Cours, master 2ième année, 1er semestre.

Algèbre générale.

Jean-Robert Belliard

année 2006–07.

Table des matières

Chapitre 1

Factorisation des applications.

1.1 Rappel de vocabulaire ensembliste.

1.1.1 Relation d'équivalence.

Définition 1.1.1 *Soit E un ensemble. On appelle relation d'équivalence sur E la donnée d'un sous-ensemble $\mathcal{R} \subset E \times E$ vérifiant :*

1. $\forall x \in E \ (x,x) \in \mathcal{R}$ *(réflexivité)*
2. $\forall x,y \in E \ (x,y) \in \mathcal{R} \implies (y,x) \in \mathcal{R}$ *(symétrie)*
3. $\forall x,y,z \in E \ ((x,y) \in \mathcal{R} \, et \, (y,z) \in \mathcal{R}) \implies (x,z) \in \mathcal{R}$ *(transitivité)*

L'usage est de noter $x \sim y$ pour $(x,y) \in \mathcal{R}$ et de dire que x et y sont équivalent pour la relation $\mathcal{R}$ ou $\sim$. L'ensemble des éléments $y \in E$ tel que $x \sim y$ s'appelle la classe de x, et se note parfois $\overline{x} \subset E$. L'ensemble de toutes les classes d'équivalence sous une relation $\sim$ est un sous-ensemble de l'ensemble $\mathbb{P}(E)$ de toutes les parties de E et se note parfois $E/\sim$. Le choix d'un, et d'un seul, élément $x_i \in \overline{x_i}$ dans chacune des classes produit ce qu'on appelle un système de représentants dans E de $E/\sim$.

Exemples

1. Sur tout ensemble l'égalité est une relation d'équivalence.
2. La relation de congruence modulo 10 dans $\mathbb{Z}$, par définition :
$$x \equiv y[10] \iff 10 \mid (x-y).$$
3. La colinéarité des vecteurs dans tout $\mathbb{K}$-espace vectoriel, par définition
$$u \sim v \iff \exists \lambda \in \mathbb{K}, \lambda \neq 0 \ \ \lambda u = v.$$
4. Dans un groupe G on peut associer à tout sous-groupe $H \subset G$ les équivalences à gauche et à droite modulo H, par définition :
$$x \sim^g_H y \iff Hx = Hy \quad \text{et} \quad x \sim^d_H y \iff xH = yH.$$
5. Un espace vectoriel semi-normé est un espace vectoriel muni d'une semi-norme φ. Une semi-norme est une application $\varphi \colon E \longrightarrow \mathbb{R}^+$ vérifiant toutes les propriétés des normes sauf l'implication $(\varphi(x)) = 0 \implies x = 0$. Sur un tel espace la relation $x \sim y \iff \varphi(x-y) = 0$ est une relation d'équivalence.

Définition 1.1.2 *Soit E un ensemble. On appelle partition de E la donnée d'une famille $(E_i)_{i\in I}$ de sous-ensembles de E, indexée par un ensemble d'indice I, et telle que :*

1. $E = \bigcup_{i\in I} E_i$
2. $i \neq j \implies E_i \cap E_j = \emptyset$

Lorsque $(E_i)_{i\in I}$ est une partition de E on note parfois $E = \amalg_{i\in I} E_i$.

Proposition 1.1.3 *Soit E un ensemble.*

1. *Etant donné une partition $E = \amalg_{i\in I} E_i$ on obtient une relation d'équivalence sur E en posant*

$$x \sim y \overset{def}{\iff} \exists i \in I,\ x \in E_i \text{ et } y \in E_i$$

2. *Réciproquement, à partir d'une relation d'équivalence $\sim$ sur un ensemble E on obtient la partition en classe :*

$$E = \coprod_{A\in E/\sim} A.$$

Démonstration. c'est évident. □

1.1.2 Applications et principe de factorisation ensembliste.

Définition 1.1.4 *Soient E et F deux ensembles.*

1. *Une application de E dans F est la donnée d'un sous-ensemble $G \subset E \times F$ (le graphe de l'application) tel que pour tout $x \in E$ il existe un unique $y \in F$ avec $(x, y) \in G$. Si on veut appeler f cette application on note alors $f\colon E \longrightarrow F$ ou bien $E \xrightarrow{f} F$, et aussi $y = f(x)$ lorsque $(x, y) \in G$. Lorsque $y = f(x)$ on dit que y est l'image de x et que x est un antécédent de y pour f.*
2. *Soit $f\colon E \longrightarrow F$. On dit que f est injective et on note $f\colon E \hookrightarrow F$ lorsque $f(x) = f(y)$ entraîne $x = y$ pour tout $x, y \in E$.*
3. *Soit $f\colon E \longrightarrow F$. On dit que f est surjective et on note $f\colon E \twoheadrightarrow F$ lorsque tout élément de F admet (au moins) un antécédent pour f.*
4. *Soit $f\colon E \longrightarrow F$. On dit que f est bijective et on note $f\colon E \xrightarrow{\sim} F$ lorsque f est à la fois injective et surjective.*

Exemples

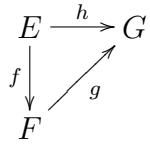
1. Sur tout ensemble E (non vide) le graphe diagonal $\{(x, x);\ x \in E\}$ qui correspond à l'application identité $\mathrm{Id}_E\colon E \longrightarrow E$ telle que $\mathrm{Id}_E(x) = x$.
2. De $\mathbb{R}$ dans $\mathbb{R}$ le graphe $\{(x, 2x); x \in \mathbb{R}\}$ qui correspond à l'application $f(x) = 2x$.
3. Étant donnée une relation d'équivalence $\sim$ sur un ensemble E pour laquelle on note $\overline{x}$ la classe de $x \in E$ le graphe $\{(x, \overline{x}); x \in E\}$ qui correspond à la surjection canonique $\pi\colon E \twoheadrightarrow E/\sim$.

4. Étant donnée une application ensembliste $f\colon E \longrightarrow F$, on obtient une relation d'équivalence sur E en posant

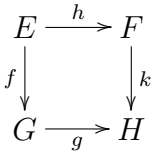
$$x \sim y \overset{\text{def}}{\iff} f(x) = f(y)$$

Exercice 1.1 *Soit $\sim_1$ une relation d'équivalence sur E. Qu'obtient-t'on si on applique la recette de l'exemple 4 à la surjection canonique associée à $\sim_1$ telle que décrite dans l'exemple 3 ? Et réciproquement ?*

Etant donnée deux applications $f\colon E \longrightarrow F$ et $g\colon F \longrightarrow G$ on obtient une troisième application $h := g \circ f\colon E \longrightarrow F$ (composée de g avec f) en posant $h(x) = g(f(x))$. La composition des applications se visualise mieux avec les diagrammes sagitaux. Par exemple :



On parle de diagramme commutatif lorsque les divers morphismes obtenus par compositions (éventuelles) suivant différents chemins coïncident. Dans le cas du triangle ci-dessus, la seule égalité sous-entendue par la commutativité du diagramme est l'égalité $h = g \circ f$. De même on dit qu'un carré de la forme

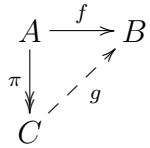


est commutatif lorsque $g \circ f = k \circ h$. Bien sur on rencontrera des diagrammes commutatifs plus complexes.

Théorème 1.1.5 (factorisation ensembliste) *Soit $f\colon A \longrightarrow B$ une application entre deux ensembles A et B et $\pi\colon A \twoheadrightarrow C$ une application surjective entre A et un troisième ensemble C. Alors l'assertion 1. ci-dessous est équivalente à l'implication 2. ci-dessous.*

1. *Il existe une unique application $g\colon C \longrightarrow B$ telle que $g \circ \pi = f$.*
2. *Pour tout $a_1, a_2 \in A$ l'égalité $\pi(a_1) = \pi(a_2)$ entraîne $f(a_1) = f(a_2)$.*

On retiendra plus facilement cet énoncé si on pense au triangle de factorisation suivant :



On parle de factorisation parce que la question posée est en quelque sorte de "diviser" au sens de la composition l'application f par l'application π. L'assertion 2 donne une condition nécessaire et suffisante à cette divisibilité.

Démonstration du théorème 1.1.5.

On montre l'implication 1. $\Longrightarrow$ 2. On suppose donc l'existence d'une application g telle que $g \circ \pi = f$. Soient $a_1, a_2 \in A$ tels que $\pi(a_1) = \pi(a_2)$. Alors en appliquant g à cette égalité on obtient $f(a_1) = g(\pi(a_1)) = g(\pi(a_2)) = f(a_2)$.

On montre l'implication 2. $\Longrightarrow$ 1. Soit $c \in C$. Puisque π est surjective $\pi^{-1}(c)$ est non vide. Soit $a \in \pi^{-1}(c)$. Alors pour tout $a' \in \pi^{-1}(c)$ on a $\pi(a') = \pi(a) = c$ et donc $f(a') = f(a)$ par 2. En conséquence l'ensemble $f(\pi^{-1}(c))$ est un singleton, et l'élément $b \in B$ tel que $f(\pi^{-1}(c)) = \{b\}$ est uniquement défini pour c fixé. Ainsi on peut dire suivant l'usage que $c \mapsto b \in f(\pi^{-1}(c))$ est une application "bien définie". Appelons g cette application. En suivant la construction de g on vérifie immédiatement que $g \circ \pi = f$. Cette égalité donne aussi l'unicité de g. En effet soit $g' \colon C \longrightarrow B$ une application telle que $g' \circ \pi = f$. Pour montrer que $g = g'$ on vérifie l'égalité $g'(c) = g(c)$ pour tout $c \in C$. On fixe c et on choisit $a \in \pi^{-1}(c) \subset A$. Puisque $g \circ \pi = f = g' \circ \pi$ on obtient $g(c) = g(\pi(a)) = f(a) = g'(\pi(a)) = g'(c)$. □

Corollaire 1.1.6 *On reprend le contexte et les notations du théorème 1.1.5. On suppose que les assertions équivalentes 1. et 2. de ce théorème sont vraies. On a alors en outre les équivalences :*

1. *g est surjective si et seulement si f l'est.*
2. *g est injective si et seulement si l'implication de l'assertion 2. est une équivalence.*

Démonstration. Exercice. □

1.2 Factorisation et suites exactes de modules.

1.2.1 Principe de factorisation.

Le principe de factorisation du théorème 1.1.5 se décline dans diverses situation et pour des objets et morphismes plus divers que le cas particulier des ensembles et des applications ensemblistes décrit plus haut. Les énoncés et les démonstrations de ce paragraphe sont valables pour toute structure et tout type de morphisme, à condition qu'il soit possible de définir les noyaux des morphismes et les objets quotients. Pour fixer les idées dans la suite on étudiera les modules à gauche sur un anneau unitaire A et les morphismes de A-modules, même si tout resterait valable *mutatis-mutandis* pour les quotients des structure que vous connaissez (groupes, anneaux, espaces vectoriels, algèbre, groupe topologique, etc...). On énonce tout de même le théorème 1.2.3 en toute généralité.

Définition 1.2.1 *Soient A et B deux ensembles munis d'une des structures ci-dessus, et soit $f \colon A \longrightarrow B$ un morphisme.*

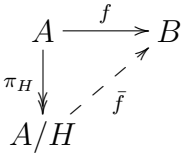
1. *On appelle noyau de f et on note* $\operatorname{Ker} f$ *le sous-objet de A image réciproque de $0 \in B$, c'est-à-dire* $\operatorname{Ker} f = \{a \in A; f(a) = 0\}$.
2. *On appelle image de f et on note* $\operatorname{Im} f$ *le sous-objet de B défini par* $\operatorname{Im} f = \{b \in B; \exists a \in A,\ f(a) = b\}$.
3. *On appelle conoyau de f et on note* $\operatorname{Coker} f$ *le quotient* $B/\operatorname{Im} f$.

Proposition 1.2.2 *Soient A et B deux ensembles munis d'une des structures ci-dessus, et soit $f\colon A \longrightarrow B$ un morphisme.*

1. *f est injective si et seulement si* $\operatorname{Ker} f = \{0\}$.
2. *f est surjective si et seulement si* $\operatorname{Im} f = B$ *si et seulement si* $\operatorname{Coker} f = \{0\}$.

Démonstration. Exercice. □

Théorème 1.2.3 *Soient A et B deux ensemble muni d'une des structures ci dessus, $H \lhd A$ un sous-objet de A tel que A/H soit lui-même muni de cette structure (i.e H est distingué dans le cas particulier de la structure de groupe). Soit $f\colon A \longrightarrow B$ un morphisme, et $\pi_H\colon A \longrightarrow A/H$ la surjection canonique.*



1. *On a équivalences entre les (a) et (b) ci-dessous :*
   (a) Il existe un unique morphisme $\bar{f}\colon A/H \longrightarrow B$ tel que $\bar{f} \circ \pi = f$.
   (b) $H \subset \operatorname{Ker} f$.
2. *Si $\bar{f}$ existe alors $\bar{f}$ est surjective si et seulement si f l'est.*
3. *Si $\bar{f}$ existe alors $\bar{f}$ est injective si et seulement si l'inclusion du (b) est une égalité.*

Démonstration. On se ramène au théorème 1.1.5 en remarquant, par exemple pour la structure de groupe, que pour $a, a' \in A$ l'équivalence $\pi_H(a) = \pi_H(a') \iff a^{-1}a' \in H$. Cette équivalence permet de traduire les inclusions de noyaux du type de l'assertion (a) du théorème 1.2.3 en des implications du type de celle de l'assertion 1. du théorème 1.1.5 (et de même les égalités de noyaux deviennent des équivalences). Ensuite si l'on suppose que f est un morphisme et puisque π_H l'est aussi on démontre au cas par cas, mais sans difficulté, que $\bar{f}$ est aussi un morphisme dès que $\bar{f}$ existe. □

1.2.2 Suites exactes de modules.

Pour fixer les idées à partir de maintenant on se donne un anneau unitaire R et on travaille dans la *catégorie* des R-modules à gauche (les morphismes sont les applications R-linéaires et le noyaux d'une application linéaire est défini comme image réciproque du neutre du module d'arrivé). Soit $N \subset M$ des R-modules. On note $\imath\colon N \longrightarrow M$ et $\pi\colon M \longrightarrow M/N$ les morphismes canoniques. Alors $\imath$ est injectif, π est surjectif, la composée $\pi \circ \imath$ est nulle et on a même l'égalité $\operatorname{Im}(\imath) = \operatorname{Ker}(\pi)$. Cette situation se produit très souvent et il est commode de parler dans ce cas de suites exactes de A-modules :

$$0 \longrightarrow N \xrightarrow{\imath} M \xrightarrow{\pi} M/N \longrightarrow 0$$

Dans cette suite de morphismes les applications $\{0\} \longrightarrow N$ et $M/N \longrightarrow \{0\}$ sont les seules possibles et on note 0 le module $\{0\}$ par abus. Plus généralement on peut parler de suite exacte de longueur quelconque.

Définition 1.2.4 *Étant donné une suite de R-module $(M_n)_{n\in\mathbb{N}}$ et de morphismes $f_n\colon M_n \longrightarrow M_{n+1}$, on dit que*

1. $\ldots M_n \overset{f_n}{\to} M_{n+1} \overset{f_{n+1}}{\to} \ldots$ *est un <u>complexe</u> lorsque $f_{n+1}\circ f_n = 0$.*
2. *On dit que la suite*

$$\cdots \longrightarrow M_{n-1} \xrightarrow{f_{n-1}} M_n \xrightarrow{f_n} M_{n+1} \longrightarrow \cdots$$

   est <u>exacte en M_n</u> lorsque $\mathrm{Im}(f_{n-1}) = \mathrm{Ker}(f_n)$.
3. *On dit que la suite*

$$\cdots \longrightarrow M_{n-1} \xrightarrow{f_{n-1}} M_n \xrightarrow{f_n} M_{n+1} \longrightarrow \cdots$$

   est <u>exacte</u> lorsqu'elle est exacte en M_n pour tout n.

Proposition 1.2.5

1. *Dire que $M \xrightarrow{\alpha} N \longrightarrow 0$ est une suite exacte de module revient à dire que α est un morphisme de modules surjectif.*
2. *Dire que $0 \longrightarrow M \xrightarrow{\beta} N$ est une suite exacte de module revient à dire que β est un morphisme de modules injectif.*
3. *Si un module M apparaît dans une suite exacte $0 \longrightarrow M \longrightarrow 0$ alors le module M est nul.*
4. *Dire que $0 \longrightarrow M \xrightarrow{\gamma} N \longrightarrow 0$ est une suite exacte revient à dire que γ est un isomorphisme.*

Démonstration. C'est immédiat. □

Proposition 1.2.6 *Soit*

$$0 \longrightarrow A \xrightarrow{\alpha} B \xrightarrow{\beta} C \longrightarrow 0$$

une suite exacte (courte) de R-module. Les assertions suivantes sont équivalentes :

(i) Le sous-module $\alpha(A)$ est facteur direct de B.
(ii) Il existe un sous-module $F \subset B$ tel que la restriction de β à F soit un isomorphisme $F \cong C$.
(iii) Il existe un morphisme $a\colon B \longrightarrow A$ tel que $a\circ\alpha = \mathrm{Id}_A$.
(iv) Il existe un morphisme $b\colon C \longrightarrow B$ tel que $\beta\circ b = \mathrm{Id}_C$.

Lorsque ces conditions sont vérifiées le morphisme $b \mapsto (a(b), \beta(b))$ est un isomorphisme $B \cong A \oplus C$.

Démonstration. Pour établir cette équivalence on montre successivement les implications $(i) \implies (ii) \implies (iv) \implies (iii) \implies (i)$.
On montre $(i) \implies (ii)$. Si $\alpha(A)$ est facteur direct soit F un supplémentaire à $\alpha(A)$ dans B. Par définition des suites exactes $\mathrm{Ker}\,\beta = \alpha(A)$ et on a donc $\mathrm{Ker}\,\beta \cap F = \{0\}$. Si $c \in C$ il existe un $b \in B$ tel que $\beta(b) = c$. Or B est somme de F et $\alpha(A)$. Il existe

donc $f \in F$ et $a \in \mathrm{Ker}(\beta)$ tel que $b = f + a$. On a donc $\beta(f) = \beta(b) = c$. La restriction de β au sous-module F est bien un isomorphisme.

On montre $(ii) \implies (iv)$. Soit F tel que $\beta\colon F \longrightarrow C$ soit un isomorphisme, soit $\gamma\colon C \longrightarrow F$ le morphisme réciproque et soit $\varepsilon\colon F \longrightarrow B$ le morphisme donné par l'inclusion. Alors $b = \varepsilon \circ \gamma$ vérifie bien $\beta \circ b = \mathrm{Id}_C$.

On montre $(iv) \implies (iii)$. Puisque α est injective il existe toujours un isomorphisme réciproque $\eta\colon \alpha(A) \longrightarrow A$. Pour $x \in B$, on pose $p(x) = x - b \circ \beta(x)$. On définit ainsi un morphisme $p\colon B \longrightarrow B$. Alors comme $\alpha(A) = \mathrm{Ker}(\beta)$ la restriction de p à $\alpha(A)$ est l'identité. Si $x \in B$ alors $\beta(p(x)) = \beta(x) - \beta \circ b \circ \beta(x) = 0$ car $\beta \circ b = \mathrm{Id}_C$. Donc l'image de p est contenu dans $\alpha(A)$. Le morphisme $a = \eta \circ p$ vérifie bien $a \circ \alpha = \mathrm{Id}_A$.

On montre $(iii) \implies (i)$. Pour ce on vérifie que $\mathrm{Ker}\, a$ est un supplémentaire de $\alpha(A)$ dans B. Soit $x \in \mathrm{Ker}\, a \cap \alpha(A)$. alors il existe $y \in A$ tel que $x = \alpha(y)$. Et comme $a \circ \alpha = \mathrm{Id}_A$ on a $0 = a(x) = a(\alpha(y)) = y$. Il suit $x = \alpha(y) = 0$. On a bien $\alpha(A) \cap \mathrm{Ker}\, a = \{0\}$. Soit $x \in B$. Alors $a(x - \alpha(a(x))) = a(x) - a(\alpha(a(x))) = 0$ puisque $a \circ \alpha = \mathrm{Id}_A$. Donc $x - \alpha(a(x)) \in \mathrm{Ker}\, a$. Donc comme $\alpha(a(x))$ appartient à $\alpha(A)$ l'élément x appartient à $\langle \alpha(A) \cup \mathrm{Ker}\, a \rangle$.

On a démontré les équivalences requises. Si ces conditions sont remplies, l'application $b \mapsto (a(b), \beta(b))$ est clairement linéaire, et son morphisme réciproque est $(x, y) \mapsto \alpha(x) + b(y)$, comme on le voit par un calcul immédiat. □

Définition 1.2.7 *Lorsque les conditions équivalentes du lemme 1.2.6 sont vérifiées on dit que la suite exacte* $0 \longrightarrow A \xrightarrow{\alpha} B \xrightarrow{\beta} C \longrightarrow 0$ *est scindée.*

Lorsque R est un corps tous les sous-espaces vectoriels sont facteurs directs et toutes les suites courtes sont scindées. Il est alors préférable d'utiliser la notion de somme directe plus facile à manier et il serait ridicule de parler de suites exactes d'espaces vectoriels. Bien entendu pour les modules il existe des suites qui ne sont pas scindées, par exemple la suite exacte de $\mathbb{Z}$-modules $0 \longrightarrow p\mathbb{Z} \xrightarrow{x \mapsto x} \mathbb{Z} \xrightarrow{x \mapsto \overline{x}} \mathbb{Z}/p\mathbb{Z} \longrightarrow 0$ n'est pas scindée (Exercice).

Soit $\ldots \longrightarrow M \xrightarrow{f} N$ une suite exacte ne terminant pas par 0. Alors la suite $\ldots \longrightarrow M \xrightarrow{f} N \xrightarrow{\pi_{f(M)}} N/f(M) \longrightarrow 0$ est une suite exacte terminant par 0.

Soit $N \xrightarrow{g} M \longrightarrow \ldots$ une suite exacte ne commençant pas par 0. Alors la suite $0 \longrightarrow \mathrm{Ker}\, g \longrightarrow N \xrightarrow{g} M \longrightarrow \ldots$ est une suite exacte qui commence par 0.

Soit

$$\cdots \longrightarrow A \longrightarrow B \xrightarrow{f} C \longrightarrow \cdots$$

une suite exacte avec plus de trois modules non nuls. Alors on peut la "couper" pour obtenir une suite exacte à trois termes non nuls (dite suite exacte courte) et les deux suites moins longues qui suivent :

$$\cdots \longrightarrow A \longrightarrow \mathrm{Ker}\, f \longrightarrow 0$$

$$0 \longrightarrow \mathrm{Ker}\, f \longrightarrow B \xrightarrow{f} \mathrm{Im}\, f \longrightarrow 0$$

$$0 \longrightarrow \operatorname{Im} f \longrightarrow C \longrightarrow \cdots$$

On peut conclure des remarques ci-dessus que l'étude des suites exactes se ramène à celle des suites exactes courtes c'est-à-dire aux modules quotients. Cependant il est plus commode et élégant lorsque c'est possible de ne considérer qu'une seule suite longue plutôt que de multiplier les suites courtes.

1.2.3 Problème : le lemme du serpent.

Soit R un anneau unitaire. On considère comme donnés les R-modules et les applications R-linéaires du diagramme ci-dessous :

$$\begin{array}{ccccccccc} & & M & \xrightarrow{m} & N & \xrightarrow{n} & P & \longrightarrow & 0\,. \\ & & \downarrow{\scriptstyle \mu} & & \downarrow{\scriptstyle \nu} & & \downarrow{\scriptstyle \rho} & & \\ 0 & \longrightarrow & M' & \xrightarrow{m'} & N' & \xrightarrow{n'} & C' & & \end{array}$$

On suppose que les deux lignes sont des suites exactes et que le diagramme est commutatif. L'objet de l'exercice est de démontrer le

Lemme 1.2.8 (lemme du serpent) *Il existe un morphisme de R-modules*

$$\delta\colon \operatorname{Ker}\rho \longrightarrow \operatorname{Coker}\mu,$$

qui s'insère dans une suite exacte (longue) de R-modules :

$$\begin{array}{ccccccc} \operatorname{Ker}\mu & \xrightarrow{\tilde{m}} & \operatorname{Ker}\nu & \xrightarrow{\tilde{n}} & \operatorname{Ker}\rho & & \\ & & & & & \searrow & \\ \operatorname{Coker}\mu & \xrightarrow[\overline{m'}]{} & \operatorname{Coker}\nu & \xrightarrow[\overline{n'}]{} & \operatorname{Coker}\rho & & \end{array}$$

De plus si n' est surjectif alors $\overline{n'}$ l'est aussi ; et si m est injectif alors $\tilde{m}$ aussi.

Étapes de la démonstration :

1. On définit les applications $\tilde{}$ par restriction.
2. On définit les applications $\bar{}$ par factorisation.
3. On utilise la commutativité du diagramme et une "chasse" au diagramme pour montrer que δ est bien définie (partie "dure" de la démonstration).
4. Vérifications (plutôt moins difficile) de la linéarité de toutes les applications et de l'exactitude de la suite elle-même.

Chapitre 2

Algèbre linéaire basique.

2.1 fondement théorique : la dimension.

On fixe k un corps (commutatif). On appelle groupe additif tout groupe commutatif dont on note $+$ la loi de groupe. On suppose connue les notions de bases concernant les groupes commutatifs (jusqu'au passage au quotient par un sous-groupe).

2.1.1 Espaces, sous-espaces et applications linéaires.

Un espace vectoriel V sur k est un groupe additif muni d'une opération externe $k \times V \longrightarrow V$ notée $(\lambda, x) \mapsto \lambda\ x$ vérifiant les axiomes de la théorie des modules à gauche sur un anneau, à savoir :

Définition 2.1.1 (structure de k-espace vectoriel) *Soit V un groupe additif.*

1. *Une opération externe à gauche de k sur V est une application notée $(\lambda, m) \mapsto \lambda m$ du produit cartésien $k \times V$ dans V.*
2. *On dit que V est un espace vectoriel sur k (ou k-espace vectoriel) lorsqu'il existe une opération externe à gauche de k sur V vérifiant les axiomes (Pour tout $v, v' \in V$ et tout $\lambda, \mu \in k$) :*
   (a) $\lambda(v + v') = \lambda v + \lambda v'$
   (b) $(\lambda + \mu)v = \lambda v + \mu v$
   (c) $1_k v = v$
   (d) $(\lambda\mu)v = \lambda(\mu v)$
3. *Soient V et W deux k-espaces. On appelle application k-linéaire un morphisme de groupes $f : V \longrightarrow W$ compatible avec l'opération de k, autrement dit tel que, pour tout $v \in V$ et tout $\lambda \in k$, on ait $f(\lambda v) = \lambda f(v)$. On note $\mathrm{Hom}_k(V, W)$ l'ensemble des applications k-linéaires de V dans W. On appelle isomorphisme d'espace vectoriel une application linéaire bijective.*
4. *Soit V un k-espace, et soit $W \subset V$. On dit que W est un sous-espace de V lorsque W est un sous-groupe de V stable pour l'opération de k, autrement dit lorsque, pour tout $\lambda \in k$ et tout $w \in W$, on a $\lambda w \in W$.*

Soit V un k-espace et $W < V$ un sous-espace et soit V/W le groupe additif quotient. Alors l'opération externe de k sur V se factorise en une opération externe de k sur V/W. Avec cette opération V/W est aussi un k-espace vectoriel : c'est l'espace vectoriel quotient.

Exemples :

1. Soit n un entier, le produit cartésien k^n muni des opérations évidentes (composantes par composantes) est un k-espace vectoriel.
2. Plus généralement si I est un ensemble et V un k-espace vectoriel l'ensemble V^I des applications de I dans V muni des opérations

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

$$\text{et} \quad \lambda(x \mapsto f(x)) = (x \mapsto \lambda f(x))$$

   est un k-espace vectoriel. L'éléments $f \in V^I$ est parfois noté $(f(i))_{i\in I}$.
3. On note $V^{(I)}$ le sous-espace de V^I contenant les applications "presque toujours nulles" c'est-à-dire les applications f telle que l'image réciproque de 0 par f soit de complémentaire fini dans I ou encore telle qu'il existe un J avec $J \subset I$, $I \backslash J$ fini et pour tout $i \in J$ $f(i) = 0$.
4. Étant donné des espaces vectoriels $(V_i)_{i\in I}$, le produit cartésien $\prod_i V_i$ muni des opérations composantes par composantes est un espace vectoriel.

2.1.2 Familles de vecteurs d'un espace vectoriel.

Définition 2.1.2 *Soit I un ensemble et V un k-espace vectoriel. On appelle famille d'éléments de V (de scalaires si $V = k$) et on note $(x_i)_{i\in I}$, la donnée d'une application $I \longrightarrow V$ notée $i \mapsto x_i$. Par abus on dit que le cardinal $\#I$ de I est le cardinal de la famille $(x_i)_{i\in I}$. Lorsque I est fini on dit que la famille $(x_i)_{i\in I}$ est une famille finie.*

Toute intersection de sous-espaces vectoriel est un sous-espace de sorte que pour toute partie d'un espace V il existe toujours un plus petit sous-espace contenant cette partie.

Définition 2.1.3 *Soit V un espace vectoriel et $\mathcal{F} = (x_i)_{i\in I}$ une famille de vecteurs de V. On appelle sous-espace engendré par $\mathcal{F}$ et on note $\langle x_i, i \in I\rangle$ le plus petit sous-espace de V contenant tous les x_i.*

Proposition 2.1.4 *$\langle x_i, i \in I\rangle$ est l'ensemble formé de toutes les combinaisons linéaires finies possibles $\sum_{i\in J} \lambda_i x_i$ où J parcourt les parties finies de I et les $(\lambda_j)_{j\in J}$ parcourent les familles finies de scalaires.*

Démonstration. C'est évident. □

Définition 2.1.5 *Soit V un espace vectoriel et $\mathcal{F} = (x_i)_{i\in I}$ une famille de vecteurs de V.*

1. *La famille $\mathcal{F}$ est dite libre lorsque pour toute partie finie $J \subset I$ et toute famille de scalaires $(\lambda_j)_{j\in J}$ l'identité $\sum_{j\in J} \lambda_j x_j = 0$ entraîne les identités $\forall j \in J,\ \lambda_j = 0$. Lorsqu'au contraire il existe une relation linéaire finie non triviale $\sum_{j\in J} \lambda_j x_j = 0$ avec au moins un λ_j non nul on dit que la famille $\mathcal{F}$ est liée.*
2. *La famille $\mathcal{F}$ est dite génératrice lorsque $\langle \mathcal{F} \rangle = V$.*
3. *On dit que la famille $\mathcal{F}$ est une base lorsqu'elle est libre et génératrice.*

Proposition 2.1.6 *Soit $(x_i)_{i\in I}$ une famille de V. On a équivalence entre les trois assertions suivantes :*

1. *La famille $(x_i)_{i\in I}$ est une base.*
2. *Pour tout vecteur v de V il existe une unique famille finie de scalaires $(v_j)_{j\in J}$ avec $J \subset I$ telle que*
$$v = \sum_{j\in J} v_j x_j.$$
3. *L'application naturelle $k^{(I)} \longrightarrow V$ définie par $(\lambda_i)_{i\in I} \mapsto \sum_{i\in I} \lambda_i x_i$ est un isomorphisme.*

Démonstration. Exercice. □

Proposition 2.1.7 *Soit $f\colon E \longrightarrow F$ une application linéaire.*

1. *$\mathrm{Im}(f)$ est un sous-espace de F et $\mathrm{Ker}(f) := f^{-1}(0)$ est un sous-espace de E.*
2. *L'image par f d'une famille génératrice de E est une famille génératrice de $\mathrm{Im}(f)$.*
3. *f est injective si et seulement si l'image de toute famille libre de E est une famille libre de F.*
4. *f est un isomorphisme si et seulement si l'image d'une base de E et une base de F si et seulement si l'image de toute base de E est une base de F.*

Démonstration. Exercice. □

Lemme 2.1.8 *Soit $\mathcal{G}$ une famille finie génératrice de $V = \langle \mathcal{G} \rangle$, et soit $\mathcal{L}$ une famille libre de V. Alors $\mathcal{L}$ est finie et $\#\mathcal{L} \leq \#\mathcal{G}$.*

Démonstration. Soit $n = \#\mathcal{G}$ et écrivons $\mathcal{G} = (g_i)_{1\leq i\leq n}$. On va montrer que toute famille de $n+1$ éléments est liée. On procède par récurrence sur n. Si $n = 1$ tout vecteur $v \in V$ s'écrit $\lambda_v g_1$ et deux vecteurs de V non-triviaux, v et w, vérifient la relation linéaire non-triviale $\lambda_v w - \lambda_w v = 0$. Pour établir l'hérédité, soit m un entier, $m \geq 2$, tel que le lemme soit vrai pour toute famille $\mathcal{G}'$ de cardinal $m-1$. Supposons $\mathcal{G}$ de cardinal m et soit $\mathcal{V} = (v_i)_{0\leq i\leq m}$ une famille de $m+1$ vecteurs de V. Pour tout j il existe une famille de scalaire $(\lambda_{i,j})_{1\leq i\leq m}$ tels que $v_j = \sum_i \lambda_{i,j} g_i$. Pour montrer que $\mathcal{V}$ est liée on peut supposer v_0 non nul et donc quite à permuter les g_i que le *pivot* $\lambda_{1,0}$ est non nul. On utilise ce pivot non nul pour éliminer la composante en g_1 et on forme les m vecteurs $w_1, \cdots w_m$ du sous-espace $\langle g_2, \cdots g_m \rangle$ ci-dessous :

$$w_j = \lambda_{1,0} v_j - \lambda_{1,j} v_0 = \sum_{i=2}^{m} (\lambda_{1,0}\lambda_{i,j} - \lambda_{1,j}\lambda_{i,0}) g_i.$$

Par hypothèse de récurrence ces m vecteurs vérifient une relation linéaire non triviale

$$0 = \sum_{j=1}^{m} \alpha_j(\lambda_{1,0}v_j - \lambda_{1,j}v_0) = (\sum_{j=1}^{m} -\alpha_j\lambda_{1,j})v_0 + \sum_{j=1}^{m} \alpha_j\lambda_{1,0}v_j.$$

Puisque $\lambda_{1,0}$ et au moins l'un des α_j est non nul la famille $v_0, \cdots, v_m$ est liée. □

Remarque : L'idée de la preuve repose sur le principe du pivot de Gauß : on passe de la matrice des v_j à celle des w_j en utilisant le pivot $\lambda_{1,0}$ pour annuler la première ligne. Par récurrence sur la dimension on obtient un système triangulaire inférieur et si il y a plus de colonnes que de lignes, les dernières colonnes sont nulles (et en particulier linéairement dépendantes).

2.1.3 Dimension des espaces vectoriels de type fini.

L'inclusion des images définit une relation d'ordre sur les familles. Cette relation d'ordre permet de caractériser les bases d'un espace vectoriel et de montrer leur existence en toute généralité.

Proposition 2.1.9 *Soit V un espace vectoriel et $\mathcal{F}$ une famille de V. Les assertions suivantes sont équivalentes :*

1. *$\mathcal{F}$ est une base de V.*
2. *$\mathcal{F}$ est une famille libre maximale parmi les familles libres de V.*
3. *$\mathcal{F}$ est une famille génératrice minimale parmi les familles génératrices de V.*

Démonstration. Supposons que $\mathcal{F}$ soit une base de V. Alors $\mathcal{F}$ est à la fois génératrice et libre. Soit $x \in V \setminus \mathcal{F}$. Alors x s'écrit comme combinaison linéaire finie d'éléments de $\mathcal{F}$ et en particulier la famille $\mathcal{F} \cup \{x\}$ n'est pas libre. On a montré que 1 entraîne 2. Soit $x \in \mathcal{F}$. Comme $\mathcal{F}$ est libre x n'appartient pas aux sous-espace engendré par $\mathcal{F} \setminus \{x\}$, et cette dernière famille n'est pas génératrice. On a montré que 1 entraîne 3.

Supposons que $\mathcal{F}$ soit une famille libre maximale. Soit $x \in V \setminus \mathcal{F}$. Alors la famille $\mathcal{F} \cup \{x\}$ contient strictement $\mathcal{F}$ et n'est donc plus libre. Il existe donc une relation de dépendance linéaire finie non triviale entre les éléments de cette famille $\lambda_x x + \sum_{y\in\mathcal{F}} \lambda_y y = 0$. Puisque $\mathcal{F}$ est libre on a aussi $\lambda_x \neq 0$. De sorte que $x = -(1/\lambda_x) \sum_{y\in\mathcal{F}} \lambda_y y \in \langle\mathcal{F}\rangle$. Cela montre que 2 entraîne 1.

Supposons que $\mathcal{F}$ soit une famille génératrice minimale. Alors une relation de dépendance linéaire non triviale entre les éléments de $\mathcal{F}$ permet d'omettre un élément de $\mathcal{F}$ qui est déjà dans le sous-espace engendré par les autres éléments. Cela contredit la minimalité de cette famille, qui est donc libre. On a montré que 3 entraîne 1. □

Définition 2.1.10 *On dit qu'un espace vectoriel V est de type fini lorsqu'il admet une famille génératrice finie.*

Théorème 2.1.11 (dimension) *Soit V un espace vectoriel de type fini.*

1. *De toute famille génératrice de V on peut extraire une base de V : en particulier V admet une base.*

2. *Toute les bases de V sont finie et ont même cardinal : ce cardinal s'appelle la dimension de V.*
3. *Toute famille libre de V se complète en une base.*

Démonstration. 1. Par hypothèse V admet un système générateur fini fixé $e_1, \cdots, e_t$. On part d'un système générateur quelconque $\mathcal{G} = (g_i)_{i\in I}$. Alors chaque e_j pour $1 \leq j \leq t$ est combinaison linéaire d'un nombre fini de g_i de sorte qu'on peut extraire de $\mathcal{G}$ un système générateur fini. On se ramène ainsi au cas d'un système générateur fini $\mathcal{G}_1 = (g_1, \cdots, g_n)$ de V. Si $\mathcal{G}$ est libre alors c'est une base. Sinon une relation linéaire non triviale $\sum_i \lambda_i g_i = 0$ avec $\lambda_j \neq 0$ montre que la famille $\mathcal{G}_2 = \mathcal{G}_1 \setminus g_j$ est encore génératrice. Par ce procédé on aboutit soit à un système générateur minimal c'est-à-dire une base, soit à un système générateur ne contenant que le seul vecteur nul. Dans ce dernier cas l'espace vectoriel tout entier est réduit à $\{0\}$ et on convient que $\emptyset$ est une base de V.

2. est une conséquence du lemme 2.1.8 et de 1. Soient $\mathcal{B}$ une base finie de V et $\mathcal{B}'$ une base de V. Alors $\mathcal{B}'$ est libre et $\mathcal{B}$ est génératrice donc $\#\mathcal{B}' \leq \#\mathcal{B}$. Ainsi $\mathcal{B}'$ est génératrice finie tandis que $\mathcal{B}$ est libre et le lemme 2.1.8 s'applique aussi pour l'inégalité réciproque.

3. Avec 1. et 2. on peut parler de la dimension finie n de V. Par le lemme 2.1.8 on sait aussi que toute famille de $n+1$ vecteurs est liée. Soit $\mathcal{L}$ une famille libre. Si $x \in V \setminus \langle \mathcal{L} \rangle$ alors la famille $\mathcal{L} \cup \{x\}$ est encore libre. Par ce procédé on aboutit en au plus n étapes à une famille libre maximale contenant $\mathcal{L}$. Cette base convient. □

Corollaire 2.1.12 *Deux espaces vectoriels de type fini sont isomorphes si et seulement si ils ont même dimension.*

Démonstration. Si deux espaces ont même dimension n ils sont tous deux isomorphes à k^n. Réciproquement si deux espaces sont isomorphes l'image d'une base de l'un est une base de l'autre par cet isomorphisme et les dimensions coïncident. □

Proposition et définition 2.1.13 *Soient F et G deux sous-espaces d'un espace vectoriel E de dimension finie. Les assertions suivantes sont équivalentes et lorsqu'elles sont remplies on dit que F et G sont supplémentaire l'un de l'autre dans E, et on note $E = F \oplus G$.*

1. *$F \cap G = \{0\}$ et $E = F + G$.*
2. *Tout $x \in E$ d'écrit de manière unique $x = f + g$ avec $f \in F$ et $g \in G$*
3. *L'application naturelle $F \times G \longrightarrow E$ définie par $(f, g) \mapsto f + g$ est un isomorphisme.*

Démonstration. Exercice. □

Corollaire 2.1.14 *Tout sous-espace $F \subset E$ d'un espace vectoriel de dimension finie admet un supplémentaire.*

Démonstration. Il suffit de compléter une base de F en une base de E puis de considérer le sous-espace vectoriel engendré par les vecteurs qui complètent cette base. □

Exercice 2.1 *Soit V un espace vectoriel de dimension finie. Soit $\mathcal{G} = (g_i)_{i\in I}$ une famille génératrice de V et $\mathcal{L} = (l_j)_{j\in J}$ une famille libre de V. Montrer qu'il existe une base $\mathcal{B}$ de V contenant $\mathcal{L}$ et telle que $\mathcal{B} \setminus \mathcal{L}$ soit contenu dans $\mathcal{G}$.*

Définition 2.1.15 *Soit $f\colon E \longrightarrow F$ une application linéaire. On appelle rang de f et on note* $\mathrm{rang}(f)$ *la dimension de l'image de f.*

Théorème 2.1.16 (rang) *Soit $f\colon E \longrightarrow F$ une application linéaire. Alors on a* $\dim(E) = \dim(\mathrm{Ker}\, f) + \mathrm{rang}(f)$.

Démonstration. On se donne des vecteurs $e_1, \cdots, e_{n+r}$ tels que les e_i pour $i = 1 \cdots n$ forment une base de $\mathrm{Ker}(f)$ et que les $f(e_i)$ pour $i = n+1, \cdots n+r$ forment une base de $\mathrm{Im}(f)$. On vérifie que la famille $(e_i)_{1\leq i\leq n+r}$ est une base de E. □

2.2 Matrices

2.2.1 Prérequis

Je considère comme connue la notion de matrices à coefficients dans un anneau commutatif unitaire A, la structure de A-module libre (de rang nm) de $M_{n,m}(A)$ avec sa base canonique $E_{i,j} \in M_{n,m}(A)$ et la notion de produit matriciel

$$M_{n,m}(A) \times M_{m,l}(A) \longrightarrow M_{n,l}(A)$$

$$([a_{i,j}], [b_{j,k}]) \longmapsto [c_{i,k}]$$

avec

$$c_{i,k} = \sum_{j=1}^{m} a_{i,j} b_{j,k}$$

Dans les écritures ci-dessus les indices i parcourent $\{1, \cdots, n\}$, les indices j parcourent $\{1, \cdots, m\}$ et les indices k parcourent $\{1, \cdots, l\}$. L'écriture des matrices sous la forme $[a_{i,j}]$ l'indice i se rapportant aux lignes et j se rapportant aux colonnes est une convention parfaitement légitimée par l'élimination des indices j dans la somme qui définit $c_{i,k}$. Il faut aussi se rappeler que l'on multiplie à gauche par une matrice A ayant autant de colonnes que la matrice de droite B a de lignes. Le résultat du produit est la matrice C qui a autant de ligne que la matrice de gauche A et autant de colonnes que la matrice de droite B. Avec ce produit l'ensemble des matrices carrés d'ordre n, noté $M_n(A)$, est une A-algèbre unitaire de neutre multiplicatif la matrice diagonale avec coefficients diagonals tous égaux à 1 notée I_n. L'anneau A s'identifie canoniquement avec le sous-anneaux AI_n de $M_n(A)$. Le groupe linéaire d'ordre n est le groupe multiplicatif des matrices inversibles d'ordre n. On le note $GL_n(A)$.

Exercice 2.2 *Montrer que A est le centre de $M_n(A)$.*

2.2.2 Représentation matricielle des morphismes.

Soit E et F deux espaces vectoriels de dimension finies rapportés aux bases $e = (e_1, \cdots, e_n)$ de E et $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_d)$ de F. Tout vecteur x de F est uniquement représenté par ses coordonnée dans la base ε c'est-à-dire par l'unique famille de scalaires $(x_i)_{1\leq i\leq d}$ telle que $x = \sum_i x_i\varepsilon_i$. On convient de représenter cette famille de scalaire par une matrice à d lignes et 1 colonne. De sorte que toute famille finie $(f_j)_{1\leq j\leq s}$ de vecteurs de F est uniquement représentée par une matrice à d lignes et s colonnes, la j-ième colonne étant la matrice des coordonnées de f_j dans la base ε.

Définition 2.2.1 *La matrice de la famille de vecteur $f_1, \cdots, f_t$ relativement à la base ε de F est la matrice à d lignes et t colonnes notée* $\mathrm{Mat}_\varepsilon(f_1, \cdots, f_t) = [m_{i,j}]$ *et définie par*

$$f_j = \sum_{i=1}^{d} m_{i,j}\varepsilon_i.$$

Soit $f\colon E \longrightarrow F$ une application linéaire. Alors f est uniquement déterminée par les images $f(e_j)$ pour $j = 1, \cdots, n$.

Définition 2.2.2 *La matrice de f relativement aux bases e de E et ε de F est la matrice du système de vecteurs $f(e_1), \cdots, f(e_n)$ relativement à la base ε. On la note*

$$\mathrm{Mat}_{\varepsilon,e}(f) = \mathrm{Mat}_\varepsilon(f(e_i)_{i=1,\cdots,n}).$$

Si chacun des $f(e_j)$ s'écrit dans la base ε comme ci-dessous

$$f(e_j) = \sum_{i=1}^{d} m_{i,j}\varepsilon_i,$$

alors la matrice de f relatives aux bases ε et e est la matrice $d \times n$

$$\mathrm{Mat}_{\varepsilon,e}(f) = \left[m_{i,j}\right]_{1\leq i\leq d,\, 1\leq j\leq n}.$$

Si G est un troisième espace, rapporté à une troisième base $\gamma = (\gamma_1, \cdots, \gamma_s)$ et si $g\colon F \longrightarrow G$ est une autre application linéaire on a

$$\mathrm{Mat}_{\gamma,e}(g \circ f) = \mathrm{Mat}_{\gamma,\varepsilon}(g)\mathrm{Mat}_{\varepsilon,e}(f).$$

La convention d'écrire d'abord la base ε de l'espace d'arrivée dans la notation $\mathrm{Mat}_{\varepsilon,e}$ permet cette élimination de l'indice ε intermédiaire.

2.2.3 Changement de bases.

Soit E un espace vectoriel de dimension n finie et soient $e = (e_1, \cdots, e_n)$ et $e' = (e'_1, \cdots, e'_n)$ deux bases de E. On a

$$\mathrm{Mat}_{e',e}(\mathrm{Id}_E) = \mathrm{Mat}_{e'}(e_1, \cdots, e_n) =: \mathrm{Mat}_{e'}(e).$$

La dernière égalité définit la notation $\mathrm{Mat}_{e'}(e)$. L'inverse de cette matrice est la matrice $\mathrm{Mat}_{e',e}^{-1} = \mathrm{Mat}_{e,e'}$. Cette matrice (dite de passage) permet de calculer les coordonnée d'un vecteur dans la base e' à partir de ses coordonnées dans la base e. Précisons : soit $x = \sum_i x_i e_i = \sum_i x'_i e'_i$. Alors on a

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = \mathrm{Mat}_{e'}(e) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Prenons maintenant en sus un espace vectoriel F de dimension finie m, deux bases ε et ε' de F et une application linéaire $f\colon E \longrightarrow F$. Alors on a un carré commutatif :

$$\begin{array}{ccc} (E,e) & \xrightarrow{f} & (F,\varepsilon) \\ \downarrow{\scriptstyle \mathrm{Id}_E} & & \downarrow{\scriptstyle \mathrm{Id}_F} \\ (E,e') & \xrightarrow{f} & (F,\varepsilon') \end{array}.$$

Ce diagramme donne les identités :

$$\mathrm{Mat}_{\varepsilon',e'}(f)\,\mathrm{Mat}_{e'}(e) = \mathrm{Mat}_{\varepsilon'}(\varepsilon)\,\mathrm{Mat}_{\varepsilon,e}(f)$$

ou encore

$$\mathrm{Mat}_{\varepsilon',e'}(f) = \mathrm{Mat}_{\varepsilon'}(\varepsilon)\,\mathrm{Mat}_{\varepsilon,e}(f)\,\mathrm{Mat}_{e}(e').$$

Définition 2.2.3 *Soient n et m des entiers. Deux matrices $M, M' \in M_{m,n}(k)$ sont dites équivalentes si il existe une application linéaire $f\colon k^n \longrightarrow k^m$ et deux couples de bases e, e' de k^n et $\varepsilon, \varepsilon'$ de k^m tels que $M = \mathrm{Mat}_{\varepsilon,e}(f)$ et $M' = \mathrm{Mat}_{\varepsilon',e'}(f)$.*

Il revient au même de dire qu'il existe une matrice $S \in GL_m(k)$ et une matrice $T \in GL_n(k)$ telles que $M = SM'T$. Comme l'indique la terminologie on définit ainsi une relation d'équivalence sur $M_{m,n}(k)$.

Définition 2.2.4 *Soit $M \in M_{m,n}(k)$ une matrice. On appelle rang de M la dimension du sous-espaces de k^m engendré par les n vecteurs colonnes de M.*

Le rang d'une matrice est bien sûr égal au rang de toute application linéaire représenté par cette matrice. Ce rang ne dépend pas des bases choisies : c'est un invariant de la classe d'équivalence de la matrice. Sur un corps, cet invariant donne à lui seul un système complet d'invariant :

Théorème 2.2.5 *Soit M une matrice de $M_{m,n}(k)$ de rang r. Alors M est équivalente à une matrice de la forme*

$$M \sim \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix},$$

où les matrices $0_{s,t}$ sont les matrices identiquement nulles avec s lignes et t colonnes (on convient que la matrice à 0 ligne ou à 0 colonne est la matrice vide). En particulier deux matrices sont équivalentes si et seulement si elles ont même rang.

Démonstration. Soit $f\colon k^n \longrightarrow k^m$ le morphisme représenté par M dans les bases canoniques de k^n et de k^m. Soient $e_1, \cdots, e_r$ des vecteurs de k^n tels que la famille $f(e_i)$ forme une base de $\mathrm{Im}(f)$. Alors la famille $f(e_i)_{1\leq i\leq r}$ est libre dans k^m et on peut la compléter en une base ε de k^m. La famille $(e_i)_{1\leq i\leq r}$ est libre et constitue une base d'un supplémentaire de $\mathrm{Ker}(f)$. Complétons la famille libre $(e_i)_{1\leq i\leq r}$ avec une base du noyau de f pour obtenir une base e de k^n. Alors la matrice $\mathrm{Mat}_{\varepsilon,e}(f)$ est équivalente à M et est de la forme voulue. □

Étant donnée une matrice concrète, la démonstration qui précède ne fournit aucune méthode pour calculer le rang d'une matrice M et encore moins les bases de l'image ou du noyau qui interviennent dans le raisonnement. La section qui suit décrit l'algorithme du pivot de Gauß qui résout de façon effective pratiquement toutes les questions calculatoires d'algèbre linéaire.

2.3 Opérations élémentaires sur les matrices.

Soit A un anneau commutatif.

Définition 2.3.1 *Soit $M \in M_{m,n}(A)$. On appelle opération élémentaire sur M l'une des transformations suivantes :*

1. *ajouter à une colonne (resp. ligne) de M le produit par un élément de A d'une* **autre** *colonne (resp. ligne) : on parle de transvection sur les lignes (resp. colonnes) de M.*
2. *permuter les colonnes (resp. lignes) de M.*
3. *multiplier une colonne (resp. ligne) de M par un élément de $A^\times$: on parle de dilatation ou d'affinité sur M.*

Ces opération élémentaires peuvent être modélisée par la multiplication matricielle de M par une matrice élémentaire. Pour retrouver ces "matrices élémentaires" l'on doit retenir quelques principes simples.

- L'action sur les lignes de M est toujours modélisée par la multiplication à gauche de M par une matrice élémentaire. La matrice M est alors remplacée par EM où E est la matrice élémentaire ad hoc. L'action sur les colonnes s'obtient elle par multiplication de M à droite par la matrice élémentaire ad hoc E. La matrice M est alors remplacée par ME.
- Les coefficients de la matrice élémentaire E elle-même s'obtiennent en appliquant à la matrice I_m (resp. I_n) l'operation élémentaire sur les lignes (resp. colonnes) que E est sensée modéliser : en effet on a toujours $EI_m = E$ et $I_nE = E$ ...
- La matrice de transvection $T_{i,j}(\lambda) \in GL_m(A)$ qui modélise l'opération élémentaire sur m lignes $L_i \leftarrow L_i + \lambda L_j$ est la transposée de la matrice de transvection $T_{j,i}(\lambda) \in GL_m(A)$ qui modélise la "même" opération élémentaire sur m colonnes $C_i \leftarrow C_i + \lambda C_j$.
- La matrice de permutation $Q(\sigma^{-1})$ qui modélise l'opération élémentaire sur m lignes $L_i \leftarrow L_{\sigma(i)}$ est l'inverse de la matrice de permutation $Q(\sigma)$ qui modélise la "même" opération élémentaire sur m colonnes.

Ceci dit en appliquant à la matrice I_s l'opération élémentaire à modéliser, on définit les notations :

Notations 2.3.2 *Soit s un entier, ε la base canonique de A^s et $(E_{i,j})_{1\leq i,j\leq s}$ la base canonique de $M_s(A)$, c'est-à-dire pour i et j fixé $E_{i,j} = [e_{l,c}]$ avec $e_{l,c} = 1$ si $i = l$ et $j = c$, $e_{l,c} = 0$ sinon.*

1. *Pour $\lambda \in A$ et $i \neq j$ on note $T_{i,j}(\lambda) = I_s + \lambda E_{i,j} \in GL_s(A)$ la matrice de transvection.*
2. *Pour $\sigma \in S_n$ on note $Q(\sigma) = \mathrm{Mat}_\varepsilon(\varepsilon_{\sigma(i)}, i = 1 \cdots s) \in GL_s(A)$ la matrice de permutation.*
3. *Pour $\mu \in A^\times$ on note $D_j(\mu) = \sum_{1\leq i\leq s,\, j\neq i} E_{i,i} + \mu E_{j,j} \in GL_s(A)$ la matrice de dilatation.*

Remarques :

1. Les applications $\lambda \mapsto T_{i,j}(\lambda)$, (resp. $\mu \mapsto D_i(\mu)$, resp. $\sigma \mapsto Q(\sigma)$) sont des homomorphismes de groupes $A \longrightarrow GL_s(A)$, (resp. $A^\times \longrightarrow GL_s(A)$, resp. $S_s \longrightarrow GL_s(A)$). En particulier
$$T_{i,j}(\lambda)^{-1} = T_{i,j}(-\lambda),\ D_i(\mu)^{-1} = D_i(\mu^{-1}) \text{ et } Q(\sigma)^{-1} = Q(\sigma^{-1}).$$
2. On appelle dilatation ou transvection un endomorphisme $f \in \mathrm{Hom}_k(V)$ dont la matrice relative à une base de V est $T_{i,j}(\lambda)$. On appelle dilatation un endomorphisme $f \in \mathrm{Hom}_k(V)$ dont la matrice dans une base de V est $D_i(\mu)$.

Maintenant pour être logiquement complet il faudrait vérifier que ces matrices agissent bien comme il se doit. J'énonce le résultat attendu et laisse la vérification au lecteur.

Proposition 2.3.3 *Soient A un anneau commutatif, et soit $M \in M_{m,n}(A)$ une matrice dont on note C_j, $j = 1 \cdots n$ les colonnes et L_i, $i = 1 \cdots m$ les lignes.*

$$M = [C_1\ C_2\ \cdots C_n] = \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{bmatrix}.$$

Pour $T_{i,j}(\lambda) \in GL_m(A)$, on a $T_{i,j}(\lambda)M = \begin{bmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_i + \lambda L_j \\ \vdots \\ L_m \end{bmatrix}$.

Pour $T_{i,j}(\lambda) \in GL_n(A)$, on a $MT_{i,j}(\lambda) = [C_1\ \cdots\ C_{j-1}\ C_j + \lambda C_i\ \cdots\ C_n]$.

Pour $\sigma \in S_m$, on a $Q(\sigma)M = \begin{bmatrix} L_{\sigma^{-1}(1)} \\ L_{\sigma^{-1}(2)} \\ \vdots \\ L_{\sigma^{-1}(m)} \end{bmatrix}$.

Pour $\sigma \in S_n$, on a $MQ(\sigma) = \begin{bmatrix} C_{\sigma(1)} & C_{\sigma(2)} & \cdots & C_{\sigma(n)} \end{bmatrix}$.

Pour $D_i(\mu) \in GL_m(A)$, on a $D_i(\mu)M = \begin{bmatrix} L_1 \\ \vdots \\ L_{i-1} \\ \mu L_i \\ \vdots \\ L_m \end{bmatrix}$.

Pour $D_j(\mu) \in GL_n(A)$, on a $MD_j(\mu) = [C_1 \ \cdots \ C_{j-1} \ \mu C_j \ \cdots \ C_n]$.

Démonstration. Exercice. □

Remarques : Il faut aussi savoir que $\det(T_{i,j}(\lambda)) = 1$, $\det(D_i(\mu)) = \mu$ et $\det(Q(\sigma)) = \varepsilon(\sigma)$ (voir le chapitre 3 pour la théorie du déterminant). En fait l'algorithme du pivot de Gauß que l'on va présenter immédiatement fournit aussi une méthode de calcul de déterminant très efficace.

Théorème 2.3.4 *Soit k un corps et $M \in M_{m,n}(k)$ une matrice de rang r dont on note les colonnes $C_1, \cdots C_n$. Alors il existe $P \in GL_m(k)$ une matrice produit de $T_{i,j}(\lambda)$, $\sigma \in S_n$ et $D_r(\mu) \in GL_r(k)$ tels que :*

$$PMQ(\sigma) = \begin{bmatrix} D_r(\mu) & * \\ 0_{m-r,r} & 0_{m-r,n-r} \end{bmatrix}.$$

Si $r < m$ on peut choisir $\mu = 1$. Si $r = n = m$ alors $\det(M) = \varepsilon(\sigma)\mu$. *La famille $(C_{\sigma(i)})_{1 \leq i \leq r}$ est une base du sous-espace de k^m engendré par les vecteurs colonnes de M. Si $C_1, \cdots, C_r$ est une famille libre, on peut choisir $\sigma = \mathrm{Id}$.*

Démonstration. Il s'agit de l'algorithme de Gauß. Pour une démonstration complète et une interprétation de la matrice $*$ voir le théorème 1.2.3.1 p. 45 du livre "Algèbre des matrices" par Jean Fresnel. J'indiquerai au tableau les grandes lignes de cet algorithme qu'il faut vraiment maîtriser. Ce théorème admet énormément de variantes et il est essentiel que vous disposiez d'une référence qui *vous* convienne à ce sujet. □

Cet algorithme a une foule d'applications pratiques. Initialement c'est une excellente méthode de résolution des systèmes linéaires. Ensuite l'algèbre linéaire se ramène, pour l'essentiel, à la résolution de ces systèmes. Je mentionnerai par exemple le critère d'inversibilité et l'inversion de matrices carrés.

Corollaire 2.3.5 *Soient k un corps et $M \in GL_n(k)$. Alors $M = PD_n(\mu) = D_n(\mu)Q$ où* $\det(M) = \mu$ *et où les matrices P et Q sont des produits de $T_{i,j}(\lambda)$. En particulier $GL_n(k)$ est engendré par les $T_{i,j}(\lambda)$ et les $D_n(\mu)$, et le sous-groupe $SL_n(k)$ formé des matrices de déterminant 1 est engendré par les $T_{i,j}(\lambda)$.*

Démonstration. Par le théorème puisque M est inversible on peut prendre $\sigma = \mathrm{Id}$ et écrire $M = P^{-1}D_n(\mu)$, mais P et donc P^{-1} est produit de $T_{i,j}(\lambda)$. On montre qu'on peut écrire $M = D_n(\mu)Q$ en utilisant la version "duale" du théorème pour laquelle on permute les lignes et on "nettoie" les colonnes. □

Chapitre 3

Déterminant.

Dans ce chapitre on définit la notion de déterminant en le replaçant dans les contexte plus général des formes multilinéaires d'un module libre sur un anneau commutatif unitaire. Le cadre des formes multilinéaires rend plus naturelle les formules de définition qui autrement paraissent parachutées. Utiliser un anneau et non un corps de scalaires donne plus de souplesse pour des applications classiques. Par exemple cela permet de définir un polynôme caractéristique (et pas une fraction rationnelle) associée à un endomorphisme de façon rigoureuse et sans contorsions ridicules. Cette présentation du déterminant peut paraître abstraite. On devra parfois penser au cadre plus intuitif des espaces vectoriels euclidiens, et dans ce cadre il est bon de se souvenir que le déterminant d'un système de vecteurs est le volume du parallélotope bordé par ces vecteurs. Ce point de vue sera rappelé au chapitre suivant consacré aux sous-groupes de $\mathbb{R}^n$.

3.1 Formes multilinéaires alternées.

On fixe un anneau commutatif unitaire A et un A-module libre de rang n noté E, dont on fixe une base $\varepsilon_1, \cdots, \varepsilon_n$.

Définition 3.1.1 *Soit p un entier*

1. *Une forme p-linéaire sur E est une application $\varphi\colon E^p \longrightarrow A$ linéaire par rapport à chacune des coordonnées. C'est à dire telle que pour tout $\lambda \in A$, tout $(x_i)_{i=1}^{i=p} \in E^p$ et tout $x \in E$ on ait*

$$\begin{aligned} f(x_1, \cdots, \lambda x_i + x, \cdots, x_p) &= \lambda f(x_1, \cdots, x_i, \cdots, x_p) \\ &+ f(x_1, \cdots, x, \cdots, x_p). \end{aligned}$$

2. *L'ensemble des formes p-linéaires sur E se note $\mathcal{L}^p(E, A)$. C'est un A module pour les opérations naturelles*

$$(\lambda\varphi + \psi)(x_1, \cdots, x_p) = \lambda\varphi(x_1, \cdots, x_p) + \psi(x_1, \cdots, x_p).$$

Pour tout entier naturel k on note $\mathbb{N}_k = \{1 \cdots, k\}$ l'ensemble des k premiers entiers naturels. Par abus on note $\mathbb{N}_n^p = \mathbb{N}_n^{\mathbb{N}_p}$ l'ensemble des n^p applications de $\mathbb{N}_p$ dans $\mathbb{N}_n$.

Pour tout $\alpha \in \mathbb{N}_n^p$ et tout $(x_1, \cdots, x_p) \in E^p$ avec $x_j = \sum_{i=1}^n x_{i,j}\varepsilon_i$ on pose :

$$e_\alpha(x_1, \cdots, x_p) = \prod_{j=1}^p x_{\alpha(j),j}.$$

Ces e_α sont manifestement des formes p-linéaires. En fait on a :

Proposition 3.1.2 *La famille e_α est une base du module $\mathcal{L}^p(E, A)$, qui est donc libre de rang n^p.*

Démonstration. Partant de la formule évidente $\varepsilon_j = \sum_{i=1}^n \delta_{i,j}\varepsilon_i$ on obtient pour tout $\alpha, \beta \in \mathbb{N}_n^p$:

$$e_\alpha(\varepsilon_{\beta(1)}, \cdots, \varepsilon_{\beta(p)}) = \prod_{j=1}^p \delta_{\alpha(j),\beta(j)} = \begin{cases} 1 \text{ si } \alpha = \beta \\ 0 \text{ sinon.} \end{cases}$$

En particulier si $\sum_{\alpha\in\mathbb{N}_n^p} \lambda_\alpha e_\alpha = 0$ alors pour tout $\beta \in \mathbb{N}_n^p$ on a l'égalité $0 = \sum_{\alpha\in\mathbb{N}_n^p} \lambda_\alpha e_\alpha(\varepsilon_{\beta(1)}, \cdots, \varepsilon_{\beta(p)}) = \lambda_\beta$, d'où la liberté de la famille e_α. Pour voir que e_α est génératrice on constate que tout forme p-linéaire φ s'écrit

$$\varphi = \sum_{\alpha\in\mathbb{N}_n^p} \varphi(\varepsilon_{\alpha(1)}, \cdots, \varepsilon_{\alpha(p)})e_\alpha.$$

Pour démontrer cette dernière égalité une façon de procéder est de "multiplier les indices avec soin", ce que je laisse à la charge des lecteurs. Alternativement on peut constater que cette égalité est immédiate lorsqu'on l'évalue contre une famille de vecteurs de la forme $(\varepsilon_{\beta(1)}, \cdots, \varepsilon_{\beta(p)})$ pour $\beta \in \mathbb{N}_n^p$. Ensuite par récurrence sur p on démontre que deux formes p-linéaires sont égales si et seulement si elles coïncident après évaluation contre ces familles de vecteurs. En effet pour $p = 1$ c'est dire que deux formes linéaires sont égales si et seulement si elles coïncident sur une base. Pour l'hérédité on utilise que pour toute forme p-linéaire φ l'application $(x_2, \cdots, x_p) \mapsto \varphi(\varepsilon_i, x_2, \cdots, x_p)$ est une forme $p-1$ linéaire, donc caractérisée par ses valeurs en les familles de vecteurs de la forme $(\varepsilon_{\beta(2)}, \cdots, \varepsilon_{\beta(p)})$ avec β parcourant $\mathbb{N}_n^p$. □

Remarque : En fait on utilise sans le dire l'isomorphisme canonique $\mathcal{L}^p(E, A) \cong \mathcal{L}(E, \mathcal{L}^{p-1}(E, A))$ défini comme suit :

$$\varphi \in \mathcal{L}^p(E, A) \mapsto (x \mapsto ((x_2, \cdots, x_p) \mapsto \varphi(x, x_2, \cdots, x_p)))$$

Pour le cas particulier des espaces vectoriels sur un corps, cet isomorphisme permet de calculer les dimensions par récurrence (on trouve bien n^p) et dispense de vérifier que la famille e_α est génératrice.

Définition 3.1.3

1. *Une forme p-linéaire φ sur E est dite <u>alternée</u> lorsque pour toute famille $(x_1, \cdots, x_p)$ de E on a $\varphi(x_1, \cdots, x_p) = 0$ dès qu'il existe $i \neq j$ avec $x_i = x_j$.*
2. *L'ensemble de toutes les formes p-linéaires alternées forme un sous-A-module de $\mathcal{L}^p(E, A)$ noté $\mathcal{A}^p(E, A)$.*

Lemme 3.1.4 *Soit $\varphi \in \mathcal{A}^p(E,A)$ alors pour toute permutation $\sigma \in S_p$ de signature $\varepsilon(\sigma)$ on a $\varphi(x_{\sigma(1)},\cdots,x_{\sigma(p)}) = \varepsilon(\sigma)\varphi(x_1,\cdots,x_p)$.*

Démonstration. Puisque les transpositions engendrent S_p et que la signature est un homomorphisme on se ramène au cas particulier $\sigma = (h\ k)$ avec $1 \leq h < k \leq p$ et $\varepsilon(\sigma) = -1$. Alors la forme bilinéaire

$$(x,y) \mapsto \varphi(x_1,\cdots,x_{h-1},x,x_{h+1},\cdots,x_{k-1},y,x_{k+1},\cdots,x_p)$$

est alternée donc antisymétrique par la proposition 5.1.10. □

Proposition 3.1.5 *Si $p > n$ alors $\mathcal{A}^p(E,A) = \{0\}$.*

Démonstration. Soit $\varphi \in \mathcal{A}^p(E,A) \subset \mathcal{L}^p(E,A)$. Alors dans la base e_α associée à la base ε on peut écrire $\varphi = \sum_\alpha \varphi(\varepsilon_{\alpha(1)},\cdots,\varepsilon_{\alpha(p)})e_\alpha$. Mais $p > n$ donc aucune application α n'est injective et on a forcément une répétition $\alpha(i) = \alpha(j)$ pour $i \neq j$. Comme φ est alternée on en déduit $\varphi(\varepsilon_{\alpha(1)},\cdots,\varepsilon_{\alpha(p)}) = 0$. □

3.2 La forme déterminant.

Théorème 3.2.1 (fondamental) *Soit E un A-module libre de rang n. Le module $\mathcal{A}^n(E,A)$ est libre de rang 1 sur A, engendré par la forme déterminant $\det_\varepsilon$ associée à toute base ε (voir définition 3.2.2).*

La suite de ce paragraphe est consacré à la démonstration de ce théorème. Soit $\varepsilon_1,\cdots,\varepsilon_n$ une base de E, et soit e_α la base de $\mathcal{L}^n(E,A)$ associée. Tout $\varphi \in \mathcal{A}^n(E,A)$ s'écrit donc $\varphi = \sum_\alpha \varphi(\varepsilon_{\alpha(1)},\cdots,\varepsilon_{\alpha(n)})e_\alpha$. Mais si $\alpha \in N_n^n$ n'est pas bijective alors α n'est pas injective et puisque φ est alternée on a $\varphi(\varepsilon_{\alpha(1)},\cdots,\varepsilon_{\alpha(n)}) = 0$ pour $\alpha \notin S_n$. D'autre part par la proposition 3.1.4 on a pour tout $\sigma \in S_n$, $\varphi(\varepsilon_{\sigma(1)},\cdots,\varepsilon_{\sigma(n)}) = \varepsilon(\sigma)\varphi(\varepsilon_1,\cdots,\varepsilon_n)$. Cela donne l'identité :

$$\varphi = \sum_{\sigma\in S_n} \varepsilon(\sigma)\varphi(\varepsilon_1,\cdots,\varepsilon_n)e_\sigma = \varphi(\varepsilon_1,\cdots,\varepsilon_n)\sum_{\sigma\in S_n}\varepsilon(\sigma)e_\sigma.$$

Définition 3.2.2 *On appelle <u>déterminant</u> relativement à la base ε et on note $\det_\varepsilon$ la forme n-linéaire*

$$\det_\varepsilon = \sum_{\sigma\in S_n}\varepsilon(\sigma)e_\sigma.$$

Explicitement la forme déterminant évaluée en le système de vecteur $x_j = \sum_{i=1}^n x_{i,j}\varepsilon_i$ s'écrit

$$\det_\varepsilon(x_1,\cdots,x_n) = \sum_{\sigma\in S_n}\varepsilon(\sigma)\prod_{i=1}^n x_{\sigma(i),i}.$$

Cette forme est non nulle (les e_σ sont A-libres) et on a vu que $\mathcal{A}^n(E,A)$ est contenu dans le sous-A-module monogène engendré par $\det_\varepsilon$.

$$\mathcal{A}^n(E,A) \subset A\det_\varepsilon.$$

Pour terminer la preuve du théorème il faut démontrer que $\det_\varepsilon$ est alternée et n'est pas de torsion.

Lemme 3.2.3 *La forme* $\det_\varepsilon$ *n'est pas de torsion.*

Démonstration. En effet si $\lambda \in A$ alors $\lambda\det_\varepsilon(\varepsilon_1, \cdots, \varepsilon_n) = \lambda$. D'où l'équivalence $\lambda\det_\varepsilon = 0 \iff \lambda = 0$. $\square$

Lemme 3.2.4 *La forme* $\det_\varepsilon$ *est alternée.*

Démonstration. On donne d'abord une démonstration plus "conceptuelle" et plus simple en supposant que 2 ne divise pas 0 dans A. On définit une action linéaire (i.e. par automorphisme) du groupe S_p sur $\mathcal{L}^p(E, A)$ en posant

$$(\sigma * \varphi)(x_1, \cdots, x_p) = \varphi(x_{\sigma(1)}, \cdots, x_{\sigma(p)}).$$

Pour cette action l'orbite des e_α est facile à décrire puisque $\sigma * e_\alpha = e_{\alpha\circ\sigma^{-1}}$. Dans notre cas particulier $n = p$, on voit que pour toute permutation γ on a

$$\begin{aligned}\gamma * \det_\varepsilon = \gamma * \sum_{\sigma\in S_n} \varepsilon(\sigma)e_\sigma &= \sum_{\sigma\in S_n} \varepsilon(\sigma)\gamma * e_\sigma = \sum_{\sigma\in S_n} \varepsilon(\sigma)e_{\sigma\circ\gamma^{-1}} \\ &= \sum_{\sigma\in S_n} \varepsilon(\sigma\circ\gamma)e_\sigma = \sum_{\sigma\in S_n} \varepsilon(\sigma)\varepsilon(\gamma)e_\sigma = \varepsilon(\gamma)\det_\varepsilon.\end{aligned}$$

On trouve donc que pour tout transposition τ on a $\tau * \det_\varepsilon = -\det_\varepsilon$. Si 2 ne divise pas zéro dans A les formes bilinéaires antisymétriques sont alternées (c'est la proposition 5.1.9) et la forme $(x, y) \mapsto \det_\varepsilon(x_1, \cdots, x, \cdots, y, \cdots, x_p)$ est antisymétrique puisque échanger x et y revient à faire agir une transposition sur $\det_\varepsilon$. Cela démontre le lemme sous l'hypothèse "2 ne divise pas 0 dans A".

En général on procède de façon plus calculatoire et l'action de S_n qui sert de fil conducteur n'a même pas besoin d'être définie. Soit $(x_1, \cdots, x_n)$ une famille de E avec $x_h = x_k$ pour $h < k$ et écrivons $x_j = \sum_{i=1}^n x_{i,j}\varepsilon_i$. Soit τ la transposition $\tau = (h\ k)$. Alors on a $S_n = A_n \amalg A_n\tau$ et on en déduit :

$$\begin{aligned}\det_\varepsilon(x_1, \cdots, x_n) &= \sum_{\sigma\in S_n} \varepsilon(\sigma)\prod_{i=1}^n x_{\sigma(i),i} \\ &= \sum_{\sigma\in A_n} \varepsilon(\sigma)\prod_{i=1}^n x_{\sigma(i),i} + \sum_{\sigma\in A_n} \varepsilon(\sigma\tau)\prod_{i=1}^n x_{\sigma(\tau(i)),i} \\ &= \sum_{\sigma\in A_n} \varepsilon(\sigma)\prod_{i=1}^n x_{\sigma(i),i} - \sum_{\sigma\in A_n} \varepsilon(\sigma)\prod_{i=1}^n x_{\sigma(\tau(i)),i} \\ &= 0.\end{aligned}$$

En effet $x_{\sigma(\tau(h)),h} = x_{\sigma(k),h} = x_{\sigma(k),k}$, par symétrie $x_{\sigma(\tau(k)),k} = x_{\sigma(h),h}$ et pour tout les autres $i \neq h, k$ on a $x_{\sigma(\tau(i)),i} = x_{\sigma(i),i}$. $\square$

Conclusion : On a vu que $\mathcal{A}^n(E, A)$ est libre monogène engendré par $\det_\varepsilon$. En outre si $\varphi \in \mathcal{A}^n(E, A)$ alors $\varphi = \varphi(\varepsilon_1, \cdots, \varepsilon_n)\det_\varepsilon$. On appelle déterminant du système de vecteurs $(x_1, \cdots, x_n)$ relativement à ε la quantité $\det_\varepsilon(x_1, \cdots, x_n)$.

3.3 Déterminant d'un endomorphisme.

Proposition et définition 3.3.1 *Soit f un endomorphisme de E. Il existe un unique scalaire* $\det(f) \in A$ *tel que pour toute forme linéaire alternée $\varphi \in \mathcal{A}^n(E)$ on ait :*

$$\varphi(f(x_1), \cdots, f(x_n)) = \det(f)\varphi(x_1, \cdots, x_n).$$

Ce scalaire s'appelle le déterminant de f.

Démonstration. On fixe une base ε de E. Puisque f est linéaire la forme

$$(x_1, \cdots, x_n) \mapsto \det_{\varepsilon}(f(x_1), \cdots, f(x_n))$$

est n-linéaire alternée et par le théorème fondamental 3.2.1 il existe un scalaire α tel que :

$$\det_{\varepsilon}(f(x_1), \cdots, f(x_n)) = \alpha \det_{\varepsilon}(x_1, \cdots, x_n),$$

pour tout $(x_1, \cdots, x_n)$ de E^n, c'est-à-dire que le scalaire α vérifie l'identité requise pour la forme $\det_{\varepsilon}$. Mais comme $\mathcal{A}^n(E, A)$ est monogène engendré par $\det_{\varepsilon}$ et comme A est commutatif, le scalaire $\alpha = \det(f)$ convient aussi pour tout $\varphi = \lambda_{\varphi} \det_{\varepsilon} \in \mathcal{A}^n(E, A)$. □

En cours de démonstration on a obtenu la formule (valable pour toute famille $x_1, \cdots, x_n$ telle que $\det_{\varepsilon}(x_1, \cdots, x_n) \in A^{\times}$) :

$$\det(f) = \det_{\varepsilon}(f(x_1), \cdots, f(x_n))\det_{\varepsilon}(x_1, \cdots, x_n)^{-1}.$$

En particulier puisque $\det_{\varepsilon}(\varepsilon_1, \cdots, \varepsilon_n) = 1$ on retrouve la formule (utile pour les calculs pratique mais parfois parachutée en guise de définition) :

$$\det(f) = \det_{\varepsilon}(f(\varepsilon_1), \cdots, f(\varepsilon_n)).$$

Proposition 3.3.2 *Pour tous $f, g \in \mathrm{End}_A(E)$ on a* $\det(fg) = \det(f)\det(g)$. *Pour $f = \mathrm{Id}$ on a* $\det(\mathrm{Id}) = 1$. *En conséquence si f est inversible alors* $\det(f)$ *aussi et on a* $\det(f)^{-1} = \det(f^{-1})$.

Démonstration. Par la proposition-définition 3.3.1 on a pour tous $(x_1, \cdots, x_n)$ de E les identités :

$$\begin{aligned} \det(fg)\det_{\varepsilon}(x_1, \cdots, x_n) &= \det_{\varepsilon}(fg(x_1), \cdots, fg(x_n)) \\ &= \det(f)\det_{\varepsilon}(g(x_1), \cdots, g(x_n)) \, . \\ &= \det(f)\det(g)\det_{\varepsilon}(x_1, \cdots, x_n) \end{aligned}$$

On en déduit $\det(fg) = \det(f)\det(g)$ en prenant $(x_1, \cdots, x_n) = (\varepsilon_1, \cdots, \varepsilon_n)$. Le reste de la proposition est immédiat. □

3.4 Déterminant d'une matrice carré.

Définition 3.4.1 *Pour une matrice carré $M = [m_{i,j}]$ on pose*

$$\det(M) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i),i}.$$

On note aussi

$$\det(M) = |m_{i,j}|\,.$$

Proposition 3.4.2 *Soit $M \in M_n(A)$ une matrice dont on note $[C_1, \cdots, C_n]$ les colonnes.*

1. $\det(M) = \det({}^tM)$
2. det *est une forme n linéaire alternée des colonnes de M (resp. des lignes de M).*
3. *Si ε est la base canonique de A^n, alors* $\det(M) = \det_\varepsilon(C_1, \cdots, C_n)$.
4. *Pour toute base ε de E et tout endomorphisme f de E on a*

$$\det(f) = \det(\mathrm{Mat}_\varepsilon(f)).$$

Démonstration. Pour 1, on part de la formule de définition et on obtient :

$$\begin{aligned}\det(M) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i),i} = \sum_{\sigma \in S_n} \varepsilon(\sigma^{-1}) \prod_{i=1}^{n} m_{\sigma^{-1}(i),i} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} m_{j,\sigma(j)} = \det({}^tM)\end{aligned}$$

Les autres affirmations sont immédiates. □

3.5 Techniques de calculs.

3.5.1 Matrices triangulaires par blocs.

Lemme 3.5.1 *Soit A un anneau commutatif et $M \in M_r(A), N \in M_{r,s}(A)$ et $P \in M_s(A)$ alors*

$$\begin{vmatrix} M & N \\ 0 & P \end{vmatrix} = |M||P|.$$

Proposition 3.5.2 *Soient $M_i \in M_{r_i}(A)$ pour $1 \leq 1 \leq r$. Alors*

$$\begin{vmatrix} A_1 & * & * & * \\ 0 & A_2 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & A_r \end{vmatrix} = \prod_{i=1}^{r} |A_i|$$

Démonstration. Voir p. 35 du livre "algèbre des matrices" de Fresnel.

3.5.2 Pivot de Gauß.

C'est l'une des plus efficaces en général. On a déjà vu ensemble le principe de fonctionnement. le plus rapide est de rendre triangulaire (supérieure ou inférieure) la matrice avec des opérations élémentaires sur les lignes ou les colonnes. Le déterminant de la matrice est alors égal au produit des coefficients de la diagonale par la proposition 3.5.2. Voir p. 38 du même livre.

3.5.3 Développement par rapport à une ligne ou une colonne.

La formule et la démonstration se trouve p. 64 du livre de R. Goblot "algèbre linéaire". Étant donnée $M = [m_{i,j}]$ une matrice de $M_n(A)$ On note $M_{i,j}$ la sous-matrice de M obtenue en enlevant la i-ième ligne et la j-ième colonne de M. Alors pour tout i et tout j on a

$$|M| = \sum_{h=1}^{n} (-1)^{h+j} m_{h,j} |M_{h,j}| = \sum_{k=1}^{n} (-1)^{k+i} m_{i,k} |M_{i,k}|.$$

La seconde égalité s'obtient en transposant la première. Pour la première on utilise la linéarité du déterminant par rapport à la j-ième colonne, puis on effectue des permutation sur lignes et colonnes (d'où le signe $(-1)^{h+j}$) pour se ramener à n matrices de la forme

$$\begin{pmatrix} 1 & * \\ 0 & M_{h,j} \end{pmatrix}.$$

Le déterminant de ces matrices est $|M_{h,j}|$ par la proposition 3.5.2. Ces formules (Cramer) ont un corollaire important (référence la p. 65 du livre de Goblot), c'est celle qui justifie l'introduction de la transposée de la comatrice d'une matrice.

Définition 3.5.3 *On appelle comatrice de M et on note $\widetilde{M}$ la matrice $\widetilde{M} = (\widetilde{m}_{i,j})$ où $\widetilde{m}_{i,j} = (-1)^{i+j} |M_{i,j}|$.*

Proposition 3.5.4

$$^t\widetilde{M} M = M \widetilde{M}^t = \det(M) I_n.$$

Corollaire 3.5.5 *Une matrice $M \in M_n(A)$ est inversible si et seulement si*

$$\det(M) \in A^\times.$$

Corollaire 3.5.6 *Un endomorphisme f de E est inversible si et seulement si*

$$\det(f) \in A^\times.$$

Corollaire 3.5.7 *Une famille de vecteurs $x_1, \cdots, x_n$ de E forme une base de E si et seulement si*

$$\det_{\varepsilon}(x_1, \cdots, x_n) \in A^\times.$$

Dans le cadre des espaces vectoriels sur un corps ces corollaires s'obtiennent sans la proposition 3.5.4 en utilisant le principe (mis en défaut avec les modules) qu'une famille libre de rang maximal est une base.

3.6 Applications classiques.

Il y quatre "applications" de la notion de déterminant qui sont incontournables. J'en donne la liste et une référence pour les trois première mais vous pouvez retrouver ces exemples développés un peu partout.

1. Déterminant de Vandermonde : La matrice de Vandermonde, en liaison avec les polynômes de Lagrange est définie p.42 du livre de Goblot. La formule du déterminant de Vandermonde est donnée en exercice p. 76. La correction de cet exercice peut se trouver en principe n'importe où.
2. Déterminant circulant : voir exercice III.2 p.76 du livre de Goblot.
3. Matrice résultante et son déterminant le résultant de deux polynômes : voir p.43 et 67 du livre de Goblot.
4. Polynôme caractéristique d'un endomorphisme d'un espace vectoriel V sur un corps k : c'est le déterminant d'une matrice à coefficient dans $k[X]$.

Chapitre 4

Dualité.

4.1 Dual d'un espace vectoriel.

Dans la suite E désignera un espace vectoriel sur un corps commutatif k, sans autre restriction de généralité. Je m'attend à ce que les lecteurs soient familiarisés avec l'algèbre linéaire sur le corps des réels $\mathbb{R}$ et éventuellement sur $\mathbb{C}$. L'un des objectifs de ce cours est vous donner accès à d'autres exemples. Les autres corps que vous pouvez utiliser comprennent (sans prétendre être exhaustif) $\mathbb{F}_p$, $\mathbb{Q}$, $\mathbb{Q}[\sqrt{3}]$, tout extension algébrique (voire clotûre) des précédents, les corps de fonctions $k(T)$ à coefficient dans l'un des corps précédents, etc...

Définition 4.1.1 *Soit E un k-espace vectoriel. On appelle dual de E et on note E^* l'espace vectoriel des applications k-linéaires de E dans k. Les éléments de E^* sont appelées formes linéaires.*

Par exemple la projection sur la i-ième composante $p_i \colon k^n \longrightarrow k$ définie par la formule $p_i(x_1, \cdots, x_n) = x_i$ est une forme linéaire.

Notation : Si φ est une forme linéaire sur E et si $x \in E$ on notera

$$\langle \varphi, x \rangle = \varphi(x)$$

Définition 4.1.2 *Soit $(e_i)_{i\in I}$ une base de E. Pour tout $i \in I$ on note e_i^* la i-ième forme linéaire coordonnée définie par*

$$e_i^*(e_j) = \delta_{i,j} = \begin{cases} 1 \ si \ i = j \\ 0 \ si \ i \neq j \end{cases}$$

Remarque : $\forall x \in E, \quad x = \sum_{i\in I} \langle e_i^*, x \rangle e_i$, et cette somme est finie.

Exercice 4.1 On suppose k de caractéristique nulle (autrement dit $n \neq 0$ dans k pour tout entier n non nul).

1. On prend $E = k[X]$ l'espace vectoriel des polynômes à coefficient dans k muni de sa base canonique $(e_n = X^n)_{n\in\mathbb{N}}$. On note $P^{(n)}(X)$ la n-ième dérivée formelle du polynôme $P(X)$. Vérifier que la forme linéaire coordonnée est donnée par la formule
$$e_n^*(P) = \frac{P^{(n)}(0)}{n!}$$

2. On prend $E = k_n[X]$ l'espace vectoriel des polynômes à coefficient dans k de degré inférieur ou égal à n. Déterminer la forme linéaire coordonnée d'indice j de la base

$$f_0(X) = 1,\ f_1(X) = X,\ \cdots,\ f_n(X) = \frac{X(X-1)\cdots(X-n+1)}{n!}.$$

Exemple : Soit E un k-espace vectoriel muni d'une base $(e_i)_{i\in I}$. Soit φ la forme linéaire définie par $\forall i \in I,\ \varphi(e_i) = 1$. Si I est fini alors $\varphi = \sum_{i\in I} e_i^*$ est dans l'espace vectoriel engendré par les e_i^*. Si au contraire I n'est pas fini alors φ n'est pas combinaison linéaire finie des e_i^* : en effet pour tout sous-ensemble fini $J \subset I$ et tout $i \in I \setminus J$ on a $\sum_{j\in J} \lambda_j e_j^*(e_i) = 0 \neq 1 = \varphi(e_i)$.

Théorème 4.1.3 *On se donne une base de E notée $\mathcal{B} = (e_i)_{i\in I}$. On considère la famille $\mathcal{B}^* = (e_i^*)_{i\in I}$. Alors $\mathcal{B}^*$ est une famille libre de E^*. C'est une base de E^* si et seulement si E est de dimension finie. Auquel cas $\mathcal{B}^*$ est appelée base duale de $\mathcal{B}$.*

Démonstration. On part d'une relation linéaire finie $\sum_{j\in J} \lambda_j e_j^* = 0$. Alors pour tout $k \in J$ on a $\lambda_k = \sum_{j\in J} \lambda_j e_j^*(e_k) = 0$. Cela montre que $\mathcal{B}^*$ est libre. Si en outre I est fini alors $\mathcal{B}$ est générateur puisque tout φ de E^* s'écrit $\varphi = \sum_{i\in I} \varphi(e_i) e_i^*$. Enfin si I n'est pas fini le contre-exemple qui précède le théorème s'applique et $\mathcal{B}^*$ n'est pas générateur. □

Proposition 4.1.4 *Le crochet de dualité $\langle \varphi, x\rangle$ satisfait les propriétés suivantes :*

1. *les applications $\varphi \mapsto \langle \varphi, x\rangle$ et $x \mapsto \langle \varphi, x\rangle$ sont k-linéaires.*
2. *La forme linéaire φ est nulle si et seulement si pour tout $x \in E$, $\langle \varphi, x\rangle = 0$.*
3. *Un élément $x \in E$ est nul si et seulement si pour tout forme linéaire $\varphi \in E^*$, $\langle \varphi, x\rangle = 0$.*

Démonstration. 1. et 2. sont évidents. Pour 3. le sens direct est immédiat. On suppose x non nul et on doit trouver une forme linéaire qui ne s'annule pas en x. Mais puisque x est non nul le théorème de la base incomplète fournit une base commençant par x, et donc une forme linéaire x^* associée à cette base vérifiant $x^*(x) = 1 \neq 0$. □

Proposition 4.1.5 *Soit $k^{\mathbb{N}}$ l'espace des suites à valeurs dans k et soit $k^{(\mathbb{N})}$ l'espace des suites ultimement nulles (à valeur dans k aussi). Le dual de $k^{(\mathbb{N})}$ est $k^{\mathbb{N}}$.*

Démonstration. Sur $k^{(\mathbb{N})}$ on dispose de la base canonique e_i définie par $e_i(n) = \delta_{i,n}$. Grâce à cette base on voit que l'application linéaire $\varphi \mapsto (\varphi(e_i))_{i\in\mathbb{N}}$ est un isomorphisme de $(k^{(\mathbb{N})})^*$ sur $k^{\mathbb{N}}$. □

4.2 bidual

Définition 4.2.1 *Soit E un espace vectoriel sur k. Le bidual de E est le dual du dual de E. On le note E^{**}.*

Proposition 4.2.2 *Soit E un espace vectoriel.*

1. *A tout x de E on associe la forme linéaire $\delta_x : E^* \longrightarrow k$ définie par $\delta_x(\varphi) = \varphi(x)$. L'application $x \mapsto \delta_x$ est linéaire et injective (on l'appelle l'injection canonique d'un espace dans son bidual). En dimension finie cette injection canonique est un isomorphisme.*
2. *On se donne une base $\mathcal{B} = (e_i)_{i\in I}$ de E. Alors l'application linéaire $e_i \mapsto e_i^*$ est une injection de E dans E^*. C'est un isomorphisme en dimension finie.*

Démonstration. Lorsque E est de dimension fini on a vu avec les bases duales que $\dim(E) = \dim(E^*)$. Les isomorphies en dimensions finies sont donc conséquences des injections et de l'égalité des dimensions. L'application du 1. est clairement bien définie et linéaire. Son injectivité provient du 3 de la proposition 4.1.4. L'application du 2. est définie par linéarité à partir d'une base. Cette application est injective puisque la famille $\mathcal{B}^*$ est libre (théorème 4.1.3). □

Remarque : L'application $E \longrightarrow E^{**}$ est dite canonique puisqu'elle ne dépend pas du choix d'une base sur E. Elle est intrinsèque à E. Par contre l'application du 2. dépend du choix de la base. Par exemple dans $\mathbb{Q}^2$ l'image du vecteur $(0, 1)$ change selon qu'on le complète en une base avec le vecteur $(1, 1)$ ou bien avec le vecteur $(1, 0)$.

4.3 Orthogonalité

Dans ce paragraphe, sauf mention explicite du contraire, E est de dimension quelconque (finie ou pas). La plupart des résultats sont énoncés dans la littérature en supposant la dimension finie, mais ils restent valables en toute généralité et cette hypothèse ne simplifie même pas les preuves.

Définition 4.3.1

1. *Soit F un sous-espace de E, on appelle orthogonal de F dans E^* et on note $F^\perp$ le sous-espace de E^* des formes linéaires qui s'annulent sur F.*

$$F^\perp = \{\varphi \in E^*;\ \forall x \in F,\ \varphi(x) = 0\}$$

2. *Soit G un sous-espace vectoriel de E^*, on appelle orthogonal de G dans E et on note G^0 l'intersection dans E des noyaux des éléments de G.*

$$G^0 = \{x \in E;\ \forall \varphi \in G,\ \varphi(x) = 0\}$$

Proposition 4.3.2 *Soit E un espace vectoriel sur k, soient F et F' deux sous-espace de E, et soit G un sous-espace de E^* alors :*

1. $(F \subset F') \implies \left((F')^\perp \subset F^\perp\right)$
2. $(F + F')^\perp = F^\perp \cap (F')^\perp$
3. $F^\perp + (F')^\perp = (F \cap F')^\perp$
4. *Toute forme linéaire de $F^\perp$ se factorise en une forme linéaire $\bar{\varphi}$ de E/F. Cela définit un isomorphisme canonique $F^\perp \xrightarrow{\sim} (E/F)^*$.*
5. $(F^\perp)^0 = F$

6. $G \subset (G^0)^\perp$.

Démonstration. 1. suit directement de la définition.

Soit $\varphi \in (F+F')^\perp$ alors puisque $F \cup F' \subset F+F'$ on a $\varphi(F) = \varphi(F') = 0$, d'où $\varphi \in F^\perp \cap (F')^\perp$. Réciproquement soit $\psi \in F^\perp \cap (F')^\perp$, et soit $x \in F+F'$. Alors x peut s'écrire $x = f + f'$ avec $f \in F$ et $f' \in F'$. Il suit $\psi(x) = \psi(f) + \psi(f') = 0$. Donc $\psi \in (F+F')^\perp$. Cela démontre 2.

Puisque $F \cap F' \subset F$ on a $F^\perp \subset (F \cap F')^\perp$. De même on montre l'inclusion $(F')^\perp \subset (F \cap F')^\perp$ et il suit $F^\perp + (F')^\perp \subset (F \cap F')^\perp$. Réciproquement soit $\varphi \in (F \cap F')^\perp$. Par le théorème de la base incomplète (valable aussi en dimension infinie) on peut écrire $E = (F \cap F') \oplus F_s \oplus F'_s \oplus S$ où l'espace S est un supplémentaire de $F+F'$ dans E, l'espace F_s un supplémentaire de $(F \cap F')$ dans F et l'espace F'_s un supplémentaire de $(F \cap F')$ dans F'. Suivant cette décomposition de E, la forme linéaire φ s'écrit comme somme[1] $\varphi = t + u + v + w$ avec $t \in (F \cap F')^*$, $u \in F_s^*$, $v \in (F'_s)^*$, et $w \in S^*$. Comme $\varphi \in (F \cap F')^\perp$ on a $t = 0$. On remarque que $u \in (F')^\perp$ tandis que $v + w \in (F)^\perp$, ce qui donne $\varphi = u + (v+w) \in (F')^\perp + F^\perp$ et démontre 3.

Par définition si $\varphi \in F^\perp$ alors $F \subset \operatorname{Ker}\varphi$ et donc φ se factorise en $(\bar{\varphi}\colon E/F \longrightarrow k) \in (E/F)^*$. On note $\pi_F\colon E \longrightarrow E/F$ la surjection canonique. Pour Montrer que l'application $\varphi \mapsto \bar{\varphi}$ est un isomorphisme de $F^\perp$ sur $(E/F)^*$ on doit vérifier :

1. que cette application est linéaire : Exercice.
2. que cette application est surjective, mais si $\psi \in (E/F)^*$ alors l'application linéaire $\psi \circ \pi_F \in F^\perp \subset E^*$ est un antécédent de ψ.
3. que cette application est injective, mais si $\bar{\varphi}$ est la forme linéaire nulle, alors on a $\varphi = \bar{\varphi} \circ \pi_F = 0$.

Cela démontre 4.

Soit $x \in F$. Alors pour tout $\varphi \in F^\perp$ on a $\varphi(x) = 0$ et donc $x \in (F^\perp)^0$. D'où l'inclusion $F \subset (F^\perp)^0$. En dimension fini on conclut la démonstration avec l'égalité des dimensions (voir la proposition 4.3.3 qui suit). En général pour montrer l'égalité $F = (F^\perp)^0$ on procède par l'absurde et on suppose l'existence d'un $x \in ((F^\perp)^0 \setminus F)$. Soit $\mathcal{F}$ une base de F, et $\mathcal{B}$ une base de E qui complète $\mathcal{F} \cup \{x\}$. Alors la forme linéaire x^* relative à la base $\mathcal{B}$ appartient à $F^\perp$ mais ne s'annule pas sur x : cela contredit $x \in (F^\perp)^0$. Cela démontre 5.

Soit $\varphi \in G$. Alors pour tout $x \in G^0$ on a $\varphi(x) = 0$ et donc $\varphi \in (G^0)^\perp$. D'où l'inclusion 6. □

Contre-exemple : On prend $E = k^{(\mathbb{N})}$ muni de sa base canonique $(e_i)_{i\in\mathbb{N}}$ comme dans la proposition 4.1.5. Soit G le sous espace de $k^{\mathbb{N}}$ engendré par les e_i^*. Alors $G^0 = 0$ et donc $(G^0)^\perp = E^*$, mais $G \subsetneq E^*$ (voir l'exemple qui suit l'exercice 4.1)

En dimension fini on peut identifier E à son bidual. Cela permet de traduire dans E^* la proposition 4.3.2 et donne le complément d'information ci-dessous.

Proposition 4.3.3 *Soit E un espace vectoriel de dimension finie, soit F un sous-espace de E et soit G un sous-espace de E^*.*

1. Étant donnée une écriture de E en somme directe $E = A \oplus B$ on peut considérer A^* comme sous-espace de E^* en prolongeant les éléments de A^* par 0 sur B. Cette injection $A^* \hookrightarrow E^*$ n'est pas canonique puisqu'elle dépend du choix du supplémentaire B. Par contre pour A et B fixé et pour tout V on a un isomorphisme canonique $\operatorname{Hom}(A \oplus B, V) \cong \operatorname{Hom}(A,V) \oplus \operatorname{Hom}(B,V)$, et cela se généralise aux sommes directes de plus de deux espaces.

1. $\dim F + \dim F^\perp = \dim E$.
2. $\dim G + \dim G^0 = \dim E^*$.
3. $G = (G^0)^\perp$.

Démonstration. En dimension finie le théorème du rang appliqué à la surjection canonique $\pi_F \colon E \longrightarrow E/F$ donne l'égalité $\dim E = \dim(E/F) + \dim F$, tandis que le 4. de la proposition 4.3.2 donne l'égalité $\dim(E/F) = \dim F^\perp$. Cela démontre 1.

la propriété 2 est la propriété duale de 1. Puisque la dualité est parfaite en dimension finie, cela suffit à démontrer 2. Pour se familiariser à cet exercice on détaille la preuve pour cette fois. Soit $\imath \colon E \longrightarrow E^{**}$ l'isomorphisme canonique. Alors par définition même on a $\imath(G^0) = G^\perp$. Il suit $\dim(G^0) = \dim G^\perp$ puis avec l'égalité 1. dans E^* on en déduit 2.

L'affirmation 6. de la proposition 4.3.2 fournit l'inclusion $G \subset (G^0)^\perp$. L'égalité suit puisque les dimensions sont les mêmes d'après 1. et 2. Cela démontre 3. □

Lemme 4.3.4 *Soit f et soit $(f_i)_{1\leq i\leq n}$ des formes linéaires sur E. Alors f est combinaison linéaire des f_i si et seulement si $\cap_{i=1}^n \operatorname{Ker} f_i \subset \operatorname{Ker} f$.*

Démonstration. Le sens direct est évident. Pour montrer la réciproque on suppose $\cap_{i=1}^n \operatorname{Ker} f_i \subset \operatorname{Ker} f$. On vérifie d'abord que pour une seule forme linéaire on a $\langle f\rangle = (\operatorname{Ker} f)^\perp$. Par le 6 de la proposition 4.3.2 on sait déjà que $\langle f\rangle \subset (\langle f\rangle^0)^\perp = (\operatorname{Ker} f)^\perp$. Réciproquement, soit $\varphi \in (\operatorname{Ker} f)^\perp$ non nulle et soit $x \in E$ tel que $\varphi(x) \neq 0$. Alors puisque le sous-espace $\operatorname{Ker} f = \operatorname{Ker}\varphi$ admet $\langle x\rangle$ comme supplémentaire dans E on a $f(x) \neq 0$ et aussi $\varphi = (\varphi(x)/f(x))f \in \langle f\rangle$. D'où l'égalité. (argument plus rapide à méditer : on passe aux quotient par $\operatorname{Ker} f$ et alors on est en dimension 1 et le 3. de la proposition 4.3.3 s'applique). Avec l'identité $\langle f\rangle = (\operatorname{Ker} f)^\perp$ et dans l'ordre le 1. et le 3. de la proposition 4.3.2 on en déduit

$$\langle f\rangle = (\operatorname{Ker} f)^\perp \subset (\bigcap_{i=1}^n \operatorname{Ker} f_i)^\perp = \sum_{i=1}^n (\operatorname{Ker} f_i)^\perp = \langle f_1, \cdots, f_n\rangle$$

□

Proposition 4.3.5 (formules de Cramer) *On suppose E de dimension finie n et on se donne une forme n-linéaire alternée ϕ non nulle sur E (i.e. telle que $\phi(e_1, \cdots, e_n) \neq 0$). Alors la forme linéaire*

$$x \mapsto \frac{\phi(e_1, \cdots, e_{i-1}, x, e_{i+1}, \cdots, e_n)}{\phi(e_1, \cdots, e_n)}$$

est la forme linéaire e_i^*. On retrouve ainsi les formules de Cramer

$$x = \sum_{i=1}^n \frac{\phi(e_1, \cdots, e_{i-1}, x, e_{i+1}, \cdots, e_n)}{\phi(e_1, \cdots, e_n)} e_i$$

Démonstration. C'est immédiat. □

4.4 Problème : codimension des noyaux

Il s'agit de démontrer la proposition 4.4.2.

Définition 4.4.1 *Soit E un espace vectoriel et $F \subset E$ un sous-espace. On appelle codimension de F dans E et on note* $\mathrm{codim}(F)$ *la dimension du quotient E/F.*

Proposition 4.4.2 *Soit E un espace vectoriel sur k.*

1. *Soient $f_1, \cdots, f_l \in E^*$. La dimension de $\langle f_1, \cdots, f_l \rangle$ est égal à la codimension de $\cap_{i=1}^{l} \mathrm{Ker}\, f_i$.*
2. *Réciproquement si H est un sous-espace de codimension r il existe r formes linéaires indépendantes $f_1, \cdots, f_r$ telles que $H = \cap_{i=1}^{r} \mathrm{Ker}\, f_i$*

Indication :

1. Sens direct :
   (a) Utiliser le lemme 4.3.4 pour montrer qu'on peut supposer les f_i linéairement indépendants.
   (b) Passer au quotient par $\cap_{i=1}^{l} \mathrm{Ker}\, f_i$ pour se ramener au cas de l formes linéaires indépendantes et $\dim E = l$.
   (c) conclure.
2. Sens réciproque : utiliser l'identification $H^{\perp} \xrightarrow{\sim} (E/H)^*$.

4.5 Transposée d'une application linéaire.

Définition 4.5.1 *Soient E et F deux espaces vectoriels et $f\colon E \longrightarrow F$ une application linéaire. La transposée de f, notée f^t est l'application linéaire*

$$f^t\colon F^* \longrightarrow E^*$$

$$\varphi \longmapsto \varphi \circ f.$$

Proposition 4.5.2 *Soient E et F deux espaces vectoriels et $f\colon E \longrightarrow F$ une application linéaire.*

1. $\forall x \in E, \forall \varphi \in F, \langle \varphi, f(x)\rangle = \langle f^t(\varphi), x\rangle$.
2. *Si $g\colon E \longrightarrow F$ est une autre application linéaire alors $(f+g)^t = f^t + g^t$.*
3. *Si $g\colon F \longrightarrow G$ est une autre application linéaire alors $(g\circ f)^t = f^t \circ g^t$*
4. *Si f est inversible alors f^t aussi et on a $(f^t)^{-1} = (f^{-1})^t$.*
5. $(\operatorname{Ker} f)^\perp = \operatorname{Im}(f^t)$ *et* $(\operatorname{Im} f)^\perp = \operatorname{Ker}(f^t)$.
6. *Le rang de f est égal à celui de f^t.*

Démonstration. 1. et 2. sont immédiats.

Par définition, pour $\varphi \in G^*$ on a $(g\circ f)^t(\varphi) = \varphi\circ g\circ f = f^t(\varphi\circ g) = f^t(g^t(\varphi)) = (f^t\circ g^t)(\varphi)$.Cela démontre 3.

Pour tout espace vectoriel E on a $(\mathrm{Id}_E)^t = \mathrm{Id}_{E^*}$. L'assertion 4. se déduit donc du 3.

Soit $\varphi \in \operatorname{Im}(f^t)$. Alors il existe $\psi \in F^*$ telle que $\varphi = \psi\circ f$, et donc pour tout $x \in \operatorname{Ker} f$ on a $\varphi(x) = \psi(f(x)) = \psi(0) = 0$. Donc $\varphi \in (\operatorname{Ker} f)^\perp$. Réciproquement soit $\varphi \in (\operatorname{Ker} f)^\perp$. Alors $\operatorname{Ker} f \subset \operatorname{Ker}\varphi$ et donc par factorisation il existe $\psi \in F^*$ telle que $\varphi = \psi\circ f$. Donc $\varphi = f^t(\psi) \in \operatorname{Im}(f^t)$. On a démontré la première égalité. Soit $\varphi \in (\operatorname{Im} f)^\perp$. Alors pour tout $x \in E$ on a $f^t(\varphi)(x) = \varphi(f(x)) = 0$. Donc $\varphi \in \operatorname{Ker}(f^t)$. Réciproquement soit $\varphi \in \operatorname{Ker}(f^t)$ et soit $y = f(x) \in \operatorname{Im} f$. Alors $\varphi(y) = \varphi(f(x)) = f^t(\varphi)(x) = 0$. Donc $\varphi \in (\operatorname{Im} f)^\perp$ et cela donne la deuxième égalité du 5.

On a déjà l'égalité $\operatorname{Im}(f^t) = (\operatorname{Ker} f)^\perp$. Par le 4. de la proposition 4.3.2 on sait que $(\operatorname{Ker} f)^\perp \cong (E/\operatorname{Ker} f)^*$. Par factorisation on obtient $(E/\operatorname{Ker} f)^* \cong (\operatorname{Im} f)^*$ et on a donc un isomorphisme canonique $\operatorname{Im}(f^t) \cong (\operatorname{Im} f)^*$. Par définition le rang de f est la dimension de son image. Ainsi f est de rang fini si et seulement si f^t est de rang fini, et dans ce cas la dimension de $(\operatorname{Im} f)^*$ est égale à celle de $\operatorname{Im} f$ et on a bien l'égalité de rangs annoncée.

□

4.6 Quelques calculs matriciels.

4.6.1 Matrice transposée

Définition 4.6.1 *Soit*

$$A = [a_{i,j}] \in \mathcal{M}_{m,n}(k)$$

une matrice à m lignes et n colonnes et à coefficients dans k. On appelle matrice transposée de A et on note A^t la matrice à n lignes et m colonnes

$$A^t = [a_{j,i}] \in \mathcal{M}_{n,m}(k).$$

Visuellement on obtient A^t à partir de A en appliquant une symétrie par rapport à la diagonale principale. Par exemple

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^t = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}.$$

Proposition 4.6.2 *Soient E et F deux espaces vectoriels de dimension finies, munis des bases $e_1, \cdots e_n$ pour E et $f_1, \cdots f_m$ pour F. Soient $e_1^*, \cdots, e_n^*$ et $f_1^*, \cdots f_m^*$ les bases duales. Soit $f\colon E \longrightarrow F$ une application linéaire. Alors la matrice transposée de la matrice de f relativement aux bases (e_i) et (f_j) est la matrice de l'application transposée f^t relativement aux bases (f_j^*) et (e_i^*).*

Démonstration. Exercice. □

4.6.2 Une utilisation du pivot de Gauß.

Soit $e_1, \cdots, e_n$ une base de E et soit F un sous-espace de E engendré par les vecteurs $u_1, \cdots, u_l$. Soit M la matrice $n \times l$ des coordonnées des u_i dans la base des e_j, c'est-à-dire

$$(u_1, \cdots, u_l) = (e_1, \cdots, e_n)M \quad .$$

En suivant l'algorithme du pivot de Gauß on peut vérifier qu'il existe une matrice inversible $A \in GL_n(k)$, une matrice de permutation $B \in GL_l(k)$ et une matrice triangulaire supérieure inversible T avec $\operatorname{rang}(T) = \dim F$ telle que

$$AMB = \begin{bmatrix} T & * \\ 0 & 0 \end{bmatrix}.$$

La matrice A correspond à une suite d'opération sur les lignes de M tandis que B correspond à une suite de permutations des colonnes de M survenant lorsqu'en cours d'algorithme on rencontre une colonne nulle avant d'avoir terminé. Soit M' la matrice $n \times (l+1)$ obtenue à partir de M en rajoutant la colonne $X_1, \cdots, X_n$, c'est-à-dire

$$(u_1, \cdots, u_l, \sum_{i=1}^{n} X_i e_i) = (e_1, \cdots, e_n)M' \ ,$$

et soit $B' \in GL_{l+1}(k)$ la matrice représentant la même permutation que B mais vue dans $\mathfrak{S}_{l+1}$ (avec $l+1$ comme point fixe). Autrement dit B' laisse fixe la dernière colonne de M' et permute les autres colonne comme B. Alors on a

$$AM'B' = \begin{bmatrix} T & * & * \\ 0 & 0 & C \end{bmatrix}.$$

Alors C est une matrice à $n-r$ lignes et une colonne et chaque ligne de C sera de la forme $\sum a_i X_i$. Ces lignes de C fournissent une base de $F^\perp$, c'est-à-dire un système

fondamental d'équation pour F en remplaçant les X_i par les e_i^* dans chaque ligne de C. En effet les colonnes de T donnent une base de F, la colonne de X_i n'est rien d'autre qu'une façon de garder en mémoire les opérations effectuée sur les lignes de M. Les formes linéaires obtenue à partir de C annulent cette base de F parce que la matrice en dessous de T dans $AM'B'$ est nulle. Ces formes linéaires sont indépendantes parce que A est inversible. On conclut avec les dimensions.

Exercice 4.2 *Dans $\mathbb{Q}^3$ avec sa base canonique on étudie les vecteurs $u_1 = (1,1,1)$, $u_2 = (2,1,1)$ et $u_3 = (3,2,2)$. Donner une base de F et de $F^\perp$ (N.B. : lorsqu'on applique l'algorithme décrit ci-dessus on ne calcule en aucun cas les matrices A ni B ni B')*

Exercice 4.3 *Trouver un exemple intéressant avec 3 vecteurs de $\mathbb{Q}^4$ engendrant un espace F de dimension 2 et réduire cet exemple suivant l'algorithme ci-dessus.*

4.7 Dualité dans les espaces euclidiens.

Dans cette section $k = \mathbb{R}$, l'espace E est un espace euclidien de dimension finie n et pour x, y dans E on note (x, y) leur produit scalaire. Deux bases $\mathcal{B}$ et $\mathcal{B}'$ sont dites de même sens lorsque $\det_{\mathcal{B}}(\mathcal{B}') > 0$. Sinon elles sont dites de sens contraire. Être de même sens est une relation d'équivalence sur l'ensemble des bases de E. Il y a deux classes d'équivalences pour cette relation. Le choix d'une de ces classes d'équivalence s'appelle une orientation de E. A orientation fixée les bases de cette orientation sont dites directes, les autres indirectes.

Proposition 4.7.1 *L'application $x \mapsto (x, .)$ est un isomorphisme de E sur son dual E^*. En particulier pour toute forme linéaire φ il existe un et un seul vecteur x_φ tel que $\langle \varphi, . \rangle = (x_\varphi, .)$.*

Démonstration. Puisque le produit scalaire est bi-linéaire $(x, .) \in E^*$ pour tout x de E et l'application $x \mapsto (x, .)$ est linéaire. Puisque le produit scalaire est non dégénéré cette application est injective. Puisque E et E^* ont même dimension cette application est un isomorphisme. □

On fixe une orientation sur E. Cela permet entre autre de définir le produit vectoriel de $n-1$ éléments :

Lemme 4.7.2 *Soit $\mathcal{B} = \{e_1, \cdots, e_n\}$ une base directe de E et $x_1, x_2, \cdots, x_{n-1}$ des vecteurs de E. Il existe un unique vecteur x dans E tel que*

$$\forall u \in \mathbb{R}^n,\ (x, u) = \det{}_{\mathcal{B}}(x_1, \cdots, x_{n-1}, u).$$

Démonstration. Comme le déterminant est multilinéaire, l'application

$$\varphi \colon u \mapsto \det{}_{\mathcal{B}}(x_1, \cdots, x_{n-1}, u)$$

est une forme linéaire (nulle si et seulement si les x_i sont liés). Alors l'unique vecteur $x = x_\varphi$ de la proposition 4.7.1 correspondant à cette forme linéaire φ convient.

Définition 4.7.3 *Soit $\mathcal{B}$ une base* **orthonormale directe** *de $E = \mathbb{R}^n$, et Soient $x_1, x_2, \cdots, x_{n-1}$ des vecteurs de E. On appelle produit vectoriel des x_i et on note*

$$x_1 \wedge x_2 \wedge \cdots \wedge x_{n-1}$$

l'élément de E défini par le lemme 4.7.2

Proposition 4.7.4 *Le produit vectoriel sur E vérifie les propriétés :*

1. *Soit $\sigma \in \mathfrak{S}_{n-1}$ une permutation, alors*

$$x_{\sigma(1)} \wedge x_{\sigma(2)} \wedge \cdots \wedge x_{\sigma(n-1)} = \varepsilon(\sigma)\ (x_1 \wedge x_2 \wedge \cdots \wedge x_{n-1}).$$

2. $$(\lambda x_1 + \mu x_1') \wedge \cdots \wedge x_{n-1} = \lambda\ (x_1 \wedge \cdots \wedge x_{n-1}) + \mu\ (x_1' \wedge \cdots \wedge x_{n-1}).$$

3. *$x_1 \wedge \cdots \wedge x_{n-1} = 0$ si et seulement si les x_i sont liés.*

Démonstration. Ces propriétés sont des conséquences de l'unicité du produit vectoriel et des propriétés analogues du déterminant. □

Exercice 4.4 *Écrire les détails de la démonstration de la proposition 4.7.4.*

Chapitre 5

Formes quadratiques et hermitiennes.

5.1 Généralités sur les formes sesquilinéaires.

Soit k un corps et soit σ un automorphisme de k. On note $\sigma(x) = x^\sigma$.

Définition 5.1.1 *Soit E un k-espace vectoriel. Une application*

$$f\colon E \times E \longrightarrow k,$$

est appelée forme sesquilinéaire ou s'il faut préciser forme σ-sesquilinéaire lorsque

1. *$\forall y \in E$, $x \mapsto f(x,y)$ est k-linéaire.*
2. *$\forall x \in E$, $y \mapsto f(x,y)$ est semi-linéaire, c'est-à-dire additive et vérifiant pour tout x, y dans E et tout λ dans k, $f(x,\lambda y) = \lambda^\sigma f(x,y)$.*

Exemple :

1. Pour $\sigma = \mathrm{Id}$ on retrouve les formes bilinéaires.
2. Pour $k = \mathbb{C}$, $E = \mathbb{C}^n$ et σ la conjugaison complexe le produit hermitien canonique
$$\langle (z_1, \cdots, z_n), (z'_1, \cdots, z'_n) \rangle = \sum_{i=1}^{n} z_i \sigma(z'_i),$$
est σ-sesquilinéaire.

Dans la suite de ce chapitre on va supposer E de dimension finie $n = \dim E$.
Représentation matricielle Soit $e_1, \cdots e_n$ une base de E et soit M la matrice $M = [f(e_i, e_j)]_{1 \leq i,j \leq n}$. Alors pour $u = \sum x_i e_i$ et $v = \sum y_i e_i$ on a

$$f(u,v) = (x_1, \cdots, x_n) \quad M \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}^\sigma$$

Cette matrice M représente aussi l'application semi-linéaire $\bar{f}\colon E \longrightarrow E^*$ définie par $\bar{f}(y) = f(.,y)$ dans les bases $e_1, \cdots, e_n$ et sa duale :

$$M = \mathrm{Mat}_{e^*,e}(\bar{f}) \ .$$

Attention l'application $\bar{f}$ est semi-linéaire. En conséquence pour $x \in E$ représenté par la matrice colonne X telle que $x = (e_1, \cdots, e_n)X$ les coordonnée de $\bar{f}(x)$ ne s'obtiennent pas avec le produit matriciel $M\ X$ mais avec le produit matriciel $M\ X^\sigma$ (en notant $[m_{i,j}]^\sigma = [m_{i,j}^\sigma]$).

Définition 5.1.2 *On dit que f est non dégénérée si $\bar{f}$ est injective. Puisqu'on est en dimension finie cela revient à dire que $\bar{f}$ est bijective. Le sous-espace* $\operatorname{Ker} \bar{f}$ *est aussi appelé noyau de f, ou selon les auteurs le radical de f (alors noté* $\operatorname{rad} f$*).*

Évidemment si $M = \operatorname{Mat}_{e^*,e}(\bar{f})$ est la matrice de f dans la base $e_1, \cdots, e_n$ alors

$$\operatorname{Ker} \bar{f} = 0 \iff \det M \neq 0.$$

Cependant ce déterminant $\det M$ n'est pas un invariant de f puisqu'il dépend de la base $e_1, \cdots, e_n$. Soit $(u_1, \cdots, u_n) = (e_1, \cdots, e_n)P$ une autre base de E. Alors $P = \operatorname{Mat}_e(u) = \operatorname{Mat}_{e,u}(\operatorname{Id}_E)$ et donc ${}^tP = \operatorname{Mat}_{u^*,e^*}(\operatorname{Id}_{E^*})$. Par le changement de base habituel on obtient

$$\operatorname{Mat}_{u^*,u}(\bar{f}) = \operatorname{Mat}_{u^*,e^*}(\operatorname{Id}_{E^*}) \operatorname{Mat}_{e^*,e}(\bar{f})(\operatorname{Mat}_{e,u}(\operatorname{Id}_E))^\sigma = {}^tPMP^\sigma.$$

Alors la matrice de f relativement à cette nouvelle base est ${}^tPMP^\sigma$ dont le déterminant est

$$\det({}^tPMP^\sigma) = \det(M)(\det(P)^{1+\sigma}).$$

Cela montre que l'élément $\det(M)(k^\times)^{1+\sigma} \in \{0\} \cup k^\times/(k^\times)^{1+\sigma}$ est un invariant de la forme f elle-même.

Définition 5.1.3 *Le discriminant de f est la classe* $\det(M)(k^\times)^{1+\sigma}$ *dans le quotient* $k/(k^\times)^{1+\sigma} := \{0\} \cup k^\times/(k^\times)^{1+\sigma}$.

Définition 5.1.4 *Une forme sesquilinéaire est réflexive lorsque*

$$\forall x, y \in E,\ f(x,y) = 0 \iff f(y,x) = 0.$$

Définition 5.1.5 *Une forme bilinéaire $f\colon E \times E \longrightarrow k$ est dite symétrique lorsque*

$$\forall x, y \in E,\ f(x,y) = f(y,x).$$

Lorsque f est une forme bilinéaire symétrique l'application $q(x) = f(x,x)$ est la forme quadratique associée à f et la forme f est la forme polaire associée à q.

Proposition 5.1.6 *On suppose k de caractéristique différente de 2. Soit f une forme bilinéaire symétrique et q sa forme quadratique associée, alors*

$$f(x,y) = \frac{q(x+y) - (q(x) + q(y))}{2}$$

Démonstration. Exercice. □

Définition 5.1.7 *Une forme bilinéaire f est dite anti-symétrique lorsque*

$$\forall x, y \in E,\ f(x,y) = -f(y,x).$$

Définition 5.1.8 *Une forme sesquilinéaire f est dite alternée lorsque*

$$\forall x \in E,\ f(x,x) = 0.$$

Proposition 5.1.9 *Si f est une forme bilinéaire anti-symétrique et si k n'est pas de caractéristique 2, alors f est alternée.*

Démonstration. Puisque f est anti-symétrique on a $f(x,x) = -f(x,x)$ et il suit $2f(x,x) = 0$ pour tout x de E. $\square$

Proposition 5.1.10 *Soit f une forme sesquilinéaire alternée non nulle. Alors $\sigma = \mathrm{Id}$ et f est bilinéaire et anti-symétrique.*

Démonstration. Soient $x, y \in E$, alors puisque f est alternée on a

$$0 = f(x+y, x+y) = f(x,x) + f(x,y) + f(y,x) + f(y,y),$$

et il suit $f(x,y) = -f(y,x)$ pour tout x et tout y dans E. Puisque f est non nulle il existe x, y dans E tels que $f(x,y) \neq 0$. En utilisant l'anti-symétrie de f on obtient $\lambda^\sigma f(x,y) = f(x, \lambda y) = -f(\lambda y, x) = \lambda(-f(y,x)) = \lambda f(x,y)$. Et puisque $f(x,y) \neq 0$ il suit $\lambda = \lambda^\sigma$. $\square$

Définition 5.1.11 *Une forme f de E dans k est dite <u>hermitienne</u> lorsque*

$$\forall x, y \in E,\ f(x,y) = f(y,x)^\sigma.$$

Exercice 5.1 *Soit f une forme σ-hermitienne non nulle. Montrer que σ est une involution.*

Théorème 5.1.12 *Soit E un k-espace vectoriel de dimension finie $n \geq 2$, et f une forme σ-sesquilinéaire non dégénérée, reflexive. Alors*

1. *σ est une involution.*
2. *Si σ est l'identité, f est bilinéaire symétrique ou antisymétrique.*
3. *Si $\sigma \neq \mathrm{Id}$, il existe un élément $\alpha \in k^\times$ tel que αf soit hermitienne.*

Démonstration. Soit $y \neq 0$ un vecteur de E. L'ensemble $H_y = \mathrm{Ker}(f(.,y))$ est un hyperplan de E. Par reflexivité de f on a

$$\forall x \in E, (f(x,y) = 0) \iff (f(y,x) = 0) \iff (f(y,x)^{\sigma^{-1}} = 0).$$

L'application $x \mapsto f(y,x)^{\sigma^{-1}}$ est une forme linéaire avec même noyau que $f(.,y)$. Par factorisation il existe $\lambda_y \in k$ tel que

$$\forall x \in E,\ f(x,y) = \lambda_y f(y,x)^{\sigma^{-1}}.$$

Soient $\bar{f}$ et $\bar{g}$ les applications de E dans E^* et définie respectivement par

$$\bar{f}(y) = (x \mapsto f(x,y)) \quad \text{et} \quad \bar{g}(y) = \left(x \mapsto f(y,x)^{\sigma^{-1}}\right).$$

Ces applications sont bijectives, la première et σ-semi-linéaire et la deuxième est σ^{-1}-semi-linéaire. Donc l'application $\bar{f}^{-1} \circ \bar{g}$ est σ^{-2}-semi-linéaire de E dans lui-même et vérifie (pour $\mu_y = (\lambda_y^{\sigma^{-1}})^{-1}$) :

$$\forall y \in E, \exists \mu_y \in k, \quad \bar{f}^{-1} \circ \bar{g}(y) = \mu_y y \ .$$

dans cette situation le lemme ci-dessous donne que μ_y ne dépend pas de y et que σ est une involution. Énonçons et démontrons le :

Lemme 5.1.13 *Soit E un k-espace vectoriel de dimension supérieure ou égale à 2, τ un automorphisme de k et u une application τ-semi-linéaire non nulle telle que pour tout x de E les vecteurs x et $u(x)$ sont liés. Alors u est une homothétie et τ est l'identité.*

Démonstration du lemme 5.1.13. On suppose que x et y engendrent un sous-espace vectoriel de dimension 2. Alors si $u(x) = \alpha x$, $u(y) = \beta y$ et $u(x+y) = \gamma(x+y) = \alpha x + \beta y$ puisque x et y sont libres on obtient $\alpha = \beta = \gamma$. Ensuite si une homothétie non-nulle est τ-semi-linéaire alors $\tau = \mathrm{Id}$. □

Fin de la preuve du théorème 5.1.12 On a obtenu $\sigma^2 = \mathrm{Id}$ et l'existence d'un $\mu \in k^\times$ tel que $\bar{g} = \mu \bar{f}$.

1. Si $\sigma = \mathrm{Id}$, alors $\mu^2 = 1$ et soit $\mu = 1$ et f est symétrique soit $\mu = -1$ et f est anti-symétrique.
2. Si $\sigma \neq \mathrm{Id}$, alors f n'est pas alternée et pour un x_0 tel que $f(x_0, x_0) \neq 0$ l'application $(f(x_0, x_0))^{-1} f$ est hermitienne.

Remarque : En dimension 1 sur le corps fini $\mathbb{F}_{27} = \mathbb{F}_{3^3}$ avec l'automorphisme de Fröbenius σ d'ordre 3 défini par $\sigma(x) = x^3$ l'application $(x, y) \mapsto xy^3$ est σ-sequilinéaire réflexive et non dégénérée mais σ n'est pas une involution.

5.2 Sous-espaces orthogonaux, isotropes.

Dans ce paragraphe §5.2 f désigne une forme sesquilinéaire réflexive non dégénérée sauf précision contraire. En outre f est supposée hermitienne si $\sigma \neq \mathrm{Id}$.

Définition 5.2.1 *Soit A une partie de E. On appelle orthogonal de A et on note $A^\perp$ l'ensemble*

$$A^\perp = \{x \in E, \forall y \in A \ f(x, y) = 0\} = \{x \in E, \forall y \in A \ f(y, x) = 0\}.$$

Les éléments de $A^\perp$ sont les éléments de E qui sont orthogonaux aux éléments de A.

L'isomorphisme (semi-linéaire) permet d'identifier E et E^*. Il vérifie en outre les égalités : $\bar{f}(A^\perp) = A^\perp \subset E^*$ (le deuxième $\perp$ réfère à la dualité entre E et E^*). Si on utilise cet isomorphisme on peut traduire les propositions 4.3.2 et 4.3.3 en terme d'orthogonalité pour f. Cela conduit à :

Proposition 5.2.2 *Soit $A \subset E$.*

1. $A \mapsto A^\perp$ est décroissante (pour l'inclusion).

2. *$A^\perp = \langle A\rangle^\perp$ est un sous-espace de E.*
3. *Si A est un sous-espace de E alors* $\dim E = \dim A + \dim A^\perp$.
4. *Si V et W sont deux sous-espaces de E alors*

$$(V+W)^\perp = V^\perp \cap W^\perp,\ (V\cap W)^\perp = V^\perp + W^\perp,\ V^{\perp\perp} = V$$

Démonstration. L'argumentation qui précède l'énoncé est déjà une démonstration complète. Il est aussi possible et instructif de démontrer cette proposition directement en suivant le cheminement des preuves des propositions citées plus haut mais sans y faire référence ni utiliser $\bar{f}$. □

Définition 5.2.3

1. *Un vecteur isotrope x de E est un vecteur non nul vérifiant $f(x,x) = 0$, c'est-à-dire tel que $x \in \{x\}^\perp$.*
2. *Un sous-espace isotrope V de E est un sous-espace tel que $V \cap V^\perp \neq \{0\}$.*
3. *Un sous-espace totalement isotrope est un sous-espace vérifiant $V \subset V^\perp$.*

Remarques : L'ensemble des vecteurs isotropes s'appelle le cône isotrope de f, parce que c'est une partie de E stable par homothéties. En général ce n'est pas un sous-espace vectoriel.

Définition 5.2.4 *Soit f une forme sesquilinéaire non dégénérée réflexive et hermitienne si $\sigma \neq \mathrm{Id}$. On appelle indice de f le maximum des dimensions des sous-espaces totalement isotropes.*

Puisque f est non dégénérée on a $\dim E = \dim V^\perp + \dim V$ et l'inclusion $V \subset V^\perp$ conduit à $\dim V \leq n/2$, pour tout sous-espace V totalement isotrope. En particulier l'indice de f est aussi inférieur à $n/2$.

Si V est isotrope, alors $V \cap V^\perp$ est totalement isotrope.

Si V n'est pas isotrope, alors $V \cap V^\perp = \{0\}$, donc $E = V \oplus V^\perp =: V \perp V^\perp$ est somme directe othogonale de V et de son orthogonal. La notation $V \perp W$ est définie par ce qui précède.

On verra plus tard que les sous-espaces totalement isotropes maximaux (pour l'inclusion), dits "setim", ont tous la même dimension (en l'occurence l'indice de f).

Les définitions de ce paragraphe ont encore un sens même si f est dégénérée, mais on réserve la notion de somme directe orthogonale aux formes non-dégénérées. Que reste-t'il de la proposition 5.2.2 ? En fait si f est dégénérée alors $\operatorname{Ker} f$ est un sous-espace vectoriel de E et f induit par factorisation une application $\tilde{f}\colon E/\operatorname{Ker} f \times E/\operatorname{Ker} f \longrightarrow k$ de même nature (hermitienne ou bilinéaire symétrique ou bilinéaire anti-symétrique) que f. Alternativement en prenant un supplémentaire V de $\operatorname{Ker} f$ on obtient une somme directe $E = V \oplus \operatorname{Ker} f$ avec $f(x+y, x'+y') = f(x,x')$ pour $x, x' \in V$ et $y, y' \in \operatorname{Ker} f$ et $f_{|V}$ non dégénérée. Les définitions qui précèdent et les propriétés obtenues pour $\tilde{f}$ ou $f_{|V}$ donnent aussi des informations pour f, mais qui doivent éventuellement être modifiée légèrement pour tenir compte de $\operatorname{Ker} f$.

Exercice 5.2 *On suppose f dégénérée avec $0 < \dim \operatorname{Ker} f = r < n$. Soit $V \subset E$. Quelle est la dimension de $V^\perp$?*

5.3 Groupes unitaires, orthogonaux, symplectiques.

On fixe σ un automorphisme du corps k, E un k-espace vectoriel de dimension n et on note $GL(E)$ le groupe linéaire des k-automorphismes de E.

5.3.1 Définitions générales.

Le groupe $GL(E)$ agit par composition à droite sur l'ensemble des formes σ-sesquilinéaires de E, concrètement cette action de groupe est définie par la formule

$$(f.\varphi)(u,v) = f(\varphi(u),\varphi(v))$$

pour f sesquilinéaire et $\varphi \in GL(E)$. Être dans la même orbite sous l'action de $GL(E)$ est une relation d'équivalence sur les formes sesquilinéaires sur E.

Définition 5.3.1 *Dans le cas bilinéaire cette relation d'équivalence et son interprétation matricielle s'appelle la congruence. En pratique f et g sont équivalentes si il existe $u \in GL(E)$ telle que $f = g.u$ ou encore*

$$f(x,y) = g(u(x),u(y)).$$

L'interprétation matricielle se déduit de l'équivalence, pour f et g bilinéaire entre "$f = g.u$" et "pour toute base e de E il existe $P \in GL_n(k)$ avec $\mathrm{Mat}_e(f) = {}^tP\,\mathrm{Mat}_e(g)P$". On reconnaît la congruence des matrices. J'ignore quelle est la terminologie pour la relation d'équivalence $M \sim {}^tPMP^\sigma$ qui correspond aux formes sesquilinéaires pour σ non trivial.

Définition 5.3.2 *Soit f une forme non dégénérée. Le stabilisateur de f est appelé :*

1. *le groupe unitaire de f et noté $U(f)$ si f est hermitienne.*
2. *le groupe orthogonal de f et noté $O(f)$ si f est symétrique.*
3. *le groupe symplectique de f et noté $Sp(f)$ si f est alternée.*

Proposition et définition 5.3.3 *Si f est symétrique ou hermitienne, on appelle isométrie pour f un élément $u \in GL(E)$ tel que $\forall x \in E\ f(x,x) = f(u(x),u(x))$. Lorsque $\mathrm{car}(k) \neq 2$ on a équivalence entre "u est une isométrie" et "u appartient au stabilisateur de f".*

Démonstration. Soit $q\colon E \longrightarrow k$ la forme quadratique (ou quadratique hermitienne si σ est non trivial) associée à f, autrement dit définie par $q(x) = f(x,x)$. Si f est symétrique l'équivalence annoncée en caractéristique impaire suit de

$$f(x,y) = \frac{1}{4}(q(x+y) - q(x-y)).$$

On suppose donc σ d'ordre 2. Alors l'application linéaire $\mathrm{Id} - \sigma$ est non nulle et si $x \in k$ vérifie $\sigma(x) \neq x$, on obtient un $a \in k^\times$ qui vérifie $\sigma(a) = -a$ en posant $a = x - \sigma(x)$. On peut ensuite vérifier l'identité hermitienne (qui permet de conclure comme dans le cas symétrique) :

$$f(x,y) = \frac{1}{4}(q(x+y) - q(x-y) - \frac{1}{a}(q(x+ay) - q(x-ay))).$$

□

Remarque : Lorsque $k = \mathbb{C}$ et σ est la conjugaison complexe on prend habituellement $a = i$ et l'identité hermitienne devient :

$$f(x,y) = \frac{1}{4}(q(x+y) - q(x-y) + i(q(x+iy) - q(x-iy))).$$

À partir d'ici et jusqu'à la fin du chapitre 5 on suppose le corps des scalaires k de caractéristique impaire (ou nulle) : $\underline{\text{car}(k) \neq 2.}$

Version matricielle : Si f est une forme de matrice $M = \text{Mat}_e(f)$ et si u est une isométrie pour f de matrice $U = \text{Mat}_{e,e}(u)$ on a ${}^tUMU^\sigma = M$. Il suit donc $\det(u)^{1+\sigma} = 1$, et en particulier lorsque $\sigma = \text{Id}$ on obtient $\det(u) = \pm 1$.

Définition 5.3.4 *Soit f une forme non dégénérée. Le stabilisateur de f est appelé :*

1. *Si f est hermitienne on appelle $\underline{\text{groupe spécial unitaire}}$ de f et on note $SU(f)$ ou $U^+(f)$ le sous-groupe de $U(f)$ formé des endomorphismes de déterminant 1.*
2. *Si f est symétrique on appelle $\underline{\text{groupe spécial orthogonal}}$ de f et noté $SO(f)$ ou $O^+(f)$ le sous-groupe de $O(f)$ formé des endomorphismes de déterminant 1. Les éléments de $SO(f)$ s'appellent des isométries positives, ou des rotations.*

Remarque : $SU(f)$ est distingué dans $U(f)$ et $SO(f)$ est distingué dans $O(f)$. En ce qui concerne les endomorphismes symplectiques, on sait démontrer que leur déterminant vaut toujours 1 (il n'y a pas de sous-groupe spécial symplectique). Pour le dévissage complet (centres, générateurs, sous-groupes dérivés, etc ...) des sous-groupes $GL(E), SL(E), O(f)$ et $SO(f)$, voir le livre de Perrin "cours d'algèbre". Ici on se contente de définir les symétries et d'en extraire un système générateur de $O(f)$ et $SO(f)$.

5.3.2 symétries orthogonales.

Définition 5.3.5 *On appelle $\underline{\text{symétrie}}$ ou $\underline{\text{involution}}$ sur E un endomorphisme $u \in GL(E)$ d'ordre divisant 2. On note E^+ et E^- les sous-espaces propres de u associés respectivement aux valeurs propres $+1$ et -1. On dit que u est une $\underline{\text{symétrie orthogonale}}$ pour f lorsqu'en outre $u \in O(f)$. On dit que u est une $\underline{\text{réflexion}}$ d'hyperplan E^+ lorsque E^+ est un hyperplan. On dit que u est un $\underline{\text{renversement}}$ lorsque* $\dim(E^-) = 2$.

Une symétrie u vérifie l'identité polynômiale $u^2 = 1$, en particulier u est diagonalisable et on a $E = E^+ \oplus E^-$. L'identité $\text{Id} \in GL(E)$ correspond au cas $\dim(E^-) = 0$, les réflexions au cas $\dim(E^-) = 1$, les retournements au cas $\dim(E^-) = 2$. La récurrence s'arrête là ...

Proposition 5.3.6 *Soit f une forme bilinéaire symétrique. Une symétrie u de E est orthogonale si et seulement si E^+ et E^- sont orthogonaux.*

Démonstration. On suppose u orthogonale, soit $x \in E^+$ et $y \in E^-$. Alors $f(x,y) = f(u(x),u(y)) = f(x,-y) = -f(x,y)$. Puisque $\mathrm{car}(k) \neq 2$ on obtient $x \perp y$. Réciproquement on suppose E^+ et E^- orthogonaux. Soit $x, y \in E$. On écrit $x = x^+ + x^-$ et $y = y^+ + y^-$ suivant la décomposition $E = E^+ \oplus E^-$. On obtient $f(x,y) = f(x^+ + x^-, y^+ + y^-) = f(x^+, y^+) + f(x^-, y^-)$ par orthogonalité de E^+ et E^-. Tandis que $f(u(x),u(y)) = f(x^+ - x^-, y^+ - y^-) = f(x^+, y^+) + f(-x^-, -y^-) = f(x^+, y^+) + f(x^-, y^-)$ aussi. □
Le même argument qu'en début de cette preuve montre que si x et y sont des vecteurs propres pour un endomorphismes orthogonal u associés à des valeurs propres λ_x et λ_y telles que $\lambda_x \lambda_y \neq 1$ alors $x \perp y$. Pour cette démonstration il n'est pas utile non plus de supposer f non dégénérée. Par contre si f est non dégénérée alors les inclusions $E^+ \subset (E^-)^\perp$ et $E^- \subset (E^+)^\perp$ deviennent des égalités par calcul de dimensions. On obtient alors $E = E^+ \perp E^-$, et en particulier ni E^+ ni E^- ne sont isotropes. Réciproquement on a :

Proposition 5.3.7 *Soit f une forme bilinéaire symétrique non dégénérée, et $F \subset E$ un sous-espace non isotrope. Alors il existe une unique symétrie orthogonale de sous-espace positif $E^+ = F$.*

Démonstration. Puisque F est non isotrope et f non dégénérée on peut décomposer E en somme directe $E = F \oplus F^\perp$. Alors l'endomorphisme $u = \mathrm{Id}_F \oplus (-\mathrm{Id}_{F^\perp})$ convient et c'est le seul. □
Par exemple la réflexion orthogonale s_H d'hyperplan H avec $H^\perp = kv$ est aussi définie par la formule

$$s_H(x) = x - 2\frac{f(v,x)}{f(v,v)}v.$$

5.3.3 Générateurs de $O(f)$ et $SO(f)$.

Dans ce sous-paragraphe on suppose que f est une forme bilinéaire symétrique non dégénérée et on notera q la forme quadratique associée à f.

Théorème 5.3.8 *Le groupe orthogonal $O(f)$ est engendré par les réflexions orthogonales.*

Démonstration. On procède par récurrence sur $n = \dim(E)$. La propriété est vraie en dimension 1, car alors $O(f) = \{\pm \mathrm{Id}\}$. On prend donc $n > 1$ et on suppose le théorème vrai pour tout espace E de dimension au plus $n-1$. On démontre d'abord deux lemmes calculatoires.

Lemme 5.3.9 *E contient des vecteurs non nul et non isotropes pour f.*

Démonstration. Soit x un vecteur isotrope pour f. Comme f est non dégénérée, il existe $y \in E$ tel que $f(x,y) \neq 0$. Alors $y = f(x,z)^{-1}z$ vérifie $f(x,y) = 1$. Si y est non isotrope alors y convient. Si y est isotrope alors $x+y$ vérifie $f(x+y, x+y) = 2 \neq 0$. □

Lemme 5.3.10 *Si $q(x) = q(y) \neq 0$ alors $(q(x+y) = 0) \implies q(x-y) \neq 0$.*

Démonstration. En effet sinon on aurait $q(x+y) = 0 = 2q(x)+2f(x,y)$ et $q(x-y) = 0 = 2q(x) - 2f(x,y)$ et en ajoutant $4q(x) = 0$. □
On reprend la démonstration du théorème 5.3.8. On se donne donc $u \in O(f)$ et on veut montrer que u est produit d'un nombre fini de réflexions orthogonales. On distingue les cas suivants :

1. L'endomorphisme u admet un vecteur fixe x non isotrope.
2. Tout les vecteurs non isotrope de E vérifient $u(x) \neq x$. On fixe alors x non isotrope dans E et on a donc $y = u(x)$ qui vérifie $q(x) = q(y)$, puis par le lemme les deux seuls sous-cas possibles :
   (a) $q(x-y) \neq 0$.
   (b) $q(x-y) = 0$ et alors $q(x+y) \neq 0$.

Démonstration dans le cas 1. Soit H l'orthogonal de $\langle x \rangle$. Alors $u(H) = H$ puisque $f(x,y) = 0$ équivaut à $f(u(x), u(y)) = 0$. On peut donc appliquer l'hypothèse de récurrence à $u_{|H}$, qui s'écrit $u_{|H} = \tau_1 \cdots \tau_r$ où les τ_i sont des réflexions de H. Mais si on pose $\sigma_i = \tau_i \oplus \mathrm{Id}_{\langle x \rangle}$ alors σ_i est une réflexion de E et $u = \sigma_1 \cdots \sigma_r$, d'où le théorème.
Démonstration dans le cas 2(a). Alors $H = \langle x-y \rangle^\perp$ contient $x+y$ parce que $f(x+y, x-y) = q(x) - q(y) + f(y,x) - f(x,y) = 0$. Soit τ_H la réflexions orthogonale d'hyperplan H. Alors $\tau_H(x-y) = y-x$ et $\tau_H(x+y) = x+y$ donne par élimination $\tau_H(y) = x$. Ainsi x est un vecteur non isotrope et fixé par $\tau_H \circ u$: on est ramené au cas 1.
Démonstration dans le cas 2(b). Le vecteur $x+y$ est non nul non isotrope. Soit $H = \langle x+y \rangle^\perp$, et τ_H la réflexion orthogonale d'hyperplan H. Le même calcul qu'en (a) conduit à $\tau_H(y) = -x$. Puis comme x n'est pas isotrope, on dispose de la réflexion τ_L d'hyperplan $L = \langle x \rangle^\perp$ qui vérifie $\tau_L(-x) = x$. Alors x est un vecteur fixe et non isotrope de $\tau_L \circ \tau_H \circ u$: cela ramène au cas 1. □
Remarque : Lorsque la forme f n'a pas de vecteurs isotropes (par exemple lorsque f est un produit scalaire euclidien) seul les cas 1 et 2(a) peuvent se produire et on voit en outre par récurrence que u est produit d'au plus p_u réflexions où p_u est la dimension des supplémentaires de l'espace des points fixes de u, $p_u = n - \dim(\mathrm{Ker}(u - \mathrm{Id}))$. En général on sait aussi que n réflexions suffisent : c'est le théorème de Cartan-Dieudonné dont une démonstration se trouve p. 190 du Perrin.

En dimension 1 il n'y a pas de renversement. En dimension 2 le seul renversement est $-\mathrm{Id}$ qui n'engendre pas $SO(f)$. En dimension supérieure on a :

Théorème 5.3.11 *Soit f une forme bilinéaire symétrique non dégénérée sur E de dimension $n \geq 3$ alors $SO(f)$ est engendré par les renversements.*

Démonstration. Il suffit de montrer que le produit de deux réflexions est un produit de renversements. On va voir que deux renversements suffisent et en particulier en dimension 3 ou plus tout $u \in SO(f)$ est produit de n renversements. Soient σ_1 et σ_2 deux réflexions orthogonale d'hyperplans (non-isotropes) H_1 et H_2. Si $H_1 = H_2$ alors $\sigma_1 \circ \sigma_2 = \mathrm{Id}$. On peut donc supposer $H_1 \neq H_2$. Si on est en dimension 3 alors $\tau_i = -\sigma_i$ est un renversement et on a $\sigma_1 \circ \sigma_2 = \tau_1 \circ \tau_2$. En dimension supérieure à 3 on prend un sous-espace non isotrope $V \subset H_1 \cap H_2$ de codimension 3 dans E. Alors $E = V \perp V^\perp$ et les σ_i se restreignent à l'identité sur V. On se ramène ainsi à $V^\perp$ de

dimension 3 puisqu'il suffit de prolonger par l'identité sur V les renversements de $V^\perp$ pour obtenir des renversements de E. Si f est anisotrope la preuve est terminée. En général il reste à démontrer l'existence de ce sous-espace V non isotrope de codimension 3. En effet même si H_1 et H_2 sont non-isotropes leur intersection peut l'être. Soient x_1 et x_2 des vecteurs engendrant les orthogonaux respectifs de H_1 et H_2. Alors $\langle x_i \rangle = H_i^\perp$ et comme H_i n'est pas isotrope x_i non plus. Donc le radical $\mathrm{rad}(f_{|(H_1 \cap H_2)})$ est égal au radical $\mathrm{rad}(f_{|\langle x_1, x_2 \rangle})$ de dimension au plus 1. Soit W un supplémentaire (de dimension $n-2$ ou $n-3$) dans $H_1 \cap H_2$ à ce radical. Alors $f_{|W}$ est non-dégénérée et W contient des sous-espaces non isotropes de dimension $n-3$: pour cela il suffit de prendre, si $\dim(W) = n-2$ l'hyperplan dans W orthogonal à un vecteur non isotrope. □

5.4 Classification des formes sesquilinéaires.

En général la classification à équivalence près des formes sesquilinéaires (ou à congruence près des formes quadratiques) est un problème difficile sur un corps quelconque. On va présenter quelques cas particuliers plus accessibles. Le point de départ est l'existence de bases orthogonales.

Définition 5.4.1 *Soit f une forme sesquilinéaire sur E. Une base $e_1, \cdots, e_n$ de E est dite base orthogonale pour f lorsque pour tout $i \neq j$ on a $f(e_i, e_j) = 0$.*

Lorsque e est une base orthogonale la matrice $\mathrm{Mat}_e(f)$ est diagonale. La seule matrice diagonale et anti-symétrique est la matrice nulle, il n'y a donc pas de base orthogonale pour une forme alternée non triviale. On n'étudiera pas dans ce cours la classification des formes alternées.

Théorème 5.4.2 *Soit f une forme symétrique ou hermitienne sur E de dimension finie. Alors il existe une base orthogonale e pour f. Et en outre on a $f(e_i, e_i)^\sigma = f(e_i, e_i)$.*

Démonstration. C'est une récurrence immédiate sur $\dim E$. □
Pour classifier à équivalence près les formes sesquilinéaires on a déjà défini trois invariants de leur classe d'équivalence : le rang, l'indice et le discriminant dans $k/(k^\times)^{1+\sigma}$. En général ces invariants ne suffisent pas. Par exemple en dimension 2 et pour $k = \mathbb{R}$ les formes quadratiques $x^2 + y^2$ et $-x^2 - y^2$ ne sont pas équivalentes et pourtant elles ont même rang, même indice et même discriminant.

Théorème 5.4.3 *Soit E un k-espace vectoriel de dimension finie n.*

1. *On suppose k algébriquement clos. Alors toutes les formes quadratiques non dégénérées sur E sont équivalentes à la forme quadratique $x_1^2 + x_2^2 + \cdots + x_n^2$. Leur indice est la partie entière de $n/2$.*
2. *On suppose $k = \mathbb{R}$. Pour toute forme quadratique q, il existe p tel que $0 \leq p \leq n$ et tel que q soit congruente à la forme*

$$q(x_1, \cdots, x_n) = \sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{n} x_i^2.$$

À congruence près il y a exactement ces $n+1$ formes quadratiques non dégénérées sur E. Le couple $(p, n-p)$ s'appelle la signature de q et c'est un système d'invariants complet pour les formes quadratiques à congruence près.

3. *On suppose $k = \mathbb{C}$ et σ est la conjugaison complexe. Pour toute forme quadratique hermitienne q non dégénérées sur E il existe p tel que $0 \leq p \leq n$ et tel que q soit équivalente à la forme*

$$q(z_1, \cdots, z_n) = \sum_{i=1}^{p} z_i \overline{z}_i - \sum_{i=p+1}^{n} z_i \overline{z}_i.$$

À équivalence près il y a exactement ces $n+1$ formes hermitiennes non dégénérées sur E.

Démonstration. Pour démontrer 1, soit e' une base orthogonale pour une forme quadratique f de rang n, et soit $a_i = f(e'_i, e'_i)$. Puisque k est algébriquement clos il existe des $\alpha_i \in k^\times$ tels que pour tout i $\alpha_i^2 = a_i$. Alors dans la base formée des $e_i = \alpha_i^{-1} e'_i$ la matrice de q est la matrice identité. On dit que la base e est orthonormale pour q.

Pour démontrer 2 on part aussi d'une base orthogonale e' pour une forme quadratique q. Quitte à renumérotter les e'_i on peut supposer $q(e'_i) = a_i > 0$ pour $1 \leq i \leq p$ et $q(e'_i) = -a_i < 0$ pour $i > p$. Puisque les a_i sont positifs il existe des α_i tels que pour tout i on ait $\alpha_i^2 = a_i$. Alors dans la base $e_i = \alpha_i^{-1} e'_i$ la forme quadratique q a la matrice requise. On doit ensuite s'assurer que deux telles formes quadratiques avec des invariants p distincts ne sont pas congruentes. En raisonnant matriciellement on peut fixer une forme quadratique q et vérifier que si deux bases orthogonales e et e' de E sont telles que :

$$q(e_1) = \cdots = q(e_p) = q(e'_1) = \cdots = q(e'_{p'}) = 1$$

$$q(e_{p+1}) = \cdots = q(e_n) = q(e'_{p'+1}) = \cdots = q(e'_n) = -1,$$

alors $p = p'$. Montrons le. Dans ce cas si F est le sous-espace de E engendré par $e_1, \cdots, e_p$ et G' le sous-espace de E engendré par $e'_{p+1}, \cdots, e'_n$ alors $q(x) > 0$ pour $x \in F \setminus \{0\}$ et $q(x) < 0$ pour $x \in G' \setminus \{0\}$. Donc F et G' sont en somme directe et on en tire $p + n - p' \leq n$ soit $p \leq p'$. Par symétrie on récupère l'autre inégalité et on a bien $p = p'$.

L'assertion 3 se démontre exactement comme l'assertion 2. Pour comprendre cette similarité il suffit de constater que dans les deux cas $\mathrm{Im}(q) \subset \mathbb{R}$ et $q(\lambda x) = \lambda^{1+\sigma} q(x)$. Ainsi on peut normaliser modulo $(k^\times)^{1+\sigma} = \mathbb{R}_{>0}^\times$ (dans les deux cas) et seul le signe de $q(e'_i) \in \mathbb{R}$ est un invariant de q. □

On va classifier maintenant les formes quadratique ou hermitienne non dégénérées sur un corps fini $k = \mathbb{F}_q$ de caractéristique impaire. Comme dans le cas précédent le groupe quadratique $\mathbb{F}_q^\times / (\mathbb{F}_q^\times)^2$ joue un rôle important. Puisque $\mathbb{F}_q^\times$ est cyclique ce groupe quadratique est d'ordre 2, et on le représente dans $\mathbb{F}_q^\times$ par le système $\{1, \alpha\}$ où α est un élément de $\mathbb{F}_q^\times$ qui n'est pas un carré.

Théorème 5.4.4 *Soit E un espace vectoriel de dimension $n \geq 1$ sur $\mathbb{F}_q$ avec $2 \nmid q$. À congruence près il y a exactement deux classes de formes quadratiques non dégénérées sur E. Le discriminant dans $\mathbb{F}_q^\times / (\mathbb{F}_q^\times)^2$ est un système d'invariant complet.*

Matriciellement ces classes de congruences sont représentées soit par la matrice I_n soit par la matrice diagonale dont tous les coefficients diagonaux sont 1 *sauf le dernier égal à α.*

Démonstration. Les deux matrices décrites plus haut ne sont pas congruentes parce qu'elles n'ont pas le même discriminant. Il s'agit donc de montrer que toute forme quadratique non dégénérée est représentée dans une base convenable par une de ces deux matrices. On procède par récurrence sur $n = \dim(E)$. Si $n = 1$ c'est évident. On suppose le théorème démontré en dimension $n-1 \geq 1$. Soit x un vecteur non isotrope de E et y un vecteur non isotrope de $\langle x \rangle^\perp$. Alors $q(\lambda x + \mu y) = \lambda^2 q(x) + \mu^2 q(y)$. On admet provisoirement l'existence d'un couple (λ, μ) tel que pour $z = \lambda x + \mu y$ on ait $q(z) = 1$. Alors l'hypothèse de récurrence appliquée à l'orthogonal de $\langle z \rangle$ permet de conclure. Il reste à démontrer le

Lemme 5.4.5 *Soient $u, v \in \mathbb{F}_q^\times$. Alors l'équation $\lambda^2 u + \mu^2 v = 1$ admet au moins une solution $(\lambda, \mu) \in \mathbb{F}_q \times \mathbb{F}_q$.*

Démonstration. Sur $\mathbb{F}_q^\times$ le noyau de $x \mapsto x^2$ est ± 1. Il y a donc $(q-1)/2$ carré non nuls dans $\mathbb{F}_q$ soit $(q+1)/2$ carrés en tout. Puisque u et v sont non nuls les applications $t \mapsto ut$ et $t \mapsto 1 - vt$ sont injectives et les ensembles $\{u\lambda^2, \lambda \in \mathbb{F}_q\}$ et $\{1 - v\mu^2, \mu \in \mathbb{F}_q\}$ comptent tous deux $(q+1)/2$ éléments. Leur intersection n'est pas vide (sinon leur réunion contiendrait $q + 1 > q$ éléments distincts). Un élément de cette intersection donne la solution requise. □

On classifie maintenant les formes hermitiennes associées à une involution σ non triviale sur k fini. Alors dans ce cas on sait que l'ordre de k est un carré puisque k est de degré 2 sur le sous-corps (fini) fixé par σ. Réciproquement $\mathbb{F}_{q^2}$ admet une unique involution non triviale $x \mapsto x^q$.

Théorème 5.4.6 *Soit σ l'involution non triviale de $k = \mathbb{F}_{q^2}$. À équivalence près il y a une seule classe de forme σ-hermitienne non dégénérée sur l'espace de dimension finie E. Cette forme hermitienne est représentée dans une base convenable par la matrice identité.*

Soit f une forme hermitienne sur E de forme quadratique hermitienne q, soient $e_1, \cdots, e_n$ une base orthogonale et notons $a_i = q(e_i)$. Puisque f est hermitienne on a $\sigma(a_i) = a_i$ et donc $a_i \in \mathbb{F}_q = k^\sigma$. Maintenant $q(\lambda e_i) = \lambda^{1+\sigma} a_i$ et pour pouvoir normaliser il faut disposer d'antécédents des a_i^{-1} pour l'application norme $N_{k/\mathbb{F}_q}(x) = x^{1+\sigma} = x^{1+q}$. Le noyau de cette norme est formé des racines de l'équation $x^{q+1} = 1$ et contient donc au plus $q+1$ éléments. L'image de N a donc au moins $q^2 - 1/q + 1 = q - 1 = o(\mathbb{F}_q^\times)$ éléments, c'est-à-dire que N est surjective. □

5.5 Théorème de Witt.

Définition 5.5.1 *On appelle espace quadratique régulier un espace E muni d'une forme quadratique non dégénérée.*

Dans toute la suite on considère des espace quadratique régulier sur un corps k de caractéristique différente de 2.

5.5.1 Plan hyperbolique.

Définition 5.5.2 *Un plan hyperbolique est un espace régulier de dimension 2 (plan régulier) admettant un vecteur isotrope.*

Proposition 5.5.3 *Soit (E, q) un espace régulier et x un vecteur isotrope de E. Alors il existe un plan P de E contenant x et tel que $(P, q_{|P})$ soit hyperbolique.*

Démonstration. Puisque q est régulière il existe $y \in E$ telle que $f(x, y) \neq 0$. Alors $P = \langle x, y \rangle$ convient. □

Proposition 5.5.4 *Soit (E, q) un plan hyperbolique. Il existe une base $e = (e_1, e_2)$ et une base $\varepsilon = (\varepsilon_1, \varepsilon_2)$ telle que*

$$\mathrm{Mat}_e(q) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ et } \mathrm{Mat}_\varepsilon(q) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

On dit que la base e est une base hyperbolique.

Démonstration. Par définition il existe $e_1 \in E$ avec $q(e_1) = 0$. Puisque q est non dégénérée il existe $y \in E$ avec $f(e_1, y) = 1$. Alors y n'est pas colinéaire à e_1 et pour tout $\lambda \in k$ on a $f(e_1, \lambda e_1 + y) = 1$ et $q(\lambda e_1 + y) = 2\lambda + q(y)$. Il suffit donc de poser $e_2 = (-q(y)/2)e_1 + y$ pour obtenir la base e voulue. Avec la matrice $\mathrm{Mat}_e(q) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ on voit que

$$f(\alpha e_1 + \beta e_2, \alpha' e_1 + \beta' e_2) = \alpha'\beta + \alpha\beta'.$$

Donc $\varepsilon_1 = (e_1 + e_2)/2$ et $\varepsilon_2 = (e_1 - e_2)/2$ convient. □

Corollaire 5.5.5 *Dans un plan hyperbolique, il y a exactement deux droites de vecteurs isotropes. Les sous-espaces totalement isotropes sont donc de dimension 1.*

Démonstration. À partir de la matrice $\mathrm{Mat}_e(q) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ on voit que $q(\alpha e_1 + \beta e_2) = 2\alpha\beta$. □

Proposition 5.5.6 *Un plan quadratique (P, q) est hyperbolique si et seulement si le discriminant de q est $-1(k^\times)^2$ dans $k/(k^\times)^2$.*

Démonstration. Si (P, q) est hyperbolique alors le discriminant de q calculé dans les bases de la proposition 5.5.4 vaut -1. Réciproquement soit (P, q) un plan quadratique de discriminant $-1(k^\times)^2$, soit f_1, f_2 une base orthogonale de P et soit $a = q(f_1)$ et $b = q(f_2)$. Puisque le discriminant modulo les carrés est un invariant il existe $\alpha \in k$ tel que $-ab = \alpha^2$ et donc $-b/a = (\alpha/a)^2$. Avec la formule $q(xf_1 + yf_2) = ax^2 + by^2$ on vérifie que $(\alpha/a)f_1 + f_2$ est isotrope. □

Corollaire 5.5.7 *Si k est algébriquement clos tout plan régulier est hyperbolique.*

5.5.2 Sous-espaces hyperboliques, seti et setim.

Définition 5.5.8 *Un sous-espace hyperbolique est un espace somme directe orthogonale de plans hyperboliques.*

Un sous-espace hyperbolique $E = P_1 \perp \cdots \perp P_r$ est donc de dimension paire $n = 2r$. Si on se donne une base hyperbolique e_i, f_i de chaque P_i alors la matrice de q dans la base $\mathcal{B} = e_1, \cdots, e_r, f_1, \cdots, f_r$ est

$$\mathrm{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} 0 & I_r \\ I_r & 0 \end{pmatrix}.$$

Dans une base orthogonale convenablement ordonnée et normalisée $\mathcal{B}'$ la matrice de q est

$$\mathrm{Mat}_{\mathcal{B}'}(q) = \begin{pmatrix} I_r & 0 \\ 0 & -I_r \end{pmatrix}.$$

Proposition 5.5.9 *Un espace vectoriel de dimension $2r$ est hyperbolique si et seulement si il est régulier et possède un sous-espace totalement isotrope de dimension r.*

Démonstration. Soit $E = P_1 \perp \cdots \perp P_r$ un espace hyperbolique (donc régulier) de dimension $n = 2r$. On choisit dans chaque P_i un vecteur isotrope e_i. Alors le sous-espace engendré par les e_i est de dimension r et totalement isotrope. Réciproquement on procède par récurrence sur r. Le cas $r = 1$ suit directement de la définition des plans hyperboliques. On suppose le théorème vrai pour r et on se donne un espace régulier E de dimension $2r + 2$ contenant F totalement isotrope de dimension $r + 1$. Soit $e_1, \cdots, e_{r+1}$ une base de F. Il existe $u \in E$ tel que $f(e_{r+1}, u) = 1$ et forcément $u \notin F$ et $P := \langle u, e_{r+1} \rangle$ est un plan hyperbolique pour la restriction de q. En particulier P est non isotrope et on a $E = P \perp P^\perp$. En outre puisque F est totalement isotrope et $e_{r+1} \in F$ on voit que $x \in F \cap P^\perp$ équivaut à $x \in \langle u \rangle^\perp \cap F$. La forme linéaire $x \mapsto f(x, u)$ est non-nulle sur F donc son noyau est de dimension r. Donc le sous-espace $F \cap P^\perp$ est totalement isotrope de dimension r. Par récurrence $P^\perp$ est hyperbolique puis $E = P \perp P^\perp$ aussi. □
On appelle parfois lagrangien un tel sous-espace totalement isotrope de dimension maximale d'un espace hyperbolique.

Proposition 5.5.10 *Un sous-espace S totalement isotrope d'un espace régulier E est contenu dans un sous-espace hyperbolique.*

Démonstration. On procède par récurrence sur $n = \dim S$. Si $n = 1$ et $S = \langle s \rangle$ alors pour tout $u \in E$ tel que $f(s, u) \neq 0$ le plan $\langle u, s \rangle$ est hyperbolique. On suppose la proposition vraie pour $\dim S = n$ et on suppose que S est totalement isotrope de dimension $n + 1$. On prend $e_{n+1} \in S$ et $y \in E$ avec $f(e_{n+1}, y) = 1$, et on considère le plan hyperbolique $P = \langle e_{n+1}, y \rangle$. Alors comme dans la preuve précédente on vérifie que $E = P \perp P^\perp$ et que $\dim(S \cap P^\perp) = n$. Par récurrence $(S \cap P^\perp)$ est contenu dans un espace hyperbolique H et alors $S \subset (H \perp P)$ aussi. □

Corollaire 5.5.11 *Si S est un sous-espace totalement isotrope de E régulier alors il existe S' totalement isotrope de même dimension que S et tel que $S \cap S' = 0$.*

Démonstration. Lors de la récurrence de la preuve de la proposition 5.5.10 on a démontré en fait que toute base $e_1, \cdots, e_n$ d'un sous-espace totalement isotrope S se complète en une base $e_1, e'_1, \cdots, e_n, e'_n$ telle que les couples (e_i, e'_i) forment n bases hyperboliques. Le sous-espace S' engendré par les e'_i convient. □
Le corollaire qui suit permet de plonger les sous-espaces isotropes dans des sous-espaces non isotropes. Ceci ramène beaucoup de questions au cas non isotrope en particulier la démonstration du théorème de Witt à suivre.

Corollaire 5.5.12 *Soit (E, q) un espace régulier, soit $F \subset E$ un sous-espace, soit $F_0 = \mathrm{rad}(q_{|F})$ et U un supplémentaire dans F de F_0 (de sorte que $q_{|U}$ soit non dégénérée). Alors il existe un sous-espace hyperbolique H de E contenant F_0 et orthogonal à U.*

Démonstration. U donc $U^\perp$ est non isotrope. Le sous-espace F_0 est contenu dans $U^\perp$ donc dans un sous-espace hyperbolique $H \subset U^\perp$. □

Proposition 5.5.13 *Soit (E, q) un espace quadratique régulier, tous les setim de E ont la même dimension.*

Démonstration. Soient S et T deux setim. On choisit un supplémentaire S_1 (resp. T_1) à $S \cap T$ dans S (resp. dans T) et on a $S = S_1 \oplus S \cap T$ et $T = T_1 \oplus S \cap T$.

Lemme 5.5.14 $T_1 \cap S_1^\perp = \{0\}$

Démonstration. Pour tout $x \in S^\perp$ on a $q(x) \neq 0$ ou $x \in S$ sinon $S + kx$ serait un seti contenant strictement S. Soit $x \in T_1 \cap S_1^\perp$. Alors $x \in T$ et donc x est orthogonal à $S \cap T$ puisque T est totalement isotrope. Ainsi x est orthogonal à S_1 et à $S \cap T$ donc à S. Comme $x \in T$ on a $q(x) = 0$ et donc $x \in S \cap T_1 = S \cap T \cap T_1 = \{0\}$. □
On reprend la preuve de la proposition. Par le lemme on a $\dim T_1 + \dim S_1^\perp \leq \dim E$ et on en tire $\dim T_1 \leq \dim S_1$ puis $\dim T \leq \dim S$. Par symétrie cette inégalité est une égalité. □

Théorème 5.5.15 *Soit E un espace quadratique régulier.*

1. *E est somme directe orthogonale d'un espace hyperbolique H avec un espace anisotrope*[1] *G.*
2. *Si E est somme directe orthogonale $E = H \perp G$ avec H hyperbolique et G anisotrope alors l'indice de q vaut exactement* $\dim H/2$.

Démonstration. Soit S un setim de E et H un sous-espace hyperbolique contenant S. Alors $E = H \perp H^\perp$. Soit $x \in H^\perp$ avec $q(x) = 0$. Alors $S + kx$ est un seti et donc $x = 0$ par maximalité de S. L'espace $G = H^\perp$ est bien anisotrope ce qui montre 1. Pour 2 il s'agit de montrer qu'un setim F de H est aussi un setim de E. Or $F \oplus G = F^\perp$ car l'inclusion $\subset$ est immédiate et les dimensions sont les mêmes. Ainsi tout $x \in F^\perp$ s'écrit $x = y + z$ avec $y \in F$ et $z \in G$. Puisque F est un seti et $F \perp G$ on a $q(x) = q(z)$. Autrement dit x isotrope et $x \in F^\perp$ si et seulement si $x \in F$. Cela démontre la maximalité de F. □

1. on dit que (G, q) est anisotrope lorsque $q(x) \neq 0$ pour tout x non nul de G.

5.5.3 Théorème de Witt.

Dans un espace vectoriel E sans structure quadratique supplémentaire $GL(E)$ agit sur les sous-espaces de E et deux sous-espaces sont dans la même orbite si et seulement si ils ont même dimension. Si on se fixe un espace quadratique (E,q) général le groupe orthogonal $O(q)$ agit aussi sur les sous-espaces et si F et F' sont dans la même orbite alors ils ont même dimension et pour la restriction de q ils sont isométrique c'est à dire que la matrice de $q_{|F}$ est congruente à celle de $q_{|F'}$. Le théorème de Witt donne l'implication réciproque et ramène l'étude des orbites des sous-espaces de E sous l'action de $O(q)$ à une question d'équivalence de formes.

Théorème 5.5.16 (Witt) *Soit (E,q) un espace quadratique, $F, F' \subset E$ des sous-espaces et $\sigma\colon (F,q) \xrightarrow{\sim} (F',q)$ une isométrie.*

1. *Si (E,q) est régulier, alors il existe $u \in O(q)$ telle que $u_{|F} = \sigma$.*
2. *Si F est non isotrope, alors il existe $u \in O(q)$ telle que $u_{|F} = \sigma$.*
3. *Si F et F' sont supplémentaire dans E et orthogonaux alors il existe $u \in O(q)$ telle que $u_{|F} = \sigma$.*

Démonstration. Pour 3 on écrit $E = F \oplus F'$. Alors l'isométrie

$$u = \sigma \oplus \sigma^{-1}\colon F \oplus F' \longrightarrow F' \oplus F = E$$

prolonge σ.

On montre d'abord que 1 se ramène à 2. On suppose E régulier et F quelconque. Il s'agit ensuite de prolonger σ en une isométrie définie sur un espace non isotrope contenant F. Soit $F_0 = \mathrm{rad}(q_{|F})$ et $F'_0 = \mathrm{rad}(q_{|F'})$. Alors $\sigma(F_0) = F'_0$ et pour tout supplémentaire U à F_0 dans F, la restriction de q à U est régulière et on a $F' = F'_0 \oplus \sigma(U)$. Par le corollaire 5.5.12 il existe H hyperbolique dans E de dimension $2\dim F_0 = 2\dim F'_0$ orthogonal à U et contenant F_0. Alors $H \perp U$ contient F et est régulier. Pour nous ramener à 2, il reste à prolonger σ en une isométrie (non surjective) $\sigma'\colon H \perp U \longrightarrow E$. Pour cela il suffit de compléter une base f de F_0 en une suite de bases hyperboliques de H, puis de compléter la base $\sigma(f)$ de F'_0 en une suite de bases hyperboliques d'un H' hyperbolique contenant F'_0 comme setim. En envoyant la première suite de bases hyperboliques sur la seconde on définit une isométrie $\varphi\colon H \longrightarrow H'$ qui prolonge $\sigma_{|F_0}$. Alors $\sigma' = \varphi \oplus \sigma_{|U}$ convient.

On montre 2. On suppose F régulier, et on procède par récurrence sur la dimension $t = \dim F$ (étonnant, non ?). Si $t = 1$ alors $F = kx$ et $q(x) \neq 0$. En reprenant la même démarche que dans la démonstration du théorème 5.3.8 on voit que quitte à composer σ à gauche par des réflexions orthogonales de E on peut supposer $\sigma(x) = x$. Et dans ce dernier cas Id_E est orthogonale et prolonge σ. On suppose 2 vrai pour tout espace de dimension $n-1 \geq 1$ et on prend F régulier de dimension n. Partant d'une base orthogonale $e_1, \cdots, e_n$ de F on peut supposer (quitte à composer σ à gauche par des réflexions orthogonales) que $\sigma(e_n) = e_n$. Alors $F' = \langle e_1, \cdots, e_{n-1}\rangle$ et $G' = \langle \sigma(e_1), \cdots, \sigma(e_{n-1})\rangle$ sont contenus dans l'hyperplan H orthogonal à e_n, de dimensions $n-1$ et isométriques par $\sigma_{|F'}$. Par récurrence il existe une isométrie u' de H qui prolonge $\sigma_{|F'}$. Ainsi on obtient u dans $O(q)$ prolongeant σ à E en posant $u = u' \oplus \mathrm{Id}_{\langle e_n\rangle}$. $\square$

Corollaire 5.5.17 *Si q est régulière alors $O(q)$ opère transitivement sur les seti de même dimension (en particulier sur les setim).*

Démonstration. Deux setis de même dimension sont isomorphes comme espace vectoriels donc isométriques. □

Corollaire 5.5.18 *Soient $F, F' \subset E$ deux sous-espaces isomorphes comme espaces quadratiques.*

1. *Si (E, q) est régulier, alors $F^\perp$ et $(F')^\perp$ sont isomorphes (comme espaces quadratiques).*
2. *Si F est non isotrope, alors $F^\perp$ et $(F')^\perp$ sont isomorphes (comme espaces quadratiques).*

Démonstration. Dans les deux cas le théorème de Witt donne une isométrie u de E prolongeant l'isométrie entre F et F'. Alors $u(F^\perp) = u(F)^\perp = (F')^\perp$ et u définit par restriction une isométrie entre les deux orthogonaux. □

Ce dernier corollaire s'appelle parfois "théorème de simplification de Witt", et il se reformule alors :

Théorème 5.5.19 (Simplification de Witt) *Soient (E, q) et (E', q') deux espace quadratiques réguliers isomorphes. On suppose que $E = A \perp B$ et $E' = A' \perp B'$ avec $(A, q_{|A})$ et $(A', q'_{|A'})$ isomorphes. Alors $(B, q_{|B})$ et $(B', q'_{|B'})$ sont aussi isomorphes.*

□

Corollaire 5.5.20 *Si E se décompose de deux façon différente $E = H \perp G = H' \perp G'$ avec H, H' hyperboliques et G, G' anisotrope, alors il existe $u \in O(q)$ telle que $u(H) = H'$ et $u(G) = G'$. En particulier la forme anisotrope $q_a = q_{|G}$ est bien définie par q à équivalence près.*

Démonstration. Les deux setim de H et H' ont même dimension, donc les deux espaces hyperboliques aussi et ils sont isomorphes (comme espaces quadratiques). Par simplification les espaces anisotropes aussi. □

Une telle écriture $E = H \perp G$ avec H hyperbolique et G anisotrope s'appelle une décomposition de Witt de q.

Corollaire 5.5.21 *Soient q et q' deux formes non dégénérées sur E, d'indices respectifs $\nu(q)$ et $\nu(q')$ et de formes anisotropes associées respectives q_a et q'_a. Alors*

$$q \sim q' \iff (\nu(q) = \nu(q') \text{ et } q_a \sim q'_a)\,.$$

Démonstration. En effet à équivalence près q est déterminée par sa décomposition de Witt, et le sous-espace hyperbolique H est déterminé par l'indice $\nu(q)$. □

5.5.4 Exercices : calculs d'indice.

Exercice 5.3 Soit q une forme quadratique sur $\mathbb{R}^n$ non dégénérée de signature $(p, n-p)$. Quel est l'indice de q ?

Exercice 5.4 Soit k un corps algébriquement clos et soit q une forme quadratique non dégénérée sur k^n. Montrer que l'indice de q est la partie entière de $n/2$ (cela termine la preuve du 1 du théorème 5.4.3).

Exercice 5.5 Soit $k = \mathbb{F}_q$ avec $2 \nmid q$ et $n = 2$.

1. À quelle condition sur q, la classe de -1 est-elle un carré ?
2. Soit $\alpha \in \mathbb{F}_q^\times$ qui n'est pas un carré. Montrer que la forme quadratique de rang 2 et de discriminant α est anisotrope si et seulement si $4 \mid q-1$.
3. Que peut-on dire de l'indice de la forme quadratique de rang 2 et de discriminant 1 ?

Chapitre 6

Réseaux.

6.0 prérequis à propos des $\mathbb{Z}$-modules.

Dans ce chapitre et contrairement aux habitudes j'impose des pré-requis assez élevés. Voici la liste des notions admises ici. Je suppose connue la définition de $\mathbb{Z}$-modules libres, de rang des $\mathbb{Z}$-modules libres, le principe qu'un sous-module de $\mathbb{Z}^n$ est libre de rang inférieur ou égal à n, l'égalité entre le rang d'un sous-module M de $\mathbb{Z}^n$ et la dimension du $\mathbb{Q}$ espace vectoriel de $\mathbb{Q}^n$ engendré par M. On utilisera aussi le théorème de la base adaptée sur $\mathbb{Z}$ dont voici un énoncé :

Théorème 6.0.1 *Soit M un sous-$\mathbb{Z}$-module de $\mathbb{Z}^n$. Alors il existe une $\mathbb{Z}$-base de $\mathbb{Z}^n$, disons $e_1, \cdots, e_n$, et une suite d'entier positifs ou nul ordonnés par divisibilité $d_1 \mid \cdots \mid d_r \mid d_{r+1} = 0 \mid \cdots \mid d_n = 0$ tels que $d_1 e_1, \cdots, d_n e_n$ soit une base de M. La suite des d_i est un invariant de la classe d'isomorphisme de $\mathbb{Z}^n/M$ et on appelle les d_i les diviseurs élémentaires de $M \subset \mathbb{Z}^n$. On a bien sur en général*

$$\mathbb{Z}^n/M \cong \bigoplus_{i=1}^{n} \mathbb{Z}/(d_i)$$

et pour $r = n$:

$$o(\mathbb{Z}^n/M) = \prod_{i=1}^{n} d_i$$

Ces notions sont strictement contenues dans le programme de l'UV de quatrième année "modules sur les anneaux principaux". J'indiquerai brièvement en cours le fil conducteur des démonstrations de cette théorie dans la cas particulier des $\mathbb{Z}$-modules. Voir les sections 1 à 4 du chapitre XIX du livre des Gras pour un exposé détaillé de cette théorie.

6.1 Sous-groupes discrets de $\mathbb{R}^n$.

On fixe un entier n et on étudie l'espace euclidien $\mathbb{R}^n$ muni de la métrique euclidienne $\|x\|$. Dans la suite on notera $B_f(a, r)$ la boule fermée de centre a et de rayon r dans R^n (resp. $B_o(a, r)$ la boule ouverte). Lorsqu'il n'est pas utile de préciser on

notera $B(a,r)$. On dit qu'un espace topologique général est discret lorsque tous ses sous-ensembles sont ouverts (et il suffit de vérifier que tous ses points sont ouverts). Dans $\mathbb{R}^n$ on a la caractérisation suivante des sous-groupes discrets (pour la topologie induite par la métrique de $\mathbb{R}^n$).

Lemme 6.1.1 *Soit $G < \mathbb{R}^n$ un sous-groupe de $\mathbb{R}^n$ muni de la topologie induite par celle de $\mathbb{R}^n$. Les assertions suivantes sont équivalentes :*

1. *G est discret.*
2. *Pour tout compact C de $\mathbb{R}^n$ l'intersection $C \cap G$ est finie.*
3. *Pour tout $\varepsilon > 0$ l'intersection $B(0,\varepsilon) \cap G$ est finie.*
4. *Il existe $\eta > 0$ tel que $B(0,\eta) \cap G = \{0\}$.*

Démonstration. Par définition de la topologie induite (la topologie trace) une famille d'ouverts élémentaire de G est donnée par les $B_o(a,r) \cap G$. Dans le groupe topologique $\mathbb{R}^n$ les translations sont bi-continues et toutes les questions topologiques se "recentrent" en 0. Cela explique l'équivalence entre 1 et 4, mais détaillons-la quand même. Par définition 4 est équivalent à "$\{0\}$ est ouvert dans G". Cela entraîne que tout singleton $\{g\} = \{0\} + g \subset G$ est image par une application bi-continue d'un ouvert. Tous les singletons donc tous les sous-ensembles de G sont alors ouverts. L'implication $2 \Rightarrow 3$ est immédiate. Pour conclure on démontre $3 \Rightarrow 4$ et $4 \Rightarrow 2$.

On suppose 3 et on cherche η tel que $B(0,\eta) \cap G = \{0\}$. On prend $\varepsilon > 0$ et on écrit $B(0,\varepsilon) \cap G = \{0, x_1, \cdots, x_k\}$ où $k = \#(B(0,\varepsilon) \cap G) - 1$. Si $k = 0$ alors $\eta = \varepsilon$ convient. Sinon $\eta = \frac{1}{2} \min \|x_i\|$ convient.

On suppose 4 et soit C un compact de $\mathbb{R}^n$. Supposons, en vue d'une contradiction, que $C \cap G$ soit infini. Alors $C \cap G$ contient une suite infinie dont on extrait par compacité une sous-suite convergente à termes deux à deux distincts $(x_i)_{i\in\mathbb{N}} \subset G \cap C$. Soit $x = \lim x_i$. On peut donc trouver $i, j \in \mathbb{N}$ avec $x_i \neq x_j$, $\|x - x_i\| < \eta/2$ et $\|x - x_j\| < \eta/2$. Mais alors $0 \neq x_i - x_j \in B(0,\eta) \cap G$, ce qui contredit 4. □

Corollaire 6.1.2 *Les sous-groupes de $\mathbb{R}$ sont denses ou discrets.*

Soit G un sous-groupe non discret de $\mathbb{R}$. En niant 4, on voit que G contient une suite (x_n) de réels tous non nuls qui converge vers 0. Comme G est un groupe on peut prendre les x_n tous positifs. Soit $x \in \mathbb{R}$ avec $x > 0$. Alors pour tout n il existe un entier r_n (la partie entière de x/x_n) tel que

$$r_n x_n \leq x < (r_n + 1)x_n.$$

En particulier la suite des $(r_n x_n) \subset G$ converge vers x. Cela montre que $G \cap \mathbb{R}^+$ est dense dans $\mathbb{R}^+$. Par symétrie G est dense dans $\mathbb{R}$. □

Exercice 6.1 *Soient $x, y \in \mathbb{R}$ linéairement indépendants sur $\mathbb{Q}$. Alors le groupe $\langle x, y\rangle$ est dense dans $\mathbb{R}$.*

Lemme 6.1.3 *Soit G un sous-groupe de $\mathbb{R}^n$. On a équivalence entre les trois assertions :*

1. *Il existe $f \in \mathrm{Aut}_{\mathbb{R}}(\mathbb{R}^n)$ avec $G = f(\mathbb{Z}^n)$.*

2. *Il existe* $M \in GL_n(\mathbb{R})$ *avec* $G = M\mathbb{Z}^n$.

3. *Il existe une base de* $\mathbb{R}^n$ *qui engendre* G *comme* $\mathbb{Z}$*-module.*

Si un groupe G *vérifie l'une de ces assertions alors il est discret et libre de rang* n.

Démonstration. L'équivalence entre 1, 2 et 3 est immédiate. Un groupe G qui vérifie 1 est libre de rang n. On doit montrer qu'un tel groupe est discret. Soit $e_1, \cdots, e_n$ la $\mathbb{R}$-base de $\mathbb{R}^n$ qui est aussi une $\mathbb{Z}$-base de G. Alors tout $x \in \mathbb{R}^n$ assez proche de 0 s'écrit $x = \sum \lambda_i e_i$ avec $\lambda_i \in]-1; 1[$ pour tout i. Mais un tel x appartient à G si et seulement si les λ_i sont entiers c'est-à-dire nuls. On obtient ainsi un $\eta > 0$ tel que $G \cap B(0, \eta) = \{0\}$. □

Remarque : L'implication réciproque (à savoir G sous-groupe discret de rang n engendre $\mathbb{R}^n$ sur $\mathbb{R}$) est vraie. Cette implication est l'enjeu du théorème de Jacobi-Bravais à suivre.

Définition 6.1.4 *On appelle* <u>réseau</u> *de* $\mathbb{R}^n$ *un* $\mathbb{Z}$*-module libre de rang* n *engendré par une* $\mathbb{R}$*-base de* $\mathbb{R}^n$.

Lemme 6.1.5 *Soit* $G = M\mathbb{Z}^n = N\mathbb{Z}^n$ *un réseau obtenu à partir de deux matrices* M, N *de* $GL_n(\mathbb{R})$. *Alors* $MN^{-1} \in GL_n(\mathbb{Z})$ *et en conséquence*

$$\det(M) = \pm \det(N).$$

Démonstration. En partant de $M\mathbb{Z}^n = N\mathbb{Z}^n$ on arrive à $N^{-1}M\mathbb{Z}^n = \mathbb{Z}^n$ qui entraîne $N^{-1}M \in M_n(\mathbb{Z})$. Par symétrie la matrice inverse $M^{-1}N$ appartient aussi à $M_n(\mathbb{Z})$, c'est-à-dire que toutes deux sont dans $GL_n(\mathbb{Z})$. □

Définition 6.1.6 *Si* $G = M\mathbb{Z}^n$ *est un réseau alors* $|\det(M)|$ *est un invariant de* G, *on l'appelle le* <u>déterminant du réseau</u> G.

Définition 6.1.7 *Soit* $G = \oplus_{i=1}^n \mathbb{Z}e_i$ *un réseau de* $\mathbb{R}^n$. *On appelle* <u>parallélotope fondamental</u> *de* G *le convexe :*

$$\mathcal{P}(e_1, \cdots, e_n) = \{x \in \mathbb{R}^n, x = \sum_{i=1}^n \lambda_i e_i, 0 \leq \lambda_i < 1\}.$$

Un parallélotope fondamental (parfois appelé domaine fondamental) d'un réseau n'est pas un invariant du réseau. Il s'agit d'un système de représentant dans $\mathbb{R}^n$ de $\mathbb{R}^n/G$. Il y autant de tels parallélotopes que de bases (à permutation près) du réseau. Cependant le volume de ce parallélotope est un invariant :

Proposition et définition 6.1.8 *Un parallélotope fondamental* $\mathcal{P}$ *d'un réseau* G *est mesurable (au sens de la mesure de Lebesgue* μ *de* $\mathbb{R}^n$*) et on a* $\mu(\mathcal{P}) = \det(G)$. *Cet invariant se note aussi* μ_G *et s'appelle le* <u>co-volume</u> *ou la* <u>mesure de la maille</u> *du réseau* G.

Démonstration. On note ε la base canonique de $\mathbb{Z}^n$. Le parallélotope fondamental $\mathcal{P}_0$ de $\mathbb{Z}^n$ associé à ε est évidement mesurable de mesure 1. On part de $\mathcal{P} = \mathcal{P}(e_1, \cdots, e_n)$ et de la matrice $M = \mathrm{Mat}_\varepsilon(e)$ telle que $M(\varepsilon) = (e)$. Alors $M(\mathcal{P}_0) = \mathcal{P}$ et

$$\mu(\mathcal{P}) = \int_{\mathcal{P}} \mathrm{d}\mu = \int_{M(\mathcal{P}_0)} \mathrm{d}\mu = |\det(M)| \int_{\mathcal{P}_0} \mathrm{d}\mu = |\det(M)|,$$

parce que le changement de variable $x \mapsto Mx$ a pour jacobienne M. □

Proposition et définition 6.1.9 *Si $H \subset G$ est un réseau de $\mathbb{R}^n$ contenu dans le réseau G on appelle indice de H dans G et on note $(G : H)$ l'ordre du quotient (fini) G/H. On a :*

$$\mu_H = \mu_G(G : H).$$

Démonstration. Il suffit de prendre pour parallélotopes fondamentaux ceux obtenus à partir d'une base de G adaptée à H pour G et de la base de H déduite de la précédente pour H. □

Le théorème fondateur de la théorie des sous-groupes discrets de $\mathbb{R}^n$ affirme que les notions de $\mathbb{Z}$-rang et de $\mathbb{R}$-dimensions coïncident pour ces sous-groupes. (Penser au sous-groupe de rang 2 engendré par 1 et $\sqrt{2}$ dans $\mathbb{R}$)

Théorème 6.1.10 (Jacobi–Bravais) *Soit G un sous-groupe discret de $\mathbb{R}^n$, soit V le sous-espace de $\mathbb{R}^n$ engendré par G et soit $r = \dim_{\mathbb{R}} V$. Alors G est $\mathbb{Z}$-libre de rang r.*

Remarque : Attention si $r < n$ le volume μ_G n'est pas défini.
Démonstration. On prend $(e_1, \cdots, e_r) \subset G$ un système de vecteurs $\mathbb{R}$-libre de rang maximal, et soit H le sous-groupe de G engendré par ces e_i. On se donne l'ensemble $\mathcal{P}_1 = \{\sum_{i=1}^r \lambda_i e_i, 0 \leq \lambda_i < 1\}$. Par construction $\mathcal{P}_1$ est un système de représentants de V/H. L'adhérence de $\mathcal{P}_1$ est compacte et comme G est discret l'intersection $G \cap \mathcal{P}_1$ est finie. Forcément G est engendré par $(e_1, \cdots, e_r) \cup (G \cap \mathcal{P}_1)$ donc est de type fini donc libre. Il reste à voir que le $\mathbb{Z}$-rang de G est au plus r. Soit $x \in G$ et écrivons $x = \sum_{i=1}^r \alpha_i e_i$ avec $\alpha \in \mathbb{R}$. Pour tout entier m on considère $x^{(m)} = mx - \sum_i E(m\alpha_i)e_i$. Alors pour tout m on a $x^{(m)} \in G \cap \mathcal{P}_1$ et comme cet ensemble est fini il existe $l \neq j$ tels que $x^{(l)} = x^{(j)}$. On en déduit $(j - l)x = \sum_i (E(j\alpha_i) - E(l\alpha_i))e_i$ c'est-à-dire que x est dans le $\mathbb{Q}$-espace vectoriel engendré par les e_i. □

Corollaire 6.1.11 *Un sous-groupe $\mathbb{Z}$-libre de $\mathbb{R}^n$ de $\mathbb{Z}$-rang supérieur ou égal à $n+1$ n'est pas discret.*

Ce corollaire immédiat "écrase" et généralise l'exercice 6.1

6.2 Théorème de Minkowski.

Lemme 6.2.1 (Minkowski) *Soit G un réseau de $\mathbb{R}^n$ et soit Σ un sous-ensemble mesurable de $\mathbb{R}^n$ tel que $\mu(\Sigma) > \mu_G$. Alors il existe $x, y \in \Sigma$ avec $x \neq y$ et $x - y \in G$.*

Démonstration. Soit $\mathcal{P}$ un parallélotope fondamental pour G. Les translatés $\mathcal{P}+g$, $g \in G$ donnent une partition de $\mathbb{R}^n$:

$$\mathbb{R}^n = \coprod_{g\in G} (\mathcal{P}+g).$$

D'où une partition $\Sigma = \coprod_{g\in G} ((\mathcal{P}+g)\cap\Sigma)$ et en passant aux mesures :

$$\mu(\Sigma) = \sum_{g\in G} \mu((\mathcal{P}+g)\cap\Sigma) = \sum_{g\in G} \mu(\mathcal{P}\cap(\Sigma-g)).$$

La dernière égalité provient de l'invariance par translation de μ. Maintenant si les ensembles $(\mathcal{P}\cap(\Sigma-g))$ sont disjoints alors la dernière somme est majorée par μ_G ce qui contredit $\mu(\Sigma) > \mu_G$. Il existe donc $g, g' \in G$ tels que $g \neq g'$ et $\mathcal{P}\cap(\Sigma-g)\cap(\Sigma-g') \neq \emptyset$, c'est-à-dire qu'il existe $x, y \in \Sigma$ avec $x-g = y-g'$. On en déduit $x-y = g-g' \in G\setminus\{0\}$ puisque $g \neq g'$ et $g, g' \in G$. □

Théorème 6.2.2 (Minkowski) *Soit G un réseau de $\mathbb{R}^n$ et $S \subset \mathbb{R}^n$ mesurable vérifiant les trois hypothèses suivantes :*

1. *S est symétrique par rapport à 0 (autrement dit $-S \subset S$).*
2. *S est convexe (autrement dit pour tout $x, y \in S$ le segment $[x, y]$ est contenu dans S).*
3. *$\mu(S) > 2^n\mu_G$ ou $\mu(S) \geq 2^n\mu_G$ avec S compact.*

Alors $S\cap G$ contient un élément non nul.

Démonstration. On suppose d'abord $\mu(S) > 2^n\mu_G$. On applique le lemme 6.2.1 à $\Sigma = \frac{1}{2}S$. On trouve x, y distincts dans $\frac{1}{2}S$ tels que $x-y \in G$. Mais $x-y = \frac{1}{2}(2x-2y) \in S$ puisque S est symétrique et convexe. D'où la conclusion lorsque $\mu(S) > 2^n\mu_G$. On suppose maintenant S compact et $\mu(S) \geq 2^n\mu_G$. En appliquant le premier cas à $(1+1/n)S$ on obtient une suite décroissante de compacts non vide

$$K_n = \left(\left(1+\frac{1}{n}\right)S\right)\cap G.$$

Alors l'intersection des K_n est non vide et si $x \in \cap_n K_n$ alors $x \in S\cap G$ par compacité de S. Alternativement les seuls compacts discrets sont les ensembles finis et une suite décroissante d'ensemble finis est stationnaire. □

6.3 Applications diophantiennes.

6.3.1 Approximations diophantiennes simultanées.

On fixe n. On note ε la base canonique de $\mathbb{Z}^n$ et on se donne $x = (x_1, \cdots, x_n) \in \mathbb{R}^n \setminus \mathbb{Q}^n$.

Lemme 6.3.1 *Soit G le $\mathbb{Z}$-module engendré par $\{\varepsilon_1, \cdots, \varepsilon_n, x\}$. Alors G n'est pas discret.*

Démonstration. Par hypothèse sur x ce $\mathbb{Z}$-module est au moins de rang $n+1$. Le corollaire 6.1.11 conclut. □

Théorème 6.3.2 *Soient $x_1, \cdots, x_n$, n nombres réels. Alors pour tout ε tel que $0 < \varepsilon < 1$, il existe un entier q indépendant de i et n entiers $(p_i)_{1 \leq i \leq n}$ tels que*

$$\left| x_i - \frac{p_i}{q} \right| < \frac{\varepsilon}{q}.$$

Démonstration. Puisque G n'est pas discret la boule fermée

$$\mathcal{B}_\varepsilon = \{(x_1, \cdots, x_n) \in \mathbb{R}^n, \sup(x_i) \leq \varepsilon\}$$

contient un élément de G non nul que l'on écrit $g = \sum a_i \varepsilon_i + ax$ et on a $|a_i + ax_i| \leq \varepsilon$ pour tout i. Puisque $\varepsilon < 1$ on a forcément $a \neq 0$ (le réseau $\mathbb{Z}^n$ lui est discret et vérifie $\mathcal{B}_\varepsilon \cap \mathbb{Z}^n = \{0\}$ pour $\varepsilon < 1$). Comme $a \neq 0$ on obtient pour tout i :

$$\left| x_i + \frac{a_i}{a} \right| \leq \frac{\varepsilon}{a}.$$

Cela montre le théorème (en prenant $\varepsilon = \varepsilon/2$, $p_i = -a_i$ et $a = q$). □

6.3.2 Equations diophantiennes linéaires.

On prend $n \geq 2$. Les équations diophantiennes linéaires sont des équations de la forme

$$\sum_{i=1}^{n} a_i x_i = b \tag{6.1}$$

données par des entiers $a_1, \cdots, a_n$ non tous nuls (on suppose $a_n \neq 0$ pour fixer les idées) et un entier b et pour lesquelles on cherche les solutions $x_1, \cdots x_n$ entières, c'est-à-dire telles que $(x_1, \cdots, x_n) \in \mathbb{Z}^n$. Lorsque les a_i et b sont rationnels on se ramène à une équation diophantienne en chassant les dénominateurs. Lorsque l'un des paramètres est irrationnels souvent l'ensemble des solutions est vide et dans tous les cas on n'a plus affaire à un problème diophantiens.

Lemme 6.3.3 *Si $x^0 = (x_1^0, \cdots, x_n^0)$ est une solution particulière de 6.1 alors toute solution x est de la forme $x = y + x^0$ où y est une solution de l'équation homogène associée 6.2 :*

$$\sum_{i=1}^{n} a_i x_i = 0 \tag{6.2}$$

Démonstration. C'est évident. □

Lemme 6.3.4 *L'équation générale 6.1 admet une (donc plusieurs) solutions si et seulement si le pgcd des a_i divise b.*

Démonstration. Dans un anneau principal un *pgcd* des a_i est un générateur de l'idéal (principal) engendré par ces a_i. Avec cette définition le lemme devient tautologique. En vue d'une méthode de calcul des solutions il faut remarquer que l'algorithme d'euclide donne à la fois ce pgcd d et des coefficients de Bezout tels que $d = \sum a_i u_i$, c'est-à-dire une solution particulière si elle existe puisque $b = b'd = \sum a_i u_i b'$. □

Lemme 6.3.5 *L'ensemble des solutions $(y_1, \cdots, y_n)$ de l'équation homogène 6.2 est un sous-groupe discret de rang $n-1$ contenant comme sous-groupe d'indice fini le sous-groupe Γ engendré par les $\gamma_i = -a_n\varepsilon_i + a_i\varepsilon_n$ pour $1 \leq i \leq n-1$.*

Démonstration. Soit G l'ensemble des solutions de 6.2. Alors G est le noyau de la forme $\mathbb{Z}$-linéaire $\varphi\colon (x_1, \cdots, x_n) \mapsto \sum a_i x_i$. Clairement les γ_i sont linéairement indépendants (la matrice des coordonnées des γ_i est triangulaire inférieure avec a_n sur la diagonale pour $i < n$), et ils appartiennent à G dont le rang est donc supérieur ou égal à $n-1$. Réciproquement comme G est dans l'hyperplan de $\mathbb{Q}^n$ associé à φ son rang est inférieur donc égal à $n-1$. □

Définition 6.3.6 *On appelle système fondamental de solutions de (6.2) une $\mathbb{Z}$-base de l'ensemble des solutions.*

Théorème 6.3.7 *Soit $e_1, \cdots, e_n$ une base de $\mathbb{Z}^n$ adaptée au sous-groupe Γ du lemme 6.3.5, c'est-à-dire telle qu'il existe des $q_1, \cdots, q_{n-1} \in \mathbb{N}$ non nuls ordonnés par divisibilité et que $q_1 e_1, \cdots, q_{n-1} e_{n-1}$ soit une base de Γ. Alors $e_1, \cdots, e_{n-1}$ est un système fondamental de solutions de l'équation 6.2 et $(G : \Gamma) = \prod_{i=1}^{n-1} q_i$.*

Démonstration. Soit $\varphi\colon \mathbb{Z}^n \longrightarrow \mathbb{Z}$ la forme linéaire $\varphi(x_1, \cdots, x_n) = \sum a_i x_i$. Par définition $G = \operatorname{Ker}\varphi$. Pour $i < n$ on a $q_i e_i \in \Gamma \subset G$ et donc $q_i \varphi(e_i) = \varphi(q_i e_i) = 0$. Puisque $\mathbb{Z}$ est sans torsion on a $e_i \in G$ pour $i < n$, et donc $\langle e_1, \cdots, e_{n-1}\rangle \subset G \subset \mathbb{Z}^n$. Tout $y \in \mathbb{Z}^n$ s'écrit $y = \sum \lambda_i e_i$ et il vient $\varphi(y) = \lambda_n \varphi(e_n)$. Puisque φ est non-nulle on doit avoir $\varphi(e_n) \neq 0$ et on en déduit $y \in G \iff \lambda_n = 0$. □

6.3.3 Théorème des deux carrés.

Définition 6.3.8 *On dit qu'un entier n est somme de 2 carrés lorsqu'il existe $a, b \in \mathbb{N}$ tels que $n = a^2 + b^2$.*

Définition 6.3.9 *Soit $n \in \mathbb{N}$ et p un nombre premier. On appelle valuation p-adique de n et on note $v_p(n)$ le plus grand entier positif ou nul tel que $p^{v_p(n)} \mid n$.*

Remarque : La factorisation d'un entier positif n en puissances de nombre premiers s'écrit donc :

$$n = \prod_{p \in \mathcal{P}} p^{v_p(n)}.$$

Proposition 6.3.10 *Un entier positif $n \equiv 3[4]$ n'est pas somme de deux carrés.*

Démonstration. Il suffit de réduire modulo 4 l'équation $a^2 + b^2 = n$. Les carrés modulo 4 sont 0 et 1 et leur somme ne peut pas valoir 3 modulo 4. □

Théorème 6.3.11 *Si p est un nombre premier congru à 1 ou 2 modulo 4, alors p est somme de deux carrés.*

Démonstration. Pour $p = 2$ on a $2 = 1 + 1 = 1^2 + 1^2$. On suppose $p \equiv 1[4]$. Alors $\mathbb{F}_p^\times$ est cyclique d'ordre divisible par 4, donc contient u avec $u^2 = -1$. Soit $R = \{(a, b) \in \mathbb{Z}^2; a \equiv ub[p]\}$. Alors R est un réseau de co-volume p puisque c'est le noyau de l'application surjective $\mathbb{Z}^2 \longrightarrow \mathbb{F}_p$ définie par $(a, b) \mapsto \overline{a} - u\overline{b}$. Par le théorème de Minkowski le disque de rayon r contient un point non nul de R dès que $r^2 > (4p/\pi)$. On prend r tel que $(4p/\pi) < r^2 < 2p$. Soit (a, b) le point de R non nul contenu dans ce disque. Alors on a $0 < a^2 + b^2 < 2p$ tandis que $a^2 + b^2 \equiv 0[p]$. Automatiquement $a^2 + b^2 = p$. □

Lemme 6.3.12 *L'anneau $A = \mathbb{Z}[i]$ est euclidien donc principal. Si n est somme de deux carrés et si $p \equiv 3[4]$ est un nombre premier divisant n alors $v_p(n)$ est pair.*

Démonstration. Montrer que $\mathbb{Z}[i]$ est euclidien pour la norme $N(a+ib) = a^2+b^2$ est un exercice bateau. Soit $p \equiv 3[4]$. Alors l'équation $N(a + ib) = p$ n'a pas de solution dans A et on peut en déduire que p est encore irréductible dans A. On note σ la conjugaison de $\mathbb{Z}[i]$. On suppose que $p \equiv 3[4]$ divise $n = a^2 + b^2 = (a + ib)(a - ib)$. Par le lemme de Gauß on obtient $p \mid a + ib$ ou $p \mid a - ib$. Soit s tel que $(a + ib) \in p^s A$ et $a + ib \notin p^{s+1}A$. Alors par conjugaison et comme $\sigma(p) = p$ on voit que $a - ib \in p^s A$ et $a - ib \notin p^{s+1}A$. D'où $n = (a + ib)(a - ib) \in p^{2s}A \setminus p^{2s+1}A$. En particulier p^{2s+1} ne divise pas n dans $\mathbb{Z}$ non plus. En passant aux normes on voit ensuite que $(p^2)^{2s}$ divise n^2 dans $\mathbb{Z}$. □

Corollaire 6.3.13 *Un entier positif n est somme de deux carrés si et seulement si les $v_p(n)$ sont pairs pour tout $p \equiv 3[4]$.*

Démonstration. En regroupant ce qui précède on voit qu'il reste à démontrer qu'un produit de deux entiers chacun somme de deux carré est aussi somme de deux carré. Cela revient à la multiplicativité de la norme dans $\mathbb{Z}[i]$ puisque

$$\begin{aligned}(a^2 + b^2)(c^2 + d^2) &= N(a + ib)N(c + id) = N((a + ib)(c + id)) \\ &= (ac - bd)^2 + (ad + bc)^2.\end{aligned}$$

□

6.3.4 Théorème des quatres carrés.

Proposition 6.3.14 *Tout nombre premier est somme de quatre carrés.*

Démonstration. On peut supposer p impair (et même $p \equiv 3[4]$ mais cela ne sert pas). Par le principe des tiroirs déjà utilisé pour démontrer le lemme 5.4.5 il existe u, v dans $\mathbb{Z}$ tels que $u^2 + v^2 + 1 \equiv 0[p]$. On considère le réseau $R = \{(a, b, c, d) \in \mathbb{Z}^4, c \equiv ua + vb[p], d \equiv ub - va[p]\}$. De même que précédemment on montre que R est d'indice p^2 dans $\mathbb{Z}^4$. le volume de la sphère de rayon r en dimension 4 est $(\pi^2 r^4)/2$ et on choisit r tel que $16p^2 < (\pi^2 r^4)/2 < (\pi^2/2)4p^2$, c'est-à-dire $(\pi^2 r^4)/2 > 16p^2$ et aussi $r^2 < 2p$. Par le théorème de Minkowski il existe un point non nul (a, b, c, d) de R contenu dans la sphère de rayon r. On en tire $0 < a^2 + b^2 + c^2 + d^2 < 2p$ tandis que $a^2 + b^2 + c^2 + d^2 \equiv 0[p]$ d'où $p = a^2 + b^2 + c^2 + d^2$. □

Théorème 6.3.15 (Lagrange) *Tout nombre entier est somme de quatre carrés.*

Démonstration. Comme pour le théorème des deux carrés il s'agit de montrer que le produit de deux sommes de quatre carrés est une somme de quatre carrés. Ici aussi cela se ramène à la multiplicativité d'une sorte de norme mais il s'agit de la norme réduite associée aux quaternions de Hamilton c'est-à-dire une $\mathbb{Q}$ algèbre non commutative. Voir le début du paragraphe 5.7 du "Théorie Algébrique des Nombres" de Pierre Samuel (le lemme 2 suffit à nos besoins). Alternativement on peut se contenter de bombarder la formule :

$$\begin{aligned}&(a^2+b^2+c^2+d^2)(A^2+B^2+C^2+D^2) = \\ &(aA-bB-cC-dD)^2+(aB+bA+cD-dC)^2 \\ &+(aC-bD+cA+dB)^2+(aD+bC-cB+dA)^2.\end{aligned}$$

□

Chapitre 7

Réduction des endomorphismes.

Étant donné un endomorphisme u d'un k-espace vectoriel E de dimension finie n, réduire u c'est trouver une base dans laquelle la matrice de u soit "le plus simple possible". Matriciellement on cherche dans la classe de similitude de $\mathrm{Mat}_\varepsilon(u)$ une matrice "simple". Par exemple diagonaliser (matrice diagonale), trigonaliser (matrice triangulaire inférieure ou supérieure), Jordaniser (matrice de Jordan) lorsque c'est possible sont des réductions. La seule réduction générale abordée dans ce cours sera la réduction de Frobenius (en bloc diagonaux de matrices compagnons) qui donne un représentant "canonique" de la classe de similitude de $\mathrm{Mat}_\varepsilon(u)$.

7.1 sous-espaces stables par u.

Définition 7.1.1 *On appelle sous-espace stable par u un sous-espace $F \subset E$ tel que $u(F) \subset F$.*

Exemples : Si $P \in k[X]$ alors $\mathrm{Im}(P(u))$ et $\mathrm{Ker}(P(u))$ sont stables par u. En particuliers les sous-espaces propres et les sous-espaces caractéristiques de u sont stables par u. Plus généralement on a le lemme

Lemme 7.1.2 *Soit $u, v \in \mathrm{End}_k(E)$ deux endomorphismes qui commutent. Alors $\mathrm{Ker}\, u$ et $\mathrm{Im}\, u$ sont stables pour v.*

Démonstration. Soit $x \in \mathrm{Ker}\, u$, alors $u(v(x)) = v(u(x)) = v(0) = 0$ et donc $v(x) \in \mathrm{Ker}\, u$. Soit $x = v(y) \in \mathrm{Im}(v)$. Alors $u(x) = u(v(y)) = v(u(y)) \in \mathrm{Im}\, v$. □

Définition 7.1.3 *On appelle polynôme minimal de u et on note $\mu_u(X)$ le générateur unitaire de l'idéal de $k[X]$ des polynômes annulant u.*

$$\mu_u(X)k[X] = \{P \in k[X];\ P(u) = 0\}.$$

L'existence de μ_u est assurée parce que $n^2 = \dim_k(\mathrm{End}(E))$ est finie et donc les endomorphismes $1, u, u^2, \cdots, u^{n^2}$ sont liés. Le théorème de Cayley-Hamilton assure en outre que le degré de μ_u est inférieur ou égal à n.

Lemme 7.1.4 *Soit $E' \subset E$ un sous-espace stable par u, soit $E'' = E/E'$ l'espace quotient et soit u' et u'' les endomorphismes de E' et respectivement E'' induits par u. Alors $\mu_{u'}$ et $\mu_{u''}$ divisent μ_u.*

Démonstration. En effet $\mu_u(u') = 0$ et $\mu_u(u'') = 0$. □

Lemme 7.1.5 *Soit $E_1 \cdots E_s$ des sous-espaces stables tels que $E = E_1 + \cdots + E_s$, et soit M_i le polynôme minimal de l'endomorphisme de E_i induit par u. Alors $\mu_u = \mathrm{ppcm}(M_i)$.*

Démonstration. Soit $M = \mathrm{ppcm}(M_i)$. Par le lemme 7.1.4 chaque M_i divise μ_u donc M aussi. Réciproquement $M(u)$ est nul sur tous les E_i donc sur E et μ_u divise M. □

Définition 7.1.6 *Pour tout $\lambda \in k$ on appelle sous-espace propre pour u associé à λ et on note E_λ le noyau $\mathrm{Ker}(\lambda \mathrm{Id}_E - u)$. Lorsque $E_\lambda \neq \{0\}$ on dit que λ est une valeur propre de u. L'ensemble des valeurs propres de u s'appelle le spectre de u et se note $\mathrm{Spec}(u)$. Un vecteur non nul de E_λ s'appelle un vecteur propre pour λ.*

Définition 7.1.7 *Soit ε une base de E. Le déterminant ci dessous ne dépend pas du choix de ε, on l'appelle polynôme caractéristique de u et on le note χ_u :*

$$\chi_u(X) = \det(XI_n - \mathrm{Mat}_\varepsilon(u)).$$

L'indépendance vis à vis de la base vient de la multiplicativité du det puisque pour $M \in GL_n(k)$ on a

$$\det(XI_n - M\,\mathrm{Mat}_\varepsilon(u)M^{-1}) = \det(M(XI_n - \mathrm{Mat}_\varepsilon(u))M^{-1}) =$$

$$\det(XI_n - \mathrm{Mat}_\varepsilon(u)).$$

Pour que $\lambda \in k$ soit valeur propre de u il faut et il suffit que l'endomorphisme $\lambda \mathrm{Id}_E - u$ ne soit pas injectif c'est-à-dire que son déterminant soit nul ou encore que λ soit racine de χ_u.

Proposition 7.1.8 *Soient E' un sous-espace stable par u, soit $E'' = E/E'$ et soient u' (resp. u'') l'endomorphisme de E' (resp. E'') induit par u. Alors $\chi_u = \chi_{u'}\chi_{u''}$.*

Démonstration. Il suffit d'écrire la matrice de u dans une base de E complétant une base de E' et d'utiliser la formule du déterminant des matrices triangulaires par blocs. □

Définition 7.1.9 *On dit que u est trigonalisable lorsqu'il existe une base ε de E telle que la matrice de $\mathrm{Mat}_\varepsilon(u)$ soit triangulaire.*

$$\mathrm{Mat}_\varepsilon(u) = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Nota Bene : Les λ_i ne sont pas supposés distincts.

Proposition 7.1.10 *L'endomorphisme u est trigonalisable sur k si et seulement si χ_u est scindé sur k.*

Démonstration. C'est une récurrence facile sur la proposition 7.1.8. En outre si il existe une base ε avec

$$\mathrm{Mat}_\varepsilon(u) = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

alors $\chi_u = \prod_i (X - \lambda_i)$. □

Théorème 7.1.11 (Cayley-Hamilton) *Dans* $\mathrm{End}_k(E)$ *on a* $\chi_u(u) = 0$ *et dans* $k[X]$ *on a les divisibilités :*

$$\mu_u \mid \chi_u \mid \mu_u^n.$$

Démonstration. Quitte à étendre les scalaires au corps de décomposition L de χ_u on peut supposer χ_u scindé. Par le lemme 7.1.4 appliqué aux sous-espaces propres associés à la valeur propre λ on montre que $X - \lambda$ divise μ_u. Cela donne alors $\chi \mid \mu_u^n$ dans $L[X]$ et donc dans $k[X]$ (par unicité du reste de la division euclidienne). Il reste à montrer que $\chi_u(u) = 0$ lorsque χ_u est scindé. Mais en utilisant une base ε telle que

$$\mathrm{Mat}_\varepsilon(u) = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

on peut voir par récurrence sur k

$$\forall j \leq k \qquad \prod_{i \leq k} (\lambda_i \,\mathrm{Id}_E - u)(e_j) = 0.$$

□

7.2 Théorème des noyaux et applications.

7.2.1 théorème des noyaux.

Lemme 7.2.1 (des noyaux) *Soit* P *un polynôme annulant* u *et décomposé en* $P = P_1 P_2$ *avec* $(P_1, P_2) = 1$*, et soit* $E_i = \mathrm{Ker}(P_i(u))$*, pour* $i = 1, 2$*. On a*

1. $E_1 = \mathrm{Im}(P_2(u))$*,* $E_2 = \mathrm{Im}(P_1(u))$ *et* $E = E_1 \oplus E_2$*.*
2. *Les projecteurs* $E \twoheadrightarrow E_i$ *appartiennent à* $k[u]$*.*

Démonstration. On pose $F_1 = \mathrm{Im}(P_2(u))$ et $F_2 = \mathrm{Im}(P_1(u))$. Comme $0 = P_1(u) \circ P_2(u) = P_2(u) \circ P_1(u)$ on a $F_i \subset E_i$. On part d'une équation de Bezout $1 = R_1 P_1 + R_2 P2$ entre les P_i. Tout $x \in E$ s'écrit $x = P_1(u)R_1(u)(x) + P_2(u)R_2(u)(x)$ donc $E = F_1 + F_2$. Si $z \in E_1 \cap E_2$ alors $z = R_1(u) \circ P_1(u)(z) + R_2(u) \circ P_2(u)(z) = 0$, et donc $E_1 \cap E_2 = \{0\}$. On en déduit $E = F_1 \oplus F_2$ et $E_i = F_i$. Enfin $P_1(u) \circ R_1(u)$ est le projecteur sur E_2 et $P_2(u) \circ R_2(u)$ est le projecteur sur E_1. □
Par récurrence on déduit du lemme des noyaux le théorème du même nom :

Théorème 7.2.2 (des noyaux) *Soit P un polynôme annulant u décomposé en un produit $P = P_1 P_2 \cdots P_s$ avec $(P_i, P_j) = 1$ pour $i \neq j$. Soit $E_i = \mathrm{Ker}(P_i(u))$ et pour tout i soit $Q_i = \prod_{j \neq i} P_j$ et $F_i = \mathrm{Ker}(Q_i(u))$. On a*

1. *pour tout i, $E_i = \mathrm{Im}(Q_i(u))$, $F_i = \mathrm{Im}(P_i(u))$.*
2. *$E = E_1 \oplus \cdots E_s$ et pour tout i, $F_i = \oplus_{j \neq i} E_j$.*
3. *Les projecteurs sur E_i sont dans $k[u]$.*

□

7.2.2 endomorphisme diagonalisable et critère de diagonalisation.

Proposition 7.2.3 *L'endomorphisme u est diagonalisable sur k si et seulement si μ_u est scindé sur k avec des racines simples.*

Démonstration. Si u est diagonalisable on vérifie que $\prod_{\lambda \in \mathrm{spec}(u)} (X - \lambda)$ est le polynôme minimal de u. Réciproquement on applique le théorème des noyaux à la factorisation de μ_u en produit de polynômes de degré 1 distincts. □

Corollaire 7.2.4 *Si u est diagonalisable et si $E' \subset E$ est stable alors les endomorphismes induits par u sur E' et E/E' sont diagonalisables.*

Démonstration. Exercice. □

Corollaire 7.2.5 *Lorsque k est fini avec q éléments u est diagonalisable si et seulement si $u^q = u$.*

Démonstration. Exercice. □

Corollaire 7.2.6 *On suppose u diagonalisable d'espaces propres $E_1, \cdots, E_p$ associés aux valeurs propres $\lambda_1, \cdots, \lambda_p$.*

1. *Un sous-espace $F \subset E$ est stable par u si et seulement si F est de la forme $F = F_1 \oplus \cdots \oplus F_p$ avec $F_i \subset E_i$.*
2. *Un endomorphisme v de E commute à u si et seulement si tous les E_i sont stables par v.*

Démonstration. 1. Les sous-espaces F de cette forme sont évidemment u-stables. Tout sous-espace $F \subset E$ contient la somme directe $F \cap E_i \oplus \cdots F \cap E_p$. Réciproquement si F est u-stable alors F est stable par tout endomorphisme de $k[u]$ et donc par les projecteurs $\pi_i \colon E \longrightarrow E_i$. Tout x de F s'écrit donc $x = \sum \pi_i(x)$ avec $\pi_i(x) \in F \cap E_i$.

2. On note $\mathcal{C}(u)$ l'algèbre (contenant $k[u]$) des endomorphismes de E qui commutent à u. On suppose que $v \in \mathcal{C}(u)$ et on prend $x \in E_i$. Alors $u(v(x)) = v(u(x)) = v(\lambda_i x) = \lambda_i v(x)$, et donc v laisse E_i stable. Réciproquement si v laisse les E_i stable alors sa restriction aux E_i est un endomorphisme qui commute avec la restriction de u puisque cette dernière est l'homothétie de rapport λ_i. On obtient $v(u(x)) = u(v(x))$ pour tout x dans E_i et pour tout i puis $v \in \mathcal{C}(u)$. □

Remarque : On pose $n_i = \dim V_i$. On déduit de 2 la décomposition

$$\mathcal{C}(u) \simeq \prod_i \mathrm{End}(V_i),$$

et donc

$$\dim \mathcal{C}(u) = \sum_i \dim(\mathrm{End}(V_i)) = \sum_i n_i^2 \geq \sum_i n_i = n.$$

Corollaire 7.2.7 *Soit $\mathcal{F} \subset \mathrm{End}_k(E)$ un ensemble d'endomorphismes diagonalisables et commutant deux à deux. Alors il existe une base de E qui les diagonalise simultanément.*

Démonstration. On peut supposer sans perte de généralité que les $v \in \mathcal{F}$ ne sont pas tous des homothéties (ces dernières sont diagonales dans toute base). On procède par récurrence sur $n = \dim E$. Si $n = 1$ il n'y a rien à démontrer. On prend $u \in \mathcal{F}$ ayant (au moins) deux espaces propres distincts et on décompose suivant le spectre de λ :

$$E = \bigoplus_{\lambda \in \mathrm{Spec}(u)} E_\lambda.$$

Puisque tous les $v \in \mathcal{F}$ commutent à u les E_λ sont v stables pour tout $v \in \mathcal{F}$. Par récurrence on trouve une base de E_λ diagonalisant simultanément tous les $v_{|E_\lambda}$ pour tous les $v \in \mathcal{F}$ et tous les $\lambda \in \mathrm{Spec}(u)$. En recollant ces bases on obtient une base de E qui diagonalise tous les $v \in \mathcal{F}$. □

Définition 7.2.8 *Soit $\lambda \in \mathrm{Spec}(u)$.*

1. *La multiplicité de λ comme racine de $\chi_u(X)$ est la multiplicité algébrique de λ.*
2. *La dimension $n_\lambda = \dim E_\lambda$ de l'espace propre associé à λ est la multiplicité géométrique de λ.*
3. *Soit m_λ la multiplicité de λ comme racine de $\mu_u(X)$. Le sous-espace*

$$E_\lambda^c = \mathrm{Ker}((\lambda \,\mathrm{Id}_E - u)^{m_\lambda}),$$

s'appelle espace caractéristique et se note E_λ^c.

Proposition 7.2.9 *On suppose χ_u scindé sur k (donc μ_u aussi).*

1. *E est somme directe des espaces caractéristiques associés à ses valeurs propres.*

$$E = \bigoplus_{\lambda \in \mathrm{Spec}(u)} E_\lambda^c.$$

2. *Pour tout $\lambda \in spec(u)$ la dimension de l'espace caractéristique associé à λ est égal à la multiplicité algébrique de λ : $\dim E_\lambda^c = n_\lambda$*

Démonstration. L'affirmation 1. vient du théorème des noyaux appliqué au polynôme μ_u. Pour 2. soit $d_\lambda = \dim E_\lambda^c$. Par le lemme 7.1.2 les sous-espaces caractéristiques associés à u sont stables et la restriction u_λ de u à E_λ^c est un endomorphisme. On a

$$\prod_{\lambda\in\mathrm{Spec}(u)} (X-\lambda)^{n_\lambda} = \chi_u(X) = \prod_{\lambda\in\mathrm{Spec}(u)} \chi_{u_\lambda}(X).$$

Il suffit donc de voir que λ est la seule racine du polynôme (forcément scindé) $\chi_{u_\lambda}(X)$. Mais par définition de E_λ^c le polynôme minimal μ_λ de u_λ divise $(X-\lambda)^{m_\lambda}$. □

Forcément l'espace propre associé à λ est contenu dans l'espace caractéristique E_λ^c. On obtient donc le critère de diagonalisation portant sur χ_u :

Proposition 7.2.10 *L'endomorphisme u est diagonalisable si et seulement si χ_u est scindé et pour tout $\lambda \in \mathrm{Spec}(u)$ les multiplicités algébrique et géométrique coïncident :*

$$\forall \lambda \in \mathrm{Spec}(u) \qquad n_\lambda = \dim E_\lambda.$$

Démonstration. Exercice. □

Exercice 7.1 *Montrer que $\mu_{u_\lambda} = (X-\lambda)^{m_\lambda}$.*

Exercice 7.2 *Montrer les deux inégalités $1 \leq m_\lambda \leq n_\lambda$. Pour chacune donner des cas d'égalités et d'autres exemples d'inégalités strictes.*

Exercice 7.3 *Soit $\lambda \in \mathrm{Spec}(u)$. Montrer que pour tout $n \in \mathbb{N}$ on a*

$$\mathrm{Ker}(u-\lambda\,\mathrm{Id}_E)^{m_\lambda+n} = \mathrm{Ker}(u-\lambda\,\mathrm{Id}_E)^{m_\lambda}$$

et que m_λ est le minimum des entiers avec cette propriété (c'est la raison pour laquelle on prend la puissance m_λ dans la définition de l'espace caractéristique).

7.2.3 La version diagonalisable plus nilpotent de Dunford.

Théorème 7.2.11 (Dunford) *On suppose χ_u scindé sur k. Il existe deux uniques endomorphismes de E noté δ et ν qui commutent avec δ diagonalisable, ν nilpotent et $u = \delta + \nu$. De plus δ (et donc ν) appartient à $k[u]$.*

Démonstration. On écrit $\mathrm{Spec}(u) = \{\lambda_1, \cdots, \lambda_p\}$, on note m_i la multiplicité de λ_i comme racine de μ_u et on note $E_1, \cdots, E_p$ les sous-espaces caractéristiques $E_i = \mathrm{Ker}(\lambda_i\,\mathrm{Id}_E - u)^{m_i}$. Par la proposition 7.2.9 on peut décomposer $E = \bigoplus_{i=1}^p E_i$. Les sous-espaces E_i sont u-stable et donc les restrictions u_i de u à E_i sont des endomorphismes et $u = u_1 \oplus \cdots \oplus u_p$. On pose $\delta_i = \lambda_i\,\mathrm{Id}_{E_i}$ et $\nu_i = u_i - d_i$. Alors δ_i et ν_i commutent, δ_i et diagonalisable et $\nu_i^{m_i} = 0$ par définition de E_i. On obtient l'existence de δ et ν en posant $\delta = \delta_i \oplus \cdots \delta_p$ et $\nu = u - \delta = \nu_1 \oplus \cdots \oplus \nu_p$.

Par la proposition 7.2.9 la projection $\pi_i \colon E \longrightarrow E_i$ appartient à $k[u]$ et il existe donc $P_i(X) \in k[X]$ tel que $P_i(u) = \pi_i$. Il suit $\delta = \sum_{i=1}^p \lambda_i P_i(u) \in k[u]$ et $\nu = u - \delta \in k[u]$. Cela permet de montrer l'unicité. En effet soit δ' et ν' deux endomorphismes commutant tels que $u = \delta' + \nu'$ et δ' diagonalisable et ν' nilpotent. Alors δ' et ν' commutent deux à deux, donc commutent avec u et donc aussi avec δ et ν qui sont polynomiaux en u, et on a $\delta - \delta' = \nu' - \nu$. Puisque δ et δ' commutent ils sont simultanément diagonalisable et $\delta - \delta'$ aussi. Puisque ν et ν' commutent et sont nilpotent leur différence $\nu' - \nu$ est nilpotente aussi. Ainsi $\delta - \delta' = \nu' - \nu$ est à la fois diagonalisable et nilpotent, c'est-à-dire nul. □

7.3 La version semi-simple plus nilpotent de Dunford.

Dans ce paragraphe on suppose que le corps de base k est parfait. Concrètement cela veut dire soit que la caractéristique de k est nulle $\mathrm{car}(k) = 0$ soit, si $\mathrm{car}(k) = p$, que le Frobenius $x \mapsto x^p$ est surjectif sur k. Par exemple les corps algébriquement clos sont parfait, les corps finis sont parfait tandis que pour tout nombre premier p, le corps $\mathbb{F}_p(T)$ n'est pas parfait puisque $T \notin (\mathbb{F}_p(T))^p$. Les extensions algébrique des corps parfait sont séparable, c'est-à-dire que les racines des polynômes irréductibles à coefficient dans k sont simples dans toute clôture algébrique de k. Cette hypothèse sert essentiellement à faciliter la description des endomorphismes semi-simple (on pourrait les appeler "potentiellement diagonalisable").

Définition 7.3.1 *Un endomorphisme $u \in \mathrm{End}_k(V)$ est dit semi-simple lorsque tout sous-espace vectoriel u-stable $F \subset E$ admet un supplémentaire stable.*

Proposition 7.3.2 *On suppose k parfait. Un endomorphisme u de E est semi-simple si et seulement si μ_u est sans facteur carré dans $k[X]$, autrement dit si et seulement si μ_u n'a que des racines simples dans un clôture algébrique de k.*

Démonstration. On suppose u semi-simple. On décompose $\mu_u(X)$ en produit de puissances d'irréductibles $\mu_u = \prod_{i=1}^{p} P_i^{\alpha_i}(X)$, avec les $P_i(X)$ irréductibles sur $k[X]$ et deux à deux distincts. Il s'agit de montrer que pour tout i on a $\alpha_i = 1$. Par le lemme des noyaux on peut décomposer $E = E_1 \oplus \cdots \oplus E_p$ avec $E_i = \mathrm{Ker}(P_i^{\alpha_i}(u))$. L'espace $\mathrm{Ker}(P_i(u))$ est stable, il admet donc un supplémentaire F dans E stable par u. Le sous-espace $F_i = F \cap E_i$ est alors un supplémentaire stable dans E_i de $\mathrm{Ker}(P_i(u))$. Avec ces notations on a l'équivalence entre $F_i = \{0\}$ et $\alpha_i = 1$. On suppose, en vue d'une contradiction, qu'il existe $x \neq 0$ dans F_i, en particulier $x \in E_i$ et $P_i(u)(x) \neq 0$. Alors le maximum l des entiers m tels que $P_i^m(u)(x) \neq 0$ vérifie $1 \leq l < \alpha_i$, et $P_i(u)P_i^l(u)(x) = 0$. Il suit $P_i^l(u)(x) \in F \cap \mathrm{Ker}(P_i(u))$ puisque ces deux espaces sont stables, et donc $P_i^l(u)(x) = 0$ ce qui contredit la définition de l.

Réciproquement, on suppose $\mu_u(X)$ sans facteurs carrés, et on le factorise en $\mu_u(X) = \prod_{i=1}^{p} P_i(X)$ avec les $P_i(X)$ deux à deux distincts et irréductibles. Soit F un sous-espace u-stable. Par le lemme des noyaux on écrit $E = E_1 \oplus \cdots \oplus E_p$ avec $E_i = \mathrm{Ker}(P_i(u))$ et des projecteurs $\pi_i \colon E \longrightarrow E_i$ polynomiaux en u. En particulier comme F est u-stable on a $F = (F \cap E_1) \oplus \cdots \oplus (F \cap E_p)$ et il suffit de trouver, pour tout i, un supplémentaire u-stable dans E_i aux sous-espace u-stable $F_i = E_i \cap F$, c'est-à-dire un supplémentaire u_i-stable si on note $u_i = u_{|E_i}$. Par construction, $\mu_{u_i} = P_i(X)$ est irréductible et donc l'algèbre $A_i := k[u_i] \equiv k[X]/(P_i(X))$ est un corps commutatif. Alors E_i est muni de la structure de A_i-espace vectoriel naturelle et les sous-A_i-espaces vectoriels de E_i sont exactement les sous-k-espaces u-stable de E_i. Comme A_i est un corps ses sous-espaces u-stables ont tous des supplémentaires u-stables. □

Remarque : Avec cette caractérisation, on constate que les polynômes minimaux des endomorphismes semi-simples ont des racines simples dans une clôture algébrique $\overline{k}$ de k. Cela signifie que leurs matrices associées a priori non diagonalisable dans $M_n(k)$ deviennent diagonalisable dans $M_n(\overline{k})$.

Exercice 7.4 *On suppose u diagonalisable. Montrer que tout sous-espace F admet un supplémentaire u-stable.*

Théorème 7.3.3 (Dunford) *On suppose k parfait et u quelconque. Il existe deux uniques endomorphismes de E noté σ et ν qui commutent avec σ semi-simple, ν nilpotent et $u = \sigma + \nu$. De plus σ (et donc ν) appartient à $k[u]$, et si χ_u est scindé alors $\sigma = \delta$ est diagonalisable.*

Démonstration. Soit $\chi_u(X)$ le polynôme caractéristique de u. Soit $P(X)$ le produit avec multiplicité 1 de tous les facteurs irréductibles de χ_u. Alors un endomorphisme annulé par $P(X)$ sera semi-simple. On va chercher, dans $k[u]$ une solution $\rho(u)$ (avec $\rho \in k[X]$) de l'équation $P(\rho(u)) = 0$ telle que $u - \rho(u)$ soit nilpotent. Alors $\sigma = \rho(u)$ et $\nu = u - \rho(u)$ conviendront et seront des éléments de $k[u]$. La preuve de l'unicité est alors exactement la même que dans la version $\delta + \nu$ du théorème (le seul endomorphisme semi-simple et nilpotent est l'endomorphisme nul). On va procéder de façon complètement explicite en utilisant une variante polynômiale de la méthode de Newton d'approximation en analyse archimédienne des zéros des fonctions analytiques. Puisque P est sans facteur carré et que k est parfait le pgcd unitaire $pgcd(P(X), P'(X))$ est égal à 1, et on dispose d'une équation de Bezout $1 = A(X)P(X) + B(X)P'(X)$. On définit par récurrence

$$\begin{aligned} \rho_0(X) &= X \\ \rho_{n+1}(X) &= \rho_n(X) - B(\rho_n(X))P(\rho_n(X)) \end{aligned}$$

Lemme 7.3.4 (Newton P-adique) *Soit $t \in \mathbb{N}$ tel que $(P(X))^t$ divise $P(\rho_m(X))$ alors $(P(X))^{2t}$ divise $P(\rho_{m+1}(X))$.*

Démonstration. On utilise un développement limité (à l'ordre 1) de $P(X)$ en X et on obtient

$$P(X + Y) = P(X) + YP'(X) + Y^2 \cdots$$

Pour $X = \rho_n(X)$ et $Y = -B(\rho_n(X))P(\rho_n(X))$ on obtient

$$P(\rho_{n+1}(X)) = P(\rho_n(X))(1 - B(\rho_n(X))P'(\rho_n(X)) + P(\rho_n(X)) \cdots)$$

Mais par construction $P(X)$ divise $1 - B(X)P'(X)$ et donc $(P(\rho_n(X)))^2$ divise $P(\rho_{n+1}(X))$. □

Par récurrence et puisque $P(X) = P(\rho_0(X))$ se divise lui-même on montre que $P^{2^m}(X)$ divise $P(\rho_m(X))$. Ainsi on démontre :

1. $P(X)$ divise $\rho_{m+1}(X) - \rho_m(X)$ pour tout m donc divise $X - \rho_m(X)$ pour tout m.
2. Pour $2^t \geq n = \dim(E)$ on a la suite de divisibilité dans $k[X]$:

$$\mu_u(X) \mid \chi_u(X) \mid P^{2^t}(X) \mid P(\rho_t(X)).$$

Autrement dit dans l'algèbre $k[u] \cong k[X]/\mu_u(X)$ où $\overline{P}(X)$ est nilpotent on a :

1. Pour tout m, $u - \rho_m(u)$ est nilpotent.
2. Pour $2^t \geq n = \dim(E)$, l'endomorphisme $P(\rho_t(u))$ est nul.

Cela donne donc l'existence dans $k[u]$ des endomorphismes $\sigma = \rho_t(u)$ et $\nu = u - \sigma$ du théorème. Lorsque simultanément k est parfait et χ_u est scindé on dispose des deux décomposition $u = \delta + \nu = \sigma + \nu'$, mais δ est diagonalisable donc semi-simple et l'unicité dans la deuxième version conduit à $\delta = \sigma$ et $\nu = \nu'$. $\square$

Remarque : Lorsque χ_u est scindé et même si k n'est pas parfait, la preuve ci-dessus fonctionne parfaitement. Cette seconde approche est à la fois plus générale et aussi fournit une méthode complète explicite pour trouver $\delta \in k[u]$ en fonction de u, sous réserve qu'on connaisse le polynôme P (produit des irréductibles divisant χ_u ou, c'est la même chose, divisant μ_u).

Je détaille les étapes principale de ce calcul. On suppose u donné par sa matrice M dans la base canonique de E.

1. On calcule $\chi_M(X)$ (éventuellement par une triangulation de Gauß pour optimiser le temps de calcul).
2. Si $\chi'_M(X) \neq 0$ (par exemple pour $\mathrm{car}(k) = 0$) alors on calcule le pgcd unitaire $(\chi_M(X), \chi'_M(X))$ par l'algorithme d'Euclide sur les polynômes et on a

$$P(X) = \chi_M(X)/(\chi_M(X), \chi'_M(X)).$$

   Lorsque $\chi'_M(X) \neq 0$ le calcul explicite de $P(X)$ peut se révéler un problème plus difficile que la réduction explicite de Dunford de M.
3. On calcule (algorithme d'Euclide) une équation de Bezout

$$1 = A(X)P(X) + B(X)P'(X).$$

4. On calcule dans $k[M]$ la suite récurrente $M_0 = M$ et

$$M_{n+1} = M_n - B(M_n)P(M_n).$$

   On s'arrête dès que cette suite est stationnaire c'est-à-dire $P(M_n) = 0$. On sait a priori qu'on devra effectuer au plus t itération pour $2^t \geq n$.

7.4 Réduction de Jordan.

Définition 7.4.1 *On appelle bloc de Jordan de taille s associé à la valeur propre λ la matrice $J_s(\lambda) = [a_{i,j}] \in M_s(k)$ telle que $a_{i,j} = \lambda\delta_{i,j} + \delta_{i+1,j}$.*

Concrètement ce sont des matrices triangulaires supérieures avec λ sur la diagonale, 1 sur la "sur-diagonale" et 0 partout ailleurs. Par exemple :

$$J_1(\lambda) = [\lambda] \quad J_2(\lambda) = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \quad J_3(\lambda) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} \cdots$$

Clairement $J_s(\lambda) - \lambda I_s$ est nilpotente d'ordre exactement s, en particulier :

$$\mu_{J_s(\lambda)} = \chi_{J_s(\lambda)} = (X - \lambda)^s.$$

Décomposer (lorsque c'est possible) un endomorphisme u en bloc de Jordan c'est trouver une base ε de E dans laquelle $\mathrm{Mat}_\varepsilon(u)$ soit "diagonale par bloc", chaque bloc diagonaux égaux à un $J_{s_i}(\lambda_i)$. Une telle matrice est alors triangulaire supérieure et toutes les valeurs propres λ_i présentes sur la diagonale doivent appartenir à k. Autrement pour que u admette une décomposition de Jordan sur k il est nécessaire que χ_u soit scindé. On va voir que cette condition est aussi suffisante, mais en attendant il y a un cas particulier où il est très facile de scinder χ_u.

7.4.1 Réduction des endomorphismes nilpotents.

Un endomorphisme nilpotent vérifie $u^t = 0$ pour un t entier. En particulier $\mu_u \mid X^t$ et donc $\chi_u = X^n$ puis $\mu_u \mid X^n$. On va étudier une façon très visuelle de trouver et décrire une décomposition de Jordan d'un tel endomorphisme u. Cette technique peut être rendue complètement explicite et conduire à une méthode de calcul comme en section précédente (on dit "algorithme") mais quelques-uns des détails seront laissés aux lecteurs.

Lemme 7.4.2 *Pour tout entier $t \geq 1$, l'endomorphisme u définit par factorisation un morphisme injectif*

$$\overline{u}\colon \frac{\operatorname{Ker} u^{t+1}}{\operatorname{Ker} u^t} \hookrightarrow \frac{\operatorname{Ker} u^t}{\operatorname{Ker} u^{t-1}}.$$

En particulier pour le rang t_u tel que $\mu_u = X^{t_u}$:

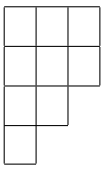
$$0 = \dim(\operatorname{Ker} u^{t_u+1}/\operatorname{Ker} u^{t_u}) \leq \dim(\operatorname{Ker} u^{t_u}/\operatorname{Ker} u^{t_u-1}) \leq \cdots \leq \dim(\operatorname{Ker} u).$$

Démonstration. En partant de l'inclusion évidente $u(\operatorname{Ker} u^{t+1}) \subset \operatorname{Ker} u^t$ on obtient un morphisme $\varphi\colon \operatorname{Ker} u^{t+1} \xrightarrow{u} \operatorname{Ker} u^t \twoheadrightarrow \operatorname{Ker} u^t/\operatorname{Ker} u^{t-1}$. Bien sur $\operatorname{Ker}\varphi = u^{-1}(\operatorname{Ker} u^{t-1}) = \operatorname{Ker} u^t$, et $\overline{u}$ est une application injective par factorisation. □

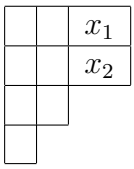
En appliquant successivement le théorème du rang on trouve aussi :

$$\begin{aligned} n = \dim(\operatorname{Ker} u^{t_u}) &= \dim(\operatorname{Ker} u^{t_u}/\operatorname{Ker} u^{t_u-1}) + \dim(\operatorname{Ker} u^{t_u-1}) \\ &= \cdots \\ &= \sum_{t=1}^{t_u} \dim(\operatorname{Ker} u^t/\operatorname{Ker} u^{t-1}) \end{aligned}$$

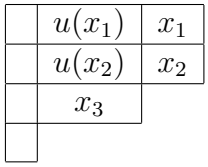
On appelle tableau de Young associé à l'endomorphisme nilpotent u un tableau (à case vide) avec $\dim(\operatorname{Ker} u^t/\operatorname{Ker} u^{t-1})$ cases dans la t-ième colonne. Dans un tel tableau le nombre de lignes diminue quand on passe d'une colonne à sa voisine de droite, et le nombre total de cases est $n = \dim E$. Par exemple voici le tableau de Young (en dimension 9) associé à un endomorphisme u tel que $\dim(\operatorname{Ker} u) = 4$, $\dim(\operatorname{Ker} u^2) = 7$ et $\operatorname{Ker} u^3 = E$ de dimension 9.



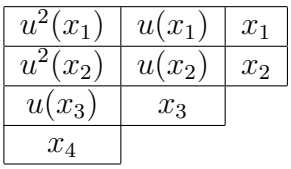
Pour décomposer en blocs de Jordan cet endomorphisme on va remplir les cases du tableau de Young colonnes par colonnes mais en commençant par la droite. On choisit une base $\overline{x}_1, \overline{x}_2$ de $\operatorname{Ker} u^3/\operatorname{Ker} u^2$ (de dimension 2) que l'on relève en x_1, x_2 dans E. On remplit la dernière colonne du tableau.



Par le lemme 7.4.2 le système $\overline{u(x_1)}, \overline{u(x_2)}$ est encore libre dans $\operatorname{Ker} u^2/\operatorname{Ker} u$ et on peut choisir $\overline{x}_3$ dans $\operatorname{Ker} u^2/\operatorname{Ker} u$ pour que $\overline{u(x_1)}, \overline{u(x_2)}, \overline{x}_3$ soit une base du quotient $\operatorname{Ker} u^2/\operatorname{Ker} u$. On relève $\overline{x}_3$ en $x_3 \in \operatorname{Ker} u^2 \subset E$ et on remplit l'avant dernière colonne :



On finit en complétant le système libre $u^2(x_1), u^2(x_2), u(x_3)$ par un vecteur x_4 pour avoir une base de $\operatorname{Ker} u$ et on obtient :



Évidemment ce processus qui se comprend parfaitement sur cet exemple est complètement général, le seul argument utilisé est l'injectivité de

$$\overline{u}\colon \frac{\operatorname{Ker} u^{t+1}}{\operatorname{Ker} u^t} \hookrightarrow \frac{\operatorname{Ker} u^t}{\operatorname{Ker} u^{t-1}},$$

qui permet en appliquant u à une base de $\operatorname{Ker} u^{t+1}/\operatorname{Ker} u^t$ d'obtenir un système libre de $\operatorname{Ker} u^t/\operatorname{Ker} u^{t-1}$, qui se complète en une base etc...Pour terminer il faut lire le tableau de Young ligne par ligne de gauche à droite (normalement quoi) et on obtient la base ε de E qui suit

$$u^2(x_1), u(x_1), x_1, u^2(x_2), u(x_2), x_2, u(x_3), x_3, x_4.$$

Et voici la matrice de u dans cette base :

$$\mathrm{Mat}_\varepsilon(u) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} u^2(x_1) \\ u(x_1) \\ x_1 \\ u^2(x_2) \\ u(x_2) \\ x_2 \\ u(x_3) \\ x_3 \\ x_4 \end{matrix}$$

$$= \begin{bmatrix} J_3(0) & 0 & 0 & 0 \\ 0 & J_3(0) & 0 & 0 \\ 0 & 0 & J_2(0) & 0 \\ 0 & 0 & 0 & J_1(0) \end{bmatrix}$$

De façon générale il y aura autant de "blocs" de Jordan que de lignes dans la première colonne et chaque bloc aura la taille correspondant au nombre de cases dans sa ligne. Cela démontre l'existence d'une décomposition de Jordan pour les matrices nilpotentes et donne une ébauche d'algorithme de calcul. Pour avoir l'algorithme complet il faudrait décrire un processus de complétion en une base de tout système libre des espaces quotients $\operatorname{Ker} u^t / \operatorname{Ker} u^{t-1}$. Cela est parfaitement élémentaire et peut se traiter comme toujours par du pivot de Gauß. On a démontré la partie existence d'une décomposition de Jordan dans la proposition ci-dessous :

Proposition 7.4.3 *Tout endomorphisme nilpotent admet une décomposition en bloc de Jordan. Deux matrices nilpotentes sont semblables si et seulement si leur décomposition de Jordan est la même (à permutation des blocs près).*

Démonstration. Deux matrices nilpotentes semblables représentent le même endomorphisme u dans des bases éventuellement différentes. Mais alors on a vu dans la partie existence que le nombre et la taille des blocs de Jordan déterminent et sont uniquement définis par la suite des dimensions $\dim(\operatorname{Ker}(u^t)/\operatorname{Ker}(u^{t-1}))$ qui elle-même ne dépend que de u. □

7.4.2 Réduction de Jordan.

Théorème 7.4.4 (Décomposition de Jordan) *On suppose χ_u scindé sur k et on écrit $\mathrm{Spec}(u) = \{\lambda_1, \cdots, \lambda_p\}$. Alors il existe une base ε de E dans laquelle $\mathrm{Mat}_\varepsilon(u)$ soit diagonale par blocs le i-ième bloc étant un bloc de Jordan $J_{s_i}(\lambda_{t_i})$ avec $1 \leq i \leq q$ pour un entier $q \geq p$ et $1 \leq t_i \leq p$.*

Démonstration. On commence par décomposer E en sous-espaces caractéristiques $E = E_1 \oplus \cdots \oplus E_p$. Comme les E_i sont stables pour u lorsqu'on recolle des bases de tous les E_i en une base de E, la matrice de u relative à cette base est diagonale par bloc, chaque bloc étant la matrice de la restriction $u_i = u_{|E_i}$. Il suffit donc de trouver une décomposition de Jordan de tous les u_i c'est-à-dire qu'on se ramène au cas $\mathrm{Spec}(u) = \{\lambda\}$. Dans ce cas on a $\chi_u(X) = (X-\lambda)^n$ et en particulier l'endomorphisme

$v = u - \lambda I_E$ est nilpotent. Par la proposition 7.4.3 l'endomorphisme v admet une décomposition de Jordan et sa matrice dans une base convenable est diagonale par blocs avec q blocs de Jordan du type $J_{s_1}(0), \cdots, J_{s_q}(0)$ pour une suite d'entiers décroissante $s_1 \geq s_2 \geq \cdots \geq s_q \geq 1$ et $\sum s_i = n$. En conséquence dans cette même base la matrice de $u = v + \lambda \mathrm{Id}_E$ est diagonale par blocs les blocs diagonaux étant $J_{s_1}(\lambda), \cdots, J_{s_q}(\lambda)$. $\square$

Remarques : Une matrice diagonale est déjà sous forme de Jordan. Réciproquement le polynôme minimal de $J_s(\lambda)$ est $(X-\lambda)^s$. En particulier un endomorphisme u tel que χ_u soit scindé est diagonalisable si et seulement si sa réduction de Jordan n'admet que des blocs de taille 1 c'est-à-dire est déjà diagonale. En fait les blocs de Jordan (à permutation près) fournissent une système complet d'invariant des classes de similitudes des matrices.

Proposition 7.4.5 *Soit A et B des matrices de $M_n(k)$ telle que χ_A et χ_B soient scindés sur k. Alors A et B sont semblables si et seulement si elles ont même décomposition de Jordan (à permutation des blocs près).*

il y a un sens évident : si deux matrices ont même décomposition de Jordan, elles sont semblables. Réciproquement si A et B sont semblable soit u l'endomorphisme de k^n représenté par A et B dans des bases distinctes. Alors la dimension des espaces caractéristiques de u détermine la somme des tailles des blocs de Jordan associé à chaque valeur propre de A, et cette taille est la même pour B. Ainsi on peut supposer que A et B sont semblables et ont une seule valeur propre (alors commune) et égale à λ. Mais dans ce cas $A - \lambda I_n$ et $B - \lambda I_n$ sont semblables et nilpotentes : elles ont même réduction de Jordan. $\square$

Remarque : Dans la construction de la proposition 7.4.3 les blocs de Jordan arrivent naturellement ordonné par taille décroissante (le nombre de cases par ligne décroît dans un tableau de Young). Dans tout ce qui précède les parenthèses (à permutation des blocs près) pourraient être supprimée si on se fixe un ordre sur le spectre commun aux matrices semblables. En conclusion la suite des tailles des blocs de Jordan donne un système complet d'invariant des classes de similitudes des matrices à coefficient dans un corps k qui scinde tous les polynômes caractéristiques c'est-à-dire algébriquement clos et la matrice réduite de Jordan (convenablement ordonnée) donne un représentant canonique de chaque classe de similitude. C'est ce même problème que la réduction de Frobenius résout mais sur un corps quelconque.

7.5 Réduction de Frobenius.

Définition 7.5.1 *Soit $P = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ un polynôme de $k[X]$. On note $C(P)$ et on appelle matrice compagnon de $P(X)$ la matrice $C(P) = [c_{i,j}] \in M_n(k)$ avec $c_{i,j} = \delta_{i+1,j}$ pour $j < n$ et $c_{i,n} = -a_{i-1}$.*

Autrement dit on a

$$C(P) = \begin{bmatrix} 0 & 0 & \cdots & & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & \ddots & & \vdots \\ & & \ddots & 0 & -a_{n-2} \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{bmatrix}.$$

L'espace vectoriel $E = k[X]/(P)$ peut être muni de sa base "canonique"

$$\bar{1}, \bar{X}, \cdots, \overline{X}^{n-1}.$$

Alors l'application $\overline{X}\colon \overline{Q(X)} \mapsto \overline{XQ(X)}$ est un endomorphisme de E et puisque modulo P on a $XX^{n-1} \equiv X^n \equiv -\sum_{i=0}^{n-1} a_i X^i$ la matrice de cette application dans la base canonique est $C(P)$. C'est pour cette raison que la matrice $C(P)$ est associée à P. D'un point de vue plus élémentaire, on a aussi le lemme :

Lemme 7.5.2 *P est le polynôme minimal de $C(P)$ et donc aussi son polynôme caractéristique.*

Démonstration. On note $\varepsilon_1, \cdots, \varepsilon_n$ la base canonique de $E = k^n$ et $u = u_P$ l'endomorphisme de E représenté par $C(P)$. alors pour $i \leq n-1$ on a $u(\varepsilon_i) = \varepsilon_{i+1}$ et donc le système $\varepsilon_1, u(\varepsilon_1), \cdots, u(\varepsilon_{n-1})$ est libre (c'est même la base canonique de E). Le degré de $\mu_{C(P)}$ est donc n et on a déjà $\mu_{C(P)} = \chi_{C(P)}$. Il reste donc à voir que $P(u) = 0$. Sur la matrice $C(P)$ on constate immédiatement :

1. Pour $i \leq n-1$ on a $u^i(\varepsilon_1) = \varepsilon_{i+1}$
2. $u^n(\varepsilon_1) = u(u^{n-1}(\varepsilon_1)) = u(\varepsilon_n) = -\sum_{i=0}^{n-1} a_i \varepsilon_{i+1}$

Il suit donc $P(u)(\varepsilon_1) = 0$, puis pour tout i avec $2 \leq i \leq n$ comme u commute à $P(u)$ on en tire $P(u)(\varepsilon_i) = P(u)(u^{i-1}(\varepsilon_1)) = u^{i-1}P(u)(\varepsilon_1) = 0$. □

Pour des raisons évidentes en termes de $k[u]$-module on dit que E est cyclique lorsqu'il existe $e \in E$ tel que $e, u(e), \cdots, u^{n-1}(e)$ engendre E sur k (cela revient à dire que e engendre E sur $k[u]$). L'exemple canonique (i.e. à isomorphisme près le seul) d'espace cyclique est $k[X]/(P(X))$ muni de l'endomorphisme $\overline{X}$. Le théorème qui suit est une conséquence immédiate de la classification des modules de torsion sur les anneaux euclidiens, mais on va le démontrer autrement.

Théorème 7.5.3 (Frobenius) *Soit M une matrice de $M_n(k)$. Alors il existe une unique suite de polynômes unitaires $P_1 \mid P_2 \mid \cdots P_t$ telle que M soit semblable à une matrice diagonale par blocs chaque bloc étant $C(P_i)$. Deux matrices sont semblables si et seulement les suites de polynômes qui leurs sont associées coïncident. Pour cette raison on appelle $P_1, \cdots, P_k$ les invariants de similitude de M.*

Automatiquement on a alors $\mu_M = P_t$ et $\chi_M = \prod_i P_i$, et ce théorème contient Cayley-Hamilton. En cours j'indiquerai oralement comment ce théorème se déduit de la classification et comment l'algorithme de Smith appliqué à la matrice caractéristique $XI_n - M$ donne une méthode de calcul complète et efficace de la suite des $P_i(X)$. À mon goût ceci est le seul bon point de vue. Cependant, si les étudiants

ne maîtrisent pas assez bien la théorie des modules sur les anneaux principaux ou euclidiens ils doivent disposer d'une approche plus élémentaire. C'est cette approche pénible et fastidieuse que j'ai extrait (et compilé) des livres de Fresnel et Goblot et que je vais suivre jusqu'à la fin de ce polycopié.

La démonstration du théorème 7.5.3 occupe la suite et la fin de cette section et se subdivise en existence et unicité.

7.5.1 Partie existence du théorème 7.5.3.

L'existence dans 7.5.3 se déduit de la proposition (voir théorème V.5 p.112 du Goblot) :

Proposition 7.5.4 *Il existe une unique suite de polynômes unitaires (les derniers éventuellement égaux à 1) $D_1, \cdots, D_n$ avec $D_n \mid D_{n-1} \mid \cdots \mid D_1$ tels que E se décompose en somme directe de sous-espaces stables cycliques $E = E_1 \oplus \cdots \oplus E_n$ avec si l'on note u_i la restriction de u à E_i l'égalité $\mu_{u_i} = D_i$.*

Soit u l'endomorphisme de k^n représenté dans la base canonique par M. On commence par un lemme :

Lemme 7.5.5 *Il existe un sous-espace F stable pour u, cyclique pour la restriction $u' = u_{|F}$ et tel que $\mu_u = \mu_{u'}$*

Démonstration. C'est le lemme 3.1.8 p.119 de l'"algèbre des matrice" de Fresnel, mais je préfère la preuve p.105 du Goblot. On suppose pour commencer que $\mu_u = P^m$ est puissance d'un seul irréductible. Il existe donc $x \in E$ avec $P^{m-1}(x) \neq 0$. Alors le sous-espace u-stable engendré par x (l'espace vectoriel engendré par les $u^i(x)$) convient. Lorsque $\mu_u(X) = \prod_{i=1}^t P_i^{e_i}(X)$ pour des P_i irréductibles deux à deux distincts on utilise le lemme des noyaux pour décomposer $E = E_1 \oplus \cdots \oplus E_t$ avec $E_i = \mathrm{Ker}(P_i(u))$ et aussi $\mu_{u_{|E_i}} = P_i^{m_i}$, et on prend dans chaque E_i un x_i tel que $P_i^{m_i-1}(u)(x_i) \neq 0$. Alors le sous-espace stable engendré par $\sum_i x_i$ convient. □

Pour démontrer la proposition 7.5.4 on procède par récurrence sur $n = \dim E$. On prend $E_1 = F$ où F est le sous-espace stable fournit par le lemme 7.5.5 et on note $e_1 \in E_1$ un vecteur tel que les $u^i(e_1)$ engendrent E_1. Si F admet un supplémentaire stable W alors on aura $\mu_{u_{|W}} \mid \mu_u = D_1$ et on peut conclure par récurrence puisque $\dim W < n$. En outre on sait, parce que le théorème 7.5.3 est démontrée par la classification des modules qu'un tel supplémentaire stable existe. Tout le problème dans cette approche élémentaire et de décrire à la main un tel supplémentaire stable W !

On considère l'endomorphisme u' induit par u sur $E' = E/F$ et $\mu_{u'} \mid \mu_u = D_1$. Par récurrence on peut décomposer $E' = E'_2 \oplus \cdots \oplus E'_n$ avec des E'_i cycliques et sur-lesquels l'endomorphisme u'_i induit par u' admet pour polynôme minimal D'_i avec $D'_n \mid \cdots \mid D'_2 = \mu_{u'} \mid D_1$. Soient $\overline{e}_i$ des générateurs des espaces cycliques E'_i et soient $x_i \in E$ des relevés des $\overline{e}_i$. La relation $D'_i(u)(\overline{e}_i) = 0$ donne $D'_i(u)(x_i) \in k[u](e_1)$ et l'existence de $S_i(X)$ tel que $D'_i(u)(x_i) = S_i(u)(e_1)$. Puisque $D'_i \mid \mu_u$ on a un polynôme N_i tel que $\mu_u = N_i D'_i$ et il suit $0 = \mu_u(u)(x_i) = N_i D'_i(u)(x_i) = N_i(u) \circ S_i(u)(e_1)$ et puisque $\mu_u = D_1$ est le polynôme minimal de la restriction de u au sous-espace stable engendré par e_1 on en tire $\mu_u = N_i D'_i \mid N_i S_i$, soit $D'_i \mid S_i$. On pose U_i pour le

polynôme $U_i = S_i/D'_i$ et $e_i = x_i - U_i(u)(e_1)$. Notons μ_{e_i} pour le polynôme minimal de l'endomorphisme induit par u sur le sous-espace stable engendré par e_i. Dans E/F le vecteur e_i s'envoie sur $\overline{e}_i$ et donc $D'_i \mid \mu_{e_i}$. Réciproquement on vérifie sur la définition de e_i que $D'_i(u)(e_i) = 0$, d'où l'égalité $\mu_{e_i} = D'_i$. On note E_i le sous-espace stable engendré par e_i. Par construction la restriction de la projection canonique $E \longrightarrow E/F$ est surjective depuis $E_2 + \cdots + E_n \longrightarrow E'_2 \oplus \cdots \oplus E'_n$, et comme $\dim(E_i)$ est majorée par le degré de $\mu_{e_i} = D'_i$ cette surjection est forcément un isomorphisme. On conclut en prenant $E_1 = F$. $\square$

7.5.2 Partie unicité du théorème 7.5.3.

On part de la décomposition $E = E_1 \oplus \cdots \oplus E_n$ en sous-espaces E_i qui soient u-stables cyclique donnée par la proposition 7.5.4. On note u_i les endomorphismes de E_i induits par i et D_i la suite de polynômes avec $D_n \mid \cdots \mid D_2 \mid D_1 = \mu_u$. Il s'agit de vérifier que cette suite $D_1, \cdots, D_n$ est intrinsèque à la paire (E, u) et ne dépend pas du choix de la décomposition $E = E_1 \oplus \cdots \oplus E_n$. On a déja $D_1 = \mu_u$. On écrit une factorisation $\mu_u = \prod_{i=1}^{t} P_i^{m_i}$ avec des P_i irréductibles deux à deux distincts. On procède par récurrence sur le nombre $\sum_i m_i$ de facteurs irréductibles de μ_u (avec multiplicité). Si $\mu_u = P_1$ est irréductible alors tous les D_i valent soit 1 soit μ_u et il existe p tel que $D_i = \mu_u$ pour $i \leq p$ et $D_i = 1$ pour $i > p$. Mais alors on a $\dim(E_i) = \deg(\mu_u)$ pour $i \leq p$ et $E_i = 0$ pour $i > p$, et donc $p = n/\deg(\mu_u)$ ne dépend que de E et u. Cela initialise la récurrence.

On montre l'hérédité. Soit P un facteur irréductible de μ_u et soit $E' = \operatorname{Im} P(u)$ et u' l'endomorphisme de E' induit par u. Une seconde de réflexion montre que $\mu_{u'} = \mu_u/P$. Pour tout i l'espace $E'_i = P(u)(E_i)$ est cyclique avec un polynôme minimal noté D'_i. On obtient la décomposition $E' = E'_1 \oplus \cdots E'_n$. Soit p l'indice tel que $P \mid D_p$ et $P \nmid D_{p+1}$. Par Bezout pour $i \geq p+1$ l'endomorphisme $P(u)$ est inversible sur E_i et on a alors $E'_i = E_i$ avec polynôme minimal D_i. Tandis que pour $i \leq p$ alors $P \mid D_i$ et les espaces E_i/E'_i et E'_i sont cycliques de polynômes minimaux P et D_i/P. Ainsi la suite de polynômes D'_i satisfait les conditions de divisibilité pour l'espace E'. Par récurrence cette suite est intrinsèque au couple (E', u') qui lui ne dépend que de (E, u) et P, mais pas des E_i. En ce qui concerne les D_i eux-mêmes on a $D_i = PD'_i$ pour $i \leq p$ et $D_i = D'_i$ pour $i > p$: il reste seulement à vérifier que l'entier p lui-même est intrinsèque. Mais on a

$$\begin{aligned} n = \dim E &= \sum_i \deg(D_i) = \sum_{i \leq p} (\deg(D'_i) + \deg(P)) + \sum_{i>p} \deg(D'_i) \\ &= p \deg(P) + \dim E' \end{aligned} .$$

$\square$