm{id}$ implique que φ est surjective. Mais φ surjective implique φ bijective (par exemple en utilisant le théorème du rang) et donc A est inversible. Par conséquent (b)$\Longrightarrow$(a).

Supposons qu'il existe $B \in M_n(K)$ telle que $BA = I_n$. Soit φ comme ci-dessus et posons $\gamma : K^n \to K^n$ telle que $(\gamma)_E^E = B$. On a alors $\gamma \circ \varphi = \mathrm{id}$ et donc φ est injective. Mais φ injective implique que φ est bijective et donc A est inversible, et nous avons que (c)$\Longrightarrow$(a). Ces implications montrent que (a), (b) et (c) sont équivalents.

On montre maintenant que (a)$\Longrightarrow$(d)$\Longrightarrow$(e)$\Longrightarrow$(f)$\Longrightarrow$(c) pour conclure.

(a)$\Longrightarrow$(d) Supposons que A est inversible avec inverse A^{-1}. On considère l'équation $AX = 0$; on multiplie à gauche par A^{-1} des deux cotés et on obtient que $A^{-1}AX = A^{-1}0$, d'où $X = 0$. Donc il n'existe que la solution triviale $X = 0$.

(d)$\Longrightarrow$(e) On sait que la dimension de l'espace des solutions du système est égale à $n - r$, où $r = \mathrm{rang}(A)$. Mais l'espace des solutions est le sous-espace nul et donc est de dimension 0. On déduit que $n = r$, i.e. la matrice A est de rang n.

(e)$\Longrightarrow$(f) Supposons maintenant que $\mathrm{rang}(A) = n$. On effectue les opérations élémentaires sur les lignes de A pour obtenir une matrice échelonnée réduite R. Le rang de R est aussi n et par conséquent $R = I_n$.

(f)$\Longrightarrow$(c) Supposons qu'il existe $E_1, \ldots, E_t$ des matrices élémentaires telles que $E_1 \cdots E_t A = I_n$. Alors la matrice $B = E_1 \cdots E_t$ satisfait à la condition de (c). □

Corollaire 6.5.2. *Toute matrice inversible est un produit de matrices correspondant aux opérations élémentaires.*

Preuve. Par (f), il existe des matrices élémentaires $E_1, \ldots, E_t$ telles que $E_1 \cdots E_t A = I_n$, ce qui montre que $A = E_t^{-1} E_{t-1}^{-1} \cdots E_1^{-1}$ et $A^{-1} = E_1 \cdots E_t$. □

Le corollaire nous donne un algorithme pour calculer facilement l'inverse d'une matrice (ou bien pour déterminer si une matrice donnée est inversible).

On fixe un corps K.

7.1. **Le groupe symétrique.**

7.1.1. *Notation en cycles.* On rappelle que le groupe symétrique de degré n est le groupe $(\mathrm{Bij}(X), \circ)$, noté aussi S_n, où $X = \{1, 2, \ldots, n\}$.

Définition 7.1.1. 1. Soit $\{a_1, a_2, \ldots, a_m\} \subseteq \{1, 2, \ldots, n\}$ une partie de m éléments distincts. On écrit $(a_1\ a_2\ \cdots\ a_m)$ pour la permutation $\sigma \in S_n$ définie par :

$\sigma(a_i) = a_{i+1}$ pour $1 \leq i \leq m-1$, $\sigma(a_m) = a_1$, et $\sigma(b) = b$ pour tout $b \in \{1, 2, \ldots, n\} \setminus \{a_1, a_2, \ldots, a_m\}$. On appelle un tel élément un *m-cycle*.

2. Un 2-cycle s'appelle une *transposition*.
3. Pour $\sigma \in S_n$, on pose $\mathrm{supp}(\sigma) := \{j \in \{1, 2, \ldots, n\} \mid \sigma(j) \neq j\}$, le *support de σ*.

Exemples 7.1.2. 1. L'élément neutre, la permutation identité, est égale au 1-cycle (1), et aussi au 1-cycle (2), etc. Noter que $\mathrm{supp}((1)) = \emptyset$.

2. Dans le groupe S_3, tout élément est soit
   - un 1-cycle, (l'élément neutre),
   - soit un 2-cycle ($\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (1\ 2)$, $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (2\ 3)$ ou $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (1\ 3)$)
   - soit un 3-cycle ($\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (1\ 2\ 3)$ ou $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (1\ 3\ 2)$). Noter que le 3-cycle (1 2 3) est égal au 3-cycle (2 3 1), qui est égal au 3-cycle (3 1 2).
3. Dans le groupe S_4, l'élément $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$ est le 4-cycle (1 2 3 4), mais pour tous $r \geq 1$, l'élément $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$ n'est pas un r-cycle.

Le dernier exemple montre que la notation introduite dans 7.1.1 ne suffit pas pour décrire tous les éléments de S_n. La proposition 7.1.4 traite des éléments généraux.

Définition 7.1.3. On dit que deux cycles $\sigma = (a_1\ a_2\ \dots\ a_m), \tau = (b_1\ b_2\ \cdots\ b_\ell) \in S_n$ sont *disjoints* si $\mathrm{supp}(\sigma)$ et $\mathrm{supp}(\tau)$ sont disjoints, c'est-à-dire que $\{a_1, \dots, a_m\} \cap \{b_1, \dots, b_\ell\} = \emptyset$.

Noter que si σ et τ sont des cycles disjoints, alors $\sigma\tau = \tau\sigma$.

Proposition 7.1.4. *Toute permutation $\sigma \in S_n \backslash \{\mathrm{id}\}$, s'écrit comme un produit de cycles disjoints, chacun de longueur au moins 2. Cette factorisation est unique à l'ordre près des cycles.*

Preuve. (facultatif) Pour $m \in \{1, 2, \dots, n\}$ et $\sigma \in S_n$, on appelle $Y = \{\sigma^j(m) \mid j \in \mathbb{Z}\}$ *l'orbite de m sous l'action de σ* ou simplement *la σ-orbite de m.*

D'abord on montre que $Y = \{m, \sigma(m), \dots, \sigma^{k-1}(m)\}$, où $k \in \mathbb{N}$ est maximal tel que $\{m, \sigma(m), \dots, \sigma^{k-1}(m)\}$ soient distincts. Soit k comme dans l'assertion. On a que $\sigma^k(m) = \sigma^j(m)$ pour un certain $0 \leq j \leq k-1$. Donc $\sigma^{k-j}(m) = m$ et par le choix de k, $j = 0$ et $\sigma^k(m) = m$. Pour $s \in \mathbb{Z}$, on écrit $s = qk + r$ pour $q, r \in \mathbb{Z}$ avec $0 \leq r < k$. Alors $\sigma^s(m) = \sigma^r((\sigma^{qk})(m)) = \sigma^r(\sigma^k(\cdots(\sigma^k(m))) = \sigma^r(m) \in \{m, \sigma(m), \dots, \sigma^{k-1}(m)\}$. Donc $Y = \{m, \sigma(m), \dots, \sigma^{k-1}(m)\}$.

Soit Y_1 la σ-orbite de 1. Si $\mathrm{Card}(Y_1) = n$ alors $\sigma = (1\ \sigma(1)\ \sigma^2(1)\ \cdots\ \sigma^{n-1}(1))$ et σ est un n-cycle. Si $\mathrm{Card}(Y_1) < n$, alors on choisit $m \in \{1, 2 \dots, n\} \setminus Y_1$ et on pose Y_m la σ-orbite de m. On note que $Y_1 \cap Y_m = \emptyset$, car par la première étape de la preuve, $Y_1 = \{1, \sigma(1), \dots, \sigma^{k-1}(1)\}$ où $\sigma^k(1) = 1$, et $Y_m = \{m, \sigma(m), \dots, \sigma^{\ell-1}(m)\}$ où $\sigma^\ell(m) = m$. Si $\sigma^r(1) = \sigma^s(m)$ pour $0 \leq r < k$ et $0 \leq s < \ell$, alors $\sigma^{r-s}(1) = m$ et $m \in Y_1$, ce qui est en contradiction avec le choix de m.

Conclusion : Maintenant, on décompose $\{1, 2, \dots, n\}$ en une réunion disjointe de σ-orbites, $Y_1 \cup \cdots \cup Y_t$ avec Y_1 l'orbite de 1. Alors si $Y_i = \{i, \sigma(i), \dots, \sigma^{k_i-1}(i)\}$ avec $|Y_i| = k_i$, on vérifie par l'action que

$$\sigma = (a_1\ \sigma(a_1)\ \cdots\ \sigma^{k_1-1}(a_1))(a_2\ \sigma(a_2)\ \cdots\ \sigma^{k_2-1}(a_2)) \cdots (a_t\ \sigma(a_t)\ \cdots\ \sigma^{k_t-1}(a_t)).$$

Comme remarqué auparavant, les cycles disjoints commutent entre eux et donc l'écriture n'est pas unique, mais est unique à l'ordre près des cycles. En effet, les orbites de σ déterminent le support des cycles et l'action de σ sur chaque orbite détermine le cycle.

Enfin, si $Y_i = \{i\}$ pour un certain i, on supprime le cycle (i) car la notation sous-entend que $\sigma(i) = i$ si i n'apparaît dans aucun cycle. □

Exemples 7.1.5. 1. Considérons la permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & \cdots & n \\ 5 & 3 & 2 & 4 & 1 & 6 & \cdots & n \end{pmatrix}$ dans le groupe symétrique S_n pour $n \geq 5$. Son écriture en cycles disjoints de longueur au moins 2 est $(1\ 5)(2\ 3)$.

2. L'écriture de $\sigma = (1\ 2\ 3)(3\ 4\ 7)(7\ 8)(6\ 5\ 7) \in S_8$ en produit de cycles disjoints de longueur au moins 2 est
$$(1\ 2\ 3\ 4\ 7\ 6\ 5\ 8).$$

3. L'écriture de $\tau = (1\ 3\ 5\ 7)(2\ 3\ 7\ 1)(2\ 3) \in S_8$ en produit de cycles disjoints de longueur au moins 2 est $(1\ 2)(3\ 5\ 7)$. Dans chaque cas, on peut vérifier l'action sur tous les nombres entre 1 et 8.

Proposition 7.1.6. *Chaque $\sigma \in S_n$ s'écrit comme un produit de transpositions.*

Preuve. Si $n = 1$, alors $S_n = \{(1)\}$ et l'élément neutre est le produit vide. Supposons $n \geq 2$. Par Proposition 7.1.4, il suffit de montrer que chaque r-cycle s'écrit comme un produit de transpositions. Il y a plusieurs façons de le faire:

$$(a_1\ a_2\ \cdots\ a_r) = (a_1\ a_r)(a_1\ a_{r-1}) \cdots (a_1\ a_2),$$

ou

$$(a_1\ a_2\ \cdots\ a_r) = (a_1\ a_2)(a_2\ a_3) \cdots (a_{r-1}\ a_r).$$

□

7.1.2. *La signature d'une permutation.* Tout d'abord, on montrera que si $\sigma \in S_n$ s'écrit comme un produit de m transpositions ainsi que comme un produit de ℓ transpositions, alors m et ℓ ont la même parité, c'est-à-dire que soit m et ℓ sont les deux paires, soit les deux impaires.

On commence par un lemme qu'on vérifie directement en comparant les images des nombres $\{1, 2, \ldots, n\}$ sous l'action de chaque permutation.

Lemme 7.1.7. *Soient $h, k \in \mathbb{Z}$, $h, k \geq 0$ et $a, b, c_1, \ldots, c_h, d_1, \ldots, d_k \in \{1, 2 \ldots, n\}$ distincts. Alors*

$$(a\ b)(a\ c_1\ \cdots\ c_h\ b\ d_1\ \cdots\ d_k) = (b\ d_1\ \cdots\ d_k)(a\ c_1\ c_2 \cdots\ c_h). \tag{3}$$

Preuve. Exercice. □

Soit $\sigma \in S_n$, $\sigma \neq (1)$. On écrit

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_t, \tag{4}$$

un produit de cycles disjoints chacun de longueur au moins 2. On suppose que σ_i est un r_i cycle pour $1 \leq i \leq t$, $r_i \geq 2$. On définit une application $N : S_n \to \mathbb{N}$ par $N((1)) = 0$ et pour σ comme dans (4),

$$N(\sigma) = r_1 - 1 + r_2 - 1 + \cdots + r_t - 1.$$

Cette application est bien définie par l'unicité (à l'ordre des facteurs près) de la factorisation en produit de cycles disjoints.

Pour a, b, c_i, d_j comme dans Lemme 7.1.7,

$$N((b\ d_1\ \cdots\ d_k)(a\ c_1\ c_2 \cdots\ c_h)) = h + k$$

et

$$N((a\ c_1\ \cdots\ c_h\ b\ d_1\ \cdots\ d_k)) = h + k + 1.$$

Ainsi, par l'égalité (3) du Lemme 7.1.7,

$$N((ab)\sigma) = \begin{cases} N(\sigma) - 1 & \text{si } \{a, b\} \subseteq \operatorname{supp}(\sigma_i) \text{ pour un certain } i \\ N(\sigma) + 1 & \text{si } a \in \operatorname{supp}(\sigma_i), b \in \operatorname{supp}(\sigma_j), i \neq j \\ N(\sigma) + 1 & \text{si } \{a, b\} \cap \operatorname{supp}(\sigma) = \emptyset \end{cases} \tag{5}$$

Proposition 7.1.8. *Si $\sigma \in S_n$ s'écrit comme un produit de m transpositions pour $m \geq 1$, alors $N(\sigma)$ et m ont la même parité.*

Preuve. On écrit $\sigma = \tau_1 \cdots \tau_m$, τ_i une transposition pour tout i. On procède par récurrence sur m. Si $m = 1$, alors $\sigma = (ab)$ pour $1 \leq a, b \leq n$ et $N(\sigma) = 1$ par définition et le résultat est vérifié. Maintenant supposons que $m \geq 2$ et que le résultat est vrai pour tout produit de moins de m transpositions. Alors $N(\tau_1 \cdots \tau_m) = N(\tau_2 \cdots \tau_m) \pm 1$, par (5). Par l'hypothèse de récurrence, $N(\tau_2 \cdots \tau_m)$ et $m - 1$ ont la même parité. Ainsi on trouve que $N(\sigma)$ et m ont aussi la même parité comme énoncé. □

Ce résultat nous permet de poser la définition suivante:

Définition 7.1.9. Soit $\sigma \in S_n$.

1. On dit que σ est *paire* si σ s'écrit comme un produit d'un nombre pair de transpositions, c'est-à-dire que $N(\sigma)$ est pair. On dit que σ est *impaire* si σ s'écrit comme un produit d'un nombre impair de transpositions, c'est-à-dire que $N(\sigma)$ est impair.
2. La *signature de* σ, notée $\varepsilon(\sigma)$, est égale à $(-1)^{N(\sigma)}$, soit 1 si σ est paire, -1 si σ est impaire.

On obtient donc

Proposition 7.1.10. *L'application* $\varepsilon : S_n \to (\{1, -1\}, \cdot)$ *est un morphisme de groupes.*

Preuve. Soit $\sigma, \tau \in S_n$. Si σ et τ sont les deux paires ou les deux impaires alors $\varepsilon(\sigma) = \varepsilon(\tau)$, $\sigma\tau$ est une permutation paire et donc $\varepsilon(\sigma\tau) = 1 = \varepsilon(\sigma)^2 = \varepsilon(\sigma)\varepsilon(\tau)$.

Si l'une des deux est paire et l'autre est impaire, alors $\sigma\tau$ est impaire et $\varepsilon(\sigma\tau) = -1 = \varepsilon(\sigma)\varepsilon(\tau)$. □

7.2. **Applications multilinéaires.**

Définition 7.2.1. Soient V et W des K-espaces vectoriels. Une application $\phi : V \times \cdots \times V \to W$, du produit cartésien de m copies de V dans W, est dite *m-multilinéaire* si pour tous $1 \leq i \leq m$, $v_1, \ldots, v_m, u \in V$ et $\lambda \in K$ on a

$$\phi(v_1, \ldots, v_{i-1}, v_i + \lambda u, v_{i+1}, \ldots, v_m) = \phi(v_1, \ldots, v_m) + \lambda\phi(v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_m).$$

Autrement dit, ϕ est K-linéaire par rapport à chaque coordonnée.

Exemple 7.2.2 (produit scalaire usuel dans $\mathbb{R}^2$)**.** L'application $\beta : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ définie par $\beta(u, v) = u \cdot v := u_1v_1 + u_2v_2$ pour tous $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{R}^2$ est une application 2-linéaire. Dans ce cas on dit plutôt *bilinéaire*.

Remarque 7.2.3. On vérifie (exercice) que si ϕ est une application m-linéaire de V dans W, alors pour tous $1 \leq i \leq m$ et pour tout $v_j \in V$, on a

$$\phi(v_1, v_2, \ldots, v_{i-1}, 0, v_{i+1}, \ldots, v_m) = 0.$$

Voir Proposition 7.2.7(a).

On peut identifier le produit cartésien $K^n \times \cdots \times K^n$ de n copies de K^n avec l'ensemble $M_n(K)$. Un élément $(v_1, \ldots, v_n) \in K^n \times \cdots \times K^n$ est associé avec la matrice dont la i-ème ligne est le vecteur v_i. On a donc la définition suivante.

Définition 7.2.4. Une application $D : M_n(K) \to K$ est dite *n-linéaire (par rapport aux lignes)* si D est n-multilinéaire lorsqu'on identifie $M_n(K)$ avec $K^n \times \cdots \times K^n$ comme précédemment. Plus précisément, si $A \in M_n(K)$ est écrit $A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix}$ avec A_i la i-ème ligne de A, D est n-linéaire si pour tout A, on a

$$D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{i-1} \\ A_i + \mu B_i \\ A_{i+1} \\ \vdots \\ A_n \end{pmatrix}\right) = D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{i-1} \\ A_i \\ A_{i+1} \\ \vdots \\ A_n \end{pmatrix}\right) + \mu \cdot D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{i-1} \\ B_i \\ A_{i+1} \\ \vdots \\ A_n \end{pmatrix}\right),$$

pour tout $B_i \in K^n$ et $\mu \in K$.

Définition 7.2.5. On dit qu'une application n-linéaire $D : M_n(K) \to K$ est *alternée* si $D(A) = 0$ à chaque fois que deux lignes de la matrice A sont égales.

Exemple 7.2.6. L'application $D : M_2(K) \to K$ définie par $D\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = 2(ad - bc)$ est bilinéaire alternée.

Proposition 7.2.7. *Soit $D : M_n(K) \to K$ une application n-linéaire.*

(a) Si une ligne de A est nulle, alors $D(A) = 0$.

(b) Si D est alternée, alors on a $D(T_{ij}A) = -D(A)$ pour tout $1 \leq i < j \leq n$.

Preuve. On montre (b) en premier. On rappelle que $T_{ij}A$ est la matrice obtenue en échangeant les lignes i et j de la matrice A. On considère la matrice $C = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_i + A_j \\ \vdots \\ A_j + A_i \\ \vdots \\ A_n \end{pmatrix}$.
Comme il y a deux lignes égales et D est alternée, on a $D(C) = 0$. Mais on peut aussi développer en utilisant la multilinéarité:

$$0 = D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_i + A_j \\ \vdots \\ A_j + A_i \\ \vdots \\ A_n \end{pmatrix}\right) = D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_j + A_i \\ \vdots \\ A_n \end{pmatrix}\right) + D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_j \\ \vdots \\ A_j + A_i \\ \vdots \\ A_n \end{pmatrix}\right).$$

Et encore une fois :

$$0 = D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix}\right) + D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix}\right) + D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_j \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix}\right) + D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix}\right).$$

Mais comme D est alternée, les deux termes du milieu sont égaux à 0 et on a que

$$D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix}\right) = -D\left(\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_j \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix}\right).$$

L'énoncé de (a) est couvert par la Remarque 7.2.3.

□

7.3. **Le déterminant.**

Définition 7.3.1. Soit $A \in M_n(K)$, $A = (a_{ij})$. Le *déterminant* de A, noté $\det(A)$, est l'élément de K défini par

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

où $\varepsilon : S_n \to \{1, -1\}$ est la signature de σ. On écrit également

$$\det(A) = |A| = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Exemples 7.3.2. Les cas particuliers de $n = 1$, $n = 2$, $n = 3$ et la règle de Sarrus, qui peut être utilisée pour le cas $n = 3$, seront abordés en cours. Attention, si vous utilisez cette règle pour calculer $\det(A)$, il faut la citer.

Soit $A \in M_n(K)$. On note A_i la i-ème ligne de A, donc $A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix}$.

Théorème 7.3.3 (multilinéarité du déterminant)**.** *Le déterminant est une application n-linéaire (linéaire par rapport à chaque ligne) ; c'est-à-dire, pour $A_1, \dots, A_n, B_i \in K^n$ et $\mu \in K$,*

$$\det \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{i-1} \\ A_i + \mu B_i \\ A_{i+1} \\ \vdots \\ A_n \end{pmatrix} = \det \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{i-1} \\ A_i \\ A_{i+1} \\ \vdots \\ A_n \end{pmatrix} + \mu \cdot \det \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{i-1} \\ B_i \\ A_{i+1} \\ \vdots \\ A_n \end{pmatrix}.$$

Preuve. On a

$$\det\begin{pmatrix} A_1 \\ \vdots \\ A_i + \mu B \\ \vdots \\ A_n \end{pmatrix} = \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{i-1\sigma(i-1)} (a_{i\sigma(i)} + \mu b_{i\sigma(i)}) \cdots a_{n\sigma(n)}$$

$$= \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{i\sigma(i)} \cdots a_{n\sigma(n)} + \mu \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots b_{i\sigma(i)} \cdots a_{n\sigma(n)}$$

$$= \det\begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix} + \mu \cdot \det\begin{pmatrix} A_1 \\ \vdots \\ B_i \\ \vdots \\ A_n \end{pmatrix}.$$

□

Lemme 7.3.4 (Lemme fondamental)**.** *Le déterminant est une application n-linéaire alternée. C'est-à-dire, si deux lignes de A sont égales, alors* $\det(A) = 0$.

Preuve. Soient $1 \leq i < j \leq n$. Supposons que $A_i = A_j$, c'est-à-dire $a_{ik} = a_{jk}$ pour tout k. Soit $\tau = (ij)$ la transposition qui échange i et j et qui fixe tout autre élément de $\{1, 2 \ldots, n\}$. Soit $H \leq S_n$ l'ensemble des permutations paires dans S_n. On note que S_n est l'union disjointe de H et $H\tau$. En effet, si $\rho \in S_n$ est impaire alors $\rho\tau = \sigma \in H$ et donc $\rho = \sigma\tau \in H\tau$. On calcule le déterminant de A:

$$\det(A) = \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\sigma\in H} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} + \sum_{\gamma\in H\tau} \varepsilon(\gamma) a_{1\gamma(1)} \cdots a_{n\gamma(n)}$$

$$= \sum_{\sigma\in H} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} + \sum_{\sigma\in H} \varepsilon(\sigma\tau) a_{1\sigma\tau(1)} \cdots a_{i\sigma\tau(i)} \cdots a_{j\sigma\tau(j)} \cdots a_{n\sigma\tau(n)}$$

$$= \sum_{\sigma\in H} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} + \sum_{\sigma\in H} \varepsilon(\sigma)(-1) a_{1\sigma(1)} \cdots a_{i\sigma(j)} \cdots a_{j\sigma(i)} \cdots a_{n\sigma(n)}.$$

Mais comme $A_i = A_j$ par hypothèse, on a $a_{i\sigma(j)} = a_{j\sigma(j)}$ et $a_{j\sigma(i)} = a_{i\sigma(i)}$, donc les deux sommes s'annulent et on obtient 0. □

Proposition 7.3.5 (permutation de lignes)**.** *Soit $D : M_n(K) \to K$ une application n-linéaire alternée. $A = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} \in M_n(K)$ et $\sigma \in S_n$. Alors*

$$D\begin{pmatrix} A_{\sigma(1)} \\ \vdots \\ A_{\sigma(n)} \end{pmatrix} = \varepsilon(\sigma) \cdot D\begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix}.$$

Preuve. On écrit σ comme produit de $k \geq 0$ transpositions et on procède par récurrence sur k. Si $k = 0$, alors $\sigma = \mathrm{id}$ et le résultat est vérifié. Si $k = 1$, alors $\sigma = \tau$ est une transposition, donc $\varepsilon(\tau) = -1$ et la Proposition 7.2.7(b) donne le résultat. Supposons maintenant que $k > 1$ et que le résultat est vérifié pour un produit de $k-1$ transpositions. On a $\sigma = \tau_1 \cdots \tau_k = \gamma\tau_k$, où $\gamma = \tau_1 \cdots \tau_{k-1}$. On a alors

$$D\begin{pmatrix} A_{\sigma(1)} \\ A_{\sigma(2)} \\ \vdots \\ A_{\sigma(n)} \end{pmatrix} = D\begin{pmatrix} A_{\gamma\tau_k(1)} \\ A_{\gamma\tau_k(2)} \\ \vdots \\ A_{\gamma\tau_k(n)} \end{pmatrix} = \varepsilon(\gamma) \cdot D\begin{pmatrix} A_{\tau_k(1)} \\ A_{\tau_k(2)} \\ \vdots \\ A_{\tau_k(n)} \end{pmatrix},$$

par l'hypothèse de récurrence. Et par la Proposition 7.2.7(b), ce dernier est égal à

$$-\varepsilon(\gamma) \cdot D\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} = \varepsilon(\gamma)\varepsilon(\tau_k) \cdot D\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} = \varepsilon(\gamma\tau_k)D\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} = \varepsilon(\sigma)D\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix},$$

car ε est un homomorphisme de groupes. □

Proposition 7.3.6. *Soit $D : M_n(K) \to K$ une application n-linéaire. Pour tout $A \in M_n(K)$, $A_r \in K^n$, et $\lambda \in K$ on a*

(a) $D\begin{pmatrix} A_1 \\ \vdots \\ \lambda A_i \\ \vdots \\ A_n \end{pmatrix} = \lambda D\begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix}$ *; et*

(b) si D *est alternée, pour* $r \neq i$, $D\begin{pmatrix} A_1 \\ \vdots \\ \lambda A_r + A_i \\ \vdots \\ A_n \end{pmatrix} = D\begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix}$.

Preuve. La partie (a) suit directement de la linéarité. Pour (b), on développe en utilisant la linéarité :

$$D\begin{pmatrix} A_1 \\ \vdots \\ \lambda A_r + A_i \\ \vdots \\ A_n \end{pmatrix} = \lambda D\begin{pmatrix} A_1 \\ \vdots \\ A_r \\ \vdots \\ A_r \\ \vdots \\ A_n \end{pmatrix} + D\begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} = D\begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix},$$

où le premier terme de la somme vaut 0 car deux lignes sont égales. □

Théorème 7.3.7 (déterminant d'une matrice triangulaire)**.** *Si* $A = (a_{ij})$ *est triangulaire supérieure, c'est-à-dire,* $a_{ij} = 0$ *pour tout* $i > j$*, alors* $\det(A) = a_{11}a_{22}\cdots a_{nn}$ *(également vrai pour les matrices triangulaires inférieures).*

Preuve. On a $\det(A) = \sum_{\sigma\in S_n} \varepsilon(\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)}$. On montre que tous les termes de la somme sont nuls sauf le terme $a_{11}\cdots a_{nn}$, où $\sigma = \mathrm{id}$.

Supposons donc que $a_{1\sigma(1)}\cdots a_{n\sigma(n)} \neq 0$ pour un $\sigma \in S_n$. Alors

$$a_{j\sigma(j)} \neq 0 \text{ pour tout } j. \tag{6}$$

Comme dans une matrice triangulaire supérieure $a_{jk} = 0$ pour tout $k < j$, on a

$$\sigma(j) \geq j \text{ pour tout } j. \tag{7}$$

En particulier $\sigma(n) = n$. Donc σ permute l'ensemble $\{1, 2, \ldots, n-1\}$ et par (7), $\sigma(n-1) = n-1$. Par récurrence sur k, on montre que $\sigma(n-k) = n-k$ pour tout $k \geq 0$, ce qui implique $\sigma(i) = i$ pour tout i, et on déduit que $\sigma = \mathrm{id}$. Donc $\det(A) = \varepsilon(\mathrm{id}) a_{11} \cdots a_{nn} = a_{11} a_{22} \cdots a_{nn}$. □

Proposition 7.3.8. *Pour tout $\lambda \in K$, $1 \leq r \neq s \leq n$, on a*

(a) $\det(I_n) = 1$.

(b) $\det(T_{rs}) = -1$.

(c) $\det(D_r(\lambda)) = \lambda$.

(d) $\det(L_{rs}(\lambda)) = 1$.

Preuve. C'est une conséquence des résultats 7.3.7, 7.3.4, 7.2.7 et 7.3.6. □

7.4. **Unicité du déterminant.**

Théorème 7.4.1. *Soit $D : M_n(K) \to K$ une application n-linéaire alternée. Alors il existe $d \in K$ tel que $D(A) = d \cdot \det(A)$ pour tout $A \in M_n(K)$. De plus $d = D(I_n)$.*

Preuve. Soit $A = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix}$ et $I_n = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$ où e_i est le i-ème vecteur de la base canonique de K^n. On a $A_i = \sum_{j=1}^{n} a_{ij} e_j$. On calcule $D(A)$, utilisant la n-linéarité de D:

$$D(A) = D \begin{pmatrix} \sum_{j_1=1}^{n} a_{1j_1} e_{j_1} \\ \vdots \\ \sum_{j_n=1}^{n} a_{nj_n} e_{j_n} \end{pmatrix} = \sum_{j_1=1}^{n} a_{1j_1} D \begin{pmatrix} e_{j_1} \\ \sum_{j_2=1}^{n} a_{2j_2} e_{j_2} \\ \vdots \\ \sum_{j_n=1}^{n} a_{nj_n} e_{j_n} \end{pmatrix} = \cdots$$

$$= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} D \begin{pmatrix} e_{j_1} \\ e_{j_2} \\ \vdots \\ e_{j_n} \end{pmatrix}.$$

Comme D est alternée, si $j_k = j_\ell$ alors $D \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix} = 0$. Donc les seuls termes non nuls de la somme sont ceux avec $\{j_1, \ldots, j_n\} = \{1, \ldots, n\}$. En plus, les n-uplets $(j_1, \ldots, j_n)$ avec coordonnées distinctes forment un ensemble complet et sans répétition des permutations de l'ensemble $\{1, \ldots, n\}$. La somme est donc égale à $\sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} D \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$.

Ce dernier est égal à $\sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \cdot \varepsilon(\sigma) \cdot D \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$, par la Proposition 7.3.5. Donc, on trouve $D(A) = \det(A) \cdot D(I_n)$ comme énoncé. □

7.5. **Multiplicativité du déterminant.**

Théorème 7.5.1. *Soient $A, B \in M_n(K)$. Alors*

(a) $\det(AB) = \det(A)\det(B)$, *et*

(b) si A est inversible, alors $\det(A) \neq 0$ *et* $\det(A^{-1}) = (\det(A))^{-1}$.

Preuve. Fixons $B \in M_n(K)$. On définit une application $D_B : M_n(K) \to K$ par $D_B(A) = \det(AB)$. On vérifie d'abord que D_B est une application n-linéaire alternée. Pour ceci, on note que la matrice $A_i B$ est une matrice $1 \times n$, égale à la i-ème ligne de la matrice

AB. Soient $\lambda \in K$ et $C \in K^n$. On a

$$D_B\begin{pmatrix} A_1 \\ \vdots \\ A_i + \lambda C \\ \vdots \\ A_n \end{pmatrix} = \det\left(\begin{pmatrix} A_1 \\ \vdots \\ A_i + \lambda C \\ \vdots \\ A_n \end{pmatrix} \cdot B\right) = \det\begin{pmatrix} A_1B \\ \vdots \\ (A_i + \lambda C)B \\ \vdots \\ A_nB \end{pmatrix}$$

$$= \det\begin{pmatrix} A_1B \\ \vdots \\ A_iB \\ \vdots \\ A_nB \end{pmatrix} + \lambda\det\begin{pmatrix} A_1B \\ \vdots \\ CB \\ \vdots \\ A_nB \end{pmatrix} = D_B(A) + \lambda D_B\begin{pmatrix} A_1 \\ \vdots \\ C \\ \vdots \\ A_n \end{pmatrix}.$$

Donc D_B est n-linéaire.

Aussi soit $A \in M_n(K)$ telle que $A_r = A_s$ pour $1 \leq r < s \leq n$. On a

$$D_B(A) = \det(AB) = \det\begin{pmatrix} A_1B \\ \vdots \\ A_rB \\ \vdots \\ A_sB \\ \vdots \\ A_nB \end{pmatrix} = 0$$

car $A_rB = A_sB$. On a donc que D_B est aussi alternée. Par le théorème de l'unicité du déterminant, on a $D_B(A) = D_B(I_n)\det(A)$ pour tout $A \in M_n(K)$. Donc $\det(AB) = D_B(A) = D_B(I_n)\det(A) = \det(B)\det(A) = \det(A)\det(B)$.

Pour (2): on suppose que A soit inversible avec inverse A^{-1}. On a $1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$, par la partie (1). Cette égalité montre que $\det(A) \neq 0$. Aussi, la même égalité montre que $\det(A^{-1}) = (\det(A))^{-1}$. □

Proposition 7.5.2 (Critère d'inversibilité)**.** *Soit $A \in M_n(K)$. Alors A est inversible si et seulement si* $\det(A) \neq 0$.

Preuve. La direction $\Longrightarrow$ est le théorème précédent.

Supposons maintenant que $\det(A) \neq 0$. Soit R une matrice échelonnée réduite ligne équivalente à A. Donc R est une matrice triangulaire supérieure de déterminant $r_{11} \cdots r_{nn}$.

Aussi il existe des matrices élémentaires $E_1, \ldots, E_t$ telles que $R = E_1 \cdots E_t A$. Par la multiplicativité du déterminant on a $\det(R) = \det(E_1 \cdots E_t)\det(A)$. Comme $E_1 \cdots E_t$ est inversible, son déterminant est non nul et par hypothèse $\det(A) \neq 0$. D'où $\det(R) \neq 0$. Donc R possède n pivots, et par conséquent est de rang n, de même que A. Par le Théorème 6.5.1, A est inversible. □

Corollaire 7.5.3. *(a) L'application* $\det : \mathrm{GL}_n(K) \to K \setminus \{0\}$ *est un morphisme de groupes, où on munit $K \setminus \{0\}$ de la loi de composition de multiplication.*

(b) Soient $A, B \in M_n(K)$ des matrices semblables. Alors $\det(A) = \det(B)$.

Preuve. Faites en cours. □

Ce dernier résultat nous permet de définir le déterminant d'une application linéaire d'un K-espace vectoriel de dimension finie.

Définition 7.5.4. (1) Soit V un K-espace vectoriel de dimension n et $\phi \in \mathcal{L}(V, V)$. On définit $\det(\phi)$ comme suit: on choisit une base C de V et on pose $\det(\phi) = \det((\phi)_C^C)$. Le corollaire précédent montre que la valeur est indépendante du choix de la base C.

(2) Le noyau de l'application $\det : \mathrm{GL}_n(K) \to K \setminus \{0\}$ s'appelle le groupe linéaire spécial et est noté $\mathrm{SL}_n(K) := \ker(\det) = \{A \in \mathrm{GL}_n(K) \mid \det(A) = 1\}$.

7.6. **La transposée.**

Théorème 7.6.1. *Soit $A \in M_n(K)$. Alors* $\det(A^t) = \det(A)$.

Preuve. Nous avons

$$\det(A^t) = \sum_{\sigma \in S_n} \varepsilon(\sigma)(A^t)_{1\sigma(1)} \cdots (A^t)_{n\sigma(n)} = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

Or $\sigma(i) = j$ si et seulement si $\sigma^{-1}(j) = i$. Donc $a_{\sigma(i)i} = a_{j\sigma^{-1}(j)}$. Aussi la somme sur $\sigma \in S_n$ est la même que la somme sur $\sigma^{-1} \in S_n$. (L'application $\sigma \mapsto \sigma^{-1}$ est une bijection de $S_n \to S_n$). Finalement, on note que $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$.

Donc on a que

$$\begin{aligned}\det(A^t) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma^{-1}) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} = \sum_{\tau \in S_n} \varepsilon(\tau) a_{1\tau(1)} \cdots a_{n\tau(n)} = \det(A).\end{aligned}$$

□

Corollaire 7.6.2. *Toutes les propriétés des déterminants relatives aux lignes sont aussi valables pour les colonnes.*

(a) *Le déterminant est linéaire par rapport à chaque colonne.*

(b) *Si une colonne est nulle, alors* $\det(A) = 0$.

(c) *Si deux colonnes sont égales, alors* $\det(A) = 0$.

(d) *Si on effectue une permutation* σ *des colonnes de* A*, le déterminant de la matrice résultante est égal à* $\varepsilon(\sigma) \cdot \det(A)$.

(e) *Si on multiplie une colonne de* A *par* $\lambda \in K$*, le déterminant est multiplié par* λ.

(f) *Si on additionne à une colonne un multiple scalaire d'une autre colonne, le déterminant ne change pas.*

7.7. **Cofacteurs.**

Définition 7.7.1. Soit $A \in M_n(K)$. On suppose $n \geq 2$.

(a) On pose $A(r|s)$ la matrice dans $M_{n-1}(K)$ obtenue à partir de A en supprimant la r-ème ligne de A et la s-ème colonne de A.

(b) $\det(A(r|s))$ s'appelle un *mineur de* A *d'ordre* $n-1$.

(c) $(-1)^{r+s}\det(A(r|s))$ s'appelle le *cofacteur du coefficient* A_{rs}.

Théorème 7.7.2 (développement par rapport à la r-ème ligne de A)**.** *On fixe* r*,* $1 \leq r \leq n$*.*

$$\det(A) = \sum_{j=1}^{n} A_{rj}(-1)^{r+j}\det(A(r|j)).$$

Preuve. Par définition $\det(A) = \sum_{\sigma\in S_n} \varepsilon(\sigma)A_{1\sigma(1)}\cdots A_{r\sigma(r)}\cdots A_{n\sigma(n)}$. On réécrit la somme :

$$\begin{aligned}\det(A) &= A_{r1}\sum_{\sigma\in S_n,\sigma(r)=1} \varepsilon(\sigma)A_{1\sigma(1)}\cdots A_{r-1,\sigma(r-1)}A_{r+1,\sigma(r+1)}\cdots A_{n\sigma(n)}\\ &+A_{r2}\sum_{\sigma\in S_n,\sigma(r)=2} \varepsilon(\sigma)A_{1\sigma(1)}\cdots A_{r-1,\sigma(r-1)}A_{r+1,\sigma(r+1)}\cdots A_{n\sigma(n)}\\ &\vdots\\ &+A_{rn}\sum_{\sigma\in S_n,\sigma(r)=n} \varepsilon(\sigma)A_{1\sigma(1)}\cdots A_{r-1,\sigma(r-1)}A_{r+1,\sigma(r+1)}\cdots A_{n\sigma(n)}\\ &=\sum_{j=1}^{n} A_{rj}\Big(\sum_{\sigma\in S_n,\sigma(r)=j} \varepsilon(\sigma)A_{1\sigma(1)}\cdots A_{r-1,\sigma(r-1)}A_{r+1,\sigma(r+1)}\cdots A_{n\sigma(n)}\Big).\end{aligned}$$

On pose

$$A'_{rj} = \sum_{\sigma\in S_n,\sigma(r)=j} \varepsilon(\sigma)A_{1\sigma(1)}\cdots A_{r-1,\sigma(r-1)}A_{r+1,\sigma(r+1)}\cdots A_{n\sigma(n)}.$$

Il faut montrer que

$$A'_{rj} = (-1)^{r+j}\det(A(r|j)).$$

Cas 1. $r = j = 1$.

Soit $H \leq S_n$ un sous-groupe tel que $H = \{\sigma \in S_n \mid \sigma(1) = 1\}$. Alors on identifie H naturellement avec le groupe des permutations de l'ensemble $\{2,\ldots,n\}$, ce qui est également identifié avec le groupe S_{n-1}.

Dans ce cas nous avons

$$\begin{aligned}A'_{11} &= \sum_{\sigma\in S_n,\sigma(1)=1} \varepsilon(\sigma)A_{2\sigma(2)}\cdots A_{n\sigma(n)} = \sum_{\sigma\in H} \varepsilon(\sigma)A_{2\sigma(2)}\cdots A_{n\sigma(n)}\\ &= \det A(1|1) = (-1)^{1+1}\det A(1|1),\end{aligned}$$

comme affirmé.

Cas 2. r et j quelconques.

Soit B la matrice obtenue à partir de A en remplaçant la r-ème ligne par la ligne $(0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0)$, où le coefficient 1 est à la j-ème place. Donc précisément nous avons

$$B_{k\ell} = \begin{cases} A_{k\ell} & \text{si } k \neq r \\ 0 & \text{si } k = r, \ell \neq j \\ 1 & \text{si } k = r, \ell = j \end{cases}.$$

Si on calcule B'_{rj} les coefficients de la ligne r n'apparaissent pas dans la somme et donc

$$B'_{rj} = A'_{rj}. \tag{8}$$

Aussi

$$B(r|j) = A(r|j), \tag{9}$$

car on supprime la ligne qui est différente.

Maintenant, $\det(B) = \sum_{k=1}^{n} B_{rk} B'_{rk} = 1 \cdot B'_{rj}$, car $B_{rk} = 0$ si $k \neq j$. Donc

$$B'_{rj} = \det(B). \tag{10}$$

Soit maintenant C la matrice obtenue à partir de B en permutant cycliquement les r premières lignes (i.e. on permute les lignes de B selon le r-cycle $(1 \ 2 \ \cdots \ r)$) et ensuite cycliquement les j premières colonnes (selon le j-cycle $(1 \ 2 \ \cdots \ j)$).

On rappelle que $\varepsilon((1 \ 2 \ \cdots \ s)) = (-1)^{s+1}$. Par conséquent, $\det(C) = (-1)^{r+1}(-1)^{j+1}\det(B)$, d'où

$$\det(C) = (-1)^{r+j}\det(B). \tag{11}$$

La r-ème ligne de B étant la première ligne de C et la j-ème colonne de B étant la première colonne de C, donc

$$C(1|1) = B(r|j) \tag{12}$$

On calcule maintenant $\det(C)$ en développant par rapport à la première ligne: $\det(C) = \sum_{k=1}^{n} C_{1k} C'_{1k} = C'_{11}$, car C_{1k} est 0 si $k \neq 1$. Et par le Cas 1, déjà traité ci-dessus,

$C'_{11} = \det(C(1|1))$. Donc nous avons

$$\det(C) = \det(C(1|1)). \tag{13}$$

On peut maintenant conclure:

$$A'_{rj} = B'_{rj} = \det(B) = (-1)^{r+j}\det(C) = (-1)^{r+j}\det(C(1|1))$$
$$= (-1)^{r+j}\det(B(r|j)) = (-1)^{r+j}\det(A(r|j)),$$

où la première égalité suit de l'égalité (8), la deuxième de l'égalité (10), la troisième de l'égalité (11), la quatrième de l'égalité (13), la cinquième de l'égalité (12), et la dernière de l'égalité (9). □

Théorème 7.7.3 (développement par rapport à une colonne)**.** *On fixe s tel que $1 \leq s \leq n$. Alors,*

$$\det(A) = \sum_{i=1}^{n} A_{is}(-1)^{i+s}\det(A(i|s)).$$

Preuve. Par Théorème 7.6.1, on a que

$$\det(A) = \det(A^t) = \sum_{j=1}^{n}(A^t)_{sj}(-1)^{s+j}\det(A^t(s|j)) = \sum_{j=1}^{n} A_{js}(-1)^{s+j}\det(A(j|s))$$

. □

Définition 7.7.4. Soit $A = (a_{ij}) \in M_n(K)$, $n \geq 2$. La *matrice des cofacteurs de A* est la matrice $\mathrm{cof}(A)$ formée des cofacteurs de la matrice A :

$$(\mathrm{cof}(A))_{ij} = (-1)^{i+j}\det(A(i|j)).$$

Remarque 7.7.5. $\mathrm{cof}(A^t) = (\mathrm{cof}(A))^t$, car

$$(\mathrm{cof}(A^t))_{ij} = (-1)^{i+j}\det(A^t(i|j)) = (-1)^{i+j}\det A(j|i) = (\mathrm{cof}(A))_{ji}.$$

Théorème 7.7.6 (La matrice des cofacteurs)**.** *Soit $A = (a_{ij}) \in M_n(K)$, $n \geq 2$. Alors*

$$A \cdot \mathrm{cof}(A)^t = \det(A)I_n = \mathrm{cof}(A)^t \cdot A.$$

Preuve. D'abord on calcule les coefficients du produit $A \cdot \mathrm{cof}(A)^t$ qui apparaissent le long de la diagonale du produit $A \cdot \mathrm{cof}(A)^t$:

$$(A \cdot \mathrm{cof}(A)^t)_{ii} = \sum_{k=1}^{n} A_{ik}(\mathrm{cof}(A)^t)_{ki} = \sum_{k=1}^{n} A_{ik}(\mathrm{cof}(A))_{ik} = \sum_{k=1}^{n} A_{ik}(-1)^{i+k}\det(A(i|k)) = \det(A),$$

par le Théorème 7.7.2.

Maintenant, on calcule les autres coefficients $(A \cdot (\mathrm{cof}A)^t)_{k\ell}$ pour $k \neq \ell$.

On définit une nouvelle matrice B comme suit:

On remplace la ℓ-ème ligne de A par la k-ème ligne de A. De ce fait, $\det(B) = 0$ et $B = (b_{ij})$ avec $b_{ij} = a_{ij}$ si $i \neq \ell$ et $b_{\ell j} = a_{kj}$ pour tout $1 \leq j \leq n$.

En termes des lignes nous avons $A = \begin{pmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_\ell \\ \vdots \\ A_n \end{pmatrix}$ et $B = \begin{pmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_k \\ \vdots \\ A_n \end{pmatrix}$. Ainsi, nous avons aussi l'égalité $B(\ell|j) = A(\ell|j)$ pour tout j.

Maintenant, on calcule $\det(B)$ en développant le long de la ℓ-ème ligne:

$$0 = \det(B) = \sum_{j=1}^{n} b_{\ell j}(-1)^{\ell+j}\det(B(\ell|j)) = \sum_{j=1}^{n} a_{kj}(-1)^{\ell+j}\det(A(\ell|j))$$

$$= \sum_{j=1}^{n} a_{kj}(\mathrm{cof}A)_{\ell j} = \sum_{j=1}^{n} a_{kj}((\mathrm{cof}(A)^t)_{j\ell} = (A \cdot \mathrm{cof}(A))^t)_{k\ell}.$$

Nous avons établi l'égalité matricielle $A \cdot (\mathrm{cof}(A))^t = \det(A) \cdot I_n$. Pour $(\mathrm{cof}(A))^t \cdot A$, on applique le cas précédent à A^t et on utilise le fait montré en exercices que $(AB)^t = B^t A^t$, et le fait que $\det(A^t) = \det(A)$. □

Corollaire 7.7.7. *Si A est inversible, $A^{-1} = \frac{1}{\det(A)} \cdot \mathrm{cof}(A)^t$.*

On fixe un corps K.

8.1. **Vecteurs propres et valeurs propres.** Ici, on étudie l'anneau unitaire $\mathcal{L}(V, V)$ et on rappelle que dans le cas d'un K-espace vectoriel de dimension finie n, cet anneau est isomorphe à l'anneau $M_n(K)$ (voir Corollaire 5.3.10).

Définition 8.1.1. Une *transformation linéaire d'un K-espace vectoriel V* est une application K-linéaire de V dans V, c'est-à-dire un élément de $\mathcal{L}(V, V)$. On dit aussi *un opérateur linéaire de V* ou un *endomorphisme de V*.

Les matrices qui sont les plus faciles à "manipuler" algébriquement sont les matrices diagonales. Ensuite, les matrices triangulaires partagent quelques propriétés utiles aussi (par exemple, facilité pour le calcul du déterminant ou du rang). Soit $A \in M_n(K)$ une matrice diagonale ou triangulaire supérieure. Alors, par l'isomorphisme entre $M_n(K)$ et $\mathcal{L}(K^n, K^n)$ associé au choix de la base canonique $C = (e_1, \ldots, e_n)$ de K^n, il existe une unique $\phi \in \mathcal{L}(K^n, K^n)$ avec $(\phi)_C^C = A$. Comme A est triangulaire supérieure ou diagonale, on note que $\phi(e_1) = a_{11}e_1$; cette observation motive la définition suivante.

Définition 8.1.2. Soit $\phi : V \to V$ une transformation linéaire d'un K-espace vectoriel V.

(1) On dit que $v \in V$ est un *vecteur propre de ϕ* si
   - $v \neq 0$, et
   - $\phi(v)$ est un multiple scalaire de v.

   Plus précisément, v est un vecteur propre de ϕ si $v \neq 0$, et qu'il existe $\lambda \in K$ tel que $\phi(v) = \lambda v$.
(2) Le scalaire λ s'appelle la *valeur propre de ϕ associée au vecteur propre v*.
(3) L'ensemble des valeurs propres de ϕ s'appelle le *spectre de ϕ*.

On a la notion analogue pour les matrices $A \in M_n(K)$.

Définition 8.1.3. Soit $A \in M_n(K)$.

On dit qu'un vecteur colonne $v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in M_{n\times 1}(K)$ est un *vecteur propre pour* A si $v \neq 0$, et qu'il existe $\lambda \in K$ tel que $Av = \lambda v$. On appelle λ la *valeur propre de* A *associée au vecteur propre* v.

Remarque 8.1.4. (1) Un vecteur propre est par définition non nul, mais la valeur propre associée à un vecteur propre peut être nulle.

(2) Si 0 est une valeur propre pour $\phi \in \mathcal{L}(V, V)$, alors un vecteur propre de valeur propre 0 est un vecteur $v \in V$, $v \neq 0$ tel que $\phi(v) = 0$. Donc $v \in \ker(\phi)$. On déduit que 0 est une valeur propre de ϕ si et seulement si ϕ est non injective.

(3) Si v, w sont des vecteurs propres de ϕ, les deux de valeur propre λ, alors pour tout $\mu \in K$, le vecteur $\mu v + w$ est soit égal à 0, soit un vecteur propre de ϕ, de valeur propre λ. En effet, on a que $\phi(\mu v + w) = \mu\phi(v) + \phi(w) = \mu(\lambda v) + \lambda w = \lambda(\mu v + w)$.

Proposition 8.1.5. *Soit* $A \in M_n(K)$. *Alors* A *est diagonale si et seulement si* $e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, *le vecteur colonne avec* 1 *à la* i*-ème coordonnée et* 0 *ailleurs, est un vecteur propre de* A *pour tout* $1 \leq i \leq n$.

Proof. Ce résultat découle directement du fait que Ae_i est la i-ème colonne de A. □

8.2. **Matrices et transformations diagonalisables et trigonalisables.** D'abord rappelons que des matrices $A, B \in M_n(K)$ sont semblables s'il existe $P \in \mathrm{GL}_n(K)$ telle que $B = P^{-1}AP$.

Définition 8.2.1. (1) On dit que $A \in M_n(K)$ est *diagonalisable* si A est semblable à une matrice diagonale.

(2) On dit que $A \in M_n(K)$ est *trigonalisable*, si A est semblable à une matrice triangulaire.

Remarque 8.2.2. On note que chaque matrice triangulaire supérieure est semblable à une matrice triangulaire inférieure, et de même, que chaque matrice triangulaire inférieure est semblable à une matrice triangulaire supérieure. En effet, si A est triangulaire supérieure, on pose B la base ordonnée de K^n, $B = (e_n, e_{n-1}, \ldots, e_1)$, où $C = (e_1, \ldots, e_n)$ est la base canonique de K^n. On vérifie que pour $P = (\mathrm{id})_C^B$, on a que PAP^{-1} est triangulaire inférieure.

Maintenant, on considère les transformations linéaires.

Définition 8.2.3. Soit V un K-espace vectoriel de dimension finie et soit $\phi \in \mathcal{L}(V, V)$. On dit que ϕ est *diagonalisable* s'il existe une base de V formée de vecteurs propres de ϕ.

Théorème 8.2.4. *Soit $\phi : V \to V$ une transformation linéaire d'un K-espace vectoriel de dimension finie n. Soit B une base de V. Alors ϕ est diagonalisable si et seulement si la matrice $(\phi)_B^B$ est diagonalisable.*

Proof. On suppose d'abord que ϕ est diagonalisable. Par définition, il existe une base F de V formée de vecteurs propres pour ϕ. Posons $F = (f_1, \ldots, f_n)$, et soit λ_i la valeur propre associée au vecteur propre f_i pour chaque i. Comme $\phi(f_i) = \lambda_i f_i$, la i-ème colonne de la matrice $(\phi)_F^F$ est le vecteur colonne avec λ_i à la i-ème coordonnée et 0 ailleurs. Donc, $(\phi)_F^F$ est une matrice diagonale (avec les valeurs propres $\lambda_1, \ldots, \lambda_n$ le long de la diagonale). Finalement, $(\phi)_B^B = Q^{-1}(\phi)_F^F Q$, avec $Q = (\mathrm{id})_B^F$, et donc $(\phi)_B^B$ est diagonalisable.

On suppose maintenant que $(\phi)_B^B$ est diagonalisable. Il existe donc $P \in \mathrm{GL}_n(K)$ telle que $P^{-1}(\phi)_B^B P$ est une matrice diagonale. Comme P est inversible, on a $P = (\mathrm{id})_E^B$ pour un certain choix de base $E = (w_1, \ldots, w_n)$ de V, et $(\phi)_E^E = (\mathrm{id})_B^E (\phi)_B^B (\mathrm{id})_E^B = P^{-1}(\phi)_B^B P$ qui est par hypothèse une matrice diagonale, ce qui implique que $\phi(w_i) = \lambda_i w_i$ pour un certain $\lambda_i \in K$. Donc, la base E est une base de V formée de vecteurs propres de ϕ et ϕ est diagonalisable. □

Définition 8.2.5. Soient V un K-espace vectoriel et $\phi \in \mathcal{L}(V, V)$.

(1) Un sous-espace vectoriel W de V est dit *stable par* ϕ ou *ϕ-invariant* si pour tout $w \in W$ on a que $\phi(w) \in W$.

(2) On suppose maintenant que V est de dimension finie n. On dit que ϕ est *trigonalisable*) s'il existe des sous-espaces vectoriels ϕ-invariants $W_0, \ldots, W_n$ de V, avec $\{0\} = W_0 \subset W_1 \subset \cdots \subset W_{n-1} \subset W_n = V$ et $\dim W_i = i$ pour tout i

Proposition 8.2.6. *Soit V un K-espace vectoriel de dimension finie n avec base B, et soit $\phi \in \mathcal{L}(V,V)$. Alors ϕ est triagonalisable si et seulement si $(\phi)_B^B$ est trigonalisable.*

Proof. Tout d'abord, on suppose que $(\phi)_B^B$ est trigonalisable. Il existe donc $P \in \mathrm{GL}_n(K)$ telle que $P^{-1}(\phi)_B^B P$ est une matrice triangulaire supérieure (voir la Remarque 8.2.2). Comme P est inversible, on a $P = (\mathrm{id})_E^B$ pour un certain choix de base $E = (f_1, \ldots, f_n)$ de V, et $(\phi)_E^E = (\mathrm{id})_B^E (\phi)_B^B (\mathrm{id})_E^B = P^{-1}(\phi)_B^B P$. Comme $(\phi)_E^E$ est triangulaire supérieure, $\phi(f_i) \in \mathrm{Vect}\,(f_1, \ldots, f_i)$ pour tout i. On pose donc $W_0 = \{0\}$ et $W_i = \mathrm{Vect}\,(f_1, \ldots, f_i)$ pour tout $1 \leq i \leq n$, des sous-espaces ϕ-invariants qui satisfont la définition 8.2.5.

Maintenant, on suppose que ϕ est trigonalisable, c'est-à-dire qu'il existe des sous-espaces vectoriels ϕ-invariants $U_0, \ldots, U_n$ avec $\dim U_i = i$, $U_i \subset U_{i+1}$ pour tout i. On choisit une base $F = (u_1, \ldots, u_n)$ de V, avec $(u_1, \ldots, u_i)$ une base de U_i pour tout $1 \leq i \leq n$. Comme $\phi(u_j) \in \mathrm{Vect}\,(u_1, \ldots, u_i)$ pour tout $j \leq i$, la matrice $(\phi)_F^F$ est triangulaire supérieure, et $(\phi)_B^B = (\mathrm{id})_F^B (\phi)_F^F (\mathrm{id})_B^F$ est trigonalisable. □

8.3. **Polynôme caractéristique et valeurs propres.** Dans ce paragraphe, on établira une méthode pour "trouver" les valeurs propres, dans le cas des matrices ou bien des transformations linéaires des espaces de dimension finie.

Théorème 8.3.1 (Caractérisation de valeurs propres)**.** *Soit V un K-espace vectoriel et $\phi \in \mathcal{L}(V,V)$. Soit encore $\lambda \in K$. Alors λ est une valeur propre de ϕ si et seulement si $\phi - \lambda \mathrm{id}_V$ n'est pas inversible, et si et seulement si* $\ker(\phi - \lambda \mathrm{id}_V) \neq 0$.

Proof. On montre les trois équivalences en même temps :

$$\lambda \text{ est une valeur propre de } \phi \iff \exists v \in V, v \neq 0 \text{ telle que } \phi(v) = \lambda v$$

$$\iff (\phi - \lambda \mathrm{id}_V)(v) = 0 \iff v \in \ker(\phi - \lambda \mathrm{id}_V) \iff \ker(\phi - \lambda \mathrm{id}_V) \neq \{0\}$$

$\iff \phi - \lambda \mathrm{id}_V$ n'est pas inversible. $\square$

Dans le cas d'un espace de dimension finie, le résultat précédent nous donne une méthode "calculatoire" pour trouver les valeurs propres:

Proposition 8.3.2. *Soient V un K-espace vectoriel de dimension finie et $\phi \in \mathcal{L}(V,V)$, et soit encore $\lambda \in K$. Fixons une base B de V et posons $A = (\phi)_B^B$. Alors λ est une valeur propre de A (ou de ϕ) si et seulement si la matrice $A - \lambda I_n$ est non inversible, et si et seulement si* $\det(A - \lambda I_n) = 0$.

Ce dernier résultat motive la définition suivante:

Définition 8.3.3. Soit $A \in M_n(K)$. Soit t une indéterminée. Alors $\det(A - tI_n)$ est un polynôme en t, appelé le *polynôme caractéristique de A*. On le dénote par $c_A(t)$. Donc

$$c_A(t) = \det(A - tI_n).$$

Par la caractérisation des valeurs propres (Théorème 8.3.1) et la Proposition 8.3.2, nous déduisons:

Proposition 8.3.4. *Soit $A \in M_n(K)$ et $\lambda \in K$. Alors λ est une valeur propre de A si et seulement si λ est une racine du polynôme caractéristique $c_A(t)$.*

Quelques cas particuliers:

(1) Si $A = (a_{ij})$ est une matrice triangulaire alors $c_A(t) = (a_{11}-t)(a_{22}-t)\cdots(a_{nn}-t)$. Par conséquent ses valeurs propres sont précisément les valeurs le long de sa diagonale.

(2) Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(K)$. Alors

$$c_A(t) = \det \begin{pmatrix} a-t & b \\ c & d-t \end{pmatrix} = t^2 - (a+d)t + (ad - bc) = t^2 - (\mathrm{Tr}(A))t + \det(A).$$

La proposition suivante montre que les propriétés soulignées dans (2) ci-dessus se généralisent.

Proposition 8.3.5. *Soit $A \in M_n(K)$. Le polynôme caractéristique de A est un polynôme de degré n. De plus, le coefficient de t^n est $(-1)^n$, le coefficient de t^{n-1} est $(-1)^{n-1}\mathrm{Tr}(A)$, et le terme constant est égal à $\det(A)$.*

Proof. Posons $A = (a_{ij})$, et on note $I_n = I$ dans la suite. Alors

$$c_A(t) = \det(A - tI) = \sum_{\sigma \in S_n} \varepsilon(\sigma)(A - tI)_{1\sigma(1)} \cdots (A - tI)_{n\sigma(n)}.$$

Le terme constant d'un polynôme est la valeur du polynôme évalué en $t = 0$, ce qui donne

$$c_A(0) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \cdot a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \det(A).$$

La plus haute puissance de t qu'on peut obtenir dans un produit de termes $(A - tI)_{1\sigma(1)} \cdots (A - tI)_{n\sigma(n)}$ a lieu lorsque tous les facteurs ont un terme avec t, donc lorsque $\sigma(i) = i$ pour tout i (comme l'indéterminée n'apparaît que dans les coefficients diagonaux de la matrice $A - tI$). Comme t apparaîtra dans chacun des facteurs $(A - tI)_{ii}$, on trouvera un terme $(-1)^n t^n$ et aucun terme de degré plus haut. Enfin, on trouve t^{n-1} dans tous les termes de la somme où $\sigma(i) \neq i$ pour au plus une valeur de i. Mais la seule permutation $\sigma \in S_n$ avec cette propriété est la permutation identité, et donc le terme t^{n-1} n'apparaîtra aussi que dans le terme de la somme $(A - tI)_{11} \cdots (A - tI)_{nn}$. C'est un exercice de montrer que le coefficient de t^{n-1} dans le polynôme $(a_{11} - t) \cdots (a_{nn} - t)$ est égal à $(-1)^{n-1}(a_{11} + \cdots + a_{nn})$. □

Grâce au résultat suivant, on peut définir le polynôme caractéristique d'une transformation linéaire (d'un K-espace vectoriel de dimension finie).

Proposition 8.3.6. *Deux matrices semblables ont le même polynôme caractéristique.*

Proof. Soient $A, B \in M_n(K)$ et $P \in \mathrm{GL}_n(K)$ telles que $B = P^{-1}AP$. On a

$$c_B(t) = \det(B - tI_n) = \det(P^{-1}AP - tP^{-1}I_nP)$$

$$= \det(P^{-1}(A - tI_n)P) = \det(P^{-1})\det(A - tI_n)\det(P) = c_A(t).$$

□

Corollaire 8.3.7. *Deux matrices semblables ont les mêmes valeurs propres.*

Corollaire 8.3.8. *Pour $A, B \in M_n(K)$, si B est semblable à A alors $Tr(B) = Tr(A)$ et* $\det(B) = \det(A)$.

Définition 8.3.9. Soit V un K-espace vectoriel de dimension finie et soit $\phi \in \mathcal{L}(V, V)$.

(1) Le *polynôme caractéristique de* ϕ est le polynôme caractéristique $c_A(t)$ où $A = (\phi)_B^B$, pour B une base ordonnée quelconque de V.

(2) La *trace de* ϕ, notée $\mathrm{Tr}(\phi)$ est la trace de A.

(3) Le *déterminant de* ϕ est le déterminant de A.

8.4. **Espaces propres, multiplicité géométrique, multiplicité algébrique.**

Définition 8.4.1. Soit $\alpha \in \mathcal{L}(V, V)$ et soit $\lambda \in K$ une valeur propre de α. *L'espace propre associé à* λ est le sous-espace vectoriel $E_\lambda = \{v \in V \mid \alpha(v) = \lambda v\}$.

On a que

$$E_\lambda = \{0\} \cup \{\text{vecteurs propres de } \alpha \text{ associés à } \lambda\},$$

et Remarque 8.1.4(3) montre que E_λ est un sous-espace vectoriel de V.

De plus, $E_\lambda = \mathrm{Ker}\,(\alpha - \lambda \cdot \mathrm{id}_V)$. Si V est de dimension n, avec base B, et $A = (\alpha)_B^B$, alors E_λ est l'ensemble des vecteurs $v \in V$ tel que $(v)_B$ est une solution du système $(A - \lambda \cdot I_n)X = 0$. En particulier, $\dim E_\lambda = n - \mathrm{rang}(A - \lambda I_n)$. On rappelle aussi que si λ est une valeur propre de α, alors λ est une racine du polynôme caractéristique de α et donc $c_\alpha(t)$ se factorise: $c_\alpha(t) = (t - \lambda)^m f(t)$. En mettant en évidence autant de facteurs $(t - \lambda)$ que possible, on peut supposer que $f(\lambda) \neq 0$.

Définition 8.4.2. Soit V un K-espace vectoriel de dimension finie et soit $\lambda \in K$ une valeur propre de $\alpha \in \mathcal{L}(V, V)$.

a) La *multiplicité algébrique de* λ, notée $m_{alg}(\lambda)$, est la multiplicité de λ comme racine du polynôme caractéristique $c_\alpha(t)$; c'est-à-dire, si $c_\alpha(t) = (t-\lambda)^m \cdot f(t)$, avec $f(\lambda) \neq 0$, alors $m_{alg}(\lambda) = m$.

b) La *multiplicité géométrique de* λ, notée $m_{geom}(\lambda)$, est la dimension de l'espace propre E_λ.

Proposition 8.4.3. *Soit V un K-espace vectoriel de dimension finie et $\lambda \in K$ une valeur propre de $\alpha \in \mathcal{L}(V,V)$. Alors $m_{geom}(\lambda) \leq m_{alg}(\lambda)$.*

Proof. Soit $m = m_{geom}(\lambda)$ et soit $(v_1, \ldots, v_m)$ une base de l'espace propre E_λ. On complète cette base en une base B de V, $B = (v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$. Alors la matrice de ϕ par rapport à la base B est de la forme $A = (\alpha)_B^B = \begin{pmatrix} D & M \\ 0 & N \end{pmatrix}$ ou $D = \lambda I_m$, $M \in M_{m\times(n-m)}(K)$ et $N \in M_{n-m,n-m}(K)$. Par un exercice,

$$c_\alpha(t) = c_A(t) = c_D(t)c_N(t) = (t-\lambda)^m c_N(t).$$

On déduit donc que $m_{alg}(\lambda) \geq m$, comme affirmé. □

8.5. **Diagonalisation.**

Proposition 8.5.1. *Soient V un K-espace vectoriel et $\lambda_1, \ldots, \lambda_r \in K$ des valeurs propres distinctes de $\alpha \in \mathcal{L}(V,V)$, et supposons que $r \geq 2$. Soient $E_{\lambda_1}, \ldots, E_{\lambda_r}$ les espaces propres associés. Alors la somme $E_{\lambda_1} + \cdots + E_{\lambda_r}$ est directe.*

Proof. On raisonne par récurrence sur r. On suppose que $r = 2$. Alors la somme $E_{\lambda_1} + E_{\lambda_2}$ est directe si et seulement si $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$. Soit $w \in E_{\lambda_1} \cap E_{\lambda_2}$. On a

$$\alpha(w) = \lambda_1 w = \lambda_2 w \Longrightarrow (\lambda_1 - \lambda_2)w = 0.$$

Comme $\lambda_1 \neq \lambda_2$ on déduit que $w = 0$ et $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$.

On suppose maintenant que $r > 2$ et que l'énoncé est vérifié pour une somme inférieure à r espaces propres. Cette fois, on doit montrer que, pour tout $1 \leq i \leq r$,

$$E_{\lambda_i} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{i-1}} + E_{\lambda_{i+1}} + \cdots + E_{\lambda_r}) = \{0\}.$$

Soit $w \in E_{\lambda_i} \cap (E_{\lambda_1} + \cdots + E_{\lambda_{i-1}} + E_{\lambda_{i+1}} + \cdots + E_{\lambda_r})$. On écrit $w = \sum_{1\leq j\leq r, j\neq i} w_j$, où $w_j \in E_{\lambda_j}$. On a $\alpha(w) = \lambda_i w$, car $w \in E_{\lambda_i}$ et on a aussi

$$\alpha(w) = \sum_{1\leq j\leq r, j\neq i} \alpha(w_j) = \sum_{1\leq j\leq r, j\neq i} \lambda_j w_j.$$

On déduit que

$$\lambda_i(\sum_{1\leq j\leq r, j\neq i} w_j) = \sum_{1\leq j\leq r, j\neq i} \lambda_j w_j$$

.

Par l'hypothèse de récurrence, la somme $E_{\lambda_1} + \cdots + E_{\lambda_{i-1}} + E_{\lambda_{i+1}} + \cdots + E_{\lambda_r}$ est directe, et par l'unicité d'expression dans une somme directe, on a que $\lambda_i w_j = \lambda_j w_j$ pour tout $j \neq i$, d'où $(\lambda_i - \lambda_j)w_j = 0$. Comme $\lambda_i \neq \lambda_j$ pour tout $j \neq i$, on trouve que $w_j = 0$ pour tout $j \neq i$, ce qui montre que $w = 0$. □

Théorème 8.5.2 (caractérisation des transformations linéaires diagonalisables)**.** *Soit V un K-espace vectoriel de dimension n, et soit $\alpha \in \mathcal{L}(V, V)$. Alors α est diagonalisable si et seulement si les deux conditions suivantes sont satisfaites:*

(1) *$c_\alpha(t)$ est scindé dans $K[t]$, c'est-à-dire $c_\alpha(t) = (-1)^n \prod_{i=1}^r (t - \lambda_i)^{m_i}$, pour $\lambda_i \in K$, $m_i \in \mathbb{N}$, $m_i \geq 1$, et*

(2) *pour chaque valeur propre λ de α, on a $m_{geom}(\lambda) = m_{alg}(\lambda)$.*

Proof. On suppose d'abord que α est diagonalisable. Par définition, il existe une base F de V, formée de vecteurs propres de α. Par conséquent, $(\alpha)_F^F$ est une matrice diagonale. Posons

$$A = (\alpha)_F^F = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}.$$

Soit $\lambda_1, \ldots, \lambda_r$ les valeurs propres distinctes de α. On suppose que la base F est ordonnée comme suit:

$$f_{11}, \ldots, f_{1m_1}, f_{21}, \ldots, f_{2m_2}, \ldots, f_{r1}, \ldots, f_{rm_r},$$

où f_{ij} est un vecteur propre de valeur propre λ_i pour tout $1 \leq i \leq r$ et par conséquent les scalaires $d_1, \ldots, d_n$ sont précisément les scalaires λ_1 (répétés m_1 fois), λ_2 (répétés m_2 fois), etc. Alors $c_\alpha(t) = \det(A - tI_n) = (-1)^n(t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_r)^{m_r}$. En particulier, $c_A(t)$ est scindé. De plus, $m_{alg}(\lambda_i) = m_i$ (rappelons que les λ_i sont distinctes) et $m_{geom}(\lambda) = \dim E_{\lambda_i} \geq m_i$ car $f_{i1}, \ldots f_{im_i}$ sont des vecteurs linéairements indépendants dans E_{λ_i}. Par Proposition 8.4.3, $m_{geom}(\lambda_i) \leq m_{alg}(\lambda_i) = m_i$. Les deux inégalités montrent que $m_{geom}(\lambda_i) = m_{alg}(\lambda_i)$, et ceci pour tout i.

Supposons maintenant que les deux conditions (1) et (2) sont satisfaites, avec $c_\alpha(t) = (-1)^n \prod_{i=1}^r (t-\lambda_i)^{m_i}$. Sans perte de généralité on suppose que les λ_i sont distinctes. Par la condition (2), $\dim E_{\lambda_i} = m_i$ pour tout i. On fixe une base $e_{i1}, \ldots, e_{im_i}$ de E_{λ_i}. Par Proposition 8.5.1, le sous-espace vectoriel $E_{\lambda_1} + \cdots + E_{\lambda_r}$ est une somme directe, et donc

$$\dim(E_{\lambda_1} + \cdots + E_{\lambda_r}) = \sum_{i=1}^r \dim E_{\lambda_i} = \sum_{i=1}^r m_i = \deg(c_A(t)) = \dim V.$$

On déduit que $E_{\lambda_1} + \cdots + E_{\lambda_r} = V$. Comme $E_{\lambda_i} = \mathrm{Vect}\,(e_{i1}, \ldots, e_{im_i})$, l'ensemble $B = \{e_{ij} \mid 1 \leq i \leq r, 1 \leq j \leq m_i\}$ est une famille génératrice de V, et donc de cardinal au moins $\dim V = \sum_{i=1}^r m_i$. Comme B possède au plus $\sum_{i=1}^r m_i$ vecteurs, B est de cardinal $\dim V$ et forme une base de V; c'est-à-dire que B est une base de vecteurs propres pour α et α est diagonalisable. □

Corollaire 8.5.3. *Soient V un K-espace vectoriel de dimension finie n et $\alpha \in \mathcal{L}(V,V)$. Si $c_\alpha(t)$ possède n valeurs propres distinctes, alors α est diagonalisable.*

Proof. Par hypothèse, $c_\alpha(t) = (\lambda_1 - t)\cdots($

$lambda_n - t$) avec comme scalaires $\lambda_1, \ldots, \lambda_n$ distincts, et par conséquent, $c_A(t)$ est scindé. Pour tout $1 \leq i \leq n$, on a $m_{alg}(\lambda_i) = 1$. Comme $1 \leq m_{geom}(\lambda_i) \leq m_{alg}(\lambda_i)$ (Prop. 8.4.3), on a $m_{geom}(\lambda_i) = m_{alg}(\lambda_i)$ pour tout i. Par le théorème précédent, α est diagonalisable. □

Application: calcul des puissances d'une matrice diagonalisable

Soit $A \in M_n(K)$. Si A est diagonalisable, alors il existe $P \in \mathrm{GL}_n(K)$ telle que

$$PAP^{-1} = D = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}.$$

Alors $A = P^{-1}DP$, et donc pour $k \in \mathbb{N}$,

$$A^k = (P^{-1}DP)(P^{-1}DP)\cdots(P^{-1}DP) = P^{-1}D^kP = P^{-1}\begin{pmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n^k \end{pmatrix} P.$$

8.6. **Trigonalisation.**

Théorème 8.6.1 (de trigonalisation)**.** *Soit V un K-espace vectoriel de dimension finie n et soit $\phi \in \mathcal{L}(V,V)$. Alors ϕ est trigonalisable si et seulement si $c_\phi(t)$ est scindé dans $K[t]$.*

Proof. Supposons d'abord que ϕ est trigonalisable. Par Proposition 8.2.6, il existe une base B de V telle que $(\phi)_B^B$ soit une matrice triangulaire (supérieure), disons

$$A = (\phi)_B^B = \begin{pmatrix} d_1 & * & \cdots & * \\ 0 & d_2 & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & d_n \end{pmatrix}.$$

Alors $c_\phi(t) = c_A(t) = \det(A - tI_n)$. Comme $A - tI_n$ est une matrice triangulaire, son déterminant est le produit des coefficients le long de la diagonale; on trouve $c_\phi(t) = (d_1 - t)\cdots(d_n - t)$, ce qui montre que $c_\phi(t)$ est scindé.

Maintenant, on procède par récurrence sur n pour montrer que toute transformation linéaire d'un K-espace vectoriel de dimension finie dont le polynôme caractéristique est scindé est trigonalisable. Si $n = 1$, toute $\phi \in \mathcal{L}(V,V)$ est trigonalisable; prenons $W_0 = \{0\}$ et $W_1 = V$ dans la Definition 8.2.5. On suppose maintenant que $n > 1$ et que le résultat est vrai pour toute transformation linéaire d'un K-espace vectoriel de dimension inférieure à n. Par hypothèse, $c_\phi(t)$ est scindé. En particulier, ϕ possède une valeur propre $\lambda \in K$. Soit $w \in V$ un vecteur propre de valeur propre λ. Posons $U_1 = \text{Vect}\,(w)$, un sous-espace ϕ-invariant de dimension 1. Fixons une base B de V avec $B = (w, w_2, \dots, w_n)$ et posons

$A = (\phi)_B^B$. Alors

$$A = \begin{pmatrix} \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Soit $A' \in M_{n-1,n-1}(K)$ la matrice $A' = \begin{pmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$. Posons $W' = \text{Vect}\,(w_2, \ldots, w_n)$, un K-espace vectoriel de dimension $n-1$ avec base $B' = (w_2, \ldots, w_n)$. On définit $\alpha \in \mathcal{L}(W', W')$ par $(\alpha)_{B'}^{B'} = A'$. On note que pour tout $u \in W'$, $\phi(u) - \alpha(u) \in \text{Vect}\,(w)$ (il suffit de vérifier sur les vecteurs de la base B'.) En utilisant le développement par rapport à la première colonne, on trouve que

$$c_\phi(t) = \det(A - tI_n) = (\lambda - t)\det(A' - tI_{n-1}) = (\lambda - t)c_\alpha(t).$$

Comme $c_\phi(t)$ est scindé, le polynôme caractéristique de α est aussi scindé. Par l'hypothèse de récurrence, α est trigonalisable. Soit $(f_2, \ldots, f_n)$ une base de W' telle que Vect $(f_2, \ldots f_i)$ est α-invariant pour tout i. (Voir la preuve de la Proposition 8.2.6.) Enfin, posons $U_i = \text{Vect}\,(w, f_2, \ldots, f_i)$ pour $2 \leq i \leq n$. On montre que $\{0\} = U_0 \subset U_1 \subset U_2 \cdots \subset U_n = V$ satisfait aux conditions de la Définition 8.2.5. Par construction, $\dim U_i = i$. Aussi, pour tout $i \geq 2$, $\phi(f_i) - \alpha(f_i) \in \text{Vect}\,(w)$. Comme Vect $(f_2, \ldots, f_j)$ est α-invariant, on trouve que $\phi(f_i) \in U_j$ pour tout $2 \leq i \leq j$, ce qui donne l'invariance par ϕ de chaque sous-espace vectoriel U_j, et de suite la tringonalisabilité de ϕ. □

Par le théorème fondamental de l'algèbre, chaque polynôme $p(t) \in \mathbb{C}[t]$ est scindé. Le théorème précédent implique alors :

Corollaire 8.6.2. *Toute transformation linéaire d'un $\mathbb{C}$-espace vectoriel de dimension finie est trigonalisable.*

Proposition 8.6.3 (sur la trace et le déterminant)**.** *Soit V un K-espace vectoriel de dimension n et soit $\phi \in \mathcal{L}(V, V)$. Supposons que $c_\phi(t)$ est scindé dans $K[t]$. Soient*

$\lambda_1, \ldots, \lambda_n$ *les racines de* $c_\phi(t)$ *(avec répétition selon leur multiplicité). Alors* $\mathrm{Tr}(\phi) = \lambda_1 + \cdots + \lambda_n$ *et* $\det(\phi) = \lambda_1 \cdots \lambda_n$.

Proof. Par le Théorème de trigonalisation, ϕ est trigonalisable. Par Proposition 8.2.6, il existe une base B de V telle que $(\phi)_B^B = A = \begin{pmatrix} d_1 & \cdots & & * \\ 0 & \ddots & & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & d_n \end{pmatrix}$. On trouve $c_\phi(t) = c_A(t) = (d_1 - t) \cdots (d_n - t) = (\lambda_1 - t) \cdots - (\lambda_n - t)$. Donc, quitte à renuméroter, on peut supposer que $d_i = \lambda_i$ pour tout i. Enfin, $\mathrm{Tr}(\phi) = \mathrm{Tr}(A) = \sum_{i=1}^n d_i = \sum_{i=1}^n \lambda_i$ et $\det(\phi) = \det(A) = d_1 \cdots d_n = \lambda_1 \cdots \lambda_n$. □

Institut de mathématiques, EPFL