Rappels d'algèbre générale

Boyer Pascal

17 juillet 2017

On réunit ici les notions d'algèbre introduites en L et nécessaire pour entamer serainement un master de mathématiques fondamentales : la plupart des énoncés sont donnés sans démonstration parfois je donne des preuves, rapides, de certains résultats intéressants.

Table des matières

1 Groupes

1.1 Définitions

Définition 1. *On appelle* groupe *un couple $(G, *)$ formé d'un ensemble G et d'une loi de composition*

$$(x, y) \in G^2 \mapsto x * y \in G$$

tels que les trois conditions suivantes soient vérifiées :

— associativité *: pour tous $x, y, z \in G$, on a $x * (y * z) = (x * y) * z$;*

— élément neutre *: il existe $e \in G$ tel que pour tout $x \in G$, on a $e * x = x * e = x$;*[1]
— symétrique *: pour tout $x \in G$ il existe $y \in G$ tel que $x * y = y * x$.*

Remarque: si de plus quels que soient $x, y \in G$, on a $x * y = y * x$ on dit que G est un groupe *commutatif* ou *abélien*.

Notation 1. *Si G est fini, on note $|G|$ son cardinal qu'on appelle aussi* son ordre.

Remarque: habituellement si la loi est commutative on la note avec un $+$ en lieu et place de $*$; sinon on préfère utiliser la notation multiplicative xy plus courte à écrire que $x * y$ et son symétrique est communément appelé son *inverse* que l'on note sous la forme x^{-1}. En ce qui concerne l'élément neutre on le note 0 dans le cas commutatif et 1 sinon.
Exemples :

— l'ensemble $\mathbb{Z}$ des entiers relatifs muni de l'addition est un groupe abélien d'élément neutre 0. En remplaçant $\mathbb{Z}$ par $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$, on obtient le groupe additif des nombres rationnels, réels ou complexes.
— L'ensemble $\mathbb{Q}^\times$ des nombres rationnels non nuls muni de la multiplication est un groupe abélien d'élément neutre 1 ; c'est le groupe multiplicatif des nombres rationnels. On définit de même $\mathbb{R}^\times$ et $\mathbb{C}^\times$.
— Si X est un ensemble, on note $\mathfrak{S}(X)$ l'ensemble des bijections de X muni de la loi de composition ; on définit ainsi un groupe non commutatif d'élément neutre l'identité que l'on appelle le groupe symétrique de X.
— Si dans l'exemple précédent, X est un $\mathbb{R}$-espace vectoriel de dimension n, et que l'on considère les bijections *linéaires* de X, on obtient le groupe linéaire $GL(X)$ isomorphe à $GL_n(\mathbb{R})$ une fois une base de X choisie.
— Si $G_1, \cdots, G_n$ sont des groupes, le produit cartésien $G = G_1 \times \cdots \times G_n$ muni de la loi produit

$$(x_1, \cdots, x_n) * (y_1, \cdots, y_n) = (x_1 y_1, \cdots, x_n y_n)$$

est un groupe appelé *le produit direct* de $G_1, \cdots, G_n$. Son élément neutre est $(1, \cdots, 1)$ et l'inverse de $(x_1, \cdots, x_n)$ est $(x_1^{-1}, \cdots, x_n^{-1})$.

Définition 2. *Pour G un groupe et X un ensemble quelconque, on note G^X l'ensemble des applications de X dans G ; la loi de groupe de G muni G^X d'une structure de groupe. Plus généralement on note $G^{(X)}$ le sous-ensemble de G^X des applications à support fini, i.e. celles telle que l'ensemble des x avec $f(x) \neq e$ est fini.*

1. un élément neutre est nécessairement unique comme le montrent les relations $e_1 * e_2 = e_1 = e_2$.

Remarque: par exemple $\mathbb{Z}^{\mathbb{N}}$ (resp. $\mathbb{Z}^{(\mathbb{N})}$) désigne l'ensemble des suites $(u_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{Z}$ (resp. telles que l'ensemble des n tels que $u_n \neq 0$ est fini).

Définition 3. *On dit qu'un sous-ensemble H d'un groupe $(G, *)$ est* un sous-groupe *si les conditions suivantes sont réalisées :*

- *l'élément neutre e appartient à H ;*
- *pour tous $x, y \in H$, l'élément xy est dans H ;*
- *pour tout $x \in H$, l'inverse x^{-1} est dans H.*

Remarque: de manière équivalente H est un sous-groupe si et seulement s'il est non vide et que pour tous $x, y \in H$ alors $xy^{-1} \in H$. Dans ce cas H muni de la loi $*$, est un groupe. Habituellement pour montrer qu'un ensemble muni d'une loi interne est un groupe, on essaie de montrer qu'il s'agit d'un sous-groupe d'un groupe déjà connu.
Exemples :

- les sous-ensembles G et $\{e\}$ de G sont clairement des sous-groupes que l'on qualifie habituellement de *triviaux* ;
- le sous-ensemble $\mathbb{R}_+^\times$ des réels strictement positifs ainsi que $\{\pm 1\}$ sont des sous-groupes de $\mathbb{R}^\times$;
- l'ensemble des nombres complexes de module 1 est un sous-groupe de $\mathbb{C}^\times$.

Remarque: l'intersection quelconque d'une famille de sous-groupes de G est aussi un sous-groupe de G ce qui permet de définir *le plus petit* sous-groupe contenant une partie X quelconque de G ; on le note $< X >$ et on l'appelle le sous-groupe engendré par X.

Définition 4. *Pour $g \in G$ si le sous-groupe $< g >= \{g^k : k \in \mathbb{Z}\}$ engendré par g est de cardinal fini, ce cardinal est appelé* l'ordre *de g ; sinon on dit que g est d'ordre infini.*

Définition 5. *Pour G un groupe et H un sous-groupe de G, on associe la relation binaire $\mathcal{R}_H$ sur G définie par*

$$x\mathcal{R}_H y \Leftrightarrow x^{-1}y \in H.$$

Cette relation est une relation d'équivalence ; l'ensemble des classes d'équivalence est noté G/H et s'appelle l'ensemble des classes à gauche modulo H.

Remarque: la classe d'équivalence d'un élément $x \in G$ est le sous-ensemble $xH = \{xh : h \in H\}$.
Remarque: on peut aussi définir la relation d'équivalence par $xy^{-1} \in H$ auquel cas les classes sont dites à droite et on note $H\backslash G$ l'ensemble des classes d'équivalence. La classe de x est alors $Hx = \{hx : h \in H\}$. On notera que la classe à gauche xH est égale à la classe à droite Hx si et seulement si $xHx^{-1} = H$; en particulier c'est toujours le cas si G est abélien.

Théorème 6. ***(de Lagrange)***
Si G est fini alors on a $|G| = |H| \times |G/H|$ et en particulier l'ordre de H divise celui de G.

Preuve : Il suffit de dénombrer les éléments de G en utilisant la partition définie par les classes à gauche de G modulo H. Comme toutes ces classes ont le même cardinal égal à l'ordre de H, la relation s'en déduit.
Remarque: si G est de cardinal un nombre premier alors ses seuls sous-groupes sont les sous-groupes triviaux $\{e\}$ et G.

Corollaire 7. *Soit G un groupe fini de cardinal n, alors pour tout $g \in G$ on a $g^n = e$.*

Remarque: autrement dit l'ordre d'un élément est un diviseur du cardinal du groupe.

On voudrait que la loi de G induise sur l'ensemble G/H une structure de groupe, i.e. on voudrait définir $(xH)*(yH) = xyH$; pour cela il faut vérifier que la formule ne dépend pas des choix de x et y, i.e. que pour $x' = xh_1$ et $y' = yh_2$ on a bien $x'y'H = xyH$ autrement dit pour tout $h_1, h_2 \in H$, il existe $h \in H$ tel que $xh_1yh_2 = xyh$ que l'on peut écrire encore sous la forme $yHy^{-1} \subset H$. On introduit alors la notion suivante.

Définition 8. *Un sous-groupe H de G est dit* distingué *si pour tout $g \in G$, on a $gHg^{-1} \subset H$. Un groupe est dit* simple *si ses seuls sous-groupes distingués sont ses sous-groupes triviaux.*

Proposition 9. *La loi de G induit sur G/H une structure de groupe si et seulement si H est un sous-groupe distingué de G.*

Remarque: en particulier si G est abélien alors tout sous-groupe est automatiquement distingué et G/H est, via la loi de G, muni d'une structure de groupe.
Exemple fondamental : reprenons la construction précédente pour le sous-groupe $n\mathbb{Z}$ de $\mathbb{Z}$. Ainsi la relation d'équivalence s'écrit :

$$x \sim_n y \Leftrightarrow n|x - y$$

et on dit que x et y sont *congruents modulo n* ; on écrit $x \equiv y \mod n$. L'ensemble quotient est $\mathbb{Z}/n\mathbb{Z}$ dont les éléments sont $\overline{x} = \{x + kn : \ k \in \mathbb{Z}\}$. On peut ainsi écrire, par exemple, $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \cdots, \overline{n-1}\}$, où un élément x appartient à $\overline{r}$ pour r le reste de la division euclidienne de x par n. On vérifie alors aisément que la définition suivante est cohérente : $\overline{x}+\overline{y} = \overline{x_0 + y_0}$ où x_0 et y_0 sont des éléments quelconques de $\overline{x}$ et $\overline{y}$ respectivement.

1.2 Morphismes

Définition 10. Un morphisme de groupes $f : G \to G'$ *est une application telle que pour tous $x, y \in G$ on a $f(xy) = f(x)f(y)$.*

Remarque: l'élément neutre de G s'envoie nécessairement sur l'élément neutre de G' ; par ailleurs on a $f(x^{-1}) = f(x)^{-1}$.
Exemples :

- la fonction logarithme népérien : $\ln : \mathbb{R}_+^\times \to \mathbb{R}$ définit un morphisme de $(\mathbb{R}_+^\times, \times)$ dans $(\mathbb{R}, +)$. De même la fonction exponentielle définit un morphisme de $(\mathbb{R}, +)$ dans $(\mathbb{R}_+^\times, \times)$.
- Pour $g \in G$, l'application $k \in \mathbb{Z} \mapsto g^k \in G$ est un morphisme dont l'image est $< g >$ le sous-groupe de G engendré par g.
- Pour $n \geq 1$, la surjection canonique $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ est un morphisme. Plus généralement si H est un sous-groupe distingué de G, l'application qui à g associe sa classe $\overline{g}$ modulo H, est un morphisme surjectif.

La composée de deux morphismes est évidemment un morphisme ; en outre si le morphisme f est bijectif, on dit alors que f est *un isomorphisme* ou que G et G' sont isomorphes via f, alors son application inverse f^{-1} est aussi un morphisme. Dans le cas où $G' = G$, on dit que f est *un automorphisme* ; l'ensemble $\mathrm{aut}(G)$ des automorphismes de G est par ailleurs un groupe pour la loi de composition.

Lemme 11. *Soit f un morphisme de G vers G'.*

- *Pour tout sous-groupe H de G, l'image $f(H)$ est un sous-groupe de G'.*
- *Pour tout sous-groupe H' de G', l'image réciproque*

$$f^{-1}(H') := \{h \in G : \ f(h) \in H'\}$$

est un sous-groupe de G.

Remarque: on se méfiera de la notation $f^{-1}(H)$ qui laisserait à penser que l'application f^{-1} existerait ce qui n'est à priori pas le cas sauf si f était un isomorphisme.

Définition 12. *On appelle noyau d'un morphisme $f : G \to G'$ et on note* $\mathrm{Ker}\, f$ *l'ensemble $f^{-1}(\{e'\}) := \{g \in G : \ f(g) = e'\}$. L'image de f est notée* $\mathrm{Im}\, f$.

Remarque: $\mathrm{Ker}\, f$ est un sous-groupe distingué de G.

Lemme 13. *Le morphisme $f : G \to G'$ est injectif si et seulement si* $\mathrm{Ker}\, f$ *est réduit à l'élément neutre.*

Théorème 14. ***(de factorisation)***
Soit $f : G \to G'$ un morphisme de groupe. Alors le groupe quotient f induit un isomorphisme de $G/\mathrm{Ker}\, f$ sur $\mathrm{Im}\, f$.

Remarque: pour π un morphisme $G \to H$, on dit que f se factorise par H ou par π s'il existe $\overline{f} : H \to G'$ tel que $f = \overline{f} \circ \pi$. On notera que f se

factorise toujours par un quotient G/H où $H \subset \mathrm{Ker}\, f$: en effet il suffit de poser $\overline{f}(\overline{g}) := f(g)$ puisque $f(gh) = f(g)$ pour tout $h \in H \subset \mathrm{Ker}\, f$.
Exemple : reprenons l'application $\mathbb{Z} \to G$ qui à k associe g^k pour $g \in G$. L'image est $< g >$ et son noyau est un sous-groupe de la forme $n\mathbb{Z}$ de sorte que $< g > \simeq \mathbb{Z}/n\mathbb{Z}$ où n est l'ordre de g. On peut ainsi voir cet ordre comme le plus petit entier strictement positif m tel que $g^m = 1$.

Définition 15. *Une application surjective $f : G \to G'$ de noyau $H = \mathrm{Ker}\, f$ se présente habituellement sous la forme* d'une suite exacte courte *:*

$$1 \to H \to G \to G' \to 1.$$

Remarque: si H est un sous-groupe distingué de G alors H est le noyau du morphisme $G \longrightarrow G/H$.

1.3 Produits semi-direct

Étant donnés deux groupes N et H, on se propose de construire un troisième groupe G contenant un sous-groupe $\overline{N}$ distingué isomorphe à N et $\overline{H} \simeq H$ tel que la surjection canonique $G \twoheadrightarrow G/N$ induisent un isomorphisme $\overline{H} \simeq G/N$. Le produit direct $N \times H$ est un tel exemple mais on aimerait en construire d'autres où $\overline{H}$ ne serait pas distingué dans G. Pour ce faire on a besoin d'un morphisme $\Psi : H \longrightarrow \mathrm{aut}(N)$.

Proposition 16. *Soit G l'ensemble $N \times H$ que l'on munit de la loi de composition interne*

$$(n, h).(n', h') := (n\Psi(h)(n'), hh').$$

Alors G muni de cette loi est un groupe où $\overline{N} := N \times \{1_H\}$ est un sous-groupe distingué de G et $\overline{H} := \{1_N\} \times H$ est isomorphe à $G/\overline{N}$ via la projection canonique.

Remarque: dans le cas où Ψ est trivial, i.e. $\Psi(h) = \mathrm{Id}_N$ pour tout $h \in H$, on retrouve la définition du produit direct $N \times H$.
Preuve : La loi de composition ainsi définie est clairement interne, vérifions son associativité :

$$\begin{aligned}
[(n_1, h_1).(n_2, h_2)].(n_3, h_3) &= (n_1\Psi(h_1)(n_2), h_1h_2).(n_3, h_3) \\
&= (n_1\Psi(h_1)(n_2)\Psi(h_1h_2)(n_3), h_1h_2h_3) \\
&= (n_1\Psi(h_1)(n_2)\Psi(h_1) \circ \Psi(h_2)(n_3), h_1h_2h_3) \\
&= (n_1\Psi(h_1)(n_2\Psi(h_2)(n_3)), h_1h_2h_3) \\
&= (n_1, h_1).(n_2\Psi(h_2)(n_3), h_2h_3) \\
&= (n_1, h_1).[(n_2, h_2).(n_3, h_3)]
\end{aligned}$$

où on utilise que $\Psi : H \longrightarrow \mathrm{aut}(N)$ est un morphisme de groupe, i.e. $\Psi(h_1h_2) = \Psi(h_1) \circ \Psi(h_2)$ et que $\Psi(h_1) \in \mathrm{aut}(N)$ est un morphisme de groupe, i.e.

$$\Psi(h_1)(n_1n_2) = \Psi(h_1)(n_1)\Psi(h_1)(n_2).$$

On vérifie aisément que $(1_N, 1_H)$ est un élément neutre pour cette loi et que l'inverse de (n,h) est $(\Psi(h^{-1})(n^{-1}, h^{-1}))$. Soient alors

$$\overline{N} = \{(n, 1_H) \in G : \ n \in N\} \text{ et } \overline{H} = \{(1_N, h) \in G : \ h \in H\}.$$

L'application naturelle $f_N : N \longrightarrow \overline{N}$ définie par $f_N(n) = (n, 1_H)$ (resp. $f_H : H \longrightarrow \overline{H}$ définie par $f_H(h) = (1_N, h)$), est clairement bijective et est un morphisme de groupe : $f_N(n_1 n_2) = (n_1 n_2, 1_H) = (n_1, 1_H).(n_2, 1_H)$ car $\Psi(1_H) = \mathrm{Id}_N$. On notera $\overline{n}$ (reps. $\overline{h}$) pour $(n, 1_H)$ (resp. $(1_N, h)$). Montrons que le sous-groupe $\overline{N}$ de G est distingué : comme $(n,h) = \overline{n}\overline{h}$, il suffit de montrer que $\overline{N}$ est laissé stable par les automorphismes intérieurs définis par les $\overline{h}$, la stabilité sous les automorphismes intérieurs associés aux éléments du type $\overline{n}$ étant évidente :

$$\overline{h}\overline{n}\overline{h}^{-1} = \overline{h}(n, h^{-1}) = (\Psi(h)(n), 1_H) = \overline{\Psi(h)(n)}. \tag{1}$$

Définition 17. *Le groupe défini dans la proposition précédente est* le produit semi-direct *de N par H via Ψ : on le note $N \rtimes_\Psi H$.*

Remarque: on a une suite exacte courte

$$1 \longrightarrow N \longrightarrow N \rtimes_\Psi H \longrightarrow H \longrightarrow 1.$$

Proposition 18. *Soient ψ, ϕ des morphismes de H vers* $\mathrm{aut}(N)$. *Considérons les deux cas suivants :*

(i) il existe $\alpha \in \mathrm{aut}(H)$ tel que $\psi = \phi \circ \alpha$;

(ii) il existe $u \in \mathrm{aut}(N)$ tel que $\forall h \in H$, $\phi(h) = u\psi(h)u^{-1}$.

Alors $N \rtimes_\psi H \simeq N \rtimes_\phi H$.

Preuve : (i) Soit $f : N \rtimes_\Psi H \longrightarrow N \rtimes_\Phi H$ défini par $f(n,h) = (n, \alpha(h))$; f est un morphisme car

$$f((n,h)(n',h')) = f(n\Psi(h)(n'), hh') = (n\Psi(h)(n'), \alpha(hh'))$$

qui est égal à

$$f(n,h)f(n',h') = (n, \alpha(h))(n', \alpha(h')) = (n\Phi(\alpha(h))(n'), \alpha(h)\alpha(h'))$$

car $\Psi = \Phi \circ \alpha$ et que $\alpha(hh') = \alpha(h)\alpha(h')$. De même on a un morphisme $g : N \rtimes_\Phi H \longrightarrow N \rtimes_\Psi H$ défini par $g(n,h) = (n, \alpha^{-1}(h))$ qui est clairement inverse de f, d'où le résultat.

(ii) Soit $f : N \rtimes_\Psi H \longrightarrow N \rtimes_\Phi H$ défini par $f(n,h) = (u(n), h)$; f est un morphisme car

$$\begin{aligned} f((n,h)(n',h')) &= f(n\Psi(h)(n'), hh') \\ &= (u(n\Psi(h)(n')), hh') = (u(n)\Phi(h)(n'), hh') \\ &= f(n,h)f(n',h'). \end{aligned}$$

Comme précédemment $g : N \rtimes_\Phi H \longrightarrow N \rtimes_\Psi H$ défini par $g(n,h) = (u^{-1}(n), h)$ est clairement le morphisme inverse de f, d'où le résultat.

Considérons à présent la question de savoir reconnaitre si un groupe G est un produit semi-direct. On a des conditions nécessaires évidentes :

— G doit posséder une sous-groupe distingué que l'on note N,

— ainsi qu'un sous-groupe H tel que la restriction à H de la surjection canonique $G \twoheadrightarrow G/N$ soit un isomorphisme. On dit que H est *un relèvement* de G/N.

La formule (1) nous suggère alors de considérer le morphisme

$$\Psi : H \longrightarrow \mathrm{aut}(N)$$

donné par la conjugaison, i.e. $\Psi(h) : \ n \in N \mapsto hnh^{-1} \in N$.

Définition 19. *On dit d'une suite exacte courte associée à $N \triangleleft G$*

$$1 \longrightarrow N \longrightarrow G \longrightarrow G/N \longrightarrow 1,$$

qu'elle est scindée, si elle admet un relèvement, i.e. G admet un sous-groupe H tel que la restriction de $G \twoheadrightarrow G/N$ à H est un isomorphisme.

Proposition 20. *Si la suite exacte courte associée à $N \triangleleft G$ est scindée alors $G \simeq N \rtimes_\Psi H$ où Ψ est donnée par la conjugaison de H sur N comme ci-dessus.*

Preuve : Le groupe N étant distingué dans G, H y agit par automorphismes intérieurs $\Psi : H \longrightarrow \mathrm{aut}(N)$ avec $\Psi(h)(n) = hnh^{-1}$. Soit alors $f : N \rtimes_\Psi H \longrightarrow G$ défini par $f(n,h) = nh$; f est un morphisme de groupe car $f(1_N, 1_H) = 1_G$ et

$$\begin{aligned} f((n,h).(n',h')) &= f(nhn'h^{-1}, hh') \\ &= nhn'h^{-1}hh' \\ &= nhn'h' \\ &= f(n,h)f(n',h'). \end{aligned}$$

Soit $(n,h) \in \mathrm{Ker}\, f$, soit $nh = 1_G$ et donc $n = h^{-1} \in N \cap H$; or $\pi : G \longrightarrow G/N$ induit un isomorphisme $\pi_{|H} : H \simeq G/N$ d'où $\pi(n) = 1_{G/N}$ et comme $n \in H$, on en déduit $n = 1_H = 1_G$ et $(n,h) = (1_N, 1_H)$, d'où l'injectivité. Pour montrer la surjectivité, soit $g \in G$ et soit $h \in H$ tel que $\pi(h) = \pi(g)$; on a alors $n = gh^{-1} \in \mathrm{Ker}\, \pi = N$, soit $g = nh$.

Exemple du groupe diédral : dans le plan affine, on considère le polygone régulier à n cotés, formé par les points d'affixe les racines n-ièmes de l'unité. On considère le sous-groupe G du groupe des isométries du plan, constitué des isométries qui laissent stable ce polygone. Alors G est le produit semi-direct $\mathbb{Z}/n\mathbb{Z} \rtimes_\psi \mathbb{Z}/2\mathbb{Z}$ où $\psi : \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathrm{aut}(\mathbb{Z}/n\mathbb{Z})$ est tel que $\psi(1) = -1$ est la multiplication par -1. Le groupe G ainsi défini est *le groupe diédral* noté D_n. En effet, classiquement toute application affine du plan qui laisse globalement stable le polygone régulier en question, laisse le barycentre invariant

et correspond à une isométrie vectorielle. L'ensemble G de ces isométries vectorielle est donc constitués des rotations d'angle $2k\pi/n$ et des réflexions par rapport aux médiatrices des segments constituants le polygone régulier. La loi sur G est bien évidemment donnée par la composition des applications linéaires ; $|G| = 2n$. Dans G, le sous-groupe N des isométries positives est cyclique engendré par exemple par la rotation r d'angle $2\pi/n$; il est distingué car c'est le noyau du déterminant ; $N \simeq \mathbb{Z}/n\mathbb{Z}$. La suite exacte $1 \longrightarrow N \longrightarrow G \longrightarrow G/N \longrightarrow 1$ est scindée puisqu'un relèvement de $G/N \simeq \mathbb{Z}/2\mathbb{Z}$ est donné par le choix d'une quelconque réflexion s de G. Ainsi on a $G \simeq \mathbb{Z}/n\mathbb{Z} \rtimes_\Psi \mathbb{Z}/2\mathbb{Z}$, où $\Psi(1)(1)$ est l'entier k modulo n tel que $srs = r^k$; classiquement on obtient $k = -1$.

1.4 Groupes résolubles

Définition 21. *Un groupe G est dit* résoluble *s'il possède une filtration croissante par des sous-groupes*

$$1 = G_0 \subset G_1 \subset \cdots \subset G_n = G$$

avec G_i distingué dans G_{i+1} et G_{i+1}/G_i commutatif.

Remarque: moralement tout ce qu'on sait faire pour un groupe commutatif, on devrait pouvoir l'étendre au cas des groupes résolubles.

Définition 22. Le groupe dérivé $D(G)$ *d'un groupe G est le groupe engendré par les commutateurs $[a, b] := aba^{-1}b^{-1}$ pour $a, b \in G$.*

Remarque: de la formule $g[a, b]g^{-1} = [gag^{-1}, gbg^{-1}]$, on en déduit que G est un groupe distingué. En outre $G/D(G)$ est commutatif et vérifie la propriété universelle suivante : tout morphisme $G \to H$ avec H commutatif se factorise par $G \twoheadrightarrow G/D(G)$.

Définition 23. *On définit par récurrence $D^0 = G$ et $D^{n+1}(G) = D(D^n(G))$ pour $n \geq 0$.*

Lemme 24. *Le groupe G est résoluble si et seulement si $D^n(G)$ est trivial pour n assez grand.*

Preuve : Supposons G résoluble et soit $G_0 \subset \cdots \subset G_n = G$ une filtration comme dans la définition 21. D'après la propriété universelle de $D(G)$, la surjection canonique $G \twoheadrightarrow G_n/G_{n-1}$ se factorise par $G/D(G)$ et donc $D(G) \subset G_{n-1}$. Par récurrence simple, on montre que $D^i(G) \subset G_{n-i}$ et donc $D^n(G)$ est trivial.

Réciproquement si $D^n(G)$ est trivial on pose $G_{n-i} = D^i(G)$ et la filtration obtenue convient.

Proposition 25. *Si*

$$1 \to G_1 \longrightarrow G_2 \longrightarrow G_3 \to 1$$

est exacte alors G_2 est résoluble si et seulement si G_1 e G_3 le sont.

Preuve : On a d'une part $D^n(G_2) \longrightarrow D^n(G_3)$ surjectif et $D^n(G_1) \longrightarrow D^n(G_2)$ injectif de sorte que G_2 résoluble implique que G_1 et G_3 le sont. Inversement si $D^n(G_3)$ est trivial, l'image de $D^n(G_2)$ dans G_3 est nul et donc $D^n(G_2)$ est contenu dans G_1. Si en outre $D^m(G_1)$ est trivial alors $D^{m+n}(G_2) \subset D^m(G_1) = 1$ d'où le résultat.
Remarque: en itérant le résultat précédent on obtient le corollaire suivant qui dit que la classe des groupes résolubles est stable par extension.

Corollaire 26. *Si G possède une filtration croissante de sous-groupes*

$$1 = G_0 \subset \cdots \subset G_n = G$$

avec G_i distingué dans G_{i+1} et G_{i+1}/G_i résoluble alors G est résoluble.

1.5 Sur le groupe symétrique

Définition 27. *L'ensemble des bijections de l'ensemble $\{1, \cdots, n\}$ muni de la loi de composition est un groupe noté $\mathfrak{S}_n$ appelé le groupe symétrique d'ordre n ; ses éléments sont appelés des permutations.*

Remarque: pour E un ensemble fini de cardinal, toute bijection de E sur $\{1, \cdots, n\}$, induit un isomorphisme du groupe $\mathfrak{S}(E)$ des bijections de E dans E, sur $\mathfrak{S}_n$.

Lemme 28. *Le cardinal de $\mathfrak{S}_n$ est égal à $n!$.*

Remarque: il n'est pas raisonnable d'espérer comprendre « parfaitement » le groupe $\mathfrak{S}_n$; en effet d'après le théorème de Cayley, en faisant, avec le vocabulaire du paragraphe suivant, opérer tout groupe fini G sur lui-même par translation à gauche, G s'identifie à un sous groupe de $\mathfrak{S}_{|G|}$. Un argument plus convaincant pour justifier l'étude plus précise des $\mathfrak{S}_n$ est l'heuristique suivante : pour comprendre un groupe G, il est en général très instructif de le faire agir sur un ensemble E, i.e. de construire un morphisme $G \to \mathfrak{S}(E)$ c'est même parfois seulement comme cela que le groupe G est défini.

Définition 29. *Les orbites de l'action du groupe engendré par σ sur $\{1, \cdots, n\}$, sont appelés ses cycles ; si $1 \leq k \leq n$ appartient à un cycle de longueur > 1, on dit qu'il appartient au support de σ. Si le support de σ est constitué d'un unique cycle de cardinal m, on dit que σ est un m-cycle.*

Remarque: autrement dit le support d'une permutation σ est l'ensemble des k tels que $\sigma(k) \neq k$. Les dérangements sont les permutations de support maximal, i.e. $\{1, \cdots, n\}$; ce sont en quelque sorte les permutations les plus compliquées celles que l'on ne peut pas identifier avec une permutation d'ordre strictement plus petit. A l'opposé, les permutations les plus simples sont les 2-cycles que l'on appelle les transpositions.

Théorème 30. *Toute permutation $\sigma \in \mathfrak{S}_n$ peut s'écrire comme la composée de cycles à supports disjoints. Cette décomposition est unique au sens où l'ordre de composition de ces cycles est indifférent. Les supports de ces cycles correspondent aux orbites de σ.*

Remarque: le résultat précédent s'appelle *la décomposition à supports disjoints* d'une permutation. En particulier en utilisant que l'ordre d'un m-cycle est m et la commutation de deux cycles à supports disjoints, on en déduit que l'ordre de σ est égal au ppcm des cardinaux de ses orbites.

Notation 2. *On utilisera la notation $(a_1 a_2 \cdots a_r)$ pour désigner le cycle de longueur r qui, pour tout $i = 1, \cdots, r-1$, envoie a_i sur a_{i+1}, et a_r sur a_1.*

Proposition 31. *Soit $c \in \mathfrak{S}_n$ un m-cycle ; pour tout $r \in \mathbb{N}$, la décomposition à support disjoints de c^r admet $m \wedge r$-cycles tous de longueur $\frac{m}{m \wedge r}$. Une permutation $\sigma \in \mathfrak{S}_n$ commute avec c si et seulement elle s'écrit sous la forme $c^r \circ \sigma'$ où le support de σ' est disjoints de celui de c.*

Remarque: le commutant de c en tant que sous-groupe de $\mathfrak{S}_n$ est ainsi isomorphe à $\mathbb{Z}/m\mathbb{Z} \times \mathfrak{S}_{n-m}$.

Si on cherche les générateurs les plus simples possibles, on se tourne vers les transpositions et on peut montrer les résultats suivants :

(i) les transpositions engendrent $\mathfrak{S}_n$;

(ii) les transpositions $(i\ i+1)$ engendrent $\mathfrak{S}_n$;

(iii) les transpositions $(1\ i)$ engendrent $\mathfrak{S}_n$.

Dans les deux derniers cas, on remarque qu'on ne peut pas enlever des transpositions : si (iii) on enlève $(1\ k)$ alors toute permutation dans le groupe engendré par les autres laisse k invariant ; dans (ii) ce sont les sous-ensembles $[1, k]$ et $[k+1, n]$ qui sont stables. On peut en fait montrer le résultat suivant.

Proposition 32. *Soit $\{\tau_1, \cdots, \tau_r\}$ un ensemble de transpositions qui engendrent $\mathfrak{S}_n$, alors $r \geq n-1$.*

Preuve : Considérons le graphe S de sommets numérotés de 1 à n et dont les arêtes sont données par les transpositions $\tau_i = (a\ b)$ reliant a à b. Comme $\mathfrak{S}_n$ agit transitivement, si $\langle \tau_1, \cdots, \tau_r \rangle = \mathfrak{S}_n$ alors S est connexe. Or un graphe connexe à n sommets admet au moins $n-1$ arêtes : en effet notons $\delta_i \geq 1$ le nombre d'arêtes issues de i de sorte que

$$\sum_{i=1}^{n} \delta_i = 2r.$$

Si pour tout $i = 1, \cdots, n$ on a $\delta_i \geq 2$ alors $2r \geq 2n$ et donc $r \geq n$. Dans le cas où il existe i tel que $\delta_i = 1$, en retirant le sommet i et l'unique arête issue de i, on obtient *par récurrence* $r - 1 \geq n - 2$ et donc $r \geq n - 1$.
Remarque: évidemment comme $\mathfrak{S}_n$ n'est pas commutatif pour $n \geq 3$, on ne peut pas trouver un seul générateur, en revanche on peut trouver deux générateurs avec par exemple $(1\ 2)$ et $(1\ 2 \cdots n)$.

Une question usuelle dans l'étude d'un groupe est de comprendre ses classes de conjugaison, autrement dit en langage savant, les orbites de l'action du groupe sur lui-même par conjugaison. Dans le cas du groupe symétrique la question est réglée par la décomposition en cycles à support disjoints via la formule

$$\sigma \circ (a_1\ \cdots\ a_r) \circ \sigma^{-1} = (\sigma(a_1)\ \cdots\ \sigma(a_r)).$$

Proposition 33. *Deux permutations de $\mathfrak{S}_n$ sont conjuguées si et seulement si elles ont le même nombre de cycles de longueur donnée, dans l'écriture de leur décomposition en cycles à supports disjoints*

Remarque: la formule précédente permet de montrer aisément que, pour $n \geq 3$, le centre de $\mathfrak{S}_n$ est réduit à l'identité. Selon le même principe si f est un morphisme de groupes de $\mathfrak{S}_n$ dans $\mathbb{C}^\times$ alors toutes les transpositions ont même image car elles sont toutes conjuguées, ainsi il y a au plus un caractère non trivial, i.e. une représentation de dimension 1, $\mathfrak{S}_n \to GL_1(\mathbb{C})$. Il reste alors à la construire.
Construction de la signature : il y a essentiellement trois façons de la définir.

- la première en imposant $\epsilon(\tau) = -1$ pour toute transposition τ puis $\epsilon(\sigma) = (-1)^r$ où σ peut s'écrire en produit de r transpositions. Il s'agit alors de vérifier que la parité de r ne dépend que de σ, par contre ainsi définie ϵ est clairement un morphisme.
- La deuxième est d'utiliser la décomposition en cycles à supports disjoints et d'imposer $\epsilon(\sigma) = (-1)^{n-L(\sigma)}$ où $L(\sigma)$ est le nombre d'orbites : cette fois ci, ϵ est bien définie par contre il faut vérifier que c'est bien un morphisme.
- Enfin la troisième, et la meilleure, consiste à poser $\epsilon(\sigma) = \prod_{i<j} \frac{\sigma(j)-\sigma(i)}{j-i}$ (parfois on dit que $\epsilon(\sigma) = (-1)^s$ où s est le nombre d'inversions i.e. de couples $i < j$ tels que $\sigma(i) > \sigma(j)$) : ϵ est bien définie et clairement un morphisme.

Définition 34. *Le noyau de la signature est un sous-groupe distingué $\mathcal{A}_n$ dit alterné ; il est de cardinal $\frac{n!}{2}$.*

Proposition 35. *La classe de conjugaison dans $\mathfrak{S}_n$ d'un élément $\sigma \in \mathcal{A}_n$ donne deux classes de conjugaison de $\mathcal{A}_n$ (resp. une unique classe de conjugaison) si et seulement si le commutateur de σ est contenu dans $\mathcal{A}_n$ (resp. sinon).*

Preuve : Tout repose sur la remarque triviale suivante : soient $\tau \in \mathfrak{S}_n \backslash \mathcal{A}_n$ avec $\sigma' = \tau \circ \sigma \circ \tau^{-1}$. Alors il existe $\rho \in \mathcal{A}_n$ tel que $\sigma' = \rho \circ \sigma \circ \rho^{-1}$ si et seulement si $\tau^{-1} \circ \rho$ appartient au commutant de σ ; on conclut alors aisément.
Remarque: le commutateur de $\sigma \in \mathcal{A}_n$ est contenu dans $\mathcal{A}_n$ si et seulement si les longueurs des cycles dans la décomposition en cycles à supports disjoints sont tous impairs sans multiplicité. En effet si c est un tel cycle de longueur paire alors il appartient au commutant et n'appartient pas à $\mathcal{A}_n$; si $c_1 = (a_1 \cdots a_{2r+1})$ et $c_2 = (b_1 \cdots b_{2r+1})$ sont deux tels cycles distincts alors $(a_1 b_1) \circ \cdots \circ (a_{2r+1} b_{2r+1})$ appartient au commutant et pas à $\mathcal{A}_n$.

Proposition 36. *Pour $n \geq 3$, $\mathcal{A}_n$ est engendré par les 3-cycles.*

Corollaire 37. *Le centre de $\mathcal{A}_n$ est réduit à l'identité pour $n \geq 3$.*

Théorème 38. *Pour $n \geq 5$, $\mathcal{A}_n$ est simple.*

Remarque: il y a plusieurs preuves possibles : soit on se ramène au cas $n = 5$, soit en considérant le nombre minimal d'éléments « dérangés ». Dans tous les cas, il s'agit, étant donné un sous-groupe distingué H non trivial de $\mathcal{A}_n$, de construire un 3-cycle dans H de sorte que comme les 3-cycles sont conjugués dans $\mathcal{A}_n$, il les contient tous et est donc égal à $\mathcal{A}_n$. La technique comme d'habitude en théorie des groupes consiste à étudier des commutateurs, i.e. les $[g_1, g_2] = g_1 g_2 g_1^{-1} g_2^{-1}$.
Remarque: via les théorèmes de Sylow, on peut aussi montrer que $\mathcal{A}_5$ est le seul groupe simple d'ordre 60.

Corollaire 39. *Le groupe dérivée de $\mathcal{A}_n$ est égal à $\mathcal{A}_n$ pour $n \geq 5$.*

Remarque: $\mathcal{A}_n$ n'est donc pas résoluble, fait qui à une application spectaculaire sur la non résolution par radicaux des équations polynomiales de degré ≥ 5 ;

De la simplicité de $\mathcal{A}_n$ pour $n \geq 5$, on montre que $\mathcal{A}_n$ est le seul sous-groupe distingué non trivial de $\mathfrak{S}_n$. Le théorème de Sylow nous apprend que $\mathfrak{S}_n$ contient tous les groupes d'ordre n qui sont donc d'indice $(n-1)!$ ce qui est très gros. En ce qui concerne les gros sous-groupes citons le résultat suivant :

Proposition 40. *Si G est un sous-groupe d'indice $1 \leq k \leq n$ de $\mathfrak{S}_n$ avec $n \geq 5$, alors $k = 1, 2, n$ et G est isomorphe à $\mathfrak{S}_n$, $\mathcal{A}_n$ ou $\mathfrak{S}_{n-1}$.*

1.6 Opération d'un groupe sur un ensemble

Définition 41. *Une action d'un groupe G sur un ensemble E est un morphisme de groupes*

$$\phi : G \longrightarrow \mathfrak{S}(E).$$

Remarque: concrètement cela signifie que pour tout $g \in G$, $\phi(g) \in \mathfrak{S}(E)$ est une bijection de E telle que $\phi(gg') = \phi(g) \circ \phi(g')$; en particulier $\phi(1) = Id_E$.
Remarque: on dit parfois que la définition précédente définit une action à gauche, une action à droite étant alors définie comme un morphisme de $G^{op} \to \mathfrak{S}(E)$ où G^{op} est l'ensemble G muni de la loi $g * h = hg$.

Définitions 42. *On considère l'action d'un groupe G sur un ensemble E.*

- *L'orbite d'un élément $e \in E$ est par définition le sous-ensemble $\mathcal{O}_G(e) = \{g.e \ / \ g \in G\}$; évidemment si $e' \in \mathcal{O}_G(e)$ alors $\mathcal{O}_G(e) = \mathcal{O}_G(e')$. On dit que E est G-*homogène*, ou que G agit* transitivement *s'il n'y a qu'une seule orbite.*
- *Le stabilisateur de $e \in E$ est par définition le sous-groupe $\mathrm{Stab}_G(e) = \{g \in G \ / \ g.e = e\}$; si $e' = g.e$ alors $\mathrm{Stab}_G(e') = g\mathrm{Stab}_G(e)g^{-1}$. On dit que G opère* fidèlement *si tous les stabilisateurs sont réduits à l'élément neutre.*

Remarque: on dit que G opère n-transitivement si pour tout $(x_i)_{1\leq i\leq n}$ (resp. $(y_i)_{1\leq i\leq n}$) distincts deux à deux, il existe $g \in G$ tel que pour tout $i = 1, \cdots, n$, $gx_i = y_i$.

Proposition 43. *Pour tout $e \in E$ dont l'orbite est fini, on a $|\mathcal{O}_G(e)| = [G : \mathrm{Stab}_G(e)]$.*

Preuve : Il suffit de noter que l'application qui à $\overline{g}$ associe ge est une bijection d'image $\mathcal{O}_G(e)$.
Remarque: en particulier si G est fini, on peut écrire $|G| = |\mathcal{O}_G(e)|.|\mathrm{Stab}_G(e)|$.
Remarque: la version topologique de l'égalité numérique de la proposition précédente, consiste à dire qu'un ensemble G-homogène est isomorphe à un quotient G/H pour l'action de G par translation à gauche.

Corollaire 44. ***(équations aux classes)***
Pour une action de G sur un ensemble E, on a

$$|E| = |E^G| + \sum_{\mathcal{O}_G(e)\in\mathcal{O} \ / \ |\mathcal{O}_G(e)|\neq 1} |\mathcal{O}_G(e)|$$

où $\mathcal{O}$ est l'ensemble des orbites et E^G désigne l'ensemble des points fixes.

Remarque: l'intérêt de cette formule est de contrôler les orbites de cardinal 1. Comme illustration prenons par exemple G de cardinal 15 et E de cardinal 7. Pour écrire 7 comme une somme de $1, 3, 5, 15$, on a trois solutions :

$$7 = 1+1+1+1+1+1+1 = 1+1+1+1+3 = 1+1+5$$

ce qui donne donc toujours au moins deux orbites de cardinal 1.

Plus intéressante est *la formule de Burnside* qui permet de compter le nombre d'orbites :

$$\sum_{g\in G} |\mathrm{Fix}(g)| = |\mathcal{O}|.|G|.$$

Exemples :

— $E = G$: il y a 3 actions classiques, translation à gauche, à droite par g^{-1} et par conjugaison. Dans ce dernier cas, les orbites sont appelées *les classes de conjugaison.*
— E est un sous-groupe distingué de G : on peut alors faire opérer G par conjugaison
— E est un quotient de G : on fait alors agir G par translation à gauche.
— E est un ensemble de sous-groupe de G, par exemple ses sous-groupes de Sylow
— E est un autre groupe : on peut demander que $G \to \mathfrak{S}(E)$ s'envoie sur $\mathrm{aut}(E)$ ce qui permet de définir la notion de *produit semi-direct* et définir par exemple le groupe diédral.
— E est un espace vectoriel : on peut demander que l'image de G soit contenue dans les applications linéaires. On arrive alors à la notion de *représentations linéaires des groupes.*

1.7 Présentation par générateurs et relations

Soit X un ensemble que nous appellerons *alphabet.* Un *mot* sur l'alphabet X est un élément $u \in \mathcal{M}(X)$ de la forme $u = x_1 \cdots x_n$ où les x_i sont des lettres de X, i.e. $x_i \in X$. La concaténation des mots munit $\mathcal{M}(X)$ d'une structure de monoide avec pour élément neutre le mot vide.

Cas particulier : $X = G_1 \coprod G_2$ où G_1 et G_2 sont deux groupes quelconques d'éléments neutres respectifs e_1, e_2. Pour $g, h \in G_i$, on dispose des mots gh de longueur 2 et de la lettre (gh) que nous souhaiterions identifier. Plus généralement, lorsque x_i, x_{i+1} appartiennent à G_1 (ou à G_2) nous dirons que les mots

$$x_1 \cdots x_{i-1} x_i x_{i+1} \cdots x_n \text{ et } x_1 \cdots x_{i-1}(x_i x_{i+1}) \cdots x_n$$

sont élémentairement équivalents et on introduit la relation d'équivalence sur $\mathcal{M}(X)$, $u \sim v$ lorsque l'on peut passer de l'un à l'autre par une suite finie d'équivalences élémentaires. On vérifie alors aisément la proposition suivante

Proposition 45. *Le quotient $\mathcal{M}(X)/\sim$ muni de la concaténation, est un groupe noté $G_1 * G_2$ appelé le produit libre de G_1 et de G_2.*

Remarque: on notera que $G_1 * G_2 \simeq G_2 * G_1$ et que $G_1 * G_2 * G_3$ ne dépend pas de l'ordre de construction, i.e est égal à $(G_1 * G_2) * G_3$ ou à $G_1 * (G_2 * G_3)$. Il est facile de vérifier que tout classe de $G_1 * G_2$ contient un unique *mot réduit* au sens où deux lettres consécutives d'un mot réduit n'appartiennent pas au même G_i.
Remarque: le morphisme canonique $p_i : G_i \longrightarrow G_1 * G_2$ qui envoie g sur le mot réduit g est clairement injectif. Il est possible de définir $G = G_1 * G_2$ à l'aide d'une propriété universelle en demandant que G soit engendré par

$p_i(G_i)$ et que pour tout groupe H muni de morphismes $q_i : G_i \longrightarrow H$, il existe un unique homomorphisme $q : G \longrightarrow H$ tel que $q_i = q \circ p_i$.
Exemple : considérons $G_1 = \mathbb{Z}$ et $G_2 = \mathbb{Z}$ alors $G = \mathbb{Z} * \mathbb{Z}$ est le groupe libre à 2 générateurs, plus généralement on définit le groupe libre à n générateurs comme $\mathbb{Z} * \cdots * \mathbb{Z}$. Pour G un groupe de type fini, i.e. admettant un nombre fini de générateurs $g_1, \cdots, g_n$, on a un morphisme surjectif

$$\mathbb{Z} * \cdots * \mathbb{Z} \twoheadrightarrow G$$

qui envoie le mot 1 image de la i-ème copie de $\mathbb{Z} \hookrightarrow \mathbb{Z} * \cdots * \mathbb{Z}$ sur g_i dont le noyau $\mathcal{R}$ est appelé *les relations* de la présentation de G par *les générateurs* $g_1, \cdots, g_n$.

Lemme 46. ***(dit du ping-pong*** *Soient G un groupe, G_1 et G_2 deux sous-groupes de G qui l'engendrent. On suppose que G_1 (resp. G_2) possède au moins 3 (resp. 2) éléments et que G opére sur un ensemble X tel qu'il existe deux sous-ensembles X_1 et X_2 non vides et disjoints de X tels que*

$$\begin{cases} \forall g \in G_1 \setminus \{1\}, & g(X_2) \subset X_1 \\ \forall g \in G_2 \setminus \{1\}, & g(X_1) \subset X_2. \end{cases}$$

Alors $G \simeq G_1 * G_2$.

Preuve : Soit w un mot réduit écrit sur l'alphabet $G_1 \coprod G_2$; il s'agit de montrer que w ne représente par le mot vide. Considérons tout d'abord le cas où $w = g_1 h_1 \cdots g_n h_n g_{n+1}$ avec $g_i \in G_1 \setminus \{1\}$ et $h_j \in G_2 \setminus \{1\}$. On a alors

$$w(X_2) = g_1 h_1 \cdots g_n h_n g_{n+1}(X_2) \subset g_1 h_1 \cdots g_n h_n(X_1) \subset \cdots \subset g_1(X_2) \subset X_1.$$

Comme X_2 et X_1 sont disjoints, w ne peut pas être le neutre de G.

Si $w = h_1 g_2 \cdots g_n h_n$, on choisit $g_1 \in G_1 \setminus \{1\}$ et on considère l'élément gwg^{-1} qui d'après le cas précédent n'est pas le neutre de G, tout comme donc w.

Si $w = g_1 h_1 \cdots g_n h_n$, on choisit $g \in G_1 \setminus \{1, g_1^{-1}\}$ puis gwg^{-1}. Enfin pour $w = h_1 g_2 \cdots g_n$, on choisit $g \in G_1 \setminus \{1, g_n\}$ et gwg^{-1}.
Remarque: pour $G = G_1 = G_2 = \mathbb{Z}/2\mathbb{Z}$ agissant sur $\mathbb{R}$ par $\bar{0} \mapsto \mathrm{Id}$ et $\bar{1} \mapsto -\mathrm{Id}$. On prend $X_1 = \mathbb{R}_-^\times$ et $X_2 = \mathbb{R}_+^\times$, on a bien $\bar{1}(X_1) \subset X_2$ et $\bar{1}(X_2) \subset X_1$ alors que $\mathbb{Z}/2\mathbb{Z} \not\simeq \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$ puisque ce dernier est infini et non abélien.

Applications : considérons l'action de $PSL_2(\mathbb{Z})$ agissant sur le demi-plan de Poincarré $\mathcal{H} = \{z \in \mathbb{C} : \ \mathrm{Im}(z) > 0\}$ par homographies, i.e. l'action de $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est donnée par la formule

$$z \in \mathcal{H} \mapsto \frac{az + b}{cz + d} \in \mathcal{H}.$$

On note

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

puis

$$B := AJ = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}.$$

On vérifie aisément que $J^2 = I_2$ et $B^3 = I_2$ dans $PSL_2(\mathbb{Z})$.

Lemme 47. *Les matrices J et B engendrent $PSL_2(\mathbb{Z})$.*

Preuve : Nous allons en fait montrer que $PSL_2(\mathbb{Z})$ est engendré par A et

$$A' := JA^{-1}J = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Pour tout $k \in \mathbb{Z}$, on a

$$A^k \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+kc & b+kd \\ c & d \end{pmatrix}$$

et

$$(A')^k \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c+ka & d+kb \end{pmatrix}.$$

Ainsi pour $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in PSL_2(\mathbb{Z})$, si $|a| \leq |c|$ (resp. $|a| < |c|$), une division euclidienne fournit $k \in \mathbb{Z}$ tel que $|a+kc| < |a|$ (resp. $|c+ka| < |c|$) de sorte qu'après un nombre fini de multiplications à gauche par des puissances de A et A' nous sommes ramenés à une matrice de la forme

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \text{ ou } \begin{pmatrix} 0 & b \\ c & d \end{pmatrix},$$

avec, le déterminant étant conservé, $a = d = 1$ dans le premier cas, et $b = -c = 1$ dans le deuxième et on reconnait A^b (resp. JA^c) dans le premier (resp. deuxième) cas.

Pour appliquer le lemme du ping pong, on étend l'action de $PSL_2(\mathbb{Z})$ à $\mathcal{H} \coprod \hat{\mathbb{R}}$ où $X := \hat{\mathbb{R}} = \mathbb{R} \coprod \{\infty\}$. Pour $X_1 =]0, \infty[\coprod \{\infty\}$ et $X_2 =]-\infty, 0]$ nous avons

$$\begin{cases} J(X_1) = X_2 \\ B(X_2) =]1, \infty[\coprod \{\infty\} \subset X_1 \\ B^2(X_2) =]0, 1] \subset X_1. \end{cases}$$

Le lemme du ping pong nous fournit alors

$$PSL_2(\mathbb{Z} \simeq \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z}.$$

Remarque: de cet isomorphisme, on en déduit que les seuls sous-groupes finis de $PSL_2(\mathbb{Z})$ sont $\mathbb{Z}/2\mathbb{Z}$ et $\mathbb{Z}/3\mathbb{Z}$ et que les autres sont d'une des formes suivantes

$$\begin{cases} \mathbb{Z}/2\mathbb{Z} * F, \\ \mathbb{Z}/3\mathbb{Z} * F, \\ F \end{cases}$$

où F est un groupe libre. Par exemple, en jouant au ping pong comme précédemment, le noyau $\Gamma(2)$ de $PSL_2(\mathbb{Z}) \twoheadrightarrow PSL_2(\mathbb{Z}/2\mathbb{Z})$ est le groupe libre à deux générateurs engendré par $A_2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ et $B_2 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$.

Lemme 48. *Notons*

$$\mathcal{D} = \{z \in \mathcal{H} : \ |\mathrm{Re}\ z| \leq \frac{1}{2} \ \textit{et}\ |z| \geq 1\}.$$

Alors $\mathcal{D}$ est un domaine fondamental pour l'action de $PSL_2(\mathbb{Z})$ sur $\mathcal{H}$, i.e.
- *toute orbite rencontre $\mathcal{D}$ en un ou deux points,*
- *si deux points de $\mathcal{D}$ sont dans une même orbite alors ils sont sur sa frontière.*

Preuve : Pour ce faire fixons $z \in \mathcal{H}$ et cherchons un point de partie imaginaire maximale dans l'orbite O_z de z. Comme $\mathrm{Im}\, M(z) \geq \mathrm{Im}\, z \Leftrightarrow |cz + d| \leq 1$, cela ne donne qu'un nombre fini de couples $(c, d) \in \mathbb{Z}^2$, ce qui ne donne donc qu'un nombre fini de points dans O_z de partie imaginaire supérieure ou égale à celle de z. Soit donc $z_1 \in O_z$ de partie imaginaire maximale, quitte à appliquer A^n qui agit par $A^n(z_1) = z_1 + n$, on peut supposer $|\mathrm{Re}\ z_1| \leq \frac{1}{2}$. Ensuite comme $J(z_1) = \frac{-1}{z_1}$ est de partie imaginaire inférieure ou égale à celle de z_1, on en déduit que $|z_1| \geq 1$ et par suite $z_1 \in \mathcal{D}$.

Soient à présent z et z' deux points de $\mathcal{D}$ dans une même orbite ; on suppose $\mathrm{Im}\, z' \geq \mathrm{Im}\, z$ et notons $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in PSL_2(\mathbb{Z})$ tel que $z' = A(z)$. On a alors $|c|\, \mathrm{Im}\, z \leq 1$ avec $\mathrm{Im}\, z > \sqrt{3}/2$ et donc $|c| < 2$.
- Si $c = 0$ alors $ad = 1$ et quitte à changer A en $-A$, on a $a = d = 1$ et $z' = z + b$ ce qui impose $b = \pm 1$ avec z et z' sur la frontière de $\mathcal{D}$, i.e. sur les droites $\mathrm{Re}\ z = \pm\frac{1}{2}$.
- Si $c = 1$ (ou $c = -1$ quitte à changer A en $-A$), la condition $|z+d| \leq 1$ impose au choix
  - $d = 0$ et $|z| = 1$ et alors $b = -\det A = -1$ et $z' = a - z^{-1}$ ce qui impose $a = 0$ ou $a = 1$ et $z = -j^{-1}$, ou $a = -1$ et $z = j$. On est bien sur la frontière avec au plus deux éléments de $\mathcal{D}$ dans cette orbite.
  - Ou bien $z = j$ et $d = 1$ avec $\det A = a - b = 1$ et $z' = a + j$ et donc $a = 0$ ou 1 et donc au plus deux points de l'orbite dans $\mathcal{D}$.
  - Ou bien encore $z = -j^{-1}$ et $d = -1$ et la discussion est similaire.

Remarque: notons que cette démonstration permet de prouver à nouveau que $PSL_2(\mathbb{Z})$ est engendré par A et J puisque le raisonnement précédent fonctionne pour le sous-groupe G de $PSL_2(\mathbb{Z})$ engendré par A et J. Ainsi pour $M \in PSL_2(\mathbb{Z})$, on fixe z dans l'intérieur de $\mathcal{D}$ et on note $z' = M(z)$. Il existe alors une matrice $M' \in G$ telle que $M'(z') = M'M(z) \in \mathcal{D}$ qui est nécessairement égal à z, puisque $\mathcal{D}$ est un domaine fondamental pour l'action de $PSL_2(\mathbb{Z})$ sur $\mathcal{H}$. On en déduit donc que $MM' = I_2$ dans $PSL_2(\mathbb{Z})$ et donc $M \in G$.

1.8 Caractères des groupes abéliens finis

Soit G un groupe abélien fini noté multiplicativement.

Définition 49. *On appelle* caractère *de G, tout morphisme de groupes de G dans $\mathbb{C}^\times$. On note $\widehat{G}$ l'ensemble des caractères de G ; c'est un sous- groupe de l'ensemble des fonctions de G dans $\mathbb{C}^\times$.*

Remarque: comme G est fini, d'après le théorème de Lagrange pour tout $g \in G$, on a $g^{\sharp G} = 1_G$. Ainsi pour tout caractère χ de G, on a $\chi(g)^{\sharp G} = 1$ i.e. les valeurs prises par χ sont des racines de l'unité. En particulier le conjugué $\overline{\chi}$ d'un caractère χ de G est égal à χ^{-1}.

Lemme 50. *Soit G un groupe cyclique d'ordre n. Alors $\widehat{G}$ est isomorphe à G.*

Preuve : Notons g un générateur de G ; tout caractère $\chi \in \widehat{G}$ est déterminé par $\chi(g)$ qui est une racine n-ième de l'unité de sorte que $\widehat{G}$ s'identifie à un sous-groupe de $\mathbb{U}_n$, le groupe des racines n-ième de l'unité dans $\mathbb{C}$.

Inversement si $\xi \in \mathbb{U}_n$ alors l'application $g^i \mapsto \xi^i$ définit un élément de $\widehat{G}$. Ainsi G s'identifie à $\mathbb{U}_n$ lequel est bien isomorphe à G.

Proposition 51. *Pour tout groupe abélien fini, $\widehat{G}$ est isomorphe à G.*

Preuve : D'après la remarque de la fin du paragraphe 6.3, G est un produit direct de groupe cyclique. Le résultat découle alors directement du lemme précédent.

Remarque: pour tout $g \in G$, l'application $\chi \mapsto \chi(g)$ définit un élément de $\widehat{\widehat{G}}$ et donc une identification canonique

$$G = \widehat{\widehat{G}}.$$

Proposition 52. *Soit G un groupe abélien fini. Pour tout caractère χ de G, on a*

$$\sum_{g \in G} \chi(g) = \begin{cases} 0 & si\ \chi \neq 1, \\ \sharp G & si\ \chi = 1. \end{cases}$$

Remarque: d'après la remarque précédente on a aussi

$$\sum_{\chi\in\widehat{G}}\chi(g)=\begin{cases}0 & \text{si } g\neq 1,\\ \sharp G & \text{si } g=1.\end{cases}$$

Preuve : Le cas $\chi=1$ est évident. Supposons donc $\chi\neq 1$ et soit $h\in G$ tel que $\chi(h)\neq 1$. Comme $g\mapsto hg$ est une permutation de g, on a

$$\sum_{g\in G}\chi(g)=\sum_{g\in G}\chi(hg)=\chi(h)\sum_{g\in G}\chi(g)$$

et donc, comme $\chi(h)\neq 1$, il vient nécessairement $\sum_{g\in G}\chi(g)=0$.

1.9 Transformation de Fourier discrète

Soient $n\geq 1$ et

$$\begin{array}{rccc}\mathbf{e}_n: & \mathbb{Z}/n\mathbb{Z} & \longrightarrow & \mathbb{C}^\times\\ & x & \mapsto & \exp(2\pi ix/n)\end{array}$$

Remarque: on rappelle que les caractères (additifs) de $\mathbb{Z}/n\mathbb{Z}$ sont les $\mathbf{e}_n^a$ pour $a\in\mathbb{Z}/n\mathbb{Z}$. Pour $q=p^r$, les caractères additifs de $\mathbb{F}_q$ sont les

$$x\mapsto\exp(2i\pi\mathrm{tr}(ax)/p)$$

pour $a\in\mathbb{F}_q$ et $\mathrm{tr}: x\in\mathbb{F}_q\mapsto x+x^p+\cdots+x^{p^{r-1}}\in\mathbb{F}_p$.

Définition 53. *Pour G un groupe fini, l'algèbre de groupe $\mathbb{C}[G]$ est le $\mathbb{C}$-espace vectoriel $\mathbb{C}^G$ des fonctions de G dans $\mathbb{C}$, muni du produit de convolution*

$$f_1,f_2\in\mathbb{C}[G]\mapsto(f_1*f_2)(g)=\sum_{h\in G}f_1(h)f_2(h^{-1}g).$$

Remarque: Le produit de convolution est commutatif, associatif et bilinéaire en chacune des variables. On calcule par exemple $\delta_g*\delta_h=\delta_{gh}$ et $(f*\delta_g)(h)=f(hg^{-1})$.

Pour toute fonction f sur $\mathbb{Z}/n\mathbb{Z}$ à valeurs complexes on note

$$\langle f,g\rangle:=\sum_{t\in\mathbb{Z}/n\mathbb{Z}}f(t)\overline{g}(t).$$

Définition 54. ***La transformée de Fourier discrète*** *$\mathcal{F}(f)$ d'une fonction $f\in\mathbb{C}[\mathbb{Z}/n\mathbb{Z}]$ est la fonction $\mathbb{Z}/n\mathbb{Z}\longrightarrow\mathbb{C}$,*

$$\zeta\mapsto\langle f,[\times\zeta]^*\mathbf{e}_n\rangle:=\sum_{t\in\mathbb{Z}/n\mathbb{Z}}f(t)\mathbf{e}_n(-t\zeta),$$

où pour toute fonction g, on note $[\times\zeta]^*g$ la translatée multiplicative $t\mapsto g(t\zeta)$.

Remarque: comme le morphisme $\zeta \mapsto [\times\zeta]^*\mathbf{e}_n$ est un isomorphisme $\mathbb{Z}/n\mathbb{Z} \longrightarrow \widehat{\mathbb{Z}/n\mathbb{Z}}$, on peut voir $\mathcal{F}(f)$ comme une fonction sur $\widehat{\mathbb{Z}/n\mathbb{Z}}$. Plus généralement si G est un groupe fini abélien, on définit la transformée de Fourier sur G comme la fonctionnelle $\mathcal{F} : \mathbb{C}[G] \longrightarrow \mathbb{C}[\hat{G}]$ qui à f associe la fonction

$$\chi \in \hat{G} \mapsto \langle f, \chi\rangle = \sum_{g\in G} f(g)\overline{\chi(g)}.$$

Lorsque $G = (\mathbb{Z}/2\mathbb{Z})^k$, cette transformée de Fourier est appelée *transformée de Walsh.*
Exemples : pour tout $x \in \mathbb{Z}/n\mathbb{Z}$ on note δ_x la fonction caractéristique de $\{x\} \subset \mathbb{Z}/n\mathbb{Z}$. Par définition on a alors

$$\mathcal{F}(\delta_{-x}) = [\times x]^*\mathbf{e}_n$$

et en utilisant l'orthogonalité des caractères $\langle[\times x]^*\mathbf{e}_n, [\times\zeta]^*\mathbf{e}_n\rangle = p[x = \zeta?]$, où $[P?]$ vaut 1 si P est vraie et 0 sinon, on obtient

$$\mathcal{F}([\times x]^*\mathbf{e}_n) = p\delta_x.$$

On en déduit alors immédiatement que $\mathcal{F}^2 = p[\times(-1)]^*$, i.e.

$$\mathcal{F}(\mathcal{F}(f))(x) = pf(-x),$$

ainsi que l'identité de Parseval

$$\langle f, g\rangle = \frac{1}{p}\langle\mathcal{F}(f), \mathcal{F}(g)\rangle.$$

Lemme 55. *La transformation de Fourier transforme produit de convolution en produit usuel, i.e.*

$$\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g).$$

Exemples : pour χ un caractère multiplicatif de $\mathbb{F}_p^\times$ qu'on étend à $\mathbb{F}_p$ en posant $\chi(0) = 0$, on obtient

$$\mathcal{F}(\chi) = g_\chi\overline{\chi} + [\chi = \mathbf{1}?](p-1)\delta_0$$

où $\mathbf{1}$ désigne le caractère trivial de $\mathbb{F}_p^\times$ et g_χ est la somme de Gauss, cf. le §**??**

$$g_\chi = \sum_{t\in\mathbb{F}_p^\times} \chi(t)\exp(2i\pi t/p).$$

En particulier si $\chi \neq \mathbf{1}$, il résulte de la formule de Parseval que

$$|g_\chi| = \sqrt{p}.$$

Pour χ_1 et χ_2 deux caractères multiplicatifs, on a

$$\chi_1 * \chi_2 = J(\chi_1, \chi_2).\chi_1\chi_2 + [\chi_1\chi_2 = \mathbf{1}?]p\chi_1(-1)\delta_0,$$

où, cf. le §**??**

$$J(\chi_1, \chi_2) = \sum_{a\in\mathbb{F}_p} \chi_1(a)\chi_2(1-a).$$

En particulier lorsque $\chi_1\chi_2 \neq \mathbf{1}$ alors en appliquant $\mathcal{F}$ à l'égalité $\chi_1 * \chi_2 = J(\chi_1, \chi_2).\chi_1\chi_2$, on obtient

$$g_{\chi_1} g_{\chi_2} = J(\chi_1, \chi_2) g_{\chi_1\chi_2}.$$

Plus généralement, pour des caractères multiplicatifs $\chi_1, \cdots, \chi_r$ tels que $\chi_1 \cdots \chi_r \neq \mathbf{1}$, alors $\chi_1 * \cdots * \chi_2 = J(\chi_1, \cdots, \chi_r)\chi_1 \cdots \chi_r$ et donc, en appliquant $\mathcal{F}$, cf. le théorème **??**

$$g_{\chi_1} \cdots g_{\chi_r} = J(\chi_1, \cdots, \chi_r) g_{\chi_1 \cdots \chi_r}$$

où, cf. la définition **??**

$$J(\chi_1, \cdots, \chi_r) := \chi_1 * \cdots * \chi_r(1) = \sum_{a_1+\cdots+a_r=1} \chi_1(a_1) \cdots \chi_r(a_r).$$

Transformée de Fourier rapide : Considérons un signal f dont on prend un échantillon $\big(f(0), \cdots, f(n-1)\big)$ et considérons

$$\mathcal{F}(f) : k \in \mathbb{Z}/n\mathbb{Z} \mapsto \sum_{i=0}^{n-1} f(i)\mathbf{e}_n(-kn)$$

dont on propose de donner un algorithme de calcul rapide appelé *transformée de Fourier rapide*. Notons tout d'abord qu'un algorithme naif qui consiste à calculer tous les $f(i)\mathbf{e}(-ki)$ et à les additionner pour calculer chacun des $\mathcal{F}(f)(k)$ nécessite de l'ordre de $2n^2$ opérations. Pour expliquer le principe de l'algorithme de Cooley-Tukey, considérons le cas $n = 2^p$. On écrit alors

$$\mathcal{F}(f)(k) = \sum_{i=0}^{2^{p-1}-1} f(2i)\mathbf{e}_n(-2ki) + \sum_{i=0}^{2^{p-1}-1} f(2i+1)\mathbf{e}_n(-k(2i+1)).$$

— Pour $k \in \{0, 1, \cdots, 2^{p-1}-1\}$, l'équation précédente s'écrit comme la somme des deux transformées de Fourier discrètes $\mathcal{F}(f)(k) = \mathcal{F}(f_0)(k) + e^{-2ik\pi/n}\mathcal{F}(f_1)(k)$ où f_0 (resp. f_1) est l'échantillonnage de f aux entiers pairs (resp. impairs).
— Pour $k \in \{2^{p-1}, \cdots, 2^p-1\}$, on pose $k' = k - 2^{p-1}$ ce qui donne $\mathcal{F}(f)(k) = \mathcal{F}(f_0)(k') - e^{-2ik'\pi/n}\mathcal{F}(f_1)(k')$.

En résumé la transformée de Fourier $\mathcal{F}(f)$ se calcule à partir de celle de f_0 et f_1, lesquelles sont de taille moitié. On procède récursivement pour calculer $\mathcal{F}(f_0)$ et $\mathcal{F}(f_1)$ de sorte que si $C(p)$ désigne la complexité de l'algorithme on a l'équation fonctionnelle

$$C(p) = 2C(p-1) + 2^{p+1}$$

ce qui donne $C(p) = O(2^p p)$, soit en revenant à n, une complexité en $O(n \log_2 n)$. Dans le cas où n n'est pas une puissance de 2, on utilise l'écriture de n en base 2 et on obtient la même complexité en $O(n \log_2 n)$.

2 Polynômes

2.1 Généralités sur les anneaux, corps et algèbres

Définition 56. *On appelle* anneau *un triplet formé d'un ensemble A et de deux lois de composition interne, une addition $(x, y) \mapsto x + y$ et une multiplication $(x, y) \mapsto xy$, tels que :*

- *$(A, +)$ est un groupe commutatif d'élément neutre noté 0 ;*
- *la multiplication est associative et possède un élément neutre noté 1 ;*
- *la multiplication est distributive par rapport à l'addition, i.e.*

$$x(y+z) = xy + xz \text{ et } (x+y)z = xz + yz \quad \forall x, y, z \in A.$$

Remarque: dans certains ouvrages, on en demande pas à A de posséder un élément neutre pour la multiplication et on parle d'anneau unitaire dans le cas où elle en possède un.
Remarque: si la multiplication est commutative, on dit que A est *un anneau commutatif.*
Exemples :

- l'ensemble $\mathbb{Z}$ des entiers relatifs muni de l'addition et de la multiplication est un anneau commutatif ; de même $\mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$ sont des anneaux commutatifs.
- Pour X un ensemble et A un anneau, l'ensemble des applications de X à valeurs dans A est un anneau.
- Pour A un anneau et $n \geq 1$ un entier, l'ensemble $\mathbb{M}_n(A)$ des matrices carrées de taille n à coefficients dans A est un anneau non commutatif.
- Pour A un anneau, l'ensemble $A[X]$ des polynômes à coefficients dans A est un anneau.
- Pour $A_1, \cdots, A_n$ des anneaux, le produit cartésien $A = A_1 \times \cdots \times A_n$ muni de l'addition et de la multiplication composante par composante, est un anneau dit *anneau produit* des A_i.

Remarque: comme dans le paragraphe précédent, habituellement pour montrer qu'un triplet $(A, +, \times)$ est un anneau, on essaie de montrer qu'il s'agit

d'un sous-anneau d'un anneau déjà construit, *un sous-anneau* étant un sous-groupe contenant l'élément neutre pour la multiplication et stable par produit.

Définition 57. *Un corps est un anneau commutatif dont tous les éléments non nuls sont inversibles.*

Remarque: un anneau non commutatif dont tous les éléments admettent des inverses à gauche et à droite est généralement appelé *une algèbre à division* : le lecteur pourra alors vérifier que les inverses à gauche et à droite coincident forcément.

Définition 58. Un idéal *à gauche (resp. à droite) I d'un anneau A est un sous-groupe de $(A, +)$ tel que pour tout $a \in A$ et pour tout $i \in I$, l'élément ai (resp. ia) de A appartienne à I. Un idéal est dit* bilatère *si c'est un idéal à gauche et à droite.*

Remarque: si l'anneau A est commutatif alors tout idéal à gauche ou à droite est bilatère.
Exemples :

- les idéaux de $\mathbb{Z}$ sont les $n\mathbb{Z}$;
- pour X un ensemble et Y un sous-ensemble, le sous-ensemble de $F(X, \mathbb{R})$ formé des applications qui s'annulent sur Y est un idéal ;
- pour $a \in A$, l'ensemble aA (resp. Aa) est un idéal à droite (resp. à gauche) ; c'est *l'idéal principal* engendré par a que l'on note (a) dans le cas où A est commutatif.
- Si A est un corps alors ses seuls idéaux sont (0) et lui même : en effet dès qu'un idéal contient un inversible il est égal à tout l'anneau.

Remarque: pour I un idéal d'un anneau A on définit comme précédemment le quotient A/I qui est donc muni d'une loi de groupe induite par celle de A. La propriété de stabilité par multiplication à gauche par les éléments de A est alors juste celle qui est nécessaire pour que la multiplication de A induise sur A/I une structure d'anneau.

Définition 59. *Un morphisme d'anneaux $f : A \to A'$ est un morphisme de groupe qui envoie l'élément neutre* 1 *de A sur celui de A' et tel que pour tout $x, y \in A$, on ait $f(xy) = f(x)f(y)$.*

Remarque: la surjection canonique $A \twoheadrightarrow A/I$ est un morphisme d'anneau.

Lemme 60. *Soit $f : A \to B$ un morphisme d'anneaux et A', B' des sous-anneaux de A et B respectivement.*

- *L'image $f(A')$ est un sous-anneau de B.*
- *L'image réciproque $f^{-1}(B')$ est un sous-anneau de A.*

Si J est un idéal de B alors $f^{-1}(J)$ est un idéal de A.

Remarque: en revanche l'image d'un idéal n'est pas nécessairement un idéal ; considérer par exemple l'inclusion de $\mathbb{Z}$ dans $\mathbb{Q}$. En revanche si f est surjective alors l'image de tout idéal de A est un idéal de B.
Remarque: on a aussi une version du théorème de factorisation pour les morphismes d'anneaux puisque Ker f est clairement un idéal.

Définitions 61. *On dira d'un anneau A qu'il est :*
- intègre *si l'égalité $xy = 0$ avec $x \neq 0$ implique $y = 0$;*
- principal *si tous ses idéaux sont principaux ;*
- noethérien *si toute chaine croissante $I_1 \subset I_2 \subset \cdots$ d'idéaux est stationnaire, i.e. il existe $n \geq 1$ tel que pour tout $r \geq 0$, $I_{n+r} = I_n$;*
- artinien *si toute chaine décroissante $\cdots \subset I_2 \subset I_1$ d'idéaux est stationnaire, i.e. il existe $n \geq 1$ tel que pour tout $r \geq 0$, $I_{n+r} = I_n$;*
- euclidien *s'il est intègre et qu'il existe $\nu : A - \{0\} \to \mathbb{N}$ tel que pour tout $a, b \neq 0 \in A$ il existe $(q, r) \in A^2$ tel que $a = bq + r$ avec $r = 0$ ou $\nu(r) < \nu(b)$.*

Définitions 62. *On dira d'un idéal I de A qu'il est :*
- maximal *s'il est pour l'inclusion, i.e. $I \subset J$ alors soit $J = I$ soit $J = A$;*
- premier *si $xy \in I$ avec $x \notin I$ implique $y \in I$;*
- primaire *si $xy \in I$ avec $x^n \notin I$ pour tout $n \geq 1$ implique qu'il existe $n \geq 1$ tel que $y^n \in I$.*

Définitions 63. *Un élément $a \in A$ est dit :*
- inversible *s'il possède un inverse pour la multiplication ; on note $A^\times$ l'ensemble des éléments inversibles de A qui est alors un groupe pour la multiplication appelé* le groupe des inversibles *de A ;*
- un diviseur de zéro, *s'il existe $b \in A$ non nul tel que $ab = 0$;*
- nilpotent *s'il existe n tel que $a^n = 0$.*

Remarque: un idéal I de A est maximal (resp. premier, resp. primaire) si et seulement si A/I est un corps (resp. intègre, resp. ses diviseurs de zéro sont nilpotents).

Définition 64. *On dit qu'un ensemble E est inductif si toute partie non vide totalement ordonnée admet un majorant dans E.*

Remarque: $\mathbb{R}$ muni de la relation d'ordre usuelle n'est pas inductif ; en revanche l'ensemble des parties d'un ensemble ordonné par l'inclusion est inductif.

Lemme 65. ***(dit de Zorn)***
Tout ensemble non vide inductif admet un élément maximal.

Remarque: ce lemme peut être vu comme un axiome de la théorie des ensembles ; il est en fait équivalent à *l'axiome du choix* qui affirme que si $(E_i)_{i \in I}$ est une famille d'ensembles non vide alors $\prod_{i \in I} E_i$ est non vide.

Proposition 66. *Tout anneau non nul admet un élément maximal.*

Preuve : Soit E la famille des idéaux propres de A ; comme A est non nul, $\{0\}$ est dans E qui est donc non nul. L'ensemble E est inductif : en effet la réunion d'une famille totalement ordonnée d'idéaux propres est encore un idéal propre qui est un majorant. On conclut alors en invoquant le lemme de Zorn.

2.2 Généralités sur les polynômes

Dans ce qui suit $\mathbb{K}$ désigne un corps quelconque que l'on pourra dans un premier temps supposé égal à $\mathbb{R}$ ou $\mathbb{C}$.

Définitions 67. *Rappelons qu'un polynôme $P \in \mathbb{K}[X]$ est par définition une suite $(a_i)_{i\geq 0}$ à support fini, i.e. il existe n tel que pour tout $m > n$, $a_m = 0$. Si le polynôme est non nul, il existe alors un tel n tel que $a_n \neq 0$ que l'on appelle* le degré *de P que l'on note* $\deg(P)$*. Il est dit* unitaire *si* le coefficient dominant $a_{\deg(P)}$ *est égal à* 1*.*

Remarque: $\mathbb{K}[X]$ s'identifie avec $\mathbb{K}^{(\mathbb{N})}$.
Remarque: par convention le degré du polynôme nul est posé égal à $-\infty$; on ordonne alors $\mathbb{N} \cup \{-\infty\}$ en rendant $-\infty$ plus petit que tout élément de $\mathbb{N}$. On prolonge ensuite l'addition de $\mathbb{N}$ en posant $(-\infty) + n = -\infty$ et $(-\infty) + (-\infty) = -\infty$. Avec ces conventions on a le lemme suivant.

Lemme 68. *Soient $P, Q \in \mathbb{K}[X]$ alors*
- $\deg(P+Q) \leq \max\{\deg(P), \deg(Q)\}$ *avec égalité si* $\deg(P) \neq \deg(Q)$ *;*
- $\deg(PQ) = \deg(P) + \deg(Q)$.

Remarque: en particulier on en déduit que $\mathbb{K}[X]$ est un anneau intègre dont les éléments inversibles sont les polynômes constants non nuls que l'on identifie à $\mathbb{K}^\times$.

Définition 69. La valuation *d'un polynôme est le plus petit m tel que $a_m \neq 0$; on définit la valuation du polynôme nul comme étant égale à $+\infty$.*

Remarque: L'anneau $K[X]$ possède exactement les mêmes propriétés arithmétiques que $\mathbb{Z}$ et, souvent en utilisant les notions de degré, valuation et dérivation, les conjectures d'arithmétiques sur $\mathbb{Z}$, transposées à $K[X]$, sont démontrées et sont parfois une source d'inspiration : cf. le §**??** avec par exemple la conjecture abc ou plus profond encore, les travaux de P. Scholze sur les perfectoides.

Théorème 70. *Soient A et B des polynômes de $K[X]$ avec $B \neq 0$. Il existe alors un unique couple $(Q, R) \in K[X]$ tel que*

$$\begin{cases} A = BQ + R \\ \deg(R) < \deg(B) \end{cases}$$

Remarque: Q s'appelle le quotient et R le reste de *la division euclidienne* de A par B.

Preuve : Elle est basée sur le fait élémentaire suivant : soient $U, V \in K[X]$ avec $k = \deg(U) \geq \deg(V) = q$ si a_k (resp. b_q) désigne le coefficient dominant de U (resp. de V) alors pour $Q = \frac{a_k}{b_q} X^{k-q}$ on a $\deg(U - VQ) < \deg(U)$.

Existence : soit $\mathcal{A} = \{A - BQ : \ Q \in K[X]\}$ et notons r le plus petit des degrés de ses éléments. Si on avait $r \geq \deg(B) \geq 0$ alors en appliquant ce qui précède, on construit un monôme Q' tel que $A - B(Q + Q') \in \mathcal{A}$ et de degré $< r$ d'où la contradiction.

Unicité : soient Q_1, Q_2 tels que $\deg(A - Q_i B) < \deg(B)$ pour $i = 1, 2$. On en déduit que

$$\deg(B) > \deg\Big((a - BQ_1) - (A - BQ_2)\Big) = \deg(B(Q_2 - Q_1))$$

et donc $Q_1 = Q_2$.

Muni de cette division euclidienne, on peut reprendre les énoncés sur $\mathbb{Z}$ et on obtient que :

- les idéaux de $K[X]$ sont principaux, i.e. engendrés par un unique polynôme ;
- notion de pgcd, ppcm ;
- relation de Bézout que l'on peut calculer via l'algorithme d'Euclide ;
- les lemmes d'Euclide et de Gauss sont vérifiés ;
- les éléments premiers sont les polynômes irréductibles et tout polynôme se décompose de manière unique aux inversibles près, comme un produit de polynômes premiers ;
- le quotient $K[X]/(P)$ est par définition l'ensemble des classes d'équivalence pour la relation d'équivalence

$$Q \sim Q' \Leftrightarrow P | (Q - Q').$$

Remarque: les lois $+, \times$ et la multiplication par un scalaire de K, munissent alors ce quotient d'une structure d'algèbre : comme dans le cas de $\mathbb{Z}$, on remarque les calculs dans le quotient sont indépendants du choix des représentants dans $K[X]$.
Toute classe d'équivalence possède un unique représentant dont le degré est strictement inférieur à celui de P : il se calcule comme le reste de la division euclidienne par P.

Proposition 71. *Soit P un polynôme non constant ; la classe $\overline{A}$ d'un polynôme $A \in K[X]$ est inversible dans $K[X]/(P)$ si et seulement si A est premier avec P.*

Preuve : Si $\overline{A}$ est inversible alors il existe $\overline{B}$ tel que $\overline{A}.\overline{B} = \overline{1}$, autrement dit il existe Q tel que $AB + PQ = 1$ et donc $A \wedge P = 1$. Réciproquement si $A \wedge P = 1$, on considère une relation de Bezout $AB + PQ = 1$ de sorte que $\overline{B}$ est l'inverse de $\overline{A}$ dans $K[X]/(P)$.

Corollaire 72. *L'anneau $K[X]/(P)$ est un corps si et seulement si P est irréductible.*

2.3 Théorème de Gauss

Dans ce qui suit A désigne un anneau factoriel, et donc intègre, de corps des fractions K.

Définition 73. *On dit qu'un polynôme $P \in A[X]$ est* irréductible *si*
- $P \notin A[X]^* = A^*$,
- $\forall Q, R \in A[X]$ *tels que* $P = QR$, *on a* $Q \in A^*$ *ou* $R \in A^*$.

Remarque: L'hypothèse de factorialité de l'anneau A implique que $A[X]$ est factoriel. Par ailleurs, lorsque k est un corps, $k[X]$ est euclidien, donc factoriel. Ainsi, on se place toujours dans un cadre qui assure l'existence et l'unicité de la décomposition d'un polynôme en produits de facteurs irréductibles.

Exemples :
- Sur un corps algébriquement clos, les seuls polynômes irréductibles sont les polynômes de degré 1.
- Sur $\mathbb{R}$, les polynômes irréductibles sont les polynômes de degré 1 et les polynômes de degré 2 de la forme $aX^2 + bX + c$, avec $a \neq 0$ et $b^2 - 4ac < 0$.
- Nous verrons plus loin que sur $\mathbb{Q}$ ou sur un corps fini, il existe des polynômes irréductibles de n'importe quel degré.

Définition 74. *Soit $P \in A[X]$. On appelle* contenu *de P le* pgcd *de ses coefficients. On le note $c(P)$. On dit que P est* primitif *lorsque $c(P) = 1$.*

Lemme 75. ***Gauss*** *Pour tous $P, Q \in A[X]$, on a $c(PQ) = c(P)c(Q)$.*

Lemme 76. *Soient $P, Q \in A[X]$ tel que P soit unitaire et PQ soit unitaire et à coefficients entiers. Alors Q est unitaire et P et Q sont à coefficients entiers.*

Preuve : Le fait que Q est unitaire est évident. Écrivons maintenant $P = X^\alpha + \frac{1}{\mu}\sum_{i=0}^{\alpha-1} p_i X^i$, où les entiers $\mu, p_0, \ldots, p_{\alpha-1}$ sont premiers entre eux dans leur ensemble (pour cela, il suffit de choisir pour μ le ppcm des dénominateurs des coefficients de P). On écrit de même $Q = X^\beta + \frac{1}{\nu}\sum_{i=0}^{\beta-1} q_i X^i$, avec $\nu, q_0, \ldots, q_\nu$ premiers entre eux dans leur ensemble. On sait alors que les polynômes μP et νQ sont à coefficients entiers et de contenu 1. On a doncv1 $= c(\mu P)c(\nu Q) = c(\mu P.\nu Q) = \mu\nu.c(PQ) = \mu\nu$, ce qui implique que $\mu = \nu = 1$, i.e. que P et Q sont à coefficients entiers.

Proposition 77. *Soit A un anneau factoriel et* $\mathrm{frac}(A)$ *son corps des fractions. Les polynômes irréductibles de $A[X]$ sont :*
- *les constantes irréductibles dans A,*
- *les polynômes non constants primitifs et irréductibles dans* $\mathrm{frac}(A)$.

Remarque: Le polynôme $2X$ et irréductible sur $\mathbb{Q}$ mais pas sur $\mathbb{Z}$.

2.4 Racines d'un polynôme

Rappelons qu'à un polynôme on associe habituellement sa fonction polynôme et que dans le cas où $\mathbb{K}$ est infini, cette dernière détermine le polynôme dont on est parti ; dans ce qui suit on cèdera à la facilité de cette identification.

Définition 78. *On dit que $a \in \mathbb{K}$ est une racine de P si $P(a) = 0$.*

Lemme 79. *Soit $P \in \mathbb{K}[X]$; alors $P(a) = 0$ si et seulement si $X - a$ divise $P(X)$.*

Preuve : Le résultat découle immédiatement de la division euclidienne $P = (X - a)Q + P(a)$.
Remarque: ainsi un polynôme irréductible n'a pas de racines. La réciproque est bien entendue fausse en général ; il faut en fait regarder les racines de P dans toutes les extensions de degré $\leq \deg P$, cf. la proposition **??**.

Définition 80. *La multiplicité dans P de $a \in \mathbb{K}$ est le plus grand entier $r \geq 0$ tel que $(X - a)^r$ divise P.*

Remarque: la multiplicité est non nulle si et seulement si a est une racine de P ; si $r = 1$ on dit que a est *une racine simple* et sinon *une racine multiple.*

Proposition 81. *La multiplicité de la racine a de P est l'entier $r \geq 1$ tel que $P(a) = P'(a) = P^{(r-1)}(a) = 0$ et $P^{(r)}(a) \neq 0$.*

Preuve : Il suffit d'appliquer la formule de Taylor en a.
Remarque: en appliquant le lemme de Gauss, si $a_1, \cdots, a_n$ sont des racines de P de multiplicité $r_1, \cdots, r_n$ alors P est divisible par $\prod_{i=1}^{n}(X - a_i)^{r_i}$. En particulier on en déduit qu'un polynôme possède au plus $\deg P$ racines comptées avec multiplicités.
Remarque: un polynôme ne possède que des racines simples si et seulement si P et P' sont premiers entre eux.

Définition 82. *Un polynôme P pouvant s'écrire sous la forme $\lambda \prod_{i=1}^{n}(X - a_i)^{r_i}$ avec $\lambda \in \mathbb{K}$ est dit* totalement décomposé.

Remarque: un corps $\mathbb{K}$ dans lequel tous les polynômes sont totalement décomposé est dit *algébriquement clos.* Le théorème de d'Alembert-Gauss affirme que $\mathbb{C}$ est algébriquement clos : c'est un résultat d'analyse qui repose de manière essentielle sur le théorème des valeurs intermédaires.

2.5 Polynômes symétriques

Définition 83. *Pour $a_1, \cdots, a_n \in \mathbb{K}$ on définit pour $1 \leq k \leq n$:*

$$\sigma_k(a_1, \cdots, a_n) := \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} a_{i_1} a_{i_2} \cdots a_{i_k}.$$

Remarque: ainsi pour $k = 1$ (resp. $k = n$) on obtient $a_1 + \cdots + a_n$ (resp. $a_1 \cdots a_n$).

Proposition 84. *Pour $a_1, \cdots, a_n \in \mathbb{K}$, on a l'égalité*

$$\prod_{i=1}^{n}(X - a_i) = X^n - \sigma_1(a_1, \cdots, a_n)X^{n-1} + \sigma_2(a_1, \cdots, a_n)X^{n-2} + \cdots + (-1)^n\sigma_n(a_1, \cdots, a_n).$$

Remarque: ainsi pour $n = 2$ on trouve $(X-a)(X-b) = X^2 - (a+b)X + ab$.

Définition 85. *Pour $\underline{i} = (i_1 < \cdots < i_r)$, on note*

$$X_{\underline{i}} = X_{i_1} \cdots X_{i_r}.$$

Ce monôme est dit de degré $\deg(\underline{i}) = r$ (resp. de poids $\mathrm{wt}(\underline{i}) = i_1 + 2i_2 + \cdots + ri_r$). Le degré (resp. le poids) d'un polynôme est le maximum du degré (resp. du poids) de ses monômes.

Théorème 86. *(i) La sous-algèbre $\mathbb{Z}[X_1, \cdots, X_n]^{\mathfrak{S}_n}$ des polynômes symétriques, est engendrée sur $\mathbb{Z}$ par les polynômes symétriques élémentaires*

$$\sigma_r = \sum_{\underline{i}:\ \deg \underline{i} = r} X_{\underline{i}}.$$

(ii) Les $\sigma_1, \cdots, \sigma_n$ son algébriquement indépendants sur $\mathbb{Z}$ de sorte que tout $P \in K[X_1, \cdots, X_n]^{\mathfrak{S}_n}$ s'écrit de façon unique

$$P(X_1, \cdots, X_n) = Q(\sigma_1, \cdots, \sigma_n).$$

En outre on a $\deg P = \mathrm{wt}Q$ et le degré partiel de P par rapport à n'importe quelle variable est $\mathrm{wt}Q$.

Preuve : (i) Considérons la relation d'ordre suivante sur les monômes $X^{\underline{m}}$:

$$X_1^{m_1} \cdots X_n^{m_n} > X_1^{m'_1} \cdots X_n^{m'_n}$$

si, soit $\deg \underline{m} = m_1 + \cdots m_n > \deg \underline{m}'$, ou en cas d'égalité il existe $1 \leq r < n$ tel que

$$\forall i \leq r :\ m_i = m'_i \text{ et } m_{r+1} > m'_{r+1}.$$

Ainsi le plus grand monôme de σ_i est $X_1 \cdots X_i$ et donc celui de $\sigma_1^{d_1} \cdots \sigma_n^{d_n}$ est

$$X_1^{d_1+d_2+\cdots+d_n} X_2^{d_2+\cdots+d_n} \cdots X_n^{d_n}. \tag{2}$$

Soit alors $X^{\underline{m}}$ le plus grand monôme de P ; comme P est symétrique on a nécessairement $m_1 \geq m_2 \geq \cdots \geq m_n$. Notons c le coefficient correspondant dans P. Alors le monôme le plus grand apparaissant dans

$$P(X_1, \cdots, X_n) - c\sigma_1^{m_1-m_2}\sigma_2^{m_2-m_3} \cdots \sigma_n^{m_n}$$

est strictement plus petit que celui $X^{\underline{m}}$ de P. En répétant ce processus, un nombre fini de fois, on écrit P comme un polynôme en les σ_i.

(ii) Raisonnons par l'absurde : soit $Q \in K[X_1, \cdots, X_n]$ tel que

$$Q(\sigma_1, \cdots, \sigma_n) = 0.$$

Soit alors $X^{\underline{m}}$ le plus grand monôme de Q : il découle de la formule (2) donnant le plus grand monôme de $\sigma_1^{d_1} \cdots \sigma_n^{d_n}$ que

$$X_1^{m_1+\cdots+m_n} X_2^{m_2+\cdots+m_n} \cdots X_n^{m_n}$$

est le plus grand monôme de $Q(\sigma_1, \cdots, \sigma_n)$ qui ne peut donc pas être le polynôme nul. L'égalité $\deg P = \mathrm{wt}Q$ découle du fait que σ_i est homogène de degré i, quant au degré de Q, d'après (2, il se lit sur le degré partiel.
Remarque: en ce qui concerne les fractions rationnelles, on a encore

$$\mathbb{K}(X_1, \cdots, X_n)^{\mathfrak{S}_n} = \mathbb{K}(\sigma_1, \cdots, \sigma_n).$$

En effet si $f = g/h$ est une fraction rationnelle invariante alors $(\prod_{\sigma \in \mathfrak{S}_n} \sigma h) f$ est clairement symétrique et appartient donc à $\mathbb{K}[\sigma_1, \cdots, \sigma_n]$ et donc $f = \frac{(\prod_{\sigma \in \mathfrak{S}_n} \sigma h) f}{\prod_{\sigma \in \mathfrak{S}_n} \sigma h}$ appartient à $\mathbb{K}(\sigma_1, \cdots, \sigma_n)$. On renvoie le lecteur au §**??** pour une interprétation de ce résultat en théorie de Galois et à son application au problème de Galois inverse.
Remarque: la théorie des invariants a pour but, pour $G \subset GL_n(\mathbb{C})$ agissant sur $R := \mathbb{C}[X_1, \cdots, X_n]$ de décrire R^G. Un théorème de Noether affirme que lorsque G est fini, R^G est engendré par un nombre fini de polynômes homogènes de degré $\leq \sharp G$. Cependant en général R^G n'est pas une algèbre de polynômes. L'exemple le plus simple est $\mathbb{Z}/2\mathbb{Z} =< g >$ agissant sur $\mathbb{C}[X, Y]$ par $g(X) = -X$ et $g(Y) = -Y$. On a alors

$$\mathbb{C}[X, Y]^{<g>} \simeq \mathbb{C}[X^2, Y^2, XY] \simeq \mathbb{C}[S, T, U]/(ST - U^2).$$

L'égalité $X^2.Y^2 = (XY)^2$ peut être vue comme deux décompositions distinctes en produit d'irréductibles dans $\mathbb{C}[X, Y]^{<g>}$: ainsi $\mathbb{C}[X, Y]^{<g>}$ n'est même pas factoriel. En fait on peut montrer que R^G est une algèbre de polynômes si et seulement si G est engendré par des pseudo-réflexions, i.e. des matrices M diagonalisables possédant exactement une valeur propre distincte de 1.

2.6 Résultant et discriminant

Soient

$$\begin{aligned} P(X) &= a_0 + a_1 X + \cdots + a_p X^p, \quad a_p \neq 0 \\ Q(X) &= b_0 + b_1 X + \cdots + b_q X^q, \quad b_q \neq = 0 \end{aligned}$$

deux polynômes à coefficients dans un anneau A.

Définition 87. *On appelle* matrice de Sylvester *de P et Q la matrice suivante :*

$$S(P,Q) = \begin{pmatrix} a_p & \dots & \dots & & a_0 & & \\ & \ddots & & & & \ddots & \\ & & a_p & \dots & \dots & & a_0 \\ b_q & \dots & & b_0 & & & \\ & \ddots & & & \ddots & & \\ & & \ddots & & & & \\ & & & b_q & & \dots & b_0 \end{pmatrix} \begin{matrix} -- \\ q \\ -- \\ -- \\ \\ p \\ \\ -- \end{matrix} \begin{matrix} \\ lignes \\ \\ \\ \\ lignes \\ \\ \\ \end{matrix} \tag{3}$$

Le résultant de P et Q, noté $\mathrm{Res}(P,Q)$ *est le déterminant de $S(P,Q)$.*

Remarque: en échangeant les q premières lignes avec les p dernières, on voit que :

$$\mathrm{Res}(P,Q) = (-1)^{pq}\mathrm{Res}(Q,P). \tag{4}$$

Lemme 88. *Supposons P et Q à coefficients dans un corps K.*

1. *Si Q divise P, on a* $\mathrm{Res}(P,Q) = 0$ *;*
2. *si Q ne divise pas P, soient R le reste de la division de P par Q, r le degré de R. Alors*

$$\mathrm{Res}(P,Q) = (-1)^{pq} b_q^{p-r} \mathrm{Res}(Q,R). \tag{5}$$

Preuve : Multiplions la i-ème colonne de la matrice $S(P,Q)$ par X^{p+q-i}. On obtient la matrice $\tilde{S}(P,Q)(X)$ suivante :

$$\begin{pmatrix} a_pX^{p+q-1} & \dots & & & a_0X^{q-1} & 0 & \dots \\ & \ddots & & & & \ddots & \\ 0 & \dots & a_pX^p & \dots & & & a_0 \\ b_qX^{p+q-1} & \dots & & b_0X^{p-1} & 0 & \dots & 0 \\ 0 & \ddots & & & \ddots & & \\ \vdots & & \ddots & & & & \\ 0 & \dots & & b_qX^q & \dots & & b_0 \end{pmatrix} \begin{matrix} -- \\ q \\ -- \\ -- \\ \\ p \\ \\ -- \end{matrix} \begin{matrix} \\ lignes \\ \\ \\ \\ lignes \\ \\ \\ \end{matrix} \tag{6}$$

telle que $\tilde{S}(P,Q)(1) = S(P,Q)$. Remarquons que dans la matrice $\tilde{S}(P,Q)(X)$, la ligne l_i est formée des monômes du polynôme $X^{q-i}P(X)$ pour $1 \le i \le q$, et des monômes du polynôme $X^{p+q-i}Q(X)$ pour $q+1 \le i \le p+q$.
Montrons maintenant (5). Si $q > p$, on a $R = P$, et le lemme est vrai par la formule (4).
Si $p \ge q$, considérons la division euclidienne :

$$P = QA + R, \quad \deg(R) < \deg(Q) \text{ ou } R = 0. \tag{7}$$

Posons :

$$A(X) = \alpha_0 + \alpha_1 X + \cdots + \alpha_{p-q} X^{p-q};$$

on a donc :

$$QA = \alpha_0 Q + \alpha_1 (XQ) + \cdots + \alpha_{p-q}(X^{p-q}Q). \tag{8}$$

- Si Q divise P, on voit ainsi en utilisant (8) que la relation $P = QA$ s'interprète en disant que la ligne l_q de la matrice $\tilde{S}(P,Q)(X)$ est une combinaison linéaire des lignes $l_{p+q}, l_{p+q-1}, \ldots, l_{2q}$ avec coefficients $\alpha_0, \ldots, \alpha_{p-q}$. Le déterminant de la matrice $\tilde{S}(P,Q)(X)$ est donc nul, ce qui implique que $R(P,Q) = 0$.

- Dans le cas général, posons

$$R(X) = c_0 + c_1 X + \cdots + c_r X^r$$

avec $c_r \neq 0$. On voit alors en utilisant (8) que la relation $P = QA + R$ s'interprète en disant que la ligne l_q de la matrice $\tilde{S}(P,Q)(X)$ est la somme de la ligne $(0, \ldots, 0, c_r X^r, \ldots, c_0)$ correspondant au polynôme $R(X)$, et d'une combinaison linéaire des lignes $l_{p+q}, l_{p+q-1}, \ldots, l_{2q}$ avec coefficients $\alpha_0, \ldots, \alpha_{p-q}$.
On peut donc remplacer la ligne l_q de $\tilde{S}(P,Q)(X)$ par la ligne $(0, \ldots, 0, c_r X^r, \ldots, c_0)$ sans changer son déterminant.
En procédant de même avec les relations

$$X^i P = X^i QA + X^i R$$

pour $0 \leq i \leq q-1$, on voit que l'on peut remplacer les q premières lignes de $\tilde{S}(P,Q)(X)$ par les lignes formées de zéros et des monômes des polynômes $X^i R$, $0 \leq i \leq q-1$, la ligne l_{q-i} étant remplacée par la ligne $(0, \ldots, 0, c_r X^{r+i}, \ldots, c_0 X^i, 0, \ldots, 0)$, ceci sans changer le déterminant.
En faisant $X = 1$ on voit alors que le déterminant de $S(P,Q)$ est égal au déterminant de la matrice :

$$\left(\begin{array}{ccccccc}
0 & \ldots & c_r & \ldots & c_0 & 0 & \ldots \\
 & & & \ddots & & \ddots & \\
0 & \ldots & & & c_r & \ldots & c_0 \\
b_q & \ldots & & b_0 & 0 & \ldots & \\
0 & \ddots & & & \ddots & & \\
\vdots & & \ddots & & & & \\
0 & \ldots & & b_q & \ldots & & b_0
\end{array}\right)
\begin{array}{ll}
-- & \\
q & lignes \\
-- & \\
-- & \\
 & \\
p & lignes \\
 & \\
-- &
\end{array}$$

d'où la relation (5).
Remarque: on peut ainsi calculer le résultant en utilisant l'algorithme d'Euclide.

Corollaire 89. *Soit K un corps. Avec les notations ci-dessus, les conditions suivantes sont équivalentes :*

(1) $\mathrm{Res}(P,Q) = 0$ *;*

(2) les polynômes P et Q ont un facteur commun de degré > 0 dans $K[X]$.

Preuve : (1) $\Rightarrow$ (2) : supposons que (2) soit faux, i.e. que P et Q n'aient pas de facteur commun dans $K[X]$. Le PGCD de P et Q est alors une constante $c \neq 0$. Le lemme précédent appliqué récursivement donne

$$\mathrm{Res}(P,Q) = \alpha \mathrm{Res}(R_s, c)$$

avec $\alpha \neq 0$, $c \neq 0$ et R_s un reste de degré $r_s > 0$. Mais alors $\mathrm{Res}(R_s, c) = c^{r_s} \neq 0$, et donc que $\mathrm{Res}(P,Q) \neq 0$.

- (2) $\Rightarrow$ (1) Si P et Q ont un facteur commun non trivial A dans $K[X]$, supposons d'abord que $P = QA$ avec A de degré $p - q > 0$. Alors le lemme précédent montre que $\mathrm{Res}(P,Q) = 0$. Dans le cas général, on se retrouve dans la situation ci-dessus en considérant le dernier reste non nul dans l'algorithme d'Euclide.

Proposition 90. *Supposons que dans $K[X]$, on ait :*

$$P = a_p(X - \alpha_1)\dots(X - \alpha_p)$$
$$Q = b_q(X - \beta_1)\dots(X - \beta_q).$$

Alors

$$\mathrm{Res}(P,Q) = a_p^q b_q^p \prod_{i,j}(\alpha_i - \beta_j) = a_p^q \prod_{1\leq i\leq p} Q(\alpha_i) = (-1)^{pq} b_q^p \prod_{1\leq j\leq q} P(\beta_j) \quad (9)$$

Preuve : Les égalités :

$$a_p^q b_q^p \prod(\alpha_i - \beta_j) = a_p^q \prod Q(\alpha_i) = (-1)^{pq} b_q^p \prod P(\beta_j)$$

sont immédiates.
Posons $R_2(P,Q) = a_p^q \prod Q(\alpha_i) = (-1)^{pq} b_q^p \prod P(\beta_j)$. Pour montrer que $R_2(P,Q) = \mathrm{Res}(P,Q)$, il suffit de montrer que R_2 satisfait à la même relation de récurrence (5) que $\mathrm{Res}(P,Q)$. On peut supposer $p \geq q > 0$ (car on a évidemment $\mathrm{Res}(P,Q) = R_2(P,Q) = b_q^p$ si $q = 0$). Si $P = QA + R$, on a $P(\beta_j) = R(\beta_j)$ pour toute racine β_j de Q, et donc :

$$R_2(P,Q) = (-1)^{pq} b_q^p \prod P(\beta_j) = (-1)^{pq} b_q^p \prod R(\beta_j) = (-1)^{pq} b_q^{p-r} R_2(Q,R),$$

ce qui est bien la même relation que (5).

Définition 91. *Soit A un anneau intègre et $P = a_p X^p + \dots + a_0 \in A[X]$ tel que $a_p \neq 0$. Alors on définit* le discriminant $D(P)$ *par la formule :*

$$D(P) = \frac{(-1)^{p(p-1)/2}}{a_p} R(P, P').$$

Remarque: cette définition a bien un sens quel que soit l'anneau intègre A, car dans la matrice de Sylvester $\mathrm{Res}(P, P')$, la première colonne est divisible par a_p, puisque $P' = pa_pX^{p-1} + \cdots + a_1$.

Proposition 92. *Si* $P(X) = a_p(X - \alpha_1)\dots(X - \alpha_p)$, *alors :*

$$D(P) = (-1)^{p(p-1)/2} a_p^{2p-2} \prod_{i \neq j} (\alpha_i - \alpha_j) = a_p^{2p-2} \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

Preuve : On a

$$P'(X) = a_p \sum_{i=1}^{p} (X - \alpha_1)\dots\widehat{(X - \alpha_i)}\dots(X - \alpha_p),$$

la notation $\widehat{(X - \alpha_i)}$ signifiant que l'on omet le terme $(X - \alpha_i)$ dans le produit. Or on a $\mathrm{Res}(P, P') = (-1)^{p(p-1)/2} a_p^{p-1} \prod_{i=1}^{p} P'(\alpha_i)$ et le résultat découle de l'égalité $P'(\alpha_i) = a_p(\alpha_i - \alpha_1)\dots(\alpha_i - \alpha_p)$, où le terme $(\alpha_i - \alpha_i))$ n'apparait pas.
Exemples :

1. Si $P = aX^2 + bX + c$, $P' = 2aX + b$, on a :

$$R(P, P') = (-1) \begin{pmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{pmatrix},$$

   d'où $D(P) = b^2 - 4ac$.
2. $P = X^3 + pX + q$, $P' = 3X^2 + p$, un petit calcul de déterminant montre facilement que $D(P) = -4p^3 - 27q^2$.

Remarque: les résultants permettent d'éliminer des variables dans des systèmes d'équations algébriques. Par exemple, notons C la courbe algébrique paramétrée par

$$\begin{cases} x = \frac{2t}{1+t^2} \\ y = \frac{1-t^2}{1+t^2} \end{cases}$$

L'ensemble des points de $C \subset \mathbb{C}^2$ de coordonnées (x, y) sont ceux pour lesquels il existe $t \in \mathbb{C} - \{\pm i\}$ solution commune des deux équations

$$\begin{cases} (1 + t^2)x = 2t \\ (1 + t^2)y = 1 - t^2 \end{cases}$$

Comme pour $t = \pm i$, il n'y a pas de solutions, on peut enlever la restriction précédente. Or d'après ce que l'on a vu, pour $P_{x,y} := xT^2 - 2T + x$ et $Q_{x,y} = (y + 1)T^2 + y - 1$, l'annulation de $\mathrm{Res}(P_{x,y}, Q_{x,y})$ équivaut soit à $(x, y) = (0, -1)$ ou bien à $P_{x,y}$ et $Q_{x,y}$ ont une racine commune, i.e. $(x, y) \in C$. Le calcul du résultant en question donne $R = 4(x^2 + y^2 - 1)$ ce qui confirme que C est le cercle unité privé du point $(0, -1)$.

Proposition 93. *(cf. [?] 5.34) Soit* $P(X) = a_0 + a_1X + \cdots + a_dX^d \in \mathbb{Z}[X]$ *avec* $a_d \neq 0$ *et notons* α_i *pour* $i = 1, \cdots, n$ *ses racines. On a alors*

$$\mathrm{sep}P = \inf_{\alpha_i \neq \alpha_j} |\alpha_i - \alpha_j| \geq (2C)^{1-\frac{d(d-1)}{2}}$$

où $C = |a_d| + \sum_{1 \leq i \leq d-1} |a_i|$.

Preuve : (a) Supposons d'abord que les racines de P sont simples. On peut supposer, quitte à changer les indices, que $\mathrm{sep}P = |\alpha_1 - \alpha_2|$. Comme $P(X) \in \mathbb{Z}[X]$, son discriminant $D(P) \in \mathbb{Z}$ de sorte qu'étant non nul on a $1 \leq |D(P)|$ et

$$1 \leq |a_d|^{2d-2} \prod_{i<j} (\alpha_i - \alpha_j)^2 \text{ soit } \frac{1}{(\alpha_1 - \alpha_2)^2} \leq |a_d|^{2d-2} \prod_{\substack{i<j \\ (i,j)\neq(1,2)}} |\alpha_i - \alpha_j|^2.$$

Or on a $|\alpha_i - \alpha_j| \leq |\alpha_i| + |\alpha_j| \leq 2\frac{C}{|a_d|}$, et il y a $\frac{d(d-1)}{2} - 1 = \frac{d^2-d-2}{2}$ facteurs $|\alpha_i - \alpha_j|^2$, ce qui donne :

$$\frac{1}{|\alpha_1 - \alpha_2|^2} \leq \frac{(2C)^{d^2-d-2}}{|a_d|^{d^2-3d}} \leq (2C)^{d^2-d-2}$$

(car $|a_d| \geq 1$ et $d^2 - 3d \geq 0$), d'où le résultat.

(b) Dans le cas général lorsque les racines de $P \in \mathbb{Z}[X]$ ne sont pas nécessairement simples, on peut supposer P primitif quitte à le diviser par son contenu (qui est un entier). On considère le polynôme $R = P \wedge P'$ que l'on peut supposer dans $\mathbb{Z}[X]$ et primitif ; on a alors $P = QR$ dans $\mathbb{Z}[X]$, P et R étant primitifs. On peut alors appliquer la méthode de (a) au polynôme Q qui a les mêmes racines que P, mais avec multiplicité 1. On trouve donc, en notant d' le degré de Q, et en utilisant que $d' \leq d$:

$$\frac{1}{(\mathrm{sep}P)^2} = \frac{1}{(\mathrm{sep}Q)^2} \leq (2C)^{d'^2-d'-2} \leq (2C)^{d^2-d-2}.$$

2.7 Polynômes cyclotomiques

Soit $m \in \mathbb{N}^*$. On note $\mathbb{U}_m = \{z \in \mathbb{C} \mid z^m = 1\}$ l'ensemble des *racines m-ièmes de l'unité* dans $\mathbb{C}$. On rappelle que $\mathbb{U}_m$ est un groupe cyclique, et on appelle *racine primitive m-ième de l'unité* tout générateur de $\mathbb{U}_m$. On note $\mathbb{U}'_m$ l'ensemble des racines primitives m-ième de l'unité.

Définition 94. *On appelle* m-ième polynôme cyclotomique *le polynôme*

$$\Phi_m = \prod_{z \in \mathbb{U}'_m} (X - z).$$

Proposition 95. *(i) Le polynôme* Φ_m *est unitaire de degré* $\varphi(m)$.

(ii) $X^m - 1 = \prod_{d|m} \Phi_d$.
(iii) Φ_m *est à coefficients dans* $\mathbb{Z}$.

Preuve : Les points (i) et (ii) découlent directement des propriétés de structure de $\mathbb{Z}/n\mathbb{Z}$. Montrons le point (iii) par récurrence :

- le résultat est immédiat pour $m = 1$ puisque $\Phi_1 = X - 1$.
- en supposant le résultat vrai pour tous les entiers inférieurs à m, on obtient que $U = \prod_{d|m, d\neq m} \Phi_d$ est un polynôme unitaire à coefficients dans $\mathbb{Z}$. On peut donc effectuer dans $\mathbb{Z}[X]$ la division euclidienne de $X^m - 1$ par U : il existe $Q, R \in \mathbb{Z}[X]$ tels que $X^m - 1 = UQ + R$ et $deg(R) < deg(U)$. L'unicité de la division euclidienne dans $\mathbb{C}[X]$ permet de conclure que $Q = \Phi_m$ et $R = 0$, ce qui implique que Φ_m est bien à coefficients entiers.

Théorème 96. *Le polynôme* Φ_m *est irréductible dans* $\mathbb{Q}[X]$.

Preuve : cf. la proposition **??**.

3 Espaces vectoriels

Dans ce qui suit $\mathbb{K}$ est un corps que l'on pourra supposer dans un premier temps égal à $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$.

3.1 Généralités

Définition 97. *Un* $\mathbb{K}$*-espace vectoriel est un triplet* $(E, +, .)$ *où*

- $(E, +)$ *est un groupe commutatif,*
- *muni d'une loi externe* $(\lambda, e) \in \mathbb{K} \times E \mapsto \lambda.e \in E$ *qui vérifie les propriétés suivantes :*
  - *pour tout* $\lambda, \mu \in \mathbb{K}$ *et pour tout* $e \in E$*, on a* $(\lambda + \mu).e = \lambda.e + \mu.e$ *;*
  - *pour tout* $\lambda \in \mathbb{K}$ *et* $e, f \in E$*, on a* $\lambda.(e + f) = \lambda.e + \lambda.f$ *;*
  - *pour tout* $\lambda, \mu \in \mathbb{K}$ *et* $e \in E$*, on a* $(\lambda\mu).e = \lambda.(\mu.e)$ *;*
  - *pour tout* $e \in E$ *on a* $1.e = e$.

Remarque: ces définitions prennent aussi sens dans le cas où $\mathbb{K}$ est simplement un anneau unitaire A, on parle alors de A-module, cf. le §6.
Exemples :

- $\mathbb{K}$ et plus généralement $\mathbb{K}^n$, $\mathbb{K}^{\mathbb{N}}$ ou $\mathbb{K}^{(\mathbb{N})}$;
- $\mathbb{M}_{m,n}(\mathbb{K})$ et $\mathbb{K}[X]$;
- les fonctions de X dans $\mathbb{K}$ où X est un ensemble quelconque ;
- le produit quelconque d'une famille d'espaces vectoriels est un espace vectoriel.

Définition 98. *Soit* E *un* $\mathbb{K}$*-espace vectoriel ; un sous-ensemble* $F \subset E$ *est* un sous-espace vectoriel *si et seulement si c'est un sous-groupe stable par la loi externe, i.e. si et seulement si* F *est non vide et pour tout* $f_1, f_2 \in F$ *et pour tout* $\lambda \in \mathbb{K}$*, on a :* $f_1 + \lambda f_2 \in F$.

Exemples :

- $\mathbb{K}_n[X] \subset \mathbb{K}[X]$ le sous-ensemble des polynômes de degré $\leq n$;
- l'ensemble des suites convergentes de $\mathbb{K}^{\mathbb{N}}$;
- $\mathbb{R} \subset \mathbb{C}$ est un sous-$\mathbb{R}$-espace vectoriel mais n'est pas un sous-$\mathbb{C}$-espace vectoriel.

Remarque: comme précédemment un sous-espace vectoriel est un espace vectoriel et habituellement on se sert de cette remarque pour tester si on est en présence d'un espace vectoriel.
Remarque: l'intersection quelconque d'une famille de sous-espaces vectoriels est un espace vectoriel ce qui permet de définir le sous-espace vectoriel engendré par un sous-ensemble $A \subset E$ que l'on note $< A >$.
Remarque: si $\mathbb{K}$ est un corps infini, toute réunion finie de sous-espaces vectoriels est un sous-espace vectoriel si et seulement s'ils sont tous contenus dans un seul. En particulier une réunion finie d'hyperplans distincts n'est pas un sous-espace vectoriel.
Exemple fondamental : soit $(e_i)_{i\in I}$ une famille *quelconque* d'éléments de E alors $< \{e_i : \ i \in I\} >$ est l'ensemble *des combinaisons linéaires* $\sum_{i\in I} \lambda_i e_i$ *à support fini.*

Définition 99. *Soient F, G des sous-espaces vectoriels d'un espace vectoriel E.* La somme $F+G$ *est le sous-espace $< F \cup G >$ engendré par F et G. On dit que F et G sont* en somme directe *et on écrit $F \oplus G$ si $F \cap G = \{0\}$.*

Remarque: on vérifie aisément que $F + G = \{f + g : \ f \in F \text{ et } g \in G\}$; en outre F et G sont en somme directe si et seulement si l'écriture d'un élément $e \in F + G$ sous-la forme $f + g$ est unique.

Définition 100. *On dit que F et G sont* supplémentaires *si $E = F \oplus G$, i.e. si la somme $F + G$ est tout l'espace et qu'ils sont en somme directe.*

Remarque: on veillera bien à ne pas confondre *supplémentaires* et *complémentaires* ; rappelons que le complémentaire d'un sous-espace vectoriel n'est jamais un sous-espace puisqu'il ne contient pas le vecteur nul !
Exercice : montrer que des sous-espaces $E_1, \cdots, E_n$ sont en somme directe si et seulement si pour tout $i = 1, \cdots, n$

$$E_i \cap \Big(\sum_{1\leq k\neq i\leq n} E_i \Big) = \{0\}.$$

3.2 Théorie de la dimension

Définition 101. *Une famille $\{(e_i)_{i\in I}\}$ de vecteurs d'un espace vectoriel E est dite* libre *si pour tout famille $(\lambda_i)_{i\in I} \in \mathbb{K}^I$*

$$\sum_{i\in I} \lambda_i e_i = 0 \Rightarrow \forall i \in I, \ \lambda_i = 0.$$

Elle est dite génératrice *si* $< \{e_i : \ i \in I\} >= E$*, i.e. si tout vecteur de* E *peut s'écrire comme une combinaison linéaire à support fini des* e_i*.*

Remarque: la famille $(X^i)_{i\in\mathbb{N}} \in \mathbb{K}[X]$ est libre et génératrice.
Remarque: la famille $(e_i)_{i\in I}$ est dite *liée* si elle n'est pas libre, i.e. s'il existe une famille $(\lambda_i)_{i\in I} \in \mathbb{K}^{(I)}$ non nulle telle que $\sum_{i\in I} \lambda_i e_i = 0$.

Définition 102. *Une famille* $(e_i)_{i\in I}$ *de vecteurs de* E *est* une base *si elle est libre et génératrice.*

Théorème 103. *dit de la base incomplète.*
Soient $\{f_1, \cdots, f_p\}$ *une famille libre de vecteurs et* $\{g_1, \cdots, g_q\}$ *une famille génératrice de* E*. Il existe alors un entier* $n \geq p$ *et une base* $\{e_1, \cdots, e_n\}$ *de* E *telle que* $e_i = f_i$ *pour* $1 \leq i \leq p$ *et* $e_j \in \{g_1, \cdots, g_q\}$ *pour* $p+1 \leq j \leq n$*.*

Preuve : Considérons une famille de cardinal maximal de la forme

$$(f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r})$$

qui soit libre. Montrons alors qu'elle est aussi génératrice en vérifiant que pour tout $1 \leq j \leq q$, le vecteur g_j appartient à l'espace vectoriel engendré par cette famille. Si j est l'un des i_k pour $1 \leq k \leq r$, c'est clair sinon, comme par maximalité, la famille $(f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r}, g_j)$ est liée, il existe une famille $(\lambda_i)_{1\leq i\leq p+r+1}$ non nulle telle que

$$\lambda_1 f_1 + \cdots + \lambda_p f_p + \lambda_{p+1} g_{i_1} + \cdots + \lambda_{p+r} g_{i_r} + \lambda_{p+r+1} g_j = 0.$$

Comme la famille $(f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r})$ est libre, nécessairement $\lambda_{p+r+1} \neq 0$ et donc $g_j \in \langle f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r}\rangle$. Ainsi donc $(f_1, \cdots, f_p, g_{i_1}, \cdots, g_{i_r})$ est à la fois libre et génératrice, c'est donc une base.
Remarque: ainsi tout espace contenant une famille génératrice finie admet une base. Autrement dit tout espace vectoriel de type fini admet des bases.

Lemme 104. *Dans un espace possédant une famille génératrice de cardinal* n*, toute famille de vecteurs non nuls de cardinal* $\geq n+1$ *est nécessairement liée.*

Preuve : On raisonne par récurrence sur n. Pour $n = 1$ et v une base, pour f, g deux vecteurs non nuls, il existe λ, μ non nuls tels que $f = \lambda v$ et $g = \mu v$, de sorte $\mu f - \lambda g = 0$ et donc (f, g) est liée.

Supposons donc le résultat acquis jusqu'au rang $n-1$ et considérons une famille de $n+1$ vecteurs non nuls $(f_1, \cdots, f_{n+1})$ d'un espace vectoriel admettant $(g_1, \cdots, g_n)$ comme famille génératrice. Pour tout $i = 1, \cdots, n+1$, on écrit alors

$$f_i = \lambda_{i,1} g_1 + \cdots + \lambda_{i,n} g_n.$$

Quitte à modifier l'ordre des g_i, supposons $\lambda_{n+1,n} \neq 0$ et on définit pour tout $i = 1, \cdots, n$

$$\tilde{f}_i = \lambda_{n+1,n} f_i - \lambda_{i,n} f_{n+1} \in \langle g_1, \cdots, g_{n-1}\rangle.$$

D'après l'hypothèse de récurrence la famille $(\tilde{f}_1, \cdots, \tilde{f}_n)$ est liée, i.e. il existe une famille non nulle $(\mu_1, \cdots, \mu_n)$ telle que

$$\mu_1 \tilde{f}_1 + \cdots + \mu_n \tilde{f}_n = 0$$

ce qui fournit la relation

$$\mu_1 \lambda_{n+1,n} f_1 + \cdots + \mu_n \lambda_{n+1,n} f_n - (\sum_{i=1}^{n} \mu_i \lambda_{i,n}) f_{n+1} = 0$$

prouvant, comme $(\mu_1 \lambda_{n+1,n}, \cdots, \mu_n \lambda_{n+1,n}, -\sum_{i=1}^{n} \mu_i \lambda_{i,n})$ n'est pas nulle, que la famille $(f_1, \cdots, f_{n+1})$ est liée.

Corollaire 105. *Soit E un espace vectoriel muni d'une base de cardinal n. Alors toute les bases de E ont pour cardinal n.*

Preuve : Soient $(e_1, \cdots, e_n)$ et $(f_1, \cdots, f_m)$ deux bases de E. En appliquant le lemme précédent à la famille génératrice $(e_1, \cdots, e_n)$ (resp. $(f_1, \cdots, f_m)$) et à la famille libre $(f_1, \cdots, f_m)$ (resp. $(e_1, \cdots, e_n)$), on en déduit $m \leq n$ (resp. $n \leq m$) et donc finalement $n = m$.

Définition 106. *Le cardinal d'une base (et donc de toute base) d'un espace vectoriel E est appelé* sa dimension *; elle est finie ou infinie.*

Corollaire 107. *Tout sous-espace vectoriel F de E est de dimension inférieure ou égale à celle de E avec égalité si et seulement si $F = E$.*

Preuve : Il suffit de remarquer qu'un base de F est une famille libre de E et d'appliquer le corollaire précédent.

Définition 108. *On appelle hyperplan d'un espace vectoriel E de dimension finie, tout sous-espace de dimension $n - 1$.*

Remarque: en dimension infinie, un hyperplan est un sous-espace tel que E/F est de dimension 1. La dimension de l'espace quotient E/F s'appelle *la codimension* de F dans E.
Remarque: la dimension de $E \times F$ est le somme des dimensions de E et F. L'espace vectoriel des applications linéaires $\mathcal{L}(E, F)$ de E vers F est de dimension $\dim E . \dim F$.
Remarque: toute famille libre est de cardinal $\leq n$ avec égalité si et seulement si c'est une base.
Remarque: la dimension de $F+G$ est inférieure ou égale à $\dim F + \dim G$ avec égalité si et seulement si F et G sont en somme directe. Plus précisément on a la formule du rang.

Théorème 109. *Soient F, G deux sous-espace de E alors*

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

Preuve : Soit $(e_1, \cdots, e_r)$ une base de $F \cap G$ que l'on complète en une base $(e_1, \cdots, e_r, f_1, \cdots, f_p)$ (resp. $(e_1, \cdots, e_r, g_1, \cdots, g_q)$) de F (resp. de G). On vérifie alors aisément que $(e_1, \cdots, e_r, f_1, \cdots, f_p, g_1, \cdots, g_q)$ est une base de $F + G$ ce qui donne la formule de l'énoncé.

Finissons ce paragraphe par un court mot sur la dimension infinie.

Proposition 110. *Soit E un $\mathbb{K}$-espace vectoriel et V, W_1, W_2 des sous-espaces tels que $V \cap W_1 = \{0\}$ et $V + W_2 = E$. Il existe alors un supplémentaire W de V contenu dans W_2 et contenant W_1.*

Preuve : Considérons l'ensemble $\mathcal{E}$ des sous-espaces de E contenant W_1 et contenus dans W_2 ; $\mathcal{E}$ n'est pas vide car $W_1 \in \mathcal{E}$. En outre $\mathcal{E}$ est partiellement ordonné par la relation d'inclusion et est inductif. Rappelons que cela signifie que toute chaine totalement ordonnée admet un majorant : ici pour une telle chaine, un majorant est simplement donné par la réunion qui est clairement un sous-espace.

D'après le lemme de Zorn, $\mathcal{E}$ admet un élément maximal, notons le W. Par définition on a donc $W \cap V = \{0\}$ et $W_1 \subset W \subset W_2$. Il reste alors à prouver que $V + W = E$; tout élément $x \in E$ s'écrit $x = v + w_2$ avec $v \in V$ et $w_2 \in W_2$. Si $w_2 \in W$ alors c'est gagné, sinon on considère le sous-espace engendré X par W et w_2. Par maximalité de W, $X \notin \mathcal{E}$ de sorte qu'il existe $0 \neq y \in X \cap V$; ainsi $y = w + \lambda w_2 \in V$ et donc $y \in W \cap V$ ce qui n'est pas.
Remarque: le lecteur notera bien l'utilisation essentielle du lemme de Zorn qui, rappelons le, est équivalent à l'axiome du choix. Ainsi notre preuve n'est pas du tout constructive.

Corollaire 111. *Tout sous-espace V de E admet un supplémentaire.*

Corollaire 112. *Tout espace vectoriel non nul admet une base.*

Preuve : Considérons l'ensemble $\mathcal{A}$ des familles libres de E ; c'est clairement un ensemble non vide, partiellement ordonné par l'inclusion et inductif. D'après le lemme de Zorn, il possède un élément maximal qui est donc une famille libre maximal c'est donc nécessairement une famille génératrice et donc une base.
Remarque: le lecteur pourra s'exercer sur $\mathbb{K}^{\mathbb{N}}$ en vérifiant que toute base est nécessairement non dénombrable.

Corollaire 113. ***(Théorème de la base incomplète)***
Soit $(e_i)_{i \in I}$ une partie génératrice de E. Soit $J \subset I$ tel que $(e_i)_{i \in J}$ est libre, il existe alors $J \subset K \subset I$ tel que $(e_i)_{i \in K}$ soit une base.

Preuve : On considère l'ensemble $\mathcal{A}$ des familles libres $(e_i)_{i \in A}$ pour $A \subset I$. C'est un ensemble non vide partiellement ordonné par l'inclusion et clairement inductif. D'après le lemme de Zorn, $\mathcal{A}$ possède une élément maximal K ; comme précédement $(e_i)_{i \in K}$ est libre et génératrice par maximalité de K.

Remarque: citons enfin le cas des espaces de Hilbert, i.e. des espaces hermitiens, au sens du paragraphe sur l'algèbre bilinéaire, qui sont complets, i.e. toutes les suites de Cauchy sont convergentes.

Définition 114. *On dit que* $(e_i)_{i\in I}$ *est* une base de Hilbert *d'un espace de Hilbert* H *si et seulement si :*

- *c'est une base orthonormée, i.e.* $< e_i, e_j >= \delta_{i,j}$ *;*
- *la famille est complète au sens que pour tout* $x \in H$ *il existe* $(\lambda_i)_{i\in I}$ *telle que* $\sum_{i\in I} \lambda_i e_i = x$*, i.e. la série correspondante dans* H *est convergente de limite* x*.*

Remarque: le lecteur vérifiera aisément qu'une base au sens de Hilbert n'est pas une base au sens classique, cf. par exemple les espaces L^2.

3.3 Application linéaires.

Définition 115. Une application linéaire *ou* un morphisme f *d'un espace vectoriel* E *dans un espace* F *est une application telle que pour tous* $\lambda \in \mathbb{K}$ *et* $x, y \in E$ *on a* $f(x + \lambda y) = f(x) + \lambda f(y)$*.*

Remarque: une application linéaire de E dans E est appelée un endomorphisme. Dans le cas où $F = \mathbb{K}$, on parle de *forme linéaire.*
Remarque: pour toute application linéaire $f : E \to F$ vérifie $f(0) = 0$ et

$$f(\sum_{i=1}^{n} \lambda_i e_i) = \sum_{i=1}^{n} \lambda_i f(e_i).$$

Notation 3. *On note* $\mathcal{L}(E, F)$ *(resp.* $\mathcal{L}(E) = \mathcal{L}(E, E)$*), l'ensemble des morphismes de* E *dans* F *(resp. des endomorphismes de* E*) ; c'est un espace vectoriel de dimension* $\dim E . \dim F$*.*

En ce qui concerne l'existence des applications linéaires, on a le résultat suivant.

Proposition 116. *Soit* $(e_i)_{1\leq i\leq n}$ *une base de* E*. Pour n'importe quel ensemble de* n *vecteurs* $\{f_1, \cdots, f_n\}$ *de* F*, il existe une unique application linéaire telle que pour tout* $i = 1, \cdots, n$*, on ait* $f(e_i) = f_i$*.*

Remarque: ainsi deux applications linéaires sont égales si et seulement si elles coincident sur une base.

Définition 117. *Pour* $f \in \mathcal{L}(E, F)$*, on note* $\operatorname{Ker} f$ *l'ensemble des* $e \in E$ *tels que* $f(e) = 0$*; c'est un sous-espace vectoriel de* E *que l'on appelle* le noyau *de* f*.*

Remarque: l'image de f est aussi un sous-espace de F que l'on note $\operatorname{Im} f$. Plus généralement l'image directe ou réciproque d'un sous-espace est un sous-espace vectoriel.

Proposition 118. *Une application linéaire f est injective si et seulement si* $\operatorname{Ker} f = \{0\}$.

Remarque: f es surjective si et seulement si l'image d'une base de E est une famille génératrice de F. Ainsi f est bijective, et on dit que f est *un isomorphisme*, si l'image d'une base est une base : c'est alors vrai pour toute base.
Remarque: une application linéaire $f : E \to F$ où $\dim E = \dim F$ est injective si et seulement si elle est surjective.

Notation 4. *On note $GL(E)$ l'ensemble des isomorphismes de E, on dit aussi* automorphisme. *C'est un groupe pour la composition.*

Remarque: pour $E = \mathbb{K}^n$ les vecteurs $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ définissent une base dite *canonique*. Tout espace vectoriel muni d'une base $(e_i)_{1 \leq i \leq n}$ de cardinal n est isomorphe à $\mathbb{K}^n$ où $f : \mathbb{K}^n \to E$ est défini par $f(x_1, \cdots, x_n) = \sum_{i=1}^{n} x_i e_i$. En particulier deux espaces vectoriels de même dimension sont toujours isomorphes.

Théorème 119. *Soit $f \in \mathcal{L}(E, F)$ alors*

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

Définition 120. *La dimension de* $\operatorname{Im} f$ *s'appelle* le rang *de f ; on le note* $\operatorname{rg} f$.

3.4 Matrices

Définition 121. Une matrice à coefficients dans $\mathbb{K}$ de taille $m \times n$ *est un tableau* $(a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq i \leq n}}$ *de scalaires $a_{i,j} \in \mathbb{K}$ placés sur la i-ème ligne et la j-ème colonne. On note $\mathbb{M}_{m,n}(\mathbb{K})$ l'ensemble de ces matrices que l'on muni d'une structure d'espace vectoriel en l'identifiant avec $\mathbb{K}^{nm}$, i.e. coefficient par coefficient.*

Remarque: une matrice ligne (resp. colonne) correspond au cas où $n = 1$ (resp. $m = 1$) ; on dit aussi vecteur ligne (resp. colonne). Les lignes (resp. les colonnes) d'une matrice sont appelées ses vecteurs lignes (resp. colonnes).
Remarque: les matrices $E_{i,j}$ dont les coefficients sont tous nuls sauf celui d'indice (i, j) égal à 1, forment une base de $\mathbb{M}_{m,n}(\mathbb{K})$.
Remarque: la matrice $(b_{i,j} = a_{j,i})_{i,j} \in \mathbb{M}_{n,m}(\mathbb{K})$ s'appelle *la matrice transposée*, on la note $B = {}^tA$ si $A = (a_{i,j})_{i,j}$.
Remarque: dans le cas où $m = n$, on parle de matrices carrées et on note $\mathbb{M}_n(\mathbb{K})$ pour $\mathbb{M}_{n,n}(\mathbb{K})$. Les éléments $a_{i,i}$ de $A = (a_{i,j})_{i,j}$ sont dits *diagonaux*. Ainsi une matrice est dite :

- *diagonale* si tous ses coefficients diagonaux sont nuls ; on parle aussi de matrice *antidiagonale* si $a_{i,j} = 0$ sauf pour $i + j = n + 1$.

— *triangulaire* supérieure (resp. inférieure) si tous les $a_{i,j}$ sont nuls pour $i > j$ (resp. $j > i$).
— *tridiagonale* si $a_{i,j} = 0$ pour tout $|j - i| > 1$.

Les matrices ne sont pas de simples tableaux de chiffres mais doivent leur introduction en ce qu'ils permettent :
— d'étudier les systèmes linéaires ;
— de manipuler les endomorphismes des espaces vectoriels.

Ainsi pour $f : E \to F$ un endomorphisme entre deux espaces vectoriels munis des bases respectives $(e_i)_{1\leq i\leq n}$ et $(f_j) =_{1\leq j\leq m}$, on lui associe la matrice $A(f) = (a_{i,j})_{\substack{1\leq i\leq n\\ 1\leq j\leq m}}$ telle que pour tout $1 \leq i \leq n$ on a

$$f(e_i) = \sum_{j=1}^{m} a_{i,j} f_j.$$

Autrement dit les vecteurs colonnes de A sont les $f(e_i)$ exprimés dans la base $(f_j)_j$.
Remarque: d'après ce qui précède, f est déterminé par sa matrice $A(f)$ de sorte que l'on doit pouvoir exprimer l'image $f(x)$ de tout vecteur $x = \sum_{i=1}^{n} x_i e_i$.

Définition 122. *Pour toute matrice $A \in \mathbb{M}_{m,n}(\mathbb{K})$ et tout vecteur colonne $X \in \mathbb{M}_{n,1}(\mathbb{K})$ on définit le vecteur colonne $Y = AX \in \mathbb{M}_{m,1}(\mathbb{K})$ par la formule :*

$$y_j = \sum_{k=1}^{n} a_{j,k} x_k.$$

Pour une matrice $B \in \mathbb{M}_{n,r}$ dont on note $C_1, \cdots, C_r$ les vecteurs colonnes, on définit la matrice $M = AB \in \mathbb{M}_{m,r}(\mathbb{K})$ dont les vecteurs colonnes sont les AC_i pour $i = 1, \cdots, r$.

Proposition 123. *Soit $f : E \to F$ et $A(f)$ sa matrice relativement à des bases $(e_i)_{1\ \leq i\leq n}$ et $(f_j)_{1\leq j\leq m}$ de respectivement E et F. Pour tout $x = \sum_{i=1}^{n} x_i e_i$, on note X le vecteur colonne $(x_{i,1})_{1\leq i\leq n}$. Alors les coordonnées de $f(x)$ dans la base $(f_j)_{1\leq j\leq m}$ sont les coordonnées $(y_{j,1})_{1\leq j\leq m}$ du vecteur colonne $A(f)X$, i.e. $f(x) = \sum_{j=1}^{m} y_j f_j$.*

Corollaire 124. *Pour tout $f : E \to F$ et $g : F \to G$ des endomophismes ; on suppose E, F, G munis de base $(e_i)_{1\leq i\leq n}$, $(f_j)_{1\leq j\leq m}$ et $(g_k)_{1\leq k\leq r}$. On note $A(f), A(g)$ et $A(g \circ f)$ les matrices associées à f, g et $g \circ f$ relativement à ces bases. On a alors*

$$A(g \circ f) = A(g)A(f).$$

En particulier $\mathcal{L}(E)$ étant une algèbre on en déduit le corollaire suivant.

Corollaire 125. *La multiplication des matrices définie plus haut, munit* $\mathbb{M}_n(\mathbb{K})$ *d'une structure d'algèbre.*

Définition 126. *Les matrices de* $\mathbb{M}_n(\mathbb{K})$ *qui s'identifient aux automorphismes de* E *sont dites* inversibles *; l'ensemble de ces matrices inversibles est noté* $GL_n(\mathbb{K})$*.*

Remarque: ainsi une matrice est inversible si et seulement si ses vecteurs colonnes forment une base.

Définition 127. *Étant donné un espace vectoriel* E *muni de deux bases* $(e_i)_{1\leq i\leq n}$ *et* $(e'_i)_{1\leq i\leq n}$*, on appelle matrice de passage de* $(e_i)_i$ *à* $(e'_i)_i$ *et on la note* $P_{e_i\leftarrow e'_i}$*, la matrice de* $\mathbb{M}_n(\mathbb{K})$ *dont la* j*-ème colonne est donnée par les coordonnées de* e'_j *dans la base* $(e_i)_i$*, i.e.* $e'_j = \sum_{i=1}^n p_{i,j}e_i$*.*

Remarque: la matrice $P_{e_i\leftarrow e'_i}$ peut aussi se voir comme la matrice de l'application de l'identité de $E \to E$ où l'espace de départ est muni de la base $(e_i)_i$ et l'espace d'arrivée de la base $(e'_i)_i$. On en déduit alors que :

- $P_{e_i\leftarrow e'_i}$ est inversible, d'inverse $P_{e'_i\leftarrow e_i}$;
- si X' est le vecteur colonne des coordonnées d'un vecteur e de E dans la base $(e'_i)_i$, alors $X = P_{e_i\leftarrow e'_i}X'$ est celui de e dans la base $(e_i)_i$;
- si $A(f)$ est la matrice de $f : E \to F$ muni des bases $(e_i)_i$ et $(f_j)_j$ de respectivement E et F alors, pour des bases $(e'_i)_i$ et $(f'_j)_j$, la matrice $A'(f)$ relativement à ces bases est $P^{-1}_{e_i\leftarrow e'_i}A(f)P_{f_j\leftarrow f'_j}$. Dans le cas particulier où $E = F$ et où $A(f)$ et $A'(f)$ représentent la matrice de f dans respectivement les bases $(e_i)_i$ et $(e'_i)_i$ alors $A'(f) = P^{-1}_{e_i\leftarrow e'_i}A(f)P_{e_i\leftarrow e'_i}$.

Exemples : étant donnée une matrice $A \in \mathbb{M}_{m,n}(\mathbb{K})$, la multiplication à gauche (resp. à droite) par la matrice

- $T_{i,j}(\lambda)$ dont les coefficients diagonaux sont égaux à 1, tous les autres étant nuls sauf $t_{i,j} = \lambda$, correspond à modifier les lignes (resp. les colonnes) de A selon la règle $L_i \to L_i + \lambda L_j$ (resp. $C_i \to C_i + \lambda C_j$) ;
- $D_i(\lambda)$ matrice diagonale dont les coefficients diagonaux sont égaux à 1 sauf $d_{i,i} = \lambda$ correspond à modifier les lignes (resp. les colonnes) de A selon la règle $L_i \to \lambda L_i$ (resp. $C_i \to \lambda C_i$) ;
- $P_{i,j} = I - E_{i,i} - E_{j,j} + E_{i,j} + E_{j,i}$, correspond à modifier les lignes (resp. les colonnes) de A selon la règle $L_i \leftrightarrow L_j$ (resp. $C_i \leftrightarrow C_j$).

Remarque: pour $i \neq j$, les matrices $T_{i,j}(\lambda)$ (resp. $D_i(\lambda)$) sont des matrices dites de transvections (resp. de dilatations) élémentaires relativement à la base canonique. Les matrices $P_{i,j}$ sont des cas particuliers des matrices de permutation. Ces trois types de matrices permettent d'effecteur les opérations élémentaires sur les lignes et les colonnes d'une matrice. Nous reviendrons sur ce point lors de l'étude des systèmes linéaires.

3.5 Rappels sur la dualité

Définition 128. *Étant donné un espace vectoriel E, l'ensemble des formes linéaires sur E est un espace vectoriel noté E^* et dit* le dual *de E.*

Remarque: une base $(e_i)_{1\leq i\leq n}$ de E étant fixée, l'application linéaire $e_i^* \in E^*$ définie par $e_i^*(e_j) = \delta_{i,j}$ est une base de E^* dite la base duale de $(e_i)_i$. On se méfiera de la notation car e_i^* dépend de toute la base $(e_i)_i$ et pas seulement du seul vecteur e_i.

Proposition 129. *Étant donné un sous-espace $F \subset E$, le sous-ensemble $F^\perp \subset E^*$ des formes linéaires s'annulant sur F est un sous-espace de dimension* $\dim E - \dim F$*, i.e. la dimension de $F^\perp$ est égale à la codimension de F.*

Définition 130. *Soit $f \in \mathcal{L}(E,F)$, on lui associe alors son* application adjointe *notée $f^* \in \mathcal{L}(F^*,E^*)$ définie par la formule*

$$y^* \in F_* \mapsto f^*(y^*) = y^* \circ f$$

au sens où $f^*(y^*)$ est la forme linéaire sur E définie par $x \mapsto y^*(f(x))$.

Proposition 131. *Si E, F sont munies de bases respectives $(e_i)_i$ et $(f_j)_j$ alors la matrice de f^* dans les bases duales associées de F^* et E^* est la transposée de la matrice de f dans les bases $(e_i)_i$ et $(f_j)_j$.*

Remarque: on notera en particulier que f et f^* ont le même rang.

Proposition 132. *Soit E un espace vectoriel de dimension finie ; alors le bidual $(E^*)^*$ s'identifie canoniquement à E.*

Remarque: l'application $E \to (E^*)^*$ est donnée par $x \mapsto (f \mapsto f(x))$.

3.6 Systèmes linéaires.

Définition 133. Une équation linéaire *en les variables $x_1, \cdots, x_n$ est une expression de la forme*

$$L(x_1, \cdots, x_n) = b \text{ où } L(x_1, \cdots, x_n) = a_1x_1 + \cdots + a_nx_n$$

dont on cherche les solutions dans $\mathbb{K}^n$. On dit que l'équation est homogène *lorsque $b = 0$.*

Remarque: on peut et on doit interpréter $L(x_1, \cdots, x_n)$ comme une forme linéaire sur $\mathbb{K}^n$ écrite dans la base canonique.

Définition 134. Un système linéaire *de m équations à n variables est une collection de m équations linéaires*

$$(S) = \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = b_n \end{cases}$$

que l'on cherche à résoudre simultanément. Il est dit incompatible *s'il ne possède pas de solutions,* compatible *sinon.*

Remarque: comme suggéré par les notations, on introduit la matrice $A_S = (a_{i,j}) \in \mathbb{M}_{m,n}(\mathbb{K})$ et on écrit le système précédent sous la forme $A_S X = B$ où X (resp. B) est le vecteur colonne de coordonnées les x_i (resp. les b_i).

Définition 135. *Deux systèmes linéaires (S) et (S') sont dit* équivalents *s'ils ont le même ensemble de solutions.*

Proposition 136. *Deux systèmes linéaires (S) et (S') sont équivalents si et seulement s'il existe une matrice inversible $P \in GL_m(\mathbb{K})$ telle que $A_S = PA_{S'}$ et $B = PB'$.*

Remarque: en utilisant que $GL_n(\mathbb{K})$ est engendré par les matrices de transvections et de dilatations (en général, par commodité, on rajoute aussi les matrices de permutation $P_{i,j}$), on doit pouvoir manipuler le système (S) pour arriver au système (S') qui lui est équivalent, encore faut-il que ce processus soit constructif, ce qui est assuré par l'algorithme de Gauss.

Définition 137. *Soit (S) un système linéaire non nécessairement homogène que l'on écrit sous forme matricielle $A_S X = B$. On introduit alors la matrice $\tilde{A}_S$ en rajoutant la colonne B à la matrice A_S.*

Définition 138. *Une matrice $M \in \mathbb{M}_{m,n}(\mathbb{K})$ est dite* échelonnée *si en dessous du premier élément non nul de chaque ligne, il n'y a que des zéros. Elle est dite en outre* échelonnée réduite *si tout premier élément non nul de chaque ligne, appelé* pivot*, est égal à 1, et que chaque pivot est le seul élément non nul de sa colonne.*

Proposition 139. *Pour toute matrice $M \in \mathbb{M}_{m,n}(\mathbb{K})$, il existe une unique matrice $P \in GL_m(\mathbb{K})$ telle que PM est échelonnée réduite.*

Remarque: la mise en place pratique de ce résultat est ce que l'on appelle, **l'algorithme de Gauss**.

Ainsi étant donné un système linéaire (S) de matrice augmentée $\tilde{A}_S$, on lui applique l'algorithme de Gauss pour obtenir l'échelonnée réduite associée,

par exemple de la forme

$$\begin{pmatrix} 0 & \cdots & \mathbf{1} & \bullet & \cdots & \bullet & \cdots & \bullet & \bullet & \bullet & \bullet & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & \mathbf{1} & \bullet & 0 & 0 & \bullet & 0 & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \mathbf{1} & 0 & \bullet & 0 & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & \mathbf{1} & \bullet & 0 & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \bullet \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \alpha \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

— Si la dernière colonne contient un pivot, alors le système est incompatible, ce qui dans l'exemple précédent correspond au cas $\alpha \neq 0$.
— Sinon, les positions des pivots fournissent les indices des variables dites *principales* alors que les autres sont dites *secondaires*. L'ensemble des solutions est alors un sous-espace affine de dimension le nombre de variables secondaires ; autrement dit quelles que soient les valeurs prises par ces variables secondaires, on obtient une unique solution pour les variables principales que l'on obtient en résolvant ce système du bas vers le haut.

Définition 140. *Le système* (S) *est dit* de Cramer *s'il possède une unique solution.*

Remarque: autrement dit, (S) est de Cramer s'il est compatible sans variable secondaire, ce qui permet de prouver la proposition suivante.

Proposition 141. (S) *est de Cramer si et seulement si* A_S *est une matrice inversible, ce qui impose en particulier que* $m = n$.

Remarque: on utilise des systèmes linéaires et leur résolution via l'algorithme de Gauss, par exemple pour trouver l'inverse d'une matrice, pour donner des équations linéaires d'un sous-espace dont on connait une famille génératrice...

4 Réduction des endomorphismes

Comme on l'a vu plus haut, à chaque endomorphisme f, on associe des matrices qui dépendent du choix des bases. On cherche alors à trouver des bases pour que la matrice soit la plus simple possible.
Remarque: on utilisera le langage de la section suivante avec les endomorphismes orthogonaux (resp. unitaires), symétriques (resp. hermitiens) dans le cadre réel (resp. complexe). On espère que la perte de logique de présentation des objets, sera compensée par l'aspect pratique de rerouver en une seule section, un ensemble de résultat portant sur la réduction.

4.1 Matrices équivalentes

Définition 142. *Deux matrices $A, A' \in \mathbb{M}_{m,n}(\mathbb{K})$ (resp. de $\mathbb{M}_n(\mathbb{K})$) sont dites* équivalentes *(resp.* semblables*) s'il existe deux matrices $P \in GL_m(\mathbb{K})$ et $Q \in GL_n(\mathbb{K})$ (resp. $P \in GL_n(\mathbb{K})$) telles que $A' = PAQ$ (resp. $A' = P^{-1}AP$).*

Remarque: A et A' sont équivalentes (resp. semblables) si elles représentent le même morphisme $f : E \to F$ (resp. $f : E \to E$) relativement à des bases différentes au départ et à l'arrivée (resp. des bases différentes mais les mêmes au départ et à l'arrivée).

Proposition 143. *Toute matrice $A \in \mathbb{M}_{m,n}(\mathbb{K})$ est équivalente à une matrice de la forme*

$$\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & 0 & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & 0 \end{pmatrix},$$

où le nombre de 1 *est égal au rang de f.*

Remarque: ainsi les classes d'équivalence dans $\mathcal{L}(E, F)$ sont données par le rang.

Considérons à présent le cas $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$) et où les espaces considérés sont munis d'un produit scalaire (resp. hermitien). On ne s'autorise alors à ne considérer que des bases orthonormées et donc des matrices de changement de base orthogonale (resp. unitaire).

Proposition 144. ***(Décomposition polaire)*** *Soit $A \in \mathbb{M}_m(\mathbb{R})$ (resp. $\mathbb{M}_m(\mathbb{C})$), on peut alors écrire A sous la forme $A = PU$ où $P \in \mathbb{M}_m(\mathbb{R})$ (resp. $P \in \mathbb{M}_m(\mathbb{C})$) est positive semi-définie de même rang que A et $U \in \mathbb{M}_m(\mathbb{R})$ (resp. $\mathbb{M}_m(\mathbb{C})$) dont les vecteurs colonnes forment une famille orthonormale, i.e. $U{}^tU = I_m$ (resp. $UU^* = I_m$). La matrice P est uniquement déterminée comme l'unique racine carré positive de A^tA (resp. de AA^*) alors que U n'est uniquement déterminée que si A est de rang m.*

Remarque: il faut voir cette décomposition comme la généralisation dans le cas $n = 1$ de l'écriture d'un nombre complexe sous la forme $\rho e^{i\theta}$.

Preuve : On traite le cas de $\mathbb{R}$, le cas de $\mathbb{C}$ se traitant de manière similaire. La matrice A^tA est symétrique positive et donc diagonalisable en base orthonormée, i.e. il existe O orthogonale telle que $OA^tA({}^tO) = \mathrm{diag}(\lambda_1, \cdots, \lambda_m)$ avec les λ_i positifs ou nuls. On note alors

$$P = {}^tO\mathrm{diag}(\sqrt{\lambda}_1, \cdots, \sqrt{\lambda}_m)O$$

qui est donc positive semi-définie. Supposons que P est inversible, i.e. que A est de rang m, alors $U = P^{-1}A$ vérifie

$$U^tU = P^{-1}(A^tA)^tP^{-1} = P^{-1}(P^2)P^{-1} = I_m.$$

Si A n'est pas de rang m, soit $(A_i)_{i\in\mathbb{N}}$ une suite de matrices de $\mathbb{M}_m(\mathbb{R})$ de rang m convergeant vers A. D'après ce qui précède, on peut écrire de manière unique $A_i = P_iU_i$: comme $(U_i)_{i\in\mathbb{N}}$ appartient au compact $O_m(\mathbb{R})$ et admet donc une valeur d'adhérence $U \in O_n(\mathbb{R})$ de sorte que la suite extraite correspondante des P_i converge vers $P := AU^{-1}$ nécessairement symétrique positive semi-définie puisque l'ensemble des matrices symétriques positives semi-définies est clairement fermé.
Remarque: en raisonnant avec des suites extraites comme dans la preuve ci-dessus, il est aisé de démontrer que la décomposition polaire dans le cas inversible, est un homéomorphisme.
Remarque: dans la preuve précédente on construit assez naturellement la matrice hermitienne positive P alors que U n'est donné qu'artificiellement par la formule $P^{-1}A$. Pour corriger cette injustice montrons le corollaire suivant.

Corollaire 145. *L'application $M \mapsto \sqrt{\frac{1}{n}\mathrm{tr}(MM^*)}$ définie une norme $||-||$ sur $\mathbb{M}_n(\mathbb{C})$ telle que $||U|| = 1$ pour toute matrice unitaire. Pour $A \in GL_n(\mathbb{C})$ de décomposition polaire $A = SU$, la matrice unitaire est uniquement déterminée par la propriété suivante :*

$$||A - U|| = \min_{U'\in\mathbb{U}_n(\mathbb{C})} ||A - U'||,$$

où $\mathbb{U}_n(\mathbb{C})$ désigne l'ensemble des matrices unitaires de $\mathbb{M}_n(\mathbb{C})$.

Remarque: une autre façon d'énoncer le résultat est de dire que U est la projection orthogonale de A sur la boule unité des matrices M telles que $||M|| \leq 1$ qui est un compact convexe.
Preuve : Calculons tout d'abord

$$\begin{aligned} ||A - U|| &= ||S - I_n|| = ||PDP^* - I_n|| \\ &= ||D - I_n|| = \sqrt{\tfrac{1}{n}\textstyle\sum_{i=1}^n(\sigma_i - 1)^2} \end{aligned}$$

où les σ_i sont les valeurs propres de S (on les appellera bientôt les valeurs singulières de A). Calculons alors

$$\begin{aligned} ||A - U'|| &= ||SU - U'|| = ||SU(U')^{-1} - I_n|| \\ &= ||PDP^*U(U')^{-1} - I_n|| = ||D(P^*U(U')^{-1}P)||. \end{aligned}$$

Notons alors $V = P^*U(U')^{-1}P$ est qui est unitaire dont on note $v_{i,j}$ les coefficients. On a alors

$$\begin{aligned} ||A - U'|| &= ||DV - I_n|| = \sqrt{\tfrac{1}{n}\textstyle\sum_{i=1}^n \sigma_i^2 + n - \sigma_i(v_{i,i} + \overline{v_{i,i}})} \\ &\leq \sqrt{\tfrac{1}{n}\textstyle\sum_{i=1}^n(\sigma_i^2 + 1 - 2\sigma_i} = ||A - U|| \end{aligned}$$

où on a utilisé la majoration $v_{i,i} + \overline{v_{i,i}} = 2\text{Re}\ (v_{i,i}) \leq |v_{i,i}| \leq 1$, puisque les vecteurs colonnes de V sont de norme 1.

Pour le cas d'égalité on doit avoir $|v_{i,i}| = 1$ et $2\text{Re}\ (v_{i,i}) = 1$ soit $v_{i,i} = 1$. Or la matrice V étant unitaire, on a $v_{i,j} = \delta_{i,j}$, i.e. $V = I_n$ soit $U(U')^{-1} = I_n$ et donc $U = U'$.

Corollaire 146. ***(valeurs singulières)***
Soit $A \in \mathbb{M}_m(\mathbb{C})$ de rang k, on peut alors l'écrire sous la forme

$$A = VDW^*,$$

où $V \in \mathbb{M}_m(\mathbb{C})$ et $W \in \mathbb{M}_m(\mathbb{C})$ sont unitaires et $D = \text{diag}(d_1, \cdots, d_m)$ avec $d_1 \geq d_2 \geq \cdots \geq d_k > d_k = \cdots = d_m = 0$. Les nombres $d_{i,i}$ sont les racines carrées positives des valeurs propres de AA^* et sont donc uniquement déterminées : on les appelle les valeurs singulières de A.

Remarque: V et W ne sont pas uniquement déterminées, on peut juste dire que les colonnes de V (resp. W) sont des vecteurs propres de AA^* (resp. A^*A).

Preuve : On part de la décomposition polaire $A = PU$ et on diagonalise $P = VDV^*$ en base orthonormée où D est comme dans l'énoncé. On pose alors $W = U^*V$ de sorte que $A = VDW^*$.

Remarque: les valeurs singulières apparaissent dans le conditionnement d'une matrice en analyse numérique. Rappelons brièvement de quoi il s'agit. Soit A une matrice inversible et B une matrice colonne, on cherche à résoudre l'équation $AX = B$ d'inconnue X. D'un point de vue pratique, B peut subir une petite perturbation δ_B due par exemple à des arrondis et on cherche à contrôler la perturbation δ_X induite sur X, i.e. $A(X + \delta_X) = B + \delta_B$, soit $A\delta_X = \delta_B$. On choisit une norme subordonnée par exemple à la classique norme euclidienne $||-||_2$ de sorte que $|||A|||_2$ est égal à la plus grande valeur propre de A^*A, i.e. à la plus grande valeur singulière. On a alors, en utilisant

$$||A\delta_X||_2 \leq |||A|||_2.||\delta_X||_2, \qquad ||A^{-1}B||_2 \leq |||A^{-1}|||_2.||B||_2 \tag{10}$$

on en déduit

$$\frac{|||\delta_X|||_2}{|||X|||_2} \leq |||A|||_2.|||A^{-1}|||_2 \frac{|||\delta_B|||_2}{|||B|||_2}.$$

Le conditionnement de A relativement à la norme $||-||_2$ est alors la quantité $|||A|||_2.|||A^{-1}|||_2$ qui est donc le quotient $\frac{\mu_n}{\mu_1}$ de la plus grande valeur singulière par la plus petite. En utilisant qu'en dimension fini, la sphère unité est compacte, les inégalités de (10) sont optimales, i.e. les cas d'égalités existent, de sorte que la majoration précédente est optimale.

Illustrons le phénomène :

— on prend $A \in GL_2(\mathbb{R})$ symétrique définie positive de sorte que les valeurs singulières sont égales aux valeurs propres.

— Soient e_1 et e_2 les vecteurs propres associés à $\lambda_1 \leq \lambda_2$ et on suppose λ_2 grand et λ_1 petit.
— On pose $B = e_2$ de sorte que $X = \frac{1}{\lambda_2} e_2$ qui est petit, et
— $\delta_B = \lambda_1 e_1$, qui est donc petit,

de sorte que $\delta_X = e_1$ est grand.
Remarque: on peut aussi perturber la matrice A en gardant B : à nouveau la perturbation de X est contrôlé par le conditionnement de A.

4.2 Vecteurs propres et espaces propres

Définition 147. *Un vecteur $v \in E$ est dit* propre *par un endomorphisme f s'il est non nul et s'il existe un scalaire $\lambda \in \mathbb{K}$ tel que $f(v) = \lambda v$; le scalaire λ est alors appelé* une valeur propre. *On note* Spec$f$ *l'ensemble des valeurs propres de f : c'est le spectre de f.*

Remarque: un vecteur propre définit une droite stable par f ; plus généralement un sous-espace F de E est dit *stable* par f si $f(F) \subset F$. Si on prend une base de F que l'on complète par en une base de E, la matrice de f dans cette base sera triangulaire par blocs, i.e. de la forme $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ avec $A \in \mathbb{M}_{n_1}(K)$ et $B \in \mathbb{M}_{n_2}(K)$ avec $n_1 + n_2 = n$.
Exemples : le noyau $\operatorname{Ker} f$ et l'image $\operatorname{Im} f$ sont des sous-espaces stables par f.

Proposition 148. *Soit E un espace vectoriel muni d'une base $(e_i)_{1 \leq i \leq n}$. Il existe alors une unique application* $\det_{(e_i)_i} : E^n \to \mathbb{K}$ *qui soit* multilinéaire alternée *et telle que* $\det_{(e_i)_i}(e_1, \cdots, e_n) = 1$.

Définition 149. *Pour $E = \mathbb{K}^n$ muni de la base canonique et $\mathbb{M}_n(\mathbb{K})$ identifié via ses vecteurs colonnes à E^n, l'application de la proposition précédente définit* $\det : \mathbb{M}_n(\mathbb{K}) \to \mathbb{K}$ *et s'appelle* le déterminant.

Remarque: par opération élémentaires sur les vecteurs colonnes d'une matrice, on montre que $\det A \neq 0$ si et seulement si $A \in GL_n(\mathbb{K})$ ainsi que le corollaire suivant.

Corollaire 150. *Pour $A, B \in \mathbb{M}_n(\mathbb{K})$ on a* $\det(AB) = \det A . \det B$.

Définition 151. Le polynôme caractéristique *d'un endomorphisme $f \in \mathcal{L}(E)$ est le déterminant $\chi_A(X) := \det(A(f) - XI_n) \in \mathbb{K}[X]$ où $A(f)$ est la matrice de f dans une base de E quelconque.*

Remarque: le fait que χ_A soit indépendant de la base provient du fait que $\det(P^{-1}AP) = \det A$ d'après le corollaire précédent.
Remarque: dans le cas où χ_A est totalement décomposé (par exemple si $\mathbb{K} = \mathbb{C}$), le produit des valeurs propres est égal à $(-1)^n$ fois le coefficient constant de χ_A et donc au déterminant de A.

Lemme 152. *Sur $\mathbb{C}$, la norme du produit des valeurs propres d'une matrice $A \in \mathbb{M}_n(\mathbb{C})$ est égal au produit des valeurs absolues de ses valeurs singulières.*

Remarque: au corollaire 253, on donnera des raffinements de cette égalité.

Preuve : Il suffit de noter que le déterminant d'une matrice unitaire est nécessairement de module 1, puisque $U^*U = I_n$ implique que $|\det U| = 1$.

Remarque: la formation du polynôme caractéristique $A \mapsto \chi_A(X)$ est clairement continue puisque polynomiale. Rappelons par ailleurs que, cf. le corollaire **??**, les racines dépendent continûment de leur polynôme, de sorte que les valeurs propres dépendent continûment de la matrice. La proposition suivante quantifie cette propriété.

Proposition 153. ***(Disques de Gershgorin)*** *Les valeurs propres de $A = (a_{i,j})_{1\leq i,j\leq n} \in \mathbb{M}_n(\mathbb{C})$ appartiennent à la réunion des disques fermés centrés en $a_{i,i}$ et de rayon $\sum_{j\neq i} |a_{i,j}|$.*

Preuve : Le résultat découle directement du lemme d'Hadamard appliqué à $A - \lambda\mathrm{Id}$. Rappelons que ce lemme dit que si pour tout $1 \leq i \leq n$, on a $|a_{i,i}| > \sum_{j\neq i} |a_{i,j}|$ alors A est inversible. En effet soit X de coordonnées $(x_i)_{1\leq i\leq n}$ dans le noyau de A et soit i_0 tel que $|x_{i_0}|$ soit maximal parmi les $|x_i|$. De l'égalité $a_{i_0,i_0}x_{i_0} = -\sum_{j\neq i_0} a_{i_0,j}x_j$ on en déduit la majoration $|a_{i_0,i_0}x_{i_0}| \leq |x_{i_0} \sum_{j\neq i_0} |a_{i_0,j}|$ et donc $x_{i_0} = 0$ soit $X = 0$.

D'un point de vue pratique, examinons la perturbation subie par les valeurs propres lorsqu'on perturbe la matrice. Commençons par l'exemple simple donné par la matrice compagnon de $X^{100} - 10^{-100}$ dont les valeurs propres sont de module égal à 1. Cette matrice est ainsi une très faible perturbation du bloc de Jordan de taille 100 dont les valeurs propres sont toutes nulles. En conclusion une perturbation d'ordre 10^{-100} nous conduit à une perturbation de l'ordre $0,1$ ce qui est énorme. Essayons de quantifier ce phénomène : pour ce faire on considère une norme matricielle $||.||$ telle que pour toute matrice diagonale $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$, on ait $||D|| = \max_i |\lambda_i|$. On peut par exemple prendre les normes $|||.|||_1$, $|||.|||_2$ ou $|||.|||_\infty$.

Proposition 154. *Soit A une matrice diagonalisable avec $\mathrm{Spec}A = \{\lambda_1, \cdots, \lambda_n\}$. Alors*

$$\mathrm{Spec}(A + \delta_A) \subset \bigcup_{i=1}^{n} \big\{ z \in \mathbb{C} : \ |z - \lambda_i| \leq \gamma(A)||\delta_A|| \big\},$$

où

$$\gamma(A) = \inf\{||P||.||P^{-1}|| : \ P^{-1}AP = \mathrm{diag}(\lambda_1, \cdots, \lambda_n).$$

Remarque: Ainsi donc le contrôle des valeurs propres est donné par le conditionnement des matrices de passage et pas par le conditionnement de la matrice de départ.

Preuve : Considérons P diagonalisant A, i.e.

$$P^{-1}AP = D := \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$$

et soit λ une valeur propre de $A+\delta_A$. Si $\lambda \in \{\lambda_1, \cdots, \lambda_n\}$, il n'y a rien à montrer, sinon $(D-\lambda I_n)$ est inversible et on peut écrire

$$\begin{aligned} P^{-1}(A+\delta_A-\lambda I_n)P &= D-\lambda I_n + P^{-1}\delta_A P \\ &= (D-\lambda I_n)(I_n + (D-\lambda I_n)^{-1}P^{-1}\delta_A P). \end{aligned}$$

La matrice $(I_n + (D-\lambda I_n)^{-1}P^{-1}\delta_A P)$ n'étant pas inversible, -1 est donc une valeur propre de $(D-\lambda I_n)^{-1}P^{-1}\delta_A P)$, de sorte que d'après **??** sa norme est ≥ 1, ce qui donne :

$$1 \leq ||(D-\lambda I_n)^{-1}P^{-1}\delta_A P)|| \leq ||(D-\lambda I_n)^{-1}||.||P^{-1}||.||\delta_A||.||P||.$$

Comme par hypothèse $||(D-\lambda I_n)^{-1}|| = \frac{1}{\min|\lambda_i-\lambda|}$, il existe au moins un indice i tel que

$$|\lambda_1 - \lambda| \leq ||P||.||P^{-1}||.||\delta_A||.$$

Remarque: En particulier si A est normale, la matrice de passage est orthogonale et son conditionnement est égal à 1. Autrement dit lorsqu'on perturbe une matrice normale, ses valeurs propres sont perturbées dans la même proportion. Le cas le plus intéressant est certainement celui où A et δ_A sont toutes deux symétriques, on renvoie au §5.6 pour une étude plus précise dans cette situation.

Définition 155. *Le sous-espace propre E_λ (resp. caractéristique $E(\lambda)$) associée à une valeur propre λ est* $\mathrm{Ker}(f-\lambda\mathrm{Id})$ *(resp.* $\mathrm{Ker}(f-\lambda\mathrm{Id})^n$ *où n est la dimension de E, ou plus simplement la multiplicité de λ dans le polynôme minimal μ_f).*

Remarque: la dimension du sous-espace caractéristique est égale à la multiplicité de λ dans le polynôme caractéristique.

Lemme 156. ***(dit des noyaux)***
Si $P = P_1P_2$ est un polynôme annulateur de f avec $P_1 \wedge P_2 = 1$ alors $E = \mathrm{Ker}\, P_1(f) \oplus \mathrm{Ker}\, P_2(f)$.

Preuve : On part d'une relation de Bezout $UP_1 + VP_2 = 1$ de sorte que pour tout $e \in E$, on a $e = e_1 + e_2$ avec $e_1 = UP_1(f)(e)$ et $e_2 = VP_2(f)(e)$. On a alors

$$P_2(f) = (e_1) = U(f) \circ P_1P_2(f)(e_1) = 0 \text{ et } P_1(f) = V(f) \circ P_1P_2(f)(e) = 0$$

et donc $e_1 \in \mathrm{Ker}\, P_2(f)$ et $e_2 \in \mathrm{Ker}\, P_1(f)$ et $E = \mathrm{Ker}\, P_1(f) + \mathrm{Ker}\, P_2(f)$. En outre si $e \in \mathrm{Ker}\, P_1(f) \cap \mathrm{Ker}\, P_2(f)$ alors $e_1 = e_2 = 0$ et $e = e_1 + e_2 = 0$.
Remarque: il est important de noter que les projecteurs sur chacun de ces sous-espaces stables parallèlement à l'autre, sont des polynômes en f.

4.3 Polynôme minimal

La théorie de la réduction d'un endomorphisme est totalement contrôlée par une série de polynômes qu'on lui associe, appelés *ses invariants de similitude*. Nous verrons que le polynôme caractéristique est le produit des invariants de similitude. En général, la réponse à une question sur un endomorphisme s'exprime en utilisant toute la puissance des invariants de similitude. Le plus gros des invariants de similitude est donné par la définition suivante où l'on rappelle qu'étant donné un endomorphisme f et un polynôme $P(X) = \sum_i a_i X^i \in \mathbb{K}[X]$, $P(f)$ désigne l'endomorphisme $\sum_i a_i f^i$ où f^i désigne $f \circ \cdots \circ f$.

Définition 157. *Pour $f \in \mathcal{L}(E)$, l'ensemble I_f des polynôme $P \in \mathbb{K}[X]$ tels que $P(f)$ est l'endomorphisme nul, est un idéal de $\mathbb{K}[X]$; cet anneau étant principal, il existe un unique polynôme unitaire μ_f, appelé* polynôme minimal *de f, tel que $I_f =< \mu_f >$.*

Remarque: comme E est de dimension finie, la famille $\mathrm{Id}, f, f^2, \cdots, f^{n^2+1}$ est liée de sorte que I_f n'est pas l'idéal nul et donc μ_f n'est pas le polynôme nul.

Théorème 158. ***(de Cayley-Hamilton)***
Le polynôme caractéristique χ_f appartient à I_f, i.e. $\chi_f(f)$ est l'endomorphisme nul.

4.4 Trigonalisation

Définition 159. *Un endomorphisme est dit* trigonalisable *s'il existe une base dans laquelle sa matrice est triangulaire supérieure.*

Remarque: une autre façon d'énoncer cette propriété est de demander qu'il existe une drapeau complet

$$\{0\} = F_0 \subset F_1 \subset \cdots \subset F_n = E$$

avec $\dim F_i = i$, stable par f, i.e. pour tout $i = 0, \cdots, n$ on a $f(F_i) \subset F_i$.

Théorème 160. *Un endomorphisme f est trigonalisable si et seulement si χ_f est scindé sur $\mathbb{K}$.*

Remarque: ainsi si K est algébriquement clos tous les endomorphismes sont trigonalisables.

On s'intéresse à présent à la question de la trigonalisation simultanée, i.e. étant donné un sous-ensemble $\mathcal{E}$ de $\mathcal{L}(E)$, on cherche des drapeaux, si possible complets, stables par tous les $u \in \mathcal{E}$. On rappelle que si u est un endomorphisme d'un espace vectoriel E et si $F \subset E$ est un sous-espace stable par u alors u induit un endomorphisme $\overline{u}$ de E/F.

Étant donnés un sous-ensemble $\mathcal{E}$ de $\mathcal{L}(E)$ et $G \subset F \subset E$ des sous-espaces stables par tous les éléments u de $\mathcal{E}$, l'ensemble des quotients de $\mathcal{E}$ pour $\{G \subset F\}$ est par définition $\{\overline{u} \in \mathcal{L}(F/G) : \ u \in \mathcal{E}\}$.

Définition 161. *Une propriété P sera dite stable par quotients si pour tout ensemble $\mathcal{E} \subset \mathcal{L}(E)$ constitués d'éléments satisfaisant P alors l'ensemble quotient de $\mathcal{E}$ pour $\{G \subset F\}$ est aussi constitué d'éléments de $\mathcal{L}(F/G)$ satisfaisant P.*

Principe général : soit $\mathcal{P}$ est un ensemble de propriétés stables par quotients et vérifiant la propriété suivante : pour tout $\mathcal{E} \subset \mathcal{L}(E)$ constitué d'éléments vérifiant $\mathcal{P}$ avec $\dim E > 1$, $\mathcal{E}$ est réductible i.e. il existe un sous-espace vectoriel non trivial F de E stable par tous les éléments de $\mathcal{E}$. Alors $\mathcal{E}$ est triangularisable.
Exemple : pour E un $\mathbb{C}$-espace vectoriel de dimension finie, tout sous-ensemble commutatif de $\mathcal{L}(E)$ est triangularisable. En effet la commutativité est clairement une propriété stable par quotient. La propriété de réductibilité découlera alors des faits suivants :

- tout endomorphisme admet une valeur propre et
- tout sous-espace propre de A est stable par toute matrice B commutant avec A.

Remarque 162. en utilisant les faits suivants :

- les valeurs propres d'une matrice triangulaires sont ses termes diagonaux ;
- les termes diagonaux d'un produit (resp. somme) de matrice triangulaire sont les produits $t_{i,i}^{(1)} t_{i,i}^{(2)}$ (resp. $t_{i,i}^{(1)} + t_{i,i}^{(2)}$) de leurs termes diagonaux.

on en déduit que si $\{A_1, \cdots, A_k\}$ sont simultanément diagonalisables et si p est un polynôme non commutatif (i.e. une combinaison linéaire de mots) en $\{A_1, \cdots, A_k\}$ alors

$$\sigma(p(A_1, \cdots, A_k)) \subset p(\sigma(A_1), \cdots, \sigma(A_k))$$

où $p(\sigma(A_1), \cdots, \sigma(A_k))$ désigne l'ensemble des $p(\lambda_1, \cdots, \lambda_k)$ où pour tout $i = 1, \cdots, k$, $\lambda_i \in \sigma(A_i)$. Dans la suite nous allons donner la réciproque à ce résultat.

Si $\mathcal{A}$ est une sous-algèbre de $\mathcal{L}(E)$ l'ensemble $\mathcal{A}.x := \{Ax : \ A \in \mathcal{A}\}$, où $x \in E$, est un sous-espace stable par $\mathcal{A}$. Si $\mathcal{A}.x = E$, on dit que x est *un vecteur cyclique* pour $\mathcal{A}$. La détermination des sous-algèbres de $\mathcal{L}(E)$ qui possèdent des sous-espaces invariants non triviaux est réglée par le théorème suivant qui s'occupe de la partie réductibilité du principe général énoncé plus haut dans le cas des sous-algèbres de $\mathcal{L}(E)$.

Théorème 163. ***(de Burnside)***
Toute sous-algèbre propre de $\mathcal{L}(E)$ est réductible.

Preuve : Soit $\mathcal{A}$ une sous-algèbre irréductible de $\mathcal{L}(E)$; comme tout endomorphisme est une somme d'endomorphisme de rang 1, nous allons montrer que tout endomorphisme de rang 1 appartient à $\mathcal{A}$.

Montrons tout d'abord que $\mathcal{A}$ contient un élément de rang 1. Soit $u_0 \in \mathcal{A}$ non nul de rang minimal ; si ce rang est strictement plus grand que 1, alors il existe des vecteurs x_1 et x_2 tels que $(u_0(x_1), u_0(x_2))$ est linéairement indépendant. Comme $\{u \circ u_0(x_1) :\ u \in \mathcal{A}\} = E$, il existe $u_1 \in \mathcal{A}$ tel que $u_1 \circ u_0(x_1) = x_2$ et donc $(u_0 \circ u_1 \circ u_0(x_1), u_0(x_1))$ est libre. Soit alors λ tel que la restriction de $u_1 \circ u_0 - \lambda \mathrm{Id}$ à $u_0(E)$ n'est pas inversible ; $(u_0 \circ u_1 - \lambda \mathrm{Id})u_0$ est non nul car l'image de x_1 est non nulle, et $(u_0 \circ u_1 - \lambda \mathrm{Id}) \circ u_0$ est de rang strictement plus petit que celui de u_0, d'où la contradiction et donc u_0 est de rang 1.

Pour y_0 dans l'image de u_0, on considère la forme linéaire φ_0 définie par $u_0(x) = \varphi_0(x)y_0$. Soit alors $u \in \mathcal{L}(E)$ défini par $u(x) = \varphi(x)y$ où $y \in E$ et $\varphi \in E^*$. Montrons alors que u appartient à $\mathcal{A}$. Pour $v \in \mathcal{A}$, on a $u_0 \circ v \in \mathcal{A}$ et $u_0 \circ v(x) = \varphi_0(v(x))y_0$. Soit alors $F' \subset E^*$, l'ensemble des formes linéaires φ telles que $x \mapsto \varphi(x)y_0$ appartienne à $\mathcal{A}$: F' est clairement un sous-espace de E^*. Si ce sous-espace était strict, il existerait $x_0 \neq 0$ tel que $\varphi(x_0) = 0$ pour tout $\varphi \in F'$ (un espace vectoriel de dimension finie est réflexif). La contradiction découle alors du fait que $\varphi_0(v(x_0)) = 0$ pour tout $v \in \mathcal{A}$ implique que x_0 est nul car $\{v(x_0) :\ v \in \mathcal{A}\} = E$. Soit donc $v_1 \in \mathcal{A}$ tel que $\varphi = \varphi_0 \circ v_1$.

De même comme $y_0 \neq 0$, alors $\{v(x_0) :\ v \in \mathcal{A}\} = E$ et donc pour tout $y \in E$ soit $v_2 \in \mathcal{A}$ tel que $v_2(y_0) = y$ et donc $u = v_2 \circ u_0 \circ v_1$.

Corollaire 164. *Les seuls idéaux bilatères de $\mathcal{L}(E)$ sont $\{0\}$ et $\mathcal{L}(E)$.*

Preuve : Soit $\mathcal{I}$ un idéal bilatère de $\mathcal{L}(E)$ non réduit à 0. Il suffit alors de montrer que $\mathcal{I}$ est irréductible. Si $u \neq 0$ appartient à $\mathcal{I}$, pour tout $0 \neq x \in E$, il existe $v \in \mathcal{L}(E)$ tel que $u \circ v(x) \neq 0$. Soit $y \in E$ et $w \in \mathcal{L}(E)$ tel que $w \circ u \circ v(x) = y$. On a $w \circ u \circ v \in \mathcal{I}$ de sorte que tout vecteur $x \neq 0$ est cyclique pour $\mathcal{I}$ et donc $\mathcal{I}$ est irréductible.

Corollaire 165. *Soit E est $\mathbb{C}$-espace vectoriel de dimension finie alors tout automorphisme d'algèbre ϕ de $\mathcal{L}(E)$ est intérieur, i.e. il existe $P \in GL(E)$ tel que pour tout $A \in \mathcal{L}(E)$, $\phi(A) = PAP^{-1}$.*

Preuve : Soit $A_0 \in \mathcal{L}(E)$ un idempotent de rang 1, $\phi(A_0)$ est alors un idempotent, montrons qu'il est aussi de rang 1. L'ensemble $\{A_0BA_0 :\ B \in \mathcal{L}(E)\}$ est un sous-espace vectoriel de $\mathcal{L}(E)$ de dimension 1 : on peut l'identifier avec $\mathcal{L}(\mathrm{Im}\, A_0)$. Son image par ϕ, $\{\phi(A_0)C\phi(A_0) :\ C \in \mathcal{L}(E)\}$ est donc aussi un sous-espace de $\mathcal{L}(E)$ de dimension 1 identifié à $\mathcal{L}(\mathrm{Im}\, \phi(A_0))$ de sorte que $\phi(A_0)$ est de rang 1. Comme tous les idempotents de rang 1 sont semblables à $\mathrm{diag}(1, 0, \cdots, 0)$, quitte à composer ϕ par $A \mapsto PAP^{-1}$, on peut supposer que $\phi(A_0) = A_0$.

Notons x_0 un vecteur directeur de $\operatorname{Im} A_0$ et soit $P \in \mathcal{L}(E)$ défini par $P(Bx_0) = \phi(B)x_0$: si $B_1x_0 = B_2x_0$ alors comme $A_0x_0 = x_0$, on a $(B_1 - B_2)A_0 = 0$ et donc $(\phi(B_1) - \phi(B_2))A_0 = 0$ de sorte que $\phi(B_1)x_0 = \phi(B_2)x_0$ et P est bien définie et évidemment linéaire. Supposons que $\phi(B)x_0 = 0$ de sorte que $\phi(B)\phi(A_0) = \phi(BA_0) = 0$ et donc $BA_0 = 0$ soit $Bx_0 = 0$ ce qui prouve l'injectivité de P et comme on est en dimension finie $P \in GL(E)$.

Soit alors $A \in \mathcal{L}(E)$, on a $P(AB)x_0 = \phi(AB)x_0 = \phi(A)\phi(B)x_0 = \phi(A)PBx_0$ et donc $PAy = \phi(A)Py$ pour tout $y = Bx_0$. Quand B décrit $\mathcal{L}(E)$, y décrit E et donc $PA = \phi(A)P$ pour tout $A \in \mathcal{L}(E)$ d'où le résultat.

Corollaire 166. *Toute algèbre d'endomorphismes nilpotents est triangularisable.*

Preuve : La propriété d'être nilpotent est stable par quotient comme en outre il existe des éléments de $\mathcal{L}(E)$ qui ne sont pas nilpotents, toute algèbre constituée d'endomorphismes nilpotents est, d'après le théorème de Burnside, réductible. La triangularisation découle alors du principe général énoncé plus haut.

Théorème 167. *Si $\mathcal{A}$ est une sous-algèbre de $\mathcal{L}(E)$ alors $\mathcal{A}$ est triangularisable si et seulement si tout commutateur $BC - CB$ avec $B, C \in \mathcal{A}$ est nilpotent.*

Preuve : Si $\mathcal{A}$ est triangularisable alors d'après la remarque 162, on a $\sigma(BC - CB) = \{0\}$ et donc $BC - CB$ est nilpotent. Réciproquement la propriété d'avoir des commutateurs nilpotents est stable par quotient et comme il existe des commutateurs non nilpotents dans $\mathcal{L}(E)$, d'après le théorème de Burnside, $\mathcal{A}$ est réductible ; la triangularisation découle alors du principe général.
Remarque: on peut ainsi voir la triangularisabilité comme une généralisation de la commutativité ; on relâche la condition d'être nul pour un commutateur, en demandant qu'il soit nilpotent.

Théorème 168. ***((McCoy)*** *La paire $\{A, B\}$ est triangularisable si et seulement si $p(A, B)(AB - BA)$ est nilpotent pour tout polynôme commutatif p en A et B.*

Preuve : Le sens direct découle de la remarque 162. Pour la réciproque d'après le principe général il suffit de montrer que l'algèbre $\mathcal{A}$ engendrée par A, B est réductible dès que $\dim E > 1$. Si $AB = BA$ c'est clair, sinon soit $x \in E$ tel que $(AB - BA)x \neq 0$ et $C \in \mathcal{L}(E)$ vérifiant $C(AB - BA)x = x$. Si $\mathcal{A}$ était irréductible alors d'après le théorème de Burnside elle serait égale à $\mathcal{L}(E)$ et donc $C \in \mathcal{A}$ et donc de la forme $p(A, B)$. La contradiction découle alors du fait que $C(AB - BA)$ n'est pas nilpotent.

Corollaire 169. *Une sous-algèbre unitaire $\mathcal{A}$ de $\mathcal{L}(E)$ est triangularisable si et seulement si $\mathcal{A}/\mathrm{Rad}\mathcal{A}$ est commutatif, où*

$$\mathrm{Rad}\mathcal{A} = \{A \in \mathcal{A} : \sigma(AB) \subset \{0\}\ \forall B \in \mathcal{A}\}$$

est le radical de A, i.e. l'intersection de tous les idéaux à droite (ou à gauche) maximaux de $\mathcal{A}$.

Preuve : Si $\mathcal{A}$ est triangularisable et si $B, C \in \mathcal{A}$ alors pour tout $A \in \mathcal{A}$ d'après la remarque 162, on a $\sigma\Big((BC - CB)A\Big) = \{0\}$ et donc $BC - CB \in \mathrm{Rad}\mathcal{A}$ i.e. $\mathcal{A}/\mathrm{Rad}\mathcal{A}$ est commutatif.

Réciproquement si $\mathcal{A}/\mathrm{Rad}\mathcal{A}$ est commutatif alors $BC - CB \in \mathrm{Rad}\mathcal{A}$ pour tout $B, C \in \mathcal{A}$ de sorte que $\mathcal{A}$ est triangularisable d'après le corollaire 166.

Nous allons nous intéresser à des sous-espaces vectoriels de $\mathcal{L}(E)$ stable par certaines multiplications non associatives comme par exemple les *algèbres de Lie* stables sous le crochet de Lie $[A, B] = AB - BA$, ou les algèbre de Jordan stable sous $(A, B) \mapsto AB + BA$. Le résultat le plus célèbre est le théorème de Engel ci- dessous.

Théorème 170. *Soit $\mathcal{N}$ un sous-ensemble du cône nilpotent de $\mathcal{L}(E)$ vérifiant la propriété suivante : pour tout $A, B \in \mathcal{N}$ il existe un polynôme non commutatif p en Aet B tel que $AB + p(A, B)A \in \mathcal{N}$. Alors $\mathcal{N}$ est triangularisable.*

Preuve : On raisonne par récurrence sur la dimension n de E ; le cas $n = 1$ étant trivial supposons donc le résultat acquis jusqu'au rang n et traitons celui de $n+1$. Soit $\mathcal{F}$ l'ensemble des sous-espaces de E qui sont des intersections de noyaux d'éléments de $\mathcal{N}$, i.e. de la forme $V_{\mathfrak{S}} = \cap_{A \in \mathfrak{S}} \mathrm{Ker}\, A$ où $\mathfrak{S} \subset \mathcal{N}$, et soit $K \in \mathcal{F}$ de dimension minimale non nulle. Notons alors $\mathcal{N}_0 = \{A \in \mathcal{N} :\ Ax = 0\ \forall x \in K\}$; l'ensemble $\overline{\mathcal{N}_0} \subset \mathcal{L}(E/K)$ vérifie les hypothèses de l'énoncé de sorte que d'après l'hypothèse de récurrence $\overline{\mathcal{N}_0}$ est triangularisable et donc $\mathcal{N}_0$ aussi car ses éléments sont nuls sur K.

Il suffit alors de montrer que $\mathcal{N}_0 = \mathcal{N}$. Dans le cas contraire soit $B \in \mathcal{N}$ et $x \in K$ tels que $Bx \neq 0$. Si K était un sous-espace stable de l'endomorphisme nilpotent B, il existerait $x_0 \in K$ tel que $Bx_0 = 0$ et $\mathrm{Ker}\, B \cap K$ serait un élément non nul de $\mathcal{F}$ de dimension strictement plus petite que celle de K ce qui n'est pas par hypothèse. Soit alors $x_1 \in K$ et $A_1 \in \mathcal{N}_0$ tel que $A_1 B x_1 \neq 0$ et soit p_1 un polynôme non commutatif en A_1 et B tel que $B_1 = A_1 B + p_1(A_1, B)A_1 \in \mathcal{N}$. Comme $B_1 x_1 \neq 0$ comme précédemment K n'est pas un sous-espace stable de B_1 ; soit alors $A_2 \in \mathcal{N}_0$ et $x_2 \in K$ tels que $A_2 B_1 x_2 \neq 0$. Soit p_2 tel que $B_2 = A_2 B_1 + p_2(A_2, B_1)A_2 \in \mathcal{N}$. En continuant le processus, on construit

$$\{A_1, A_2, \cdots, A_{n+1}\} \subset \mathcal{N}_0, \qquad \{B_1, B_2, \cdots, B_n\} \subset \mathcal{N}$$

et des vecteurs $\{x_1, \cdots, x_{n+1}\}$ tels que

$$A_{n+1}A_n \cdots A_2 A_1 B x_{n+1} = A_{n+1} B_n x_{n+1} \neq 0.$$

Comme $\mathcal{N}_0$ est triangularisable, tout produit de $n+1$ de ses éléments est nul de sorte que $A_{n+1} \cdots A_1 = 0$ ce qui contredit $A_{n+1} \cdots A_1 B x_{n+1} \neq 0$. On en déduit donc que $\mathcal{N}_0 = \mathcal{N}$ qui est trigonalisable.

Corollaire 171. *Un ensemble $\mathcal{N}$ d'éléments nilpotents de $\mathcal{L}(E)$ qui vérifie une des propriétés suivantes est trigonalisable :*

- ***théorème de Jacobson** : pour tout $A, B \in \mathcal{N}$, il existe un scalaire c tel que $AB - cBA \in \mathcal{N}$;*
- ***théorème de Engel** : pour tout $A, B \in \mathcal{N}$, $[A, B] = AB - BA \in \mathcal{N}$;*
- *pour tout $A, B \in \mathcal{N}$, $AB + BA \in \mathcal{N}$.*

Corollaire 172. *Soit $\mathcal{E} \subset \mathcal{L}(E)$ stable par le crochet de Lie. Alors $\mathcal{E}$ est triangularisable si et seulement si tous ses commutateurs sont nilpotents.*

Preuve : Le sens direct découle de la remarque 162. Réciproquement comme la propriété est clairement stable par quotient d'après le principe général il suffit de montrer que $\mathcal{E}$ admet un sous-espace stable. Soit $\mathcal{N}$ l'ensemble des commutateurs de $\mathcal{E}$; si $\mathcal{N} = \{0\}$ alors $\mathcal{E}$ est commutatif admet donc un sous-espace stable. Sinon d'après le corollaire précédent $\mathcal{N}$ est trigonalisable de sorte que $K = \cap_{N \in \mathcal{N}} \operatorname{Ker} N$ est un sous-espace non nul stable par tous les éléments de $\mathcal{E}$. En effet pour $x \in K$ et $B \in \mathcal{E}$, pour tout $A \in \mathcal{N}$, on a $Ax = 0$ et $(AB - BA)x = 0$ de sorte que $ABx = 0$ soit $Bx \in K$, d'où le résultat.

Théorème 173. ***(Levitzki)*** *Tout semi-groupe S d'éléments nilpotents de $\mathcal{L}(E)$ est triangularisable.*

Preuve : La nilpotence étant une propriété stable par quotient, il suffit d'après le principe général de montrer la réductibilité de S dès que $\dim E > 1$. La trace est une forme linéaire qui s'annule sur S et donc sur l'algèbre engendrée par S qui est simplement l'espace vectoriel engendré par S. Le résultat découle alors du théorème de Burnside et du fait qu'il existe des éléments de $\mathcal{L}(E)$ de trace non nulle.
Remarque: on aurait aussi pu déduire ce résultat du théorème de Jacobson pour $c = 0$.

Théorème 174. ***(Kolchin)*** *Si tout élément du semi-groupe S est unipotent alors S est triangularisable.*

Preuve : L'unipotence étant une propriété stable par quotient il suffit d'après le principe général de montrer la réductibilité de S dès que $\dim E > 1$. Si tous les commutateurs sont nuls alors S est abélien et donc réductible. Sinon soit $C = AB - BA \neq 0$ et soit $\mathcal{A}$ l'algèbre engendrée par S. L'idéal

bilatère engendré par C est alors contenu dans le noyau de la forme linéaire trace : en effet XCY s'écrit comme une combinaison linéaire de produit DCE avec $D, E \in S$ de sorte que $\mathrm{tr}(DCE) = \mathrm{tr}(DABE) - \mathrm{tr}(DBAE) = n - n = 0$. Comme $\mathcal{L}(E)$ n'admet pas d'idéal propre non nul, on en déduit que $\mathcal{A}$ n'est pas égal à $\mathcal{L}(E)$ et est donc réductible d'après le théorème de Burnside.

Si $\mathcal{E} \subset \mathcal{L}(E)$ est triangularisable alors la fonction trace est permutable sur $\mathcal{E}$, i.e. pour tout $A_1, \cdots, A_m \in \mathcal{E}$ et pour toute permutation $\sigma \in \mathfrak{S}_m$, on a

$$\mathrm{tr}(A_1 A_2 \cdots A_m) = \mathrm{tr}(A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(m)}).$$

La réciproque est vraie et découle simplement du théorème 167.

Proposition 175. *Soit $\mathcal{E} \subset \mathcal{L}(E)$, alors $\mathcal{E}$ est triangularisable si et seulement si la trace est permutable sur $\mathcal{E}$.*

Preuve : Pour tout $A, B, C \in \mathcal{E}$, on a $\mathrm{tr}\Big((AB - BA)C\Big) = 0$. En particulier on en déduit que pour tout $k > 0$, $\mathrm{tr}(AB - BA)^k = 0$ et donc $AB - BA$ est nilpotent d'où le résultat d'après 167.

Corollaire 176. *Si S est un semi-groupe qui vérifie une des propriétés suivantes, est triangularisable :*

(i) ***Kaplansky*** *: la trace est constante sur S ;*

(ii) la trace est multiplicative sur S.

En outre toutes dans la première situation les termes diagonaux dans une telle triangularisation ne dépendent que de S et sont soit égaux à 0 ou 1 ; dans la deuxième il existe j qui ne dépend que de S tels que tous les termes diagonaux dans une triangularisation, hors celui en (j, j), sont nuls.

Preuve : (i) La triangularisabilité découle de la proposition précédente. Soit alors $A \in S$ de sorte que les $\mathrm{tr}A^k$ sont constants pour tout $k \geq 1$ égaux à c. Notons $\lambda_1, \cdots, \lambda_m$, les valeurs propres non nulles de A ; il suffit alors de montrer que pour tout $\mathrm{tr}A = m$. En effet d'après les relations de Newton, l'ensemble $\{\lambda_1, \cdots, \lambda_m\}$ est déterminé uniquement par les $S_k = \lambda_1^k + \cdots + \lambda_m^k$ pour $k = 1, \cdots, m$; si tous les S_k sont égaux à 1 alors $\lambda_1 = \cdots = \lambda_m = 1$ convient trivialement. Notons σ_i les fonctions symétriques usuelles des λ_j de sorte que l'on a

$$S_{m+1} - S_m \sigma_1 + S_{m-1} \sigma_2 + \cdots + (-1)^m S_1 = 0 \sigma_n$$

$$S_m - \sigma_1 S_{m-1} + \cdots + (-1)^m m \sigma_m = 0$$

de sorte que comme $\sigma_m \neq 0$, on obtient $c = m$.

Ainsi les valeurs propres de A sont 0 ou 1 ; si les termes diagonaux de $A, B \in S$ n'étaient pas égaux alors on aurait $\mathrm{tr}(ST) < \mathrm{tr}S$ ce qui n'est pas.

(ii) La triangularisabilité découle de la proposition précédente. Soit $A \in S$ de valeurs propres $\lambda_1, \cdots, \lambda_n$. Par hypothèse on a $\sum_i \lambda_i^k = (\sum_i \lambda_i)^k$. Si

$\sum_i \lambda_i = 0$ alors tous les λ_i sont nuls d'où le résultat. Si $\sum_i \lambda_i = b \neq 0$ alors pour tout $k \geq 1$, $\sum_i (a_i/b)^k = 1$ de sorte que d'après la preuve de (i), il existe un unique j tel que $\lambda_j \neq 0$. La multiplicativité de la trace implique que ce j est le même pour tous les $A \in S$.

Corollaire 177. *Soit G un sous-groupe de $GL(E)$; les propriétés suivantes sont alors équivalentes :*

(i) G est triangularisable ;

(ii) pour tout $g \in G$, la trace est constante sur $gD(G)$;

(iii) la trace est constante sur $D(G)$;

(iv) pour tout $A \in D(G)$, $\sigma(A) = \{1\}$, i.e. A est unipotent.

Preuve : (i) $\Rightarrow$ (ii) : d'après la proposition précédente, la trace est permutable sur G de sorte que pour tout $g \in G$ et $h \in D(G)$, on a $\mathrm{tr}(gh) = \mathrm{tr}(g\mathrm{Id}) = \mathrm{tr}g$.

(ii) $\Rightarrow$ (iii) : immédiat

(iii) $\Leftrightarrow$ (iv) : découle du théorème de Kaplansky prouvé ci-dessus en utilisant que les éléments de G sont inversibles.

(iv) $\Rightarrow$ (i) : comme la propriété (iv) est clairement stable par quotient, il suffit donc de montrer que G est réductible. D'après le théorème de Kolchin, $D(G)$ est triangularisable : soit $\{0\} = F_0 \subset F_1 \subset \cdots \subset F_{n-1} \subset F_n = E$ le drapeau complet correspondant. Supposons $D(G) \neq \{\mathrm{Id}\}$ car sinon G est commutatif et le résultat est clair. Soit F le sous-espace vectoriel engendré par $\bigcup_{h \in D(G)} (h - \mathrm{Id})(E)$. Comme pour tout $h \in D(G)$, $(h - \mathrm{Id})(E) \subset F_{n-1}$, F est un sous-espace strict de E qui de plus est invariant sous G : en effet soit $g \in G$ alors pour tout $h \in D(G)$, on a $g(h - \mathrm{Id}) = (ghg^{-1} - \mathrm{Id})g$ et comme $ghg^{-1} \in D(G)$, on a $g\Big((h - \mathrm{Id})(E)\Big) \subset (ghg^{-1} - \mathrm{Id})(E) \subset F$, d'où le résultat.

Corollaire 178. *Soit G un sous-groupe de $GL(E)$ tel que pour tout $A, B, C \in G$, $\sigma(ABC) = \sigma(BAC)$ alors G est triangularisable.*

Remarque: comme $\sigma(AB) = \sigma(BA)$, la propriété de l'énoncé revient à dire que le spectre est permutable, i.e. pour tout $\{A_1, \cdots, A_m\} \subset G$, et pour toute permutation $\sigma \in \mathfrak{S}_m$, on a $\sigma(A_1 A_2 \cdots A_m) = \sigma(A_{\sigma(1)} \cdots A_{\sigma(m)})$. Le lecteur notera bien que cette propriété plus faible que celle de permutabilité de la trace car ici on demande juste que les ensembles de valeurs propres sont égaux, sans tenir compte des multiplicités.

Preuve : Le résultat découle directement de l'implication (iv) $\Rightarrow$ (i) dans le corollaire précédent. En effet pour tout $h \in D(G)$, on a $\sigma(h) = \sigma(\mathrm{Id}) = \{1\}$ de sorte que h est unipotent.

4.5 Noyaux emboités

Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n et $u \in \mathcal{L}(E)$. Pour tout $\lambda \in \mathbb{K}$ et $r \geq 1$, on note

$$K_r(\lambda) := \mathrm{Ker}(u - \lambda \mathrm{Id})^r \quad \text{et} \quad I_r(\lambda) := \mathrm{Im}(u - \lambda \mathrm{Id})^r,$$

et on note $dK_r(\lambda) := \dim_{\mathbb{K}} K_r(\lambda)$ et $dI_r(\lambda) := \dim_{\mathbb{K}} I_r(\lambda)$. On pose aussi $dK_0(\lambda) = 0$ et $dI_0(\lambda) = n$.

Proposition 179. *La suite $dK_r(\lambda)$ (resp. $dI_r(\lambda)$) est tout d'abord strictement croissante (resp. décroissante) puis stationnaire à partir d'un indice r_0 (resp. le même indice r_0). Par ailleurs la suite*

$$\delta_r(\lambda) := dK_r(\lambda) - dK_{r-1}(\lambda)$$

pour $r \geq 1$ est décroissante jusqu'au rang r_0 puis stationnaire égale à 0.

Preuve : Quitte à considérer $u - \lambda \mathrm{Id}$, on suppose $\lambda = 0$ et on note simplement K_r pour $K_r(0)$. Soit alors r tel que $K_r = K_{r+1} \subset K_{r+2}$. Pour $x \in K_{r+2}$ on a $u(x) \in K_{r+1} = K_r$ et donc $u^{r+1}(x) = 0$, i.e. $x \in K_{r+1}$ et donc $K_{r+2} = K_{r+1}$, ce qui montre la première partie de l'énoncé puisqu'avec le théorème du rang $dK_r(\lambda) + dI_r(\lambda) = n$.

En ce qui concerne δ_r, considérons l'endomorphisme $K_r \longrightarrow K_{r-1}/K_{r-2}$ qui envoie x sur l'image de $u(x) \in K_{r-1}/K_{r-2}$. Son noyau est clairement K_{r-1} de sorte qu'on a une injection

$$K_r/K_{r-1} \hookrightarrow K_{r-1}/K_{r-2}$$

et donc $\delta_r \geq \delta_{r-1}$, d'où le résultat.

Proposition 180. *Soit $u \in \mathcal{L}(E)$ dont le polynôme caractéristique est scindé. On note*

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

la décomposition de E en sous-espace caractéristiques. Pour chaque $\lambda \in$ Spec$u$, il existe une base de $E(\lambda)$ dans laquelle la matrice de u est diagonale par blocs du type

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

où la taille des blocs sont données par les longueurs des lignes du tableau de Young, cf. la figure **??**, *associé à u et λ dont les colonnes sont de taille $\delta_r(\lambda)$.*

Une façon de construire ce tableau de Young est la suivante : on prend un vecteur e_1 de $K_r - K_{r-1}$ et on note pour $i = 1, \cdots, r-1$, $e_{i+1} = u^i(e_1)$. Si $\dim K_r/K_{r-1} > 1$, on choisit un vecteur $e_{r+1} \in K_r$ tel que les images de e_1, e_{r+1} dans K_r/K_{r-1} soient libres et on pose pour $i = 1, \cdots, r-1$, $e_{r+1+i} = u^i(e_{r+1})$. On continue le procédé jusqu'a obtenir une base $e_1, e_{r+1}, \cdots, e_{kr+1}$ de K_r/K_{r-1}. On choisit alors un vecteur $e_{(k+1)r+1}$ de K_{r-1} tel que les images de $u(e_1), \cdots, u_(e_{kr+1}), e_{(k+1)r+1}$ forment une famille libre de K_{r-1}/K_{r-2} et on pose pour tout $i = 1, r-2$, $e_{(k+1)r+1+i} = u(e_{(k+1)r+1})$. On continue ce procédé jusqu'à épuiser tous les K_i. Parallèlement on remplit le tableau de Young comme dans la figure **??** dans laquelle l'image de e_1 est une base de $\operatorname{Ker} u^6/\operatorname{Ker} u^5$, les images de $u(e_1), e_7, e_{12}$ forment une base de $\operatorname{Ker} u^5/\operatorname{Ker} u^4$, les images de $u^4(e_1), u^3(e_7), u^3(e_{12}), e_{17}$ forment une base de $\operatorname{Ker} u^2/\operatorname{Ker} u$ et

$$u^5(e_1), u^4(e_7), u^4(e_{12}), u(e_{17}), e_{19}$$

forment une base de $\operatorname{Ker} u$.

Définition 181. *La maison de u est l'ensemble des tableaux de Young de u pour $\lambda \in \operatorname{Spec}(u)$.*

Remarque: sur un corps algébriquement clos, la classe de conjugaison d'un endomorphisme correspond à sa maison au sens de la définition précédente. Une application intéressante de ce résultat est le théorème de Brauer suivant.

Théorème 182. ***(de Brauer)*** *Soit K un corps quelconque et pour $\sigma \in \mathfrak{S}_n$, on note $M(\sigma)$ la matrice de permutation associée, i.e. définie par $M(\sigma)(e_i) = e_{\sigma(i)}$ pour tout $i = 1, \cdots, n$. Alors σ et σ' sont conjugués dans $\mathfrak{S}_n$ si et seulement si $M(\sigma)$ et $M(\sigma')$ sont semblables dans $GL_n(K)$.*

Preuve : Le sens direct est évident puisque si $\sigma' = \tau\sigma\tau^{-1}$ alors $M(\sigma') = PM(\sigma)P^{-1}$ avec $P = M(\tau)$.

Réciproquement supposons que $M(\sigma)$ et $M(\sigma')$ sont semblables. Notons $V^\sigma = \operatorname{Ker}(M(\sigma) - \mathrm{Id})$ l'espace des invariants sous $M(\sigma)$, i.e. l'ensemble des $v = \sum_{i=1}^n \lambda_i e_i$ tels que λ_i ne dépend que de l'orbite de i sous l'action de σ. Ainsi $\dim V^\sigma$ est égal au nombre d'orbites sous σ. Soit alors $\sigma = c_1 \cdots c_r$ la décomposition en cycles à supports disjoints de σ et notons, pour $k = 1, \cdots, n$, $n_k(\sigma)$ le nombre de cycles c_i de longueurs k. On a alors

$$\dim V^\sigma = \sum_{k=1}^n n_k(\sigma) = \sum_{k=1}^n n_k(\sigma') = \dim V^{\sigma'}.$$

De même comme $M(\sigma^r) = M(\sigma)^n$, on a aussi $\dim V^{\sigma^r} = \dim V^{(\sigma')^r}$ et comme, si c est un cycle de longueur k alors c^r s'écrit comme le produit de $k \wedge r$- cycles à supports disjoints tous de même longueur $\frac{k}{k\wedge r}$, on en déduit que pour tout r, on a

$$\sum_{k=1}^n (k \wedge r) n_k(\sigma) = \sum_{k=1}^n (k \wedge r) n_k(\sigma'),$$

ce qui s'écrit matriciellement $SX = SX'$ où $S := (i \wedge j)_{1 \leq i,j \leq n}$ est la matrice des pgcd et X (resp. X') est la matrice colonne des $n_k(\sigma)$ (resp. $n_k(\sigma')$). Soit alors $A = (a_{i,j})_{1 \leq i,j \leq n}$ la matrice définie par $a_{i,j} = 1$ si j divise i et 0 sinon. La relation $\sum_{d|m} \varphi(d) = m$ s'écrit alors matriciellement sous la forme

$$A \mathrm{diag}(\varphi(1), \cdots, \varphi(n))^t A = S$$

de sorte que la matrice S est inversible et donc $X = X'$, i.e. σ et σ' ont « la même » décomposition en cycles à supports disjoints et sont donc conjugués.

4.6 Endomorphismes cycliques

Soit E un K-espace vectoriel de dimension finie.

Définition 183. *Un endomorphisme $f \in \mathcal{L}(E)$ est dit cyclique si et seulement s'il existe $v \in E$ tel que $E = \{P(f)(v) : \ P \in K[X]\}$.*

Remarque: si n est la dimension de E alors $E = \{P(f)(v) : \ P \in K_{n-1}[X]\}$ et comme $(1, X, \cdots, X^{n-1})$ est une base de $K_{n-1}[X]$, on en déduit que $\big(v, f(v), \cdots, f^{n-1}(v)\big)$ est une base de V. Dans cette base la matrice de f est de la forme

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & a_0 \\ 1 & 0 & & \vdots & a_1 \\ 0 & \ddots & \ddots & & \vdots \\ \cdots & & 1 & 0 & a_{n-2} \\ 0 & \cdots & \cdots & 1 & a_{n-1} \end{pmatrix}.$$

On vérifie aisément que le polynôme caractéristique de cette matrice est $P(X) = X^n - a_{n-1}X^{n-1} - \cdots - a_0$ et on dit que la matrice précédente est *la matrice compagnon* de $P(X) = X^n - a_{n-1}X^{n-1} - \cdots - a_0$.

Lemme 184. *Le polynôme minimal d'un endomorphisme cyclique est égal à son polynôme caractéristique.*

Preuve : Comme $v, f(v), \cdots, f^{n-1}(v)$ est une famille libre, tout polynôme Q de degré $\leq n-1$ vérifie alors $Q(f)(v) \neq 0$, de sorte que le polynôme minimal est de degré $\geq n$. Comme par ailleurs il divise le polynôme caractéristique, lequel est de degré n, on en déduit qu'il lui est égal.

Proposition 185. *Soit $g \in \mathcal{L}(E)$ tel que $gf = fg$; il existe alors $P \in K[X]$ tel que $g = P(f)$.*

Remarque: autrement dit le commutant d'un endomorphisme cyclique est $\{P(f) : \ P \in \mathbb{K}[X]/(\pi_f)\}$.

Preuve : On écrit $g(v) = \alpha_0 v + \cdots + \alpha_{n-1} f^{n-1}(v)$ et on pose $Q(X) = \alpha_0 + \cdots + \alpha_{n-1}X^{n-1}$. Pour vérifier l'égalité $g = Q(f)$ il suffit de la vérifier sur la base $(v, f(v), \cdots, f^{n-1}(v))$, soit

$$g(f^i(v)) = f^i(g(v)) = f^i(Q(f)(v)) = Q(f)(f^i(v)),$$

d'où le résultat.

4.7 Invariants de similitude

Soit V un K-espace vectoriel de dimension finie et $f \in \mathcal{L}(E)$. L'idée nouvelle est alors de considérer V muni de l'endomorphisme f, comme un $K[X]$-module comme suit, et d'utiliser le théorème de structure des modules de type fini sur un anneau principal, cf. le §6.3.

Définition 186. *On munit V d'une structure de $K[X]$-module en posant $X.v := f(v)$ et par linéarité pour tout $P \in K[X]$, on a $P.v = P(f)(v)$. On notera V_f l'espace V muni de cette structure de $K[X]$-module.*

Proposition 187. *Deux endomorphismes f et g sont semblables si et seulement si les deux structures de $K[X]$-module induites sur V sont isomorphes, i.e. $V_f \simeq V_g$.*

Preuve : Soit $h \in \mathcal{L}(E)$ tel que $g = hfh^{-1}$ alors pour $v \in V_f$, on a $h(X.v) = hf(v) = g(h(v) = X.h(v)$, i.e. h induit un isomorphisme de $K[X]$-modules : $V_f \simeq V_g$.

Réciproquement si $h : V_f \simeq V_g$ alors $h(X.v) = X.h(v)$, i.e. $h(f(v) = g(h(v))$ et donc $hf = gh$ soit $g = hfh^{-1}$ et donc f et g sont semblables.

De la théorie générale des modules sur un anneau principal, cf. le théorème 265, on en déduit le théorème suivant qu'il faut comprendre comme une décomposition de l'espace en somme directe de sous-espaces cycliques.

Théorème 188. *Il existe une unique suite de polynômes non constants $P_1(X)|\cdots|P_r(X)$ tels que*

$$V_f \simeq K[X]/(P_1) \times \cdots \times K[X]/(P_r). \tag{11}$$

En particulier on a $P_r = \pi_f$ et $P_1 \cdots P_r = \chi_f$.

Exemple : si $V_f \simeq K[X]/(X^n)$ alors la matrice de f dans la base d'image $(1, X, \cdots, X^{n-1})$ est la matrice de Jordan $J_n(0)$. Ainsi si P_r est totalement décomposé, par application du lemme chinois, on passe de la décomposition (11) à celle de Jordan.

Proposition 189. ***(Décomposition de Dunford)*** *Soit $f \in \mathcal{L}(E)$ tel que son polynôme caractéristique est séparable. Alors f s'écrit de manière unique sous la forme $d + n$ où*

— *n est nilpotent et d est semi-simple,*

— *d et n commutent.*

En outre d et n sont des polynômes en f.

Remarque: d'après la définition **??**, un polynôme est séparable si tous ses facteurs irréductibles le sont. Par ailleurs un polynôme irréductible est séparable si et seulement si ses racines sont simples dans son corps de décomposition qui est alors une extension galoisienne du corps de départ. Rappelons enfin qu'en caractéristique nulle tout polynôme est séparable.

Remarque: comme le polynôme caractéristique et le polynôme minimal de f ont les mêmes facteurs irréductibles, ils sont soit tous les deux séparables ou pas.

Preuve : Commençons par l'existence : en utilisant (11) et le lemme chinois, il suffit de traiter le cas où f est cyclique avec $\chi_f = \pi^r$ pour π un polynôme irréductible de $K[X]$. Notons $L = \mathrm{Dec}_K(\pi)$ un corps de décomposition de π sur K de sorte que l'extension L/K est galoisienne. On écrit

$$L[X]/(\pi^r) \simeq L[X]/(X-\lambda_1)^r \times \cdots \times L[X]/(X-\lambda_n)^r,$$

où les λ_i sont les racines de π dans L, i.e. $\pi(X) = \prod_{i=1}^n (X-\lambda_i)$. Sur chaque composante $L[X]/(X-\lambda_i)^r$, on pose $d_i = \lambda_i \mathrm{Id}$ et $n_i = f - \lambda_i$ qui est nilpotent d'ordre r. On note alors, d'après le lemme chinois, P *le* polynôme unitaire de degré $< nr$ tel que $P(X) \equiv \lambda_i \mod (X-\lambda_i)^r$ pour tout $i = 1, \cdots, n$. Soit alors $\sigma \in \mathrm{Gal}(L/K)$ qui permute les λ_i : en particulier on remarque que $\sigma(P)$ vérifie les mêmes congruences de sorte que $\sigma(P) = P$ et donc finalement $P(X) \in K[X]$ et donc la décomposition $f = d + n$ est bien définie sur K où d, et donc aussi n, est un polynôme en f.

Considérons enfin le problème de l'unicité : soit $d' + n' = d + n$ avec $d'n' = n'd'$ et où d, n est défini comme ci-avant, i.e. d et n sont des polynômes en f. Ainsi d' commute avec d et donc d, d', et donc aussi $d - d'$, sont simultanément diagonalisables dans une extension finie. De même n' commute avec n et donc $n - n'$ est nilpotent et semi-simple, puisque égal à $d' - d$ ce qui impose que $d' - d = 0 = n - n'$, i.e. $d = d'$ et $n = n'$, d'où le résultat.

Contre exemple : considérons $K = \mathbb{F}_p((T))$ et le K-espace vectoriel $E = K[X]/(X^p - T)$ muni de l'endomorphisme f défini par la multiplication par X. Notons que $X^p - T$ étant irréductible E est un corps de sorte que pour tout $Q \in K[X]$, l'endomorphisme $Q(f)$ égal donc à la multiplication par $Q(X)$ est soit nulle soit un isomorphisme. Ainsi si l'énoncé précédent était valable, l'endomorphisme nilpotent $n = Q(f)$ est nécessairement nul et donc f serait semi-simple, i.e. diagonalisable dans $E' := K'[X]/(X^p - T)$ où $K' = \mathbb{F}_p((T^{\frac{1}{p}}))$. Or dans K' on a $X^p - T = (X - T^{\frac{1}{p}})^p$ et donc f vu comme endomorphisme de E' admet une unique valeur propre : s'il était diagonalisable ce serait donc une homothétie, ce qui n'est visiblement pas le cas.

Un défaut de la preuve précédente est qu'elle nécessite la connaissance des racines du polynôme minimal π_f de f et n'est donc pas constructive, i.e. on ne peut pas programmer un ordinateur pour calculer la décomposition de Dunford en se basant sur la preuve précédente. On propose à présent, en adaptant la méthode de Newton classique, une construction algorithmique de d en remarquant qu'il doit annuler le polynôme $P(X)$ défini, en caractéristique nulle, par

$$P(X) := \frac{\pi_f(X)}{\pi_f(X) \wedge \pi_f'(X)}.$$

On introduit ainsi la suite $(f_n)_{n\in\mathbb{N}}$ définie par récurrence

$$f_0 = f, \qquad f_{n+1} = f_n - P(f_n)P'(f_n)^{-1}.$$

— Vérifions tout d'abord que cette suite est bien définie, i.e. que $P'(f_n)$ est une matrice inversible. Pour ce faire on raisonne par récurrence et on ajoute à l'hypothèse de récurrence la propriété suivante :

$$P(f_n) = P(f)^{2^n} Q_n(f), \quad Q_n \in K[X],$$

de sorte que $P(f_n)$ est nilpotent puisque par $P(f)$ l'est : en effet pour r plus grand que la plus grande multiplicité d'une racine de π_f, le polynôme P^r est divisible par π_f et donc $P(f)^r$ est nul.

— Au rang $n = 0$, en posant $Q_0 = 1$ on a bien $P(f) = P(f)^{2^0} Q_0(f)$. On écrit ensuite une relation de Bezout entre P et P' qui par hypothèse sont premier entre eux ce qui donne $U(f)P(f)+V(f)P'(f) = \mathrm{Id}$. Comme $P(f)$ est nilpotent, on en déduit que $\mathrm{Id} - U(f)P(f)$ est inversible et donc aussi $P'(f)$.

— Supposons donc le résultat acquis jusqu'au rang n, de sorte que f_{n+1} est bien défini : on notera au passage que puisque $P'(f_n)$ est un polynôme en f_n, on a $P(f_n)P'(f_n)^{-1} = P'(f_n)^{-1}P(f_n)$. On calcule alors $P(f_{n+1})$ en utilisant la formule de Taylor :

$$\begin{aligned} P(f_{n+1}) &= P(f_n) + (f_{n+1} - f_n)P'(f_n) + (f_{n+1} - f_n)^2 Q(f_n) \\ &= (f_{n+1} - f_n)Q(f_n) \\ &= P(f_n)^2 \big(P'(f_n)^{-2} Q(f_n) \end{aligned}$$

qui est donc bien de la forme $P(f)^{2^{n+1}} Q_{n+1}(f)$, d'après l'hypothèse de récurrence et en utilisant que f_n est un polynôme en f.

— Ainsi donc pour n assez grand, $P(f_n)$ est nul et la suite f_n devient stationnaire égale à d qui par construction est un polynôme en f et vérifie $P(d) = 0$. Ainsi d est semi-simple. Par ailleurs on a

$$d - f = (f_n - f_{n-1}) + (f_{n-1} - f_{n-2}) + \cdots + (f_1 - f_0)$$

où chacun des $f_{i+1} - f_i = -P(f_i)P'(f_i)^{-1}$ est un polynôme en f et nilpotent. Ainsi ces endomorphismes nilpotents commutent entre eux deux à deux et leur somme est donc nilpotente, d'où le résultat.

4.8 Sous-espaces stables

Un sous-espace $W \subset V$ est stable par f si et seulement si W est un sous-$K[X]$-module de V_f.

Lemme 190. *Soit f un endomorphisme cyclique. Les sous-espaces stables par f sont les $\Im P(f) = \mathrm{Ker}\, \frac{\chi_f}{P}(f)$ où P décrit les diviseurs de χ_f.*

Remarque: en particulier un endomorphisme cyclique n'admet qu'un nombre fini de sous-espaces stables. On renvoie à l'exercice 49 pour la réciproque.
Preuve : Tout sous-espace stable de $V_f \simeq K[X]/(\chi_f)$ sont ses sous-modules et donc aux idéaux $(P(X))$ pour P un diviseur de χ_f, d'où le résultat.

Définition 191. *Un endomorphisme est dit* indécomposable *si on ne peut pas décomposer l'espace en une somme directe de deux sous-espaces stables stricts. Il est dit* semi-simple *si tout sous-espace stable admet un supplémentaire stable.*

Proposition 192. *Un endomorphisme f est indécomposable si et seulement s'il est cyclique et son polynôme caractéristique est la puissance d'un irréductible.*

Preuve : D'après la décomposition (11) en sous-espaces cycliques, l'endomorphisme doit être nécessairement cyclique. En outre d'après le lemme chinois le polynôme caractéristique doit être la puissance d'un irréductible.

Réciproquement supposons que f est cyclique avec χ_f la puissance d'un irréductible. Si on avait $V = V_0 \oplus V_1$ alors pour $W_0 \simeq K[X]/(P_0)$ (resp. $W_1 \simeq K[X]/(P_1)$) un sous-espace cyclique de V_0, on devrait avoir, d'après le théorème chinois $P_0 \wedge P_1 = 1$ et P_0P_1 qui diviser χ_f ce qui n'est pas.

Proposition 193. *Un endomorphisme f est semi-simple si et seulement s'il est sans facteur carré.*

Preuve : Supposons que f est semi-simple et, par l'absurde que $\pi_f = P^2Q$. Le sous espace $W = \operatorname{Ker} P(u)$ est stable et admet donc un supplémentaire stable W'. Montrons alors que $PQ(f)$ est nul sur W et W' et est donc divisible par π_f ce qui sera contradictoire. Clairement $QP(f)$ s'annule sur W et pour $w' \in W'$, on a $P(f) \circ (PQ(f))(w') = 0$ et donc $PQ(f)(w') \in W$: mais comme W' est stable par f, on a aussi $PQ(f)(w') \in W'$ et donc comme $W \cap W' = \{0\}$, $PQ(f)(w') = 0$.

Supposons à présent que π_f est irréductible de sorte que V est un E-espace vectoriel où $E = K[X]/(\pi_f)$ est un corps. Un sous-espace stable de V est alors un E-espace vectoriel qui admet donc un E-supplémentaire qui est un K-espace vectoriel stable par f et donc f est bien semi-simple.

Soit enfin f tel que $\pi_f = \prod_i P_i$ est sans facteur carré. La décomposition (11) s'écrit aussi sous la forme $V = \bigoplus_i V_i$ où $V_i = \operatorname{Ker} P_i(f)$ est un E_i-espace vectoriel où $E_i = K[X]/(P_i)$ est une extension finie de K. Soit alors W un sous-espace stable de V et soit $W_i = W \cap V_i = \operatorname{Ker} P_i(f)_{|W}$ de sorte que d'après le lemme des noyaux

$$W = \bigoplus_i W_i.$$

Comme $f_{|V_i}$ est semi-simple, on note W_i' un supplémentaire stable de W_i dans V_i de sorte que $W' := \bigoplus_i W_i'$ est un supplémentaire stable de W.

Remarque: sur un corps algébriquement clos, semi-simple est équivalent à la notion plus classique d'endomorphisme diagonalisable au sens de la définition suivante.

Définition 194. *Un endomorphisme est dit* diagonalisable *s'il existe une base de vecteurs propres, i.e. s'il existe une base dans laquelle sa matrice est diagonale.*

Remarque: sur $\mathbb{R}$ un endomorphisme semi-simple est semblable à une matrice diagonale par blocs dont les blocs sont soient de dimension 1 soit de dimension 2 de la forme $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Théorème 195. *L'endomorphisme f est diagonalisable si et seulement si son polynôme minimal est scindé à racines simples.*

Remarque: par exemple si le polynôme caractéristique est scindé à racines simples alors π_f aussi ; plus généralement si P est un polynôme annulateur de f scindé à racines simples alors μ_f le sera aussi, puis que μ_f divise tout polynôme annulateur.

Nous avons vu que dans le cas d'un endomorphisme cyclique f, le commutant de f est $\{Q(f) : \ Q \in K[X]\}$. Considérons le cas $V_f \simeq K[X]/(P) \oplus K[X]/(PQ)$ et l'endomorphisme $g = A_1(f_1) \oplus A_2(f_2)$ où f_1 (resp. f_2) est la restriction de f au premier facteur $K[X]/(P)$ (resp. $K[X]/(PQ)$). On choisit alors A_2 tel que $A_2 \not\equiv A_1 \mod P$. S'il existe $B \in K[X]$ tel que $g = B(f)$ alors $B \equiv A_1 \mod P$ et $B \equiv A_2 \mod PQ$ ce qui n'est pas puisque $A_2 \not\equiv A_1 \mod P$. On a ainsi construit un endomorphisme g commutant avec f mais qui n'est pas un polynôme en f. En particulier on en déduit que si f est tel que son commutant est $\{Q(f) : \ Q \in K[X]\}$ alors f est un endomorphisme cyclique.

Proposition 196. *Avec les notations précédentes, on suppose g commute avec f et que tout sous-espace stable F par f est aussi stable par g. Alors g est de la forme $Q(f)$ pour $Q \in K[X]$.*

Preuve : Considérons la décomposition de

$$V_f = K[X]/(P_1) \oplus K[X]/(P_2) \oplus \cdots \oplus K[X]/(P_r)$$

en sous-espace cyclique. D'après le cas cyclique, comme la restriction g_i de g à chacun des facteurs $V_i := K[X]/(P_i)$ commute à celle f_i de f, on en déduit qu'il existe des polynômes $Q_1, \cdots, Q_r$ tels que $g_i = Q_i(f_i)$ et il s'agit de montrer que $Q_r \equiv Q_i \mod P_i$ pour tout $i = 1, \cdots, r$. Notons v_i un vecteur engendrant l'espace cyclique V_i et soit $w_i = v_i + \frac{P_r}{P_i}(f)(v_r)$ de sorte que l'espace cyclique engendré par w_i est isomorphe à $K[X]/(P_i)$. Comme

par hypothèse g stabilise cet espace cyclique, sa restriction est de la forme $Q(f)$, i.e.

$$\begin{aligned} g(w_i) &= Q(f)(w_i) = Q(f)(v_i) + Q(f)\big(\tfrac{P_r}{P_i}(f)(v_r)\big) \\ &= Q_i(f)(v_i) + Q_r(f)\big(\tfrac{P_r}{P_i}(f)(v_r)\big), \end{aligned}$$

et donc, dans $V_i \oplus V_r$,

$$Q \equiv Q_i \mod P_i \text{ et } Q \equiv Q_r \mod P_i,$$

d'où le résultat.

4.9 Classes de congruences

Définition 197. *Deux matrices carrées A et B de $\mathbb{M}_n(\mathbb{C})$ sont dites congruentes, s'il existe une matrice inversible P telle que $B = P^*AP$.*

La relation de congruence est clairement une relation d'équivalence. Cette relation n'est réellement utilisée que pour des matrices hermitiennes et correspond à l'effet d'un changement de base sur la matrice associée à un produit hermitien, cf. la proposition 202. Dans ce cadre, nous verrons que les classes sont paramétrées par la signature, cf. le théorème 231 dans le cas réel et 240 dans le cas hermitien. En revanche contrairement au cas des classes de similitudes, on a un énoncé de congruence simultanée suivant.

Proposition 198. *Soit deux matrices hermitiennes A et B avec B définie positives. Il existe alors une matrice inversible P vérifiant*

$$P^*AP = D \qquad et \qquad P^*BP = I_n,$$

où D est une matrice diagonale.

Remarque: on notera bien que la matrice P n'est pas unitaire de sorte que A et B ne sont pas simultanément diagonalisable !

Preuve : On note Q une matrice dont les vecteurs colonnes forment une base orthonormée pour le produit hermitien défini par B, i.e. $(X, Y) \mapsto X^*BY$. On a ainsi $Q^*BQ = I_n$. Comme Q^*AQ est visiblement hermitienne, on la diagonalise en base orthonormée pour le produit hermitien canonique, i.e. il existe une matrice unitaire R telle que $R^*(Q^*AQ)R = D$. On a alors $R^*(Q^*BQ)R = R^*R = I_n$, de sorte que la matrice $P = QR$ convient.

4.10 Classes de similitudes unitaires

On s'intéresse ici aux classes de similitudes unitaires. Notons tout d'abord que si A et B sont dans la même classe de similitude unitaire alors

$$\sum_{i,j=1}^{n} |b_{ij}|^2 = \sum_{i,j=1}^{n} |a_{i,j}|^2$$

En effet cela découle de l'égalité $\sum_{i,j=1}^{n} |a_{i,j}|^2 = \mathrm{tr}(A^*A)$. Plus généralement étant donnés un mot $M(s,t) = s^{m_1}t^{n_1}s^{m_2}t^{n_2}\cdots s^{m_k}t^{n_k}$ en deux variables s,t avec $m_1, n_1, \cdots, m_k, n_k \geq 0$: le degré de $M(s,t)$ est par définition égal à $m_1 + n_1 + \cdots + m_k + n_k$. Pour $A \in \mathbb{M}_n(\mathbb{C})$, on pose

$$M(A, A^*) = A^{m_1}(A^*)^{n_1}\cdots A^{m_k}(A^*)^{n_k}$$

On remarque alors que pour tout $M(s,t)$, $\mathrm{tr}M(A, A^*)$ est constant sur la classe de similitude unitaire.

Théorème 199. ***(Specht)*** *Deux matrices A, B sont unitairement semblables si et seulement si $\mathrm{tr}M(A, A^*) = \mathrm{tr}M(B, B^*)$ pour tout mot $M(s,t)$.*

L'inconvénient évident du théorème de Specht est qu'il faut vérifier une infinité de conditions. Le théorème de Pearcy suivant permet de se ramener à un nombre fini.

Théorème 200. *(Pearcy [**?**]) Deux matrices A, B sont unitairement semblables si et seulement si $\mathrm{tr}M(A, A^*) = \mathrm{tr}M(B, B^*)$ pour tout mot $M(s,t)$ de degré au plus $2n^2$.*

Remarque: un rapide calcul montre qu'il y a au plus 4^{n^2} mots distincts de degré au plus $2n^2$. Pour $n = 2$, il est facile de montrer qu'il suffit en fait de considérer les mots s, s^2 et ts. Pour $n = 3$, on peut montrer qu'il suffit de considérer les 9 mots, $s, s^2, ts, ts^2, t^2s^2, tsts, ts^2ts, ts^2t^2s$ au lieu de tous les mots de degré au plus 18.

5 Algèbre bilinéaire

5.1 Formes sesquilinéaires : généralités

Rappelons qu'étant donné un automorphisme σ du corps $\mathbb{K}$, par exemple la conjugaison complexe de $\mathbb{C}$, *une application semi-linéaire* est une application θ telle que pour tout $x, y \in E$ et $\lambda \in \mathbb{K}$ on a

$$\theta(x + \lambda y) = \theta(x) + \lambda^\sigma \theta(y)$$

où par convention on note λ^σ pour $\sigma(\lambda)$.

Définition 201. *On appelle* forme σ-sesquilinéaire *toute application $\phi : E \times E \to \mathbb{K}$ vérifiant les conditions suivantes :*

- *pour tout $x \in E$, l'application $\phi_x : y \in E \mapsto \phi(x,y)$ est linéaire ;*
- *pour tout $y \in E$ l'application $\phi_y : x \in E \mapsto \phi(x,y)$ est σ-linéaire.*

Remarque: les notations ϕ_x et ϕ_y ne sont pas exemplaires, on veillera à ne pas se mélanger.

Proposition 202. *Soit E un espace vectoriel muni d'une base $(e_i)_{1\leq i\leq n}$. On note $A_\phi = \big(\phi(e_i,e_j)\big) \in \mathbb{M}_n(\mathbb{K})$ la matrice associée à la forme sesquilinéaire ϕ, relativement à la base $(e_i)_i$. Pour tous vecteurs $x, y \in E$ de vecteur colonne coordonnées X et Y, on a alors*

$$\phi(x,y) = {}^tX^\sigma A_\phi Y.$$

Remarque: si $P_{(e_i)\leftarrow(e'_i)}$ est la matrice de changement de base de $(e_i)_i$ à $(e'_i)_i$ alors la matrice A'_ϕ relativement à cette nouvelle base est telle que

$$A'_\phi = {}^tP^\sigma_{(e_i)\leftarrow(e'_i)_i} A_\phi P_{(e_i)\leftarrow(e'_i)_i}.$$

En particulier le déterminant de A_ϕ, qu'on appelle le discriminant de ϕ, n'est défini que comme élément de $\mathbb{K}/N(\mathbb{K})$ où $N(\mathbb{K}) = \{\lambda\lambda^\sigma,\ \lambda\in\mathbb{K}\}$.

Définition 203. *Pour M une partie de E, on note*

$$M^\perp = \{y\in E,\ \phi(M,y)=0\}, \qquad {}^\perp M = \{x\in E,\ \phi(x,M)=0\}.$$

On dit que $M^\perp$ (resp. ${}^\perp M$) est l'orthogonal à droite (resp. à gauche) de M.

Remarque: $M^\perp$ et ${}^\perp M$ sont clairement des sous-espaces de E.

Lemme 204. *Si M est un sous-espace de E alors*

$$\dim M + \dim M^\perp = \dim E + \dim(M\cap {}^\perp E).$$

Preuve : Soit $f : E \longrightarrow E^*$ l'application semi-linéaire qui à x associe $y\mapsto \phi(x,y)$. Par définition $M^\perp$ est l'orthogonal au sens usuel, cf. la proposition 129, de $f(M)$ de sorte que

$$\dim M^\perp = \dim E - \dim f(M) = \dim E - (\dim M - \dim(M\cap {}^\perp E))$$

puisque ${}^\perp E = \operatorname{Ker} f$.
Remarque: si on applique la formule à $M = E$, on obtient $\dim E^\perp = \dim {}^\perp E$.

Définition 205. *On dit que ϕ est* non dégénérée *si $E^\perp = \{0\}$ (resp. ${}^\perp E = \{0\}$. Le rang de A_ϕ est appelé* le rang *de ϕ, il est égal à la codimension de $E^\perp$ et ${}^\perp E$.*

Remarque: lorsque ϕ est non dégénérée, l'application $x\mapsto \phi_x$ induit un isomorphisme semi-linéaire *canonique* de E vers son dual E^*.

Lemme 206. *On a ${}^\perp(M^\perp) = M + {}^\perp E$.*

Remarque: En particulier si ϕ est non dégénérée, on retrouve la propriété habituelle de bidualité ${}^\perp(M^\perp) = M$.

Preuve : On a clairement $M + {}^{\perp}E \subset {}^{\perp}(M^{\perp})$ de sorte qu'il suffit de montrer l'égalité des dimensions. En raisonnant comme dans le lemme précédent, on a

$$\dim N + \dim {}^{\perp}N = \dim E + \dim(N \cap E^{\perp}).$$

On applique la formule à $N = M^{\perp}$ de sorte qu'en remarquant que $E^{\perp} \subset M^{\perp}$, on obtient

$$\dim {}^{\perp}(M^{\perp}) = \dim E - \dim M^{\perp} + \dim E^{\perp} = \dim M + \dim {}^{\perp}E - \dim(M \cap {}^{\perp}E)$$

car $\dim E^{\perp} = \dim {}^{\perp}E$ d'après ce qui précède, et on reconnait la formule de Grassman qui donne la dimension de $M + {}^{\perp}E$.

Définition 207. *Une forme σ-sesquilinéaire est dite* réflexive *si pour tout $x, y \in E$, $\phi(x, y) = 0$ équivaut à $\phi(y, x) = 0$.. Elle est dite hermitienne (resp. antihermitienne) si $\phi(x, y) = \epsilon\Big(\phi(y, x)\Big)^{\sigma}$ avec $\epsilon = 1$ (resp. $\epsilon = -1$).*

Remarque: pour une forme hermitienne ou antihermitienne, σ est nécessairement une involution ; dans le cas antihermitien en caractéristique différente de 2, on a même $\sigma = \mathrm{Id}$ et on dit simplement que ϕ est *anti-symétrique.*

Proposition 208. *On suppose que ϕ est une forme hermitienne ou antihermitienne non dégénérée. Pour tout $u \in \mathcal{L}(E)$, il existe un unique endomorphisme $u^* \in \mathcal{L}(E)$, appelé* adjoint *de u, tel que pour tous $x, y \in E$:*

$$\phi(u(x), y) = \phi(x, u^*(y)).$$

En outre on a aussi $\phi(x, u(y)) = \phi(u^*(x), y)$.

Remarque: la donnée d'une forme σ-sesquilinéaire non dégénérée, induit un isomorphisme semi-linéaire canonique entre E et E^* de sorte que l'adjoint habituel d'un endomorphisme vu dans $\mathcal{L}(E^*)$ se voit comme un endomorphisme de E. L'égalité $\phi(u(x), y) = \phi(x, u^*(y))$ de la proposition est alors une simple traduction de l'isomorphisme entre $\mathcal{L}(E^*)$ et $\mathcal{L}(E)$ induit par ϕ.
Preuve : Considérons pour y fixé, la forme linéaire $x \mapsto \phi(y, u(x))$; comme ϕ est non dégénérée, il existe un vecteur dépendant de y que l'on note $u^*(y)$ telle que pour tout x on ait $\phi(u^*(y), x) = \phi(y, u(x))$ et donc aussi, en utilisant la nature hermitienne de ϕ, $phi(u(x), y) = \phi(x, u^*(y))$. Enfin on vérifie aisément que $y \mapsto u^*(y)$ est linéaire.

On suppose à présent que ϕ est une forme hermitienne ou antihermitienne, auquel cas la caractéristique est en outre supposée différente de 2.

Définitions 209.
- *Un vecteur x de E est dit* isotrope *si $\phi(x, x) = 0$.*
- *Un sous-espace F de E est dit* isotrope *si $F \cap F^{\perp} \neq \{0\}$.*
- *Un sous-espace F de E est dit* totalement isotrope *et on écrit* SETI, *si $F \subset F^{\perp}$.*

— *Un séti est dit maximal et on écrit* SETIM*, si pour tout SETI G contenant F alors $G = F$.*

Remarque: comme on est en dimension finie, tout SETI est contenu dans un SETIM.
Remarque: si F est non isotrope alors $E = F \oplus F^{\perp}$.

Proposition 210. *Tous les SETIM ont la même dimension appelée* indice de ϕ.

Preuve : Soient U et V deux SETIM et introduisons M et N qui sont respectivement des supplémentaires de $U \cap V$ dans U et V. On suppose par l'absurde $\dim U > \dim V$ de sorte que $r := \dim M > s := \dim N$. Pour $f_1, \cdots, f_s$ une base de N, considérons l'application $M \longrightarrow \mathbb{K}^s$ définie par

$$m \mapsto \big(\phi(f_1, m), \cdots, \phi(f_s, m)\big).$$

Son noyau est $M \cap N^{\perp}$ de sorte que d'après le théorème du rang

$$\dim(M \cap N^{\perp}) = \dim M - \dim \Im \geq r - s > 0.$$

Considérons alors un vecteur non nul $x \in M \cap N^{\perp} \subset U$. Montrons que $x \in V^{\perp}$: soit donc $v \in V$ que l'on décompose $v = u + n$. On a alors $\phi(x, v) = \phi(x, u) + \phi(x, n)$ avec $\phi(x, u) = 0$ car U est un SETI et $\phi(x, n) = 0$ car $x \in N^{\perp}$. Considérons alors $W = V + \mathbb{K}x$: pour tout $w = v + \lambda x$ on a

$$\phi(w, w) = \phi(v, v) + \lambda^2 \phi(x, x) + \bar{\lambda}\phi(v, x) + \lambda\phi(x, v),$$

où $\phi(v, v) = 0$ (resp. $\phi(x, x) = 0$) car V (resp. U) est un SETIM, et $\phi(v, x) = \phi(x, v) = 0$ car $x \in V^{\perp}$. Ainsi donc W est un SETI contenant strictement V alors que ce dernier était supposé maximal, d'où la contradiction.

Proposition 211. *Si ϕ est non dégénérée, il existe alors une décomposition dite de Witt de l'espace $E = F \oplus F' \oplus G$ avec F, F' des sétim et G un sous-espace non isotrope telle que la matrice de ϕ dans une base adaptée soit de la forme*

$$\begin{pmatrix} 0 & I_r & 0 \\ \epsilon I_r & 0 & 0 \\ 0 & 0 & B \end{pmatrix}.$$

Remarque: on donnera une preuve plus loin de ce résultat dans le cas où $\sigma = \mathrm{Id}$, i.e. pour les formes quadratiques.

5.2 Endomorphismes remarquables

On suppose à présent E muni d'une formes hermitienne non dégénérée en caractéristique différente de 2.

Définition 212. *Un automorphisme u de E est dit* unitaire *si pour tous $x, y \in E$:*

$$\phi(u(x), u(y)) = \phi(x, y).$$

L'ensemble des automorphismes unitaires est un sous-groupe de $GL(E)$ noté U_ϕ et appelé le groupe unitaire *de ϕ. Le noyau du morphisme déterminant d'image $\mathbb{H} = \{\lambda \in \mathbb{K} : \ \lambda\lambda^\sigma = 1\}$, est noté SU_ϕ et s'appelle* le groupe spécial unitaire *de ϕ.*

Remarque: u est unitaire si et seulement si $u^{-1} = u^*$. Matriciellement la matrice U de u est unitaire si et seulement si ${}^tU^\sigma A_\phi U = A_\phi$. Dans le cas où $A_\phi = I_n$, on retrouve la condition usuelle ${}^tU^\sigma U = I_n$.

Définition 213. *Une similitude de rapport $\lambda \in \mathbb{K}^\times$ est un automorphisme tel que pour tous $x, y \in E$:*

$$\phi\big(u(x), u(y)\big) = \lambda\phi(x, y).$$

Le groupe des similitudes se note GU_ϕ.

Remarque: une définition classique d'une similitude consiste à demander :

$$\phi(x, y) = 0 \Rightarrow \phi\big(u(x), u(y)\big) = 0.$$

Définition 214. *Un endomorphisme u est dit* auto-adjoint *s'il vérifie $u = u^*$.*

Remarque: dans le cas réel (resp. complexe) on dit aussi *symétrique* (resp. *hermitien*).

Définition 215. *Un endomorphisme u d'un espace hermitien est dit* normal *si $u \circ u^* = u^* \circ u$.*

Théorème 216. *Un endomorphisme normal est semi-simple.*

Preuve : Nous allons montrer plus précisément que si F est stable par u normal alors $F^\perp$ est aussi u-stable. Considérons une base orthonormée de F que l'on complète en une base orthonormée de $F^\perp$. La matrice de u dans cette base est de la forme $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ et on calcule

$$\begin{aligned} MM^* &= \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}\begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix} = \begin{pmatrix} AA^* + BB^* & BC^* \\ CB^* & CC^* \end{pmatrix} \\ M^*M &= \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix}\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \begin{pmatrix} A^*A & A^*B \\ B^*A & B^*B + C^*C \end{pmatrix} \end{aligned}$$

ce qui fournit les égalités

$$\begin{cases} AA^* + BB^* = A^*A \\ BC^* = A^*B \\ CC^* = B^*B + C^*C. \end{cases}$$

Comme $\mathrm{tr}(AA^*) = \mathrm{tr}(A^*A)$ on en déduit que $\mathrm{tr}(BB^*) = 0$ et donc $B = 0$ d'où le résultat
Remarque: pour f normal, on a $||f(x)|| = ||f^*(x)||$ pour tout x et donc $\mathrm{Ker}\, f = \mathrm{Ker}\, f^*$. Plus généralement si λ est une valeur propre de f, en utilisant que $f - \lambda\mathrm{Id}$ est aussi normal, on en déduit que le sous-espace propre $\mathrm{Ker}(f - \lambda\mathrm{Id})$ est égal au sous-espace propre $\mathrm{Ker}(f^* - \bar{\lambda}\mathrm{Id})$. Leur orthogonal commun est donc stabilisé à la fois par f et f^*. Ainsi dans le cas où le corps est algébriquement clos, on peut utiliser ces arguments pour prouver que f et f^* sont simultanément diagonalisables.

5.3 Formes quadratiques

On reprend les définitions du paragraphe précédent dans le cas où $\sigma = \mathrm{Id}$, on parle alors de forme bilinéaire symétrique ϕ et $q(x) := \phi(x, x)$ est *une forme quadratique*. En caractéristique différente de deux, la formule de polarisation

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

permet d'identifier les formes bilinéaires symétriques et les formes quadratiques. On garde le vocabulaire du paragraphe précédent avec les notions d'isotropie, de forme non dégénérée permettant d'identifier canoniquement l'espace E avec son dual E^*. La traduction matricielle de $\phi(x, y)$ est ${}^tXAY$ et le changement de base s'exprime $A' = {}^tPAP$.

Définition 217. *Étant donné deux formes bilinéaires symétriques (V, ϕ) et (V', ϕ'), la somme orthogonale $(V, \phi) \perp (V', \phi')$ est la forme bilinéaire symétrique $\phi \perp \phi'$ définie sur $V \oplus V'$ par la formule*

$$(\phi \perp \phi')(x + x', y + y') = \phi(x, y) + \phi'(x', y'), \qquad \forall x, y \in V,\ \forall x', y' \in V'.$$

Remarque: si V est muni d'une forme bilinéaire symétrique ϕ et de deux sous-espaces U, W tels que $V = U \oplus W$ et $\phi(u, w) = 0$ pour tout $u \in U$ et $w \in W$, alors

$$(V, \phi) \simeq (U, \phi_{|U}) \perp (W, \phi_{|W}).$$

En particulier on a $(V, \phi) \simeq (V^\perp, 0) \perp (V/V^\perp, \phi')$ avec ϕ' non dégénérée.

Lemme 218. *Soit (V, ϕ) une forme bilinéaire symétrique et W un sous-espace de V tel que $b_{|W}$ est non dégénérée. On a alors*

$$(V, \phi) = (W, \phi_{|W}) \perp (W^\perp, b_{|W^\perp}).$$

Preuve : Commençons par montrer que $V = W \oplus W^\perp$. Comme ϕ_W est non dégénérée, on a $W \cap W^\perp = \{0\}$. Soit alors $v \in V$ et $f_v \in W^*$ défini par

$$f_v : w \in W \mapsto \phi(w, v).$$

Comme $\phi_{|W}$ est non dégénérée l'application $w \in W \mapsto f_w \in W^*$ est un isomorphisme et il existe donc $w \in W$ tel que $f_v = f_w$ et donc $v - w \in W^\perp$ d'où l'affirmation. L'isomorphisme de l'énoncé découle alors de la remarque précédente.

Notation 5. *Pour $a_1, \cdots, a_n \in K^\times$, on note $\langle a_1, \cdots, a_n \rangle$ la forme bilinéaire symétrique sur K^n définie par*

$$(x, y) \in K^n \mapsto \sum_{i=1}^{n} a_i x_i y_i \in K.$$

Remarque: la matrice de $\langle a_1, \cdots, a_n \rangle$ dans la base canonique est la matrice diagonale $\mathrm{diag}(a_1, \cdots, a_n)$. Comme les a_i sont non nuls, cette matrice est inversible et $\langle a_1, \cdots, a_n \rangle$ est non dégénérée.

Théorème 219. *Toute forme bilinéaire symétrique est équivalente à une forme $\langle a_1, \cdots, a_r \rangle$ pour des $a_i \in K^\times$ et où r est le rang de la forme.*

Preuve : On raisonne par récurrence sur la dimension de l'espace ; le cas d'une droite étant évident. La forme bilinéaire ϕ étant non nulle, de la formule de polarisation, on en déduit l'existence d'un vecteur v tel que $q(v) \neq 0$ de sorte que la restriction de ϕ à $W = K.v$ est non dégénérée. D'après le lemme précédent, on a $(V, \phi) = (W, \phi_{|W}) \perp (W^\perp, \phi_{W^\perp})$ et on conclut par récurrence.
Remarque: l'énoncé signifie qu'il existe une base $(e_1, \cdots, e_n)$ telle que $q(e_i) = a_i$ où pour $i > r$ on a posé $a_i = 0$. Notons $l_1, \cdots, l_n$ la base duale associée, on a alors

$$q(x) = a_1 l_1(x)^2 + \cdots + a_r l_r(x)^2.$$

L'algorithme de décomposition de Gauss permet de trouver directement les formes linéaires indépendantes $l_1, \cdots, l_r$. Notons en outre que changer e_i en λe_i modifie a_i par $\lambda^2 a_i$, de sorte que les a_i ne sont à priori bien définis que dans $K^\times / K^{\times,2}$. Pour autant ce ne sont pas nécessairement des invariants au sens où $\langle a_1, \cdots, a_n \rangle$ pet être isomorphe à $\langle a'_1, \cdots, a'_n \rangle$ même, modulo permutations, les $a_i a'_i \notin K^{\times,2}$. Pour

- $K = \mathbb{C}$, comme tout élément non nul est un carré, $\phi \simeq \langle 1, \cdots, 1 \rangle$ où le nombre de 1 est égal au rang de ϕ, qui est bien un invariant.
- Pour $K = \mathbb{R}$, on a $\phi \simeq \langle 1, \cdots, 1, -1, \cdots, -1 \rangle$. On note r (resp. s) le nombre 1 (resp. de -1) ; le couple (r, s) s'appelle la signature de ϕ. Pour montrer que ce sont des invariants, il suffit par exemple de noter que r (resp. s) est la dimension maximale d'un sous-espace sur lequel la restriction de ϕ est définie positive (resp. négative).
- Sur un corps fini $K = \mathbb{F}_q$, en dimension n on dispose de deux classes de formes bilinéaires symétriques non dégénérées à savoir $\langle 1, \cdots, 1 \rangle$ et $\langle 1, \cdots, 1, \alpha \rangle$ où α n'est pas un carré de $\mathbb{F}_q^\times$. Pour le démontrer on raisonne par récurrence sur n, le cas $n = 1$ étant évident. Soit

alors $n \geq 2$; pour tout $\alpha, \beta \in \mathbb{F}_q^\times$, les ensembles $\{\alpha x^2 : \ x \in \mathbb{F}_q\}$ et $\{1 - \beta y^2 : \ y \in \mathbb{F}_q\}$ sont de cardinal $\frac{q+1}{2}$ et ne peuvent pas être disjoints. Ainsi l'équation $\alpha x^2 + \beta y^2 = 1$ admet une solution, i.e. il existe v tel que $q(v) = 1$. On applique alors l'hypothèse de récurrence à $(Kv)^\perp$.

Définition 220. *Étant donné un vecteur v tel que $q(v) \neq 0$, la réflexion τ_v par rapport à $(K.v)^\perp$ est définie par*

$$\tau_v(x) = x - 2\frac{\phi(x,v)}{q(v)}v.$$

Pour$x = x_v + x' \in (Kv) \oplus (Kv)^\perp$, on a $\tau_v(x) = -x_v + x'$ et τ_v est une isométrie relativement à q, i.e. $\phi(x,y) = \phi(\tau_v(x), \tau_v(y))$.

Lemme 221. *Soit ϕ une forme bilinéaire symétrique sur un K-espace vectoriel V et soient $x, y \in V$ tels que $q(x) = q(y) \neq 0$. Il existe alors une isométrie τ de V pour ϕ telle que $\tau(x) = y$.*

Preuve : Soit $u = \frac{x+y}{2}$ et $v = \frac{x-y}{2}$ avec donc $x = u + v$ et $y = u - v$. Comme $q(x) = q(y)$, on calcule $\phi(u, v) = 0$ et $0 \neq q(x) = q(u) + q(v)$ de sorte que, quitte à échanger le rôle de x et y, on peut supposer que $q(v) \neq 0$. La réflexion τ_v de la définition précédente vérifie alors $\tau_v(x) = y$ d'où le résultat.

Théorème 222. ***(de Witt)*** *Soient (V_1, ϕ_1), (V_2, ϕ_2) et (V, ϕ) trois espaces munis d'une forme bilinéaire symétrique avec ϕ non dégénérée. Alors*

$$(V_1, \phi_1) \perp (V, \phi) \simeq (V_2, \phi_2) \perp (V, \phi) \Leftrightarrow (V_1, \phi_1) \simeq (V_2, \phi_2).$$

Preuve : Comme ϕ est non dégénérée, elle est diagonalisable et il suffit donc de traiter le cas où $(V, \phi) = (Kv, \langle a \rangle)$. Pour $i = 1, 2$, on note $v_i = (0, v) \in V_i \perp V$ et $v_2' = f(v_1)$ où $f : (V_1, \phi_1) \perp (Kv, \langle a \rangle)) \simeq (V_2, \phi_2) \perp (Kv, \langle a \rangle))$. On applique alors le lemme précédent à v_2 et v_2' avec $\tau(v_2') = v_2$ de sorte que $\tau \circ f$ envoie v_1 sur v_2 et induit donc un isomorphisme entre leurs orthogonaux i.e. entre (V_1, ϕ_1) et (V_2, ϕ_2) d'où le résultat.

Proposition 223. *Soit V un K-espace vectoriel de dimension 2 muni d'une forme bilinéaire symétrique ϕ non dégénérée. Les propriétés suivantes sont équivalentes :*

(i) ϕ est isotrope ;
(ii) $\det \phi = -1 \in K^\times / K^{\times,2}$ *;*
(iii) $\phi \simeq \langle 1, -1 \rangle$ *;*
(iv) il existe une base dans laquelle la matrice de ϕ est $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Preuve : Nous allons suivre la chaine d'implications (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) $\Rightarrow$ (iv) $\Rightarrow$ (iii). Les seule qui ne sont pas évidentes sont (ii) $\Rightarrow$ (i) $\Rightarrow$ (iv).

- Montrons donc (ii) $\Rightarrow$ (i) : on a $\phi\langle\alpha,\beta\rangle$ avec $\alpha,\beta \in K^\times$ et $\alpha\beta = -1 \in K^\times/K^{\times,2}$. On a donc $\phi \simeq \langle\alpha,-\alpha\rangle$ qui est bien isotrope puisque $q(e_1+e_2)=0$.

- Montrons (i) $\Rightarrow$ (iv) : soit alors $x \neq 0$ tel que $q(x)=0$. Comme ϕ est non dégénérée, il existe y tel que $\phi(x,y) \neq 0$ et $\phi(x,x+\lambda y) = \lambda\phi(x,y)$. On peut donc choisir y tel que $\phi(x,y)=1$ avec donc nécessairement y non colinéaire à x. On note $\alpha = q(y)$ et on pose $z = y - \frac{\alpha}{2}x$ de sorte que $q(z)=0$ et $\phi(x,z)=1$. Dans la base (x,z), la matrice de ϕ est celle de (iv).

Définition 224. *Une forme qui satisfait l'une des propriétés équivalentes de la proposition précédente est appelée* plan hyperbolique *: on le note H. Une base telle que la matrice de ϕ soit comme dans (iv) est dite hyperbolique. On appelle enfin* forme hyperbolique *une somme orthogonale de plans hyperboliques.*

Lemme 225. *Soit ϕ une forme bilinéaire symétrique non dégénérée sur un K-espace vectoriel V. Les assertions suivantes sont équivalentes :*

(i) la forme ϕ est hyperbolique ;

(ii) il existe un sous-espace W de V avec $2\dim W = \dim V$ avec $\phi_{|W} = 0$;

(iii) il existe un sous-espace W de V tel que $W^\perp W$.

Preuve : - (i) $\Rightarrow$ (ii) : notons $(e_1,\cdots,e_{2n})$ la base de V dans laquelle $\phi \simeq \langle 1,-1,\cdots,1,-1\rangle$. Le sous-espace vectoriel W engendré par $e_{2i-1}+e_{2i}$ pour $i=1,\cdots,n$ vérifie alors (ii).

- L'implication (ii) $\Rightarrow$ (iii) vérifions (iii) $\Rightarrow$ (i). Soit $(e_1,\cdots,e_n)$ une base de W. Pour tout $i=2,\cdots,n$, le sous-espace $(Ke_1)^\perp$ est de dimension $2n-1$ et contient W : d'après le théorème du rang il existe alors y orthogonal à $e_2,\cdots,e_n$ et n'appartenant pas à W. D'après la proposition précédente, le plan H engendré par e_1 et y est alors hyperbolique et $W \cap H^\perp$ vérifie (iii) relativement à $H^\perp$. On conclut alors par récurrence sur la dimension.
Remarque: en particulier $(V,\phi) \perp (V,-\phi)$ est une forme hyperbolique puisque le sous-espace $W = \{(v,v) : v \in V\}$ vérifie le point (ii) du lemme précédent.

Lemme 226. *Soit ϕ une forme hyperbolique et ψ une forme non dégénérée. Alors $\phi\otimes\psi$ est hyperbolique.*

Preuve : Par hypothèse $\phi = \langle 1,-1\rangle \perp \cdots \perp \langle 1,-1\rangle$ de sorte que

$$\phi\otimes\psi \simeq (\psi \perp -\psi) \perp \cdots \perp (\psi \perp -\psi)$$

et le résultat découle de la remarque précédente.

Proposition 227. ***(Décomposition de Witt)*** *Soit ϕ une forme non dégénérée isotrope. Il existe alors une forme non dégénérée anisotrope (V_a, ϕ_a) et une forme hyperbolique (V_h, ϕ_h) telles que*

$$(V, \phi) \simeq (V_h, \phi_h) \perp (V_a, \phi_a).$$

Une telle décomposition est en outre unique à isomorphisme près.

Remarque: autrement dit pour étudier une forme bilinéaire symétrique non dégénérée, on se ramène aux formes anisotropes.

Preuve : Si ϕ est anisotrope il n'y a rien à faire. Sinon soit $x \neq 0$ tel que $q(x) = 0$ puis, cf. la preuve de la proposition précédente, y tel que $\phi(x, y) = 1$. Comme (x, y) est nécessairement libre, ils engendrent un plan hyperbolique W et la restriction ϕ' de ϕ à $W^\perp$ est non dégénérée ce qui permet d'itérer la construction jusqu'à obtenir une forme anisotrope.

Le fait que la décomposition soit unique à isomorphisme près découle du théorème de Witt 222 et du fait qu'une forme hyperbolique est de la forme $H \perp \cdots \perp H$.

Définition 228. *Le nombre de facteurs H dans*

$$(V_h, \phi_h) \simeq H \perp \cdots \perp H$$

est appelé l'indice de Witt *de ϕ.*

Notation 6. *Soit $\phi \simeq \langle a_1, \cdots, a_n \rangle$ et $\phi' \simeq \langle a'_1, \cdots, a'_m \rangle$ des formes bilinéaires symétriques non dégénérées. On note $\phi \otimes \phi'$ la forme bilinéaire symétrique non dégénérée isomorphe à*

$$\langle a_1 a'_1, \cdots, a_1 a'_m, a_2 a'_1, \cdots, a_n a'_m \rangle.$$

Remarque: la notation ci-dessus est un artifice pour ne pas avoir à définir canoniquement le produit tensoriel de deux espaces vectoriels.

Définition 229. *On note $\mathcal{M}_K$ l'ensemble des classes d'isomorphismes de formes bilinéaires symétriques non dégénérées sur K et on le munit de la relation d'équivalence*

$$\phi \sim \phi' \Leftrightarrow \phi_a \simeq \phi'_a.$$

On note $W(K)$ l'ensemble quotient.

Remarque: on vérifie aisément que $W(K)$ muni de la somme directe orthogonale $\perp$ et du produit tensoriel $\otimes$ est un anneau commutatif.

Exemples :

— dans $\mathbb{C}$, -1 est un carré et donc $\langle 1, \cdots, 1 \rangle \simeq \langle 1, -1, 1 \cdots \rangle$ qui est une forme hyperbolique si et seulement si la dimension de l'espace est pair. Ainsi $W(\mathbb{Z}) \simeq \mathbb{Z}/2\mathbb{Z}$ où la flèche est donnée par la dimension de l'espace modulo 2.

— Le même raisonnement dans le cas $K = \mathbb{R}$, montre que $W(\mathbb{R}) \simeq \mathbb{Z}$ où la flèche est donnée par $r - s$ où (r, s) est la signature.
— Pour $K = \mathbb{F}_p$ où $p \equiv 1 \mod 4$ comme -1 est un carré, on se ramène comme précédemment aux formes suivantes

$$\langle 1\rangle,\ \langle \alpha\rangle,\ \langle 1, \alpha\rangle,\ H$$

où $\alpha \in \mathbb{F}_p^\times - \mathbb{F}_p^{\times,2}$. Ces quatre formes sont distinctes dans $W(\mathbb{F}_p)$ et on vérifie aisément qu'elles sont toutes d'ordre 2

$$\langle 1\rangle \perp \langle 1\rangle = H \sim 0, \quad \langle \alpha\rangle \perp \langle \alpha\rangle \simeq 0, \quad \langle 1, \alpha\rangle \perp \langle 1, \alpha\rangle = \langle 1, 1\rangle \perp \langle \alpha, \alpha\rangle \sim 0.$$

Ainsi $W(\mathbb{F}_p) \simeq (\mathbb{Z}/2\mathbb{Z})^2$.
— Terminons par le $K = \mathbb{F}_p$ avec $p \equiv 3 \mod 4$, on obtient alors les formes

$$\langle 1\rangle,\ \langle -1\rangle,\ \langle 1, 1\rangle,\ H = \langle 1 - 1\rangle$$

qui sont distinctes dans $W(\mathbb{F}_p)$. Or $\langle 1\rangle \perp \langle 1\rangle = \langle 1, 1\rangle \not\sim H$ et donc $W(\mathbb{F}_p) \simeq \mathbb{Z}/4\mathbb{Z}$.

5.4 Le cas réel

Dans ce cas on a nécessairement $\sigma = \mathrm{Id}$ et on revient donc sur les formes quadratiques du paragraphe précédent avec les notions de positivité.

Définition 230. *Une forme bilinéaire symétrique est dite :*
— *positive (resp. négative) si pour tout $x \in E$, on a $\phi(x, x) \geq 0$ (resp. $\phi(x, x) \leq 0$) ;*
— *définie positive (resp. définie négative) si elle est positive (resp. négative) et que $\phi(x, x) = 0$ si et seulement si x est le vecteur nul.*

Théorème 231. ***(Loi d'inertie de Sylvester)***
Soit ϕ une forme bilinéaire symétrique.
(i) Il existe alors une décomposition

$$E = E^\perp \oplus E^+ \oplus E^-$$

telle que la restriction de ϕ à E^+ (resp. E^-) est définie positive (resp. définie négative). Une telle décomposition n'est pas unique mais les dimensions r de E^+ et s de E^- sont les mêmes pour toute telle décomposition et sont respectivement égale au maximum des dimensions des sous-espaces F de E tels que la restriction de ϕ y soit définie positive (resp. négative). On dit que le couple (r, s) est la signature *de ϕ.*
(ii) Le rang de ϕ est égal à $s + r$ et son indice est égal à $(n - \mathrm{rg}\phi) + \min\{s, r\}$.

Preuve : (i) L'existence de la décomposition a été vue plus haut dans un cadre général, il reste alors à vérifier la caractérisation de r et s. Soit donc un espace F sur lequel q est défini positif et supposons par l'absurde que sa dimension est $> r$. D'après le théorème du rang, il intersecte alors non trivialement $E^\perp \oplus E^-$ espace sur lequel est $q(x) \leq 0$ d'où la contradiction. Ainsi r est bien égal à la dimension maximale d'un espace sur lequel q est définie positive.

(ii) Le rang de ϕ est clairement égal à $r+s$. Rappelons que son indice est la dimension d'une SETIM. Notons $(e_1^+, \cdots, e_r^+)$ (resp. $(e_1^-, \cdots, e_s^-)$) une base de E^+ (resp. E^-) ainsi que $(e_1^0, \cdots, e_t^0)$ une base de $E^\perp$. Considérons alors

$$F = \text{Vect}(e_1^+ + e_1^-, \cdots, e^+_{\min(r,s)} + e^-_{\min(r,s)}, e_1^0, \cdots, e_t^0).$$

On vérifie aisément que F est totalement isotrope de dimension $(n - \text{rg}\phi) + \min(r,s)$.

Supposons alors par l'absurde qu'il existe un SETI F de dimension $> (n - \text{rg}\phi) + \min(r,s)$. D'après le théorème du rang il intersecte alors non trivialement E^+ si $r > s$ et E^- sinon alors que pour tout x non nul de E^+ (resp. E^-) on a $q(x) > 0$ (resp. $q(x) < 0$), d'où la contradiction.
Application : tout sous-$\mathbb{R}$-espace vectoriel du cône nilpotent, i.e. de l'ensemble des matrices nilpotentes de $\mathbb{M}_n(\mathbb{R})$, est de dimension inférieur à $\frac{n(n-1)}{2}$. En effet on considère la forme quadratique q définie sur l'espace des matrices qui à X associe $\text{tr}X^2$. De manière évidente le cône isotrope est constitué de vecteurs isotropes de sorte que l'espace vectoriel en question sera totalement isotrope. Par ailleurs, si $X \neq 0$ est symétrique (resp. antisymétrique) alors $q(X) > 0$ (resp. $q(X) < 0$) de sorte que la signature de q est $(\frac{n(n+1)}{2}, \frac{n(n-1)}{2})$. Ainsi un sous-espace totalement isotrope est de dimension inférieure ou égale à $\frac{n(n-1)}{2}$. L'égalité est clairement atteinte pour les matrices strictement triangulaires supérieures.

Définition 232. *Soit F un sous-espace non isotrope de E ; l'unique involution unitaire u telle que $F = \text{Im}(u + \text{Id})$ s'appelle* la symétrie orthogonale par rapport au sous-espace non isotrope F. *Si F est un hyperplan, on dit que u est* une réflexion *; si F est de codimension 2 on dit que u est* un retournement *ou* un renversement.

Remarque: si v est un vecteur normé orthogonal à un hyperplan H non isotrope, alors la réflexion par rapport à H est l'application $x \mapsto x - 2\phi(x,v)v$.

Théorème 233. ***(de Cartan-Dieudonné)***
Soit q une forme quadratique non dégénérée sur E.

(i) Tout élément u de $O(q)$ peut s'écrire comme un produit de $p := \dim \Im(u - \text{Id})$ réflexions ; en outre tout écriture de u comme composé de réflexions exige au moins p réflexions différentes.

(ii) En dimension ≥ 3, tout élément de $SO(q)$ est peut s'écrire comme un produit de q retournements avec $q \leq \dim E$.

Remarque: En dimension 2, on a un isomorphisme de groupes $\mathbb{R}/2\pi\mathbb{Z} \to SO(2)$ qui à θ associe la matrice

$$R(\theta) := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

On remarque alors qu'il n'est pas possible d'exhiber une sous-famille génératrice particulière.

Preuve : (i) Notons tout d'abord que si r_H désigne la réflexion orthogonale relativement à l'hyperplan H, alors $u = r_{H_1} \circ \cdots r_{H_r}$ vérifie $H_1 \cap \cdots \cap H_r \subset \mathrm{Ker}(u - \mathrm{Id})$ de sorte que $\dim \Im(u - \mathrm{Id}) = n - \dim \mathrm{Ker}(u - \mathrm{Id}) \leq r$ et donc toute écriture de u comme un composé de réflexions nécessite au moins $\dim \Im(u - \mathrm{Id})$ facteurs.

Montrons alors par récurrence sur $p = \dim \Im(u - \mathrm{Id})$, qu'il existe une écriture $u = r_1 \circ \cdots r_q$ avec $q \leq p$. Pour $p = 0$ c'est clair puisqu'alors $u = \mathrm{Id}$. Supposons le résultat acquis jusqu'à $p-1$. Considérons alors $x \in \mathrm{Ker}(u-\mathrm{Id})^\perp$ tel que $u(x) \neq x$ et soit r_0 la réflexion orthogonale relativement à l'hyperplan H orthogonal à $u(x) - x$. Comme $u \in O(q)$, on a $u(x) \in \mathrm{Ker}(u - \mathrm{Id})^\perp$ et donc $\mathrm{Ker}(u - \mathrm{Id}) \subset H$ et donc $\mathrm{Ker}(u - \mathrm{Id}) + \mathrm{Vect}(x) \subset \mathrm{Ker}(r_0 \circ u - \mathrm{Id})$. Ainsi $\dim \Im(r_0 \circ u - \mathrm{Id}) \leq p-1$ et d'après l'hypothèse de récurrence, il existe une écriture $r_0 \circ u = r_1 \circ \cdots \circ r_q$ avec $q \leq p-1$ et le résultat s'en déduit en composant à gauche par r_0.

(ii) Il suffit de noter que pour r une réflexion orthogonale relativement à un hyperplan H, alors $-r$ est un retournement. Ainsi pour $u = (r_1 \circ r_2) \circ \cdots \circ (r_{2k-1} \circ r_{2k})$, où on notera que le nombre de facteurs est nécessairement pair pour $u \in SO(q)$, en écrivant $r_{2i-1} \circ r_{2i} = (-r_{2i-1}) \circ (-r_{2i})$, on fait bien apparaitre u comme un composé de retournements.

En dimension 3 toute matrice de $SO(3)$ est semblable à une matrice de la forme

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

On a aussi un isomorphisme de groupes entre $SO(3)$ et le groupe des quaternions de norme 1 quotienté par $\{\pm 1\}$. Cette description est particulièrement utile quand il s'agit de composer des rotations de l'espace.

Proposition 234. *Un endomorphisme u est normal si et seulement si dans une base orthonormée il admet une matrice diagonale par blocs de taille* 1 *ou* 2 *de la forme* $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

Preuve : Comme sur $\mathbb{R}$ un polynôme irréductible est de degré au plus 2, il résulte de (11) et en utilisant que les sous-modules de $\mathbb{R}[X]/(P)$ correspondent aux diviseurs de P, que u admet soit une droite soit un plan

stable. Si on a une droite stable et comme d'après 216 un endomorphisme normal est semi-simple, on conclut par récurrence sur la dimension. Si on a un plan stable et comme la restriction de u à un sous-espace stable est encore normal, pour $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, l'égalité $MM^* = M^*M$ donne

$$\begin{cases} b^2 = c^2 \\ ac + bd = ab + dc \end{cases}$$

et donc

- soit $c = b$ auquel cas la matrice est symétrique avec un polynôme caractéristique scindé, i.e. u admet une droite stable ;
- soit $c = -d$ et $a = d$ et on retrouve la matrice de l'énoncé.

On conclut encore par récurrence sur la dimension en utilisant la semi-simplicité de u.

Réciproquement, il suffit de vérifier que la matrice $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ est normale, ce qui a déjà été fait ci-avant.

Corollaire 235. *Un endomorphisme u de E est*

- *symétrique si et seulement si il est diagonalisable en base orthonormée.*
- *anti-symétrique si et seulement si, dans une base orthonormée, sa matrice est diagonale par blocs nuls ou de la forme* $\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$.
- *orthogonal si et seulement si, dans une base orthonormée, sa matrice est diagonale par blocs où les blocs sont soit* ± 1 *soit des matrices de rotation* $\begin{pmatrix} cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

Remarque: on peut utiliser ce théorème de réduction pour montrer, par exemple, que le groupe spécial orthogonal est connexe par arcs, en reliant toute matrice orthogonal positive O à la matrice identité. Pour ce faire, on transforme toute matrice de rotation de O et d'angle θ, en une matrice de rotation d'angle $t\theta$ pour $t \in [0, 1]$, et en regroupant les blocs de -1 deux par deux pour les identifier à une matrice de rotation d'angle π, que l'on transforme en une matrice d'angle $t\pi$.

Proposition 236. *Tout sous-groupe compact G de $GL_n(\mathbb{R})$ est contenu dans un conjugué du groupe orthogonal.*

Preuve : Il s'agit donc de montrer que G stabilise une forme quadratique définie positive. En effet si A est la matrice d'une telle forme, elle est diagonalisable en base orthonormée ${}^tPAP = D$ avec ${}^tP = P^{-1}$ et $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ où les λ_i sont strictement positifs. On peut alors écrire

$D = D_1^2$ où D_1 est une matrice diagonale à valeurs propres strictement positives de sorte que $A = B^2$ avec $B = {}^tPD_1P$. Ainsi pour tout $M \in G$, on a ${}^tBB = {}^tMB^2M$ soit

$$ {}^t(BMB^{-1})(BMB^{-1}) = 1$$

i.e. BMB^{-1} est orthogonale.

Considérons alors l'action de G sur l'ensemble $\mathfrak{S}_n^{++}$ des formes quadratiques définies positives selon la formule ${}^tBB \mapsto {}^tM{}^tBBM = {}^t(BM)BM$. L'idée est d'utiliser un théorème de point fixe ce qui nécessite de se restreindre à un convexe compact. Pour avoir un compact il suffit de considérer $\{{}^tBB : \ B \in G\}$ et pour la convexité de prendre l'enveloppe convexe de ce compact qui reste donc compact d'après le théorème de Carathéodory. On a ainsi l'action d'un groupe compact sur un convexe compact et le lemme 238 ci-après assure l'existence d'un point fixe. Il ne reste plus qu'à vérifier alors que le point fixe correspond à une forme définie positive, ce qui découle du lemme suivant.

Lemme 237. *L'ensemble $\mathfrak{S}_n^{++}$ des formes quadratiques définies positives est convexe.*

Preuve : L'idée est de faire de la congruence simultanée. Soient donc A et B deux matrices symétriques définies positives : A définit alors un produit scalaire ϕ_A. La matrice $U = A^{-1}B$ définit un endomorphisme symétrique relativement à ϕ_A puisque $AU = {}^tUA$ et donc il existe une base orthonormée pour ϕ_A diagonalisant U, i.e. ${}^tPAP = I_n$ et $P^{-1}UP = D$ et donc $(P^{-1}A^{-1}{}^tP^{-1}){}^tPBP = D$ soit ${}^tPBP = D$. Ainsi pour $0 \leq t \leq 1$, on a

$$ tA + (1-t)B = {}^tP^{-1}(tI_n + (1-t)D)P^{-1}$$

qui est bien définie positive.

On conclut à présent avec le lemme de point fixe nécessaire pour finir de prouver la proposition précédente.

Lemme 238. *Soit $G \subset GL(E)$ un groupe compact agissant sur un compact convexe K d'un espace vectoriel euclidien E. Il existe alors un point de K fixé par tous les éléments de G.*

Preuve : On introduit l'application

$$ x \in E \mapsto N_G(x) = \sup_{g \in G} ||g(x)||,$$

où le sup est en fait un maximum, i.e. est atteint, car G est compact. On vérifie aisément que N_G est une norme qui est de plus strictement convexe puisque

$$ N_G(x+y) = ||g_0(x) + g_0(y)|| \leq ||g_0(x) + g_0(y)|| \leq N_G(x) + N_g(y)$$

et pour avoir égalité il faut en particulier que la première inégalité soit une égalité et donc qu'il existe $\lambda > 0$ tel que $g_0(x) = \lambda g_0(y)$ et comme g_0 est linéaire et inversible, $x = \lambda y$, d'où l'affirmation.

Ainsi il existe un unique $x_0 \in K$ minimisant N_G et comme trivialement pour tout x on a $N_G(g(x)) = N_G(x)$, d'après l'unicité du minimum on a $g(x_0) = x_0$, d'où le résultat.

5.5 Le cas hermitien

Dans ce paragraphe pour $\mathbb{K} = \mathbb{C}$, pour ne pas répéter le cadre des formes quadratiques, on considère le cas où σ est la conjugaison complexe. Pour $A \in GL_n(\mathbb{C})$, on note A^* pour ${}^t\overline{A}$. Notons en particulier que tout forme hermitienne ϕ vérifie $\phi(x,x) \in \mathbb{R}$. On dit alors qu'elle est *positive* (resp. négative) si $\phi(x,x) \geq 0$ (resp. ≤ 0) pour tout $x \in E$ et on dit qu'elle est en outre *définie* si $\phi(x,x) = 0 \Rightarrow x = 0$.
Remarque: si E est muni d'une forme hermitienne définie positive on dit que E est *un espace hermitien*.

Proposition 239. *Si ϕ est positive alors*

$$\phi(x,y)\overline{\phi(x,y)} \leq q(x)q(y).$$

Théorème 240. *Comme dans le cas réel,*

- *il existe une décomposition $E = E^{\perp} \oplus E^+ \oplus E^-$ telle que la restriction de ϕ à E^+ (resp. E^-) est définie positive (resp. négative). En outre la dimension s de E^+ et t de E^- sont indépendantes de cette décomposition et le couple (s,t) s'appelle* la signature *de ϕ.*
- *Il existe une base $(e_i)_{1\leq i\leq n}$ telle que*

$$\phi\Big(\sum_{i=1}^{n}\lambda_i e_i, \sum_{j=1}^{n}\mu_j e_j\Big) = \sum_{i=1}^{s}\lambda_i\overline{\mu_i} - \sum_{i=s+1}^{s+t}\lambda_i\overline{\mu_i}.$$

- *Le rang de ϕ est $s+t$ et son indice $n-(s+t)+\min\{s,t\}$.*

Terminons ces rappels algébriques par un exemple fondamental de matrice hermitienne.

Définition 241. *Soit $(x_1, \cdots, x_m)$ une famille de vecteurs d'un espace hermitien E.* La matrice de Gram *définie par*

$$Gram(x_1, \cdots, x_m) = \Big(\phi(x_i, x_j)\Big)_{1\leq i,j\leq m}$$

est une matrice hermitienne dont on note $G(x_1, \cdots, x_m)$ le déterminant.

Proposition 242. *Pour tout $(x_1, \cdots, x_m)$, le déterminant de Gram $G(x_1, \cdots, x_m)$ appartient aux réel positifs ; il est non nul si et seulement si la famille $(x_1, \cdots, x_m)$ est libre.*

Corollaire 243. *Soient $x \in E$ et F un sous-espace de E ; si $(x_1, \cdots, x_m)$ est une base de F alors la distance $d(x, F)$ de x à F est donnée par*

$$d(x, F)^2 = \frac{G(x, x_1, \cdots, x_m)}{G(x_1, \cdots, x_m)}.$$

Proposition 244. *Un endomorphisme est normal si et seulement s'il est diagonalisable en base orthonormée.*

Preuve : Si u est normal, il est alors semi-simple d'après 216 : pour x un vecteur propre de u, on a $E = \langle x \rangle \oplus \langle x \rangle^\perp$ et on conclut par récurrence sur la dimension.

La réciproque est évidente et découle du fait qu'une matrice diagonale commute avec son adjointe diagonale.

Corollaire 245. *Un endomorphisme d'un espace hermitien est*

- *hermitienne (resp. anti-hermitienne) si et seulement si, dans une base orthonormée, sa matrice est diagonale réelle (resp. imaginaire pure).*
- *unitaire si et seulement si, dans une base orthonormée, sa matrice est diagonale avec des coefficients de module* 1.

5.6 Valeurs propres de matrices hermitiennes

Pour une matrice complexe générale, les valeurs propres sont les racines du polynôme caractéristique. Si la matrice est hermitienne, les valeurs propres sont en outre les solutions de plusieurs problèmes d'optimisation : ce paragraphe est consacré à la présentation de quelques uns de ceux-ci, le lecteur intéressé pourra consulter [**?**] §4.2. Dans la suite A désigne une matrice complexe hermitienne dont les valeurs propres réelles sont classées par ordre croissant :

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Théorème 246. *(Courant-Fisher) Pour tout $1 \leq k \leq n$, on a*

$$\lambda_k = \min_{F \in \mathcal{E}_k} \max_{0 \neq x \in F} \frac{x^* A x}{x^* x} = \max_{F \in \mathcal{E}_{n-k+1}} \min_{0 \neq x \in F} \frac{x^* A x}{x^* x}$$

où $\mathcal{E}_k$ désigne l'ensemble des sous-espaces de dimension k de $\mathbb{C}^n$.

Remarque: le cas de λ_1 et λ_n est connu sous le nom du théorème de Rayleight-Ritz.

Preuve : Notons pour tout $1 \leq k \leq n$, x_k un vecteur propre unitaire pour la valeur propre λ_k. Soit alors $F \in \mathcal{E}_k$ de sorte que

$$F \cap \mathrm{vect}(x_k, x_{k+1}, \cdots, x_n)$$

est de dimension supérieure à 1. Pour $x = \sum_{i=k}^{n} \alpha_i x_i \neq 0$ un vecteur unitaire de cette intersection, on a

$$x^* A x = \sum_{i=k}^{n} \lambda_i |\alpha_i|^2 \geq \lambda_k \sum_{i=k}^{n} |\alpha_i|^2 = \lambda_k.$$

On en déduit alors l'inégalité

$$\lambda_k \geq \min_{F \in \mathcal{E}_k} \max_{0 \neq x \in F} \frac{x^* A X}{x^* x}.$$

Par ailleurs pour $F = \mathrm{vect}(x_1, \cdots, x_k)$, on a

$$\lambda_k = \max_{0 \neq x \in F} \frac{x^* A X}{x^* x},$$

d'où l'égalité. L'autre cas se traite de manière strictement identique.

Citons quelques applications du théorème précédent, cf. [**?**] §4.3.

Théorème 247. *(Weyl) Soient A, B des matrices hermitiennes de valeurs propres $\lambda_i(A), \lambda_i(B)$ classées par ordre croissant. En classant de même les valeurs propres de $A + B$, on obtient :*

(a) $\lambda_k(A)+\lambda_1(B) \leq \lambda_k(A+B) \leq \lambda_k(A)+\lambda_n(B)$, *pour tout* $k = 1, \cdots, n$;

(b) si B est de rang au plus r :

— $\lambda_k(A + B) \leq \lambda_{k+r}(A) \leq \lambda_{k+2r}(A + B)$ *pour* $k = 1, \cdots, n - 2r$;

— $\lambda_k(A) \leq \lambda_{k+r}(A + B) \leq \lambda_{k+2r}(A)$, *pour* $k = 1, \cdots, n - 2r$;

(c) $\lambda_{j+k-n}(A + B) \leq \lambda_j(A) + \lambda_k(B)$, *pour tout* $1 \leq j, k \leq n$ *tels que* $j + k \geq n + 1$;

(d) $\lambda_j(A) + \lambda_k(B) \leq \lambda_{j+k-1}(A + B)$, *pour tout* $1 \leq j, k \leq n$ *tel s que* $j + k \leq n + 1$.

Preuve : (a) On a $\frac{x^*(A+B)x}{x^*x} = \frac{x^*Ax}{x^*x} + \frac{x^*Bx}{x^*x}$; le résultat découle alors simplement de l'encadrement $\lambda_1(B) \leq \frac{x^*Bx}{x^*x} \leq \lambda_n(B)$ et du théorème 246.

(b) On écrit B sous la forme $\alpha_1 u_1 u_1^* + \cdots + \alpha_r u_r u_r^*$ où les vecteurs u_i de $\mathbb{C}^n$ ne sont pas nécessairement indépendants. On a alors

$$\begin{aligned} \lambda_{k+2r}(A + B) &= \min_{F \in \mathcal{E}_{k+2r}} \max_{0 \neq x \in F} \frac{x^*(A+B)x}{x^*x} \\ &\geq \min_{F \in \mathcal{E}_{k+2r}} \max_{0 \neq x \in F \cap \mathrm{vect}(u_1, \cdots, u_r)^\perp} \frac{x^*(A+B)x}{x^*x}. \end{aligned}$$

En notant $\mathcal{E}'_{k+r}$ l'ensemble des sous-espaces de $\mathrm{vect}(u_1, \cdots, u_r)^\perp$ de dimension $k + r$, le dernier terme ci-dessus est égal à

$$\min_{F \in \mathcal{E}'_{k+r}} \max_{0 \neq x \in F} \frac{x^* A x}{x^* x} \geq \min_{F \in \mathcal{E}_{k+r}} \max_{0 \neq x \in F} \frac{x^* A x}{x^* x} = \lambda_{k+r}(A).$$

Selon le même schéma, on a

$$\begin{aligned}\lambda_k(A+B) &= \max_{F\in\mathcal{E}_{n-k+1}} \min_{0\neq x\in F} \frac{x^*(A+B)x}{x^*x}\\ &\leq \max_{F\in\mathcal{E}_{n-k+1}} \min_{0\neq x\in F\cap \mathrm{vect}(u_1,\cdots,u_r)^\perp} \frac{x^*(A+B)x}{x^*x}\\ &= \max_{F\in\mathcal{E}'_{n-k+1-r}} \min_{0\neq x\in F} \frac{x^*Ax}{x^*x}\\ &\leq \max_{F\in\mathcal{E}_{n-k+1-r}} \min_{0\neq x\in F} \frac{x^*Ax}{x^*x}\\ &= \lambda_{k+r}(A)\end{aligned}$$

Ces deux familles d'inégalités fournissent alors celles de l'énoncé.

(c) On diagonalise $A = UD_1U^*$ et $B = VD_2V^*$ et on note u_i (resp. v_i) les vecteurs colonnes de la matrice unitaire U (resp. V). Pour un couple (j,k) vérifiant les conditions de l'énoncé, on note pour β assez grand tel que pour tout $j+1 \leq i \leq n$ et pour tout $k+1 \leq i' \leq n$, $\lambda_i(A) - \beta < \lambda_j(A)$ et $\lambda_{i'}(B) - \beta < \lambda_k(B)$:[2]

$$A_j = \beta(u_nu_n^* + \cdots + u_{j+1}u_{j+1}^*) \qquad B_k = \beta(v_n^* + \cdots + v_{k+1}v_{k+1}^*).$$

En remarquant que $(A - A_j)u_i = (\lambda_i - \beta)u_i$ pour $i = j+1, \cdots, n$, et est égal à $\lambda_i u_i$ pour $i = 1, \cdots, j$, on note que $\lambda_n(A - A_j) = \lambda_j(A)$. De même on a $\lambda_n(B - B_k) = \lambda_k(B)$.

Par ailleurs comme A_j (resp. B_k) est de rang $n-j$ (resp. $n-k$), $A_j + B_k$ est de rang au plus $2n - j - k$ et donc d'après (b)

$$\begin{aligned}\lambda(A - A_j + B - B_k) &= \lambda_n(A + B - (A_j + B_k))\\ &\geq \lambda_{n-(2n-j-k)}(A+B)\\ &= \lambda_{j+k-n}(A+B)\end{aligned}$$

D'après (a) pour $k = n$, on a aussi

$$\lambda_n(A - A_j + B - B_k) \leq \lambda_n(A - A_j) + \lambda_n(B - B_k)$$

de sorte que

$$\begin{aligned}\lambda_k(A) + \lambda_k(B) &= \lambda(A - A_j) + \lambda_n(B - B_k) \geq \lambda_n(A - A_j + B - B_k)\\ &= \lambda_n((A+B) - (A_j + B_k)) \geq \lambda_{j+k-n}(A+B)\end{aligned}$$

(d) Le résultat découle directement de (c) en considérant $-A$ et $-B$ et en notant que $\lambda_i(-A) = -\lambda_{n-i+1}(A)$.
Remarque: dans (b) pour le cas $r = 1$, les valeurs propres sont entrelacées, i.e.

$$\lambda_k(A+B) \leq \lambda_k(A) \leq \lambda_{k+2}(A+B) \qquad \lambda_k(A) \leq \lambda_{k+1}(A+B) \leq \lambda_{k+2}(A)$$

Le même phénomène se produit dans la situation suivante.

2. On fera attention que dans [**?**] théorème 4.3.6 (c), l'auteur oublie que $\lambda_j(A)$ peut être strictement négatif, de sorte que $\lambda_n(A - B)$ serait nul et non égal à $\lambda_{n-r}(A)$.

Théorème 248. *Soit $A \in \mathbb{M}_{n+1}(\mathbb{C})$ hermitienne de valeurs propres $\lambda_1 \leq \cdots \leq \lambda_{n+1}$. Alors les valeurs propres $\lambda'_1 \leq \cdots \leq \lambda'_n$ d'une matrice extraite principale*[3] *de A vérifie les inégalités suivantes :*

$$\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \lambda'_2 \leq \cdots \leq \lambda'_{n-1} \leq \lambda_n \leq \lambda'_n \leq \lambda_{n+1}$$

Preuve : Par souci de simplicité supposons que $i = n+1$, on a alors

$$\begin{aligned}
\lambda_{k+1} &= \min_{F\in\mathcal{E}_{k+1}} \max_{0\neq x\in F} \frac{x^*Ax}{x^*x} \\
&\geq \min_{F\in\mathcal{E}_{k+1}} \max_{0\neq x\in F\cap \mathrm{vect}(e_{n+1})^\perp} \frac{x^*Ax}{x^*x} \\
&= \min_{F'\in\mathcal{E}'_k} \max_{0\neq x'\in F'} \frac{(x')^*A'x'}{(x')^*x'} \\
&= \lambda'_k
\end{aligned}$$

où on a écrit pour $x \in \mathbb{C}^{n+1}$, $x' \in \mathbb{C}^n$ est le vecteur obtenu en supprimant la dernière coordonnée : on a adopté des notations similaires pour F' et $\mathcal{E}'_k$ désigne les sous-espaces de dimension k de $\mathbb{C}^n$. La majoration $\lambda_k \leq \lambda'_k$ se prouve de la même manière en considérant $\max_{F\in\mathcal{E}_{n-k+1}} \min_{0\neq x\in F}$.
Remarque: évidemment la démonstration s'adapte simplement au cas où l'on considère une matrice extraite principale A_r de A, où l'on a supprimé r-lignes et les r-colonnes correspondantes. Le résultat est alors :

$$\lambda_k(A) \leq \lambda_k(A_r) \leq \lambda_{k+n-r}(A).$$

Le théorème précédent admet aussi une réciproque :

Théorème 249. *Soient pour un entier $n \geq 1$, des nombres réels tels que :*

$$\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \lambda'_2 \leq \cdots \leq \lambda'_{n-1} \leq \lambda_n \leq \lambda'_n \leq \lambda_{n+1}$$

Soit $A' = \mathrm{diag}(\lambda'_1, \cdots, \lambda'_n)$, il existe alors un réel a et un vecteur $y \in \mathbb{R}^n$ tels que $\{\lambda_1, \cdots, \lambda_{n+1}\}$ est l'ensemble des valeurs propres de la matrice symétrique : $A = \begin{pmatrix} A' & y \\ {}^t y & a \end{pmatrix}$.

Preuve : Le calcul de la trace donne $a = \sum_{i=1}^{n+1} \lambda_i - \sum_{i=1}^{n} \lambda'_i$. Pour tout t distincts des λ_i, on a l'égalité suivante :

$$\begin{pmatrix} I & 0 \\ {}^t((tI-A')^{-1}y) & 1 \end{pmatrix} \begin{pmatrix} tI-A' & -y \\ -{}^t y & t-a \end{pmatrix} \begin{pmatrix} I & (tI-A')^{-1}y \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} tI-A' & 0 \\ 0 & (t-a) - {}^t y(tI-A')^{-1}y \end{pmatrix}$$

En prenant les déterminants, on obtient alors l'égalité suivante entre les polynômes caractéristiques :

$$\chi_A(t) = \chi_{A'}(t)\Big[(t-a) - \sum_{i=1}^{n} y_i^2 \frac{1}{t-\lambda_i}\Big]$$

3. i.e. pour un indice $1 \leq i \leq n+1$, on enlève à A sa i-ème colonne et sa i-ème ligne

avec donc $\chi_{A'}(t) = \prod_{i=1}^{n}(t-\lambda_i)$. Il s'agit alors de monter l'existence de n réels y_i tels que $\chi_A(\lambda_k) = 0$ pour tout $k = 1, \cdots, n+1$. On considère alors la division euclidienne $\xi_A(t) := \prod_{i=1}^{n+1}(t-\lambda_i) = Q(t)\chi_{A'}(t) + R(t)$ avec donc $Q(t) = t - a$ et pour tout $1 \leq k \leq n$, $R(\lambda'_k) = \xi_A(\lambda'_k)$, ce qui détermine uniquement le polynôme R de degré inférieur ou à n en utilisant par exemple les polynômes interpolateurs de Lagrange. Supposons pour simplifier que tous les λ'_i sont distincts, on a alors

$$R(t) = \sum_{i=1}^{n} \xi_A(\lambda'_i) \frac{\chi_{A'}(t)}{\chi'_{A'}(t)(t-\lambda'_i)}$$

de sorte que

$$\frac{\xi_A(t)}{\chi_{A'}(t)} = (t-a) - \sum_{i=1}^{n} \frac{-\xi_A(\lambda'_i)}{\chi'_{A'}(t)} \frac{1}{t-\lambda'_i}$$

Il suffit alors de montrer que pour tout $i = 1, \cdots, n$, $\xi_A(\lambda'_i)\chi'_{A'}(\lambda'_i) \leq 0$ de sorte qu'en posant $y_i^2 = -\frac{\xi_A(\lambda'_i)}{\chi'_{A'}(\lambda'_i)}$, on aura $\chi_A(t) = \xi_A(t)$.

Il s'agit alors d'utiliser l'hypothèse d'entrelacement des valeurs propres : on remarque ainsi que

$$\xi_A(\lambda'_i) = (-1)^{n-i+1} \prod_{j=1}^{n+1} (\lambda'_i - \lambda_j) \qquad \chi'_{A'}(\lambda'_i) = (-1)^{n-i} \prod_{\substack{j=1 \\ j\neq i}}^{n} (\lambda'_i - \lambda'_j)$$

sont effectivement de signes opposés.
Remarque: dans le cas où certains des λ'_i sont égaux, par exemple $\lambda'_1 = \lambda'_2 = \cdots \lambda'_k < \lambda'_{k+1} \leq \cdots$, on remarque alors que $(t-\lambda'_1)^{k-1}$ divise $\xi_A(t)$ et on reprend le raisonnement précédent en divisant $\xi_A(t)$ et $\chi_{A'}(t)$ par $(t-\lambda'_1)^{k-1}$.

Corollaire 250. *Soit $A \in \mathbb{M}_n(\mathbb{C})$ hermitienne ; pour $1 \leq r \leq n$, U désigne une matrice de $\mathbb{M}_{n,r}(\mathbb{C})$ telle que ses vecteurs colonnes forment une famille orthonormale, i.e. $U^*U = I \in \mathbb{M}_r(\mathbb{C})$. On a alors les propriétés suivantes :*
(i) pour tout $k = 1, 2, \cdots, r$, $\lambda_k(A) \leq \lambda_k(U^*AU) \leq \lambda_{k+n-r}(A)$;
(ii) $\lambda_1(A) + \cdots + \lambda_r(A) = \min_{U^*U=I\in\mathbb{M}_r(\mathbb{C})} \operatorname{tr}(U^*AU)$ et

$$\lambda_{n-r+1}(A) + \cdots + \lambda_n(A) = \max_{U^*U=I\in\mathbb{M}_r(\mathbb{C})} \operatorname{tr}(U^*AU).$$

Remarque: (i) est connu comme le théorème de séparation de Poincaré et est utilisé en mécanique quantique où on a accès aux calculs de $u_i^*Au_j$ pour une famille orthonormée $(u_i)_{1\leq i\leq r}$.

Preuve : (i) si $r < n$, on complète les vecteurs colonnes de U en une base orthonormée ; la matrice U' est alors unitaire et $(U')^*AU'$ a les mêmes valeurs propres que A et U^*AU en est une matrice extraite principale, le résultat découle alors de 248, ou plutôt de la remarque qui suit.

(ii) les majorations découlent directement de (i), les égalités sont alors obtenues si les colonnes de U correspondent aux vecteurs propres des r plus petites valeurs propres.

On a le même genre de résultat pour les valeurs singulières.

Théorème 251. *Soit $A \in \mathbb{M}_{m,n}(\mathbb{C})$ et $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q$ ses valeurs singulières pour $q = \min\{m,n\}$. Pour $1 \leq k \leq q$, on a*

$$\sigma_k = \min_{F \in \mathcal{E}_k} \max_{0 \neq x \in F} \frac{||Ax||_2}{||x||_2} = \min_{F \in \mathcal{E}_{n-k+1}} \max_{0 \neq x \in F} \frac{||Ax||_2}{||x||_2}$$

Remarque: suivant le même schéma que précédemment, on peut obtenir des résultats similaires sur l'enchevêtrement des valeurs singulières. Par exemple pour $A \in \mathbb{M}_{m,n}(\mathbb{C})$ avec $m \geq n$, pour $1 \leq i \leq n$, $A^{(i)}$ désigne la matrice extraite de A en enlevant sa i-ème colonne et ligne. On note σ_i (resp. σ'_i) les valeurs singulières de A (resp. A'), de sorte que l'on a

$$\sigma_1 \geq \sigma'_1 \geq \sigma_2 \geq \sigma'_2 \geq \cdots \geq \sigma'_{n-1} \geq \sigma_n \geq 0$$

On en déduit alors cf. [**?**] 3.1.3, que si A_r désigne une sous-matrice de A obtenue en lui ôtant r lignes et/ou colonnes alors $\sigma_k(A) \geq \sigma_k(A_r) \geq \sigma_{k+r}(A)$.

Corollaire 252. *Soit $A \in \mathbb{M}_n(\mathbb{C})$ de valeurs singulières $\sigma_1 \geq \cdots \geq \sigma_n$ et soit $H(A) = \frac{1}{2}(A + A^*)$ sa partie hermitienne de valeurs propres $\lambda_1 \geq \cdots \lambda_n$. Pour tout $k = 1, \cdots, n$, on a $\lambda_k \leq \sigma_k$.*

Preuve : Pour $x \in \mathbb{C}^n$ unitaire, on a $x^*H(A)x = \text{Re}\,(x^*Ax) \leq |x^*Ax| \leq ||x||_2.||Ax||_2 = ||Ax||_2$. Ainsi on a

$$\lambda_k = \min_{F \in \mathcal{E}_k} \max_{\substack{0 \neq x \in F \\ ||x||_2 = 1}} x^*H(A)x \leq \min_{F \in \mathcal{E}_k} \max_{\substack{0 \neq x \in F \\ ||x||_2 = 1}} ||Ax||_2 = \sigma_k$$

Corollaire 253. *(cf. [**?**] 3.3.2) Soit $A \in \mathbb{M}_n(\mathbb{C})$ de valeurs singulières $\sigma_1 \geq \cdots \geq \sigma_n$ et de valeurs propres $\{\lambda_1, \cdots, \lambda_n\}$ ordonnées de sorte que $|\lambda_1| \geq \cdots \geq |\lambda_n|$. On a alors*

$$|\lambda_1 \cdots \lambda_k| \leq \sigma_1 \cdots \sigma_k \quad \forall k = 1, \cdots, n$$

avec égalité pour $k = n$.

Preuve : Soit U unitaire telle que $U^*AU = T$ est triangulaire supérieure et où la diagonale de T est $(\lambda_1, \cdots, \lambda_n)$. Soit $U_k \in \mathbb{M}_{n,k}(\mathbb{C})$ la matrice extraite de U constituée par ses k premières colonnes de sorte que $U^*AU = \begin{pmatrix} U_k^*AU_k & * \\ * & * \end{pmatrix}$. La matrice $T_k = U_k^*AU_k$ est donc triangulaire supérieure de diagonale égale à $\lambda_1, \cdots, \lambda_k$ de sorte que $|\det T_k| = |\lambda_1 \cdots \lambda_k|$ est égal au produit $\sigma_1(T_k) \cdots \sigma_k(T_k)$ des valeurs singulières de T_k. Le résultat découle alors de la remarque qui suit le théorème 251, i.e. $\sigma_1(T_k) \cdots \sigma_k(T_k) \leq \sigma_1 \cdots \sigma_k$.

6 Modules

Dans la suite A désignera un anneau commutatif que l'on supposera rapidement principal. Le lecteur est vivement encouragé, dans un premier temps, à le considérer égal à $\mathbb{Z}$ puis $K[X]$ pour K un corps.

6.1 Généralités

Dans ce paragraphe A est un anneau commutatif quelconque.

Définition 254. - *Un A-module est un groupe commutatif $(M,+)$ muni d'une application $A \times M \to M$, où l'on note ax l'image de (a,x), telle que :*

1. *$\forall a \in A$ et $x, y \in M$, $a(x+y) = ax + ay$;*
2. *$\forall a, b \in A$ et $x \in M$, $(a+b)x = ax + bx$;*
3. *$\forall a, b \in A$ et $x \in M$, $1x = 1$ et $a(bx) = (ab)x$.*

- Un sous-module d'un A-module M est un sous-groupe N de M stable par l'action de A.
- Un morphisme de A-modules $M \to N$ est un morphisme des groupes additifs $(M,+) \to (N,+)$ qui est de plus A-linéaire.

Remarque: la notion de A-module est formellement identique à celle de K-espace vectoriel sauf que l'action externe est relative à un anneau A plutôt qu'à un corps K.
Exemples : les constructions habituelles sur les espaces vectoriels se généralisent au cas des A-modules (quotient, somme, intersection, A-module engendré...). On utilisera dans la suite plus spécifiquement les exemples suivants.
- Si $(G,+)$ est un groupe commutatif, il est canoniquement muni d'une structure de $\mathbb{Z}$-module, en définissant, pour $n \geq 0$, ng comme $\overbrace{g+g+\cdots+g}^{n}$, et $(-1)g$ comme $-g$.
- Si V est un K-espace vectoriel et $u \in \mathcal{L}(V)$ un endomorphisme de V, on munit V d'une structure de $K[X]$-module en posant pour tout $P \in K[X]$ et pour tout $\overrightarrow{v} \in V$, $P.\overrightarrow{v} := P(u)(\overrightarrow{v})$.
- L'anneau A est lui-même un A-module. Les sous-A-modules de A sont ses idéaux.

Définition 255. *Un sous-ensemble S d'un A-module M est dit*

- libre *si l'égalité $\sum_{s \in S} a_s s = 0$ où la famille $(a_s)_{s \in S}$ est supposée à support fini, implique que pour tout $s \in S$, on a $a_s = 0$.*
- générateur *si tout élément $m \in M$ peut s'écrire sous la forme $\sum_{s \in S} a_s s$ où la famille $(a_s)_{s \in S}$ est à support fini.*
- une base *si S est à la fois libre et générateur.*

On dit que M est

- libre *si M admet une base.*
- de type fini *s'il admet un sous-ensemble fini S générateur.*

— de torsion *si l'ensemble des éléments $\lambda \in A$ qui annulent M i.e. tels que $\forall m \in M$ on ait $\lambda m = 0$, est un idéal non nul de A. Cet idéal est appelé* annulateur *de M et noté* $\mathrm{Ann}(M)$.

Exemple : l'annulateur de $M = A/I$ est l'idéal I.
Remarque: la donnée d'une base d'un A-module libre de type fini est équivalent à la donnée d'un isomorphisme $A^n \longrightarrow M$.

Proposition 256. *Soit M un A-module libre de type fini. Alors toutes ses bases ont le même cardinal.*

Preuve : Soient $(e_1, \cdots, e_n)$ une base de M et $\mathcal{M} \in A$ un idéal maximal de sorte que le quotient $A/\mathcal{M}$ est un corps k. On note $\mathcal{M}M = \{\sum_{i=1}^n m_i e_i : m_i \in \mathcal{M}\}$ de sorte que $V := M/\mathcal{M}M$ est un k-espace vectoriel, i.e. un groupe muni d'une action externe de $A/\mathcal{M}$. Notons par ailleurs que $(\overline{e}_i)_{i=1,\cdots,n}$ est une base de V. C'est clairement une famille génératrice. Pour la liberté, $\sum_{i=1}^n \lambda_i \overline{e}_i = 0$ s'écrit aussi $\sum_{i=1}^n \mu_i e_i = \sum_{i=1}^n m_i e_i$ où $\overline{\mu}_i = \lambda_i$ et les $m_i \in I$. La famille $(e_i)_{i=1,\cdots,e_n}$ étant libre, on en déduit que $\mu_i = m_i$ et donc $\lambda_i = \overline{\mu}_i = 0$.

Ainsi n est la dimension de l'espace vectoriel $M/\mathcal{M}M$ et est donc le cardinal de toute base du A-module M.

Proposition 257. *Un sous-module d'un module de type fini est de type fini.*

Preuve : Soit $f : A^n \twoheadrightarrow M$ définie par la donnée d'une famille génératrice de cardinal n de M. Pour un sous-module N de M, il suffit de montrer que $f^{-1}(N)$ est de type fini, i.e. on est ramené au cas où $M = A^n$. On raisonne alors par récurrence sur n : dans le cas $n = 1$, un sous-module de A est un idéal qui est donc principal et donc libre de rang 1.
Supposons alors le résultat acquis jusqu'au rang $n-1$ et traitons le cas de n. Considérons alors l'application $g : N \hookrightarrow A^n \twoheadrightarrow A$ où la deuxième flèche est donnée par la première projection $(a_1, \cdots, a_n) \mapsto a_1$. L'image de g est de la forme $a_1 A$ et notons $n_1 \in N$ un antécédent puis $N' = N \cap A^{n-1}$ où A^{n-1} est le noyau de la première projection. Notons que $n_1 A \cap N' = (0)$ puisque si $g(\lambda n_1) = 0$ alors $\lambda = 0$. En outre pour $n \in N$, on peut écrire $n = \lambda n_1 + (n - \lambda n_1)$ où $f(n) = \lambda a_1 = f(\lambda n_1)$ et donc $n - \lambda n_1 \in N'$. Autrement dit on a $N = A n_1 \oplus N'$ avec $N' \subset A^{n-1}$. Par récurrence N' est de type fini et donc N aussi.

6.2 Calculs matriciels dans un anneau principal

Rappelons que $\mathbb{M}_n(A)$ désigne l'ensemble des matrices carré de taille n à coefficients dans A.

Lemme 258. *La matrice $M \in \mathbb{M}_n(A)$ est inversible si et seulement si* $\det M \in A^\times$.

Preuve : Si la matrice M est inversible, alors en appliquant le déterminant à l'égalité $M.M^{-1} = I_n$ on obtient que l'inverse de $\det M$ est $\det(M^{-1})$. Inversement, notons $\tilde{M}$ la transposée de la matrice des cofacteurs de M. De l'égalité $\tilde{M}.M = \det M$ on en déduit que si $\det M \in A^\times$ alors $(\det M)^{-1}\tilde{M}$ est l'inverse de M.

Le lemme suivant est l'ingrédient calculatoire essentiel pour les calculs matriciels de la suite.

Lemme 259. *Soient $x, y \in A$ non tous deux nuls, z un pgcd de x et $ux + vy = z$ une relation de Bézout. Alors la matrice $M := \begin{pmatrix} u & v \\ -y/z & x/z \end{pmatrix}$ de déterminant 1 vérifie l'équation*

$$M \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} z \\ 0 \end{pmatrix}.$$

Nous utiliserons la matrice de taille (n, n) suivante :

$$L_{j,k}(x,y) = \begin{pmatrix}
1 & 0 & \dots & & & & & & \dots & 0 \\
0 & \ddots & & & & & & & 0 & \vdots \\
 & & 1 & & & & & & & \\
\vdots & 0 & \dots & u & 0 & \dots & & v & \vdots & \\
 & \vdots & & & 1 & & & & & \\
 & & & & & \ddots & & & & \\
 & & & & & & 1 & & & \\
 & & & -y/z & 0 & \dots & & x/z & & \\
 & & & & & & & & 1 & \\
 & & & & & & & & \ddots & \\
0 & \dots & & & & & & & & 1
\end{pmatrix}$$

u étant en (j, j), v en (j, k), $-y/z$ en (k, j) et x/z en (k, k).
Remarque: comme d'habitude, en notant l_i la ligne d'indice i d'une matrice M, la multiplication à gauche d'une matrice $M \in \mathbb{M}_{n,m}(A)$ par la matrice $L_{j,k}(x, y)$ remplace l_j et l_k par respectivement $\alpha l_j + \beta l_k$ et $\gamma l_j + \delta l_k$.

Proposition 260. *Soit $M \in \mathbb{M}_{n,m}(A)$. Il existe alors une matrice $L \in SL_n(A)$ telle que LM soit triangulaire supérieure.*

Preuve : La démonstration consiste à appliquer plusieurs fois le lemme 259 pour faire apparaitre des zéros sous la diagonale principale.
a) Soit $M = (a_{ij})$ $(1 \leq i \leq n,\ 1 \leq j \leq m)$. Multiplions M à gauche par la matrice $L_1 = L_{1,2}(a_{1,1}, a_{2,1})$ de sorte que la première colonne de $M_1 = L_1 M$ commence par $\begin{pmatrix} d \\ 0 \end{pmatrix}$ avec $d = a_{11} \wedge a_{21}$.

b) On multiplie ensuite M_1 à gauche par une matrice $L_2 = L_{1,3}(d, a_{3,1})$ de façon à faire apparaitre un zéro à la place $(3,1)$ et à remplacer d par $d_1 = a_{1,1} \wedge a_{2,1} \wedge a_{3,1}$. On continue ainsi jusqu'à ce que l'on obtienne la matrice $M_{n-1} = L_{n-1} \cdots L_1 M$ dont la première colonne est $(d_{n-1}, 0 \dots 0)$ avec d_{n-1} le pgcd des éléments de la première colonne de M.
c) On continue de même avec la seconde colonne, en commençant par multiplier à gauche par une matrice $L_{2,3}$ de façon à laisser la première ligne inchangée et à ne manipuler que les lignes $l_2, \dots, l_n$. En procédant ainsi on obtient une deuxième colonne de la forme $(a, b, 0, \cdots, 0)$. En continuant ainsi pour toutes les colonnes, on obtient une matrice triangulaire supérieure.

Définition 261. *Une matrice $M \in \mathbb{M}_{n,m}(A)$ est dite* réduite *si*

$$M = \begin{pmatrix} a_{1,1} & 0 & \dots & & 0 \\ 0 & a_{2,2} & 0 & \dots & \vdots \\ \vdots & & \ddots & & \\ & & & a_{n,n} & \dots \; 0 \end{pmatrix}$$

avec

$$a_{i,i} \mid a_{i+1,i+1}, \quad 1 \leq i \leq \inf(n,m) - 1.$$

Remarque: on a représenté une matrice M avec $n < m$. Il est à noter que les derniers $a_{i,i}$ peuvent être nuls et que tous les éléments non sur la diagonale sont nuls.

Théorème 262. *Soit $M \in \mathbb{M}_{n,m}(A)$. Il existe alors $L \in SL_n A$ et $R \in SL_m(A)$ telles que $M' = LMR$ soit réduite.*

Remarque: l'énoncé analogue sur un corps K est que toute matrice $M \in \mathbb{M}_{n,m}(K)$ est équivalente à une matrice de la forme $M' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Preuve : Nous avons vu comment en manipulant les lignes, on faisait remonter le pgcd de chaque colonne sur la première ligne. On commence donc par faire cela pour chacune des colonnes. Puis, comme on a opéré sur les lignes, on opère sur les colonnes de sorte à ramener le pgcd de la première ligne, et donc celui de tous les coefficients de la matrice, en position $(1,1)$.

À ce stade, le coefficient $a_{1,1}$ est le pgcd de tous les $a_{i,j}$. On recommence alors à opérer sur les lignes de façon à obtenir la première colonne égale à $(a_{1,1}, 0, \cdots, 0)$: on note en particulier que la première ligne n'est pas modifiée. On opère ensuite de même sur les colonnes pour que la première ligne soit égale à $(a_{1,1}, 0, \cdots, 0)$. Comme précédemment, on ne modifie pas la première ligne de sorte qu'à ce stade la matrice est diagonale par blocs avec un premier bloc de taille 1 et le deuxième de taille $(n-1, m-1)$.

On conclut alors par récurrence.

6.3 Théorème de la base adaptée

Théorème 263. *Soit N un sous-module d'un A-module L libre de type fini. Alors N est un sous-module libre de type fini et il existe une base, dite* adaptée, *$(f_1, \ldots, f_n)$ de L ainsi que des éléments $a_i \in A$, $1 \leq i \leq n$ tels que :*

$$\begin{cases} a_1 \mid a_2 \mid \cdots \mid a_n, \\ \text{les } (a_i f_i) \text{ tels que } a_i \neq 0 \text{ forment une base de } N. \end{cases}$$

De plus, la suite des idéaux (a_i) satisfaisant ces conditions est unique.

Preuve : D'après la proposition 257, N est de type fini, notons alors $(g_1, \ldots, g_m)$ une famille génératrice de N et écrivons la matrice de passage M des g_i pour $1 \leq i \leq m$ dans une base $(e_1, \cdots, e_n)$ de L. D'après 262, il existe $P \in SL_n A$ et $Q \in SL_m(A)$ telles que $M' = PMQ$ soit réduite avec des éléments $a_{i,i}$ sur la diagonale que l'on note simplement a_i.
La matrice P (resp. Q) s'interprète comme une matrice de changement de base de L (resp. de changement de famille génératrice de N). Notons $(f_1, \cdots, f_n)$ la nouvelle base de L ; l'écriture matricielle de M' s'interprète alors en disant que $(a_1 f_1, \cdots, a_r f_r)$ est une famille génératrice de N, où on a noté a_r le dernier des a_i non nuls. On note alors que cette nouvelle famille génératrice de N est libre, i.e. N est aussi libre avec $M/N \simeq A/(a_1) \times \cdots \times A/(a_r) \times A^{n-r}$. Montrons alors l'unicité des (a_i). Notons tout d'abord que $n - r$ ne dépend que de N. Pour ce faire considérons un irréductible p ne divisant pas a_r de sorte que pour tout $1 \leq i \leq n$, a_i est inversible modulo p et donc pour $M_i = A/(a_i)$ on a M_i/pM_i est nul. On voit ainsi que $n-r$ est la dimension du $A/(p)$ espace vectoriel $(M/N)/p(M/N)$. Considérons alors

$$\frac{A}{(a_1)} \times \cdots \times \frac{A}{(a_q)} \simeq \frac{A}{(a'_1)} \times \cdots \times \frac{A}{(a'_s)},$$

les (a_i) et les (a'_i) vérifiant les propriétés de divisibilité de l'énoncé et sont tous non nuls. Alors $(a_r) = \mathrm{Ann}(M/N) = (a'_s)$ et donc $\frac{A}{(a_1)} \times \cdots \times \frac{A}{(a_{q-1})} \simeq \frac{A}{(a'_1)} \times \cdots \times \frac{A}{(a'_{s-1})}$. En procédant de même de manière, on identifie de proche en proche les a_{q-i} avec les a'_{s-i} jusque obtenir $s = q$.

Définition 264. *On dit que $m \in M$ est un* élément de torsion *si $m \neq 0$ et s'il existe $\lambda \in A, \lambda \neq 0$, tel que $\lambda m = 0$. L'ensemble des éléments de torsion de M est noté M_t. Si $M_t = \{0\}$, on dit que M est* sans torsion.

Remarque: M_t est un sous-module de M. Par ailleurs M est de torsion si et seulement si $M = M_t$. Avec ces notions la preuve du théorème précédent se réécrit.

Théorème 265. *Soit M un A-module de type fini, M_t son sous-module de torsion. Alors il existe un sous-module $L \subset M$ libre de rang r tel que $M = M_t \oplus L$ ainsi que des éléments $a_1|a_2|\cdots|a_q$ de A tels que*

$$M_t \simeq A/(a_1) \oplus A/(a_2) \oplus \cdots \oplus A/(a_q).$$

Définition 266. *Le rang de la partie libre de M s'appelle le* rang *de M. Les idéaux non nuls (a_i) pour $1 \leq i \leq q$ sont les* facteurs invariants *de M.*

Remarque: le corps des rationnels $\mathbb{Q}$ est un exemple de $\mathbb{Z}$-module sans torsion et non libre (si $q_1 = a/b$ et $q_2 = c/d$, sont deux rationnels non nuls, on a la relation $bcq_1 - adq_2 = 0$). Ainsi $\mathbb{Q}$ n'est pas un $\mathbb{Z}$-module de type fini.

Définition 267. *Un A-module M est dit* indécomposable *s'il n'est pas isomorphe à la somme directe de deux A-modules non nuls.*

Proposition 268. *Soit M un A-module de type fini. Les conditions suivantes sont équivalentes :*

1. *le module M est indécomposable ;*
2. *$M \simeq A$, ou il existe un élément irréductible $p \in A$, un entier $\alpha > 0$ tels que $M \simeq A/(p^\alpha)$.*

Preuve : 1. $\Rightarrow$ 2.
D'après le théorème 265 on peut supposer $M = A/(a)$. Si l'élément a a au moins deux facteurs irréductibles, il résulte du lemme chinois que M n'est pas indécomposable.
2. $\Rightarrow$ 1.
L'anneau A étant intègre, il est clair que le A-module A est indécomposable. Si $\alpha > 0$, les sous-modules de $\tilde{M} = A/(p^\alpha)$ sont engendrés par les images dans $\tilde{M}$ des éléments p^γ pour $\gamma \leq \alpha$. Si M_1 et M_2 sont deux tels sous-modules, on a toujours $M_1 \subset M_2$ ou $M_2 \subset M_1$; ils ne peuvent donc pas être en somme directe.

Définition 269. *Soient M un A-module, $p \in \mathcal{P}$ un élément irréductible. On note $M(p)$ l'ensemble des éléments $x \in M$ de p-torsion, i.e. annulés par une puissance de p.*

Remarque: en appliquant le théorème chinois au théorème 265, on obtient la décomposition canonique en indécomposable donnée par le théorème suivant.

Théorème 270. *Soit M un A-module de torsion de type fini, $(a) = \mathrm{Ann}(M)$ son annulateur. Alors :*

1. *$M = \bigoplus_{p_i \in \mathcal{P},\, p_i|a} M(p_i)$ et $M(p_i) \neq (0)$ pour chaque élément irréductible p_i tel que $p_i|a$*

2. *pour chaque élément irréductible $p_i \in \mathcal{P}$, $p_i|a$, il existe une suite d'entiers $\nu_{i1} \leq \nu_{i2} \leq \cdots \leq \nu_{ik}$ unique telle que :*

$$M(p_i) \simeq \prod_{j=1}^{k} A/(p_i^{\nu_{ij}})).$$

3. *la décomposition $M \simeq \prod_{i,j} A/(p_i^{\nu_{ij}})$ est l'unique décomposition de M en produit de modules indécomposables.*

Exemple : prenons $A = \mathbb{Z}$, $M = M_t = \mathbb{Z}/96\mathbb{Z} \times \mathbb{Z}/72\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z}$. On a $96 = 2^5 \times 3$, $72 = 2^3 \times 3^2$, d'où :

$$M \simeq \mathbb{Z}/32\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$$

par le théorème chinois. On a donc

$$\begin{array}{lcccccc} M(2) & \simeq & \mathbb{Z}/32\mathbb{Z} & \times & \mathbb{Z}/8\mathbb{Z} & \times & \mathbb{Z}/2\mathbb{Z} \\ M(3) & \simeq & \mathbb{Z}/9\mathbb{Z} & \times & \mathbb{Z}/3\mathbb{Z} & & \\ M(5) & \simeq & \mathbb{Z}/5\mathbb{Z}. & & & & \end{array} \tag{12}$$

Pour trouver la décomposition en indécomposable (resp. en facteurs invariants), on lit le tableau ci-dessus en lignes (resp. en colonnes), et on trouve $M = M(2) \oplus M(3) \oplus M(5)$ (resp. $a_3 = 32 \times 9 \times 5 = 1440$, $a_2 = 8 \times 3 = 24$, $a_1 = 2$, d'où la décomposition $M \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/1440\mathbb{Z}$.
Remarque: en considérant un groupe abélien comme un $\mathbb{Z}$-module, on peut donner la classification des groupes abéliens finis d'ordre n donné. On procède de la manière suivante. Soit G un groupe d'ordre n.

1. On écrit $n = p_1^{\nu_1} \dots p_s^{\nu_s}$ avec p_i premiers, ν_i entiers ;
2. on a alors $G \simeq G(p_1) \oplus \cdots \oplus G(p_s)$;
3. pour chaque entier i, $1 \leq i \leq s$ il existe une suite (ν_{ij}) unique d'entiers > 0 tels que $\sum_j \nu_{ij} = \nu_i$ et $G(p_i) \simeq \bigoplus_j \mathbb{Z}/p_i^{\nu_{ij}}\mathbb{Z}$;
4. deux groupes d'ordre n sont isomorphes si et seulement si tous les p_i et les entiers ν_{ij} sont les mêmes.

Exemple : donnons à isomorphisme près tous les groupes abéliens d'ordre $108 = 2^2 \times 3^3$. Soit $G = G(2) \oplus G(3)$ avec $G(2)$ d'ordre 4 et $G(3)$ d'ordre 27. Il y a à isomorphisme près deux possibilités pour $G(2)$, $\mathbb{Z}/4\mathbb{Z}$ et $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, et trois pour $G(3)$, $\mathbb{Z}/27\mathbb{Z}$, $\mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ et $(\mathbb{Z}/3\mathbb{Z})^3$. Il y a donc à isomorphisme près six groupes abéliens d'ordre 108.

7 Sous-groupes de $\mathbb{R}^n$

7.1 Généralités sur les réseaux

Parmi les sous-groupes de $\mathbb{R}$, les exemples les plus simples sont d'une part $\{0\}$, $\mathbb{Z}$ et plus généralement $\delta\mathbb{Z}$ pour $\delta \in \mathbb{R}$, et d'autre part $\mathbb{Z} + \sqrt{2}\mathbb{Z}$,

$\mathbb{Q}$ et $\mathbb{R}$. Les premiers sont *discrets* au sens où pour tout compact K de $\mathbb{R}$ l'intersection $K \cap G$ est finie, alors que les deuxièmes sont denses. Par ailleurs étant donnés deux sous-groupes G_1 et G_2 de respectivement $\mathbb{R}^{n_1}$ et $\mathbb{R}^{n_2}$, le groupe produit $G_1 \times G_2$ est un sous-groupe de $\mathbb{R}^{n_1+n_2}$. Nous allons voir dans une certaine mesure qu'on obtient ainsi tous les sous-groupes de $\mathbb{R}^n$.

Lemme 271. *Un sous-groupe G de $\mathbb{R}^n$ est discret dans $\mathbb{R}^n$ si et seulement s'il existe un ouvert U de $\mathbb{R}^n$ contenant 0 tel que $G \cap U$ soit discret.*

Preuve : Si G est discret on peut prendre $U = \mathbb{R}^n$. Réciproquement si G n'est pas discret alors il existe $z \in \mathbb{R}^n$ qui est un point d'accumulation d'éléments de G, i.e. pour tout $\epsilon > 0$, il existe $x_1 \neq x_2 \in G$ tel que $0 \leq |z - x_i| < \epsilon$ pour $i = 1, 2$ et donc $0 < |x_1 - x_2| < 2\epsilon$ ce qui montre que 0 est un point d'accumulation de G car $x_1 - x_2 \in G$.
Remarque: en particulier un sous-groupe non discret de $\mathbb{R}$ est partout dense.

Définition 272. *Un sous-groupe discret d'un espace vectoriel est appelé un* sous-réseau *de V ; s'il engendre V on dit que c'est* un réseau.

Voici une caractérisation des réseaux parmi les sous-réseaux.

Lemme 273. *Soit G un sous-réseau de V. Pour que G engendre V, il faut et il suffit qu'il existe un ensemble borné B de V tel que*

$$V = \bigcup_{g \in G} (B + g).$$

Preuve : Si G contient une base $(e_1, \cdots, e_n)$ de V alors

$$B = \{\sum_{i=1}^{n} x_i e_i,\ 0 \leq x_i < 1\}$$

convient. Réciproquement si G est contenu dans un sous-espace strict V' de V, notons $p : V \longrightarrow W$ la projection de V sur un supplémentaire W de V' dans V. Alors

$$p\Big(\bigcup_{g \in G} (B + g)\Big) = p(B).$$

Comme B est borné et que $W = p(V)$ est de dimension ≥ 1, on a $p(B) \neq p(V)$ et donc

$$\bigcup_{g \in G} (B + g) \neq V.$$

Proposition 274. *Soit G un sous-groupe discret de $\mathbb{R}^n$. Il existe alors $1 \leq t \leq n$ et des éléments $e_1, \cdots, e_t \in G$ linéairement indépendants tels que $G = \mathbb{Z}e_1 + \cdots + \mathbb{Z}e_t$.*

Preuve : Soit $f_1, \cdots, f_t$ une base de l'espace vectoriel V engendré par G de sorte que $G' = \mathbb{Z}f_1 + \cdots + \mathbb{Z}f_t$ est un sous-groupe de G. Montrons que G' est d'indice fini dans G. Considérons le compact $K = \left\{u_1 f_1 + \cdots + u_t f_t : \ 0 \leq u_i \leq 1\right\}$. Soit $g \in G \subset V$ que l'on écrit sous la forme

$$g = x_1 f_1 + \cdots + x_t f_t = \sum_{i=1}^{t} \lfloor x_i \rfloor f_i + k$$

avec $k \in G \cap K$. Comme G est supposé discret $G \cap K$ est fini de sorte que G/G' l'est aussi. Notons $s := [G : G']$ et soit $f_i' = \frac{1}{s} f_i$. On a alors

$$\mathbb{Z}f_1 + \cdots + \mathbb{Z}f_1 \subset G \subset \mathbb{Z}f_1' + \cdots + \mathbb{Z}f_t',$$

et le résultat découle du théorème de la base adaptée 263.

Théorème 275. *Soit G un sous-groupe additif de $\mathbb{R}^n$. Il existe alors un plus grand sous-espace vectoriel V de $\mathbb{R}^n$ contenu dans l'adhérence de G et il existe un sous-groupe discret G' de G tel que G soit la somme directe*

$$G = (G \cap V) \oplus G'.$$

Si t désigne le rang de G' et d la dimension de V alors $d+t$ est la dimension de l'espace vectoriel engendré par G.

Preuve : Pour $\rho > 0$, notons $B(0, \rho) = \{x \in \mathbb{R}^n, \ ||x|| \leq \rho\}$ et soit V_ρ le $\mathbb{R}$-espace vectoriel engendré par $G \cap B(0, \rho)$. L'application $\rho \mapsto \dim V_\rho$ est croissante à valeurs entières positives de sorte qu'il existe $\rho_0 > 0$ tel que pour tout $0 < \rho \leq \rho_0$, $V_\rho = V_{\rho_0}$. Posons $V := V_{\rho_0}$ et montrons que $G' = G \cap V$ est dense dans V. Soit $\epsilon > 0$ et soit $x \in V$. Posons $\rho = \min\{\epsilon/d, \rho_0\}$ et soit $(e_1, \cdots, e_d)$ une base de V avec $e_i \in G \cap B(0, \rho)$. Pour $x = x_1 e_1 + \cdots + x_d e_d$, on pose $m_i = \lfloor x_i \rfloor$ et $y = m_1 e_1 + \cdots + m_d e_d$. Alors $y \in G'$ vérifie $||x - y|| \leq \epsilon$. Soit W le sous-espace engendré par G ; comme il contient V sa dimension est $d + t$ avec $t \geq 0$. Notons V' un supplémentaire de V dans W et soit $p : W \longrightarrow V'$ la projection de noyau V. Montrons que $p(G)$ est un sous-groupe discret de V'. Dans le cas contraire il existerait $z \in p(G)$ tel que $0 < ||z|| < \epsilon$ avec $\epsilon = \rho_0/2$. Soit $w \in G$ tel que $z = p(w)$; on a $u = z - w \in V$. Comme G' est dense dans V, il existe $w' \in G'$ tel que $||u - w'|| < \epsilon$ et $||w - w'|| < \rho_0$ puis $p(w - w') = z \neq 0$, i.e. $w - w' \in G$ vérifie $w - w' \notin V$ ce qui contredit le fait que $V = V_{\rho_0}$.

Alors $p(G)$ est un sous-groupe discret de V' de rang t, donc un réseau de V'. On en prend une base $p(y_1), \cdots, p(y_t)$ et on pose $G'' = \mathbb{Z}y_1 + \cdots + \mathbb{Z}y_t$ de sorte que $G = G' \oplus G''$. Enfin comme G'' est discret, V est bien le plus grand sous-espace vectoriel de $\mathbb{R}^n$ contenu dans l'adhérence de G.

Proposition 276. *Soit Γ un réseau de $\mathbb{R}^n$, $(e_1, \cdots, e_n)$ une $\mathbb{Z}$-base de Γ. La famille $(v_1, \cdots, v_n)$ est alors une base de Γ si et seulement si la matrice de passage P de e à v appartient à $GL_n(\mathbb{Z})$.*

Preuve : Dans le sens direct si $(v_1, \cdots, v_n)$ est une $\mathbb{Z}$-base de Γ alors la matrice de passage de v à e est à coefficient dans $\mathbb{Z}$ et est l'inverse de la matrice P de l'énoncé qui appartient donc à $GL_n(\mathbb{Z})$.

Réciproquement si $P \in GL_n(\mathbb{Z})$ alors son inverse est à coefficient dans $\mathbb{Z}$ de sorte que pour tout $i = 1, \cdots, n$, le vecteur e_i appartient à $\mathbb{Z}v_1 + \cdots + \mathbb{Z}v_n$ et donc $\Gamma \subset \mathbb{Z}v_1 + \cdots + \mathbb{Z}v_n$. L'inclusion inverse étant évidente puisque les $v_i \in \Gamma$, on en déduit que v est une base de Γ.
Remarque: rappelons aussi le théorème de la base adaptée : soient Γ un réseau de V et $\Lambda \subset \Gamma$ un sous-groupe alors Λ est un sous-réseau de V et il existe une $\mathbb{Z}$-base $(e_1, \cdots, e_n)$ de Γ, $1 \leq s \leq n$ et $a_1, \cdots, a_s \in \mathbb{Z}^\times$ tels que

— $(a_1 e_1, \cdots, a_s e_s)$ est une $\mathbb{Z}$-base de Λ,

— pour $1 \leq i < s$, a_i divise a_{i+1}

Ainsi si le corps est $\mathbb{Q}$, il existe $d \in \mathbb{N}\backslash\{0\}$ tel que $d\Gamma \subset \Lambda$. En outre $\Gamma + \Lambda$ et $\Gamma \cap \Lambda$ sont des réseaux de V.

7.2 Domaines fondamentaux

On munit désormais $\mathbb{R}^n$ de sa structure euclidienne canonique et on note $\epsilon_1, \cdots, \epsilon_n$ sa base canonique orthonormée.

Définition 277. *Soit Γ un réseau de $\mathbb{R}^n$ et $e = (e_1, \ldots, e_n)$ une $\mathbb{Z}$-base de Γ. On note $P_e = \{t_1 e_1 + \ldots + t_n e_n \mid \forall i \in \{1, \ldots, n\}, t_i \in [0, 1]\}$ le* parallèlogramme fondamental *du réseau Γ associé à la $\mathbb{Z}$-base e. On appelle* mesure *du réseau Γ, et on note $\mu(\mathbb{R}^n/\Gamma)$, le volume de P_e (pour la mesure de Lebesgue).*

Remarque: Si A_e est la matrice de $M_n(\mathbb{R})$ dont les vecteurs colonnes sont les vecteurs e_i (représentés dans la base canonique de $\mathbb{R}^n$), alors $\mu(P_e) = |\det(A_e)|$. Comme toute matrice de changement de $\mathbb{Z}$-base du réseau Γ a pour déterminant 1 ou -1, ce volume ne dépend pas de la $\mathbb{Z}$-base e choisie.

Définition 278. *Une partie $\mathcal{D}$ de $\mathbb{R}^n$ est un domaine fondamental de Γ, si $\mathcal{D}$ est μ-mesurable et si ses translatés par les vecteurs de Γ forment une partition de $\mathbb{R}^n$.*

Proposition 279. *$D_{\mathbf{e},\Gamma}$ est un domaine fondamental et $\mu(\mathcal{D}) = \mu(\mathbb{R}^n/\Gamma)$ pour tout domaine fondamental $\mathcal{D}$ de Γ.*

Preuve : Soient $\mathcal{D}_1$ et $\mathcal{D}_2$ sont des domaines fondamentaux quelconques ; en considérant $\mathcal{D}_2$ comme un domaine fondamental, on écrit

$$\mathcal{D}_1 = \coprod_{v \in \Gamma} \mathcal{D}_1 \cap (v + \mathcal{D}_2),$$

Γ étant dénombrable et μ étant invariante par translation, on a

$$\begin{aligned} \mu(\mathcal{D}_1) &= \textstyle\sum_{v \in \Gamma} \mu(\mathcal{D}_1 \cap (v + \mathcal{D}_2)) \\ &= \textstyle\sum_{v \in \Gamma} \mu((-v + \mathcal{D}_1) \cap \mathcal{D}_2) \end{aligned}$$

Or comme $-\Gamma = \Gamma$, on en déduit $\mu(\mathcal{D}_1) = \sum_{v\in\Gamma} \mu(\mathcal{D}_2 \cap (v + \mathcal{D}_1)) = \mu(\mathcal{D}_2)$, la dernière égalité découlant du fait que $\mathcal{D}_1$ est un domaine fondamental.
Remarque: d'après le théorème de la base adaptée, si $\Lambda \subset \Gamma$ sont des réseaux alors Γ/Λ est fini et

$$\mu(\mathbb{R}^n/\Lambda) = \mathrm{card}(\Gamma/\Lambda)\mu(\mathbb{R}^n/\Gamma)$$

7.3 Théorème de Minkowski

Théorème 280. ***(Minkowski)*** *Soient Γ un réseau de $\mathbb{R}^n$ et A une partie μ-mesurable, convexe, symétrique par rapport à O et vérifiant $\mu(A) > 2^n\mu(\mathbb{R}^n/\Gamma)$, alors $A \cap \Gamma \neq \{O\}$.*

Preuve : Soit $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^n/\Gamma$ la surjection canonique associé au réseau Γ et soit F une partie de $\mathbb{R}^n$, μ-mesurable vérifiant $\mu(F) > \mu(\mathbb{R}^n/\Gamma)$; la restriction de φ à F n'est alors pas injective. En effet soit $\mathcal{D}$ un domaine fondamental :

$$\mu(F) = \sum_{\gamma\in\Gamma} \mu(F \cap (\gamma + \mathcal{D})) = \sum_{\gamma\in\Gamma} \mu((F - \gamma) \cap \mathcal{D}) > \mu(D)$$

d'où il en résulte que les $(F - \gamma) \cap \mathcal{D}$ pour $\gamma \in \Gamma$ ne sont pas deux à deux disjoints. Soient donc $x, y \in F$ et $\alpha \neq \beta \in \Gamma$ vérifiant $x - \alpha = y - \beta$ soit $x - y = \alpha - \beta \in \Gamma\backslash\{O\}$ et donc φ non injective.

Soit alors $F = \frac{1}{2}A$; on a donc $\mu(F) > \mu(\mathbb{R}^n/\Gamma)$. D'après ce qui précède, il existe $x, y \in F$ tels que $x-y \in \Gamma\backslash\{O\}$. En outre $2x$ et $-2y$ appartiennent à A d'après la propriété de symétrie de A par rapport à O, et donc $x-y = \frac{(2x-2y)}{2}$ appartient à A d'après la propriété de convexité de A, d'où le résultat.

Corollaire 281. *Soit C un convexe compact de $\mathbb{R}^n$, symétrique par rapport à O tel que $\mu(C) \geq 2^n\mu(\mathbb{R}^n/\Gamma)$ alors $C \cap \Gamma \neq \{O\}$.*

Preuve : Soit $C_r = (1 + 1/r)C$ pour $r \geq 1$; $C = \cap_{r\geq 1} C_r$ et $\mu(C_r) > 2^n\mu(\mathbb{R}^n/\Gamma)$. D'après la question précédente, soit $x_r \in C_r \cap (\Gamma\backslash\{O\}) \subset K := 2C \cap (\Gamma\backslash\{O\})$; K étant fini, on peut extraire de la suite $(x_r)_{r\geq 1}$ une sous-suite convergente, donc stationnaire, d'où le résultat.
Remarque: notons v_n le volume de la boule unité fermée de $\mathbb{R}^n$; on déduit du corollaire précédent qu'il existe $\gamma \in \Gamma$ différent de O et de norme inférieure ou égale à deux fois la racine n-ième de $v_n^{-1}\mu(\mathbb{R}^n/\Gamma)$.

Le théorème suivant implique non seulement le théorème de Minkowski mais possède aussi de nombreuses applications qui ne découlent pas de ce dernier.

Théorème 282. ***(de Blichfeldt)*** *Soit C une partie mesurable de $\mathbb{R}^2$ d'aire $> n$; il existe un translaté de C tel que $C \cap \mathbb{Z}^2$ est de cardinal $\geq n + 1$.*

Preuve : Le réseau $\mathbb{Z}^2$ découpe C en parties $C_1, \cdots, C_k$ telles que pour tout $1 \leq i \leq k$, C_i appartient à un translaté R_i du carré fondamental $[0,1]^2$. On considère alors les parties C_i' du carré fondamental obtenu par la translation qui identifie R_i avec $[0,1]^2$. Remarquons alors qu'il existe un point $P \in [0,1]^2$ tel que l'ensemble I des $1 \leq i \leq k$ tels que $P \in C_i'$ est de cardinal $\geq n+1$. En effet dans le cas contraire la somme des aires des C_i' pour $1 \leq i \leq k$ est inférieure ou égale à n ce qui contredit l'hypothèse que l'aire de C est $> n$. Notons pour tout $i \in I$, P_i le point de C_i d'image P par la translation identifiant R_i avec $[0,1]^2$. Soit alors la translation t qui envoie P sur l'origine, on remarque l'image de C contient les $n+1$ points de $\mathbb{Z}^2$ obtenus comme l'image par t des P_i pour $i \in I$.
Remarque: le résultat précédent est valable sous la forme plus générale suivante : un ensemble C de mesure A contient à translation près strictement plus de A points de $\mathbb{Z}^2$. Ce résultat est optimal comme le montre l'exemple du carré de coté $1-\epsilon$ avec $\epsilon > 0$ petit.

Plutôt que de redémontrer le théorème de Minkowski, nous proposons l'amélioration suivante.

Corollaire 283. *Soit C un convexe compact symétrique par rapport à l'origine et d'aire $4A$; alors C contient $\lfloor A \rfloor - 1$ paire de points de $\mathbb{Z}^2$ autres que l'origine O.*

Preuve : Notons $n = \lfloor A \rfloor$; d'après le théorème de Blichfeldt, il existe une translation t de $D = \frac{1}{2}C$ telle que $D' = t(D)$ contient n points de $\mathbb{Z}^2$. On choisit t tel qu'il existe un point $P \in \mathbb{Z}^2$ sur la frontière de D' ; notons D_1' le symétrique de D' par rapport à P. Si $Q \in D' \cap \mathbb{Z}^2$ alors son symétrique Q_1 est un point de $D_1' \cap \mathbb{Z}^2$. Si le segment $[QQ_1]$ n'appartient pas à la frontière de D' alors Q_1 n'est pas un point de D' ; dans le cas contraire, notons $[AB]$ la frontière commune à D' et D_1' sur la droite QQ', en prenant pour P le point le plus proche de A, cette situation n'est plus de sorte que l'on se ramène au cas où $D' \cup D_1'$ contient $2(n-1)+1$ points distincts de $\mathbb{Z}^2$. La translation qui envoie $t(O)$ sur P composée avec l'homothétie de centre P et de rapport 2, composée avec la translation qui envoie P sur O, transforme D en C et contenant $2(n-1)$ points de $\mathbb{Z}^2$ distincts de O.
Remarque: l'un des thèmes favoris des travaux de Blichfeldt concernait les valeurs minimales prises par une forme quadratique définies positives. Il a en particulier prouvé le résultat suivant.

Théorème 284. *Soit q une forme quadratique définie positive de discriminant $d > 0$, sur $\mathbb{R}^n$; il existe alors un vecteur $v \in \mathbb{R}^n$ tel que*

$$q(v) \leq \frac{2}{\pi}\Big(\Gamma(1+\frac{n+2}{2})\Big)^{2/n} d^{1/n}.$$

7.4 Bases d'un réseau

Etant donné une famille génératrice, il est aisé d'en trouver une base, au sens où il existe un algorithme en temps polynomial. On construit la matrice A des vecteurs générateurs dans la base canonique, par opérations élémentaires sur les colonnes, il existe $U \in GL_n(\mathbb{Z})$ telle que $B = AU$ où B est sous une forme normale de Hermite, i.e. $B = (0\ C)$ où C est triangulaire supérieure, $c_{i,i} > 0$ et pour tout $j > i$, on a $0 \leq c_{i,j} < c_{i,i}$.
Remarque: la méthode du pivot permet de la même façon de déterminer si un vecteur est un point du réseau.

L'inconvénient de la méthode précédente est qu'elle peut mener à des vecteurs de très grande norme par rapport à la mesure du réseau. Une problématique naturelle est de pouvoir construire la meilleur base possible au sens suivants :

— les vecteurs de base le plus courts possibles : réduction de Minkowski.
— la famille de vecteurs de base la plus orthogonale possible : réduction de Korkine et Zolotarev.

Remarque: en dimension 2 c'est très simple, on opère des translations successives. Par contre en général on conjecture que trouver un vecteur de norme minimal est un problème NP-dur : il convient donc de mettre de l'eau dans son vin.

Rappelons tout d'abord l'inégalité d'Hadamard.

Proposition 285 (Inégalité d'Hadamard)**.** *Soit Γ un réseau de $\mathbb{R}^n$ et $e = (e_1, \ldots, e_n)$ une $\mathbb{Z}$-base de Γ. Alors*

$$\mu(\mathbb{R}^n/\Gamma) \leq \prod_{i=1}^{n} \|e_i\|,$$

avec égalité si et seulement si $(e_1, \ldots, e_n)$ est une base orthogonale de $\mathbb{R}^n$.

Lemme 286. *Soit Γ un réseau de $\mathbb{R}^n$ et $e_1 \in \Gamma$ tel que $\|e_1\| = \min_{u \in \Gamma \setminus \{0\}} \|u\|$. Soit p la projection orthogonale sur $H = e_1^{\perp}$ et $\Lambda = p(\Gamma)$. Alors*

1. *Λ est un réseau de H.*
2. *soit $(f_2, \ldots, f_n)$ une $\mathbb{Z}$-base de Λ et $e_2, \ldots, e_n \in \Gamma$ tels que pour tout $2 \leq i \leq n$, on a $p(e_i) = f_i$. Alors $(e_1, \ldots, e_n)$ est une $\mathbb{Z}$-base de Γ.*
3. *pour tout $f \in \Lambda$, il existe $e \in \Gamma$ tel que $\|e\| \leq \|f\| \sqrt{\frac{4}{3}}$.*

Preuve : Si $(e_1, \ldots, e_n)$ est une $\mathbb{Z}$-base de Γ, alors $(p(e_2), \ldots, p(e_2))$ est une base de H. En effet, pour tout $2 \leq i \leq n$ on peut écrire $e_i = p(e_i) + \alpha_i e_1$ avec $\alpha_i \in \mathbb{R}$. Par conséquent, pour tout $\beta_2, \ldots, \beta_n \in \mathbb{R}$, si $\sum_{i=2}^{n} \beta_i p(e_i) = 0$, alors $(\sum_{i=2}^{n} \alpha_i \beta_i)\, e_1 + \sum_{i=2}^{n} \beta_i e_i = 0$, ce qui impose $\beta_2 = \ldots = \beta_n = 0$. On en déduit que $\Lambda = \mathbb{Z}p(e_2) \oplus \ldots \oplus \mathbb{Z}p(e_n)$ est un réseau de H.

Soit $(f_2, \ldots, f_n)$ une $\mathbb{Z}$-base de Λ et $e_2, \ldots, e_n \in \Gamma$ tels que pour tout $2 \leq i \leq n$, on a $p(e_i) = f_i$. On a clairement $\sum_{i=1}^{n} \mathbb{Z}e_i \subset \Gamma$. Réciproquement, pour

tout $x \in \Gamma$, il existe $\alpha_2, \ldots, \alpha_n \in \mathbb{Z}$ tels que $p(x) = \sum_{i=2}^n \alpha_i f_i$. On a alors $x - \sum_{i=2}^n \alpha_i e_i \in (\mathbb{R}e_1 \cap \Gamma)$, donc il existe $\alpha_1 \in \mathbb{Z}$ tel que $x - \sum_{i=2}^n \alpha_i e_i = \alpha_1 e_1$. On en déduit que $x = \sum_{i=1}^n \alpha_i e_i \in \sum_{i=1}^n \mathbb{Z}e_i$. Par conséquent, $(e_1, \ldots, e_n)$ est une $\mathbb{Z}$-base de Γ.

Enfin, pour tout $f \in \Lambda$, il existe $g \in \Gamma$ tel que $f = p(g)$. On a $g = f + \alpha e_1$, avec $\alpha \in \mathbb{R}$. Il existe $\beta \in \mathbb{Z}$ tel que $|\alpha - \beta| \leq 1/2$. On pose alors $e = f + (\alpha - \beta)e_1$. Clairement, $e \in \Gamma$ et par le théorème de Pythagore, $\|f\|^2 = \|e\|^2 - |\alpha - \beta|^2 \|e_1\|^2 \geq 3\|e\|^2/4$. On a donc $\|e\| \leq \|f\|\sqrt{\frac{4}{3}}$, ce qui termine la preuve du lemme.

On en déduit par récurrence le résultat suivant.

Théorème 287. ***(Hermite)*** *Tout réseau Γ de $\mathbb{R}^n$ admet une $\mathbb{Z}$-base $e = (e_1, \ldots, e_n)$ telle que*

$$\prod_{i=1}^n \|e_i\| \leq \left(\frac{4}{3}\right)^{\frac{n(n-1)}{4}} \mu(\mathbb{R}^n/\Gamma).$$

Définition 288. *On appelle* minima successifs *d'un réseau Γ de $\mathbb{R}^n$ les réels $\lambda_1(\Gamma) \leq \ldots \leq \lambda_n(\Gamma)$, où $\lambda_i(\Gamma)$ est le plus petit réel t tel qu'il existe i vecteurs libres de norme inférieure où égale à t.*

On dit qu'une famille libre $e = (e_1, \ldots, e_n)$ de Γ est une famille de vecteurs courts *de Γ si pour tout $1 \leq i \leq n$, on a $\|e_i\| = \lambda_i(\Gamma)$. On dit que e est une $\mathbb{Z}$-base de vecteurs courts de Γ si c'est une $\mathbb{Z}$-base et une famille de vecteurs courts de Γ.*

Contrairement à ce que l'on pourrait penser, une famille de vecteurs courts d'un réseau Γ n'est pas forcément une $\mathbb{Z}$-base de Γ, dès que $n \geq 4$. Par exemple, considérons le réseau Γ de $\mathbb{R}^4$ donné par la $\mathbb{Z}$-base

$$e = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right).$$

Notons que Γ est le sous-réseau de $\mathbb{Z}^4$ constitué des points dont la somme des coordonnées est paire. Par conséquent, les minima successifs de Γ sont $\sqrt{2}, \sqrt{2}, \sqrt{2}, \sqrt{2}$. La famille

$$f = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right)$$

est donc une famille de vecteurs courts de Γ mais n'est pas une $\mathbb{Z}$- base de Γ (puisque $e_1 + e_3 \notin \bigoplus_{i=1}^4 \mathbb{Z}f_i$).

Plus étrangement encore, pour $n \geq 5$, il existe des réseaux qui n'admettent aucune base de vecteurs courts. Par exemple, considérons le réseau Λ de $\mathbb{R}^5$ donné par la $\mathbb{Z}$-base

$$e = \left(\begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right).$$

On vérifie que les minima successifs de Λ sont $2, 2, 2, 2, 2$, et que la famille

$$f = \left(\begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix} \right)$$

est la seule famille (à changement de signe près) de vecteurs courts, et n'est pas une $\mathbb{Z}$-base de Λ (puisque $e_5 \notin \bigoplus_{i=1}^{5} \mathbb{Z} f_i$).

Ainsi, il n'est pas possible de donner un sens à une notion de réduction optimale d'un réseau. On se contente de réductions approchées : on s'intéresse par exemple aux $\mathbb{Z}$-bases dont les vecteurs sont relativement courts (ie. leur norme est proche des minima successifs), ou dont les vecteurs sont quasi-orthogonaux. De telles notions ne sont pertinentes d'un point de vue algorithmique que si l'on peut déterminer de telles $\mathbb{Z}$-bases en temps raisonnable. Dans le paragraphe suivant, on présente la réduction LLL : il s'agit de bases dont les vecteurs sont relativement courts, et que l'on peut trouver en temps polynomial par l'algorithme LLL.

7.5 Algorithme LLL

Définition 289. *Soit $b = (b_1, \ldots, b_n)$ une base de $\mathbb{R}^n$. On appelle* base orthogonalisée de Gram-Schmidt *associée à b la base $b^* = (b_1^*, \ldots, b_n^*)$, où pour tout $1 \leq i \leq n$, le vecteur b_i^* est la projection orthogonale de b_i sur le supplémentaire orthogonal de $\sum_{j=1}^{i-1} \mathbb{R} b_j$.*

En particulier, la base b^* est orthogonale et on a pour tout $1 \leq i \leq n$, $\sum_{j=1}^{i-1} \mathbb{R} b_j = \sum_{j=1}^{i-1} \mathbb{R} b_j^*$. Par ailleurs, on a la formule

$$b_i^* = b_i - \sum_{j=1}^{i-1} \frac{\langle b_i | b_j^* \rangle}{\|b_j^*\|^2} b_j^* = b_i - \sum_{j=1}^{i-1} \mu_{i,j} b_j^*.$$

La base orthogonalisée de Gram-Schmidt associée à une base d'un réseau permet de minorer ses minima successifs :

Lemme 290. *Soit $b = (b_1, \ldots, b_n)$ une $\mathbb{Z}$-base d'un réseau Γ. Alors pour tout $1 \leq i \leq n$, le i-ème minima de Γ est minoré par*

$$\lambda_i(\Gamma) \geq \min_{i \leq j \leq n} \|b_j^*\|.$$

Preuve : En effet, soit $e = (e_1, \ldots, e_n)$ une famille libre de vecteurs courts de Γ. Pour tout $1 \leq i \leq n$, on note $(\alpha_{i,j})_{1 \leq j \leq n}$ (resp. $\beta_{i,j}$) les coordonnées de e_i sur la $\mathbb{Z}$-base b (resp. la base b^*) et $\ell_i = \max\{j \mid e_{i,j} \neq 0\}$ de sorte que l'on a

$$e_i = \sum_{j=1}^{\ell_i} \alpha_{i,j} b_j = \sum_{j=1}^{\ell_i} \beta_{i,j} b_j^*.$$

Puisque la famille est libre, il est clair qu'il existe $j \leq i$ tel que $\ell_j \geq i$. On a donc

$$\lambda_i(\Gamma) \geq \lambda_j(\Gamma) = \|e_j\| = \|\sum_{k=1}^{\ell_j} \beta_{j,k} b_j^*\| \geq |\beta_{j,\ell_j}| \|b_j^*\|.$$

Mais $\beta_{j,\ell_j} = \alpha_{j,\ell_j} \in \mathbb{Z}^*$, d'où le résultat.

Définition 291. *Soit b une $\mathbb{Z}$-base de Γ et $1/4 < \delta < 1$. On dit que b est LLL-réduite à un facteur δ si*

1. *elle est faiblement réduite, i.e. si pour tout $1 \leq j < i \leq n$, on a $|\mu_{i,j}| \leq 1/2$.*
2. *pour tout $1 \leq i \leq n-1$, on a $\|b_{i+1}^* + \mu_{i+1,i} b_i^*\|^2 \geq \delta \|b_i^*\|^2$ (condition de Lovász).*

En général, si b est une base et β est la base obtenue à partir de b en intervertissant les deux vecteurs b_i et b_{i+1}, alors $\beta_i^* = b_{i+1}^* + \mu_{i+1,i} b_i^*$. Ainsi, la condition de Lovász permet d'assurer que l'on ne gagne pas trop en changeant l'ordre de la base.

Proposition 292. *Soit $1/4 < \delta < 1$ et $\alpha = \frac{1}{\delta - 1/4}$. Si b est une $\mathbb{Z}$-base LLL-réduite à u facteur δ du réseau Γ, alors pour tout $1 \leq i \leq n$,*

$$\|b_i\| \leq \alpha^{\frac{d-1}{2}} \lambda_i(\Gamma).$$

Pour prouver cette proposition, il faut essentiellement remarquer que la condition de Lovász, et la réduction faible donnent

$$\|b_{i+1}^*\|^2 \geq (\delta - \mu_{i+1,i}^2) \|b_i^*\|^2 \geq (\delta - 1/4) \|b_i^*\|^2.$$

On en déduit que pour tout $1 \leq i \leq j \leq n$, $\|b_j^*\|^2 \geq (\delta - 1/4)^{j-i} \|b_i^*\|^2$. D'où

$$\begin{aligned} \|b_i\|^2 &= \left\| b_i^* + \textstyle\sum_{k=1}^{i-1} \mu_{i,j} b_j^* \right\|^2 \\ &= \|b_i^*\|^2 + \textstyle\sum_{k=1}^{i-1} |\mu_{i,j}|^2 \|b_j^*\|^2 \\ &\leq \left(\alpha^{j-i} + \frac{1}{4} \textstyle\sum_{k=1}^{i-1} \alpha^{j-k} \right) \|b_j^*\| \\ &\leq \alpha^{j-1} \|b_j^*\| \end{aligned}$$

Le lemme 290 permet donc de conclure.

Ainsi, si une base est LLL-réduite, elle est certainement proche des minima du réseau. Reste à savoir calculer efficacement une telle base. L'algorithme LLL (du nom de ses inventeurs Lenstra, Lenstra et Lovász) permet de le faire en temps polynomial :

8 Exercices

Exercice 1. *Montrer qu'un groupe G tel que pour tout $x \in G$, on a $x^2 = e$, est nécessairement commutatif.*

Exercice 2. *Soient G un groupe et H, K deux sous-groupes. On suppose que soit H soit K est distingué, montrer alors que $HK = \{hk,\ h \in H \text{ et } k \in K\}$ est un sous-groupe de G.*

Exercice 3. *Notons D_n le nombre de dérangements de $\mathfrak{S}_n$ c'est à dire l'ensemble des permutations de $\mathfrak{S}_n$ sans point fixe. Montrer les propriétés suivantes :*

(1) $\sum_{k=0}^{n} \binom{k}{n} D_{n-k} = n!$;

(2) $D_n = n!\left(\sum_{k=0}^{n} \frac{(-1)^k}{k!}\right)$;

(3) la série génératrice $\sum_{k\geq 0} \frac{D_k z^k}{k!}$ *est égale à* $\frac{e^{-z}}{1-z}$*, son rayon de convergence est e ;*

(4) $D_k = \lfloor \frac{k!}{e} + \frac{1}{2} \rfloor$;

(5) $D_{n+1} = n(D_n + D_{n-1})$;

(6) $D_n = nD_{n-1} + (-1)^n$.

Exercice 4. *Un paysan a $2n+1$ vaches telles, l'une quelconque de ses vaches étant mises de côté, il peut répartir les $2n$ restantes en deux sous-troupeaux de même poids total. Montrer que toutes les vaches ont le même poids.*

Exercice 5. *Une fonction $f : \mathbb{Z} \to \mathbb{Z}$ est appelée un quasi-endomorphisme s'il existe une constante C_f telle que*

$$\forall m, n \in \mathbb{Z},\ |f(n+m) - f(n) - f(m)| \leq C_f.$$

On note $\mathcal{Q}$ l'ensemble des quasi-endomorphismes de $\mathbb{Z}$.

(1) Montrer que l'addition usuelle des fonctions munie $\mathcal{Q}$ d'une loi de groupe abélien.

(2) Montrer que $\mathcal{Q}$ est stable par la composition des fonctions ; $(\mathcal{Q}, +, \circ)$ est-il un anneau ?

(3) Montrer que pour tout $k \geq 2$, on a

$$\left| f(n_1 + \cdots + n_k) - f(n_1) - \cdots - f(n_k) \right| \leq (k-1)C_f.$$

(4) *Montrer que la suite* $\left(\frac{f(n)}{n}\right)_{n\geq 1}$ *est convergente dans* $\mathbb{R}$.

Indication : on pourra utiliser une majoration de $\left|\frac{f(nm)}{nm} - \frac{f(n)}{n}\right|$ *afin de montrer que la suite* $\left(\frac{f(n)}{n}\right)_{n\geq 1}$ *est de Cauchy.*

(5) *Soit* $l : \mathcal{Q} \to \mathbb{R}$ *qui à* f *associe* $\lim_{n\to\infty} \frac{f(n)}{n}$. *Montrer que* l *est un morphisme surjectif de groupes dont le noyau* $\mathcal{K}$ *est l'ensemble des quasi-endomorphismes bornés.*

(6) *Montrer que pour tout* $f, g \in \mathcal{Q}$, *on a* $l(g \circ f) = l(g)l(f)$.

Remarque: le quotient $\mathcal{Q}/\mathcal{K}$ *est donc isomorphe à* $\mathbb{R}$. *Notons qu'il est possible de construire* $\mathbb{R}$ *en utilisant les quasi-endomorphismes de* $\mathbb{Z}$. *Il s'agit alors de montrer que le quotient* $\mathcal{Q}/\mathcal{K}$ *muni de* $\circ$ *est un anneau que l'on peut munir d'une relation d'ordre total :* $f \leq g$ *si et seulement si* $f - g$ *est borné. On montre enfin que* $\mathcal{Q}/\mathcal{K}$ *est un corps satisfaisant la propriété de la borne supérieure. Pour une preuve complète on renvoie le lecteur à http ://www.math.mq.edu.au/∼street/EffR.pdf*

Exercice 6. *Soient* G *un groupe fini, d'élément neutre* e, *et* x *un élément de* G *d'ordre* m.

1. *Montrer que pour tout entier* k, *l'ordre de* x^k *est* $\frac{m}{(m\wedge k)}$.
2. *Soit* n *un entier* ≥ 1. *Montrer que les conditions suivantes sont équivalentes :*
   — *on a* $m = n$.
   — *On a* $x^n = e$ *et pour tout diviseur premier* p *de* n, *on a* $x^{\frac{n}{p}} \neq e$.

Exercice 7.

1. *Montrer que dans un groupe fini d'ordre impair, tout élément est un carré.*
2. *Soient* G *un groupe cyclique d'ordre* n *pair, d'élément neutre* e. *Montrer que* G *possède exactement* $\frac{n}{2}$ *éléments qui sont des carrés.*
3. *Soit* a *un élément de* G. *Montrer l'équivalence*

$$a \text{ est un carré dans } G \iff a^{\frac{n}{2}} = e.$$

4. *Supposons que* n *soit une puissance de* 2. *Montrer que l'ensemble des générateurs de* G *est l'ensemble des éléments qui ne sont pas des carrés.*

Exercice 8. ***Puissances dans un groupe cyclique*** *Soient* G *un groupe cyclique d'ordre* n, *d'élément neutre* e, *et* a *un élément de* G.

1. *Soit* k *un entier naturel. Montrer que pour qu'il existe* $x \in G$ *tel que* $x^k = a$ *il faut et il suffit que l'on ait*

$$a^{\frac{n}{d}} = e \quad \text{où} \quad d = (k \wedge n). \tag{1}$$

2. *Soit k un entier naturel tel que la condition (1) soit satisfaite. Soit x_0 un élément de G tel que $x_0^k = a$. Montrer que l'ensemble des éléments $x \in G$ tels que $x^k = a$ est*

$$\Big\{x_0 z \mid z \in G \text{ et } z^d = e\Big\},$$

et que son cardinal est d.

3. *On prend pour G le groupe additif $\mathbb{Z}/25\mathbb{Z}$. Déterminer dans G l'ensemble des solutions de l'équation $5x = \overline{15}$.*

Exercice 9. *Soit A un anneau intègre.*

1. *Montrer qu'un idéal $\mathcal{P}$ de A est premier si et seulement s'il vérifie la propriété suivante :*

$$xy \in \mathcal{P} \text{ et } x \notin \mathcal{P} \Rightarrow y \in \mathcal{P}.$$

2. *Soient $\mathcal{P}$ un idéal premier de A et $I_1, \cdots, I_r$ des idéaux tels que $I_1.\cdots.I_r \subset \mathcal{P}$. Montrer que $\mathcal{P}$ contient l'un des I_k.*

3. *Soit I un idéal non premier de A, montrer qu'il existe deux idéaux I_1 et I_2 tels que $I \subset I_1$, $I \subset I_2$ et $I_1.I_2 \subset I$.*
En utilisant le lemme de Zorn, montrer l'existence d'idéaux premiers minimaux pour l'inclusion. En supposant A noethérien, montrer qu'il existe un nombre fini d'idéaux premiers minimaux.

4. *Un idéal $\mathcal{Q}$ sera dit primaire s'il vérifie :*

$$\forall x, y \in A \; xy \in \mathcal{Q},\; x \notin \mathcal{Q} \Rightarrow \exists n \; y^n \in \mathcal{Q}.$$

Si $\mathcal{Q}$ est primaire que peut-on dire de $A/\mathcal{Q}$? Pour tout idéal I de A, on pose

$$\sqrt{I} = \{x \in A \;/\; \exists n \; x^n \in I\}.$$

Montrer que Q primaire entraine que sa racine est un idéal premier. Réciproquement : soit $I = (X),\; n > 1$ $J = (X, Y)^n$ dans $A = \mathbb{C}[X, Y]$. Montrer que $Q = I \cap J$ n'est pas primaire bien que son radical soit premier.

Exercice 10. *Soient $\mathfrak{A}$ et $\mathfrak{B}$ des idéaux d'un anneau A. On définit alors*

$$\mathfrak{A} : \mathfrak{B} = \{a \in A : \; ab \in \mathfrak{A} \; \forall b \in \mathfrak{B}\}$$

Montrez que $\mathfrak{A} : \mathfrak{B}$ est un idéal de A tel que :

1. $(\mathfrak{A} : \mathfrak{C}) + (\mathfrak{B} : \mathfrak{C}) \subset (\mathfrak{A} + \mathfrak{B}) : \mathfrak{C}$;
2. $\mathfrak{A} : (\mathfrak{B} + \mathfrak{C}) = (\mathfrak{A} : \mathfrak{B}) \cap (\mathfrak{A} : \mathfrak{C})$;
3. $(\mathfrak{A} : \mathfrak{B}) : \mathfrak{C} = \mathfrak{A} : (\mathfrak{B}\mathfrak{C})$.

Exercice 11. *Quels sont les idéaux bilatères de $\mathcal{L}(E)$?*

Exercice 12. *Quels sont les idéaux à gauche de $\mathcal{L}(E)$?*

Exercice 13. *Quels sont les idéaux à droite de $\mathcal{L}(E)$?*

Exercice 14. *Soit k un corps et $A = k[[X]]$ l'algèbre des séries formelles à coefficients dans k.*

1. *Montrer que A est intègre et déterminer $A^\times$.*
2. *Montrer que tout idéal non nul de A est de la forme $X^n A$, $n \in \mathbb{N}$. En déduire que A est principal et déterminer ses éléments irréductibles.*
3. *Montrer que A est euclidien*

Exercice 15. *Soient $\epsilon > 0$ et $(\alpha_1, \cdots, \alpha_n) \in \mathbb{R}^n \backslash \mathbb{Q}^n$. Montrer qu'il existe $p_1, \cdots, p_n \in \mathbb{Z}$ et $q \in \mathbb{N}$ non nul tels que pour tout $1 \leq i \leq n$, on ait $|\alpha_i - \frac{p_i}{q}| < \frac{\epsilon}{q}$.*

Exercice 16. *On considère le réseau $\mathbb{Z}^2$ de $\mathbb{R}^2$.*

1. *Soit D une droite du plan d'équation $y = ax + b$ avec $a = m/n \in \mathbb{Q}$ écrit sous forme irréductible. Montrer que D intersecte $\mathbb{Z}^2$ si et seulement si $b = r/s \in \mathbb{Q}$, pour $r \wedge s = 1$, avec $s|n$. Expliciter alors $D \cap \mathbb{Z}^2$*
2. *Dans le cas où $a \notin \mathbb{Q}$, montrer que $D \cap \mathbb{Z}^2$ est soit vide soit réduit à un unique point.*
3. *Dans le cas où $a \notin \mathbb{Q}$, montrer que pour tout $\epsilon > 0$, il existe une infinité de points $P \in \mathbb{Z}^2$ tels que $d(P, D) < \epsilon$.*
4. *Soit $\tan\theta = m/n$ avec $m \wedge n = 1$ et $n \neq 0$. Montrer qu'il existe alors une bande de largeur $d = (m^2 + n^2)^{-1/2}$ dans la direction de θ ne contenant aucun point de $\mathbb{Z}^2$ dans son intérieur et toute bande de largeur $> d$ contient des points de $\mathbb{Z}^2$.*
5. *On s'intéresse aux solutions positives de $ax+by = n$ avec donc $(a, b) \in \mathbb{N}^2$ et premiers entre eux. Notons $A(n/a, 0)$ et $B(0, n/b)$ les points de D sur les axes. Montrer que le nombre de solutions est de la forme $\lfloor \frac{n}{ab} \rfloor + \epsilon$ avec $\epsilon = -1, 0, 1$ selon les cas (si $n < ab - a - b$, si a ou b ou ab divise n ou non).*

Exercice 17. ***Rectangles recouvrant*** *: ce sont des rectangles tels que quelque soit la translation qui leur soit appliquée, ils contiennent toujours un point de $\mathbb{Z}^2$ dans leur adhérence. Montrer qu'un rectangle de côté a, b est recouvrant si et seulement si $a \geq 1$ et $b \geq \sqrt{2}$.*

Exercice 18. *Montrer que pour tout entier n, il existe un carré dont le bord passe par exactement n points de $\mathbb{Z}^2$.*

Exercice 19. *Montrer que pour tout entier n, il existe un carré contenant exactement n points de $\mathbb{Z}^2$.*

Exercice 20. *Montrer que pour tout entier n, il existe un cercle qui passe par exactement n points de $\mathbb{Z}^2$.*

Exercice 21. *Montrer que pour tout entier n, il existe un disque qui contient exactement n points de $\mathbb{Z}^2$.*

Exercice 22. *Soient a, b, c, m quatre entiers tels que*

(i) $(x, y) \mapsto ax^2 + bxy + cy^2$ *est une forme quadratique définie positive,*

(ii) il existe $f \in \mathbb{Z}$ *tel que* $af^2 + bf + c \equiv 0\ [m]$,

(iii) $\sqrt{4ac - b^2} < \pi$.

Montrer qu'il existe $x, y \in \mathbb{Z}$ *tels que* $ax^2 + bxy + cy^2 = m$.

Exercice 23. *Dans ce exercice, I désigne un intervalle de la forme $]a, b[$ avec $-\infty \leq a < b \leq +\infty$. Soient $f_1, \cdots, f_n$ des fonctions d'une variable réelle t, dérivables $n-1$ fois sur I ; leur Wronskien $W(f_1, \cdots, f_n)$ est le déterminant*

$$W(f_1, \cdots, f_n)(t) = \begin{vmatrix} f_1(t) & f_2(t) & \cdots & f_n(t) \\ f_1'(t) & f_2'(t) & \cdots & f_n'(t) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(t) & f_2^{(n-1)}(t) & \cdots & f_n^{(n-1)}(t) \end{vmatrix}.$$

1. *Montrer que si $f_1, \cdots, f_n$ sont linéairement dépendantes sur I alors $W(f_1, \cdots, f_n)$ est identiquement nulle sur cet intervalle.*
2. *Réciproquement montrer que si $W(f_1, \cdots, f_n)$ est la fonction nulle sur I alors que $f_1, \cdots, f_n$ sont des solutions de l'équation différentielle linéaire homogène*

$$y^{(n)} + p_{n-1}(t)y^{(n-1)} + \cdots + p_1(t)y' + p_0(t)y = 0$$

   avec $p_0, \cdots, p_{n-1}$ des fonctions définies et continues sur I, alors $f_1, \cdots, f_n$ sont linéairement dépendantes sur I.
3. *Soient $f_1(t) = t^2$ et $f_2(t) = t|t|$; montrer que $W(f_1, f_2) = 0$ alors que f_1 et f_2 ne sont pas linéairement dépendantes sur $\mathbb{R}$.*
4. *Dans cette question nous allons prouver le résultat suivant : si $W(f_1, \cdots, f_n) \equiv 0$ sur I, il existe alors un sous-intervalle $J = [b, c] \subset I$ avec $b < c$, sur lequel $f_1, \cdots, f_n$ sont linéairement dépendantes.*
   (a) Soit g une fonction dérivable $(n-1)$-fois, montrer que

$$W(f_1 g, f_2 g, \cdots, f_n g) = g^n W(f_1, \cdots, f_n).$$

(b) Pour f_1 une fonction qui ne s'annule pas, montrer que

$$W(f_1,\cdots,f_n) = f_1^n W\Big((\frac{f_2}{f_1})',(\frac{f_3}{f_1})',\cdots,(\frac{f_n}{f_1})'\Big)$$

(c) Prouver le résultat annoncé plus haut.

(d) Quelle conclusion en tirez-vous si on suppose que les f_i sont des fonctions polynômes ?

Exercice 24. *Soient P et Q deux polynômes non constants de $\mathbb{C}[X]$ tels que l'ensemble des racines de P (resp. $P-1$) soit égal à l'ensemble des racines de Q (resp. $Q-1$). Montrer qu'alors $P=Q$.*

Exercice 25. *Soit $P(X) = a_0 + a_1X + \cdots + a_dX^d \in \mathbb{Z}[X]$ avec $a_d \neq 0$ et notons α_i pour $i = 1,\cdots,n$ ses racines. Montrer alors que*

$$\mathrm{sep}P = \inf_{\alpha_i\neq\alpha_j} |\alpha_i - \alpha_j| \geq (2C)^{1-\frac{d(d-1)}{2}}$$

où $C = |a_d| + \sum_{1\leq i\leq d-1} |a_i|$.
sol : on pourra commencer par traiter le cas où les racines sont simples.

Exercice 26. *Comptage des polynômes irréductibles*

1. *Soit $p \geq 3$. Montrer que la proportion des polynôme de degré d dans $\mathbb{F}_p[X]$ qui sont irréductibles unitaires est au moins égale à $\frac{1}{2d}$. (cf. aussi l'exercice **??**)*
2. *Montrer que parmi les polynômes $P \in \mathbb{Z}[T]$ unitaires de degré $d \geq 1$, à coefficients dans $[-N,N]$, la proportion de ceux qui sont irréductibles, tend vers 1 lorsque N tend vers $+\infty$.*

Exercice 27. ***Irréductibilité du déterminant*** *On considère le déterminant* det *comme un élément de $A = \mathbb{Z}[x_{1,1},\cdots,x_{n,n}]$; montrer alors qu'il est irréductible.*

Exercice 28. ***Pfaffien*** *(cf. [?] 2.3) Soit $X = [x_{i,j}] \in \mathbb{M}_n(K)$ une matrice antisymétrique.*

1. *Montrer que $\det X$ est un carré de K nul si n est impair.*
2. *On suppose à présent $n = 2m$ pair. Montrer que $\det X \in \mathbb{Z}[(x_{i,j})_{1\leq i<j\leq n}]$ est le carré d'un polynôme $Pf \in \mathbb{Z}[(x_{i,j})_{1\leq i<j\leq n}]$ qui vérifie $Pf(\mathrm{diag}(J,\cdots,J) = 1$ et $Pf({}^tQXP) = Pf(X)\det Q$.*
3. *Montrer que Pf est linéaire en la première ligne ou en la première colonne et en déduire que $Pf \in \mathbb{Z}[(x_{i,j})_{1\leq i<j\leq n}]$ est irréductible.*

Exercice 29. *Montrer que les seules matrices de Bourdaud (i.e. les valeurs propres se lisent sur la diagonale) réelles symétriques sont les matrices diagonales.*

Exercice 30. *Soit G un sous-groupe de $O(n)$ alors G fini si et seulement si G est d'exposant fini si et seulement si l'ensemble des traces des éléments de G est fini.*

Exercice 31. *Montrer que u est diagonalisable de spectre réel si et seulement si u est le produit de deux endomorphismes hermitiens, l'un au moins d'entre eux étant défini positif*

Exercice 32. *Pour $n \geq 2$ et A hermitienne non nulle de taille n, montrer que A est définie positive ou négative si et seulement si pour toute matrice B hermitienne, AB est diagonalisable.*

Exercice 33. *Soient $\lambda_1, \cdots, \lambda_n$ les valeurs propres d'une matrice complexe $A = (a_{i,j})$. Montrer que A est normale si et seulement si $\sum_{i,j} |a_{i,j}|^2 = \sum_i |\lambda_i|^2$.*

Exercice 34. *Montrer que A est normale si et seulement si $\mathrm{tr}(AA^*)^2 = \mathrm{tr}A^2A^{*2}$.*

Exercice 35. *Montrer que A est normale si et seulement si A^* est un polynôme en A*

Exercice 36.
1. *Montrer que le sous-espace vectoriel engendré par le cône nilpotent (ou par les matrices nilpotentes de rang 1) est l'hyperplan des matrices de trace nulle.*
2. *En déduire que le sous-espace vectoriel engendré par les matrices d'une classe de similitude quelconque de matrices nilpotentes est l'hyperplan des matrices de trace nulle.*

Exercice 37. *Une matrice $M = [m_{i,j}] \in \mathbb{M}_n(K)$ est dite de Hessenberg si $m_{j,k} = 0$ pour tout (j,k) tel que $j - k \geq 2$.*

1. *(cf. [?] 10.1.1) Montrer que pour tout $M \in \mathbb{M}_n(\mathbb{C})$, il existe une matrice unitaire $U \in GL_n(\mathbb{C})$ telle que $U^{-1}MU$ est de Hessenberg. Montrer que si M est réelle alors on peut alors prendre $U \in O_n$.*
2. *On notera que si M est hermitienne alors $A = U^{-1}MU$ es tridiagonale avec $a_{j,j+1} = \overline{a_{j+1,j}}$ et $a_{j,j} \in \mathbb{R}$. On écrit*

$$A = \begin{pmatrix} m & \overline{a} & 0 & \cdots & 0 \\ a & & & & \\ 0 & A_1 & & & \\ \vdots & & & & \\ 0 & & & & \end{pmatrix}$$

et on note χ_{A_0}, χ_{A_1} les polynômes caractéristiques respectifs de A et A_1. Justifier que ces polynômes sont réels puis montrer que

$$\chi_A(X) = m\chi_{A_1}(X) - |a|^2\chi_{A_2}(X)$$

où A_2 est la matrice extraite de A_1 constituée de ses $n-2$ dernières lignes et colonnes.

3. *Montrer que la suite $(\chi_{A_j}(X))_{0\leq j\leq n-1}$ est une suite de Sturm et en déduire un algorithme permettant de calculer le nombre de racines de $\chi_{A_0}(X)$ dans un intervalle $[a,b]\subset\mathbb{R}$ quelconque. Quel est l'intérêt de cette méthode ?*

Exercice 38. *(cf. [?] 4.1.7) Soit $A\in\mathbb{M}_n(\mathbb{C})$; montrer que les propriétés suivantes sont équivalentes :*

(a) A est semblable à une matrice $B\in\mathbb{M}_n(\mathbb{R})$;
(b) A est semblable à A^* ;
(c) A est semblable à A^* via une matrice de passage hermitienne ;
(d) $A=HK$ avec H,K hermitiennes, au moins une des deux étant inversible ;
(e) $A=HK$ avec H,K hermitienne.

Exercice 39. *Soit A une matrice hermitienne, montrer que A est définie positive si et seulement si tous se mineures principaux $\det A_i$ sont strictement positifs pour $i=1,\cdots,n$ où on rappelle que $A_i\in\mathbb{M}_i(\mathbb{C})$ désigne la matrice extraite de A constitué des i premières lignes et colonnes.*

Exercice 40. *Soit $A\in\mathbb{M}_n(\mathbb{C})$ et $A=PU$ sa décomposition polaire. Montrer que A est normale si et seulement si $PU=UP$.*

Exercice 41. *Montrer que si un ouvert de $\mathbb{M}_n(\mathbb{C})$ contient les matrices diagonales et est stable par similitude, alors il est égal à $\mathbb{M}_n(\mathbb{C})$ tout entier.*

Exercice 42. *Montrer que si une classe de similitude de $\mathbb{M}_n(\mathbb{C})$ n'est constituée que de matrices normales, elle est alors réduite à un seul élément. Même question si on suppose que la classe de similitude ne contient qu'un nombre fini de matrices normales.*

Exercice 43. *Quelles sont les classes de similitudes bornées de $\mathbb{M}_n(\mathbb{C})$?*

Exercice 44. *Montrer que sur $\mathbb{C}$, toute classe de similitude contient une matrice symétrique et que sur $\mathbb{R}$, pour qu'une classe de similitude contienne des matrices symétriques, il faut et il suffit qu'elle contienne une matrice diagonale. En déduire en particulier que l'ensemble des matrices symétriques d'une classe de similitude est compacte connexe.*

Exercice 45. *Montrer que l'ensemble des matrices diagonalisables de $\mathbb{M}_n(\mathbb{C})$ est connexe et dense. Quel est son intérieur ? Ce dernier est-il encore connexe ?*

Exercice 46. *Montrer que l'ensemble des matrices de rang r est connexe et son adhérence est l'ensemble des matrices de rang inférieur ou égale à r.*

Exercice 47. *Soit $M=[m_{i,j}]\in\mathbb{M}_n(K)$ une matrice de Hessenberg telle que $m_{j+1,j}\neq 0$ pour tout $j=1,\cdots,n$. Montrez que les sous-espaces propres de M sont de dimension 1.*

Exercice 48. *Soit u un endomorphisme de $V \simeq K^n$ dont on note χ_u et π_u respectivement les polynômes caractéristique et minimal.*

1. *Montrer que χ_u est irréductible si et seulement si V n'a pas de sous-espace stable par u ;*
2. *Montrer que u est cyclique avec π_u une puissance d'un polynôme irréductible si et seulement si V est indécomposable sous u.*
3. *Donner un algorithme pour tester si u est semi-simple.*

Exercice 49. *Quels sont les endomorphismes complexes u qui ne possèdent qu'un nombre fini de sous-espaces stables ?*

Exercice 50. *Quels sont les endomorphismes u tels que tout sous-espace stable est de la forme $\operatorname{Ker} P(u)$ ou $\operatorname{Im} P(u)$ pour P un polynôme.*

Exercice 51. *Étant donné un drapeau et le parabolique associé montrez que les sous-espaces stables par tous les éléments de sous-groupe parabolique sont ceux du drapeau.*

Exercice 52.

1. *Soit E un K-espace vectoriel et F un sous-espace. Montrer que l'ensemble des supplémentaires de F dans E est un espace affine de direction $\operatorname{Hom}_K(E/F, F)$.*
2. *Soit u un endomorphisme de E et soit F un sous-espace stable par u. Montrer que l'ensemble des supplémentaires de F stables par u est, quand il est non vide, un espace affine de direction le sous-espace vectoriel de $\operatorname{Hom}_K(E/F, F)$ des s tels que $s \circ \overline{u} - u_{|F} \circ s = 0$.*
3. *Soit F un sous-espace stable sous M ; on suppose qu'il existe un supplémentaire G stable $E = F \oplus G$. Donnez une condition nécessaire et suffisante pour que G soit l'unique supplémentaire stable de F.*

Exercice 53. ***Endoscopie***

1. *Pour $K \subset L$ une extension de corps, montrer que deux matrices de $GL_n(K)$ sont semblables dans $GL_n(L)$ si et seulement si elles le sont dans $GL_n(K)$.*
2. *Soit $R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Pour $\theta \in \mathbb{R}\backslash\pi\mathbb{Z}$, montrer que $R(\theta)$ et $R(-\theta)$ sont semblables dans $SL_2(\mathbb{C})$ et $GL_2(\mathbb{R})$ mais ne sont pas semblables dans $SL_2(\mathbb{R})$;*

1 On écrit $(xy)^2 = e$ ce qui donne $xyxy = e$ soit $yx = x^{-1}y^{-1} = xy$ car l'égalité $x^2 = e$ (resp. $y^2 = e$) donne $x^{-1} = x$ (resp. $y^{-1} = y$).

2 Soient h_1k_1 et h_2k_2 des éléments de KH ; si H (resp. K) est distingué, pour tout $h \in H$ et $k \in K$ on a $hk = kh'$ (resp. $hk = k'h$) pour $h' \in H$ (resp. $k' \in K$) de sorte que $h_1k_1h_2k_2 \in HK$. En ce qui concerne l'inverse on écrit $k^{-1}h^{-1}$ sous la forme $h'k^{-1}$ (resp. $h^{-1}k'$) et donc HK est un sous-groupe de G.

3 (1) La formule découle directement de la partition de $\mathfrak{S}_n$ selon le cardinal du support

(2) La relation découle directement de la première question ou se prouve en utilisant la formule du crible à l'égalité

$$D_n = n! - \sharp\Big(\bigcup_{i=1}^{n} U_i\Big)$$

où U_i désigne le sous-ensemble de $\mathfrak{S}_n$ des permutations fixant i.

(3) La formule se prouve en utilisant la première question ou alors directement par une preuve purement de combinatoire algébrique.

(4) L'égalité découle de

$$\frac{k!}{e} + \frac{1}{2} = D_k + \frac{1}{2} + k!R_k, \qquad R_k = \sum_{n=k+1}^{+\infty} \frac{(-1)^n}{n!}$$

et de la minoration des séries alternées $|R_k| \leq \frac{1}{(k+1)!}$.

(5) La relation de récurrence se prouve comme suit : soit l'orbite de $n+1$ est de cardinal 2 ce qui donne nD_{n-1} possibilité pour un tel dérangement et sinon nD_n.

(6) La relation se prouve à partir de la question précédente par récurrence ou peut se montrer purement combinatoirement mais de manière détournée pour l'instant.

4 L'énoncé se traduit matriciellement comme suit : il existe une matrice $A \in \mathbb{M}_{2n+1}(\mathbb{R})$ telle que

— $a_{i,i} = 0$ (on met la i-ème vache de côté) ;
— $a_{i,j} = \pm 1$ si $j \neq i$ (le signe dépend dans quel sous-troupeau on met la j-ème vache quand la i-ème est de côté) ;
— $\sum_{j=1}^{2n+1} a_{i,j} = 0$ (les deux sous-troupeaux sont de même cardinal) ;
— $\sum_{j=1}^{2n+1} a_{i,j}p_j = 0$, où p_j est le poids de la j-ème vache (les deux sous-troupeaux ont même poids total).

Le résultat découle alors directement du fait suivant : toute matrice $A \in \mathbb{M}_{2n+1}(\mathbb{R})$ à coefficients diagonaux nuls, les autres étant égaux à ± 1 est de rang $2n$. En effet comme le vecteur n'ayant que des 1 est dans le noyau, le

vecteur des p_i qui est aussi dans le noyau lui sera proportionnel et donc tous les p_i seront égaux. Considérons la matrice extraite B obtenue en ôtant la dernière ligne et colonne et montrons qu'elle est inversible. Son déterminant est

$$\delta = \sum_{\sigma \in \mathfrak{S}_{2n}} \epsilon(\sigma) a_{1,\sigma(1)} \cdots a_{2n,\sigma(2n)}.$$

Les termes diagonaux étant nuls, les seuls termes non nuls sont ceux pour lesquels σ est un dérangement. Par ailleurs chacun de ces termes étant égal à ± 1, il suffit de montrer que $D_n \equiv 1 \mod 2$. Comme $D_{2n-1} = (2n-2)(D_{2n-2} + D_{2n-3})$ il est pair et $D_{2n} = (2n-1)(D_{2n-1} + D_{2n-2})$ est de la même parité que D_{2n-2} ; on conclut par récurrence.

5 (1) C'est clair en posant $C_{f+g} = C_f + C_g$ et en utilisant l'inégalité triangulaire.

(2) On écrit

$$\Big|g \circ f(m+n) - g \circ f(n) - g \circ f(m)\Big| \leq$$
$$\Big|g \circ f(n+m) - g\big(f(n) + f(m)\big)\Big| + \Big|g\big(f(n) + f(m)\big) - g(f(n)) - g(f(m))\Big|$$

Le deuxième terme du membre de droite est majoré par C_g et quant au premier en écrivant $f(n+m) = f(n) + f(m) + \delta$ avec $|\delta| \leq C_f$ et en notant $M = \max_{|i| \leq M} |g(i)|$, on obtient

$$\Big|g \circ f(n+m) - g\big(f(n) + f(m)\big)\Big| \leq M + C_g$$

et donc en posant $C_{g \circ f} = M + 2C_g$, on a bien $g \circ f \in \mathcal{Q}$.
Remarque: comme $\circ$ n'est pas distributive par rapport à l'addition, $(\mathcal{Q}, +, \circ)$ n'est pas un anneau.

(3) Par récurrence sur k ; supposons le résultat acquis jusqu'au rang $k-1$, on écrit alors

$$\begin{aligned} f(n_1 + \cdots + n_k) - f(n_1) - \cdots - f(n_k) = f(n_1 + \cdots + n_k) - f(n_1 + \cdots + n_{k-1}) \\ - f(n_k) + f(n_1 + \cdots + n_{k-1}) - f(n_1) - \cdots - f(n_{k-1}) \end{aligned}$$

(4) D'après la question précédente, on a $|f(nm) - mf(n)| \leq mC_f$ et $|f(nm) - nf(m)| \leq C_f$, de sorte que

$$\Big|\frac{f(nm}{nm} - \frac{f(n)}{n}\Big| \leq \frac{C_f}{n} \text{ et } \Big|\frac{f(nm}{nm} - \frac{f(m)}{m}\Big| \leq \frac{C_f}{n}$$

et par inégalité triangulaire on obtient

$$\Big|\frac{f(m}{m} - \frac{f(n)}{n}\Big| \leq \frac{C_f}{n} + \frac{C_f}{m}$$

et donc $\left(\frac{f(n)}{n}\right)_{n\geq 1}$ est de Cauchy.

(5) Pour la surjectivité prendre par exemple $f(x) = \lfloor nx \rfloor$ pour $x \in \mathbb{R}$. Par ailleurs l est clairement un morphisme de groupe ; soit $f \in \mathcal{K}$, on va montrer que $|f(n)| \leq C_f$, pour tout $n \geq 0$, ce qui est déjà le cas pour $n = 0$. On raisonne par l'absurde en considérant $n_0 > 0$ tel que $|f(n_0)| > C_f$; on a alors $|f(kn_0) - kf(n_0)| \leq (k-1)C_f$ et donc

$$|f(kn_0)| \geq |f(n_0)| + (k-1)\Big(|f(n_0)| - C_f\Big).$$

Ainsi la suite extraite $\left(\frac{f(kn_0)}{kn_0}\right)_{k\geq 1}$ ne converge pas vers 0 ce qui n'est pas. Pour les $n < 0$, on écrit $f(0) - f(n) - f(-n)| \leq C_f$ et donc $|f(n)| \leq |f(0)| + |f(-n)| + C_f$ et donc $|f(n)| \leq 3C_f$.

(6) Si $l(f) = 0$ alors f est bornée et donc $g \circ f$ aussi et donc $l(g \circ f) = 0 = l(g)l(f)$. Si $l(f) > 0$ alors comme $f(n) \to +\infty$, il existe n_0 tel que pour tout $n \geq n_0$, $f(n) > 0$ de sorte que le résultat découle de l'égalité suivante que l'on passe à la limite :

$$\frac{g \circ f(n)}{n} = \frac{g \circ f(n)}{f(n)} \frac{f(n)}{n}.$$

Finalement le raisonnement est identique si $l(f) < 0$ en remarquant que $f(n)/n$ tend vers $l(f)$ quand $n \to -\infty$, car $|f(n) + f(-n)| \leq |f(0)| + C_f$.

6 (1) Soit d le plus grand commun diviseur de m et k. On a d'abord $(x^k)^{\frac{m}{d}} = (x^m)^{\frac{k}{d}} = e$, où e est l'élément neutre de G (car $x^m = e$). Considérons alors un entier u tel que $(x^k)^u = e$. L'entier m divise uk (car m est l'ordre de x) et donc m/d divise aussi uk/d. Les entiers m/d et k/d étant premiers entre eux, il en résulte que m/d divise u, ce qui prouve notre assertion.

(2) La première condition entraine la seconde car m est le plus petit entier $k \geq 1$ tel que $x^k = e$. Inversement, supposons la condition 2 réalisée. Il existe un entier $k \geq 1$ tel que l'on ait $n = mk$: on a $n = mk + r$ avec $k \in \mathbb{Z}$ et $0 \leq r < m$, d'où $x^r = e$ puis $r = 0$. Supposons $k \geq 2$. Soit p un diviseur premier de k. On a alors les égalités

$$x^{\frac{n}{p}} = \left(x^m\right)^{\frac{k}{p}} = e,$$

ce qui contredit l'hypothèse faite. Par suite, on a $k = 1$, puis $m = n$.

7 (1) Soit G un groupe fini d'ordre $2n-1$. Pour tout élément $x \in G$, on a $x^{2n-1} = e$ (e est l'élément neutre de G), d'où $x = x^{2n} = (x^n)^2$.

(2) Soit $f : G \to G$ l'application définie par $f(x) = x^2$. C'est un morphisme de groupes. Puisque G est cyclique, G a un unique élément d'ordre 2, et le noyau de f est donc d'ordre 2. Par suite, l'ensemble G^2 des carrés de G est un groupe d'ordre $\frac{n}{2}$.

(3) Soit H le sous-groupe de G formé des éléments x tels que $x^{\frac{n}{2}} = e$. Puisque G est cyclique, H est l'unique sous-groupe d'ordre $\frac{n}{2}$ de G. D'après la question précédente, G^2 et H ont le même ordre. On en déduit que $G^2 = H$, d'où l'équivalence annoncée.

Remarque. Si x est un élément de G qui ne soit pas un carré, $x^{\frac{n}{2}}$ est l'unique élément d'ordre 2 de G.

(4) Supposons $n = 2^t$ avec $t \geq 1$. Dans ce cas, G^2 est de cardinal 2^{t-1} et son complémentaire aussi. Par ailleurs, il y a exactement $\varphi(2^t) = 2^{t-1}$ générateurs dans G (φ est la fonction indicatrice d'Euler). De plus, un générateur de G n'est évidemment pas un carré. Cela entraine le résultat.
Remarque: On peut aussi procéder comme suit : soit x un élément de G qui n'est pas un carré dans G. Si y est un générateur de G, il existe m tel que $x = y^m$. D'après l'hypothèse faite, m est impair, donc x est un générateur, car les générateurs de G sont précisément les éléments de la forme y^k avec k impair (ce sont les entiers k premiers avec l'ordre de G).

8 (1) Considérons le morphisme de groupes $\psi : G \to G$ défini par $\psi(x) = x^k$. Vérifions que son noyau est d'ordre d. Soit x un élément de $\mathrm{Ker}(\psi)$. On a $x^k = e$ et $x^n = e$, d'où en utilisant le théorème de Bézout, $x^d = e$. On en déduit que les éléments de $\mathrm{Ker}(\psi)$ sont exactement les éléments $x \in G$ pour lesquels on a $x^d = e$. Puisque G est cyclique, on a donc $|\mathrm{Ker}(\psi)| = d$ et l'ordre de l'image de ψ est n/d. Par suite, si a est dans l'image de ψ, on a $a^{n/d} = e$. Inversement, si on a l'égalité $a^{n/d} = e$, puisque G est cyclique, a appartient à l'unique sous-groupe de G d'ordre n/d, qui est précisément l'image de ψ, d'où la condition (1) de l'énoncé.

(2) Si $x \in G$ vérifie l'égalité $x^k = a$, on a $(xx_0^{-1})^k = e$, d'où $x = x_0 z$ avec $z^k = e$, et comme on l'a constaté ci-dessus, on a alors $z^d = e$. Inversement, pour tout $z \in G$ tel que $z^d = e$, on a $(x_0 z)^k = a$ car d divise k, d'où l'ensemble des solutions annoncé. Par ailleurs, G étant cyclique, il y a exactement d éléments $z \in G$ tels que $z^d = e$. Cela établit le résultat.

(3) On remarque que $x_0 = \overline{3}$ est une solution particulière. Par ailleurs, les éléments $x \in G$ qui vérifient $5x = \overline{0}$ sont les classes de 0, 5, 10, 15 et 20. L'ensemble des solutions cherché est donc $\{\overline{3}, \overline{8}, \overline{13}, \overline{18}, \overline{23}\}$.

9 (1) La traduction est immédiate : la relation $\overline{xy} = \overline{0}$ est équivalente à $xy \in \mathcal{P}$ pour $x \in \overline{x}$ et $y \in \overline{y}$; ainsi si $\mathcal{P}$ est premier on a x ou y appartient à $\mathcal{P}$ soit $\overline{x} = \overline{0}$ ou $\overline{y} = \overline{0}$. Réciproquement si $xy \in \mathcal{P}$ alors si $A/\mathcal{P}$ est intègre, on a $\overline{x} = \overline{0}$ ou $\overline{y} = \overline{0}$ et donc x ou y appartient à $\mathcal{P}$.

(2) Soit $\mathcal{P}$ premier et $I_1 \cdots I_r \subset \mathcal{P}$, et supposons que pour tout $1 \leq k \leq r$, $I_k \not\subset \mathcal{P}$; on fixe ainsi pour tout k, un élément $x_k \in I_k$ et $x_k \notin \mathcal{P}$. Par hypothèse $x_1 \cdots x_r \in \mathcal{P}$ avec $x_1 \notin \mathcal{P}$ soit $x_2 \cdots x_r \in \mathcal{P}$ et par récurrence $x_r \in \mathcal{P}$ d'où la contradiction.

(3) Soit I non premier, et soit $x, y \in A \backslash I$ avec $xy \in I$. On pose $I_1 = (I \cup \{x\})$ et $I_2 = (I \cup \{y\})$; on a $I_1 I_2 \subset I$ avec $I \subset I_1 \cap I_2$.

On considère la relation d'ordre sur l'ensemble $\mathcal{I}$ des idéaux premiers, donnée par $I \leq J \Leftrightarrow J \subset I$. Cette relation d'ordre est à nouveau inductive, un majorant d'une chaine totalement ordonnée C étant donné par l'intersection $M = \cap_{I \in C} I$. En effet on vérifie comme précédemment que M est un idéal, le fait qu'il soit premier se montre aisément :

— soit $xy \in M$ et donc $xy \in I$ pour tout $I \in C$;
— comme I est premier, x ou y appartient à I ;
— supposons que $y \notin I$, alors pour tout $J \subset I \in C$, on a $y \notin J$ et donc $x \in J$ soit $x \in M$.

Le lemme de Zorn donne alors l'existence d'éléments maximaux qui sont donc des idéaux premiers minimaux pour l'inclusion.

On suppose A noethérien et on considère l'idéal (0) ; s'il est premier alors c'est le seul idéal premier minimal, sinon soit I_1 et I_2 comme ci-dessus. Si I_1 et I_2 sont premiers, ce sont les seuls idéaux premiers minimaux ; en effet soit $\mathcal{P}$ un idéal premier ; on a $(0) = I_1 I_2 \subset \mathcal{P}$ et donc $I_i \subset \mathcal{P}$ pour $i = 1$ ou 2, d'après ce qui précède. Si I_1 n'est pas premier, soit $I_{1,1}$ et $I_{1,2}$ comme ci-dessus ; on construit ainsi un arbre binaire dont la racine est l'idéal (0), tous les sommets sont des idéaux qui contiennent le produit des idéaux de ses deux fils et tel que tout chemin filial définit une chaine totalement ordonnée pour l'inclusion. Si on suppose A noethérien, l'arbre est fini et les idéaux premiers minimaux sont les feuilles.

(3) La traduction dans $A/\mathbb{Q}$ est à nouveau immédiate : $\mathbb{Q}$ est primaire si et seulement si dans $A/\mathbb{Q}$ les seuls diviseurs de 0 sont les éléments nilpotents, i.e. ceux tels qu'il existe n tels qu'élevés à la puissance n, ils donnent 0. Montrons en premier lieu que $\sqrt{I}$ est un idéal. Soient donc $x, y \in \sqrt{I}$, n, m des entiers tels que $x^n \in I$ et $y^m \in I$. On a alors

$$(x+y)^{n+m-1} = \sum_{k=0}^{n+m-1} C_{n+m-1}^k x^k (-y)^{n+m-1-k}.$$

Or $k < n$ si et seulement si $n + m - 1 - k \geq m$ et donc pour tout $0 \leq k \leq n + m - 1$, au moins un parmi x^k et $y^{n+m-1-k}$ appartient à I et donc $x - y \in \sqrt{I}$. Si $a \in A$ alors $(ax)^n \in I$ et donc $ax \in \sqrt{I}$ et donc finalement $\sqrt{I}$ est un idéal de A.

Exemple : dans $\mathbb{Z}$, la racine de l'idéal $n\mathbb{Z}$ avec $n = \prod_i p_i^{\alpha_i}$ est l'idéal engendré par $\prod_i p_i$.

Soit $\mathbb{Q}$ un idéal primaire et $\mathcal{P} = \sqrt{\mathbb{Q}}$; soit $x, y \in A$ tels que $xy \in \mathcal{P}$ et $x \notin \mathcal{P}$. Soit donc un entier n tel que $x^n y^n \in \mathbb{Q}$; comme $x \notin \mathcal{P}$ alors $x^n \notin \mathbb{Q}$ et donc $\mathbb{Q}$ étant primaire, soit m un entier tel que $(y^n)^m \in \mathbb{Q}$ soit $y \in \mathcal{P}$, et donc $\mathcal{P}$ est un idéal premier.

On considère $A = \mathbb{C}[X, Y]$ l'anneau des polynômes en deux variables à coefficients dans $\mathbb{C}$ et soit $I = (X)$ et $J = (X, Y)^n$ avec $n > 1$. On pose $\mathbb{Q} = I \cap J$ qui est l'idéal engendré par $X^n, X^{n-1}Y, \cdots, XY^{n-1}$; on a

$\sqrt{\mathbb{Q}} = (X)$ qui est premier alors que $\mathbb{Q}$ n'est pas primaire car $X^{n-1}Y \in \mathbb{Q}$ avec $X^{n-1} \notin \mathbb{Q}$ et $Y^m \notin \mathbb{Q}$ pour tout entier m.

10 (1) Vérifions tout d'abord que $\mathfrak{A} : \mathfrak{B}$ est un idéal de A : soient $a_1, a_2 \in \mathfrak{A} : \mathfrak{B}$ et $a \in A$, pour tout $b \in \mathfrak{B}$ on a $(r_1 + ar_2)b = r_1 + ar_2b \in \mathfrak{A}$, d'où le résultat. Soit alors $a = a_1 + a_2 \in \mathfrak{A} : \mathfrak{C} + \mathfrak{B} : \mathfrak{C}$ de sorte que pour tout $c \in \mathfrak{C}$, $a_1c \in \mathfrak{A}$ et $a_2c \in \mathfrak{B}$ et donc $ac \in \mathfrak{A} + \mathfrak{B}$ pour tout $c \in \mathfrak{C}$ soit $a \in (\mathfrak{A} + \mathfrak{B}) : \mathfrak{C}$.

(2) Soit $a \in \mathfrak{A} : (\mathfrak{B} + \mathfrak{C})$ alors $a(b + c) \in \mathfrak{A}$ pour tout $b \in \mathfrak{B}$ et $c \in \mathfrak{C}$. En particulier en prenant $c = 0$, on a $ab \in \mathfrak{A}$ pour tout $b \in \mathfrak{B}$ et donc $a \in \mathfrak{A} : \mathfrak{B}$. En procédant de même avec $\mathfrak{C}$, on obtient l'inclusion $\mathfrak{A} : (\mathfrak{B} + \mathfrak{C}) \subset (\mathfrak{A} : \mathfrak{B}) \cap (\mathfrak{A} : \mathfrak{C})$. Réciproquement soit $a \in (\mathfrak{A} : \mathfrak{B}) \cap (\mathfrak{A} : \mathfrak{C})$ alors $ab \in \mathfrak{A}$ pour tout $b \in \mathfrak{G}$ et $ac \in \mathfrak{A}$ pour tout $c \in \mathfrak{C}$ de sorte que $a(b+c) \in \mathfrak{A}$, ce qui donne l'inclusion réciproque.

(3)Soit $a \in (\mathfrak{A} : \mathfrak{B}) : \mathfrak{C}$ de sorte que pour tout $c \in \mathfrak{C}$, $ac \in \mathfrak{A} : \mathfrak{B}$ et donc pour tout $b \in \mathfrak{B}$, on a $acb \in \mathfrak{A}$. Comme $\mathfrak{A}$ est stable par l'addition, on en déduit donc que $ad \in \mathfrak{A}$ pour tout $d \in \mathfrak{B}\mathfrak{C}$ et donc $(\mathfrak{A} : \mathfrak{B}) : \mathfrak{C} \subset \mathfrak{A} : (\mathfrak{B}\mathfrak{C})$. Réciproquement soit $a \in \mathfrak{A} : (\mathfrak{B}\mathfrak{C})$ de sorte que pour tout $d \in \mathfrak{B}\mathfrak{C}$, on a $ad \in \mathfrak{A}$. En particulier pour $d = bc$, c fixé et b décrivant $\mathfrak{C}$, on obtient que $ac \in \mathfrak{A} : \mathfrak{B}$. Comme ce fait est vrai pour tout c, on en déduit que $a \in (AF : \mathfrak{B}) : \mathfrak{C}$.

11 On se ramène à $\mathbb{M}_n(K)$; soit I un idéal bilatère de $\mathbb{M}_n(K)$ et $M = (m_{i,j})_{1 \leq i,j \leq n} \in I$ non nulle. Soit (i_0, j_0) tel que $m_{i_0,j_0} \neq 0$. On a

$$E_{i,i_0} M E_{j_0,j} = \sum_{1 \leq k,l \leq n} m_{k,l} E_{i,i_0} E_{k,l} E_{j_0,j} = \sum_{k=1}^{n} m_{k,j_0} E_{i,i_0} E_{k,l} = m_{i_0,j_0} E_{i,j}$$

de sorte que pour tout (i, j), $E_{i,j} \in I$ et donc $I = \mathbb{M}_n(K)$.

12 Soit I un idéal de $\mathbb{M}_n(\mathbb{C})$, on va montrer que $I = \mathbb{M}_n(\mathbb{C})A = \{M \in \mathbb{M}_n(\mathbb{C}) \ / \ \operatorname{Ker} A \subset \operatorname{Ker} M\}$. Soit $M = PI_rQ \in I$, on a alors $Q^{-1}P^{-1}M \in I$ et donc I contient un projecteur. Pour tout f, on note I_f l'ensemble des endomorphismes qui s'annulent sur $\operatorname{Ker} f$: $I_f = \mathbb{M}_n(\mathbb{C})f$. De l'écriture $I = p + (Id - p)$, on en déduit que $\mathbb{M}_n(\mathbb{C}) = I_p \oplus I_{Id-p}$.

Soit alors p un projecteur de rang maximal dans I, alors $I \cap I_{Id-p} = 0$. En effet il suffit de montrer que l'intersection ne contient aucun projecteur q. Soit donc un projecteur q qui s'annule sur $\operatorname{Ker}(Id - p) = \operatorname{Im} p$. Dans une base convenable on a $p = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ et $q = \begin{pmatrix} 0 & B \\ 0 & D \end{pmatrix}$. Le projecteur $r = \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix}$ appartient évidemment à I ainsi donc que $p + r$ de sorte que $D = 0$ et donc $\operatorname{tr} q = 0$ soit $q = 0$. Ainsi on a $I = I_p$ d'où le résultat.

13 La réponse est $\{M \in \mathbb{M}_n(\mathbb{C}) \ / \ \operatorname{Im} M \subset \operatorname{Im} A\}$, la démonstration est parallèle à la précédente.

14 (1) Soit ν la valuation sur $k[[X]]$ définie par

$$\nu(\sum_{i=0}^{+\infty} a_i X^i) = \min\{i \in \mathbb{N},\ a_i \neq 0\}$$

avec la convention $\nu(0) = +\infty$. Ainsi si $a = \sum_i a_i X^i$ et $b = \sum_i b_i X^i$ sont de valuations respectives α, β, alors ab est de valuation $\alpha + \beta$ car $a_\alpha b_\beta \neq 0$. Montrons que $a = \sum_i a_i X^i$ est inversible si et seulement si $a_0 \neq 0$; supposons $a_0 \neq 0$, la recherche d'un inverse se ramène à la résolution du système triangulaire suivant : $a_0 b_0 = 1$ et pour tout $k \geq 1$, $\sum_{i=0}^{k} a_i b_{k-i} = 0$; une solution se calculant facilement par récurrence sur k. Réciproquement si a a pour inverse $b = \sum_i b_i X^i$, on a alors $a_0 b_0 = 1$ et donc $a_0 \neq 0$.

(2) Soit I un idéal non nul de A et soient $n = \min_{x \in I} \nu(x)$ et $a \in I$ tel que $\nu(a) = n$; $a = X^n b$ avec $\nu(b) = 0$ de sorte que b est inversible soit $X^n \in I$ et donc $(X^n) \subset I$; l'inclusion réciproque étant évidente, on en déduit $I = (X^n)$. L'anneau A est donc principal, donc factoriel. Soit $p \in A$ un élément irréductible, l'idéal (p) est alors premier et maximal. Or tout idéal I est contenu dans (X) qui est maximal car si $b \notin (X)$ alors b est inversible ; ainsi on a $(a) = (X)$ et a est associé à X de sorte qu'aux inversibles près, il n'y a qu'un seul irréductible, à savoir X.

(3) Montrons que A est euclidien pour le stathme ν. Soient donc $(a, b) \in A \times A^\times$; $b = X^n \beta$ avec $\beta \in A^\times$. On écrit $a\beta^{-1} = X^\beta q + c$ avec $\deg c < \beta$, et donc $a = c\beta + bq$ avec $c\beta = 0$ ou $\nu(c\beta) < \nu(b) = \beta$, d'où le résultat.

15 Considérons le groupe Γ engendré par les vecteurs de la base canonique et le vecteur $(\alpha_1, \cdots, \alpha_n)$. Comme Γ n'est pas un réseau, il n'est donc pas discret et il existe $\gamma = q(a_1, \cdots, a_n) - \sum_{i=1}^{n} p_i e_i$ dont la norme infinie est $\leq \epsilon$.

16 (1) Si $D \cap \mathbb{Z}^2 \neq \emptyset$ alors $b = r/s \in \mathbb{Q}$ avec $r \wedge s = 1$ et pour $a = m/n$ écrit sous forme irréductible on a

$$s(nq - mp) = nr$$

et donc $s|n$. Réciproquement si $s|n$, on écrit une relation de Bezout entre n et m soit $un + vm = 1$ de sorte qu'en posant $q_0 = r\frac{n}{s}u$ et $p_0 = -r\frac{n}{s}$, le point (p_0, q_0) appartient à $D \cap \mathbb{Z}^2$. On a alors

$$D \cap \mathbb{Z}^2 = \{(p_0 + kn, q_0 + km);\ k \in \mathbb{Z}\}.$$

(2) Si (p_1, q_1) et (p_2, q_2) étaient des points distincts de $D \cap \mathbb{Z}^2$, on aurait $a = \frac{q_2 - q_1}{p_2 - p_1} \in \mathbb{Q}$ ce qui n'est pas par hypothèse. Les deux situations se présentent : considérer par exemple $y - q = a(x - p)$ qui passe par (p, q) et $y = ax + b$ avec $b \in \mathbb{Q} - \mathbb{Z}$ alors pour $(p, q) \in \mathbb{Z}^2$, $b \neq q - ap$ qui soit entier soit irrationnel.

(3) Le sous-groupe de $\mathbb{R}$ engendré par 1 et $a \notin \mathbb{Q}$ étant dense, il existe une infinité de $(p,q) \in \mathbb{Z}^2$ tels que $|pa - q - b| < \epsilon$ et donc pour $P = (p,q)$, $d(P,D) \leq ap + b - q < \epsilon$ d'où le résultat.

(4) Rappelons que la distance d de (p,q) à la droite D_a d'équation $ny - mx = a$ est donnée par la formule $d = \frac{|ny-mx-a|}{\sqrt{m^2+n^2}}$. Considérons alors a et a' maximal tel que la bande définies par D_a et $D_{a'}$ ne contienne aucun point de $\mathbb{Z}^2$; par maximalité D_a et $D_{a'}$ contiennent des points de $\mathbb{Z}^2$ et donc, d'après ce qui précède, $a, a' \in \mathbb{Z}$ avec $|a - a'| = 1$ ce qui donne le résultat.

(4) D'après les questions précédentes, les points de D sont de la forme $(p_0 + kb, q_0 - ka)$ où $(p_0, q_0) \in D \cap \mathbb{Z}^2$ est une solution particulière. On rappelle par ailleurs que si $n > ab - a - b$, il existe alors une solution $(p_0, q_0) \in \mathbb{N}^2$ et que l'ensemble des $0 \leq n \leq ab - a - b$ tels qu'il existe une solution $(p_0, q_0) \in \mathbb{N}^2$ est de cardinal $(ab - a - b + 1)/2$. Par ailleurs comme la distance entre deux solutions consécutives est $\sqrt{a^2 + b^2}$ et que $AB = \frac{n}{ab}\sqrt{a^2 + b^2}$ on en déduit que le nombre de solutions est de la forme $\lfloor \frac{n}{ab} \rfloor + \epsilon$ avec $\epsilon = -1, 0, 1$ selon les cas (si $n < ab - a - b$, si a ou b ou ab divise n ou non).

17 La condition $a \geq 1$ est clairement nécessaire car quel que soit $b \geq 0$, le rectangle, dont un des côtés est formé des points $(\epsilon, 0)$ et $(a + \epsilon, 0)$ avec $0 < \epsilon < 1 - a$, ne contient aucun point de $\mathbb{Z}^2$. De même pour $a \geq 1$ et $1 \leq a \leq b < \sqrt{2}$, le rectangle dont les sommets sont $(\frac{1}{2}, -\frac{1}{2})$, $(\frac{3}{2}, \frac{1}{2})$, $(\frac{1}{2}, \frac{3}{2})$ et $(-\frac{1}{2}, \frac{1}{2})$ contient un tel rectangle et n'est pas recouvrant.

Réciproquement, remarquons tout d'abord qu'une bande délimitée par deux droites parallèles à distance $\sqrt{2}$ contient une infinité de points de $\mathbb{Z}^2$ (considérer l'intersection de cette bande avec les axes). On raisonne par l'absurde de sorte qu'il existe un rectangle de côté 1 et $\sqrt{2}$ qui ne soit pas recouvrant. On déplace alors ce rectangle le long de son petit côté ce qui délimite une bande de largeur $\sqrt{2}$ qui contient donc une infinité de points de $\mathbb{Z}^2$. En partant d'une situation non recouvrante, on considère la première situation ou le translaté contient un point de $\mathbb{Z}^2$ qui doit donc nécessairement se trouver sur l'un de ses côtés avec aucun point dans son intérieur. Pour montrer le résultat il suffit alors de prouver la propriété suivante : pour tout rectangle $ABCD$ avec$AB = 1$ et $AD = \sqrt{2}$ tel que $P \in \mathbb{Z}^2$ appartienne à l'un de ses côtés disons $[AD]$, alors un des 4 points P_1, P_2, P_3, P_4, de $\mathbb{Z}^2$ à distance 1 de P appartient à l'intérieur de $ABCD$. Si les côtés sont parallèles aux axes, le résultat est immédiat, on supposera dans la suite que ce n'est pas le cas.

Considérons le cercle Γde centre P et de rayon 1 ; (AD) définit un diamètre de Γ qui partitionne $\{P_1, P_2, P_3, P_4\}$ en deux sous-ensemble de cardinal 2 : notons P_1, P_2 ceux qui sont dans le demi-cercle contenant le rectangle $ABCD$. Comme $P_1P = 1$, on en déduit que le projeté orthogonal de P_1 sur la parallèle à AB passant par P appartient à $ABCD$. Si le projeté orthogonal de P_1 sur la parallèle à AD passant par P appartient aussi

à $ABCD$ alors P_1 aussi et sinon P_1 n'appartient pas à $ABCD$ auquel cas comme $P_1P_2 = \sqrt{2}$, P_2 est du même côté de la droite AD que P_4, P, B, C et donc P_1 appartient à $ABCD$.

18 Soit $n \in \mathbb{N}$.

(1) Si n est pair, notons $U = [\frac{1}{2}, \frac{n+1}{2}] \times [0, \frac{n}{2}]$ et V la frontière de U. Alors $U \cap \Gamma = \{(i, 0) \mid i = 1, \ldots, \frac{n}{2}\} \cup \{(i, \frac{n}{2}) \mid i = 1, \ldots, \frac{n}{2}\}$ est de cardinal n.

(2) Si n est impair, notons $U = [0, \frac{n}{2}]^2$ et V la frontière de U. Alors $U \cap \Gamma = \{(i, 0) \mid i = 0, \ldots, \frac{n-1}{2}\} \cup \{(0, i) \mid i = 1, \ldots, \frac{n-1}{2}\}$ est de cardinal n.

19 Considérons l'application

$$f : \left\{ \begin{array}{ccc} \mathbb{Z}^2 & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & |x + y\sqrt{3} - 1/3| + |x\sqrt{3} - y - 1/\sqrt{3}|. \end{array} \right.$$

La fonction f est injective. En effet, soient $a, b, c, d \in \mathbb{Z}$ tels que $f(a, b) = f(c, d)$. On note $\alpha, \beta, \gamma, \delta$ les signes respectifs de $a+b\sqrt{3}-1/3$, $a\sqrt{3}-b-1/\sqrt{3}$, $c + d\sqrt{3} - 1/3$ et $c\sqrt{3} - d - 1/\sqrt{3}$. On a

$$\alpha(a+b\sqrt{3}-1/3)+\beta(a\sqrt{3}-b-1/\sqrt{3}) = \gamma(c+d\sqrt{3}-1/3)+\delta(c\sqrt{3}-d-1/\sqrt{3}),$$

et en utilisant l'irrationalité de $\sqrt{3}$, on obtient

$$\text{ie.} \quad \alpha a-\beta b-\gamma c+\delta d-(\alpha-\gamma)/3 = 0 \quad \text{et} \quad -\alpha b-\beta a+\gamma d+\delta c+(\beta-\delta)/3 = 0.$$

Mais comme $\alpha, \gamma \in \{\pm 1\}$, on a $\alpha - \gamma \in \{0, \pm 1, \pm 2\}$. Pour que le terme de gauche soit entier, il faut donc que $\alpha = \gamma$. De même, $\beta = \delta$. En multipliant la première égalité par α et la seconde par β, on obtient donc

$$(a - c) + \alpha\beta(d - b) = 0 \quad \text{et} \quad \alpha\beta(d - b) - (a - c) = 0.$$

On en déduit que $a = c$ et $d = b$.

Soient $\iota : \mathbb{N} \to \mathbb{Z}^2$ et $\kappa : \mathbb{N} \to \mathbb{R}$ telles que $\mathbb{Z}^2 = \iota(\mathbb{N})$, et pour tout $i < j$, $f(\iota(i)) = \kappa(i) < \kappa(j) = f(\iota(j))$. Soient $g, h : \mathbb{R}^2 \to \mathbb{R}$ définies par

$$g(x, y) = x(1+\sqrt{3})+y(\sqrt{3}-1)-1/3 \quad \text{et} \quad h(x, y) = x(1-\sqrt{3})+y(\sqrt{3}+1)-1/3+1/\sqrt{3}.$$

On définit enfin pour tout $n \in \mathbb{N}$ l'ensemble

$$C_n = \{(x, y) \in \mathbb{R}^2 \mid |g(x, y)| \leq \kappa(n) \text{ et } |h(x, y)| \leq \kappa(n)\}.$$

On vérifie aisément que C_n est un carré. Par ailleurs, pour tout $(x, y) \in \mathbb{Z}^2$, on a

$$\begin{array}{rl} (x, y) \in C_n & \Leftrightarrow |g(x, y)| \leq \kappa(n) \text{ et } |h(x, y)| \leq \kappa(n) \\ & \Leftrightarrow \left|\frac{g(x,y)+h(x,y)}{2}\right| + \left|\frac{g(x,y)-h(x,y)}{2}\right| \leq \kappa(n) \\ & \Leftrightarrow f(x, y) \leq \kappa(n). \end{array}$$

Par conséquent, il y a bien n points de $\mathbb{Z}^2$ dans le carré C_n.

20 Plus précisément, si $n \in \mathbb{N}$ alors

(i) si $n = 2k$, le cercle de centre $(\frac{1}{2}, 0)$ et de rayon $\frac{5^{\frac{k-1}{2}}}{2}$ passe par n points de $\mathbb{Z}^2$,

(ii) si $n = 2k+1$, le cercle de centre $(\frac{1}{3}, 0)$ et de rayon $\frac{5^k}{3}$ passe par n points de $\mathbb{Z}^2$.

Nous ne démontrons ici que le (i), le (ii) se prouvant de la même manière. On considère les solutions de l'équation $x^2 + y^2 = 5^{k-1}$. D'après l'exercice **??**, il y en a $4k$. Ensuite, il est clair que si (x, y) est une solution de cette équation, la parité de x et celle de y sont différentes. Par ailleurs, (y, x) est alors aussi une solution. On en déduit qu'il y a exactement $2k = n$ solutions (x, y) pour lesquelles x est impair et y est pair.

Or un point $(u, v) \in \mathbb{Z}^2$ est situé sur le cercle de centre $(\frac{1}{2}, 0)$ et de rayon $\frac{5^{\frac{k-1}{2}}}{2}$ si et seulement si

$$\left(u - \frac{1}{2}\right)^2 + v^2 = \frac{5^{k-1}}{4}, \quad \text{ie. si et seulement si} \quad (2u-1)^2 + (2v)^2 = 5^{k-1}.$$

21 Considérons le point $\Omega = (\sqrt{3}, 1/3)$ du plan. Montrons que l'application $f : \mathbb{Z}^2 \to \mathbb{R}$ définie par $f(x) = \|x - \Omega\|^2$ est injective : pour cela, considérons $a, b, c, d \in \mathbb{Z}$ tels que

$$(a - \sqrt{3})^2 + (b - 1/3)^2 = (c - \sqrt{3})^2 + (d - 1/3)^2.$$

En regroupant les parties rationnelles et irrationnelles, on obtient

$$a^2 + b^2 - 2b/3 - c^2 - d^2 + 2d/3 = 2(a - c)\sqrt{3},$$

ce qui impose $a = c$ et $b = d$, puisque $\sqrt{3}$ est irrationnel. Par conséquent, on peut définir $\theta : \mathbb{N} \to \mathbb{Z}^2$ de telle sorte que $\mathbb{Z}^2 = \theta(\mathbb{N})$ et que pour tout $i < j$, $f(\theta(i)) < f(\theta(j))$. Ainsi, le disque de centre Ω et passant par $\theta(n)$ contient exactement n points du réseau.

22 On considère d'une part l'ensemble $\Gamma = \{(x, y) \in \mathbb{Z}^2 \mid x \equiv fy \ [m]\}$. Il est clair que Γ est un réseau de $\mathbb{R}^2$ de volume m. On considère d'autre part l'ellipse

$$C = \{(x, y) \in \mathbb{R}^2 \mid ax^2 + bxy + cy^2 \leq \frac{2m\sqrt{4ac - b^2}}{\pi}\}.$$

La partie C est convexe, compacte, symétrique par rapport à l'origine, μ-mesurable et

$$\mu(C) = \frac{2\pi}{\sqrt{4ac - b^2}} \frac{2m\sqrt{4ac - b^2}}{\pi} = 4m = 4\mu(\mathbb{R}^2/\Gamma).$$

Le théorème de Minkowski assure donc qu'il existe un élément (x, y) non nul de Γ dans C. On a alors

(i) $ax^2 + bxy + cy^2 \equiv (af^2 + bf + c)y^2 \equiv 0\ [m]$,

(ii) $0 < ax^2 + bxy + cy^2 < \frac{2m\sqrt{4ac-b^2}}{\pi} < 2m$,

donc $ax^2 + bxy + cy^2 = m$.

23 (1) Si $f_1, \cdots, f_n$ sont linéairement dépendantes alors les vecteurs colonnes de la matrice dont $W(f_1, \cdots, f_n)$ est le déterminant sont clairement liées quelque soit t de sorte que $W(f_1, \cdots, f_n) \equiv 0$.

(2) Dans l'autre sens, s'il existe un réel t_0 à l'intérieur de l'intervalle de définition tel que $W(f_1, \cdots, f_n)(t_0) = 0$, il existe alors des réels $c_1, \cdots, c_n$ tels que $c_1 f_1 + \cdots + c_n f_n$ et 0 sont des solutions d'une équation différentielle avec les mêmes conditions initiales au point t_0 ; le résultat découle alors du théorème de Cauchy.

(3) On a $W(f_1, f_2) \equiv 0$ alors que f_1 et f_2 ne sont pas linéairement dépendantes sur $\mathbb{R}$.

(4-a) On a

$$W(f_1 g, \cdots, f_n g) = \begin{vmatrix} f_1 g & f_2 g & \cdots & f_n g \\ f_1' g + f_1 g' & f_2' g + f_2 g' & \cdots & f_n' g + f_n g' \\ f_1" g + 2 f_1' g' + f_1 g" & \cdots & & \\ \vdots & & & \end{vmatrix}.$$

On factorise g dans la première ligne puis on soustrait à la deuxième ligne g' fois la première ce qui donne

$$W(f_1 g, \cdots, f_n g) = g \begin{vmatrix} f_1 & f_2 & \cdots & f_n \\ f_1' g & f_2' g & \cdots & f_n' g \\ f_1" g + 2 f_1' g' + f_1 g" & \cdots & & \\ \vdots & & & \end{vmatrix}.$$

On factorise g sur la deuxième ligne puis on soustrait à la troisième ligne $g"$ fois la première et $2g'$ fois la seconde. En continuant ce procédé on arrive au résultat.

(4-b) D'après la question précédente si f_1 est non nul alors on a

$$W(f_1, \cdots, f_n) = f_1^n W(1, \frac{f_2}{f_1}, \cdots, \frac{f_n}{f_1})$$

ce qui en développant $W(1, \frac{f_2}{f_1}, \cdots, \frac{f_n}{f_1})$ selon la première colonne donne le résultat.

(4-c) On raisonne par récurrence sur n. Le cas $n = 1$ étant trivial supposons le résultat acquis jusqu'au rang $n-1$ et $W(f_1, \cdots, f_n) \equiv 0$ sur un intervalle d'intérieur non vide. Si $f_1 \equiv 0$ sur cet intervalle, le résultat est clair, sinon il existe un sous-intervalle d'intérieur non vide sur lequel f_1 ne s'annule pas. D'après la question précédente on a $W\Big((\frac{f_2}{f_1})', (\frac{f_3}{f_1})', \cdots, (\frac{f_n}{f_1})'\Big) \equiv 0$

sur cet intervalle de sorte que d'après l'hypothèse de récurrence, il existe un sous-intervalle d'intérieur non vide ainsi que des nombres réels $c_2, \cdots, c_n$ non tous nuls tels que

$$c_2(\frac{f_2}{f_1})' + c_3(\frac{f_3}{f_1})' + \cdots + c_n(\frac{f_n}{f_1})' \equiv 0$$

i.e. $\left(\frac{c_2f_2+\cdots+c_nf_n}{f_1}\right)' \equiv 0$ de sorte qu'il existe une constante c_1 tel que $c_2f_2 + \cdots + c_nf_n = c_1f_1$ d'où le résultat.

(4-d) Une égalité polynomiale vérifiée sur un intervalle d'intérieur non vide, possède donc une infinité de zéros de sorte que le polynôme en question est nulle ; moralité $W(f_1, \cdots, f_n) = 0$ si et seulement si les f_i sont linéairement dépendants.

24 Soit $R = P - Q$, on a $\deg R \leq n$ et on va montrer que R a plus de $n + 1$ racines ce qui impliquera que R est nul. Par hypothèse R s'annule sur les racines de P et de $P - 1$. Le nombre de racines distinctes de P est égal à $\deg P - \deg P \wedge P'$. Or comme P et $P - 1$ sont premier entre eux, $P \wedge P'$ et $(P - 1) \wedge P'$ sont deux diviseurs distincts de P' de sorte que $\deg P \wedge P' + \deg(P - 1) \wedge P' \leq n - 1$. Il en résulte que R s'annule sur plus de $2n - (n - 1) = n + 1$ racines distinctes.

25 (a) Supposons d'abord que les racines de P sont simples. On peut supposer, quitte à changer les indices, que $\mathrm{sep}P = |\alpha_1 - \alpha_2|$. Comme $P(X) \in \mathbb{Z}[X]$, son discriminant $D(P) \in \mathbb{Z}$ de sorte qu'étant non nul on a $1 \leq |D(P)|$ et

$$1 \leq |a_d|^{2d-2} \prod_{i<j} (\alpha_i - \alpha_j)^2 \text{ soit } \frac{1}{(\alpha_1 - \alpha_2)^2} \leq |a_d|^{2d-2} \prod_{\substack{i<j \\ (i,j)\neq(1,2)}} |\alpha_i - \alpha_j|^2.$$

Des inégalités $|\alpha_i - \alpha_j| \leq |\alpha_i| + |\alpha_j| \leq 2\frac{C}{|a_d|}$, on en déduit $\frac{d(d-1)}{2} - 1 = \frac{d^2-d-2}{2}$ facteurs $|\alpha_i - \alpha_j|^2$, ce qui donne :

$$\frac{1}{|\alpha_1 - \alpha_2|^2} \leq \frac{(2C)^{d^2-d-2}}{|a_d|^{d^2-3d}} \leq (2C)^{d^2-d-2}$$

(car $|a_d| \geq 1$ et $d^2 - 3d \geq 0$), d'où le résultat.

(b) Dans le cas général lorsque les racines de $P \in \mathbb{Z}[X]$ ne sont pas nécessairement simples), on peut supposer P primitif quitte à le diviser par son contenu (qui est un entier). On considère le polynôme $R = P \wedge P'$ que l'on peut supposer dans $\mathbb{Z}[X]$ et primitif ; on a alors $P = QR$ dans $\mathbb{Z}[X]$, P et R étant primitifs. On peut alors appliquer la méthode de (a) au polynôme Q qui a les mêmes racines que P, mais avec multiplicité 1. On trouve donc, en notant d' le degré de Q, et en utilisant que $d' \leq d$:

$$\frac{1}{(\mathrm{sep}P)^2} = \frac{1}{(\mathrm{sep}Q)^2} \leq (2C)^{d'^2-d'-2} \leq (2C)^{d^2-d-2}.$$

26 (1) Un polynôme irréductible unitaire de degré d sur $\mathbb{F}_p$ fournit d éléments primitifs de $\mathbb{F}_{p^d}$ i.e. n'appartenant à aucun sous-corps stricts de $\mathbb{F}_{p^d}$. Le nombre de ces éléments est égal à

$$p^d - \sum_{\substack{m|d \\ m\neq d}} p^m \geq p^d - \sum_{m=1}^{d-1} p^m > p^d - \frac{p^d}{p-1} = p^d.\frac{p-2}{p-1}.$$

Comme le nombre de polynômes unitaires de degré d est égal à p^d, le résultat découle de l'inégalité $\frac{p-2}{p-1} \geq \frac{1}{2}$.

(2) Pour $p_1, \cdots, p_r$ des nombres premiers distincts, le lemme chinois

$$\mathbb{Z}/p_1\cdots p_r\mathbb{Z}[T] \longrightarrow \mathbb{Z}/p_1\mathbb{Z}[T] \times \cdots \times \mathbb{Z}/p_r\mathbb{Z}[T]$$

est un morphisme d'anneau de sorte que pour un polynôme P, les conditions « être réductible modulo p_i » sont indépendantes. Ainsi d'après la question précédente, la proportion des polynômes réductibles de $\mathbb{Z}/p_1\cdots p_r\mathbb{Z}[T]$ unitaires de degré d est majorée par $(1-\frac{1}{2d})^r$. En outre le nombre de polynômes unitaires de degré d à coefficients dans $[-N, N]$ dont la réduction modulo $p_1\cdots p_r$ est donnée, est au plus $(\frac{2N+1}{p_1\cdots p_r}+1)^d$. Ainsi le nombre de polynômes réductibles comme dans l'énoncé est majorée par

$$(1-\frac{1}{2d})^r(p_1\cdots p_r)^d(\frac{2N+1}{p_1\cdots p_r}+1)^d$$

En outre pour $N \geq p_1\cdots p_r$, cette quantité est majorée par $(\frac{3}{2})^d(1-\frac{1}{2d})^r$ ce qui en faisant tendre r vers $+\infty$ donne le résultat.

27 On raisonne par récurrence sur n ; le cas $n = 1$ étant évident supposons donc le résultat acquis jusqu'au rang $n-1$ et traitons le cas de n. Notons $A = B[x_{1,1}]$ de sorte que $\det = x_{1,1}P + Q$ avec $P, Q \in B$. Comme det est un polynôme de degré 1 en $x_{1,1}$, toute factorisation est de la forme $(x_{1,1}R + S)T$ avec $R, S, T \in B$ et donc $RT = P$. Comme P est un déterminant de taille $n-1$, d'après l'hypothèse de récurrence il est irréductible dans $C = \mathbb{Z}[x_{2,2}, \cdots, x_{n,n}]$ et donc dans $B = C[x_{1,2}, \cdots, x_{1,n}, x_{2,1}, \cdots, x_{n,1}]$ de sorte que soit R soit T est égal à 1. Si la factorisation est non triviale alors $T \neq 1$ et donc $T = P$ qui divise det. On considère alors la matrice suivante :

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

pour laquelle P est nul alors que son déterminant ne l'est pas car les vecteurs colonnes forment une famille étagée.

Remarque: une autre façon de conclure est de dire que tous les mineurs de taille $n-1$ divisent $\det$; comme par hypothèse de récurrence ils sont irréductibles et visiblement distincts, leur produit aussi. La contradiction découle alors de l'examen des degrés : $n^2(n-1) > n$.

28 (1) D'après le théorème de réduction des matrices antisymétriques, il existe $Q \in GL_n(K)$ tel que $X = {}^tQ\mathrm{diag}(J, \cdots, J, 0, \cdots, 0)Q$ où $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. En particulier si n est impair $\det X$ est forcément nul. Si $n = 2m$ est pair avec X inversible, on a alors $\det X = (\det Q)^2$.

(2) On applique le résultat précédent avec $K = \mathbb{Q}\Big((x_{i,j})_{1\leq i<j\leq n}\Big)$ de sorte qu'il existe $f, g \in \mathbb{Z}[(x_{i,j})_{1\leq i<j\leq n}]$ tel que $\det X = \frac{f^2}{g^2}$ avec f et g premiers entre eux dans l'anneau factoriel $\mathbb{Z}[(x_{i,j})_{1\leq i<j\leq n}]$. De l'égalité $g^2 \det X = f^2$ on en déduit que g divise f et donc $g = \pm 1$ soit $\det X = f^2$. Finalement on a $Pf = \pm f$ le signe étant déterminé par la valeur sur $\mathrm{diag}(J, \cdots, J)$. Par ailleurs on a $Pf({}^tQXQ)^2 = \Big(Pf(X)\det Q\Big)^2$ et donc comme $\mathbb{Z}[(x_{i,j})_{1\leq i<j\leq n}]$ est intègre, on en déduit l'égalité suivante de polynômes : $Pf({}^tQXQ) = \pm Pf(X)\det Q$; le signe étant finalement donné par le cas $Q = I_n$.

(3) Pour un anneau R et $A = [a_{i,j}] \in \mathbb{M}_{n-1}(R)$ antisymétrique, soit $\varphi_R : R^{n-1} \times R^{n-1} \longrightarrow R$ l'application qui à $x = [x_i]$ et $y = [y_i]$ associe le déterminant de la matrice $M = [m_{i,j}] \in \mathbb{M}_n(R)$ définie par : $m_{i,j} = a_{i-1,j-1}$ pour $2 \leq i, j \leq n$, $m_{1,j} = x_{j-1}$ pour $2 \leq j \leq n$, $m_{i,1} = -y_{i-1}$ pour $2 \leq i \leq n$, et $m_{1,1} = 0$. On vérifie immédiatement que φ est bilinéaire (noter que $m_{1,1} = 0$), symétrique (car $n = 2m$ est pair) de sorte que $\varphi(x, x) = Pf(M)^2$ est une forme quadratique. Comme Pf est polynomiale, c'est un polynôme de degré 1 en les x_i et donc une forme linéaire (car elle s'annule pour le vecteur nul). Ainsi on peut écrire $Pf(X) = a_1x_{1,1} + \cdots + a_nx_{1,n}$ avec $a_i \in \mathbb{Z}\Big[(x_{i,j})_{2\leq i<j\leq n}\Big]$. Soit alors une factorisation $Pf = fg$ de sorte que soit f soit g ne contient aucun des $x_{1,i}$ pour tout $2 \leq i \leq n$: pour fixer les choses supposons qu'il s'agisse de g. On reprend l'argument avec les deuxièmes lignes et colonnes et comme $x_{1,2}$ apparait dans f, on en déduit qu'aucun de $x_{2,i}$ n'apparait dans g. Finalement en considérant toutes les lignes, on obtient que g est une constante qui divise $Pf(\mathrm{diag}(J)) = 1$ et donc $g = \pm 1$ i.e. est inversible ce qui prouve l'irréductibilité de Pf.

29 Soit λ_1 la plus grande des valeurs propres, on a $\lambda_1 = \sum_{||x||=1}(A(x), x)$ et là où le sup est atteint, on est en présence d'un vecteur propre associé à λ_1 : en effet la différentielle s'annule en x_0 sur l'espace tangent à la sphère en x_0 (extrema liés), soit $(A(x_0), y) = 0$ pour tout y tel que $(x_0, y) = 0$ de sorte que $A(x_0)$ est colinéaire à x_0.

On remarque ensuite qu'il existe k tel que $\lambda_1 = a_{k,k} = (A(e_k), e_k)$ et on raisonne par récurrence sur l'orthogonal à e_k.

Remarque: On aurait aussi pu utiliser la forme quadratique $A \mapsto \mathrm{tr}({}^tAA) = \sum_{i,j} \lambda_{i,j}^2$ qui est clairement invariante sous l'action par conjugaison du groupe orthogonal de sorte que $\mathrm{tr}({}^tAA) = \sum_i \lambda_{i,i}^2$, d'où le résultat.

30 Si G est fini il est clairement d'exposant fini. Si G est d'exposant fini alors tout élément de G est semblable à une matrice diagonale par blocs avec sur la diagonale I_r, $-I_s$ et des matrices de taille 2, de rotations dont les angles sont alors de la forme $\frac{2k\pi}{n}$, ce qui donne donc un ensemble fini de traces possibles.

Si l'ensemble des traces est fini, considérons vect$G$ le sous-espace vectoriel de $\mathcal{L}_n(\mathbb{R})$ engendré par les éléments de G dont on fixe une base $g_1, \cdots, g_r$. On considère aussi le produit scalaire $(A, B) := \mathrm{tr}({}^tAB)$. On note $a_i(g)$ la composante de g sur g_i, soit $g = \sum_i a_i(g) g_i$. On compose à droite avec g_j^{-1} et on prend la trace. Comme $g_i^{-1} = {}^tg_j$, on a $\mathrm{tr}(g_i g_j^{-1}) = (g_i, g_j)$ et donc

$$\mathrm{tr}(g g_j^{-1}) = \sum_i a_i(g)(g_i, g_j)$$

et comme les g_i sont linéairement indépendants, la matrice M dont les éléments sont les (g_i, g_j) est inversible donc

$$a_i(g) = \sum_j (M^{-1})_g \mathrm{tr}(g g_j^{-1})$$

de sorte que l'on obtient un nombre fini de $a_i(g)$ et donc de g.

31 Supposons que u est diagonalisable de spectre réel : soit $(e_1, \cdots, e_n)$ une base orthonormée et soit $(x_1, \cdots, x_n)$ une base formée de vecteurs propres de u de valeurs propres réelles $\lambda_1, \cdots, \lambda_n$. On définit f et g par $f(x_i) = e_i$ et $g(e_i) = \lambda_i e_i$ de sorte que $u = f{-}1 \circ g \circ f$. Soit alors $f = qr$ la décomposition polaire de f avec q unitaire et r hermitienne définie positive. On obtient ainsi en posant $l = r \circ l \circ r^{-1} = q^* \circ g \circ q$ qui est hermitienne, alors $u = r^{-1} \circ l \circ r = (r^{-1} \circ l \circ r^{-1}) \circ r^2$ avec donc $r^{-1} \circ l \circ r^{-1}$ et r^2 hermitiennes.

Réciproquement si $u = v \circ w$ avec w (resp. v) hermitienne (resp. hermitienne définie positive). Notons $l = v^{1/2}$ qui est hermitienne définie positive, on a $u = l \circ (l \circ w \circ l) \circ l^{-1}$. Comme $l \circ w \circ l$ est hermitienne, elle est diagonalisable à spectre réel et il en est donc de même de u qui lui est semblable.

32 Le sens direct découle de l'exercice précédent. Supposons donc que pour toute matrice hermitienne B, AB est diagonalisable. Soit P unitaire telle que $A = PDP^*$ avec D diagonale réelle. Comme $AB = P(DP^*BP)P^{-1}$, on voit que AB est diagonalisable si et seulement si DP^*BP l'est. Par ailleurs comme P^*BP est hermitienne, on peut supposer $A = D$.

On se ramène aisément en dimension 2 avec $A = \mathrm{diag}(x^2, -y^2)$ avec x, y réels. Soient alors $B = \begin{pmatrix} y^2 & xy \\ xy & x^2 \end{pmatrix}$ et $B' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Si $y \neq 0$, on a

$AB \neq 0$ et $(AB)^2 = 0$. Si $y = 0$ alors $AB' \neq 0$ et $(AB')^2 = 0$. Ainsi AB (resp. AB') est nilpotente non nulle et donc non diagonalisable.

33 Rappelons que $\mathrm{tr}(AA^*) = \sum_{i,j} |a_{i,j}|^2$ est $U(n)$-invariante de sorte que si A est normale elle est alors unitairement semblable à la matrice diagonale des λ_i d'où le sens direct. Réciproquement toute matrice complexe est unitairement semblable à une matrice triangulaire T de diagonale formée des λ_i. L'égalité implique alors que les termes qui ne sont pas sur la diagonale sont nulle, i.e. que T est diagonale.

34 D'après l'exercice précédent, la matrice hermitienne $H = AA^* - A^*A$ est nulle si et seulement si $\mathrm{tr}H^2 = 0$. Ainsi A est normale si et seulement si la trace de $(AA^*)^2 - A(A^*)^2A - A^*A^2A^* + (A^*A)^2$ est nulle. Or rappelons que $\mathrm{tr}AB = \mathrm{tr}BA$ de sorte que $\mathrm{tr}(A^*A)^2 = \mathrm{tr}(AA^*)^2$ et $\mathrm{tr}A^*A^2A^* = \mathrm{tr}A(A^*)^2A = \mathrm{tr}A^2A^{*2}$, d'où le résultat.

35 La réciproque étant évidente, montrons le sens direct. Soit $(e_1, \cdots, e_n)$ une base orthonormée de vecteurs propres pour A : $Ae_i = \lambda_i e_i$. De même on a $A^*e_i = \overline{\lambda_i}e_i$. On construit alors un polynôme interpolateur de Lagrange P tel que $P(\lambda_i) = \overline{\lambda_i}$ de sorte que $A^* = P(A)$.

36 (1) L'inclusion est immédiate. Considérons le petit calcul en dimension 2 suivant : $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ est semblable à $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ en considérant la nouvelle base $e_1 + e_2$ et $e_1 - e_2$.

Soit alors A une matrice de trace nulle ; en ajoutant une combinaison linéaire de matrice nilpotente de rang 1, on se ramène à A diagonale $\mathrm{diag}(a_1, \cdots, a_b)$ avec $\sum_i a_i = 0$ que l'on écrit sous la forme

$$\mathrm{diag}(a_1, -a_1, 0, \cdots, 0) + \mathrm{diag}(0, a_2 + a_1, a_3, \cdots, a_n).$$

D'après le calcul précédent la première matrice est semblable à une combinaison linéaire de matrice nilpotentes de rang 1 ; la deuxième aussi par hypothèse de récurrence.

(2) L'orbite d'une classe de similitude quelconque contient dans son adhérence la classe de similitude des matrices nilpotentes de rang 1. Le résultat découle de la question précédente

37 (1) On va construite une suite $M_1 = M, M_2, \cdots, M_{n-1}$ de matrices unitairement semblables telles que M_{n-r} est de la forme $\begin{pmatrix} H & Z' & B \\ O_{r,n-r-1} & Z & N \end{pmatrix}$ avec $(HZ') \in \mathbb{M}_{n-r}(\mathbb{C})$ de Hessenberg et $Z \in \mathbb{C}^r$. On passe de M_{n-r} à M_{n-r+1} de la façon suivante : si Z est colinéaire au premier vecteur e^1 de la base canonique de $\mathbb{C}^r$ alors il n'y a rien à faire, sinon soit $X \in \mathbb{C}^r$ unitaire tel que SZ est colinéaire à e^1 où $S = I_m - 2XX^*$ est la matrice unitaire

de la symétrie par rapport à l'hyperplan $X^\perp$: $SX = -X$ et pour $Y \in X^\perp$ i.e. $X^*Y = 0$, $SY = Y$. On considère alors la matrice de Householder, $V = \begin{pmatrix} I_{n-r} & 0_{n-r,r} \\ 0_{r,n-r} & S \end{pmatrix}$ et on pose

$$M_{n-r+1} = V^{-1} M_{n-r} V = \begin{pmatrix} H & Z' & BS \\ 0_{n,n-r-1} & SZ & SNS \end{pmatrix}$$

qui est bien de la forme demandée.
Remarque: la détermination de U est algorithmique et nécessite $5n^3/3 + O(n^2)$ multiplications et $4n^3/3 + O(n^2)$ additions.

(2) Les polynômes χ_{A_0} et χ_{A_1} sont clairement réels car les racines carrées d'une matrice hermitienne sont réelles. La relation demandée s'obtient en développant par rapport à la première colonne.

(3) Le fait que $(\chi_{A_j}(X))_{0 \le j \le n-1}$ découle directement de la relation de la question précédente. La méthode de Sturm du théorème **??**, permet alors de calculer le nombre de racines distinctes dans un intervalle $[a, b]$ quelconque et donc de localiser, par dichotomie, les valeurs propres de A avec une précision aussi grande que souhaitée.

38 (a) $\Rightarrow$ (b) : on a $P^{-1}AP = B = Q^{-1}{}^t BQ = Q^{-1}B^*Q = Q^{-1}P^*A^*(P^{-1})^*Q$ et donc $A^* = (PQ^{-1}P^*)^{-1}A(PQ^{-1}P^*)$.

(b) $\Rightarrow$ (a) : par hypothèse A et A^* ont les mêmes blocs de Jordan et comme A est semblable à ${}^tA$ on en déduit que si $J_k(\lambda)$ est un bloc de Jordan de A alors $J_k(\overline{\lambda})$ aussi. Ainsi les blocs de Jordan de A associées aux valeurs propres non réelles et de leurs conjuguées arrivent par paire d'où le résultat.

(b) $\Rightarrow$ (c) :le procédé est classique, $P^{-1}AP = A^*$ implique que $Q^{-1}AQ = A^*$ avec $Q = \alpha P$ pour tout $\alpha = re^{i\theta} \in \mathbb{C}$. En additionnant les égalités $AQ = QA^*$ et $AQ^* = Q^*A$, on obtient $A(Q + Q^*) = (Q + Q^*)A$. Il suffit alors de choisir α pour que $Q + Q^*$ soit inversible. On remarque alors que $Q + Q^*$ est non singulière si et seulement si $Q^{-1}(Q + Q^*) = I + Q^{-1}Q^*$ l'est, i.e. si et seulement si -1 n'est pas une valeur propre de $Q^{-1}Q^* = e^{-2i\theta}P^{-1}P^*$ ce qui ne pose pas de difficultés.

(c) $\Rightarrow$ (d) $\Rightarrow$ (e) : $P^{-1}AP = A^*$ avec P hermitienne de sorte que $A = P(A^*P^{-1})$ et $(A^*P^{-1})^* = P^{-1}A = A^*P^{-1}$. La deuxième implication est évidente.

(e) $\Rightarrow$ (b) : si $A = HK$ avec H inversible, alors $H^{-1}AH = KH = (HK)^* = A^*$ ce qui donne (b). Supposons alors H et K non inversible. Soit alors U qui diagonalise H :$U^*HU = H' = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ de sorte que $U^*AU = \begin{pmatrix} DK' & * \\ 0 & 0 \end{pmatrix}$. Le terme DK' est alors le produit de deux matrices hermitiennes avec D inversible de sorte que d'après ce qui précède A est semblable à une matrice de la forme $B = \begin{pmatrix} C & * \\ 0 & 0 \end{pmatrix}$ avec C réel. Le résultat

découle alors du fait que le blocs de Jordan de B sont ceux de C plus des blocs de Jordan de taille 1 associés à la valeur propre nulle : en effet les blocs de Jordan de B pour $\lambda \neq 0$ sont déterminés par $\dim \mathrm{Ker}(B - \lambda I) = \dim \mathrm{Ker}(C - \lambda I)$. Ainsi les blocs de Jordan des valeurs propres non réelles et de leurs conjugués arrivent par pair d'où le résultat.

39 La condition nécessaire est classique ; rappelons que pour S un sous-ensemble de cardinal m de $\{1, \cdots, n\}$, la matrice $A(S) \in \mathbb{M}_m(\mathbb{C})$ extraite de A constituée des colonnes et des lignes indicées par $i \in S$, est définie positive. En effet si $x \in \mathbb{C}^n$ est un vecteur dont les composantes d'indice $i \notin S$ sont nulles et si $x(S) \in \mathbb{C}^m$ désigne le vecteur obtenu à partir de x en supprimant les composantes d'indice n'appartenant pas à S, alors

$$x(S)^* A(S) x(S) = x^* A x > 0.$$

Le résultat découle alors simplement du fait que le déterminant est le produit des valeurs propres qui dans le cas d'une matrice définie positive sont réelles strictement positives.

Pour la réciproque, on raisonne par récurrence sur i variant de 1 à n, le cas $i = 1$ étant immédiat. Si pour $k < n$, A_k est définie positive ses valeurs propres sont alors strictement positives de sorte que d'après le théorème 248, les valeurs propres de A_{k+1} sont toutes strictement positives sauf éventuellement la première. Or comme par hypothèse $\det A_{k+1} > 0$ et qu'il est égal au produit des valeurs propres, on en déduit que la plus petite valeur propres est aussi strictement positive et donc A_{k+1} est définie positive.

40 Si P et U commutent alors $AA^* = PUU^*P^* = P^2 = U^*P^*PU = A^*A$.

Réciproquement si A est normale alors $P^2 = AA^* = A^*A = U^*P^2U$. Or P^2 et U^*P^2U sont positive semi-définie de racine carrée P et U^*PU. D'après l'unicité de la racine carrée positive on obtient $P = U^*PU$ soit $UP = PU$.

41 Soit F le fermé complémentaire ; s'il était non vide il contiendrait une matrice $M = S + N$, sa décomposition de Dunford, et contiendrait aussi sa partie semi-simple S, laquelle est dans l'adhérence de la classe de similitude de M, d'où la contradiction.

42 Deux matrices normales sont semblables si et seulement si elles ont même polynôme caractéristique, elle sont alors unitairement semblables de sorte que la classe de similitude est bornée, on applique alors l'exercice précédent.

De la connexité de $U(n)$ on en déduit que l'ensemble des matrices normales dans une classe de conjugaison est connexe de sorte que s'il est fini il est réduit à un unique élément qui commute avec $U(n)$ et qui est donc scalaire.

43 En conjuguant M par $I_n + \lambda E_{i,j}$ on obtient

$$(I_n + \lambda E_{i,j})M(I_n - \lambda E_{i,j}) = M + \lambda(E_{i,j}M - ME_{i,j}) + \lambda^2 E_{i,j}ME_{i,j}$$

L'orbite étant bornée, il faut en particulier que $E_{i,j}ME_{i,j} = m_{i,j}E_{i,j}$ soit nul et donc que M soit diagonale ainsi que toutes les matrices de son orbite. Or les matrices diagonales d'une même classe de similitude sont en nombre fini (sur la diagonale, on trouve les valeurs propres) de sorte que l'orbite est finie, comme elle est connexe, elle est alors réduite à un point.
Remarque: Si on ne veut pas évoquer la connexité, on regarde pour T triangulaire inférieure ou supérieure, $TDT^{-1} = D$ de sorte que D commute avec toutes les triangulaires et donc avec leurs sommes et donc D est dans le centre ce qui donne D scalaire.

44 Sur $\mathbb{C}$, en utilisant la théorie de la réduction, on est ramené au cas d'un bloc de Jordan J_n. Il s'agit donc de trouver P telle que $PJ_nP^{-1} = {}^tP^{-1}{}^tJ_n{}^tP$ ou encore $({}^PP)J_n({}^tPP)^{-1} = {}^tJ_n$. Quand P varie, les matrices ${}^tPP$ décrivent toutes les matrices symétriques inversibles (en effet le rang classifie les classes de congruence des matrices symétriques complexes). On est ainsi ramené à trouver une matrice symétrique inversible qui conjugue J_n et sa transposée et on vérifie que la matrice co-unité convient, i.e. celle dont les seuls coefficients non nuls ceux de la co-diagonale qui valent 1.

Sur $\mathbb{R}$, les matrices symétriques réelles étant diagonalisables, la condition est nécessaire ; la réciproque est triviale. Ainsi la classe de similitude est fermée et son intersection avec l'ensemble des matrices symétriques est alors une orbite sous l'action de $SO(n)$ d'où le résultat.

45 *Connexité* : on a une application surjective $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ sur l'ensemble des matrices diagonalisables : on envoie $(P, (a_1, \cdots, a_n))$ sur $P\mathrm{diag}(a_1, \cdots, a_n)P^{-1}$. L'ensemble $GL_n(\mathbb{C}) \times (\mathbb{C}^\times)^n$ étant connexe, il en est de même de l'ensemble des matrices diagonalisables.

On rappelle que $GL_n(\mathbb{C})$ est connexe : soient P_1, P_2 deux matrices inversibles. On considère le polynôme $\det(P_1z + (1-z)P_2)$. Le complémentaire de l'ensemble (fini) des zéros de ce polynôme est connexe ; on considère alors un chemin qui relie 0 à 1 dans ce complémentaire, ce qui fournit un chemin de P_1 à P_2 dans $GL_n(\mathbb{C})$.

Densité : soit A une matrice complexe que l'on trigonalise $PAP^{-1} = T$. Soit alors $\epsilon_1, \cdots, \epsilon_n$ petits tels que les $t_{i,i} + \epsilon_i$ sont tous distincts. La matrice $T + \mathrm{diag}(\epsilon_1, \cdots, \epsilon_n)$ est alors diagonalisables car elle a n valeurs propres distinctes.

Intérieur : étant donné une matrice A diagonalisable avec une valeur propre multiple ; $P^{-1}AP = \mathrm{diag}(a_1, a_2, \cdots, a_n)$ avec $a_1 = a_2$, alors $A + PE_{1,2}P^{-1}$ n'est plus diagonalisable.

Réciproquement si A est diagonalisable à valeurs propres distinctes, vu que le polynôme caractéristique dépend continûment de A, et que, d'après

??, les racines d'un polynôme dépendent continûment de ses coefficients, on en déduit que si A' est proche de A, il aura aussi n valeurs propres distinctes et sera donc diagonalisable.

Par ailleurs l'ensemble des matrices à valeurs propres distinctes est encore connexe. En effet c'est le complémentaire des zéros du polynôme en n^2 variable définit comme le discriminant du polynôme caractéristique.

46 Sur $\mathbb{C}$ on a une surjection $GL_n(\mathbb{C})^2$ sur l'ensemble des matrices de rang r : on envoie (P, Q) sur PI_rQ où I_r est la matrice diagonale dont les r premiers termes sont égaux à 1, les autres étant nuls. La conclusion découle de la connexité de $GL_n(\mathbb{C})$. Sur $\mathbb{R}$, pour $r < n$, on remarque que $PI_rQ = P'I_rQ = PI_rQ' = P'I_rQ'$, où P' (resp. Q') est obtenue à partir de P en multipliant sa dernière ligne (resp. colonne) par -1 de sorte que parmi P, P' (resp. Q, Q') une exactement appartient à $GL_n(\mathbb{R})^+$ qui est connexe.

En ce qui concerne l'adhérence, on note que $\mathrm{diag}(\epsilon_1, \cdots, \epsilon_r, 0, \cdots, 0)$ est équivalente à I_r puisque de même rang, de sorte que l'adhérence contient l'ensemble des matrices de rang inférieur ou égal à r. Par ailleurs ce dernier ensemble est fermé puisqu'il correspond à l'annulation de tous les mineures d'ordre $r + 1$.

47 Soit $\lambda \in \overline{K}$, la matrice extraite de $M - \lambda I_n$ en supprimant sa première ligne et sa dernière colonne, est triangulaire inversible de sorte que le rang de $M - \lambda I_n$ est au moins égal à $n - 1$, d'où le résultat.

48 (1) Si V a un sous-espace stable W par u, en complétant une base de W en une base de V, la matrice de u y est diagonale par bloc et son polynôme caractéristique est divisible par celui de $u_{|W}$.

Réciproquement si χ_u est de la forme PQ avec P et Q premier entre eux le lemme des noyaux décompose l'espace en une somme directe de $\mathrm{Ker}\, P(u)$ et de $\mathrm{Ker}\, Q(u)$. Si $\chi_u = P^r$ avec P irréductible, on a alors $E = \mathrm{Ker}\, P$ i.e. $\pi_u = P$. Si on prend x quelconque non nul, l'espace vectoriel engendré par $x, u(x), u^2(x), \cdots$ est donc au plus de dimension $\deg P$ (en fait on a égalité), et par hypothèse est donc égal à l'espace tout entier, i.e. $r = 1$.

(2) Dans le sens direct, en utilisant la structure de $K[X]$-module sur V induite par u, on a $V \simeq K[X]/(\pi_u)$. Si V était décomposable il serait en tant que $K[X]$-module isomorphe à un produit direct $K[X]/P_1 \times K[X]/P_2$, ce qui impose $P_1 = P^r$ et $P_2 = P^s$ avec P irréductible et $\pi_u = P^{r+s}$. Or le polynôme minimal de ce produit direct est visiblement $P^{\max(r,s)}$ soit donc $\min(r, s) = 0$.

Réciproquement si V est indécomposable alors u est cyclique. Par ailleurs si son polynôme minimal n'était pas une puissance d'un polynôme irréductible, alors le lemme des noyaux contredirait l'indécomposabilité de V.

(3) On a vu que u est semi-simple si et seulement si π_u est sans multiplicité, i.e. est premier avec π_u'. On peut tester si u est semi-simple de manière

algorithmique : on calcule le polynôme caractéristique χ_u et on teste si $\frac{\chi_u}{\chi_u \wedge \chi_u'}$ annule u.

49 On se place évidemment sur un corps infini. Si le corps est algébriquement clos cela découle simplement du fait que les sous-espaces propres sont de dimension 1 sinon, il y aurait une infinité de droite dans un de ces sous-espaces propres, qui sont bien évidemment stables. Les sous-espaces propres étant de dimension 1, on en déduit que l'endomorphisme est cyclique (c'est vrai sur chaque sous-espace caractéristique, on applique ensuite le théorème chinois).

De manière générale, soit $E = F_1 \oplus \cdots \oplus F_r$ une décomposition de l'espace en sous- espaces stables cycliques de polynômes caractéristiques respectifs $\pi_1 | \cdots | \pi_r$. On choisit pour tout $i = 1, \cdots, r$, un vecteur $e_i \in F_i$ tel que $K[X].e_i = F_i$. Notons $e_2' = \frac{\pi_2}{\pi_1}.e_2$ et pour tout $\lambda \in K$, le sous-espace stable $F_\lambda = K[X].(e_1 + \lambda e_2')$ est cyclique de polynôme caractéristique π_1 ; par ailleurs pour tout $\lambda \neq \mu$, on a $F_\lambda \cap F_\mu = \{0\}$. En effet si $P(u)e_1 + \lambda P(u)e_2 = Q(u)e_1 + \mu Q(u)e_2$ alors $(P-Q)(u)e_1 = (\mu Q - \lambda P)(u)e_2$ et donc comme $F_1 \cap F_2 = \{0\}$, on en déduit que $P \equiv Q \mod \pi_1$ et $\mu Q \equiv \lambda P \mod \pi_1$ et donc comme $\lambda \neq \mu$, $P \equiv Q \equiv 0 \mod \pi_1$, d'où le résultat

L'espace muni de la structure de $K[X]$-module induite par u, est alors isomorphe à $K[X]/(P(X))$ et les sous-espaces stables sont en bijection avec les diviseurs de P.

50 On peut déjà remarquer que les $\operatorname{Ker} P(u)$ et $\operatorname{Im} P(u)$ décrivent un ensemble fini de sous-espaces : en effet si P est premier avec le polynôme minimal de u alors $\operatorname{Ker} P(u) = (0)$ et $\operatorname{Im} P(u) = E$. Comme dans l'exercice précédent, il faut que les sous-espaces propres soient de dimension 1, et alors l'espace étant cyclique, les sous-espaces stables seront les $\operatorname{Ker} Q(X) = \operatorname{Im} \frac{P(X)}{Q(X)}$.

51 Soit E un sous-espace stable et soit i tel que $V_i \subset E \subsetneqq V_{i+1}$. Supposons que $E \neq V_i$ et soit $x \in (V_{i+1} \backslash V_i) \cap E$ et soit $y \in V_{i+1} \backslash (V_i \cup E)$. Il existe alors $g \in P_W$ tel que $g_{|V_i} = \mathrm{Id}$ et $g(x) = y$. Or comme E est stable, on a $y \in E$ d'où la contradiction.

52 (1) On identifie les supplémentaires de F dans E avec les sections s de la surjection canonique $\pi : E \longrightarrow E/F$. L'espace affine des supplémentaires de F dans E est alors l'ensemble des solutions de l'équation linéaire avec second membre $\pi \circ s = \mathrm{Id}_{E/F}$. La direction de cet espace affine est donc l'ensemble des s tel que $\pi \circ s = 0$ soit donc l'ensemble des $s' : E/F \longrightarrow F$.

(2) Un tel supplémentaire sera alors stable si et seulement si $s \circ \overline{u} - u \circ s = 0$. La direction est alors l'intersection des deux espaces vectoriels : $\pi \circ s = 0$ et $s \circ \overline{u} - u \circ s = 0$, soit le sous-espace de $\operatorname{Hom}_K(E/F, F)$ des s tels que $s \circ \overline{u} - u_{|F} \circ s = 0$.

(3) D'après la question précédente, la direction de l'espace affine des supplémentaires stables de F est celle de l'espace vectoriel des matrices

rectangulaires $X \in \mathbb{M}_{p,n-p}(K)$ telles que $AX - XB = 0$ avec $p = \dim F$. On va montrer qu'un condition nécessaire et suffisante est que les polynômes caractéristiques de A et B sont premiers entre eux.

Soit μ_A et μ_B les polynômes minimaux respectifs de A et B et soit Q un facteur irréductible de leur pgcd. D'après la théorie de la réduction, on peut décomposer F (resp. G) en sous-espace stable par A (resp. B) sous la forme $F_A \oplus F'_A$ (resp. $G_B \oplus G'_B$), tels que F_A (resp. G_B) est cycliques relativement à A (resp. B) de polynôme caractéristique Q. Soit alors $X : G \longrightarrow F$ défini comme suit : il induit un isomorphisme de G_B sur F_A et est nul sur G'_B. On a alors $AX = XB$.

Supposons que μ_A et μ_B sont premiers entre eux ce qui est équivalent au fait que leur polynôme caractéristique le sont. Par linéarité on a $\mu_A(A)X = X\mu_A(B) = 0$ d'après Cayley-Hamilton. Or comme μ_A est premier avec μ_B, une relation de Bézout $P_A\mu_A + P_B\mu_B = 1$ donne que $\mu_A(B)$ est inversible d'inverse $P_A(B)$ de sorte que X est nulle.

53 (1) Le résultat découle simplement du fait que les invariants de similitudes sont indépendants du corps dans lequels ils sont considérés.

(2) Sur $\mathbb{C}$ les matrices $R(\theta)$ et $R(-\theta)$ sont semblables à $\mathrm{diag}(e^{i\theta}, e^{-i\theta})$ et quitte à changer la base de diagonalisation par $(f_1, f_2) \mapsto (\lambda f_1, f_2)$, on peut prendre la matrice de passage de déterminant 1, i.e. dans $SL_2(\mathbb{C})$.

En conjuguant $R(\theta)$ par la matrice associée à une réflexion orhtogonale, on obtient, classiquement, la matrice $R(-\theta)$ et donc ces deux matrices sont semblables dans $GL_2(\mathbb{R})$.

Enfin soit $g \in GL_2(\mathbb{R})$ tel que $gR(\theta)g^{-1} = R(-\theta) = sR(\theta)s^{-1}$ où $s = \mathrm{diag}(1, -1)$. On a alors gs qui appartient au commutant de $R(\theta)$. Comme $R(\theta)$ est un endomorphisme cyclique son commutant est

$$\{P(R(\theta)) : \ P \in \mathbb{R}[X]\} = \{aR(\theta) + b\mathrm{Id} : \ (a, b) \in \mathbb{R}^2\}.$$

Or le déterminant de $aR(\theta) + b\mathrm{Id}$ est $a^2(e^{i\theta} + \lambda)(e^{-i\theta} + \lambda) \geq 0$ et donc $\det(gs) < 0$. Ainsi donc $R(\theta)$ et $R(-\theta)$ ne sont pas semblables dans $SL_2(\mathbb{R})$.