Algèbre 2

2023-2024

Yves Laszlo

Yves.Laszlo@universite-paris-saclay.fr

L3 Magistère, Semestre 2

Version bêta du 1er avril 2024 avec coquilles

Table des matières

Chapitre 1

Introduction

En 1872, Felix Klein se pose la question suivante. « Étant donné une multiplicité et un groupe, en étudier les êtres au point de vue des propriétés qui ne sont pas altérées par les transformations du groupe... ce que l'on peut encore exprimer ainsi : on donne une multiplicité et un groupe de transformations ; développer la théorie des invariants relatifs à ce groupe » ([Kle74]).

Felix Klein

Dans ces notes concernant la géométrie vectorielle, quadratique et hermitienne, nous illustrons ce point de vue visionnaire en classifiant des objets géométriques via des invariants sous des actions de groupes variées (facteurs invariants, invariants de similitude, discriminant, indice, signature...).

Nous nous efforçons de le faire de manière « concrète », ie avec des méthodes donnant lieu à des algorithmes. Mieux vaut en effet savoir construire un objet que de simplement connaître son existence. L'objet du cours n'est en revanche pas de donner des programmes optimisés ente termes d'efficacité (c'est un autre sujet, d'ailleurs intéressant !), mais de s'interroger sur le « comment faire ». On tombe d'ailleurs rapidement sur les défauts numériques des algorithmes type pivot.
Il ne s'agit pas pour non plus de donner des méthodes formellement constructivistes ([Bis67]) mais de donner autant que faire se peut des théorèmes d'existence qui peuvent explicitement amener à la construction de l'objet en question, par exemple grâce à un ordinateur.

Nous ne saurions trop conseiller au lecteur d'implémenter sur machine les divers algorithmes : ceci lui permettra de vérifier qu'il a compris en profondeur les preuves. Nous avons de notre côté utilisé le programme SAGEMATH, basé sur Python.

J'adresse mes chaleureux remerciements à Peter Haïssinki qui m'a aimablement donné ses jolies notes sur la partie quadratique, notes sur lesquelles je me suis beaucoup appuyé pour une première cersion du

texte, et à Olivier Debarre pour ses exemples de réduction d'endomorphismes.
Crédits photo : ChronoMaths, Flickr user Duncan, Patrick Fradin, Marcel Gotlib, UQAM, Wikipedia.

1.1 Conventions

Sauf mention expresse du contraire, les anneaux seront supposés commutatifs et unitaires, en général notés R. On les supposera sauf contre-ordre explicite de plus non nuls, *i.e.*$1 \neq 0$. Leur groupe multiplicatif des inversibles est noté $\mathrm{R}^\times$. Ceci leur octroie la propriété suivante :
Tout anneau admet un idéal propre maximal pour l'inclusion, résultat que nous verrons comme un axiome (dans cette généralité, ceci équivaut à l'axiome du choix).

Max Zorn

Sinon, le lecteur le démontera sans peine en appliquant le lemme de Zorn l'ensemble des idéaux propres de R (2.7.0.7). En pratique, on peut s'en passer le plus souvent si on y tient vraiment. Bien entendu, il ne sera utilisé que pour des théorèmes d'existence : il n'a pas de valeur algorithmique. Le lemme de Zorn permet aussi de démontrer, essentiellement formellement, que, à l'instar de **Q** contenu dans **C**, tout corps **k** est contenu dans un corps algébriquement Ω.

On l'utilisera sans plus de précision. La clef de ce résultat est le fait élémentaire que tout polynôme à coefficients dans **k** admet une racine dans un corps K éventuellement plus grand. L'existence de Ω découle alors formellement de l'existence d'idéaux maximaux dans des anneaux non nuls quelconques. Toutefois, le lecteur n'aimant pas l'axiome du choix vérifiera que l'existence des corps K précédents suffit pour nous et que l'existence de Ω est juste une commodité de langage de fait.

1.2 Prérequis

Nous ne supposons d'autre connaissance en algèbre linéaire que les bases de la théorie de la dimension, le lien entre matrices et endomorphismes et les propriétés élémentaires du déterminant (notion de polynôme caractéristique et valeur propre comprise). On suppose le lecteur familier avec la méthode du pivot de Gauss. Le lecteur qui aurait vu la théorie dans le cadre d'espaces vectoriels réels ou complexes fera l'effort d'accepter (ou de vérifier) que rien ne change sur un corps quelconque.
De manière générale, on rappelle que les opérations de ligne et de colonnes sur les matrices rectangulaires à coefficients dans un anneau R sont obtenues par multiplication à droite ou à gauche par des transvections $\mathrm{T}_{i,j}(r) = \mathrm{Id} + r\mathrm{E}_{i,j}, i \neq j$ (où $\mathrm{E}_{i,j}$ est la matrice -ici carrée- standard dont tous les coefficients sont nuls sauf celui à la ligne i et colonne j qui vaut 1), les permutations de lignes ou de colonne par des matrices

de permutation M_σ, ces matrices étant inversibles (de déterminant ± 1). La multiplication d'un pivot principal par un scalaire r s'obtient par produit avec une dilatation $D(r) = Id + (r-1)E_{1,1}$ qui est inversible dès que r l'est.

D'un point de vue général, on suppose le lecteur familier avec les définitions générales d'anneaux et d'anneaux quotients. Plus spécifiquement, outre la notion de corps, est supposée connue celle d'anneau principal (intègre dont tous les idéaux sont engendré par un élément), au moins dans le cas de $\mathbf{Z}$ et $\mathbf{k}[T]$. Pour rendre la lecture plus aisée, on donnera une preuve des principaux résultats dans le chapitre sur les anneaux factoriels (10). Pour l'essentiel, nous utiliserons deux choses : l'identité de Bézout et le fait qu'un anneau principal est factoriel (10.2.1.5) (existence et unicité à l'ordre près de la décomposition en facteurs irréductibles) ce qui permet de relier la notion de PGCD à la décomposition en facteurs irréductibles d'une part, à l'identité de Bézout d'autre part. Ainsi, le résultat clef pour nous qui en découle est le suivant, version à peine généralisée de $\mathbf{k}[T]$ (resp. $\mathbf{Z}$) à un anneau principal du célèbre lemme des noyaux en algèbre linéaire usuelle (resp. du lemme chinois usuel sur les entiers). Voir aussi l'exercice 2.7.0.12.

Proposition 1.2.0.1 (Lemme chinois). *Soit* R *un anneau principal et* $r = \prod r_i \in R$ *avec* $\mathrm{PGCD}(r_i, r_j) = 1$ *si* $i \neq j$.

1. *Il existe* $u_i \in R$ *tels que* $\sum u_i r/r_i = 1$.
2. *Posons* $e_i = (u_i r/r_i \mod r) \in R/(r)$. *On a* $e_i e_j = \delta_{i,j} e_i$ *et la projection*

$$\begin{cases} R/(r) & \to & \prod R/(r_i) \\ x & \mapsto & (x \mod r_i)_i \end{cases}$$

   est un isomorphisme d'anneau d'inverse $\varphi : (x_i) \mapsto \sum x_i e_i$ *de sorte que l'*idempotent $e_i = \varphi(0, \ldots, 0, 1, 0, \ldots 0)$ *est bien défini. En d'autres termes, chaque projection sur* $R/(r_i)$ *parallèlement aux* $R/(r_j), j \neq i$ *est définie par la multiplication par* e_i.
3. *De plus, le noyau* $\mathrm{Ker}(R/(r) \xrightarrow{r_i} R/(r)) \simeq R/(r_i)$ *de la multiplication par* r_i *est* $e_i R/(r)$.

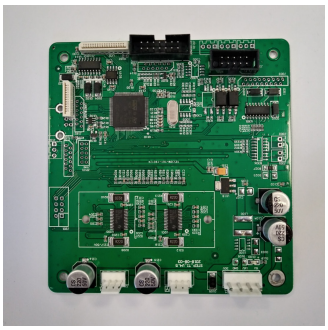
Démonstration (Sketch). *Pour le (1), observons que les facteurs irréductibles de* $\mathrm{PGCD}(r/r_i)$ *sont les facteurs* r_i *de* r. *Mais comme* $r/r_i = \prod_{j \neq i} r_j$ *et que* r_i *est premier à tous les* $r_j, j \neq i$, *il est premier à* r/r_i. *Ainsi,* $\mathrm{PGCD}(r/r_i) = 1$ *et le premier point découle de l'identité de Bézout.*

Comme $r_i | (r/r_j)$ *pour* $j \neq i$, *on a* $r = r_i(r/r_i) | (r/r_j)(r/r_i)$ *et donc* $e_j e_i = 0$ *si* $i \neq j$. *Mais comme* $\sum e_i = 1$ *par projection sur* $R/(r)$ *de l'identité de Bézout précédente, on obtient en multipliant par* e_i *la formule manquante* $e_i^2 = e_i$. *Le reste en découle immédiatement.*

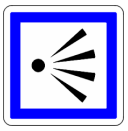
■

Chapitre 2

Généralités sur les modules



2.1 Point de vue



Ce chapitre expose de manière aussi légère que possible le langage des modules et des diagrammes. On propose au lecteur de le parcourir une première fois en se concentrant sur la résolution des exercices pour ensuite s'y familiariser avec son utilisation dans les chapitres suivants de manière concrète.

Ainsi, on le ne consultera ensuite que si besoin absolu : l'idée est que toutes les constructions formelles des espaces vectoriels ou des groupes abéliens s'appliquent *mutatis mutandis* à ce cadre général en acceptant d'avoir des scalaires à valeurs dans un anneau et plus dans un corps (ou des entiers pour les groupes abéliens).

Comme on le verra ici et dans l'ensemble du texte, le point de vue diagrammatique (cf. 2.3) une fois qu'on s'y est familiarisé est extrêmement précieux, unificateur et simplificateur. Paradoxalement, cet effort d'abstraction, outre ouvrir les portes de mathématiques modernes et profondes, les rend bien souvent extrêmement concrètes voire calculables et algorithmiques.

C'est ce que nous illustrerons notamment dans les chapitres 3,4, 5, 7, et 8 dédiés à l'étude du groupe linéaire et aux classe de similitudes de matrices carrées. Contrairement aux méthodes usuelles d'algèbre linéaire qui dépendent largement de l'étude des valeurs propres d'endomorphismes, nous allons nous focaliser les polynômes et leur action sur les endomorphismes. Si bien entendu les polynômes annulateurs

jouent un rôle particulier, leurs racines ne sont en fait pas importantes pour décider si deux endomorphismes sont semblables par exemple. Le gain est qu'en général... on ne sait pas calculer les racines des polynômes. Pire, les constructions d'algèbre linéaire sont souvent discontinues en les coefficients des matrices et donc supportent mal l'approximation numérique de ces racines. Bien entendu, la notion de valeur propre demeure essentielle comme on le verra à de nombreuses reprises.

2.2 Vocabulaire et premiers exemples

On sait qu'un espace vectoriel sur un corps $\mathbf{k}$ est un groupe abélien M muni d'une loi externe $\mathbf{k} \times \mathrm{M} \to \mathrm{M}$ vérifiant pour tout $a, a' \in \mathbf{k}$ et $m, m' \in \mathrm{M}$ (à gauche disons) les quatre compatibilités usuelles.

1. $a(m+m') = am + am'$
2. $(a+a')m = am + a'm$
3. $1m = m$
4. $a(a'm) = (aa')m$

La notion de module s'obtient exactement de même, en permettant au corps $\mathbf{k}$ d'être un anneau R (pour nous commutatif unitaire donc) :

Définition 2.2.0.1. *Un module* M *sur un anneau unitaire* R *est un groupe abélien muni d'une loi* $\mathrm{R} \times \mathrm{M} \to \mathrm{M}$ *vérifiant les propriétés de compatibilité précédentes.*

Exemple(s) 2.2.0.2. *Par définition, les modules sur les corps sont les espaces vectoriels. Donnons des exemples plus intéressants.*

1. *Les* $\mathbf{Z}$*-modules s'identifient aux groupes abéliens grâce à la multiplication externe*
$$n.m = \mathrm{signe}(n) \sum_{i=0}^{|n|} m, \ n \in \mathbf{Z}, m \in \mathrm{M}.$$
2. *Si* V *est un* $\mathbf{k}$*-espace vectoriel, l'ensemble des polynômes formels*[1] *à coefficients dans* V *est naturellement un* $k[\mathrm{T}]$*-module. Si* $(\mathrm{e}_i)_{1 \leq i \leq n}$ *est une base de* V*, les* e_i *vus comme polynômes constants de* V[T] *forment une base de* V[T]*, module qu'on identifiera donc à* $\mathbf{k}[\mathrm{T}]^n$ *par ce biais (exercice). On remarquera que la formule* $(\sum_j \lambda_{i,j} \mathrm{T}^j)_i = \sum_j (\lambda_{i,j})_i \mathrm{T}^j$ *permet d'identifier* $\mathbf{k}[\mathrm{T}]^n$ *et* $\mathbf{k}^n[\mathrm{T}]$ *ce qu'on fera désormais.*
3. *Si* R *est* intègre *et* M *un module, l'ensemble* M_{tors} *des éléments de* M *annulés par un élément non nul de* M *est un sous-module dit module de torsion .*
4. *En général, si* M *est un* R*-module arbitraire, on note* $\mathrm{Ann}_{\mathrm{M}}(r) = \mathrm{Ker}(r : \mathrm{M} \to \mathrm{M})$ *et* $\mathrm{M}[r] = \cup_{n>0} \mathrm{Ker}(r^n : \mathrm{M} \to \mathrm{M})$*, qui est bien un sous-module comme union croissante de sous-module.*

5. *L'ensemble* $C_c(T, \mathbf{R})$ *des fonctions continues à support compact d'un espace topologique* T *dans* $\mathbf{R}$ *est un module sur l'anneau des fonctions continues de* T *dans* $\mathbf{R}$. *Si* T *est un métrique non compact,* $C_c(T, \mathbf{R})$ *est un idéal mais n'est pas un anneau (exercice). Cet idéal n'est pas de type fini par exemple si* $T = \mathbf{R}^n$ *(exrecice).*
6. *Soit* $M_i, i \in I$ *une famille de modules. Comme en algèbre linéaire, le groupe abélien produit* $\prod M_i$ *a une structure naturelle de module : c'est l'unique structure tel que toutes les projections* $\pi_j : \prod M_i \to M_j$ *sont linéaires. En d'autre termes,* $a.(m_i) = (am_i)$ *(cf. 2.5.1.1).*
7. *Avec les notations précédentes, le sous-ensemble* $\oplus M_i$ *de* $\prod M_i$ *constitués des familles presque nulle est un sous-module dit « somme directe de* M_i *». La famille (de support fini)* (m_i) *est souvent notée* $\sum m_i$. *Si* I *est de plus fini, on a* $\oplus M_i = \prod M_i$.

On résume dans le tableau suivant la façons dont les constructions formelles d'algèbres linéaires s'adaptent aux modules. Pour alléger les notations, les lettres grecques $\lambda, \mu \ldots$ désignent des éléments d'un anneau R tandis que les éléments des modules sont des lettres latines $x, m, n \ldots$ pour les éléments des modules . Les énoncés sont implicitement quantifiés de manière universelle. Ainsi on écrit $\lambda(\mu x) = (\lambda\mu)x$ pour $\forall \lambda, \mu \in R$ et $\forall x \in M$, on a $\lambda(\mu x) = (\lambda\mu)x$.

Généralités

Propriété/Définition	Espace vectoriel	Module
Scalaires R	R = corps	R = anneau
Addition	$(M, +)$ groupe abélien	
Multiplication externe	$\lambda(\mu x) = (\lambda\mu)x$ et $1x = x$	
Distributivité	$\lambda(x + y) = \lambda x + \lambda y,\ (\lambda + \mu)x = \lambda x + \mu x$	
Combinaison linéaire	$\sum_{finie} \lambda_i x_i$	
Sous-espace N	N stable par combinaisons linéaires	
Exemples	noyau, image, quotient [2]	
Sous-espace engendré $\langle x_i \rangle$	$\langle x_i \rangle$ =\{combinaisons linéaires des x_i\}	
Somme de sous-espaces N_i	$+N_i$ =\{combinaisons linéaires des $x_i \in N_i$\}	
Produit [3] des N_i	$\prod N_i = \{(x_i),\ x_i \in N_i\}$	
Somme[3] directe des N_i	$\oplus N_i = \{(x_i) \in \prod N_i \mid \mathrm{Card}\{i \mid x_i \neq 0\} < \infty\}$	
$R^{(I)}, R^n$	$R^{(I)} = \oplus_I R,\ R^n = \oplus_{i=1}^n R = \prod_{i=1}^n R$	

1. C'est-à-dire des sommes $\sum_{i \geq 0} v_i T^i$ avec $v_i = 0$ si i est assez grand.

La notion d'application linéaire se transpose en celle de morphismes de modules comme dans le tableau suivant.

Exemple(s) 2.2.0.3. *Les morphismes de* $\mathbf{Z}$*-modules sont les morphismes de groupes abéliens. Voir 2.2.4 pour le cas de* V_a*.*

Généralités		
Propriété/Définition	Espace vectoriel	Module
Morphisme $f \in \mathrm{Hom}_{\mathrm{R}}(\mathrm{M}, \mathrm{M}')$	morphismes de groupes| $f(\lambda x) = \lambda f(x)$	
Isomorphisme	Morphisme bijectif	
$\mathrm{Hom}_{\mathrm{R}}(\mathrm{R}^n, \mathrm{M})$	$\mathrm{Hom}_{\mathrm{R}}(\mathrm{R}^n, \mathrm{M}) = \mathrm{M}^n$	
Matrices	$\mathrm{Hom}_{\mathrm{R}}(\mathrm{R}^n, \mathrm{R}^m) = \mathrm{M}_{m,n}(\mathrm{R})$	

Précisément on a

Lemme 2.2.0.4. *Si* M, N *sont deux* R*-modules, l'ensemble des morphismes* $\mathrm{Hom}_{\mathrm{R}}(\mathrm{M}, \mathrm{N})$ *est naturellement un module. Si* $\mathrm{M} = \mathrm{R}^n$*, l'application naturelle*

$$\begin{cases} \mathrm{Hom}_{\mathrm{R}}(\mathrm{R}^n, \mathrm{N}) & \to & \mathrm{N}^n \\ f & \mapsto & (f(\delta_{i,j})_j) \end{cases}$$

est un isomorphisme. En particulier, $\mathrm{Hom}_{\mathrm{R}}(\mathrm{R}^n, \mathrm{R}^m) = \mathrm{M}_{m,n}(\mathrm{R})$.

2.2.1 Quotient, conoyau

Le problème auquel on s'attaque est le suivant. Soit $f : \mathrm{M} \to \mathrm{N}$ un morphisme de R-modules. L'injectivité de f est caractérisée par la nullité du noyau $\mathrm{Ker}(f)$ de f. Peut-on trouver un module dont la nullité mesure la surjectivité ?
Définissons une relation sur N par la condition

$$n \sim n' \text{ si et seulement si } \exists m \text{ tel que } n - n' = f(m).$$

C'est une relation d'équivalence grâce à la linéarité de f pour la loi $+$. La classe d'équivalence de $n \in \mathrm{N}$ est

$$\overline{n} = \{n + f(m),\ m \in \mathrm{M}\} = n + f(\mathrm{M})$$

2. Voir 2.2.1.
3. Voir 2.5.1.

On note $\mathrm{Coker}(f)$ l'ensemble des classes d'équivalences de $\sim$. Ainsi, en tant qu'ensemble,

$$\mathrm{Coker}(f) = \{n + f(\mathrm{M}),\ n \in \mathrm{N}\}$$

et l'application $\pi :\ \mathrm{N} \to \mathrm{Coker}(f)$ définie par $n \mapsto \pi(n) = \overline{n}$ est surjective . L'énoncé suivant est aussi immédiat qu'important.

Proposition 2.2.1.1. *Il existe une unique structure de* R*-module sur* $\mathrm{Coker}(f)$ *telle que* π *soit un morphisme. Elle est caractérisée par* $\overline{n} + \overline{n'} = \overline{n+n'}$ *et* $\lambda\overline{n} = \overline{\lambda n}$ *; son neutre est* $\overline{0}$ *noté simplement* 0*. En outre,* f *est surjective si et seulement si on* $\mathrm{Coker}(f) = \{0\}$*.*

Ainsi, on a résolu notre problème. Un cas particulier, fondamental, est le cas où f est injectif. Dans ce cas, f induit un isomorphisme de M sur son image $f(\mathrm{M})$ qui est donc un sous-module N' de N.

Définition 2.2.1.2. *Soit* N' *un sous-module de* N *et notons* j *l'inclusion de* N' *dans* N*. On dit que* $\mathrm{Coker}(j)$ *est le quotient de* N *par* N' *et on le note* N/N'*.*

Il est important de caractériser le conoyau, à isomorphisme canonique près, par ses propriétés plus que par sa construction. C'est ce qui est expliqué en 2.5.2.1.

Remarque(s) 2.2.1.3. *D'une manière générale, nous nous intéressons aux modules à isomorphismes près. Ainsi, nous identifierons deux modules entre lesquels existe un isomorphisme canonique, c'est-à-dire ne dépendant d'auun choix. Le lecteur est par exemple habitué en algèbre linéaire à identifier un espace vectoriel de dimension finie avec son bidual (cf. 6.4.0.1), un espace euclidien avec son dual (cf. plus généralement 11.3.3), une matrice carrée de dimension* 1 *avec son unique coefficient (sa trace en fait)... De même, comme en algèbre linéaire, nous identifierons le plus souvent un morphisme injectif* $j : \mathrm{M} \to \mathrm{N}$ *avec le sous module image* $j(\mathrm{M})$ *car* j *définit un isomorphisme canonique* $\mathrm{M} \simeq j(\mathrm{M})$ *et on dira simplement (mais un peu abusivement) que* M *est un sous-module de* M*. On verra d'autres exemples.*

Le résultat suivant est formel mais important (comparer avec 2.5)

Proposition 2.2.1.4. *Soit* $f \in \mathrm{Hom}_{\mathrm{R}}(\mathrm{M}, \mathrm{N})$*. Alors* f *induit un isomorphisme canonique* $\overline{f} : \mathrm{M}/\mathrm{Ker}(f) \simeq \mathrm{Im}(f)$*.*

Démonstration. On définit

$$\overline{f}(\overline{m}) = \overline{f}(m + \mathrm{Ker}(f)) = f(m + \mathrm{Ker}(f)) = f(m) + f(\mathrm{Ker}(f)) = f(m) \in \mathrm{Im}(f).$$

Ainsi, $\overline{f}$ est bien définie et linéaire. Elle est surjective. Si $\overline{m}$ est dans le noyau, $\overline{f}(\overline{m}) = f(m) = 0$ et donc $m \in \mathrm{Ker}(f)$ soit $\overline{m} = 0$. □

2.2.2 Propriétés à manier avec précaution

Si les définitions des familles libres, génératrices ou des bases ne changent pas tout comme celle de supplémentaire, **l'essentiel des théorèmes d'existence deviennent faux dans le cas des modules** comme résumé dans le tableau plus bas. Ceci vient bien souvent des phénomènes de *torsion* : il arrive, fréquemment comme on le verra, que l'équation $am = 0$ n'entraîne pas a ou m nul. Nous y reviendrons.

 Bases, dimension, supplémentaire		
Propriété/Définition	Espace vectoriel	Module
Famille libre $(x_i)_{i\in \mathrm{I}}$	$\sum \lambda_i x_i = 0 \Rightarrow \lambda_i \equiv 0$ ou $\mathrm{R}^{(\mathrm{I})} \xrightarrow{\lambda_i \mapsto \sum \lambda_i x_i} \mathrm{M}$ injectif	
	$x \neq 0$ libre	$x \neq 0$ rarement libre
x de torsion $= x$ non libre	$x = 0$	$\exists \lambda \neq 0 \vert \lambda x = 0$
Famille génératrice $(x_i)_{i\in \mathrm{I}}$	$\langle x_i \rangle = \mathrm{M}$ ou $\mathrm{R}^{(\mathrm{I})} \xrightarrow{\lambda_i \mapsto \sum \lambda_i x_i} \mathrm{M}$ surjectif	
Base $(x_i)_{i\in \mathrm{I}}$	(x_i) libre et génératrice ou $\mathrm{R}^{(\mathrm{I})} \xrightarrow{\lambda_i \mapsto \sum \lambda_i x_i} \mathrm{M}$ bijectif	
	tout ev a une base	un module a rarement une base
Module libre M	$\mathrm{M} \simeq \mathrm{R}^{(\mathrm{I})}$ *i.e.* M admet une base	
Bases module libre	toutes les bases ont même cardinal [4]	
	tout ev est libre	un module est rarement libre
Supplémentaire S de N dans M	$\mathrm{M} = \mathrm{N} \oplus \mathrm{S}$	
	tout sev a un supplémentaire	un sous-module a rarement un supplémentaire [5]

Les (rares) modules admettant des bases sont dit *libres*. Comme en algèbre linéaire, la donnée d'une application linéaire d'un module libre dans un module quelconque est équivalent à se donner les images d'une base. De même, les applications linéaires entre modules libres munis de base s'identifient à des matrices à coefficients dans R de taille adéquate (cf. 2.2.0.4 plus haut).

5. Voir 2.7.0.4
6. Lorsque c'est le cas, on dit que N est facteur direct.

Exemple(s) 2.2.2.1. *1. La multiplication fait de* R *est un module (libre) sur lui-même (de base* 1*) et ses sous-modules sont les idéaux de* R*.*

2. $\mathrm{R}_{<n}[\mathrm{T}]$ *est un* R*-module libre de base* $\mathrm{X}^i, i < n$ *donc de rang* n *pour* $n \in \overline{\mathbf{N}} = \mathbf{N} \cup \{\infty\}$*.*

3. La multiplication par les éléments de R *fait de* $\mathrm{M}_{n,m}(\mathrm{R})$ *est un module libre de base les matrices standards* $(\mathrm{E}_{i,j})_{1\leq i\leq n, 1\leq j\leq m}$*.*

4. Le module R^m *est libre de base (canonique)* $(\mathrm{e}_j = (\delta_{i,j}, i = 1, \cdots, m))_{1\leq j\leq m}$ *(cf. 2.5.1.1).*

2.2.3 Modules cycliques

On sait que les sous-groupes d'un groupe cyclique sont cycliques et que les sous-groupes de $\mathbf{Z}/n\mathbf{Z}$ sont de la forme $n/d\mathbf{Z}/n\mathbf{Z}$ avec $d|n$. En remplaçant $\mathbf{Z}$ par un anneau principal, on obtient

Lemme 2.2.3.1 (Modules cycliques)**.** *Soit* R *un anneau principal et* $\mathrm{M} = \mathrm{R}m$ *un module* cyclique *(ou monogène) et soit* (r) *un générateur de l'idéal* $\mathrm{Ann}_{\mathrm{R}}(m)$*.*

— *On a* $\mathrm{M} \simeq \mathrm{R}/(r)$*.*

Soit N *un sous-module de* N *et* ρ' *un générateur de l'idéal* $[\mathrm{N} : \mathrm{M}] = \{x \in \mathrm{R} | x\mathrm{M} \subset \mathrm{N}\}$*. On a*

— $\rho'|r = \rho\rho'$ *et* $\mathrm{R}/(\rho) \xrightarrow{x\mapsto x\rho' m} \mathrm{N}$ *est un isomorphisme de* R*-modules.*

— r, ρ, ρ' *sont bien définis à un inversible près. En particulier, les sous-modules de* M *sont en nombre fini dès que* r *est non nul.*

Démonstration. Comme m est un générateur de M, l'homotéthie de rapport m sur M est surjective. Comme son noyau est l'idéal $\mathrm{Ann}_{\mathrm{R}}(m) = (r)$ on a $\mathrm{M} \simeq \mathrm{R}/(r)$ d'après 2.2.1.4.

Le morphisme

$$\begin{cases} [\mathrm{N} : \mathrm{M}] & \to & \mathrm{N} \\ x & \mapsto & xm \end{cases}$$

est surjectif car m engendre M et son noyau est précisément $\mathrm{Ann}_{\mathrm{R}}(m) = (r) \subset [\mathrm{N} : \mathrm{M}] = (\rho')$ de sorte que $[\mathrm{N} : \mathrm{M}]/\mathrm{Ann}_{\mathrm{R}}(m) \simeq \mathrm{N}$ d'après 2.2.1.4. Comme $r \in (\rho')$, on a bien $\rho'|r = \rho\rho'$ de sorte que la multiplication par m induit un isomorphisme $(\rho')/(\rho\rho') \simeq \mathrm{N}$. Mais alors, la multiplication par ρ' induit à sont tour un isomorphisme $\mathrm{R}/(\rho) \simeq (\rho')/(\rho\rho')$ d'où le second point. Le troisième découle du fait qu'à inversible près, le nombre de diviseurs de r est $\prod n_i$ où n_i est l'exposant d'un facteur irréductible p_i dans une décomposition en produits d'irréductibles distincts de r (cf. 10).

□

7. *i.e.*des suites presque nulles.

2.2.4 Le $\mathbf{k}[\mathrm{T}]$-module V_a

Si $\mathrm{R} = \mathbf{k}[\mathrm{T}]$ et M est un R-module, la multiplication par les éléments de $\mathbf{k}$ vus comme polynômes constants fait de M un $\mathbf{k}$-espace vectoriel. Par ailleurs, la multiplication par T définit $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{M})$: l'homotéthie de rapport T. Inversement, si V est un $\mathbf{k}$-espace vectoriel et $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$, on définit une structure de R-module V_a sur V par la formule $\mathrm{T}.v = a(v)$ et par linéarité

(i) $$\mathrm{P}(\mathrm{T}).v = \mathrm{P}(a)(v) \text{ pour tout } \mathrm{P} \in \mathrm{R} = \mathbf{k}[\mathrm{T}], v \in \mathrm{V}_a = \mathrm{V}$$

Ces deux constructions sont inverses l'une de l'autre :

Les $\mathbf{k}[\mathrm{T}]$-modules s'identifient aux paires $(\mathrm{V}, a), a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$.
Les sous-modules de V_a s'identifient alors aux sous-espaces de V stables par a (exercice).

Du point de vue des morphismes, l'identification fonctionne comme suit. Si $\mathrm{N} = \mathrm{W}_b$ est un second module associé à un endomorphisme $b \in \mathrm{End}_{\mathbf{k}}(\mathrm{W})$, un morphisme $f \in \mathrm{Hom}_{\mathrm{R}}(\mathrm{M}, \mathrm{N}) = \mathrm{Hom}_{\mathbf{k}[\mathrm{T}]}(\mathrm{V}_a, \mathrm{V}_b)$ est défini par $f \in \mathrm{Hom}_{\mathbf{k}}(\mathrm{V}, \mathrm{W})$ tel que

$$f \circ a(m) = f(\mathrm{T}m) = \mathrm{T}f(m) = b \circ f(m) \text{ pour tout } m \in \mathrm{M}$$

i.e.

(ii) $$\mathrm{Hom}_{\mathbf{k}[\mathrm{T}]}(\mathrm{V}_a, \mathrm{V}_b) = \{f \in \mathrm{Hom}_{\mathbf{k}}(\mathrm{V}, \mathrm{W}) \text{ tels que } b \circ f = f \circ a\}$$

Corollaire 2.2.4.1. *Si $f \in \mathrm{Isom}_{\mathbf{k}[t]}(\mathrm{V}_a, \mathrm{V}_b)$ si et seulement si $a = f^{-1} \circ b \circ f$ de sorte que V_a et V_b sont isomorphes si et seulement si a et b sont semblables.*

Remarque(s) 2.2.4.2. *Suivant le principe général de transposition formelle, le lecteur aura deviné que $\mathrm{Hom}_{\mathrm{R}}(\mathrm{M}, \mathrm{N})$ désigne l'espace des applications R-linéaires de M dans N, ditto pour $\mathrm{End}_{\mathrm{R}}(\mathrm{M}), \ldots$. Lorsque le contexte sera clair, on omettra la mention de l'anneau en indice.*

En particulier, lorsque $a = b$, on a

(iii) $$\mathrm{End}_{\mathbf{k}[\mathrm{T}]}(\mathrm{V}_a) = \mathrm{Com}(a)$$

où $\mathrm{Com}(a)$ est le commutant de a, ensemble des endomorphismes de V qui commutent avec a.

2.3 Suites exactes et diagrammes

2.3.1 Suites exactes

Si $f \in \mathrm{Hom}(\mathrm{M}, \mathrm{N})$ un morphisme de modules ; on a une suite canonique de morphismes

$$\mathrm{Ker}(f) \xrightarrow{\iota} \mathrm{M} \xrightarrow{f} \mathrm{N} \xrightarrow{\pi} \mathrm{Coker}(f).$$

On constate que les composés de deux morphismes succesifs $d \circ \delta$ (à savoir $f \circ \iota$ et $\pi \circ f$) sont nuls, ce qui équivaut aux inclusions $\mathrm{Im}(\delta) \subset \mathrm{Ker}(d)$. Mais on a mieux : ces inclusions sont des égalités ! Ceci amène à la définition suivante

Définition 2.3.1.1. *Soit $d_i \in \mathrm{Hom}(\mathrm{M}_i, \mathrm{M}_{i+1})$ des morphismes, notée comme une « suite » :*

$$\cdots \mathrm{M}_{i-1} \xrightarrow{d_{i-1}} \mathrm{M}_i \xrightarrow{d_i} \mathrm{M}_{i+1} \cdots$$

.

— *On dit que la suite est un complexe (en i) si $d_i \circ d_{i-1} = 0$ ie $\mathrm{Im}(d_{i-1}) \subset \mathrm{Ker}(d_i)$.*
— *On dit que la suite est exacte (en i) si de plus $\mathrm{Im}(d_{i-1}) \supset \mathrm{Ker}(d_i)$ ie $\mathrm{Ker}(d_i) = \mathrm{Im}(d_{i-1})$.*

Une suite exacte est donc un complexe particulier.

Exercice(s) 2.3.1.2. *Soit $f \in \mathrm{Hom}(\mathrm{M}, \mathrm{N})$.*

— *Montrer que $0 \to \mathrm{M} \xrightarrow{f} \mathrm{N}$ est exacte si et seulement si f injective. Quel est l'analogue pour la surjectivité ?*
— *Montrer que la suite $0 \to \mathrm{K} \to \mathrm{M} \xrightarrow{f} \mathrm{N}$ est exacte si et seulement si K s'identifie (canoniquement) au noyau de f. Comparer avec 2.4.0.2 infra.*
— *Montrer que le produit ou la somme directe de suites exactes est encore exacte.*

2.3.2 Une suite exacte fondamentale

Exemple(s) 2.3.2.1. *Soit $d \in \mathrm{R}$. Alors, la suite*

$$\mathrm{R} \xrightarrow{r \mapsto dr} \mathrm{R} \xrightarrow{r \mapsto r \mod d} \mathrm{R}/(d) \to 0$$

est exacte. Plus généralement, pour $(d_i) \in \mathrm{R}^\nu$, la suite « diagonale »

$$\mathrm{R}^\nu \xrightarrow{(r_i) \mapsto (d_i r_i)} \mathrm{R}^\nu \xrightarrow{(r_i) \mapsto (r_i \mod d_i)} \prod_{i=1}^{\nu} \mathrm{R}/(d_i) \to 0$$

est exacte (par exemple comme produit de suites exactes).

Généralisons l'exemple précédent au cas des matrices $\mathrm{D} \in \mathrm{M}_{n,m}(\mathrm{R})$ « diagonale » au sens où ses coefficients $d_{i,j}$ sont nuls si $i \neq j$. On a donc une décomposition par blocs (éventuellement vides)

$$\mathrm{D} = \begin{pmatrix} \mathrm{diag}(d_i)_{\nu,\nu} & 0_{\nu,m-\nu} \\ 0_{n-\nu,\nu} & 0_{n-\nu,m-\nu} \end{pmatrix}$$

avec $\nu = \min(m,n)$ et $d_i = d_{i,i},\ i = 1, \cdots, \nu$ (et où on remarque que $0_{n-\nu,m-\nu}$ est la matrice ... vide !). On a deux suites exactes : la première

$$\mathrm{R}^\nu \xrightarrow{(r_i)\mapsto(d_i r_i)} \mathrm{R}^\nu \xrightarrow{(r_i)\mapsto(r_i \mod d_i)} \prod_{i=1}^{\nu} \mathrm{R}/(d_i) \to 0$$

d'après l'exemple précédent, la seconde

$$\mathrm{R}^{m-\nu} \xrightarrow{0_{n-\nu,m-\nu}} \mathrm{R}^{n-\nu} \xrightarrow{\mathrm{Id}_{n-\nu}} \mathrm{R}^{n-\nu} \to 0$$

car la première flèche est... nulle !

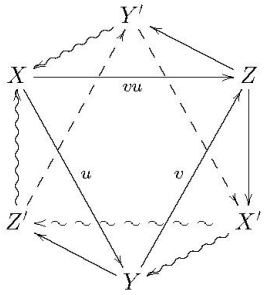
La somme de ces deux suites reste exacte : on en déduit le lemme important

Lemme 2.3.2.2. *La suite*

$$\mathrm{R}^m \xrightarrow{\mathrm{D}} \mathrm{R}^n = \mathrm{R}^\nu \times \mathrm{R}^{n-\nu} \xrightarrow{((r_i),r')\mapsto((r_i \mod d_i),r')} \prod_{i=1}^{\nu} \mathrm{R}/(d_i) \times \mathrm{R}^{n-\nu} \to 0$$

est exacte.

2.3.3 Diagrammes commutatifs



On veut voir les propriétés des morphismes en termes de diagrammes. Par exemple, dire que $f, g \in \mathrm{Hom}_k(\mathrm{V},\mathrm{W})$ sont des endomorphismes équivalents au sens de l'algèbre linéaire, c'est dire l'existence d'endomorphismes p, q de W, V tels que $p \circ f = g \circ q$ avec p, q isomorphismes. La première condition $p \circ f = g \circ q$ (resp. les deux conditions) se tradui(en)t alors en disant que le diagramme

$$\begin{array}{ccc} \mathrm{V} & \xrightarrow{p} & \mathrm{V} \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle f} \\ \mathrm{W} & \xrightarrow{q} & \mathrm{W} \end{array} \quad \text{resp.} \quad \begin{array}{ccccccc} 0 & \longrightarrow & \mathrm{V} & \xrightarrow{p} & \mathrm{V} & \longrightarrow & 0 \\ & & {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle f} & & \\ 0 & \longrightarrow & \mathrm{W} & \xrightarrow{q} & \mathrm{W} & \longrightarrow & 0 \end{array}$$

est *commutatif* avec des lignes [8] exactes (cette dernière condition étant vide pour le premier diagramme). Une définition générale formelle (qu'on encourage le lecteur à ne pas lire !) pourrait être

8. Par convention, les lignes d'un diagramme sont horizontales, les colonnes verticales.

Définition 2.3.3.1. *Soit* $\mathrm{G} = (\mathrm{S}, \mathrm{A})$ *un graphe orienté de sommets* S *et d'arêtes* A.

— *Un diagramme est la donnée pour tout sommet* $\Sigma \in \mathrm{S}$ *d'un module* M_Σ *et pour toute arête* $a : \Sigma_> \to \Sigma_<$ *de* A *d'un morphisme* $f_a : \mathrm{M}_{\Sigma_>} \to \mathrm{M}_{\Sigma_<}$.

— *Le diagramme est dit commutatif si pour toute couple de sommets* Σ, Σ'*, le composé des* f_a *associé à un chemin orienté de* Σ *vers* Σ' *ne dépend que des sommets et pas du chemin choisi.*

En pratique, on n'aura affaire qu'avec des diagramme composés de carré ou de triangles pour lesquels la définition de commutativité sera évidente.

2.4 Fonctorialité et chasses au diagramme

Bien que très simples, les énoncés de fonctorialités suivants sont cruciaux. C'est une forme très commode de formuler les propriétés universelles des noyaux et conoyaux (cf. §2.5).

Proposition 2.4.0.1 (Fonctorialité I)**.** *Supposons qu'on ait un diagramme commutatif de* R*-modules où la ligne horizontale supérieure est exacte et la ligne inférieure est un complexe.*

$$\begin{array}{ccccccc} \mathrm{M}_1 & \xrightarrow{\mu_1} & \mathrm{M}_2 & \longrightarrow & \mathrm{M}_3 & \longrightarrow & 0 \\ \downarrow & & \downarrow & & & & \\ \mathrm{N}_1 & \xrightarrow{\nu_1} & \mathrm{N}_2 & \longrightarrow & \mathrm{N}_3 & \longrightarrow & 0 \end{array}$$

Alors il existe un unique morphisme

$$f_3 : \mathrm{M}_3 \to \mathrm{N}_3$$

rendant commutatif le diagramme complété

$$\begin{array}{ccccccc} \mathrm{M}_1 & \xrightarrow{\mu_1} & \mathrm{M}_2 & \longrightarrow & \mathrm{M}_3 & \longrightarrow & 0 \\ \downarrow & & \downarrow & & \downarrow{\scriptstyle f_3} & & \\ \mathrm{N}_1 & \xrightarrow{\nu_1} & \mathrm{N}_2 & \longrightarrow & \mathrm{N}_3 & \longrightarrow & 0 \end{array}$$

Si de plus, la ligne complexe inférieure est une suite exacte et que les deux flèches $\mathrm{M}_i \to \mathrm{N}_i$*,* $i = 1, 2$ *sont des isomorphismes, alors* f_3 *est un isomorphisme..*

Démonstration. *On s'intéresse à l'existence et l'unicité du diagramme commutatif*

$$\begin{array}{ccccccc} \mathrm{M}_1 & \xrightarrow{\mu_1} & \mathrm{M}_2 & \xrightarrow{\mu_2} & \mathrm{M}_3 & \longrightarrow & 0 \\ \downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_2} & & \vdots{\scriptstyle f_3} & & \\ \mathrm{N}_1 & \xrightarrow{\nu_1} & \mathrm{N}_2 & \xrightarrow{\nu_2} & \mathrm{N}_3 & & \end{array}$$

Si on a deux flèches f_3 *et* f_3' *qui conviennent, on a* $f_3 \circ \mu_2 = \nu_2 \circ f_2 = f_3' \circ \mu_2$ *de sorte que* f_3 *et* f_3' *coïncident sur* $\mu_2(\mathrm{M}_2) = \mathrm{M}_3$ *et donc sont égales, d'où l'unicité.*

Pour l'existence, soit $m_3 \in \mathrm{M}_3$ *et considérons* m_2 un *antécédent par* μ_2. *Si* m_2 *n'est pas unique, il est défini à* $\mathrm{Ker}(\mu_2) = \mathrm{Im}(\mu_1)$ *près. Par linéarité, l'image* $\nu_2 \circ f_2(m_2)$ *est bien définie à* $\nu_2 \circ f_2 \circ \mu_1(\mathrm{M}_1)$ *près. Mais par commutativité du carré de gauche, on a* $\nu_2 \circ f_2 \circ \mu_1 = \nu_2 \circ \nu_1 \circ f_1 = 0$ *car* $\nu_2 \circ \nu_1 = 0$ *par hypothèse. Ainsi,* $\nu_2 \circ f_2(m_2)$ *est bien défini,* i.e. *ne dépend que de* m_3. *On pose alors* $f_3(m_3) = \nu_2 \circ f_2(m_2)$ *dont on vérifie qu'il convient.*

Pour la seconde partie, on peut vérifier facilement à la main que la bijectivité est de f_1, f_2 *entraîne celle de* f_3 *(exercice).*

Donnons une preuve « catégorique », preuve qui a l'avantage de se généraliser à d'autres contextes. Sous les hypothèses de bijectivité de f_1, f_2 *, on veut prouver que* f_3 *adent un inverse à gauche* g_3 *et un inverse à droite* d_3. *De* $g_3 \circ f_3 = \mathrm{Id}_{\mathrm{M}_3}$ *on obtient alors en composant à droite par* d_3 *l'égalité* $g_3 = d_3$ *et donc que* f_3 *est inversible.*

Montrons l'existence de g_3. *Appelons* g_1, g_2 *les inverses de* f_1, f_2. *Comme* $f_2 \circ \mu_1 = \nu_1 \circ f_1$, *en composant à gauche par* g_2 *et à droite par* g_1 *on* $\nu_2 \circ g_1 = g_2 \circ \nu_1$ *de sorte qu'on a un diagramme commutatif à lignes exactes*

$$\begin{array}{ccccccc}
\mathrm{M}_1 & \xrightarrow{\mu_1} & \mathrm{M}_2 & \xrightarrow{\mu_2} & \mathrm{M}_3 & \longrightarrow & 0 \\
\downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_2} & & \downarrow{\scriptstyle f_3} & & \\
\mathrm{N}_1 & \xrightarrow{\nu_1} & \mathrm{N}_2 & \xrightarrow{\nu_2} & \mathrm{N}_3 & \longrightarrow & 0 \\
\downarrow{\scriptstyle g_1} & & \downarrow{\scriptstyle g_2} & & & & \\
\mathrm{M}_1 & \xrightarrow{\nu_1} & \mathrm{M}_2 & \xrightarrow{\nu_2} & \mathrm{M}_3 & \longrightarrow & 0
\end{array}$$

qu'on peut compléter de manière unique en un diagramme commutatif à lignes exactes d'après le premier point

$$\begin{array}{ccccccc}
\mathrm{M}_1 & \xrightarrow{\mu_1} & \mathrm{M}_2 & \xrightarrow{\mu_2} & \mathrm{M}_3 & \longrightarrow & 0 \\
\downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_2} & & \downarrow{\scriptstyle f_3} & & \\
\mathrm{N}_1 & \xrightarrow{\nu_1} & \mathrm{N}_2 & \xrightarrow{\nu_2} & \mathrm{N}_3 & \longrightarrow & 0 \\
\downarrow{\scriptstyle g_1} & & \downarrow{\scriptstyle g_2} & & \downarrow{\scriptstyle g_3} & & \\
\mathrm{M}_1 & \xrightarrow{\nu_1} & \mathrm{M}_2 & \xrightarrow{\nu_2} & \mathrm{M}_3 & \longrightarrow & 0
\end{array}$$

Mais en regardant le carré externe, tenant comte de $g_1 \circ f_1 = \mathrm{Id}_{\mathrm{M}_1}$ *et* $g_2 \circ f_2 = \mathrm{Id}_{\mathrm{M}_2}$, *on a un diagramme commutatif à lignes exactes*

$$\begin{array}{ccccccc}
\mathrm{M}_1 & \xrightarrow{\nu_1} & \mathrm{M}_2 & \xrightarrow{\nu_2} & \mathrm{M}_3 & \longrightarrow & 0 \\
\downarrow{\scriptstyle \mathrm{Id}} & & \downarrow{\scriptstyle \mathrm{Id}} & & \downarrow{\scriptstyle g_3 \circ f_3} & & \\
\mathrm{M}_1 & \xrightarrow{\nu_1} & \mathrm{M}_2 & \xrightarrow{\nu_2} & \mathrm{M}_3 & \longrightarrow & 0
\end{array}$$

Mais on a aussi un diagramme commutatif

$$\begin{array}{ccccccc}
\mathrm{M}_1 & \xrightarrow{\nu_1} & \mathrm{M}_2 & \xrightarrow{\nu_2} & \mathrm{M}_3 & \longrightarrow & 0 \\
\downarrow{\scriptstyle \mathrm{Id}} & & \downarrow{\scriptstyle \mathrm{Id}} & & \downarrow{\scriptstyle \mathrm{Id}} & & \\
\mathrm{M}_1 & \xrightarrow{\nu_1} & \mathrm{M}_2 & \xrightarrow{\nu_2} & \mathrm{M}_3 & \longrightarrow & 0
\end{array}$$

ce qui, grâce à l'unicité dans le premier point, donne $g_3 \circ f_3 = \mathrm{Id}_{\mathrm{M}_3}$. *En échangeant les rôles de* M, N, *on construit l'inverse à droite de* f_3. ∎

On obtient exactement de la même manière l'énoncé suivant obtenu par « renversement du sens des flèches » [9]

Proposition 2.4.0.2 (Fonctorialité II)**.** *Supposons qu'on ait un diagramme commutatif de* R*-modules où la ligne horizontale inférieure est exacte et la ligne supérieure est un complexe.*

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathrm{M}_1 & \longrightarrow & \mathrm{M}_2 & \xrightarrow{\mu_2} & \mathrm{M}_3 \\ & & & & \downarrow & & \downarrow \\ 0 & \longrightarrow & \mathrm{N}_1 & \longrightarrow & \mathrm{N}_2 & \xrightarrow{\nu_2} & \mathrm{N}_3 \end{array}$$

Alors il existe un unique morphisme

$$\iota_1 : \mathrm{M}_1 \to \mathrm{N}_1$$

rendant commutatif le diagramme complété

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathrm{M}_1 & \longrightarrow & \mathrm{M}_2 & \xrightarrow{\mu_2} & \mathrm{M}_3 \\ & & \downarrow{\scriptstyle \iota_1} & & \downarrow & & \downarrow \\ 0 & \longrightarrow & \mathrm{N}_1 & \longrightarrow & \mathrm{N}_2 & \xrightarrow{\nu_2} & \mathrm{N}_3 \end{array}$$

Si de plus, la ligne complexe inférieure est une suite exacte et que les deux flèches $\mathrm{M}_i \to \mathrm{N}_i$, $i = 2, 3$ *sont des isomorphismes, alors* ι_3 *est un isomorphisme.*

Une généralisation parfois bien utile est le célèbre (et formel) lemme des cinq

Exercice(s) 2.4.0.3. *Considérons un diagramme commutatif de modules à lignes exactes*

$$\begin{array}{ccccccccc} \mathrm{M}_1 & \longrightarrow & \mathrm{M}_2 & \longrightarrow & \mathrm{M}_3 & \longrightarrow & \mathrm{M}_4 & \longrightarrow & \mathrm{M}_5 \\ \downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_2} & & \downarrow{\scriptstyle f_3} & & \downarrow{\scriptstyle f_4} & & \downarrow{\scriptstyle f_5} \\ \mathrm{N}_1 & \longrightarrow & \mathrm{N}_2 & \longrightarrow & \mathrm{N}_3 & \longrightarrow & \mathrm{N}_4 & \longrightarrow & \mathrm{N}_5 \end{array}$$

— *Si* f_2, f_4 *injectives et* f_1 *surjective, alors* f_3 *injective.*

— *Si* f_2, f_4 *surjectives et* f_5 *injective, alors* f_3 *bijective.*

On l'utilise le plus souvent sous la forme affaiblie suivante : Considérons un diagramme commutatif de modules à lignes exactes

9. une injection $0 \to \mathrm{M} \to \mathrm{N}$ étant donc remplacée par une surjection $\mathrm{M} \to \mathrm{N} \to 0$ et réciproquement ! C'est un phénomène général : tout énoncé formel impliquant diagrammes commutatifs, complexes et suites exactes donne lieu à un énoncé analogue par renversement du sens des flèches. On peut donner un sens précis à cet énoncé valable dans toute « catégorie abélienne ». On se contentera, et c'est bien suffisant, de voir cela comme un méta-principe.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathrm{M}_2 & \longrightarrow & \mathrm{M}_3 & \longrightarrow & \mathrm{M}_4 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f_2} & & \downarrow{\scriptstyle f_3} & & \downarrow{\scriptstyle f_4} & & \\ 0 & \longrightarrow & \mathrm{N}_2 & \longrightarrow & \mathrm{N}_3 & \longrightarrow & \mathrm{N}_4 & \longrightarrow & 0 \end{array}$$

Si f_2, f_4 bijectives f_3 bijective.

Par construction du conoyau, on a donc une suite exacte canonique

$$\mathrm{M}_1 \xrightarrow{\mu_1} \mathrm{M}_2 \to \mathrm{Coker}(\mu_1) \to 0 \tag{0}$$

On a alors la caracétrisation importante du conoyau (comparer avec l'exercice 2.3.1.2.)

Proposition 2.4.0.4. *Montrer que la suite* $\mathrm{M}_1 \xrightarrow{\mu_1} \mathrm{M}_2 \xrightarrow{\mu_2} \mathrm{M}_3 \to 0$ *est exacte si et seulement si* M_3 *s'identifie (canoniquement) au conoyau de* μ_1*.*

Démonstration. *Il suffit d'appliquer la fonctorialité 2.4.0.1 au diagramme commutatif à lignes exactes*

$$\begin{array}{ccccccccc} \mathrm{M}_1 & \xrightarrow{\mu_1} & \mathrm{M}_2 & \longrightarrow & \mathrm{Coker}(\mu_1) & \longrightarrow & 0 \\ \downarrow{\scriptstyle \mathrm{Id}} & & \downarrow{\scriptstyle \mathrm{Id}} & & & & \\ \mathrm{M}_1 & \xrightarrow{\mu_1} & \mathrm{M}_2 & \xrightarrow{\mu_2} & \mathrm{M}_3 & \longrightarrow & 0 \end{array}$$

■

Exercice(s) 2.4.0.5. *Énoncer et démontrer le résultat obtenu en renversant le sens des flèches.*

2.5 Propriétés universelles

La question posée est de caractériser les divers modules M en question par le « calcul » de

$$h(\mathrm{T}) = \mathrm{Hom}(\mathrm{T}, \mathrm{M}) \text{ ou } h^\vee(\mathrm{T}) = \mathrm{Hom}(\mathrm{M}, \mathrm{T})$$

pour T un « module test » arbitraire. Ainsi, T est vu comme une variable et $h, h^\vee$ comme une fonction de T dont les valeurs sont des ensembles. On devrait dire foncteur : la composition avec $f \in \mathrm{Hom}_{\mathrm{R}}(\mathrm{M}, \mathrm{N})$ définit une application (linéaire) $h_f(\mathrm{T}) : \ h_{\mathrm{M}}(\mathrm{T}) \to h_{\mathrm{N}}(\mathrm{T})$ (resp. $h_f^\vee : h^\vee(\mathrm{N}) \to h_{\mathrm{M}}^\vee(\mathrm{T})$) qui est compatible à la composition [10]. Le bon cadre général pour formuler ce qui suit est celui du lemme de Yoneda dans les catégories, mais on allons rester dans le cadre des modules pour les exemples qui nous intéressent pour éviter le formalisme inutile.

10. Le lecteur reconnaîtra la notion habituelle de « restriction » d'un morphisme pour $h_f(\mathrm{T})$ et dualement de « transposée » pour $h^\vee(f)$

2.5.1 Somme et produit

Soit $M_i, i \in I$ une famille de modules. On note $M_i \xrightarrow{\varphi_i} \oplus M_i$ les injections canoniques et $\prod M_i \xrightarrow{\pi_i} M_i$ les projections canoniques. Si T est un module test on a deux applications tautologiques

$$\underline{h}^\vee(T) : \begin{cases} \mathrm{Hom}_R(\oplus M_i, T) & \to & \prod \mathrm{Hom}(M_i, T) \\ f & \mapsto & (\varphi_i \circ f) \end{cases}$$

et

$$\underline{h}(T) : \begin{cases} \mathrm{Hom}_R(T, \prod M_i) & \to & \prod \mathrm{Hom}(T, M_i) \\ g & \mapsto & (g \circ \pi_i) \end{cases}$$

Lemme 2.5.1.1 (Propriétés universelles de la somme et du produit)**.** *Les applications* $\underline{h}(T)$ *et* $\underline{h}^\vee(T)$ *sont bijectives.*

La preuve est immédiate et laissée en exercice. Dans le cas de la somme directe, le sens du lemme est que se donner un morphisme $f : \oplus M_i \to T$ équivaut à se donner une collection de morphismes $f_i : M_i \to T$ (grâce à la formule $f(\sum m_i) = \sum f_i(m_i)$ qui est bien définie car la somme est en fait finie).

2.5.2 Noyau et conoyau

Soit $f : M \to N$ un morphisme de modules. Par construction, on a deux suites exactes

$$0 \to \mathrm{Ker}(f) \xrightarrow{j} M \to N$$

et

$$M \to N \xrightarrow{p} \mathrm{Coker}(f) \to 0$$

qui caractérisent noyau et conoyau (2.3.1.2 et 2.4.0.4).
Si T est un module test on a deux applications tautologiques

$$h^\vee(T) : \begin{cases} \mathrm{Hom}(\mathrm{Coker}(f), T) & \to & \mathrm{Hom}_0(N, T) = \{\psi \in \mathrm{Hom}(N, T) | \psi \circ f = 0\} \\ \varphi & \mapsto & \varphi \circ p \end{cases}$$

et

$$h(T) : \begin{cases} \mathrm{Hom}(T, \mathrm{Ker}(f)) & \to & \mathrm{Hom}_0(T, M) = \{\psi \in \mathrm{Hom}(T, M) | f \circ \psi = 0\} \\ \varphi & \mapsto & j \circ \varphi \end{cases}$$

Lemme 2.5.2.1 (Propriétés universelles du noyau et du conoyau)**.** *Les applications* $h(\mathrm{T})$ *et* $h^\vee(\mathrm{T})$ *sont bijectives.*

Démonstration. *Prouvons par exemple la propriété universelle du conoyau ie construisons l'inverse de* $h^\vee(\mathrm{T})$*. Observons qu'on a une suite exacte* $0 \to \mathrm{T} \xrightarrow{Id} \mathrm{T} \to 0$*. Soit alors* $\psi \in \mathrm{Hom}_0(\mathrm{N}, \mathrm{T})$*. La condition* $\psi \circ f = 0$ *assure précisément la commutativité du diagramme diagramme*

$$\begin{array}{ccccccc} \mathrm{M} & \xrightarrow{f} & \mathrm{N} & \xrightarrow{p} & \mathrm{Coker}(f) & \longrightarrow & 0 \\ \downarrow & & \downarrow{\scriptstyle \psi} & & & & \\ 0 & \longrightarrow & \mathrm{T} & \xrightarrow{\mathrm{Id}} & \mathrm{T} & \longrightarrow & 0 \end{array}$$

de sorte que 2.4.0.1 assure l'existence d'un unique φ *faisant commuter le diagramme*

$$\begin{array}{ccccccc} \mathrm{M} & \xrightarrow{f} & \mathrm{N} & \xrightarrow{p} & \mathrm{Coker}(f) & \longrightarrow & 0 \\ \downarrow & & \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \varphi} & & \\ 0 & \longrightarrow & \mathrm{T} & \xrightarrow{\mathrm{Id}} & \mathrm{T} & \longrightarrow & 0 \end{array}$$

On vérifie que l'application $\psi \mapsto \varphi$ *est l'inverse de* $h^\vee(\mathrm{T})$*.*

Le sens du lemme est que se donner un morphisme φ du conoyau dans T équivaut à se donner un morphisme ψ de N dans T tel que le composé de $\psi \circ f$ est nul, ou encore ψ se factorise à travers le quotient (ou passe au quotient) en φ si et seulement si $\psi\circ = 0$ (et l'analogue pour le noyau en renversant le sens des flèches). D'un point de vue diagrammatique, on résume souvent en ne gardant que le sens non formel de l'énoncé :

$$\text{Si } \psi \circ f = 0 \text{ alors} \qquad \begin{array}{ccccc} \mathrm{M} & \xrightarrow{f} & \mathrm{N} & \longrightarrow & \mathrm{Coker}(f) \\ & & {\scriptstyle \psi}\searrow & & \uparrow{\scriptstyle \exists! \varphi} \\ & & & & \mathrm{T} \end{array}$$

Une autre manière de dire, en termes des foncteurs h et $h^\vee$ est que les suites de morphismes de modules qu'ils définissent

$$0 \to \mathrm{Hom}(\mathrm{Coker}(f), \mathrm{T}) \to \mathrm{Hom}(\mathrm{N}, \mathrm{T}) \to \mathrm{Hom}(\mathrm{M}, \mathrm{T})$$

et

$$0 \to \mathrm{Hom}(\mathrm{T}, \mathrm{Ker}(f)) \to \mathrm{Hom}(\mathrm{T}, \mathrm{M}) \to \mathrm{Hom}(\mathrm{T}, \mathrm{N})$$

sont exactes.

2.6 Une variante du lemme chinois

« Quand le général Han Ting range ses soldats par trois, il reste deux soldats, quand il les range par cinq, il en reste trois et quand il les range par sept, il en reste deux. Combien l'armée de Han Ting

comporte-t-elle de soldats ? », Sun Zi, autour du IVe siècle.

Armée de terre cuite

Mausolée de l'empereur Qin

Soit M un module sur un anneau principal R. Pour p irréductible, on pose

$$\mathrm{M}[p] = \cup_{n>0}\mathrm{Ker}(p^n : \mathrm{M} \to \mathrm{M}),$$

la *composante p-primaire de* M. C'est un sous-module, comme union croissante de sous-module. On suppose ici qu'il existe $r \in \mathrm{R}$ annulant M. Rappelons la définition de l'annulateur

$$\mathrm{Ann}_{\mathrm{M}}(r) = \mathrm{Ker}(r : \mathrm{M} \to \mathrm{M}).$$

Lemme 2.6.0.1 (Décomposition primaire)**.** *Soit* $r = \prod r_i \in \mathrm{R}$. *On suppose* $\mathrm{PGCD}(r_i, r_j) = 1$ *si* $i \neq j$. *Soit* M *un module annulé par* r.

1. *Il existe* $u_i \in \mathrm{R}$ *(indépendants de* M*) tels que* $\sum u_i r/r_i = 1$.
2. *Alors,*
$$\mathrm{M} = \oplus\, \mathrm{Ann}_{\mathrm{M}}(r_i)$$
et la projection p_i *sur* $\mathrm{Ann}_{\mathrm{M}}(r_i)$ *parallèlement à* $\oplus j \neq i\, \mathrm{Ann}_{\mathrm{M}}(r_j)$ *est l'homothétie de rapport* $u_i r/r_i \in \mathrm{R}$.
3. *Les* p_i *forment une famille orthogonale de projecteurs de* M i.e.$\sum p_i = \mathrm{Id}$ *et* $p_i p_j = \delta_{i,j} p_i$.
4. *Supposons de plus* $r_i = p_i^{n_i}$ *avec* p_i *irréductible*[1] *. Alors*
$$\mathrm{Ann}_{\mathrm{M}}(p^{n_i}) := \mathrm{Ker}(p_i^{n_i} : \mathrm{M} \to \mathrm{M}) = \mathrm{M}[p_i]$$
et[2] *:* $\mathrm{M} = \oplus\, \mathrm{Ker}(p_i^{n_i} \mathrm{M} \to \mathrm{M})$.

Démonstration. *Les* r/r_i *sont globalement premiers entre eux de sorte que le premier point est l'identité de Bézout.*

Pour le second point, prouvons déjà que la somme des $\mathrm{Ann}_{\mathrm{M}}(r_i)$ *est directe. Supposons donc* $\sum m_i = 0$ *avec* $m_i \in \mathrm{Ann}_{\mathrm{M}}(r_i)$. *Pour tout* j, *on réécrit* $m_j = -\sum_{i \neq j} m_i$. *On déduit que l'idéal* I_j *annulateur de* m_j

1. De sorte que $r = \prod p_i^{n^i}$ est une décomposition en facteur irréductibles de r.
2. $\mathrm{Ann}_{\mathrm{M}}(p^{n_i}) := \mathrm{Ker}(p_i^{n_i}$ s'appelle parfois l'espace caractéristique associé à p_i par extension du cas de l'algèbre linéaire où $\mathrm{M} = \mathrm{V}_a$ et $r = \mu_a \in \mathbf{k}[\mathrm{T}]$.

contient r_j (membre de gauche de l'égalité) et $\prod_{i\neq j} p_i = r/r_i$ (membre de droite) et donc leur PGCD *par Bézout. Comme r_j et r/r_j sont premiers entre eux,* $\mathrm{PGCD}(r_i, r/r_i) = 1 \in \mathrm{I}_j$ *et* $1m_j = m_j = 0$ *pour tout j. Soit alors $m \in \mathrm{M}$. On a*

$$r = \sum u_i r/r_i m$$

et $r_i(u_i r/r_i m) = u_i r m = 0$ donc $m_i \in \mathrm{Ann}_{\mathrm{M}}(r_i)$. L'orthogonalité est évidente car $1 = \sum u_j r/r_j$ et chaque $u_j r/r_j$ est divisible par r_i si $j \neq i$ et

Le troisième point est un cas particulier du second. ■

En particulier, les projections $\mathrm{M} \to \mathrm{M}[p]$ sont « fonctorielles » au sens suivant : soit $f \in \mathrm{Hom}_{\mathrm{R}}(\mathrm{M}, \mathrm{N})$; si r comme dans la proposition annule à la fois N et M, on a un diagramme *commutatif* où les flèches verticales sont les projections (donc les homotéthies de rapport $u_i r/r_i$)

(iv)

$$\begin{array}{ccc} \mathrm{N} & \xrightarrow{f} & \mathrm{M} \\ {\scriptstyle u_i r/r_i}\downarrow & \circlearrowleft & \downarrow{\scriptstyle u_i r/r_i} \\ \mathrm{N}[p_i] & \xrightarrow{f} & \mathrm{M}[p_i] \end{array}$$

Exemple(s) 2.6.0.2. *Si $\mathrm{R} = k[\mathrm{T}]$ avec $\mathrm{P} = \prod \mathrm{P}_i \in$ avec P_i deux à deux premiers entre eux et $\mathrm{M} = \mathrm{V}_a$ (2.2.4), alors $\mathrm{Ann}_{\mathrm{M}}(\mathrm{P}) = \mathrm{Ker}(\mathrm{P}(a))$ et on retrouve le lemme des noyaux habituel*

$$\mathrm{Ker}(\mathrm{P}(a)) = \oplus \mathrm{Ker}(\mathrm{P}_i(a)).$$

Exercice(s) 2.6.0.3. *Soit* N *un sous-module de* M. *Montrer l'égalité* $\mathrm{N}[p] = \mathrm{N} \cap \mathrm{M}[p]$.
Si M *est le* $\mathrm{R} = k[\mathrm{T}]$*-module* V_a *(2.2.4) avec χ_a scindé, montrer que* $\mathrm{M}[\mathrm{P}] = \mathrm{Ker}(a - \lambda\mathrm{Id})^v$ *si* $\mathrm{P} = \mathrm{T} - \lambda$ *avec* $\chi_a(\lambda) = 0$ *et* $\mathrm{M}[\mathrm{P}] = 0$ *sinon. En d'autres termes, les composantes primaires de* V_a *sont ses espaces caractéristiques. Que retrouve-t-on comme énoncé sur les espaces stables d'un endomorphisme d'un espace vectoriel ?*

Remarque(s) 2.6.0.4. *Si* R *est euclidien, le calcul des u_i est algorithmique. La plupart des résultats que nous démontrerons pour $k[\mathrm{T}]$ de manière algorithmique se transposent mutatis mutandis aux anneaux euclidiens. Ils restent vrais dans le cadre principal, mais sans algorithme général (on utilise alors explicitement ou non des décompositions en facteurs irréductibles pour trouver des couples de Bézout, notamment dans la généralisation infra du pivot de Gauss). Cette différence est en fait profonde : c'est une fenêtre vers la* K*-théorie. Pour uné généralisation du lemme chinois au cas non principal, voir 2.7.0.12.*

Exemple(s) 2.6.0.5. *L'anneau* $\mathbf{Z}[\sqrt{-19}]$ *est principal mais non euclidien ; i on préfère la géométrie, il en est de même de l'anneau de fonctions sur le cercle de rayon* $\sqrt{-1}$*, l'anneau* $\mathbf{R}[x,y]/(x^2+y^2+1)$ *(voir 3.4) pour des références.*

2.7 Exercices supplémentaires

Exercice(s) 2.7.0.1. *1. Montrer qu'un groupe abélien est fini si et seulement si le* $\mathbf{Z}$*-module associé est de type fini et de torsion.*

2. Montrer que si V_a *correspond à* (V, a) *(2.2.4), alors* V *est de dimension finie ssi* V_a *est de type fini et de torsion.*

Exercice(s) 2.7.0.2. *Soit* R *un anneau commutatif,* M, N *deux* R*-modules et* M′ *un sous-module de* M*. On note* π *la surjection canonique* $\pi : \mathrm{M} \twoheadrightarrow \mathrm{M}/\mathrm{M}'$

1. Quels sont les sous-modules du R*-module* R *? Que peut-on dire dans ce cas du quotient ?*

2. Construire à partir de π *une bijection entre l'ensemble des sous-*R*-modules de* M *contenant* M′ *d'une part et l'ensemble des sous-*R*-modules de* M/M′*.*

Soit $f : \mathrm{M} \to \mathrm{N}$ *un morphisme de* R*-modules (i.e. une application* R*-linéaire).*

3. Montrer que $\operatorname{Ker} f$ *et* $\operatorname{Im} f$ *sont des* R*-modules, ainsi que* $\operatorname{Coker} f = \mathrm{N}/\operatorname{Im} f$*. Montrer que l'on a un isomorphisme de* R*-modules*

$$\mathrm{M}/\operatorname{Ker} f \xrightarrow{\sim} \operatorname{Im} f.$$

4. Soit l'application $f : \mathrm{R}^n \to \mathrm{R}^m$ *associée à la matrice* $\mathrm{A} = (a_{i,j})$*, avec*

$$a_{i,j} = \begin{cases} 0 & \text{si } i \neq j \\ d_i & \text{si } i = j. \end{cases}$$

Donner la structure comme R*-module de* $\operatorname{Coker} f$*.*

Exercice(s) 2.7.0.3. *On considère une suite exacte de modules* $0 \to \mathrm{M}_1 \xrightarrow{f_1} \mathrm{M}_2 \xrightarrow{f_2} \mathrm{M}_3 \to 0$*. On dit que* $\sigma \in \operatorname{Hom}_{\mathrm{R}}(\mathrm{M}_3, \mathrm{M}_2)$ *est une section de* f_2 *si* $f_2 \circ \sigma = \mathrm{Id}_{\mathrm{M}_3}$*. Quand une telle section existe, on dit que la suite est scindée.*

1. On suppose qu'il existe une telle section. Montrer que l'application $(m_1, m_3) \mapsto f_1(m_1) + \sigma(m_3)$ *définit un isomorphisme* $\mathrm{M}_1 \oplus \mathrm{M}_3 \simeq \mathrm{M}_2$*. En déduire que* $\mathrm{M}_1 \simeq f_1(\mathrm{M}_1)$ *admet alors un supplémentaire.*

2. Inversement, supposons que $\mathrm{M}_1 \simeq f_1(\mathrm{M}_1)$ *admette un supplémentaire* S*. Montrer que* f_3 *définit un isomorphisme* $\mathrm{S} \simeq \mathrm{M}_3$*.*

3. Montrer qu'un sous-module N *de* M *est facteur direct si et seulement si la suite exacte*

$$0 \to \mathrm{N} \to \mathrm{M} \to \mathrm{M}/\mathrm{N} \to 0$$

est scindée. Dans ce cas, montrer que tout supplémentaire de N *est isomorphe à* M/N*.*

4. *Montrer que si* $n > 1$, *la suite exacte canonique* $0 \to \mathbf{Z} \to \mathbf{Z} \to \mathbf{Z}/n\mathbf{Z} \to 0$ *n'est pas scindée.*
5. *Soit* $\pi : \mathrm{R}^{n+m} \to \mathrm{R}^m$ *la projection sur les* m *dernières coordonnées. Montrer qu'on a une suite exacte*
$$0 \to \mathrm{R}^n \to \mathrm{R}^{n+m} \xrightarrow{\pi} \mathrm{R}^m \to 0$$
et que cette suite est scindée.
6. *Supposons qu'il existe trois matrices carrées* $\mathrm{A}, \mathrm{B}, \mathrm{C}$ *à coefficients dans* R *de taille* $n, n+m, m$ *rendant commutatif le diagramme*
$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathrm{R}^n & \longrightarrow & \mathrm{R}^{n+m} & \longrightarrow & \mathrm{R}^n & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{A}} & & \downarrow{\scriptstyle \mathrm{B}} & & \downarrow{\scriptstyle \mathrm{C}} & & \\ 0 & \longrightarrow & \mathrm{R}^n & \longrightarrow & \mathrm{R}^{n+m} & \longrightarrow & \mathrm{R}^n & \longrightarrow & 0 \end{array}$$
Montrer que B *est triangulaire par blocs et identifier les blocs diagonaux. Énoncer et démontrer une réciproque.*

Exercice(s) 2.7.0.4. *Soit* M *un* R*-module.*

1. *Montrer qu'un idéal propre* I *de* R *est maximal si et seulement si* R/I *est un corps.*
2. *Montrer que* M *est de type fini si et seulement si il existe une surjection* R*-linéaire* $\mathrm{R}^n \to \mathrm{M}$ *pour un certain* $n \in \mathbf{N}$.
3. *Montrer que si* $f \in \mathrm{Hom}_{\mathrm{R}}(\mathrm{R}^m, \mathrm{R}^n) = \mathrm{M}_{n,m}(\mathrm{R})$ *est surjective alors* $m \geq n$.
Indication : Considérer un idéal maximal I *de* R *et voir qu'après réduction modulo* I, *l'application* f *reste surjective modulo* I.
4. *Montrer que si* f *est un isomorphisme, alors* $n = m$.
5. *Montrer qu'un module libre de type fini* L *admet une base finie et que toutes ses bases ont même cardinal : le rang de* L.
6. *Montrer que le rang de* L *est le cardinal minimal d'une famille génératrice finie.*

Exercice(s) 2.7.0.5. *Soit* V *un* $\mathbf{R}$*-espace vectoriel de dimension* 2 *et soit* D *une droite de* V.
Supposons que la droite D *soit donnée sous* ***forme paramétrée****, c'est-à-dire que que l'on se donne un vecteur directeur* v *de* D *i.e. un vecteur* $v \in \mathrm{V}$ *tel que* $\mathrm{D} = \mathbf{R} \cdot v$.

1. *On définit l'application linéaire* $\varphi : t \in \mathbf{R} \mapsto t \cdot v \in \mathrm{V}$, *montrer que la suite de* $\mathbf{R}$*-espace vectoriels suivante est exacte :*
$$\{0\} \longrightarrow \mathbf{R} \xrightarrow{\varphi} \mathrm{V}.$$
2. *Quelle est l'image du morphisme* $\mathbf{R}$*-linéaire* φ *?*

Supposons désormais que la droite D *est donnée sous* ***forme implicite****, i.e. que l'on se donne une équation de la droite* D, *c'est-à-dire une forme linéaire* $f \in \mathrm{V}^*$ *telle que* $\mathrm{D} = \mathrm{Ker}(f)$.

resume Montrer que la suite de $\mathbf{R}$*-espace vectoriels suivante est exacte :*
$$\mathrm{V} \xrightarrow{f} \mathbf{R} \longrightarrow 0.$$

resume Compléter cette suite en une suite exacte courte :

$$\{0\} \longrightarrow \mathbf{R} \xrightarrow{\varphi} \mathrm{V} \xrightarrow{f} \mathbf{R} \longrightarrow \{0\}.$$

resume Généraliser l'exercice à un corps quelconque, un espace vectoriel V *de dimension (finie) arbitraire et à des sous-espaces vectoriels quelconques.*

Exercice(s) 2.7.0.6. *Soit* $\mathbf{k}$ *un corps et* R *un anneau.*

- *Montrer que les inversibles de* $\mathbf{k}[\mathrm{T}]$ *sont les polynômes constantes non nulle de* $\mathbf{k}^*$
- *Montrer qu'une matrice de* $\mathrm{M}_n(\mathrm{R})$ *est inversible si et seulement si son déterminant est un inversible de* $\mathrm{R}^\times$. *En déduire que* $\mathrm{M} \in \mathrm{M}_n(\mathbf{k}[\mathrm{T}])$ *est inversible si et seulement si* $\det(\mathrm{M}) \in \mathbf{k}^*$.

Exercice(s) 2.7.0.7. *On rappelle le lemme de Zorn. Soit* I *un ensemble ordonné non vide qu'on suppose inductif (toute sous-ensemble totalement ordonné admet un élément maximal). Le lemme de Zorn assure que* I *admet un élément maximal. Montrer que le lemme de Zorn entraîne l'existence d'idéaux maximaux (*i.e.*idéaux propres maximaux).*

Exercice(s) 2.7.0.8 (Lemme du serpent)**.** *Considérons un diagramme commutatif de modules à lignes exactes :*

$$\begin{array}{ccccccccc} & & \mathrm{A} & \xrightarrow{i} & \mathrm{B} & \xrightarrow{p} & \mathrm{C} & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & \mathrm{A}' & \xrightarrow{i'} & \mathrm{B}' & \xrightarrow{p'} & \mathrm{C}' & & \end{array}$$

1. *Montrer que* i *envoie* $\operatorname{Ker} f$ *dans* $\operatorname{Ker} g$ *et que* p *envoie* $\operatorname{Ker} g$ *dans* $\operatorname{Ker} h$.
2. *Montrer que* i' *induit un morphisme* $\operatorname{Coker} f \to \operatorname{Coker} g$ *et que* p *induit un morphisme* $\operatorname{Coker} g \to \operatorname{Coker} h$.
3. *Montrer qu'il existe un unique morphisme* $\delta : \operatorname{Ker} h \to \operatorname{Coker} f$ *tel que la suite suivante soit exacte :*

$$\operatorname{Ker} f \longrightarrow \operatorname{Ker} g \longrightarrow \operatorname{Ker} h \xrightarrow{\delta} \operatorname{Coker} f \longrightarrow \operatorname{Coker} g \longrightarrow \operatorname{Coker} h.$$

Montrer que si i *est injective et* p *est surjective, alors la suite suivante est exacte :*

$$0 \longrightarrow \operatorname{Ker} f \longrightarrow \operatorname{Ker} g \longrightarrow \operatorname{Ker} h \xrightarrow{\delta} \operatorname{Coker} f \longrightarrow \operatorname{Coker} g \longrightarrow \operatorname{Coker} h \longrightarrow 0.$$

4. *(Bonus) Retrouver le lemme des cinq à partir du lemme du serpent.*

Exercice(s) 2.7.0.9. *On va à montrer que si l'anneau* R *n'est* pas *supposé commutatif, alors il peut arriver que les* R*-modules* R^n, $n \geq 1$ *soient tous isomorphes. À cet effet, on fixe un espace vectoriel réel* V *muni d'une base dénombrable* $(\mathrm{e}_k)_{k\in\mathbf{N}}$ *et on note* R *l'anneau des applications linéaires sur* V *(muni de la composition), identifiés à des « matrices infinies » de* $\mathbf{R}^{\mathbf{N}\times\mathbf{N}}$. *On définit deux applications linéaires* T *et* T′ *sur* V *par les relations suivantes pour* $n \in \mathbf{N}$ *:*

$$\begin{cases} \mathrm{T}(\mathrm{e}_{2n}) = \mathrm{e}_n, \\ \mathrm{T}(\mathrm{e}_{2n+1}) = 0, \end{cases} \qquad \textit{et} \qquad \begin{cases} \mathrm{T}'(\mathrm{e}_{2n}) = 0, \\ \mathrm{T}'(\mathrm{e}_{2n+1}) = \mathrm{e}_n. \end{cases}$$

Écrire les '« matrices » de T *et* T'. *Étant donné* $n \in \mathbf{N}^*$, *on considère* R^n *comme un* R*-module pour la multiplication scalaire :*

$$\mathrm{R} \times \mathrm{R}^n \to \mathrm{R}^n, \left(r, \begin{pmatrix} \mathrm{T}_1 \\ \mathrm{T}_2 \\ \vdots \\ \mathrm{T}_n \end{pmatrix} \right) \mapsto \begin{pmatrix} r \circ \mathrm{T}_1 \\ r \circ \mathrm{T}_2 \\ \vdots \\ r \circ \mathrm{T}_n \end{pmatrix}.$$

1. *Donner une base à un élément du* R*-module* R^1.
2. *Montrer que* $(\mathrm{T}, \mathrm{T}')$ *est également une base du* R*-module* R^1.
3. *Montrer que* R^1 *et* R^2 *sont isomorphes comme* R*-modules puis que* R^n *est isomorphe à* R *pour tout* $n \in \mathbf{N}^*$.

Exercice(s) 2.7.0.10. *Soit* $d \geq 1$ *un entier naturel,* R *un anneau principal et* $\mathrm{M} = \mathrm{R}^d$. *Soit* N *un sous-module de* M. *On se propose de montrer par récurrence sur* d *que* N *est isomorphe à* R^δ *avec* $\delta \leq d$. *Supposons* $d \geq 1$ *et le théorème prouvé pour les sous-modules de* $\mathrm{R}^{d'}$ *si* $d' < d$.

1. *Soit* $\underline{\nu} = (\nu_1, \cdots, \nu_d) \in \mathrm{N}^d - \{0\}$ *et* i *tel que* $n_i \neq 0$. *Montrer que la projection*

$$\pi_i : (x_1, \cdots, x_d) \mapsto x_i$$

induit une suite exacte

$$0 \to \mathrm{K} \to \mathrm{N} \xrightarrow{\pi_i} \mathrm{C} \to 0 \tag{v}$$

où C *est un sous-module non nul de* A *et* $\mathrm{K} \subset \mathrm{R}^{d-1}$.

2. *Montrer qu'il existe* $d' < d$ *et une suite exacte*

$$0 \to \mathrm{R}^{d'} \xrightarrow{j} \mathrm{N} \xrightarrow{\pi} \mathrm{R} \to 0.$$

3. *Montrer qu'il existe une* section $\sigma = \mathrm{A} \to \mathrm{N}$ *de* π, i.e.*vérifiant* $\pi \circ \sigma = \mathrm{Id}_\mathrm{A}$.
4. *Montrer que l'application* $\begin{cases} \mathrm{R}^{d'} \oplus \mathrm{R} & \to & \mathrm{N} \\ (x, y) & \mapsto & j(x) + \sigma(y) \end{cases}$ *est un isomorphisme.*
5. *Conclure.*

Exercice(s) 2.7.0.11. *TBD*

Exercice(s) 2.7.0.12. *Soit* I_i, $1 \leq i \leq n$ *un nombre fini d'idéaux d'un anneau* R. *On suppose* $\mathrm{I}_i + \mathrm{I}_j = \mathrm{R}$. *Montrer par récurrence sur* n *la généralisation suivante du lemme chinois. On a*

1. $\sum \mathrm{I}_i = \mathrm{R}$.
2. *La projection naturelle* $\mathrm{R} \mapsto \prod \mathrm{R}/\mathrm{I}_i$ *est surjective.*
3. *Son noyau* $\mathrm{I}_1 \cap \ldots \mathrm{I}_n$ *est l'idéal produit* $\mathrm{I}_1 \ldots \mathrm{I}_n$ *engendré par les produits de* n *éléments dans* $\mathrm{I}_1, \ldots, \mathrm{I}_n$ *respectivement.*

Exercice(s) 2.7.0.13 (Résultant). *Soit* R *un anneau et* $\mathrm{P},\mathrm{Q}\in\mathrm{R}[\mathrm{T}]$ *deux polynômes de degré* $p,q>0$. *On note* $\mathrm{Res}(\mathrm{P},\mathrm{Q})$ *le résultant de* P *et* Q, *égal par définition au déterminant dans les bases canoniques (cf. 2.2.4) de l'application linéaire entre modules libres de rang* $p+q$

$$\rho(\mathrm{P},\mathrm{Q}):\left\{\begin{array}{ccc}\mathrm{R}_{<q}[\mathrm{T}]\times\mathrm{R}_{<p}[\mathrm{T}] & \to & \mathrm{R}_{<p+q}[\mathrm{T}]\\ (\mathrm{A},\mathrm{B}) & \mapsto & \mathrm{AP}+\mathrm{BQ}\end{array}\right.$$

1. *Calculer* $\mathrm{Res}(\mathrm{P},\mathrm{Q})$ *si* P *est de degré* 1.
2. *En considérant la comatrice de* $\rho(\mathrm{P},\mathrm{Q})$, *montrer qu'il existe* $\mathrm{A},\mathrm{B}\in\mathbf{R}[\mathrm{T}]$ *de degré* q,p *tels que* $\mathrm{AP}+\mathrm{BQ}=\mathrm{R}(\mathrm{P},\mathrm{Q})$. *En déduire que si* P,Q *ont une racine commune dans* R, *alors* $\mathrm{R}(\mathrm{P},\mathrm{Q})=0$.
3. *Si* P,Q *sont de plus unitaires, montrer que* $\rho(\mathrm{P},\mathrm{Q})$ *est la matrice de la multiplication* $\mu:\ \mathrm{R}[\mathrm{T}]/(\mathrm{Q})\times\mathrm{R}[\mathrm{T}]\to\mathrm{R}[\mathrm{T}]/(\mathrm{PQ})$ *dans les bases canoniques (des classes de monômes* T^i*).*
4. *Toujours avec* P,Q *sont de plus unitaires, montrer qu'on a un diagramme commutatif à lignes exactes*

$$\begin{array}{ccccccccc}0 & \longrightarrow & \mathrm{R}[\mathrm{T}]/(\mathrm{PQ}) & \xrightarrow{(\mathrm{T}-r)} & \mathrm{R}[\mathrm{T}]/((\mathrm{T}-r)\mathrm{PQ}) & \xrightarrow{ev_r} & \mathrm{R} & \longrightarrow & 0\\ & & \uparrow{\scriptstyle \rho(\mathrm{P},\mathrm{Q})} & & \uparrow{\scriptstyle \rho((\mathrm{T}-r)\mathrm{P},\mathrm{Q})} & & \uparrow{\scriptstyle \mathrm{Q}(r)} & & \\ 0 & \longrightarrow & \mathrm{R}[\mathrm{T}]/(\mathrm{Q})\times\mathrm{R}[\mathrm{T}]/(\mathrm{P}) & \xrightarrow{(1,(\mathrm{T}-r))} & \mathrm{R}[\mathrm{T}]/(\mathrm{Q})\times\mathrm{R}[\mathrm{T}]/((\mathrm{T}-r)\mathrm{P}) & \xrightarrow{ev_{\mathrm{Q}}(r)} & \mathrm{R} & \longrightarrow & 0\end{array}$$

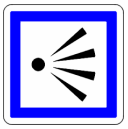
où $ev(\mathrm{A})=\mathrm{A}(r)$ *et* $ev_{\mathrm{Q}}(\mathrm{A},\mathrm{B})=\mathrm{A}(r)$. *En déduire que* $\rho((\mathrm{T}-r)\mathrm{P},\mathrm{Q})$ *est triangulaire par blocs de diagonale* $\mathrm{diag}(\rho(\mathrm{P},\mathrm{Q}),\mathrm{Q}(r))$ *puis que* $\mathrm{Res}((\mathrm{T}-r)\mathrm{P},\mathrm{Q})=\mathrm{Q}(r)\,\mathrm{Res}(\mathrm{P},\mathrm{Q})$.

5. *Si* Q *unitaire, montrer* $\mathrm{Res}(\prod(\mathrm{T}-r_i),\mathrm{Q})=\prod\mathrm{Q}(r_i)$. *Que se passe-t-il si* Q *n'est pas supposé unitaire ?*
6. *Si* $\mathrm{R}=\mathbf{k}$ *est un corps, montrer que* $\deg(\mathrm{PGCD}(\mathrm{P},\mathrm{Q}))>0$ *si et seulement si il existe* $\mathrm{A},\mathrm{B}\in\mathbf{k}[\mathrm{T}]$ *non nuls de degré* $<q$ *et* $<p$ *respectivement tels que* $\mathrm{AP}=\mathrm{BQ}$. *En déduire que* P,Q *sont premiers entre eux si et seulement si leur résultant* $\mathrm{Res}(\mathrm{P},\mathrm{Q})\neq 0$.

Chapitre 3

Classes d'équivalence dans $\mathrm{M}_{p,q}(\mathbf{k}[\mathrm{T}])$.

3.1 Point de vue



Nous exposons la théorie d'un point de vue aussi concret et algorithmique que possible en généralisant les techniques de pivot de Gauss classiques sur les matrices à coefficients dans un corps au cas des anneaux de polynômes sur $\mathbf{k}[\mathrm{T}]$ (ou d'un anneau euclidien).

Comme on l'évoque en fin de chapitre, il y a de bonnes raisons à considérer le pivot à valeurs dans des anneaux R : la présence sous-jacente d'un groupe caché nouveau, le groupe de K-théorie algébrique $\mathrm{SK}_1(\mathrm{R})$.

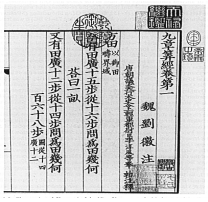
3.2 Introduction

Le lecteur qui a suivi un cours de base de théorie des groupes avec la classification des groupes abéliens de types fini reconnaîtra dans cette section une simple adaptation de ce qui a été vu pour les matrices à coefficients entiers. Il s'agira donc alors d'un simple « rappel ». Pour les autres, partons à la découverte.

Rappelons que deux matrices A, B de $\mathrm{M}_{p,q}(\mathrm{R})$ sont équivalentes (on note $\mathrm{A} \sim \mathrm{B}$) si et seulement si il existe $\mathrm{Q} \in \mathrm{GL}_q(\mathrm{R}), \mathrm{P} \in \mathrm{GL}_p(\mathrm{R})$ telles que $\mathrm{A} = \mathrm{QBP}^{-1}$. Ceci définit bien... une relation d'équivalence. Dans le cas $\mathrm{R} = \mathbf{k}$, on sait que deux matrices sont équivalentes si et seulement si elles ont même rang (application immédiate du pivot par exemple ou du théorème de la base incomplète comme on préfère).

On n'utilisera cette notion d'équivalence de matrices essentiellement uniquement dans le cas de $\mathrm{R} = \mathbf{k}[\mathrm{T}]$. Nous allons exhiber dans chaque classe d'équivalence de $\mathrm{M}_{p,q/\sim}(\mathbf{k}[\mathrm{T}])$ un représentant canonique (3.3.2.3) Le lecteur généralisera les énoncés de cette section à tout anneau muni d'une division euclidienne par simple substitution de $\mathbf{k}[\mathrm{T}]$ par un tel anneau. Soit $\mathrm{A} \in \mathrm{M}_{p,q}(\mathbf{k}[\mathrm{T}])$ une matrice rectangulaire à coefficients dans $\mathbf{k}[\mathrm{T}]$.

3.3 Diviseurs élémentaires



Les neuf chapitres

Karl Friedriech Gauss

La méthode du pivot a été redécouverte par Gauss et Jordan au XIXème. Mais elle était connue des chinois au moins au Ier siècle avant notre ère : cf. la réédition commentée « Les neuf chapitres Le classique mathématique de la Chine ancienne et ses commentaires » par Karine CHEMLA et Shuchun GUO chez Dunod en 2005.

3.3.1 Existence

Proposition 3.3.1.1. *Il existe une famille de polynômes unitaires* $\underline{\mathrm{P}} = (\mathrm{P}_r|\cdots|\mathrm{P}_2|\mathrm{P}_1)$ *telle que* A *est équivalente à la matrice*[1] *diagonale*

$$\Delta(\underline{\mathrm{P}}) = \begin{pmatrix} \mathrm{diag}(\mathrm{P}_r, \cdots, \mathrm{P}_1) & 0_{r,q-r} \\ 0_{p-r,r} & 0_{p-r,q-r} \end{pmatrix}$$

Démonstration. *On utilise librement les opérations élémentaires sur les matrices car elles laissent invariante la classe d'équivalence. On peut supposer* A *non nulle. On procède par récurrence sur* $p+q \geq 2$. *Si* $p+q=2$, *il n'y a rien à démontrer. Supposons l'énoncé prouvé si* $p+q \leq n$ *et soit* A *non nulle avec* $p+q=n+1$.

Soit $d \geq 0$ *le degré minimal d'un coefficient parmi tous les coefficients non nuls de la classe d'équivalence de* A. *On peut supposer que ce degré est atteint pour un coefficient de* A.

- *Quitte à permuter lignes et ou colonnes, on peut supposer que ce coefficient est* $a_{1,1}$.
- $a_{1,1}$ *divise nécessairement tous les* $a_{1,l}$ *et* $a_{l,1}$ *pour* $l > 1$ *(on peut par opération élémentaire remplacer ces coefficients par leur reste par la division euclidienne par* $a_{1,1}$ *qui est nul par minimalité*

de d). Par le même argument, on peut donc supposer $a_{1,l} = a_{l,1} = 0$ pour $l > 1$.

— *$a_{1,1}$ divise tout coefficient $a_{i,j}$ avec $i, j > 1$. En effet, on peut par opération élémentaire $\mathrm{L}_1 \mapsto \mathrm{L}_1 + \mathrm{L}_i$ mettre $a_{i,j}$ sur la première ligne. Puis par opération élémentaire $\mathrm{C}_j \mapsto \mathrm{C}_j - q\mathrm{C}_1$ avec q quotient de la division de $a_{i,j}$ par $a_{1,1}$ aboutir à un degré $< d$ et donc à un reste nul par minimalité de d.*

— *Ainsi $a_{1,1} = \mathrm{P}_r$ est le* PGCD *des coefficients et* A *s'écrit par blocs*

$$\mathrm{P}_r \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{B} \end{pmatrix}$$

avec $\mathrm{B} \in \mathrm{M}_{p-1,q-1}(\mathbf{k}[\mathrm{T}])$. On conclut par récurrence.

■

Remarque(s) 3.3.1.2. *Notons que r est visiblement le rang de* A *vue comme matrice à coefficients dans le corps des fractions de $\mathbf{k}[\mathrm{T}]$. Il ne dépend donc que de* A.
On peut rendre facilement cette preuve algorithmique (4.13.4 ou, au choix, [Jac85] ou [vdW50]). On invite fortement le lecteur à l'implémenter lui-même sur ordinateur (par exemple en utilisant le logiciel open source, basé sur Python, SageMath[2]).

3.3.2 Quelle unicité ?

On définit

$$\delta_n(\mathrm{A}) = \mathrm{PGCD}\left(\wedge^n(\mathrm{A})\right)$$

où $\wedge^n \mathrm{A}$ est l'idéal engendré par tous les mineurs d'ordre $n \geq 1$ de A.

Lemme 3.3.2.1. *Si*

$$\Delta(\underline{\mathrm{P}}) = \begin{pmatrix} \mathrm{diag}(\mathrm{P}_r, \cdots, \mathrm{P}_1) & 0_{r,q-r} \\ 0_{p-r,r} & 0_{p-r,q-r} \end{pmatrix} \mathrm{P}_r | \cdots | \mathrm{P}_2 | \mathrm{P}_1 \text{ unitaires}$$

on a

$$\delta_n(\mathrm{A}) = \mathrm{P}_r \cdots \mathrm{P}_{r-n+1}$$

avec ici la convention $\mathrm{P}_n = 0$ si $n \leq 0$.

Démonstration. *Tous les mineurs $\Delta_{\mathrm{I,J}}$ de $\Delta = \Delta(\underline{\mathrm{P}})$ sont triangulaires avec au moins un élément diagonal nul si $\mathrm{I} \neq \mathrm{J}$. Si $\mathrm{I} = (n \geq i_1 > \cdots, > i_n)$, on a $\det(\Delta_{\mathrm{I,I}}) = \mathrm{P}_{i_n} \cdots \mathrm{P}_{i_1}$ si $n \leq r$ et est nul sinon.*

2. Voir pour une prise en main complète A. Casamayou, N. Cohen, G. Connan, T. Dumont, L. Fousse, F. Maltey, M. Meulien, M. Mezzarobba, C. Pernet, N. M. Thiéry, P. Zimmermann, Calcul mathématique avec Sage, `https://www.sagemath.org/sagebook/french.htm`)

Si $n \leq r$, on a $i_j \leq r+1-j$ de sorte que $\mathrm{P}_r \cdots \mathrm{P}_{r-n+1} | \mathrm{P}_{i_n} \cdots \mathrm{P}_{i_1}$ à cause de la décroissance des P_i pour la divisibilité. ■

Lemme 3.3.2.2. *Soient $\mathrm{A}, \mathrm{B} \in \mathrm{M}_{p,q}(\mathbf{k}[\mathrm{T}])$. Si A et B sont équivalentes, on a*

$$\delta_n(\mathrm{A}) = \delta_n(\mathrm{B}) \text{ pour tout } n \geq 0.$$

Démonstration. *Le déterminant d'une matrice étant égale à celui de la transposée, on a $\delta_n(\mathrm{A}) = \delta_n({}^t\mathrm{A})$ pour tout n. On en déduit qu'il suffit de montrer que pour toute matrice $\mathrm{P} \in \mathrm{M}_{q,r}(\mathbf{k}[\mathrm{T}])$ (inversible ou pas) on a*

$$\wedge^n(\mathrm{AP}) \subset \wedge^n(\mathrm{A})$$

Le lecteur savant invoquera la formule générale de Binet-Cauchy

$$\det((\mathrm{AP})_{\mathrm{I,J}} = \sum_{\mathrm{K}|\ \mathrm{Card}(\mathrm{K})=n} \det(\mathrm{A}_{\mathrm{I,K}}) \det(\mathrm{P}_{\mathrm{K,J}})$$

de calcul des mineurs d'un produit de matrices quelconques. Mais on n'a pas besoin de cette précision. On peut procéder ainsi. Toute colonne de AP est une combinaison linéaire des colonnes de A. La multilinéarité du déterminant assure alors que le mineur $(\mathrm{AP})_{\mathrm{I,J}}$ est une combinaison linéaire de déterminants de matrices extraites de taille n où les colonnes sont des colonnes de A (éventuellement égales) et les lignes sont indexées par I. Si deux colonnes sont égales, le déterminant est nul (le déterminant est alterné). Sinon, l'ensemble des colonnes en question est indexé par un ensemble K de cardinal n et le déterminant en question est de la forme $\mathrm{A}_{\mathrm{I,K}}$ ce qui implique que $\det(\mathrm{AP})_{\mathrm{I,J}}$ est une combinaison linéaire des $\det(\mathrm{A}_{\mathrm{I,K}})$ avec $\mathrm{Card}(\mathrm{K}) = n$ et donc est bien dans $\wedge^n(\mathrm{A})$. ■

Du calcul précédent dans le cas diagonal (3.3.2.1) on tire

Théorème 3.3.2.3 (Diviseurs élémentaires d'une matrice polynômiale)**.** *Soit A une matrice polynômiale $\mathrm{A} \in \mathrm{M}_{p,q}(\mathbf{k}[\mathrm{T}])$*

- *Il existe une unique suite polynômes unitaires $\underline{\mathrm{P}} = (\mathrm{P}_r \cdot \cdots \cdot, \mathrm{P}_1)$ associés à A tels que pour tout n on a $\delta_n(\mathrm{A}) = \mathrm{P}_r \cdots \mathrm{P}_{r-n+1}$. On les appelles les* diviseurs élémentaires de A.
- *Deux matrices de $\mathrm{M}_{p,q}(\mathbf{k}[\mathrm{T}])$ sont équivalentes si et seulement si elles ont les mêmes diviseurs élémentaires.*
- *La suite des diviseurs élémentaires de $\Delta(\underline{\mathrm{P}})$ (cf. 3.3.2.1) est $\underline{\mathrm{P}}$.*
- *Si $\underline{\mathrm{P}}$ est la suite des diviseurs élémentaires de A, on a $\mathrm{A} \sim \Delta(\underline{\mathrm{P}})$. Cette suite peut se calculer algorithmiquement par pivot de Gauss (cf. 3.3.1.1).*

Exercice(s) 3.3.2.4. *Soient $\mathrm{P}, \mathrm{Q} \in \mathbf{k}[\mathrm{T}]$ unitaires et $\mathrm{A} = \begin{pmatrix} \mathrm{P} & 0 \\ 0 & \mathrm{Q} \end{pmatrix}$. Calculer $\delta_1(\mathrm{A})$ et $\delta_2(\mathrm{A})$ et en déduire que les invariants de similitude de A sont $\mathrm{PGCD}(\mathrm{P}, \mathrm{Q}), \mathrm{PPCM}(\mathrm{P}, \mathrm{Q})$. Retrouver ce résultat par le pivot.*

En déduire un autre algorithme que celui du pivot pour calculer les invariants de similitude d'une matrice diagonale de $\mathbf{k}[\mathrm{T}]$. *[Si* $\mathrm{Q}_i, i \in \mathrm{I}$ *sont ses coefficients diagonaux, on pourra s'intéresser à* $\mathrm{PGCD}(\mathrm{Q}_{i_1} \cdots \mathrm{Q}_{i_r})$ *lorsque* $\{i_1, \cdots, i_r\}$ *décrit les parties à* r *éléments de* I*].*

Remarque(s) 3.3.2.5. *Le lecteur adaptera sans peine le théorème précédent au cas de l'équivalence pour les matrices à coefficients dans un anneau euclidien (muni d'une division euclidienne). Il faut juste pour cela accepter une unicité des diviseurs élémentaires à multiplication par un inversible près. L'énoncé d'unicité se généralise sans changement. Si l'existence 3.3.1.1 reste vraie dans un anneau principal, sa preuve par pivot elle ne fonctionne plus (cf. exercice infra). Or, il existe des anneaux principaux non euclidiens : c'est par exemple le cas de* $\mathbf{Z}[\frac{1+\sqrt{-19}}{2}]$ *(cf. [Per88]) ou de* $\mathbf{R}[x,y]/(x^2+y^2+1)$ *(cf. [Bev16]). Dans le cas principal, il faut pour cela ajouter une opération permise de plus (3.3.2.6). Je dis bien qu'en général* il faut. *Cette différence algorithmique est une fenêtre sur la* K*-théorie algébrique (cf. 3.4).*

Exercice(s) 3.3.2.6. *Soit* R *un anneau principal et considérons des éléments de* R *tels que* $au - bv = 1$. *On définit les opérations de Bézout sur les matrices à coefficients dans* R *comme étant les multiplications à gauche ou à droite par les matrices inversibles diagonales par blocs de type*

$$\begin{pmatrix} \begin{pmatrix} a & v \\ b & u \end{pmatrix} & 0_{2,n} \\ 0_{n,2} & \mathrm{Id}_n \end{pmatrix}$$

Généraliser la preuve de 3.3.1.1 en autorisant en plus des opérations élémentaires les opérations de Bézout.

3.3.3 Les classes d'équivalence de $\mathrm{M}_{p,q}(\mathbf{k}[\mathrm{T}])$

Nous avons donc résolu notre problème initial (3). En effet, si A est dans une classe d'équivalence de $\mathrm{M}_{p,q}(\mathbf{k}[\mathrm{T}])_{/\sim}$, ses diviseurs élémentaires $\underline{\mathrm{P}}(\mathrm{A}) = \underline{\mathrm{P}} = \mathrm{P}_i$ sont bien définis et ne dépendent que de la classe $(\mathrm{A} \mod \sim)$. En effet, le théorème 3.3.2.3 assure que le quotient $\mathrm{M}_{p,q}(\mathbf{k}[\mathrm{T}])_{/\sim}$ quotient s'identifie à l'ensemble des suites $\underline{\mathrm{P}}$ de polynômes unitaires décroissantes pour la divisibilité de longueur $r \leq \inf(p,q)$ et l'application quotient s'identifie à $\mathrm{A} \mapsto \underline{\mathrm{P}}(\mathrm{A})$.

3.4 Complément : fenêtre sur la K-théorie

Cette section est culturelle et vise à introduire une idée importante en mathématique : comment mesurer l'obstruction à ce qu'un résultat soit vrai. Ici, la question posée est comment mesurer l'éventuelle impossibilité de « diagonaliser » au sens de 3.3.1.1 les matrices par pivot de Gauss à coefficients dans un anneau R quelconque.

La question précise à laquelle on s'attaque naturellement est alors la suivante : le groupe $\mathrm{GL}_n(\mathrm{R})$ est-il engendré par les matrices élémentaires de transvections de type pivot(cf. 1.2) -on verra après pour les matrices de permutation, qui sont anodines, et de dilatations qui se traitent via l'application déterminant).

Le premier pas est de s'affranchir de n : pour cela, on voit $\mathrm{GL}_n(\mathrm{R})$ comme le sous groupe de $\mathrm{GL}_{n+1}(\mathrm{R})$ des matrices diagonales par blocs $\mathrm{diag}(\mathrm{M}, 1),\ \mathrm{M} \in \mathrm{GL}_n(\mathrm{R})$ ce qui permet de considérer leur union infinie $\mathrm{GL}(\mathrm{R})$, vue comme ensemble de matrices de taille infinie donc qui contient donc tous les groupes linéaires de taille finie. On définit alors $\mathrm{E}(\mathrm{A})$ comme le sous-groupe de $\mathrm{GL}(\mathrm{A})$ engendré par toutes les transvections qui mesure les matrices de déterminant 1 qu'on peut atteindre par pivot (quitte à s'autoriser à agrandir les matrices donc).

Le premier résultat est à la fois simple et remarquable, surtout dans la preuve qu'en a donné [Mil66].

Lemme 3.4.0.1 (Whitehead)**.** *Pour tout anneau* R*, le groupe* $\mathrm{E}(\mathrm{R})$ *est le groupe dérivé* $[\mathrm{GL}(\mathrm{R}), \mathrm{GL}(\mathrm{R})]$ *engendré par le commutateurs* $[\mathrm{A}, \mathrm{B}] = \mathrm{ABA}^{-1}\mathrm{B}^{-1}$ *de matrices de* $\mathrm{GL}(\mathrm{R})$*.*

Démonstration. *Il suffit d'écrire, encore fallait-il les voir !- les deux formules*

$$\begin{pmatrix} [\mathrm{A},\mathrm{B}] & 0 \\ 0 & \mathrm{I}_n \end{pmatrix} = \begin{pmatrix} \mathrm{A} & 0 \\ 0 & \mathrm{A}^{-1} \end{pmatrix} \begin{pmatrix} \mathrm{B} & 0 \\ 0 & \mathrm{B}^{-1} \end{pmatrix} \begin{pmatrix} (\mathrm{BA})^{-1} & 0 \\ 0 & \mathrm{BA} \end{pmatrix}$$

et

$$\begin{pmatrix} \mathrm{B} & 0 \\ 0 & \mathrm{B}^{-1} \end{pmatrix} = \begin{pmatrix} \mathrm{I}_n & \mathrm{B} \\ 0 & \mathrm{I}_n \end{pmatrix} \begin{pmatrix} \mathrm{I}_n & 0 \\ \mathrm{I}_n - \mathrm{B}^{-1} & \mathrm{I}_n \end{pmatrix} \begin{pmatrix} \mathrm{I}_n & -\mathrm{I}_n \\ 0 & \mathrm{I}_n \end{pmatrix} \begin{pmatrix} \mathrm{I}_n & 0 \\ \mathrm{I}_n - \mathrm{B} & \mathrm{I}_n \end{pmatrix}$$

ramène à prouver que toute matrice de type $\begin{pmatrix} \mathrm{I}_n & \mathrm{B} \\ 0 & \mathrm{I}_n \end{pmatrix}$ *est produit de transvections ce qu'on montre sans problème par pivot de Gauss (exercice).* ∎

En particulier, $\mathrm{E}(\mathrm{A})$ est distingué et le quotient $\mathrm{K}_1(\mathrm{R}) = \mathrm{GL}(\mathrm{R})/[\mathrm{GL}(\mathrm{R}), \mathrm{GL}(\mathrm{R})]$ est un groupe commutatif puisque c'est l'abélianisé de $/\mathrm{GL}(\mathrm{R})$! C'est le groupe de K-théorie algébrique [3] de degré 1. Comme le déterminant de tout commutateur vaut 1, le morphisme déterminant passe au quotient (2.5.2.1) pour définir le groupe spécial de K-théorie algébrique de degré 1

$$\mathrm{SK}_1(\mathrm{R}) : \mathrm{Ker}(\mathrm{GL}(\mathrm{R}) \xrightarrow{\det} \mathrm{R}^\times)$$

qui permet d'éviter de considérer les dilatations et matrices de permutation qui ne jouent pas de rôle déterminant dans le pivot. L'inclusion $\mathrm{R}^\times = \mathrm{GL}_1(\mathrm{R}) \to \mathrm{GL}(\mathrm{R})$ suivie de la projection quotient $\mathrm{GL}(\mathrm{R}) \to \mathrm{K}_1(\mathrm{R})$ permet de définir un morphisme

$$\mathrm{R}^\times \times \mathrm{SK}_1(\mathrm{R}) \to \mathrm{K}_1(\mathrm{R})$$

qui est visiblement un isomorphisme.

3. Comme défini par Bass–Schanuel. Le lecteur voulant aller plus loin pourra étudier [Ros94].

Le groupe $SK_1(R)$ est visiblement l'obstruction à ce que l'algorithme de pivot (« infini ») permette de diagonaliser les matrices. Et nos résultats prouvent que si R est euclidien, $SK_1(R) = 0$. Il est à remarquer que cette obstruction est très subite. Par exemple, dans le cas de l'anneau principal non euclidien $R = \mathbf{Z}[\frac{1+\sqrt{-19}}{2}]$, on a $SK_1(R) = \{1\}$ (c'est un théorème général très difficile cf. [BMS67]). Autrement dit, ce n'est pas un exemple où le pivot avec les matrices élémentaires est insuffisant, au moins en se permettant d 'augmenter la taille des matrices. Trouver R principal tel que $SK_1(R)$ est non nul est difficile. Un exemple est donné dans [Gra81] : on prend le sous anneau de $\mathbf{Z}(T)$ engendré par $\mathbf{Z}[T]$ et les $(T^m - 1)^{-1}$, $m \geq 1$ c'est un anneau principal (!) dont le SK_1 est même infini.

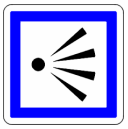
3.5 Exercices supplémentaires

Exercice(s) 3.5.0.1. *Soit* R *un anneau euclidien. Montrer que* $SL_n(R)$ *est engendré par les transvections.*

Chapitre 4

Classes de similitude de $\mathrm{M}_n(\mathbf{k})$

4.1 Point de vue



Dans notre objet d'étude, l'algèbre linéaire, on propose une vision algorithmique de la réduction des endomorphsimes (en dimension finie) sur un corps quelconque. On adopte systématiquement le dictionnaire entre $\mathbf{k}[\mathrm{T}]$-modules et endomorphismes (cf. 2.2.4).

Une motivation est d'une part la simplicité de la théorie quand on accepte le vocabulaire des modules, et, surtout, le fait que la théorie de la réduction habituelle utilise de manière plus ou moins explicite les racines du polynôme caractéristique, qu'en général on ne sait pas calculer... Aussi étrange que cela puisse paraître, le point de vue module rend la théorie algorithmique et transparente, en s'affranchissant de la connaissance de ces racines. Bien entendu, les valeurs propres jouent un rôle important sous-jacent, rôle que l'on explicitera (cf. 5.2 et 7.3).

4.2 Introduction

Règles de la méthode de Descartes [1] :

1. *Ne recevoir aucune chose pour vraie que je ne la connusse évidemment être telle.*

1. R. Descartes, *Discours de la méthode* (1637), Gallimard (2009).

René Descartes

Fidèle à l'enseignement de Descartes, nous allons classifier complètement les endomorphismes (4.9.0.2) et les réduire deux classes très simples : les bien-nommés endomorphismes semi-simples et les nilpotents (5.3.2.1).

2. *Diviser chacune des difficultés que j'examinerais, en autant de parcelles qu'il se pourrait et qu'il serait requis pour les mieux résoudre.*
3. *Conduire par ordre mes pensées, en commençant par les objets les plus simples et les plus aisés à connaître pour monter peu à peu, comme par degrés, jusqu'à la connaissance des plus composés.*
4. *Faire partout des dénombrements si entiers, et des revues si générales, que je fusse assuré de ne rien omettre ". C'est la règle du dénombrement. Faire une revue entière, générale des objets ce qui fait intervenir la prudence, la circonspection.*

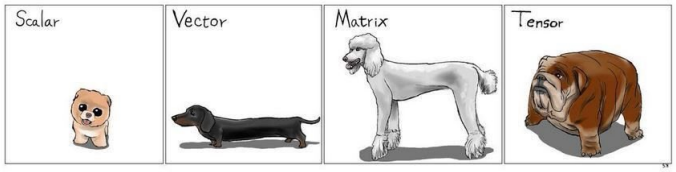
Nous utiliserons librement les propriétés habituelles des anneaux principaux (identité de Bézout, factorialité...) et le fait que les anneaux euclidiens sont principaux. Les exemples clefs pour nous sont $\mathbf{Z}$ et $\mathbf{k}[\mathrm{T}]$. Le lecteur pourra si besoin se référer (sans cercle vicieux) au chapitre 10 pour la factorialité des anneaux principaux.
Un résultat bien utile est une généralisation de la division euclidienne lorsque le diviseur est à coefficient dominant unitaire : c'est le résultat *infra* dont la preuve est une simple relecture de la preuve usuelle (exercice).

Lemme 4.2.0.1 (Division euclidienne généralisée)**.** *Soit* A, B *deux polynômes à coefficients dans un anneau unitaire* $\mathcal{R}$ *non nécessairement commutatif avec* B *non nul. On suppose que le coefficient dominant de* B *est inversible à droite (resp. à gauche). Alors, il existe* $\mathrm{Q}_d, \mathrm{R}_d \in \mathcal{R}[\mathrm{T}]$ *(resp.* $\mathrm{Q}_g, \mathrm{R}_g$*) tels que* $\mathrm{A} = \mathrm{BQ}_d + \mathrm{R}_d$ *avec* $\deg(\mathrm{Q}_d) < \deg(\mathrm{R}_d)$ *division euclidienne à droite (resp.* $\mathrm{A} = \mathrm{Q}_g\mathrm{B} + \mathrm{R}_g$ *avec* $\deg(\mathrm{Q}_g) < \deg(\mathrm{R}_g)$ *division euclidienne à gauche). Si de plus* $\mathcal{R}$ *est intègre à gauche(resp. à droite), on a unicité à gauche (resp. à droite).*

Bien qu'il sera nécessaire de batailler avec des matrices nilpotentes par exemple, bien loin d'être diagonalisables, il ne faut pas perdre de vue que ces matrices sont en fait pathologiques : dans le cas complexe par exemple, une matrice tirée au hasard est presque sûrement avec des valeurs propres distinctes !
Une raison est donnée à titre d'échauffement en exercice (4.14.0.3). Pour autant, les mathématiques fournissent naturellement bien des matrices improbables.

4.2.1 Notations

Dans ce chapitre, on désigne (cf. 2) par



— V un espace de dimension finie n sur un corps $\mathbf{k}$ arbitraire.

— V[T] le $\mathbf{k}[T]$-module des polynômes à coefficients dans V.

— $V_a, a \in \mathrm{End}_{\mathbf{k}}(V)$ le K[T]-module $V = V_a$ caractérisé (2.2.4) par

$$Tv = a(v) \text{ pour tout } x \in V = V_a$$

et plus généralement $P(T)v = P(a)(v)$.

— On désignera par $A, B \cdots$ les matrices de $a, b \cdots$ après choix d'une base de V.

— On note χ_a (resp. μ_a) les polynômes caractéristique (resp. minimal) [2] de a.

— $\pi_a \in \mathrm{Hom}_{\mathbf{k}[T]}(V[T], V_a)$ défini par

$$\pi_a(\sum v_i T^i) = \sum a^i(v_i)$$

la surjection canonique prolongeant l'identité l'identité de V vu comme ensemble des polynômes constants de V[T].

— $\mathbf{k}_\lambda = \mathbf{k}[T]/(T - \lambda)$ le module V_a avec $V = \mathbf{k}$ et a l'homotéhie de rapport $\lambda \in \mathbf{k}$.

— $\tilde{a} \in \mathrm{End}_{\mathbf{k}[T]}(V[T])$ caractérisé par

$$\tilde{a}(vT^i) = a(v)T^i$$

l'unique prolongement $\mathbf{k}[T]$ linéaire de $a \in \mathrm{End}_{\mathbf{k}}(V)$ à V[T],

— C(P) la matrice compagnon du polynôme unitaire $P = T^n + \sum_{i=0}^{n-1} a_i T^i$

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in \mathrm{M}_n(\mathbf{k}).$$

Ainsi, C(P) est la matrice vide si $P = 1$,

2. Formellement, il n'est pas indispensable à ce stade de savoir ce qu'est le polynôme minimal de a car la théorie de la réduction va redonner son existence et ses propriétés.

— $\mathrm{J}_n = \mathrm{C}(\mathrm{T}^n)$ le bloc de Jordan standard

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \in \mathrm{M}_n(\mathbf{k})$$

de taille n.

Dès lors qu'une base de V aura été choisie, on identifiera sans plus de précision V à $\mathbf{k}^n$ et V[T] à $\mathbf{k}[\mathrm{T}]^n$ (cf. 2.2.4 (2)) de sorte que a et $\tilde{a}$ ont même matrice.

4.3 Stratégie

Rappelons que deux matrices carrées A, B de $\mathrm{M}_n(\mathbf{k})$ sont semblables (dans $\mathbf{k}$!) (on note $\mathrm{A} \approx \mathrm{B}$) si et seulement si il existe $\mathrm{P} \in \mathrm{GL}_n(\mathbf{k})$ telles que $\mathrm{A} = \mathrm{PBP}^{-1}$. Ceci définit bien... une relation d'équivalence (pour ne pas la confondre avec l'équivalence précédente des matrices polynômiales on notera $\overline{\overline{\mathrm{A}}}$ la classe de similitude d'une matrice carrée scalaire).

Nous allons exhiber dans chaque classe d'équivalence un représentant canonique. Pour ce faire, nous allons exhiber pour toute matrice carré $\mathrm{A} \in \mathrm{M}_n(\mathbf{k})$ un représentant canonique dans sa classe de similitude en deux temps à l'instar de ce que nous avons fait pour les matrices polynômiales.

1. Notons que $\mathrm{P}(\mathrm{T\,Id} - \mathrm{A})\mathrm{P}^{-1} = \mathrm{T\,Id} - \mathrm{PA} - \mathrm{P}^{-1}$ de sorte que si $\mathrm{A} \approx \mathrm{B}$, on a $\mathrm{T\,Id} - \mathrm{A} \sim \mathrm{T\,Id} - \mathrm{B}$. Le miracle est que la réciproque est vraie. Nous allons donc montrer que l'application
$$\begin{cases} \mathrm{M}_n(\mathbf{k})_{/\approx} & \hookrightarrow & \mathrm{M}_n(\mathbf{k}[\mathrm{T}])_{/\sim} \\ \overline{\overline{\mathrm{A}}} & \mapsto & \overline{\mathrm{T\,Id} - \mathrm{A}} \end{cases}$$
est injective : c'est le corollaire 4.4.0.2.
2. De manière analogue à ce qui précède, nous exhiberons grâce à $\Delta(\overline{\mathrm{T\,Id} - \mathrm{A}})$ un représentant canonique $\mathrm{C}(\underline{\mathrm{P}})$ dans toute classe de similitude $\mathrm{M}_n(\mathbf{k})_{/\approx}$) à l'aide de matrices compagnons associées aux diviseurs $\underline{\mathrm{P}}$ de $\mathrm{T\,Id} - \mathrm{A}$: c'est le théorème de décomposition de Frobenius 4.9.0.2.

4.4 Invariance par équivalence de $\mathrm{T\,Id} - \mathrm{A}$ du module V_a et applications

Soit $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ et V_a le $\mathbf{k}[\mathrm{T}]$-module associé (2.2.4).

Lemme 4.4.0.1. *La suite*

(i) $$0 \to \mathrm{V}[\mathrm{T}] \xrightarrow{\mathrm{TId} - \tilde{a}} \mathrm{V}[\mathrm{T}] \xrightarrow{\pi_a} \mathrm{V}_a \to 0$$

est exacte.

Démonstration. *Soit* $v \in \mathrm{V}$. *On a*

$$\pi_a \circ (\mathrm{TId} - \tilde{a})(v) = \mathrm{T}\pi_a(v) - a(v) = a(v) - a(v) = 0$$

de sorte que $\pi_a \circ (\mathrm{TId} - \tilde{a}) = 0$ *puisque* V *engendre* $\mathrm{V[T]}$ *et donc* $\mathrm{Im}(\mathrm{TId} - \tilde{a}) \subset \mathrm{Ker}(\pi_a)$.
Inversement, soit $v(\mathrm{T}) = \sum_{i \geq 0} \mathrm{T}^i v_i \in \mathrm{Ker}(\pi_a)$, *ie*

$$v_0 + \sum_{i \geq 1} a^i(v_i) = 0.$$

On a donc

$$v(\mathrm{T}) = \sum_{i \geq 1} (\mathrm{T}^i \mathrm{Id} - \tilde{a}^i)(v_i).$$

Mais comme TId *et* $\tilde{a}$ *commutent, on a (somme progression géométrique)*

$$\mathrm{T}^i \mathrm{Id} - \tilde{a}^i = (\mathrm{TId} - \tilde{a}) \circ (\sum_{j=0}^{i-1} \mathrm{T}^j \tilde{a}^{i-1-j})$$

et donc $v(\mathrm{T}) \in \mathrm{Im}(\mathrm{TId} - \tilde{a})$. *D'où l'exactitude au milieu. L'exactitude à gauche, facile et inutile pour nous, est laissée en utile exercice.*

En d'autres termes, on a

(ii) $$\mathrm{Coker}(\mathrm{TId} - \tilde{a}) = \mathrm{V}_a$$

ou si on est puriste $\mathrm{Coker}(\mathrm{TId} - \tilde{a}) = \pi_a$ (2.4.0.1). ■

Le choix d'une base de V identifie V à $\mathbf{k}^n$ et $\mathrm{V[T]}$ à $\mathbf{k}[\mathrm{T}]^n$ (2.2.4). La matrice de $\tilde{a}$ est alors A, matrice de a la base choisie. La suite exacte précédente (i) s'identifie alors à

(iii) $$0 \to (\mathbf{k}[\mathrm{T}])^n \xrightarrow{\mathrm{TId}-\mathrm{A}} (\mathbf{k}[\mathrm{T}]))^n \xrightarrow{\pi_\mathrm{A}} \mathrm{V_A} = (\mathbf{k}^n)_\mathrm{A} \to 0$$

avec $\pi_\mathrm{A}(\sum \mathrm{X}_i \mathrm{T}^i) = \sum \mathrm{A}^i \mathrm{X}_i$ et $\mathrm{T.X} = \mathrm{AX}$ pour tout $\mathrm{X}_i, \mathrm{X} \in \mathbf{k}^n$. On en déduit le résultat important

Corollaire 4.4.0.2. *Soit* $\mathrm{A}, \mathrm{B} \in \mathrm{M}_n(\mathbf{k})$ *les matrices de* $a, b \in \mathrm{End}_\mathbf{k}(\mathrm{V})$ *dans une base. Les propositions suivantes sont équivalentes.*

1. A *et* B *sont semblables dans* $\mathrm{M}_n(\mathbf{k})$.
2. $\mathrm{T\,Id} - \mathrm{A}$ *et* $\mathrm{T\,Id} - \mathrm{B}$ *sont équivalentes dans* $\mathrm{M}_n(\mathbf{k}[\mathrm{T}])$.
3. *Les* $\mathbf{k}[\mathrm{T}]$*-modules* V_a *et* V_b *sont isomorphes.*

De plus, si $\mathrm{T\,Id} - \mathrm{A} \simeq \Delta \in \mathrm{M}_n(\mathbf{k}[\mathrm{T}])$, *on a* $\mathrm{V}_a \simeq \mathrm{Coker}(\Delta : \mathbf{k}[\mathrm{T}]^n \to \mathbf{k}[\mathrm{T}]^n)$.

Démonstration. $1 \Rightarrow 2$. *Si* $\mathrm{P} \in \mathrm{GL}_n(\mathbf{k})$ *vérifie* $\mathrm{PAP}^{-1} = \mathrm{B}$, *a* $\mathrm{P}(\mathrm{T\,Id} - \mathrm{A})\mathrm{P}^{-1} = \mathrm{T\,Id} - \mathrm{B}$ *et donc* $\mathrm{T\,Id} - \mathrm{A} \sim \mathrm{T\,Id} - \mathrm{B}$.

$2 \Rightarrow 3$. *Il existe donc* $\mathrm{P(T)}, \mathrm{Q(T)} \in \mathrm{GL}_n(\mathbf{k}[\mathrm{T}])$ *telles que* $\mathrm{P(T)(TId - A)Q(T)}^{-1} = \mathrm{T\,Id - B}$. *Dans ce cas,* $\mathrm{P(T)}, \mathrm{Q(T)}$ *définissent d'après ce qui précède un diagramme commutatif à lignes exactes et colonnes isomorphismes*

$$\begin{array}{ccccccc} \mathbf{k}[\mathrm{T}]^n & \xrightarrow{\mathrm{TId-A}} & \mathbf{k}[\mathrm{T}]^n & \xrightarrow{\pi_\mathrm{A}} & \mathrm{V_A} & \longrightarrow & 0 \\ \downarrow{\scriptstyle \mathrm{Q(T)}} & & \downarrow{\scriptstyle \mathrm{P(T)}} & & & & \\ \mathbf{k}[\mathrm{T}]^n & \xrightarrow{\mathrm{TId-B}} & \mathbf{k}[\mathrm{T}]^n & \xrightarrow{\pi_\mathrm{B}} & \mathrm{V_B} & \longrightarrow & 0 \end{array}$$

et donc par fonctorialité du conoyau (2.4.0.1) un unique isomorphisme $\mathbf{k}[\mathrm{T}]$ *linéaire*

$$\iota : \mathrm{V_A} \to \mathrm{V_B},$$

c'est à dire (2.2.4.1) une matrice inversible

$$\mathrm{S} : \mathrm{V_A} = \mathbf{k}^n \to \mathbf{k}^n = \mathrm{V_B}$$

vérifiant $\mathrm{SA} = \mathrm{BS}$ *(car* $\iota(\mathrm{T}.v) = \iota(a(v)) = \mathrm{T}.\iota(v) = b(\iota(v))$ *cf. TD).*
$2 \Rightarrow 3$. *On a déjà remarqué (2.2.4.1) que l'existence de l'isomorphisme* $\iota : \mathrm{V_A} = \mathbf{k}^n \to \mathbf{k}^n = \mathrm{V_B}$ *définissait* $\mathrm{S} \in \mathrm{GL}_n(\mathbf{k})$ *tel que* $\mathrm{SA} = \mathrm{BS}$. ■

L'équivalence des deux premiers points se réécrit $\mathrm{A} \approx \mathrm{B}$ si et seulement si $\mathrm{T\,Id - A} \sim \mathrm{T\,Id - B}$, soit l'injectivité cherchée en (2.2.4.)

4.5 Invariants de similitude de $a \in \mathrm{End}_\mathbf{k}(\mathrm{V})$

D'après le corollaire 4.4.0.2, il est raisonnable de poser la définition suivante.

Définition 4.5.0.1. *Les diviseurs élémentaires de* $\mathrm{TId} - \mathrm{A}, \mathrm{A} \in \mathrm{M}_n(\mathbf{k})$ *s'appellent les invariants de similitude de* A.

Le corollaire 4.4.0.2 se réécrit alors

Théorème 4.5.0.2 (Invariants de similitude)**.** *Soit* $a, b \in \mathrm{End}_\mathbf{k}(\mathrm{V})$. *Les propositions suivantes sont équivalentes.*

1. *a et b ont mêmes invariants de similitude* $\underline{\mathrm{P}}$.
2. *a et b sont semblables dans* $\mathrm{End}_\mathbf{k}(\mathrm{V})$.
3. $\mathrm{T\,Id - A}$ *et* $\mathrm{T\,Id - B}$ *sont équivalentes à* $\Delta(\underline{\mathrm{P}})$
4. *Les* $\mathbf{k}[\mathrm{T}]$*-modules* V_a *et* V_b *sont isomorphes.*
5. *Les* $\mathbf{k}[\mathrm{T}]$*-modules* V_a *et* $\mathrm{Coker}(\Delta(\underline{\mathrm{P}}))$*sont isomorphes.*

Corollaire 4.5.0.3. *Soit* $A, B \in M_n(\mathbf{k})$ *et soit* K *un surcorps de* $\mathbf{k}$. *Alors,*

— A *et* B *sont semblables sur* K *si et seulement si elles sont semblables sur* $\mathbf{k}$.

— A *et* ${}^tA$ *sont semblables.*

Démonstration. *Le premier point découle par exemple du fait que les invariants de similitude de* A *se calculent par pivot de Gauss sur* $T\,Id - A$, *algorithme indépendant du surcorps où on calcule. Le second résulte de deux observations :* $A \sim B$ *entraîne* ${}^tA \sim {}^tB$ *(écrire la définition de l'équivalence) et* $A \sim \Delta(\underline{(P)}) = {}^t\Delta(\underline{P})$ *car* Δ *diagonale carrée.* ■

Remarque(s) 4.5.0.4. *Une autre de manière dire la même chose est la suivante : soit* $(A, B), (A', B')$ *des matrices carrées avec* A *inversible. Alors, il existe* P, Q *inversibles telles que* $PAQ = A'$ *et* $PBQ = B'$ *si et seulement* A' *est également et* $A^{-1}B$ *et* $A'^{-1}B'$ *ont mêmes invariants de similitude. En effets, la partie directe est évidente car dans ce cas* A' *est inversible* $(A'^{-1}PA)(A^{-1}B+T)Q = (A'^{-1}B'+T)$. *Inversement, si* $A^{-1}B$ *et* $A'^{-1}B'$ *ont mêmes invariants de similitude, il existe* Π *inversible tel que* $\Pi^{-1}A^{-1}B\Pi = A'^{-1}B'$ *et on pose* $P = A'\Pi^{-1}A^{-1}, Q = \Pi$.

Avant de passer au second point annoncé en (2.2.4), la décomposition de Frobenius 4.9.0.2, donnons quelques propriétés spécifiques aux invariants de similitude A liées au caractère bien spécifique de la matrice polynômiale $T\,Id - A$ et tirons-en quelques corollaires délicieux.

4.6 Calcul de V_a et applications

Soit A une matrice de a dans une base. Comme $V_a = \mathrm{Coker}(T\,Id - A)$, il ne dépend à isomorphisme près que de la classe d'équivalence de $T\,Id - A$ donc de ses diviseur élémentaires $\underline{P}$, qui par définition sont les invariants de similitude de a. Comme $T.\,Id - A \sim \Delta(\underline{P})$, il suffit de calculer $V_a = \mathrm{Coker}(\Delta(\overline{P}))$ dans ce cas diagonal .

Proposition 4.6.0.1. *Soit $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ et $\underline{\mathrm{P}} = (\mathrm{P}_r|\cdots|\mathrm{P}_1)$ ses invariants de similitude.*

1. *On a $r = n$ et $\prod_{i=1}^{n} \mathrm{P}_i = \chi_a(\mathrm{T})$.*
2. *On a $\Delta(\underline{\mathrm{P}}) = \mathrm{diag}(\mathrm{P}_1, \cdots, \mathrm{P}_n)$.*
3. *On a $\mathrm{P}_1|\chi_a|\mathrm{P}_1^n$ de sorte que χ_a et P_1 ont les mêmes facteurs irréductibles (et donc les mêmes racines dans toute extension de $\mathbf{k}$).*
4. *Le $\mathbf{k}[\mathrm{T}]$-module V_a est isomorphe à $\oplus_{i=1}^{n}\mathbf{k}[\mathrm{T}]/(\mathrm{P}_i)$.*
5. *P_1 est le minimal μ_a de a.*
6. *$\chi_a(a) = 0$ (Cayley-Hamilton).*

Démonstration. *Soit $\mathrm{A} \in \mathrm{M}_n(\mathbf{k})$ la matrice de a dans une base V de sorte que $\mathrm{T\,Id} - \mathrm{A} \sim \Delta(\underline{\mathrm{P}})$ (3.3.2.3). Il existe donc $\mathrm{P}(\mathrm{T}), \mathrm{Q}(\mathrm{T}) \in \mathrm{GL}_n(\mathbf{k}[\mathrm{T}])$ telles que*

$$\mathrm{P}(\mathrm{T}) \begin{pmatrix} \mathrm{diag}(\mathrm{P}_1, \cdots, \mathrm{P}_r) & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r,n-r} \end{pmatrix} \mathrm{Q}(\mathrm{T}) = \mathrm{TId} - \mathrm{A}.$$

Comme les déterminants de $\mathrm{P}(\mathrm{T})$ et $\mathrm{Q}(\mathrm{T})$ sont des scalaires non nuls et que tant les P_i que le polynôme caractéristique χ_{A} sont unitaires on a en prenant le déterminant de l'identité précédente $r = n$ et $\chi_{\mathrm{A}}(\mathrm{T}) = \mathrm{P}_1 \cdots \mathrm{P}_n$ d'où (1) et (2).

Tenant compte du fait que P_1 est un multiple de tous les P_i, on a $\mathrm{P}_1|\chi_{\mathrm{A}}|\mathrm{P}_1^n$. ce qui donne les points 1 et 2. Comme P_1 est multiple de chaque P_i, en prenant le produit, on a que P_1^n multiple de χ_a de sorte qu'on a $\mathrm{P}_1|\chi_a|\mathrm{P}_1^n$ d'où (3).

D'après (2.3.2.1), la suite

$$(\mathbf{k}[\mathrm{T}])^n \xrightarrow{\Delta} (\mathbf{k}[\mathrm{T}])^n \to \oplus_{i=1}^{n}\mathbf{k}[\mathrm{T}]/(\mathrm{P}_i) \to 0$$

est exacte et s'identifie par fonctorialité du conoyau (cf. 4.5.0.2)

$$\mathrm{Coker}(\Delta) = \oplus_{i=1}^{n}\mathbf{k}[\mathrm{T}]/(\Delta_{i,i}) = \mathrm{V}_a$$

d'où (4).

Comme P_1 multiple de chaque P_i, il annule tous les $\mathbf{k}[\mathrm{T}]_{(/\mathrm{P}_i)}$ donc également V_a. Par ailleurs, si P est de degré $< \deg(\mathrm{P}_1)$ il n'annule pas la classe de 1 dans, $\mathbf{k}[\mathrm{T}](\mathrm{P}_1)$ et donc $\mathrm{P}(a)$ n'annule pas l'antécédent de cette classe dans V_a de sorte que P_1 est bien le minimal μ_a de a d'où les points (5) et (6). ■

Remarque(s) 4.6.0.2. *Comme on le verra infra (par exemple 4.8.0.1), les derniers P_i sont bien souvent égaux 1. Ils contribuent par le module nul à V_a.*

4.7 Diagonalisation

On garde les notations de 4.6.0.1.

Corollaire 4.7.0.1. *L'endomorphisme a est diagonalisable si et seulement si son polynôme minimal $\mathrm{P}_1 = \mu_a$ est scindé sur $\mathbf{k}$ à racines simples. En particulier, la restriction d'un diagonalisable à sous-espace stable est diagonalisable.*

Démonstration. *Supposons P_1, et donc tous les P_i (qui le divisent) scindés à racines simples. Alors, d'après (4.6.0.1) et le lemme des noyaux 2.6.0.2, on a*

$$\mathrm{V}_a \simeq \oplus_i \mathbf{k}[\mathrm{T}]/\mathrm{P}_i \simeq \oplus_i \oplus_{\lambda|\mathrm{P}_i(\lambda)=0} \mathbf{k}[\mathrm{T}]/(\mathrm{T}-\lambda) = \oplus_i \oplus \mathbf{k}_\lambda$$

ce qui diagonalise a car $\mathrm{T}v = a(v) = \lambda v$ pour $v \in \mathbf{k}_\lambda$. La réciproque est claire.
Si W est stable par a, la restriction a_W de a à W est annulée par μ_a qui est donc un multiple de son minimal. Ainsi, μ_{a_W} à racines simple comme μ_a. ■

Remarque(s) 4.7.0.2. *On utilise souvent ce critère sous la forme équivalente suivante : l'endomorphisme a est diagonalisable si et seulement si il admet un polynôme annulateur scindé sur $\mathbf{k}$ à racines simples. Par exemple, toute matrice complexe vérifiant $\mathrm{A}^\mathrm{N} = \mathrm{Id}$ est diagonalisable.*

Les matrices d'une famille de de matrices diagonales commutent deux à deux. Inversement, il est remarquable et important que la réciproque soit vraie.

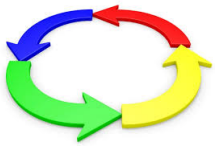
Corollaire 4.7.0.3. *Soit (a_i) une famille arbitraire d'endomorphismes diagonalisables de V. Alors, si $f_i \circ f_j = f_j \circ f_i$ pour tout i, j, il existe une base de diagonalisation commune à tous les f_i.*

Démonstration. *On fait un récurrence sur $n = \dim(\mathrm{V}) \geq 0$. On peut supposer $n > 0$ et l'enoncé vrai en dimension $< n$. Si tous les f_i sont des homothéthies $\lambda_i \mathrm{Id}$, n'importe quelle base convient. Sinon, soit i tel que f_i n'est pas une homotéthie. Alors, f_i a au moins deux valeurs propres de sorte que tous ses espaces propres $\mathrm{E}_i(\lambda)$ sont des de dimension $< n$. Mais ils sont stables par tous les f_j et leurs restrictions $f_j(\lambda)$ à chaque $\mathrm{E}_i(\lambda)$ est diagonalisable pour tout j. Pour chaque λ, on choisit alors une base de diagonalisation commune aux $f_j(\lambda)$ et la réunion de ces bases convient.* ■

Passons au second point annoncé en (4.3). On cherche donc un représentant canonique $\mathrm{C}(\underline{\mathrm{P}})$ dans toute classe de similitude $\overline{\overline{\mathrm{A}}}$ de même qu'on a trouvé le représentant $\Delta(\underline{\mathrm{P}})$ dans $\overline{\mathrm{T}\,\mathrm{Id} - \mathrm{A}}$. La difficulté est que $\Delta(\underline{\mathrm{P}})$ n'est pas de la forme $\mathrm{T}.\,\mathrm{Id} - \mathrm{A}'$. Mais ce n'est pas grave, comme nous allons le voir. Commençons par le cas où un seul des invariants de similitude est de degré > 0.

4.8 Endomorphismes cycliques

Soit V de dimension n et $\mathrm{P} = \mathrm{T}^n + \sum_{i=0}^{n-1} a_i \mathrm{T}^i \in \mathbf{k}[\mathrm{T}]$.



Proposition 4.8.0.1. *Soit $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$. Les propositions sont équivalentes.*

1. *La matrice de* A *dans une base convenable est la matrice compagnon* C(P)
2. $\mu_a = \chi_a = \mathrm{P}$
3. *Les invariants de similitude sont* $1, \cdots, 1, \mathrm{P}$
4. V_a *et* $\mathbf{k}[\mathrm{T}]/(\mathrm{P})$ *sont des* $\mathbf{k}[\mathrm{T}]$*-modules isomorphes.*
5. V_a *est monogène et* $\mu_a = \mathrm{P}$.

Démonstration. $1 \Rightarrow 2$. *Si* $\mathrm{e}_i, 0 \leq i \leq n-1$ *est la base en question, on a* $\mathrm{e}_i = u^i(\mathrm{e}_0)$ *et*

$$u^n(\mathrm{e}_0) = -\sum_{i<n} a_i \mathrm{e}_i = \sum_{i<n} a_i a^i(\mathrm{e}_0).$$

On a donc $\mathrm{V}_a = \mathbf{k}[\mathrm{T}].\mathrm{e}_0$ *et* $\mathrm{P}(\mathrm{T}).\mathrm{e}_0 = 0$ *et ainsi* $\mu_a | \mathrm{P}$. *Comme les* $u^i(\mathrm{e}_0), i < n$ *sont libres, il n'existe pas de polynôme unitaire* Q *de degré* $< n$ *tel que* $\mathrm{Q}(a) = 0$ *car sinon* $\mathrm{Q}(a)(\mathrm{e}_0) = 0$ *serait une relation de liaison et donc* $\mu_a = \mathrm{P}$ *et* $\mu_a = \chi_a$ *pour des raisons de degré puisque* $\mu_a | \mathrm{P}$.
$2 \Rightarrow 3$. *Dire* $\mu_a = \mathrm{P}$, *c'est dire* $\mathrm{P}_1 = \chi_a$ *et donc* $\mathrm{P}_i = 1$ *pour* $i > 1$ *d'après 1. de la proposition 4.6.0.1.*
$3 \Rightarrow 4$ *d'après 4. de la proposition 4.6.0.1*
$4 \Rightarrow 5$ *est tautologique : on a déjà vu* $\mathrm{P} = \mu_a$ *dans la preuve du point 5 de 4.6.0.1. Si* ι *est alors un isomorphisme* $\mathbf{k}[\mathrm{T}]/(\mathrm{P})$ *sur* V_a, *l'élément* $\iota(1 \mod \mathrm{P})$ *engendre* V_a.
$5 \Rightarrow 1$. *Soit* e_0 *un générateur de* V_a. *Le noyau de l'unique* $\mathbf{k}[\mathrm{T}]$ *morphisme surjectif* $\mathbf{k}[\mathrm{T}] \to \mathrm{V}_a$ *qui envoie* 1 *sur* e_0 *est l'annulateur de* e_0 *dans* V_a, *et donc contient* $\mu_a = \mathrm{P}$. *Mais si son générateur unitaire était* Q *de degré* $< n$, *on aurait* $\mathrm{Q}(\mathrm{T})\mathrm{e}_0 = 0$ *mais aussi* $\mathrm{Q}(\mathrm{T})\mathrm{V}_a = (0)$ *puisque* $\mathrm{V}_a = \mathbf{k}[\mathrm{T}]\mathrm{e}_0$. *Donc* $\mu_a = \mathrm{P}$ *diviserait* Q, *ce qui est impossible car* $\deg(\mathrm{P}) = n$. ■

Un vecteur v tel que $\mathbf{k}[\mathrm{T}].v$ est de dimension $\deg(\mu_a)$ est dit *cyclique* tout comme l'espace $\mathbf{k}[\mathrm{T}].v$ qu'il engendre. Autrement dit, $\mathbf{k}[\mathrm{T}].v$ est cyclique si c'est un espace monogène de dimension maximale.

4.9 Décomposition de Frobenius I

Ferdinand Georg Frobenius

On peut maintenant terminer avec le second point annoncé en (4.3). On cherche donc un représentant canonique $\mathrm{C}(\underline{\mathrm{P}})$ dans toute classe de similitude $\overline{\overline{\mathrm{A}}}$.

Définition 4.9.0.1. *Soit* $\underline{\mathrm{P}} = (\mathrm{P}_n, \cdots | \mathrm{P}_1)$ *une suite de polynômes unitaires. On définit* $\mathrm{C}(\underline{\mathrm{P}}) = \mathrm{diag}(\mathrm{P}_i)$ *la matrice compagnon généralisée (de taille* $n = \sum \deg(\mathrm{P}_i)$*).*

Notons que C(1) est... la matrice vide, comme toute matrice d'endomorphisme de $\mathbf{k}[\mathrm{T}]/(1) = (0)$!

Théorème 4.9.0.2 (Réduction de Frobenius)**.** *Soit* $\underline{P}=(\mathrm{P}_n|\cdots|\mathrm{P}_1)$, $i = 1, \cdots, n$ *des polynômes unitaires de* $\mathbf{k}[\mathrm{T}]$ *et* $\mathrm{A} \in \mathrm{M}_n(\mathbf{k})$*.*

1. *La famille des invariants de similitude de* $\mathrm{C}(\underline{\mathrm{P}})$ *est* $\underline{\mathrm{P}}$*.*
2. *Si* $\underline{\mathrm{P}}$ *est la famille des invariants de similitude de* A*, alors* A *est semblable à* $\mathrm{C}(\underline{\mathrm{P}})$*.*

Démonstration. *Soit* r *le plus grand indice* i *tel que* $d_i = \deg(\mathrm{P}_i) > 0$*. D'après la caractérisation des endomorphismes cycliques (4.8.0.1), pour tout* $i \leq r$*, la matrice* $\mathrm{T\,Id} - \mathrm{C}(\mathrm{P}_i)$ *est équivalente à* $\mathrm{diag}(1, \cdots, 1, \mathrm{P}_i)$ *(avec* 1 *répété* $d_i - 1$ *fois) et donc s'écrit* $\mathrm{Q}'_i \,\mathrm{diag}(1, \cdots, 1, \mathrm{P}_i)\mathrm{Q}_i^{-1}$ *avec* $\mathrm{Q}_i, \mathrm{Q}'_i \in \mathrm{GL}_{\deg(\mathrm{P}_i)}(\mathbf{k})$ *tandis que* $\mathrm{C}(\mathrm{P}_i)$ *est vide pour* $i > r$*. On a donc avec* $\mathrm{Q} = \mathrm{diag}(\mathrm{Q}_i), \mathrm{Q}' = \mathrm{diag}(\mathrm{Q}'_i)$, $i \leq r$

$$\mathrm{T\,Id} - \mathrm{A} = \mathrm{Q}\,\mathrm{diag}_i\Big(\mathrm{diag}(1, \cdots, 1, \mathrm{P}_i)\Big)\mathrm{Q}'^{-1} \sim \mathrm{diag}(\mathrm{P}_1, \cdots, \mathrm{P}_r, 1, \cdots, 1)$$

avec 1 *répété* $\sum_{i \leq r}(d_i - 1) = n - r$ *fois de sort* $\mathrm{diag}(\mathrm{P}_1, \cdots, \mathrm{P}_r, 1, \cdots, 1) = \mathrm{diag}(\mathrm{P1}, \cdots, \mathrm{P}_n)$*. Par unicité des diviseurs élémentaires, (1) en découle.*
Pour (2), de (1) découle que A *et* $\mathrm{C}(\underline{\mathrm{P}})$ *ont même invariants de similitude, donc sont semblables.* ■

Exercice(s) 4.9.0.3. *Soit* $\alpha, \beta \in \mathbf{k}$ *et* $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$*. Calculer les invariants de similitude de* $\alpha a + \beta\,\mathrm{Id}$ *en fonction de* α, β *et des invariants de* a*.*

4.9.1 Formulation équivalente

Avec les notations précédentes, on a $\mathrm{V}_a = \oplus \mathrm{V}_i$ où $\mathrm{V}_i \simeq \mathbf{k}[\mathrm{T}]/(\mathrm{P}_i)$ comme $\mathbf{k}[\mathrm{T}]$-modules. En particulier, l'antécédent $v_i \in \mathrm{V}_i$ de $1 \in \mathbf{k}[\mathrm{T}]/(\mathrm{P}_i)$ engendre V_i : c'est un vecteur cyclique de la restriction de a à V_i. Autrement dit, le générateur unitaire μ_{a,v_i} de l'idéal des polynômes P tels que $\mathrm{P}(a)(v_i) = 0$ est de degré maximal, à savoir le degré du minimal de $a_{|\mathrm{V}_i}$. Ou encore, de manière équivalente, $\mu_{a,v_i} = \mu_{a_{|\mathrm{V}_i}}$. On peut réécrire le théorème de Frobenius en disant qu'il existe une décomposition $\mathrm{V}_a = \oplus \mathrm{V}_i$ où chaque V_i est cyclique de polynôme minimal P_i avec la condition de divisibilité habituelle. De plus, s'il existe une telle décomposition, les P_i sont les invariants de similitude.

4.10 Résumé

En rassemblant ce qu'on a prouvé, on a les résultats suivants.
Soit $\mathrm{A}, \mathrm{B} \in \mathrm{M}_n(\mathbf{k})$ et $\underline{\mathrm{P}} = (\mathrm{P}_n|\cdots|\mathrm{P}_1)$ une famille de polynômes unitaires.

— A et B sont semblables si et seulement si elles ont mêmes invariants de similitude.

— La famille des invariants de similitude de $\mathrm{C}(\underline{\mathrm{P}})$ est $\underline{\mathrm{P}}$.

Si $\underline{\mathrm{P}}$ est la famille des invariants de similitude A, on a

— A et $\mathrm{C}(\underline{\mathrm{P}})$ sont semblables.

— On a $\mathrm{V_A} \simeq \oplus \mathbf{k}[\mathrm{T}]/(\mathrm{P}_i)$ où A désigne aussi l'endomorphisme de $\mathrm{V} = \mathbf{k}^n$ associé.

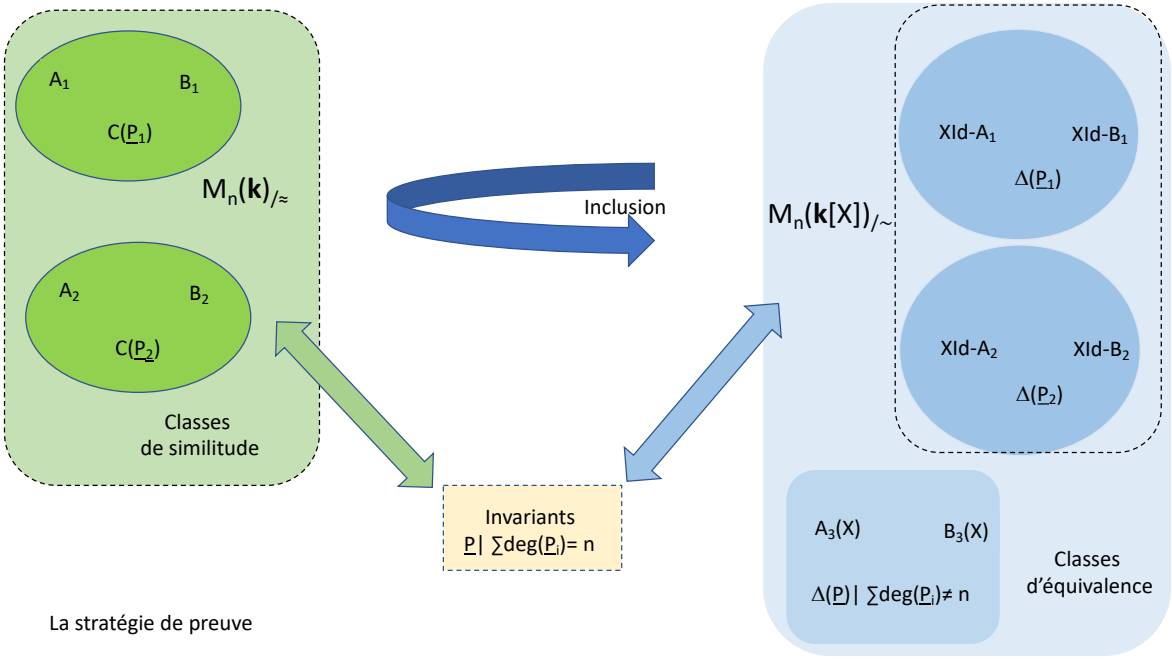
— $\mathrm{TId} - \mathrm{A}$ est équivalente à $\mathrm{diag}(\mathrm{P}_1, \cdots, \mathrm{P}_n)$.

— $\underline{\mathrm{P}}$ se calcule par pivot de Gauss en « diagonalisant » $\mathrm{TId} - \mathrm{A}$ dans $\mathrm{M}_n(\mathbf{k}[\mathrm{T}])$.

— On a $\chi_\mathrm{A} = \mathrm{P}_1 \cdots \mathrm{P}_n$ et $\mathrm{P}_1 = \mu_\mathrm{A}$.

— Les invariants de similitude de $\mathrm{C}(\mathrm{P})$ sont $(1, \cdots, 1, n)$.

La stratégie de preuve est illustrée par le schéma suivant.



La stratégie de preuve

4.11 Application : Commutant

Il est alors aisé d'étudier le commutant (cf. iii)

$$\mathrm{Com}(a) = \mathrm{End}_{\mathbf{k}[\mathrm{T}]}(\mathrm{V}_a) \simeq \mathrm{End}_{\mathbf{k}[\mathrm{T}]}(\oplus \mathbf{k}[\mathrm{T}]/(\mathrm{P}_i)).$$

par exemple de calculer sa dimension.

Proposition 4.11.0.1. *La dimension du commutant de a est $\sum(2i-1)\deg(\mathrm{P}_i)$. En particulier $\dim \mathrm{Com}(a) \geq n$ avec égalité si et seulement si a est cyclique.*

Démonstration. *On a*

$$\mathrm{End}_{\mathbf{k}[\mathrm{T}]}(\oplus \mathbf{k}[\mathrm{T}]/(\mathrm{P}_i)) = \oplus_{i,j} \mathrm{Hom}_{\mathbf{k}[\mathrm{T}]}(\mathbf{k}[\mathrm{T}]/(\mathrm{P}_i), \mathbf{k}[\mathrm{T}]/(\mathrm{P}_j))$$

Comme $\mathbf{k}[\mathrm{T}]/(\mathrm{P}_i)$ *est monogène engendré par la classe de* 1*, un élément de*

$$\mathrm{Hom}_{\mathbf{k}[\mathrm{T}]}(\mathbf{k}[\mathrm{T}]/(\mathrm{P}_i), \mathbf{k}[\mathrm{T}]/(\mathrm{P}_j))$$

est déterminé par son image $\mathrm{P} \mod \mathrm{P}_j$ *où* P *vérifie*

$$(*) \qquad \mathrm{P}_i\mathrm{P} = 0 \mod \mathrm{P}_j$$

(propriété universelle du quotient). Si $i \leq j$*, on a* $\mathrm{P}_j|\mathrm{P}_i$*, et cette condition est automatiquement vérifiée de sorte que*

$$\mathrm{Hom}_{\mathbf{k}[\mathrm{T}]}(\mathbf{k}[\mathrm{T}]/(\mathrm{P}_i), \mathbf{k}[\mathrm{T}]/(\mathrm{P}_j)) \simeq \mathbf{k}[\mathrm{T}]/(\mathrm{P}_j) \text{ si } i \leq j$$

Si $i > j$*, on a* $\mathrm{P}_i|\mathrm{P}_j$ *de sorte que la condition* $(*)$ *s'écrit* $\mathrm{P} = 0 \mod \mathrm{P}_j/\mathrm{P}_i$ *de sorte que*

$$\mathrm{Hom}_{\mathbf{k}[\mathrm{T}]}(\mathbf{k}[\mathrm{T}]/(\mathrm{P}_i), \mathbf{k}[\mathrm{T}]/(\mathrm{P}_j)) \simeq \mathrm{P}_j/\mathrm{P}_i\mathbf{k}[\mathrm{T}]/(\mathrm{P}_j) \simeq \mathbf{k}[\mathrm{T}]/(\mathrm{P}_i) \text{ si } i > j$$

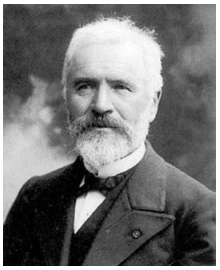
On a donc

$$\begin{aligned}
\dim_{\mathbf{k}}(\mathrm{Com}(a)) &= \sum_{i\leq j} \deg(\mathrm{P}_j) + \sum_{i>j} \deg(\mathrm{P}_i) \\
&= \sum_j j\deg(\mathrm{P}_j) + \sum_i (i-1)\deg(\mathrm{P}_i) \\
&= \sum (2i-1)\deg(\mathrm{P}_i)
\end{aligned}$$

On a alors $\dim \mathrm{Com}(a) - n = 2\sum_{i=1}^{n}(i-1)\deg(\mathrm{P}_i) \geq 0$*. De plus, l'égalité entraîne* $(i-1)\deg(\mathrm{P}_i) = 0$ *pour tout* i*, donc* $\deg(\mathrm{P}_i) = 0$ *si* $i > 1$ *de sorte que l'égalité équivaut à la cyclicité de* a*.* ■

Exercice(s) 4.11.0.2 (Bicommutant, difficile)**.** *Montrer que l'inclusion* $\mathbf{k}[a] \subset \mathrm{Com}(\mathrm{Com}(a))$ *est un égalité où* $\mathrm{Com}(\mathrm{Com}(a))$ *est l'ensemble des endomorphismes qui commutent avec tous les éléments de* $\mathrm{Com}(a)$*.*

4.12 Application : réduite de Jordan



Camille Jordan

Expliquons pourquoi la réduction de Frobenius entraîne immédiatement la réduction de Jordan des endomorphismes de polynôme caractéristique scindé. On conserve les notations précédentes (et on rappelle qu'une matrice de taille ≤ 0 est une matrice vide).

Soit donc $\mathrm{A} \in \mathrm{M}_n(\mathbf{k})$ et $\underline{\mathrm{P}}$ les invariants de similitude de A. On suppose χ_{A} scindé sur $\mathbf{k}$ et on note Λ l'ensemble de ses racines distinctes de sorte que

$$\chi_{\mathrm{A}}(\mathrm{T}) = \prod\nolimits_{\Lambda} (\mathrm{T} - \lambda)^{d_\lambda}.$$

Si on spécialise au cas $\chi_\mathrm{A} = \mathrm{T}^n$, on a $\mathrm{P}_i = \mathrm{T}^{d_i}$ avec $d_i \geq 0$ décroissante et $\sum d_i = d$.

Définition 4.12.0.1. *Une partition d'une entier $n \geq 0$ est une suite décroissante $\underline{d} = (d_i)_{1 \leq i \leq n}$ d'entiers ≥ 0 telle que $\sum d_i = n$.*

Comme chaque P_i divise χ_A, on a

(iv) $$\mathrm{P}_i = \prod_\Lambda (\mathrm{T} - \lambda)^{d_{\lambda,i}} \text{ où } \underline{d_\lambda} = (d_{\lambda,i})_i \text{ est une partition de } d_\lambda.$$

En appliquant le lemme chinois 1.2.0.1, on a

$$\mathrm{V_A} = \oplus_\Lambda \oplus_i \oplus_\Lambda \mathbf{k}[\mathrm{T}]/((\mathrm{T} - \lambda)^{d_{\lambda,i}}).$$

La formule $\mathrm{T}(\mathrm{T} - \lambda)^j = (\mathrm{T} - \lambda)^{j+1} + \lambda_j(\mathrm{T} - \lambda)^j$ assure que la matrice de la multiplication par T sur chacun des $\mathbf{k}[\mathrm{T}]/((\mathrm{T} - \lambda)^{d_{\lambda,i}}$ dans sa base $((\mathrm{T} - \lambda_j) \mod (\mathrm{T} - \lambda)^{d_{\lambda,i}})_{j < d_{\lambda,i}}$ est $\lambda + \mathrm{J}_{d_{\lambda,i}}$ où $\mathrm{J}_m = \mathrm{C}(\mathrm{T}^m)$ est le bloc de Jordan standard de taille m. Ainsi, on a

Théorème 4.12.0.2 (Réduction de Jordan)**.** *Sous les hypothèses et notations précédentes, on a :*

1. A *est semblable à une unique matrice diagonale* $\mathrm{diag}(\lambda + \mathrm{J}_{d_{i,\lambda}})$ *avec pour tout* λ *la suite* $(d_{i,\lambda})_i$ *partition de* d_λ.
2. *En particulier, si* $\chi_\mathrm{A} = \mathrm{T}^n$ *(ie* A *nilpotent), il existe une unique partition* $\underline{d} = (d_i)$ *de* n *vérifiant* A *semblable à la matrice diagonale par blocs* $\mathrm{J}_{\underline{d}} = \mathrm{diag}(\mathrm{J}_{d_n}, \cdots, \mathrm{J}_{d_1})$. *Les invariants de similitude de* A *sont* $\mathrm{T}^{d_n}, \mathrm{T}^{d_{n-1}}, \cdots, \mathrm{T}^{d_1}$.

Remarque(s) 4.12.0.3. *L'unicité découle du fait que la réduction de Jordan étant donnée, le calcul de ses invariants de similitude en découle. En effet si a forme de Jordan est composée de blocs du type* $\lambda \mathrm{Id}_r + \mathrm{J}_r$. *À chaque tel bloc est associé le polynôme* $(\mathrm{T} - \lambda)^r$. *Pour chaque valeur propre* λ, *on classe en ordre décroissant les blocs qui apparaissent, et on écrit en colonne les polynômes correspondants*

$$\begin{matrix} (\mathrm{T} - \lambda_1)^{d_{1,1}} & (\mathrm{T} - \lambda_2)^{d_{1,2}} & \cdots \\ (\mathrm{T} - \lambda_1)^{d_{2,1}} & (\mathrm{T} - \lambda_2)^{d_{2,2}} & \cdots \\ \vdots & \vdots & \end{matrix}$$

avec $d_{i+1,j} \leq d_{i,j}$. *On lit alors sur les lignes (en partant de la dernière) les facteurs invariants* $\mathrm{P}_1, \mathrm{P}_2$, *etc.*

Exercice(s) 4.12.0.4. *Soit* $\mathrm{M} \in \mathrm{M}_n(\mathbf{k})$ *une matrice nilpotente.*

1. *Montrer* $\mathrm{rg}(\mathrm{M}) = n-1$ *si et seulement si la réduite de Jordan est* J_n.

2. *Si* $\mathbf{k} = \mathbf{R}$*, montrer que l'ensemble des matrices nilpotentes de rang* $n-1$ *est le plus grand ouvert de l'ensemble des matrices nipotentes sur lequel la réduction de Jordan est continue (avec la topologie définie par une norme de* $\mathrm{M}_n(\mathbf{R})$*).*

3. *Montrer que* $\mathrm{rg}(\mathrm{M}) = n-2$ *si et seulement si* M *exactement deux blocs de Jordan* $\mathrm{J}_p, \mathrm{J}_{n-p}$ *où* p *est l'indice de nilpotence de* M*. Montrer que* $p \geq n/2$.

4. *Soit* $p \geq n/2$*, un entier* $q = n-p$ *et posons pour* $t \in \mathbf{k}$*, soit* $\mathrm{M}_t = \mathrm{diag}(\mathrm{J}_p, \mathrm{J}_q) + t\mathrm{E}_{p+q,p}$ *(on ajoute* t *en bas de la* p*-ième colonne). Calculer l'indice de nilpotence de* M_t *en fonction de* t*. En déduire que la réduite de Jordan de* M_t *est* $\mathrm{diag}(\mathrm{J}_{p+1}, \mathrm{J}_{q-1})$ *si* $t \neq 0$ *et* $\mathrm{diag}(\mathrm{J}_p, \mathrm{J}_q)$ *sinon.*

5. *Supposons* $\mathbf{k} = \mathbf{R}$*. Quelle est l'ensemble de continuité de l'application de réduction de Jordan restreint au sous-ensemble des matrices nilpotentes de rang* $n-2$ *(avec la topologie définie par une norme de* $\mathrm{M}_n(\mathbf{R})$*) ?*

4.12.1 Exemples

(1) Les diviseurs élémentaires de la réduite de Jordan

$$\begin{pmatrix} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{pmatrix}$$

(où $\lambda \neq \mu$), sont

$$\begin{array}{ll} (\mathrm{T}-\lambda)^2 & (\mathrm{T}-\mu) \\ (\mathrm{T}-\lambda)^2 & \\ (\mathrm{T}-\lambda). & \end{array}$$

Les invariants de similitude sont donc

$$(\mathrm{T}-\lambda)\ ,\quad (\mathrm{T}-\lambda)^2\ ,\quad (\mathrm{T}-\lambda)^2(\mathrm{T}-\mu).$$

(2) Si $\mathrm{M} = \begin{pmatrix} 0 & 4 & 2 \\ -1 & -4 & -1 \\ 0 & 0 & -2 \end{pmatrix}$, on a

$$\mathrm{TI}-\mathrm{M} = \begin{pmatrix} \mathrm{T} & -4 & -2 \\ 1 & \mathrm{T}+4 & 1 \\ 0 & 0 & \mathrm{T}+2 \end{pmatrix}.$$

Faisons des opérations élémentaires selon l'algorithme -ou plutôt son ébauche- décrit dans la démonstra-

tion de la proposition 3.3.1.1 :

$$\begin{pmatrix} \mathrm{T} & -4 & -2 \\ 1 & \mathrm{T}+4 & 1 \\ 0 & 0 & \mathrm{T}+2 \end{pmatrix} \xrightarrow{\mathrm{L}_1 \leftrightarrow \mathrm{L}_2} \begin{pmatrix} 1 & \mathrm{T}+4 & 1 \\ \mathrm{T} & -4 & -2 \\ 0 & 0 & \mathrm{T}+2 \end{pmatrix}$$

$$\xrightarrow{\mathrm{L}_2 \to \mathrm{L}_2 - \mathrm{T}\mathrm{L}_1} \begin{pmatrix} 1 & \mathrm{T}+4 & 1 \\ 0 & -4-\mathrm{T}(\mathrm{T}+4) & -2-\mathrm{T} \\ 0 & 0 & \mathrm{T}+2 \end{pmatrix} \overset{\substack{\mathrm{C}_2 \to \mathrm{C}_2 - (\mathrm{T}+4)\mathrm{C}_1 \\ \mathrm{C}_3 \to \mathrm{C}_3 - \mathrm{C}_1}}{\longrightarrow} \begin{pmatrix} 1 & 0 & 0 \\ 0 & (\mathrm{T}+2)^2 & -2-\mathrm{T} \\ 0 & 0 & \mathrm{T}+2 \end{pmatrix}$$

$$\xrightarrow{\mathrm{L}_2 \to \mathrm{L}_2 + \mathrm{L}_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & (\mathrm{T}+2)^2 & 0 \\ 0 & 0 & \mathrm{T}+2 \end{pmatrix} \overset{\substack{\mathrm{C}_1 \leftrightarrow \mathrm{C}_2 \\ \mathrm{L}_1 \leftrightarrow \mathrm{L}_2}}{\longrightarrow} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathrm{T}+2 & 0 \\ 0 & 0 & (\mathrm{T}+2)^2 \end{pmatrix}.$$

Les invariants de similitude sont donc $\mathrm{T}+2$ et $(\mathrm{T}+2)^2$ et la réduite de Jordan est $\begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$.

Un endomorphisme de matrice M n'est pas cyclique.

(3) Si $\mathrm{M} = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{pmatrix}$, on obtient comme réduite pour $\mathrm{TI} - \mathrm{M}$ la matrice

$$\begin{pmatrix} (\mathrm{T}-1)^2 & 0 & 0 & 0 \\ 0 & (\mathrm{T}-1)^2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Les facteurs invariants sont $(\mathrm{T}-1)^2$ et $(\mathrm{T}-1)^2$ et la réduite de Jordan est $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. Un endomorphisme de matrice M n'est pas cyclique.

(4) Un endomorphisme est cyclique si et seulement si, pour chaque valeur propre, il n'y a qu'un seul bloc de Jordan.

4.12.2 Complément sur les matrices nilpotentes

Soit A une matrice nilpotente et $\underline{d}$ la partition associée (4.12.0.2). Comme le bloc de Jordan J_p est la matrice dans la base canonique de la multiplication par T de $\mathbf{k}[\mathrm{T}]/(\mathrm{T}^p)$, l'image de J_p^i s'identifie à $\mathrm{T}^i\mathbf{k}[\mathrm{T}]/(\mathrm{T}^p)/ \simeq \mathbf{k}[\mathrm{T}]/(\mathrm{T}^{p-i})$. On déduit l'égalité $\mathrm{rg}(\mathrm{J}_p^i) = (p-i)_+$ et plus généralement

$$\mathrm{rg}(\mathrm{A}^i) = \sum_j (d_j - i)_+ \tag{v}$$

On pose

$$d_i^* = \dim(\mathrm{Im}(\mathrm{A}^{i-1})/\mathrm{Im}(\mathrm{A}^i)) = \mathrm{rg}(\mathrm{A}^{i-1}) - \mathrm{rg}(\mathrm{A}^i),\ i = 1, \cdots, n$$

(avec $\mathrm{A}^0 = \mathrm{Id}$) de sorte qu'on a $\sum d_i^* = n$ et $d_i^* \geq 0$. De plus, la multiplication par A induit une surjection $\mathrm{Im}(\mathrm{A}^{i-1})/\mathrm{Im}(\mathrm{A}^i) \to \mathrm{Im}(\mathrm{A}^i)/\mathrm{Im}(\mathrm{A}^{i+1})$ de sorte que d_i^* décroît. On a par construction $\mathrm{rg}(\mathrm{A}^i) = \sum_{j>i} d_j^*$.

(vi)
$$n - \mathrm{rg}(\mathrm{A}^i) = \sum_{j \leq i} d_j^*$$

Définition 4.12.2.1. *La partition $\underline{d}^* = (d_i^*)$ est dite partition duale de $\underline{d}$.*

On peut simplifier l'écriture de la dualité des partitions, qui de plus est involutive !

Lemme 4.12.2.2. *Avec les notations précédentes on a $d_i^* = \mathrm{Card}\{j | d_j \geq i\}$ et $\underline{d}^{**} = \underline{d}$.*

Démonstration. *On écrit d'abord*

$$\begin{aligned} d_i^* &= \sum_j (d_j - i + 1)_+ - (d_j - i)_+ \\ &= \sum_{j | d_j \geq i} (d_j - i + 1)_+ - (d_j - i)_+ \\ &= \sum_{j | d_j \geq i} 1 \\ &= \mathrm{Card}\{j | d_j \geq i\} \end{aligned}$$

d'où la première égalité. Pour la seconde, on écrit

$$\begin{aligned} d_i^{**} &= \mathrm{Card}\{j | d_j^* \geq i\} \\ &= \mathrm{Card}\{j | \mathrm{Card}\{k | d_k \geq j\} \geq i\} \end{aligned}$$

Mais $\mathrm{Card}\{k | d_k \geq j\} \geq i$ *si et seulement si* $d_i \geq j$. *En effet, si il existe un ensemble ordonné d'indices* K *de cardinal* $\geq i$ *tel que* $k \in \mathrm{K} \Rightarrow d_k \geq j$, *alors son* i*-ième élément* k *est* $\geq i$ *et* $d_i \geq d_k \geq i$ *par décroissance de* $\underline{d}$. *Inversement, si* $d_i \geq j$, *on a* $d_k \geq j$ *pour* $k \leq i$ *toujours par décroissance et donc* $\mathrm{Card}\{i | d_i \geq j\} \geq i$. *On a donc* $\mathrm{Card}\{j | \mathrm{Card}\{k | d_k \geq j\} \geq i\} = \mathrm{Card}\{j | d_i \geq j\} = d_i$. ∎

Remarque(s) 4.12.2.3. *On présente usuellement ceci à l'aide de tableaux de Young avec des preuves, plus ou moins convaincantes, d'ordre graphique. Il est inutile pour nous d'introduire ces notations supplémentaires.*

4.13 Appendices

4.13.1 Algorithme pour passer de l'équivalence à la similitude

On sait donc que si $\mathrm{TId}-\mathrm{A}$ et $\mathrm{TId}-\mathrm{B}$ sont équivalentes, ie s'il existe $\mathrm{P(T)}, \mathrm{Q(T)}$ matrices polynômiales et inversibles telles que

$$\mathrm{P(T)(TId}-\mathrm{A)} = \mathrm{(TId}-\mathrm{B)Q(T)}^{-1},$$

alors il existe $\mathrm{P} \in \mathrm{GL}_n(\mathbf{k})$ telles que $\mathrm{B} = \mathrm{PAP}^{-1}$.

Proposition 4.13.1.1 (Merci à O. Debarre)**.** *Il existe un algorithme permettant de calculer un tel* P.

Démonstration. *On peut (4.2.0.1) effectuer les divisions*

$$\begin{aligned} \mathrm{P(T)} &= \mathrm{(TId}-\mathrm{B)P_1(T)} + \mathrm{P_0}, \\ \mathrm{Q(T)}^{-1} &= \tilde{\mathrm{Q}}_1\mathrm{(T)(TId}-\mathrm{A)} + \tilde{\mathrm{Q}}_0, \end{aligned}$$

avec $\mathrm{P_0}$ *et* $\tilde{\mathrm{Q}}_0$ *dans* $\mathrm{M}_n(\mathbf{k})$ *(la difficulté provient du fait qu'on n'est pas dans un anneau commutatif). On obtient en remplaçant*

$$(\mathrm{(TId}-\mathrm{B)P_1(T)} + \mathrm{P_0})\mathrm{(TId}-\mathrm{A)} = \mathrm{(TId}-\mathrm{B)}(\tilde{\mathrm{Q}}_1\mathrm{(T)(TId}-\mathrm{A)} + \tilde{\mathrm{Q}}_0)$$

ou encore

$$\mathrm{(TId}-\mathrm{B)}(\mathrm{P_1(T)} - \tilde{\mathrm{Q}}_1\mathrm{(T)})\mathrm{(TId}-\mathrm{A)} = \mathrm{(TId}-\mathrm{B)}\tilde{\mathrm{Q}}_0 - \mathrm{P_0(TId}-\mathrm{A}).$$

Le membre de gauche est donc de degré au plus 1 *en* T*, ce qui n'est possible que si* $\mathrm{P_1(T)} = \tilde{\mathrm{Q}}_1\mathrm{(T)}$*. On a donc* $\mathrm{(TId}-\mathrm{B)}\tilde{\mathrm{Q}}_0 = \mathrm{P_0(TId}-\mathrm{A})$ *(raisonner par l'absurde et regarder le terme de plus haut degré). L'égalité des coefficients de* T *donne* $\tilde{\mathrm{Q}}_0 = \mathrm{P_0}$*, celle des coefficients constants donne* $\mathrm{B}\tilde{\mathrm{Q}}_0 = \mathrm{P_0A}$*. Il reste à montrer que* $\tilde{\mathrm{Q}}_0$ *est inversible. On refait une division*

$$\mathrm{Q(T)} = \mathrm{Q_1(T)(TId}-\mathrm{B)} + \mathrm{Q_0}$$

et on écrit

$$\begin{aligned} \mathrm{Id} &= \mathrm{Q(T)}^{-1}\mathrm{Q(T)} \\ &= (\tilde{\mathrm{Q}}_1\mathrm{(T)(TId}-\mathrm{A)} + \tilde{\mathrm{Q}}_0)\mathrm{Q(T)} \\ &= \tilde{\mathrm{Q}}_1\mathrm{(T)(TId}-\mathrm{A)Q(T)} + \tilde{\mathrm{Q}}_0\mathrm{Q(T)} \\ &= \tilde{\mathrm{Q}}_1\mathrm{(T)P(T)}^{-1}\mathrm{(TId}-\mathrm{B)} + \tilde{\mathrm{Q}}_0(\mathrm{Q_1(T)(TId}-\mathrm{B)} + \mathrm{Q_0}) \\ &= \big(\tilde{\mathrm{Q}}_1\mathrm{(T)P(T)}^{-1} + \tilde{\mathrm{Q}}_0\mathrm{Q_1(T)}\big)\mathrm{(TId}-\mathrm{B)} + \tilde{\mathrm{Q}}_0\mathrm{Q_0}. \end{aligned}$$

De nouveau, comme $\tilde{\mathrm{Q}}_0\mathrm{Q}_0$ *est constant, le facteur de* $\mathrm{TId}-\mathrm{B}$ *est nul et* $\tilde{\mathrm{Q}}_0\mathrm{Q}_0 = \mathrm{Id}$*, d'où la conclusion.*

■

4.13.2 Réduction de Jordan par dualité des nilpotents sans les modules

On va donner une preuve classique du théorème de réduction de Jordan 4.12.0.2 par récurrence sur la dimension qui utilise des méthodes standards de dualité (6) en algèbre linéaire dans le cas nilpotent, le cas général se prouvant par réduction aux espaces caractéristiques (7) grâce au lemme des noyaux 2.6.0.2
.

On amorce la récurrence en dimension 0. Soit donc a un endomorphisme nilpotent sur V de dimension $n \geq 1$ et $d = \deg \mu_a \geq 1$ son indice de nilpotence. Comme $a^{d-1} \neq 0$, on peut choisir v tel que $a^{d-1}(v) \neq 0$. Comme le vecteur est non nul, on choisit de plus $\varphi \in \mathrm{V}^*$ tel que $\langle \varphi, a^{d-1}(v) \rangle \neq 0$.

Par construction, les espaces $\mathrm{W} = \mathbf{k}[a].v$ et $\mathrm{W}_* = \mathbf{k}[{}^t a].\varphi$ sont stables par a et ${}^t a$ respectivement engendrés par $a^i(v), i \leq d-1$ et ${}^t a^i(\varphi), i \leq d-1$ respectivement. Leur dimension est donc $\leq d$. On vérifie facilement que ces familles génératrices sont libres, de sorte que qu'il sont de dimension d. En particulier, W est cyclique et la matrice de (la restriction à W de) a dans la base précédente est le bloc de Jordan standard J_d. Comme W^* est stable par ${}^t a$, son orthogonal $\mathrm{W}_*^\perp \subset \mathrm{V}^{**} = \mathrm{V}$ est stable par ${}^{tt} a = a$, donc est de dimension $n - d < n$. On peut donc appliquer l'hypothèse de récurrence à la restriction de a à W'. Reste à vérifier que la somme $\mathrm{W} + \mathrm{W}_*^\perp$ est directe donc que $\mathrm{W} \cap \mathrm{W}_*^\perp = \{0\}$ puisque les dimensions sont complémentaires. Soit donc $\sum_{i<d} \lambda_i a^i(v)$ dans l'intersection. Si un des λ_i est non nul, choisissons j le plus petit indice des coefficients non nuls et appliquons ${}^t a^{d-1-j} \varphi \in \mathrm{W}_*$. On a donc

$$0 = \langle {}^t a^{d-1-j} \varphi, \sum_{i<d} \lambda_i a^i(v) \rangle = \langle \varphi, \sum_{i \geq j} \lambda_i a^{d-1-j+i}(v) \rangle = \lambda_j \langle \varphi, a^{d-1}(v) \rangle \neq 0,$$

une contradiction. ■

4.13.3 Décomposition de Frobenius sans les modules

On va donner une preuve du théorème de Frobenius 4.9.0.2 par récurrence sur la dimension qui utilise des méthodes standards d'algèbre linéaire (voir aussi 4.9.1). C'est en fait juste une adaptation de la preuve précédente 4.13.2 du théorème de réduction de Jordan. Elle ne sera en revanche pas algorithmique [3] à cause de l'utilisation du lemme des noyaux pour le lemme clef suivant.

Lemme 4.13.3.1. *Tout endomorphisme en dimension finie admet un vecteur cyclique.*

Démonstration. *Supposons que μ est la puissance P^d d'un polynôme irréductible. Chaque minimal $\mu_{a,v}$ est de la forme P^{d_v} avec $d_v \leq d$. Si pour tout v, $d_v \leq d-1$, on aurait $\mathrm{P}^{d-1}(a)(v) = 0$, une contradiction avec $\mu_a = \mathrm{P}^d$.*

Dans le cas général, décomposons $\mu_a = \prod \mathrm{P}_i^{d_i}$ en puissances de facteurs irréductibles deux à deux premiers entre eux. Sur chaque noyau $\mathrm{K}_i = \mathrm{Ker}(\mathrm{P}^{d_i}(a))$, le minimal de $a_{|\mathrm{K}_i}$ est P^{d_i}, de sorte qu'il existe grâce à ce qui précède un vecteur cyclique v_i pour $a_{|\mathrm{K}_i}$. Il reste à appliquer le lemme des noyaux 2.6.0.2 pour se convaincre que la somme des v_i est un vecteur cyclique pour a. ■

3. Pour le lecteur intéressé, voir `http://www.lix.polytechnique.fr/~augot/CRAS_94.pdf` pour un algorithme. Comparer avec l'exrecice 4.13.3.2 *infra*.

Exercice(s) 4.13.3.2. *En utilisant le théorème 4.5.0.2 et l'algorithme correspondant, écrire un algorithme permettant de trouver un vecteur cyclique. L'implémenter avec SAGEMath.*

Pour l'existence de la décomposition de Frobenius, on va donc adapter la démonstration de la section précédente. Supposons donc que a est un endomorphisme arbitraire de V de dimension finie de polynôme minimal μ_a de degré d et choisissons v un vecteur cyclique pour a. Le sous-espace $\mathrm{W} = \mathbf{k}[a].v$ est cyclique, stable de dimension d de sorte que le minimal de $a_{|\mathrm{W}}$ est μ_a. Il en est de même de sa transposée ${}^t a_{|\mathrm{W}} \in \mathrm{End}_{\mathbf{k}}(\mathrm{W}^*)$; soit $\tilde{\varphi} \in \mathrm{W}^*$ un vecteur cyclique pour ${}^t a_{|\mathrm{W}} \in \mathrm{End}_{\mathbf{k}}(\mathrm{W}^*)$. Comme le degré de son minimal est d qui est aussi la dimension de W^*, on a $\mathbf{k}[{}^t a_{|\mathrm{W}}].\tilde{\varphi} = \mathrm{W}^*$.
Soit $\varphi \in \mathrm{V}^*$ un prolongement linéaire quelconque de $\tilde{\varphi} \in \mathrm{W}^* = \mathrm{Hom}_{\mathbf{k}}(\mathrm{W}, \mathbf{k})$ à V et posons $\mathrm{W}_* = k[{}^t a].\varphi \subset \mathrm{V}^*$. Comme tout sous-espace monogène, W_* est de dimension $\leq d$. Mais comme la restriction des formes à W envoie surjectivement $\mathrm{W}_* = k[{}^t a].\varphi \subset \mathrm{V}^*$ sur $= k[{}^t a_{|\mathrm{W}}].\tilde{\varphi} = \mathrm{W}^*$ puisque qu'elle envoie φ sur $\varphi_{|\mathrm{W}} = \tilde{\varphi}$, cette restriction est un isomorphisme $\mathrm{W}_* \simeq \mathrm{W}^*$. En particulier φ est cyclique pour ${}^t a$.
Comme plus haut, W et $\mathrm{W}_*^{\perp}$ sont stables de dimension complémentaires avec W cyclique. Par ailleurs, le minimal de $a_{|\mathrm{W}_*^{\perp}}$ divise μ_a. Reste à prouver que la somme de W et $\mathrm{W}_*^{\perp}$ est directe pour conclure par récurrence. Or, si $w \in \mathrm{W}$ est orthogonal à W_*, on a pour tout $\psi \in \mathrm{W}_*$ la nullité de $\langle \psi, w \rangle$ qui n'est autre que $\langle \psi_{|\mathrm{W}}, w \rangle$. Comme la restriction $\mathrm{W}_* \to \mathrm{W}^*$ est un isomorphisme, on déduit que w est orthogonal à toute forme de son dual, donc est nul. ■
Pour l'unicité, supposons, avec des notations évidentes, qu'on ait deux décompositions de Frobenius

$$\mathrm{V} = \oplus \mathrm{V}_i(\mathrm{P}_i) = \oplus \mathrm{W}_i(\mathrm{Q}_i).$$

Montons par récurrence (forte) que $\mathrm{P}_i = \mathrm{Q}_i$ pour tout i. On a déjà nécessairement $\mathrm{P}_1 = \mu_a = \mathrm{Q}_1$. Supposons maintenant $i > 1$ et $\mathrm{P}_1 = \mathrm{Q}_1, \cdots, \mathrm{P}_{i-1} = \mathrm{Q}_{i-1}$.
On a d'une part

$$\mathrm{P}_i.\mathrm{V} = \oplus_{j<i} \mathrm{P}_i.\mathrm{V}_j$$

car P_i divise $\mathrm{P}_j = \mu_{a|\mathrm{V}_j} | \mathrm{P}_i$ si $j \geq i$. On a d'autre part

$$\mathrm{P}_i.\mathrm{V} = \oplus_j \mathrm{P}_i.\mathrm{W}_j$$

et

$$\dim \mathrm{P}_i.\mathrm{V}_j = \dim \mathrm{P}_i.\mathrm{W}_j \text{ si } j < i$$

car, dans des bases convenables, les matrices des restrictions de a à V_j et W_j les mêmes matrices compagnons associées à $\mathrm{P}_j = \mathrm{Q}_j$ si $j < i$, et donc il en est de même de celles de $\mathrm{P}_i(a)$. En calculant la dimension de $\mathrm{P} - i.\mathrm{V}$ de deux manières, il vient $\dim \mathrm{P}_i \mathrm{W}_i = 0$ et donc $\mathrm{Q}_i | \mathrm{P}_i$ puisque Q_i est le minimal de a sur W_i. Par symétrie des rôles, on a $\mathrm{P}_i = \mathrm{Q}_i$. ■

4.13.4 Implantations en Sage

a) **Pivot de Gauss polynomial** SageMath version 9.7, Release Date : 2022-09-19, Using Python 3.10.5.

Remarque(s) 4.13.4.1. *Le programme est rapidement mis en défaut quand on remplace par exemple* **Q** *par* **R**. *La raison est l'instabilité numérique structurelle du pivot de Gauss. Lorsque qu'on fait du pivot sur des matrices scalaires, on la compense(partiellement) en prenant à chaque fois le plus grand pivot en valeur absolue pour tenter de ne pas faire exploser les coefficients et de dépasser les capacités de la machine, mais cela devient impossible quand on fait du pivot polynomial. C'est un joli sujet de réflexion de voir comment on pourrait outrepasser cette difficulté. On a choisi de reprogrammer un certain nombre de fonctions natives sur Sage comme les opérations élémentaires par exemple pour bien illustrer l'algorithme.*

```
1
2
3  #diviseurs élémentaires dans Q[T]
4  import time
5  R.<t> = PolynomialRing(QQ)
6
7  n = 5
8  #a = random_matrix(R,n,n)
9  a = random_matrix(QQ,n,n)
10 a = t*identity_matrix(n)-a
11
12 tmps1 = time.time()
13
14
15 def zero(R,m,n):
16 return matrix(R,m,n)
17
18 #pivot deg min matrice non nulle
19
20 def pivot(a):
21 n = a.nrows()
22 m = a.ncols()
23 L = [(i,j) for i in range(n) for j in range(m) if a[i,j] !=0]
24 mind = min([a[i,j].degree() for (i,j) in L])
25 nn = next((i,j) for (i,j) in L if a[i,j].degree() == mind)
26 ii = nn[0]
27 jj = nn[1]
28 return [ii,jj]
29
30 def pivotage(a):
31 i = pivot(a)[0]
32 j = pivot(a)[1]
33 a.swap_rows(0,i)
34 a.swap_columns(0,j)
35 return a
36
37 def zerotage(a):
```

```
38  a=pivotage(a)
39  m = a.nrows()
40  n = a.ncols()
41  Lc0 = [i for i in range(m-1) if a[i+1,0] !=0]
42  Lr0 = [j for j in range(n-1) if a[0,j+1] !=0]
43  if Lr0+Lc0 == []:
44  return a
45  for i in range(m-1):
46  a.add_multiple_of_row(i+1,0,-a[i+1,0]//a[0,0])
47  for j in range(n-1):
48  a.add_multiple_of_column(j+1,0,-a[0,j+1]//a[0,0])
49  return zerotage(a)
50  #divelem d'une diagonale >0
51
52  def divdiag(d):
53  if len(d)<2:
54  return d
55  d.sort()
56  a = d[0]
57  d.remove(a)
58  maxi = [x%a for x in d if x%a !=0]
59  #si le pivot d[0] divise tout le monde, youpi
60  if  maxi == []:
61  d = [x//a for x in d]
62  d.sort()
63  d = divdiag(d)
64  return [a]+[x*a for x in d]
65  #sinon
66  ii = next(i for i in range(len(d)) if d[i]%a !=0)
67  b = d[ii]
68  dd = [a*d[ii]//gcd(a,d[ii]),gcd(a,d[ii])]
69  d.remove(d[ii])
70  d = dd+d
71  d.sort()
72  return divdiag(d)
73
74  def divelem(a):
75  m = a.nrows()
76  n = a.ncols()
77  d=[]
78  #matrices de taille <=(1,1)
79  if m*n == 0 or a == zero(ZZ,m,n):
80  return d
81  a = zerotage(a)
82  d=d+[a[0,0]]
83  #matrices lignes ou colonnes
84  if (m-1)*(n-1) == 0:
85  return d
86  #matrices non ligne ou colonne
```

```
d = d+divelem(a[1:m,1:n])
#par récursivité, d défini et a équivalente à diag(d)
return divdiag(d)

def divelem_norm(a):
return [p/p.leading_coefficient() for p in divelem(a) if p !=0]

d = divelem_norm(a)
tmps2 = time.time()
print((tmps2-tmps1)*1000,'ms')
```

b) **Jordan-Chevalley**

```

#Bravo à Antoine Castellani pour ce code SAGE calculant la réduction de Dunford
l=5
d=4
k.<u> = GF(l^d)
#k=QQ Décommenter et commenter supra pour changer de corps
#il faut adapter si le corps n'est pas parfait : bon exrecice

R.<x>=PolynomialRing(k) # où A est à remplacer par QQ ou k.<u> = GF(p^n)

def enlever_facteurs_carres(P):#version pour éviter la commande factor() de SAGE
p=k.characteristic()
if p==0:
return P/gcd(P,diff(P))
else:
u=gcd(P,diff(P))
v=P/u # Renvoie les termes sans carré lorsque la puissance n'est pas divisée par p
if v==P:
return v
else:
w=u/gcd(u,v^(P.degree())) # Renvoie les autres termes, ie ceux dont la puissance est
    divisée par p. On peut donc prendre la racine p-ième
racinew=(w.numerator()).nth_root(p) # Renvoie la racine p-ième de w et on itère
return v*enlever_facteurs_carres(racinew)

def Hensel(P,x_0,n):
# P polynome
# x_0 du lemme
# n entier
solutions=[x_0]
N = valuation(n,2)+2
for j in range(N):
i=P(solutions[j])
r=(diff(P)(solutions[j])).inverse()
solutions.append(solutions[j]-r*i)
```

```
35  return solutions[N]
36
37  def Jordan(a):
38  # a matrice carrée
39  n=a.nrows()
40  pi=enlever_facteurs_carres(minpoly(a))
41  Delta=Hensel(pi,a,n)
42  return "D=",Delta,"N=",a-Delta, "test=",a*Delta-Delta*a
43
44  #test
45  #a = matrix(k,[[1,k.random_element(),1],[0,1,1],[0,0,2]])
46  #Jordan(a)
47
48
```

4.14 Exercices supplémentaires

Exercice(s) 4.14.0.1. *Montrer que $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ est trigonalisable si et seulement si χ_a est scindé. Montrer qu'une famille d'endomorphismes qui commutent et qui sont trigonalisables admet une base de trigonalisation commune (S'inspirer de la preuve de 4.7.0.3).*

Exercice(s) 4.14.0.2. *Soit G un sous-groupe commutatif de $\mathrm{GL}_n(\mathbf{R})$ dont tous les éléments sont de carré Id.*

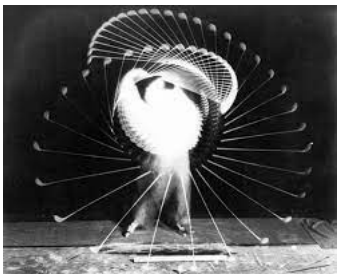
1. *Montrer que G est fini de cardinal $\leq n$ (utiliser 4.7.0.3).*
2. *Montrer que si $\mathrm{GL}_n(\mathbf{R})$ et $\mathrm{GL}_m(\mathbf{R})$ sont isomorphes, alors $n = m$.*
3. *Pouvez-vous généraliser à d'autres corps ?*
4. *Que se passe-t-il si on ne suppose plus G commutatif ?*

Exercice(s) 4.14.0.3. *On adopte les notations et résultats de 2.7.0.13.*

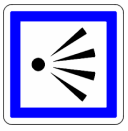
1. *Soit $\mathrm{P} \in \mathbf{C}[\mathrm{T}]$. Montrer que les racines de P sont simples si et seulement si le discriminant de P défini par $\mathrm{Res}(\mathrm{P}, \mathrm{P}')$ est non nul.*
2. *En considérant le discriminant du polynôme caractéristique d'une matrice carrée complexe, montrer que l'ensemble des matrices $\mathrm{M}_n(\mathbf{C})$ à valeurs propres distinctes est dense et que son complémentaire est de mesure de Lebesgue nulle.*

Chapitre 5

Semi-simplicité dans $\mathrm{M}_n(\mathbf{k})$



5.1 Point de vue



On montre de manière algorithmique comment l'étude des classes de similitude se ramène aux classes de matrice de $\mathrm{M}_n(\mathbf{k})$ nilpotentes (4.12.0.2) et à celle des matrices diagonalisables sur un corps algébriquement clos Ω contenant $\mathbf{k}$, les matrices (absolument) semi-simples.

La décomposition de Jordan-Chevalley [1] (sur les corps de caractéristique nulle et plus généralement parfaits est la clef de voute de la théorie des groupes algébriques et des algèbres de Lie. Dans le cas complexe, on précisera aussi en quoi cette décomposition est discontinue en général, continue si le polynôme caractéristique est fixé.

Dans ce chapitre, sauf mention expresse du contraire,
$\mathbf{k}$ désigne un corps parfait (5.2.3.1) et Ω un corps algébriquement clos qui le contient.

5.2 Semi-simplicité

La semi-simplicité est la bonne généralisation de la diagonalisabilité dans le cas des corps parfaits comme on le verra. Commençons par quelques généralités formelles.

1. On trouve étrangement la dénomination décomposition de Dundford pour décomposition de Jordan-Chevalley dans la littérature française.

Ryoan-ji, Kyoto

5.2.1 Modules semi-simples généraux

Dans ce paragraphe, R désigne un anneau commutatif unitaire arbitraire.

Définition 5.2.1.1. *Un* R*-module est dit*

- *semi-simple si tout sous-module a un supplémentaire ;*
- *simple s'il est non nul et n'a pas de sous-module non trivial.*

Un endomorphisme $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ *est dit semi-simple si le* $\mathbf{k}[\mathrm{T}]$*-module* V_a *l'est.*

Exercice(s) 5.2.1.2. *Montrer les points suivants.*

1. *Un espace vectoriel est semi-simple, et est simple si et seulement si il est de dimension* 1*.*
2. *Il existe* $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ *tel que* V_a *n'est pas semi-simple -prendre* a *nilpotent en dimension* 2 *par exemple- (cf. 5.2.4.1).*
3. *Le* $\mathbf{Z}$*-module (ie groupe abélien)* $\mathbf{Z}/4^2\mathbf{Z}$ *n'est pas semi-simple (cf. 5.2.1.3 infra).*
4. *Un anneau principal qui n'est pas un corps n'est jamais semi-simple en tant que module sur lui même.*

Les observations suivantes sont élémentaires mais très utiles.

Proposition 5.2.1.3. *Soit* N *un sous-module d'un module semi-simple* M*.*

- M *est isomorphe à* $\mathrm{N} \oplus \mathrm{M/N}$*.*
- *Tous sous-module est isomorphe à un module quotient et tout module quotient est isomorphe à un sous-module.*
- N *et* M/N *sont semi-simples.*

Démonstration. *Soit* S *un supplémentaire de* N *dans* M*. La surjection canonique* $f : \mathrm{M} \to \mathrm{M/N}$ *définit par restriction un isomorphisme* $\mathrm{S} \xrightarrow{\sim} \mathrm{M/N}$ *d'où le premier point. Mais alors, la première projection* $\mathrm{M} = \mathrm{N} \oplus \mathrm{M/N}$ *est surjective et identifie* N *au quotient* $\mathrm{Coker}(p)$*. De même pour les quotients avec*

l'inclusion $\mathrm{M/N} \hookrightarrow \mathrm{M} = \mathrm{N} \oplus \mathrm{M/N}$*, d'où le second point. Pour le dernier point, si* M' *est sous-module de* M/N*, on choisit un supplémentaire* S *du sous-module* $f^{-1}(\mathrm{M}')$ *et on vérifie que* $f(\mathrm{S})$ *est un supplémentaire de* M' *dans* M/N *de sorte que* M/N *est semi-simple. Mais comme* N *s'identifie à un quotient de* M*, il en est de même de* N. ■

Exercice(s) 5.2.1.4. *Montrer que* M *est semi-simple si et seulement si toute suite exacte courte est scindée (cf. 2.7.0.3).*

5.2.2 Modules semi-simples sur R principal

Si m est un élément de torsion, son annulateur $\mathrm{Ann_A}(m)$ a un générateur non nul, bien défini à inversible près : « son » minimal μ_m.

Proposition 5.2.2.1. *Soit* M *un module sur* R *principal qui n'est pas un corps.*

1. *Il existe un élément irréductible* p *de* R.
2. $\mathrm{R}/(p^2)$ *et donc* R *n'est pas semi-simple.*
3. M *est semi-simple si et seulement* M *est de torsion et si le minimal de tout élément est sans facteur irréductible carré.*

Démonstration. *Comme* R *n'est pas un corps,* R *a un élément non nul non inversible dont un quelconque de ses facteurs irréductibles répond à (1).*

Si $\mathrm{R}/(p^2)$ *était semi-simple, la suite exacte* $0 \to \mathrm{R}/(p) \xrightarrow{p} \mathrm{R}/(p^2) \to \mathrm{R}/(p) \to 0$ *serait scindée puisque* $p\mathrm{R}/(p)$ *aurait un supplémentaire dans* $\mathrm{R}/(p^2)$ *et donc* $\mathrm{R}/(p^2) \simeq \mathrm{R}/(p) \oplus \mathrm{R}/(p)$ *(2.7.0.3). Mais ceci entraînerait que* p *annule* $\mathrm{R}/(p^2)$*, ce qui n'est pas d'où (2). Passons à (3).*

$\Rightarrow$ *Si* $m \in \mathrm{M}$ *a un annulateur trivial,* $\mathrm{A} \xrightarrow{m} \mathrm{A}$ *est injective de sorte que* A *est un sous-module de* M*et donc devrait être semi-simple, ce qui n'est pas (5.2.1.2)). Donc, tout élément est de torsion. Soit alors* m *dont le minimal* μ_m *est divisible par* p^2 *avec* p *un irréductible de sorte que* $\mathrm{R}/(\mu_m)$ *est un sous-module de* M*. Supposons par l'absurde* M *semi-simple. Alors,* $\mathrm{R}/(\mu_m)$ *est aussi un quotient de* M *(5.2.1.3) et donc de même pour* $\mathrm{R}/(p^2)$ *(en tant que quotient de* $\mathrm{R}/(\mu_m)$ *donc de* M*) qui donc serait semi-simple, ce qui n'est pas d'après (2).*

$\Leftarrow$ *Supposons que le minimal de tout élément soit sans facteur irréductible carré et soit* N *sous-module de* M*. Alors, pour tout* p *irréductible, on a* $\mathrm{M}[p] = \bigcup_{n\geq 1} \mathrm{Ann_M}(p^n) = \mathrm{Ann_M}(p)$ *(2.6.0.1). Comme* M *est de torsion, le lemme chinois (2.6.0.1) assure* $\mathrm{M} = \oplus_p \mathrm{Ann_M}(p)$ *pour* p *décrivant les irréductibles à unité près (exercice). Mais la structure de* R*-module de* $\mathrm{Ann_M}(p)$ *se factorise à travers* $\mathrm{R} \to \mathrm{R}/p\mathrm{R} = \mathbf{k}(p)$ *qui est un corps car* R *est principal :* $\mathrm{Ann_M}(p)$ *est un* $\mathbf{k}(p)$*-espace vectoriel. De même, on a* $\mathrm{N} = \oplus_p \mathrm{Ann_N}(p)$*. Soit alors pour tout* p *un supplémentaire* S_p *du* $\mathbf{k}(p)$*-sous-espace vectoriel* $\mathrm{Ann_N}(p)$ *de* $\mathrm{Ann_M}(p)$*. Le* R*-module* $\oplus \mathrm{S}_p$ *est un supplémentaire de* N. ■

5.2.3 « Rappel » sur les corps parfaits

Sur un corps K général, il peut arriver qu'un polynôme sans facteur carré ait des racines multiples dans un corps plus gros. C'est par exemple le cas de $\mathrm{T}^2 + t$ dans $\mathrm{K} = \mathbf{F}_2(t)$ le corps des fractions de l'anneau de polynôme $\mathbf{F}_2(t)$. Ceci n'arrive pas dans les corps parfaits. Soit p un nombre premier et R un anneau tel que $p\mathrm{R} = \{0\}$. La divisibilité bien connue $p| \binom{p}{n}$ pour $1 \leq n \leq p-1$ et la formule du binôme assure que l'application $\mathrm{F} : r \mapsto r^p$ est un morphisme d'anneaux dit morphisme de Frobenius. Si R est un corps, il est de plus injectif comme tout morphisme de corps.

Définition 5.2.3.1. *Un corps de caractéristique p est dit parfait si $p = 0$ ou si tout élément admet une racine p-ième,* i.e. *si son morphisme de Frobenius est un isomorphisme.*

Toute corps fini est donc parfait puisqu'une injection entre ensembles finis est bijective. On doit donc montrer l'énoncé suivant.

Lemme 5.2.3.2. *Soit* $\mathbf{k}$ *un corps parfait et* $\mathrm{P} \in \mathbf{k}[\mathrm{T}]$. *Alors,* P *est sans facteur carré si et seulement si* $\mathrm{PGCD}(\mathrm{P}, \mathrm{P}') = 1$. *En particulier, si* $\mathbf{k}$ *est parfait et* P *irréductible, on a* $\mathrm{PGCD}(\mathrm{P}, \mathrm{P}') = 1$.

Démonstration. *Le sens $\Leftarrow$ découle immédiatement de l'identité de Bézout. Voyons le sens direct. Soit donc* P *sans facteur carré et écrivons* $\mathrm{P} = \prod \mathrm{P}_i$ *avec* P_i *irréductible. Si* $\mathrm{PGCD}(\mathrm{P}, \mathrm{P}') \neq 1$, *un des* P_i *divise* $\mathrm{P}' = \sum_i \mathrm{P}_i' \prod_{j \neq i} \mathrm{P}_j$ *et donc* $\mathrm{P}_i | \mathrm{P}_i'$. *En comparant les degrés, on a* $\mathrm{P}_i' = 0$. *Ceci impose que la caractéristique de* $\mathbf{k}$ *est un nombre premier p et que tous les coefficients de* P_i *d'indices non multiples de p soient nuls :* $\mathrm{P}_i = \sum_n a_{np} \mathrm{T}^{np}$. *Mais dans ce cas, on a* $\mathrm{P}_i = (\sum_n a_{np}^{1/p} \mathrm{T}^n)^p$ *car le Frobenius de* $\mathbf{k}[\mathrm{T}]$ *est un morphisme d'anneaux. Une contradiction avec l'irréductibilité de* P_i ∎

Exercice(s) 5.2.3.3. *Soit* V *un* $\mathbf{k}$*-espace vectoriel de dimension finie et φ un automorphisme de* $\mathbf{k}$. *On note* $[\varphi] \otimes \mathrm{V}$ *l'espace vectoriel de groupe sous-jacent* V *et de loi externe* $\lambda._{[\varphi]} v = \varphi(\lambda) v$. *Montrer* $\dim(\mathrm{V}) = \dim([\varphi] \otimes \mathrm{V})$. *En déduire que tout corps de dimension finie sur un corps parfait est encore parfait.*

5.2.4 Critère de semi-simplicité de V_a

Le calcul du PGCD de polynômes ne dépend pas du corps de base (par exemple car l'algorithme d'Euclide n'en dépend pas) pas plus que celui du minimal du matrice. D'après 4.7.0.1, la condition $\mathrm{PGCD}(\mu_a, \mu_a') = 1$ équivaut donc à ce que la matrice de a est diagonalisable dans $\mathrm{M}_n(\Omega)$. Dans le cas de V_a, on résume alors ce qui précède ainsi.

Proposition 5.2.4.1. *Soit $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ (avec $\mathbf{k}$ parfait) de matrice $\mathrm{A} \in \mathrm{M}_n(\mathbf{k})$ dans une base donnée. Les propositions suivantes sont équivalentes .*

1. *Le minimal μ_a de a est sans facteur carré.*
2. $\mathrm{PGCD}(\mu_a, \mu_a') = 1$.
3. *A est diagonalisable dans $\mathrm{M}_n(\Omega)$.*
4. *V_a est semi-simple.*
5. *Tout sous-module de V_a est semi-simple.*

Si ces conditions équivalentes sont satisfaites, on dit que a est semi-simple (dito pour une matrice de a).

Les familles commutantes d'endomorphismes diagonalisables étant simultanément diagonalisable (4.7.0.3), on déduit

Corollaire 5.2.4.2. *Soit $a, b \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ avec a, b semi simples qui commutent ($\mathbf{k}$ parfait) et $\mathrm{P} \in \mathbf{k}[\mathrm{X}, \mathrm{Y}]$. Alors, $\mathrm{P}(a, b)$ est semi-simple.*

Ce corollaire est faux dans le cas imparfait (cf. cex-ss-imp).

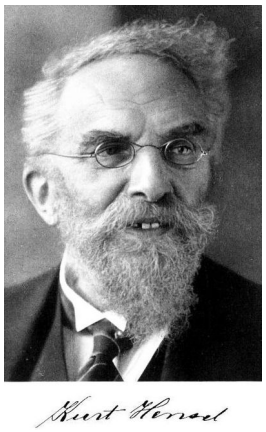
Remarque(s) 5.2.4.3. *Lorsque le corps de base K n'est pas parfait, il existe des matrices semi-simples sur K qui, considérées dans un surcorps, ne le sont plus. Par exemple, c'est le cas de $\mathrm{A} = \begin{pmatrix} 0 & t \\ 1 & 0 \end{pmatrix}$ le corps de fractions de $\mathbf{F}_2[t]$ qui est semi-simple sur K car $\chi_{\mathrm{A}}(\mathrm{T}) = \mathrm{T}^2 + t$ est irréductible sur K mais ne l'est plus sur $\mathrm{K}(t^{1/2}) = \mathrm{K}[\tau]/(\tau^2 - t)$ et a fortiori sur $\Omega \supset \mathrm{K}$. D'ailleurs, $\mathrm{A} + t^{1/2}\,\mathrm{Id}$ est même nilpotente ! La bonne notion dans le cas non parfait est celle d'absolue simplicité définie par la condition $\mathrm{PGCD}(\mu_a, \mu_a') = 1$, plus forte que la semi-simplicité.*

Exercice(s) 5.2.4.4. *Soit p premier, K le corps des fractions de $\mathbf{F}_p[\mathrm{T}]$ et $\mathrm{V} = \mathrm{K}[\mathrm{X}, \mathrm{Y}]/(\mathrm{X}^p - \mathrm{T}, \mathrm{Y}^p - \mathrm{T})$. Montrer que V est de dimension finie sur K et que les K-endomorphismes de V de multiplication par X et Y respectivement sont semi-simples, commutent mais que leur différence est nilpotente (c'est l'exercice 14 chapitre VII.5 [Bou07] réécrit sans produit tensoriel). Démontrer sans recours à la diagonalisation simultanée que la somme de deux matrices absolument semi-simples est absolument semi-simple en utilisant l'exercice 4.14.0.3.*

5.3 Décomposition de Jordan-Chevalley

Commençons par un résultat très important bien que de démonstration aisée qui permet de construire des racines de polynômes de proche en proche (adaptation de la méthode de Newton).

5.3.1 Lemme de Hensel et existence



Kurt Hensel

Isaac Newton

Lemme 5.3.1.1 (Hensel-Newton). *Soit* I *un idéal nilpotent* ($I^N = 0$) *d'un anneau* R *arbitraire et* $P \in R[T]$. *On suppose qu'il existe* $x_0 \in R$ *tel que* $P(x_0) \equiv 0 \mod I$ *et* $P'(x_0) \mod I$ *inversible. Alors, il existe* $x \in R$ *tel que* $x \equiv x_0 \mod I$ *et* $P(x) = 0$.

Démonstration. *Observons d'abord que si* $a \mod I$ *est inversible, alors* a *est inversible dans* a. *En effet, si* $b \mod I$ *est son inverse,* $ab = 1 - i$ *avec* $i \in I$. *En développant formellement* $1/(1-i)$ *en série, on déduit que* $1 - i$ *est inversible d'inverse* $\sum_{k<N} i^k$ *puisque* $i^k = 0$ *pour* $k \geq N$ *et donc* $b/(1-i)$ *est l'inverse de* a.

On va calculer (algorithmiquement) une racine approchée

$$x_k \mod I^{2^k} | P(x_k) \equiv 0 \mod I^{2^k} \text{ et } x_k \equiv x_0 \mod I$$

par approximations successives. On procède par récurrence sur $k \geq 0$*(avec initialisation tautologique). Supposons la propriété vraie au rang* k. *On cherche donc* x_{k+1} *sous la forme* $x_{k+1} + \varepsilon$, $\varepsilon \in I^{2^k}$ *de sorte que* x_{k+1} *est bien une approximation de de* $x_k \mod I^{2^k}$.

La formule de Taylor (entière !) donne

$$P(x_{k+1}) = P(x_k) + \varepsilon P'(x_k) + \varepsilon^2 Q(x_k, \varepsilon)$$

avec $Q[T, Y] \in R[T, Y]$ *(le vérifier !). Comme* $x_k \equiv x_0 \mod I$, *on a* $P'(x_k) \equiv P'(x_0) \mod I$ *et donc* $P'(x_k) \mod I^{2^k}$ *est inversible. On pose alors* $\varepsilon = -P(x_k)/P'(x_k)$. *On a* $\varepsilon \in I^{2^k}$ *par construction de* x_k.

Comme $\varepsilon^2 \in \mathrm{I}^{2^{k+1}}$, *ce choix convient. Pour terminer, on choisit* k *tel que* $2^k \geq \mathrm{N}+1$ *et on pose* $x = x_k$ *: l'algorithme converge exponentiellement !*

Corollaire 5.3.1.2 (Existence)**.** *Soit* $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ *(avec* $\mathbf{k}$ *parfait). Il existe* $d, \nu \in \mathbf{k}[a] \subset \mathrm{End}_{\mathbf{k}}[a]$ *tels que* $a = d + \nu$ *et* d *absolument semi-simple,* ν *nilpotent. En particulier,* d *et* ν *commutent.*

Démonstration. *Soit* $\pi \in \mathbf{k}[\mathrm{T}]$ *le produit des facteurs irréductibles du minimal* μ_a *de* a*. Comme il est sans facteur carré, il est premier avec sa dérivée. Choisissons* $\alpha, \beta \in \mathbf{k}[\mathrm{T}]$ *tels que* $\alpha\pi + \beta\pi' = 1$*.*

Soit I *l'idéal* $\pi(a)\mathbf{k}[a]$ *de* $\mathbf{k}[a]$*. On a* $\mu_a | \pi^n$ *et donc* $\pi^n(a) = 0$ *de sorte que* $\mathrm{I}^n = 0$ *. Par ailleurs, on a* $\beta(a)\pi'(a) = 1 \mod \mathrm{I}$ *et donc* $\pi'(a) \mod \mathrm{I}$ *inversible. En posant* $x_0 = a \in \mathbf{k}[a]$*, on déduit l'existence de* $x \in \mathbf{k}[a]$ *tel que* $x = a \mod \mathrm{I}$ *et* $\pi(x) = 0 \mod \mathrm{I}^n = (0)$*. Posons alors* $d = x$ *et* $\nu = a - \mathrm{P}(a)$*. Comme* $\pi(d) = 0$*, on a* d *absolument semi-simple. Comme* $\nu = a - \mathrm{P}(a) \in \mathrm{I}$ *et* $\mathrm{I}^n = 0$*, on a* ν *nilpotent.* ■

Remarque(s) 5.3.1.3. *C'est essentiellement la preuve de Chevalley. Outre son caractère algorithmique (très rapide), elle est importante car elle permet de définir les parties semi-simples et nilpotentes dans le cadre des algèbres de Lie et des groupes alégbriques (sur un corps parfait), cf. par exemple l'excellent [Bor91].*

5.3.2 Unicité

Théorème 5.3.2.1 (Jordan-Chevalley)**.** *On suppose toujours* $\mathbf{k}$ *parfait.*

1. *Soit* $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ *Il existe un unique couple* (d, ν) *avec* d *semi-simple,* ν *nilpotent,* d *et* ν *qui commutent avec* $a = d + \nu$*.*
2. *Soit* $\chi \in \mathbf{k}[\mathrm{T}]$ *unitaire de degré* n*. Il existe* $\mathrm{P} \in \mathbf{k}[\mathrm{X}]$ *(ne dépendant que de* χ*) tel que si* $\chi_a = \chi$*, on a* $d = \mathrm{P}(a)$ *et en particulier* $d, \nu \in \mathrm{R} = \mathbf{k}[a] \subset \mathrm{End}_{\mathbf{k}}[a]$*.*

Démonstration. *Seule l'unicité demande un argument vu ce qui précède. Soit donc* d, ν *comme dans le théorème et un couple* $d', \nu' \in \mathbf{k}[a]$ *comme dans le corollaire 5.3.1.2. Comme* d, ν *commutent entre eux, ils commutent avec* $d + \nu = a$*. Ils commutent donc avec* d', ν' *car ce sont des polynômes en* a*. Mais* $d + \nu = d' + \nu'$ *ie* $d - d' = \nu' - \nu$*. Mais* $\nu' - \nu$ *est nilpotent (comme somme de nilpotents qui commutent) et* $d - d'$ *semi-simple (comme somme de semi-simples qui commutent, 5.2.4.2) ; un endomorphisme à la fois-semi-simple et nilpotent étant nul puisque de polynôme minimal sans facteur carré et divisant* T^n*, on a bien* $d = d'$ *et* $u = u'$*.* ■

Un endomorphisme a diagonalisable se décompose donc en $d = a$ et $\nu = 0$. Donc $a = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$ se décompose en $a + 0$ et non en $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ comme on pourrait être tenté de l'écrire.

Par ailleurs, l'hypothèse $\mathbf{k}$ parfait ne peut être relâchée : la matrice $\begin{pmatrix} 0 & t \\ 1 & 0 \end{pmatrix}$ de 5.2.4.3 n'a pas de décomposition de Jordan-Chevalley. Si on veut une telle décomposition dans le cas imparfait, il faut se restreindre aux endomorphismes à polynômes caractéristiques *séparables* et remplacer semi-simple par absolument semi-simple. La preuve est alors identique.

5.3.3 Classe de similitude des composantes

On garde les notations précédentes. $a = d + \nu$. Les facteurs invariants de la partie semi-simple d sont entièrement déterminés par χ_a puisque deux endomorphismes diagonalisables de même polynômes caractéristiques sont semblables sur Ω et que les invariants ne dépendent pas du corps de base (cf. 5.3.4.1). De même, les invariants de similitude de a déterminent le type nilpotent $\underline{d}_a$ de ν. Une manière de voir est d'observer que les parties nilpotentes de deux matrices semblables ont des parties nilpotentes semblables par unicité de la décomposition de Jordan-Chevalley.

5.3.4 Appendice : Quid du caractère algorithmique de la décomposition ?

En relisant les preuves *supra*, on se convainc sans peine que trouver d et n est algorithmique dès qu'on connaît le produit π des facteurs irréductibles distincts de P_n. SageMath sait très bien faire cela grâce à la commande $factor$. Mais comment faire si cette commande n'exitait pas. En caractéristique nulle, on se convainc facilement de la formule

$$\pi = \mathrm{P}_n / \operatorname{PGCD}(\mathrm{P}_n, \mathrm{P}'_n)$$

de sorte que le procédé est algorithmique grâce à l'algorithme d'Euclide de calcul du PGCD dans $\mathbf{k}[\mathrm{T}]$. En caractéristique $p > 0$, c'est plus compliqué car il existe des polynômes de dérivée nulles : les polynômes en T^p. L'exercice suivant donne un « algorithme » pour trouver π pour un corps parfait de caractéristique $p >$. Les guillemets sont justifiées par le fait qu'on suppose connu algorithmiquement l'inverse du Frobenius [2] $\mathrm{F} : x \mapsto x^p$ de $\mathbf{k}$.

Exercice(s) 5.3.4.1. *Soit* $\mathbf{k}$ *un corps et* $\chi = \prod \pi_i^{n_i}$ *la décomposition en facteurs irréductibles unitaires de* P *unitaire de degré* n. *On note* $\chi_{red} = \prod \pi_i$. *Dans les 4 premières questions,* $\mathbf{k}$ *est supposé parfait de caractéristique* $p > 0$ *et est* I *l'ensemble des indices* i *tels que* n_i *premier à* p.

1. *Montrer que* $\chi / \operatorname{PGCD}(\chi, \chi') = \prod_{i \in \mathrm{I}} \pi_i$.
2. *Montrer que* $\prod_{i \notin \mathrm{I}} \pi_i$ *est une puissance* p*-ième dans* $\mathbf{k}[\mathrm{T}]$.
3. *Écrire un algorithme calculant* $\prod_{i \in \mathrm{I}} \pi_i$ *et* $\prod_{j \notin \mathrm{I}} \pi_j^{n_j/p}$.

2. Ce qui est par exemple le cas pour les corps finis.

4. *En déduire un algorithme calculant* χ_{red}.
5. *Que vaut* χ_{red} *en caractéristique nulle ?*
6. *Programmer l'algorithme sur* $\mathbf{F}_p$ *? Sur* $\mathbf{F}_{p^n}$ *? Sur un corps parfait général ?*
7. *Comment généraliser sur un corps non parfait ?*
8. *Toujours pour* $\mathbf{k}$ *un corps quelconque, Considérons la suite de polynômes* $\underline{\chi}_{red} = (\chi_i)_{1\leq i\leq n}$ *définis par* $\chi_1 = \chi_{red}$, $\chi_{i+1} = (\chi/(\prod_{j\leq i}\chi_j)_{red}$. *Montrer que* $\underline{\chi}_{red}$ *est la suite des facteurs invariants des endomorphismes semi-simples de polynôme caractéristique* χ.
9. *Supposons de nouveau* $\mathbf{k}$ *parfait et soit* D, N *la décomposition de Jordan-Chevalley de* $\mathrm{M} \in \mathrm{M}_n(\mathbf{k})$. *Quels sont les invariants de similitude de* D *en fonction des invariants* $\underline{\mathrm{P}}$ *de* M *[Utiliser la question précédente] ? Pouvez-vous décrire de même les invariants de* N *en fonction des* P_i *[Se placer dans* $\overline{k}$ *et étudier l'application* $\mathrm{P}_i \mapsto \mathrm{P}_i/\mathrm{P}_{i,red}$ *et ses itérés] ? Programmer l'algorithme obtenu par exemple sur* $\mathbf{F}_p$.

S'agissant du lemme de Hensel, l'écriture même de la preuve est un algorithme qui vit dans $\mathbf{k}[a] \subset \mathrm{M}_d(\mathbf{k})$ où $d = \dim(\mathrm{V})$. Il impose de calculer l'inverse de $\mathrm{P}'(x_n)$ tant que $2^n < d$. C'est un nombre de fois faible, mais si les matrices sont grandes, le calcul est lourd. Une manière de l'alléger est de considérer l'isomorphisme d'algèbres $k[\mathrm{T}]/\mu_a \xrightarrow{\sim} k[a]$ qui envoie T sur a (exercice) et de travailler dans ce quotient, ce qui est moins gourmand en calcul.

Malgré tout, ces algorithmes sont très instables. Pour deux raisons. La première est que le pivot de Gauss est un algorithme numériquement instable. Et travailler avec des coefficients polynomiaux n'arrange rien. La seconde est plus sérieuse. Comme on le verra plus bas, les invariants de similitude ne varient pas continûment avec les coefficients de la matrice (voir par exemple le théorème 8.2.0.2). Dès lors, approximer des valeurs des coefficients devient périlleux. Lorsque les matrices sont à coefficients rationnels, ou dans les corps finis, on peut en étant très soigneux maitriser la hauteur des coefficients et ainsi travailler avec de véritables égalités. Même si ces algorithmes ont tendance à faire exploser les tailles des entiers en jeu... Bref, un vrai sujet de réflexion, une des motivations qui nous a poussé à inclure l'étude topologique des classes de similitude au chapitre 8.

Chapitre 6

Compléments sur la dualité en dimension finie

6.1 Rappels

Dans ce chapitre, V désigne un $\mathbf{k}$ espace vectoriel, qui, sauf mention expresse du contraire, est de dimension finie et $\mathrm{V}^* = \mathrm{Hom}(\mathrm{V}, \mathbf{k})$ désigne son dual ; l'espace vectoriel des applications linéaires de V dans $\mathbf{k}$, *i.e.* des formes linéaires de V.

Si $\varphi \in \mathrm{V}^*, v \in \mathrm{V}$, on note $\langle v, \varphi \rangle = \varphi(v)$ le crochet de dualité [1] $\mathrm{V} \times \mathrm{V}^* \to \mathbf{k}$.

Un hyperplan est le noyau d'une forme linéaire $\varphi \neq 0$. L'hyperplan détermine φ à multiplication par un scalaire non nul près.

On rappelle que si $\mathcal{B} = (\mathrm{e}_i)$ est une base (finie donc) de V, on définit la base duale $\mathcal{B}^* = (\mathrm{e}_i^*)$ de V^* par la formule $\langle \mathrm{e}_i, \mathrm{e}_j^* \rangle = \delta_i, j$. Autrement dit, e_i^* est la i-ème fonction coordonnée et on a $v = \sum_j \langle v, \mathrm{e}_j^* \rangle \mathrm{e}_j$.

En particulier, $\dim(\mathrm{V}^*) = \dim(\mathrm{V})$.

1. Attention, le dual agit à droite sur les vecteurs, cf. [Bou70].

Si $V = k^n = M_{n,1}(\mathbf{k})$ (vecteurs colonnes), on a $M_{1,n}(\mathbf{k}) = \mathbf{k}^n = V^*$ (vecteurs lignes) avec $\langle L, C\rangle = L{}^tC$ où $L \in V^*$ est une ligne et $C \in V$ une colonne. Si $\mathcal{B} = (e_i = [\delta_{i,j}]_{1\leq j\leq n})$ est la base canonique de $k^n = M_{n,1}(\mathbf{k}) = V$, sa base duale $\mathcal{B}^*$ est formée des lignes $e_i^* = {}^te_i$, qui est la base canonique de $M_{1,n}(\mathbf{k}) = \mathbf{k}^n = V^*$.

Si W est un sous-espace de V (voire une partie), on rappelle que son orthogonal est défini par

$$W^\perp = \{\varphi \in V^* | \langle w, \varphi\rangle = 0 \text{ pour tout } w \in W\} \subset V^*.$$

Si maintenant W_* est un sous-espace de V^* (voire une partie) sa polaire dans V est défini par

$$W_*^\circ = \{v \in V | \langle v, \varphi\rangle = 0 \text{ pour tout } \varphi \in W_*\} \subset V.$$

Exemple(s) 6.1.0.1. *Un exemple important vient de la géométrie différentielle. Si f est une fonction régulière sur un ouvert Ω de $\mathbf{R}^n$, sa différentielle en $\omega \in \Omega$ est une forme linéaire sur $T_\omega\Omega = \mathbf{R}^n$: la différentielle $df(\omega)$. Dans la base canonique $(\frac{d}{dx_i}(\omega))_i$ de $T_x\Omega$, cette forme est la jacobienne $J(\omega) = (\frac{df}{dx_j}(\omega))_j$ vue donc comme matrice ligne. Le noyau de $df(\omega)$ n'est autre que l'hyperplan tangent en ω à l'hypersurface d'équation $f = 0$ dès lors que la différentielle est non nulle en ce point. La généralisation à plusieurs fonctions est contenue dans la notion de sous-variété de dimension supérieure.*

6.2 Motivation

Deux manières utiles se concurrencent pour définir un sous-espace vectoriel W de $V = k^n$.

1. Via des générateurs $v_i \in V$: $W = \text{Vect}\{v_i\}$.
2. Via des équations $eq_i \in V^*$: $W = \{v | \langle v, eq_i\rangle = 0\}$ avec

$$\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, eq_i\rangle = \sum_j a_{i,j}x_j = (a, \cdots, a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

On s'intéresse ici d'abord au second point de vue, donc au dual V^* et à l'ensemble de toutes les équations possibles de W : l'orthogonal $W^\perp = \{\varphi \in V^* | \varphi(W) \equiv 0\}$ et au lien avec le premier point de vue.

6.3 Biorthogonalité formelle

Que V soit de dimension finie ou non, tout sous-espace W est évidemment contenu dans l'espace défini par l'ensemble de ses équations

(i) $$W \subset (W^\perp)^\circ = \{v | (\langle v, \varphi\rangle = 0 \text{ pour tout } \varphi \in W^\perp\}.$$

De manière générale, cette inclusion est toujours une égalité

(ii) $$\mathrm{W} = (\mathrm{W}^{\perp})^{\circ} = \{v | (\langle v, \varphi \rangle = 0 \text{ pour tout } \varphi \in \mathrm{W}^{\perp}\}.$$

En effet, si $v \notin \mathrm{W}$, on peut choisir un suppélémentaire S de $\mathrm{W} \oplus kv$ dans W et définir par exemple $\varphi \in \mathrm{W}^{\perp}$ par les conditions $\langle \mathrm{W}, \varphi \rangle = \langle \mathrm{S}, \varphi \rangle = \{0\}$ et $(\langle v, \varphi \rangle = 1$ de sorte que $v \notin (\mathrm{W}^{\perp})^{\circ}$.

6.4 Base ante-duale : bidualité

Désormais, dans tout ce chapitre, V est de dimension finie.

Proposition 6.4.0.1. *Soit* V *de dimension* $n\infty$*. Alors*

1. *L'application linéaire d'évaluation*

$$ev : \begin{cases} \mathrm{V} & \to & \mathrm{V}^{**} \\ v & \mapsto & (\varphi \mapsto (\langle v, \varphi \rangle) \end{cases}$$

   est un ismorphisme.

2. *Pour toute base* $\mathcal{B}_*$ *de* V^* *, il existe une unique base* $\mathcal{B}$ *de* V *dite ante-duale dont la duale est* $\mathcal{B}_*$,i.e. *telle que* $\mathcal{B}^* = \mathcal{B}_*$.

Démonstration. *Pour (1), on remarque que* ev *est injective entre espaces de même dimension finie. Pour (2), on remarque que* $\mathcal{B} = ev^{-1}(\mathcal{B}_*)$ *est l'unique solution au problème posé.* ∎

6.5 Orthogonal, polaire en dimension finie

Proposition 6.5.0.1. *Soit* W, W_* *deux sous espaces de* V, V^* *respectivement. On a*

1. $\dim(\mathrm{W}) + \dim(\mathrm{W}^{\perp}) = n$.
2. $\dim(\mathrm{W}_*) + \dim(\mathrm{W}_*^{\circ}) = n$.
3. $\mathrm{W}_* = (\mathrm{W}_*^{\circ})^{\perp}$.
4. $\mathrm{W} = (\mathrm{W}^{\perp})^{\circ}$.
5. $ev(\mathrm{W}_*^{\circ}) = \mathrm{W}_*^{\perp}$.
6. $ev(\mathrm{W}) = \mathrm{W}^{\perp\perp}$.

Démonstration. *Pour (1), choisisissons une base* $(\mathrm{e}_i, 1 \leq i \leq d$ *de* W *qu'on comlète en une base* $\mathcal{B} = (\mathrm{e}_i, 1 \leq i \leq n$ *de* V*. Si* $\mathcal{B}^* = (\mathrm{e}_i^*)$ *est la base duale, on a alors par construction* $\mathrm{W}^{\perp} = Vect(\mathrm{e}_i, i > d)$. *Pour (2), choisisissons une base* $(\varphi_i, 1 \leq i \leq d$ *de* W_* *qu'on comlète en une base* $\mathcal{B}_* = (\varphi_i, 1 \leq i \leq n$ *de* V^**. Si* $\mathcal{B} = (\mathrm{e}_i)$ *est la base ante-duale, on a alors par construction* $\mathrm{W}_*^{\circ} = Vect(\varphi_i, i > d)$.

En appliquant l'argument de (1) à $\mathrm{W} = \mathrm{W}_*^\circ$ *et à la base* $\varepsilon_i = \mathrm{e}_{n-i}$*, on a alors* $\mathrm{W}^\perp = (\mathrm{W}_*^\circ)^\perp = Vect(\varphi_i, i \leq d) = \mathrm{W}_*$ *ce qui donne (3).*
Le (4) est mis pour mémoire et n'utilise pas la dimension finie (ii).
Pour (5), si $\varphi \in \mathrm{W}_*^\circ$ *et* $w \in \mathrm{W}$*, on a* $ev(v)(\varphi) = \varphi(w)$ *qui est nul car* $\varphi \in \mathrm{W}_*^\circ$ *et donc* $ev(\mathrm{W}_*^\circ) \subset \mathrm{W}^\perp$*. Comme ces deux espaces ont même dimension d'après ce qui précède, cette inclusion est une égalité.*
Pour (6),i $w \in \mathrm{W}$*, et* $\varphi \in \mathrm{W}^\perp$*, on a* $ev(v)(\varphi) = \langle v, \varphi \rangle = 0$ *de sorte que* $\mathrm{W} \subset \mathrm{W}^{\perp\perp}$*. Comme ces deux espaces ont même dimension d'après ce qui précède, cette inclusion est une égalité.* ■

Remarque(s) 6.5.0.2. *Notons qu'orthogonalité et polarité sont des applications strictement décroissantes pour l'inclusion.*

Corollaire 6.5.0.3. *Soit* $\varphi_i \in \mathrm{V}^*$*,* $i = 1, \cdots, m$*. Alors, le rang de* $Vect\{\varphi_i\}$ *est celui de l'application d'évaluation* $\begin{cases} \mathrm{V} & \to & k^m \\ v & \mapsto & (\varphi_i(v))_i \end{cases}$

Démonstration. *Il sufffit d'observer que le noyau de l'évaluation est la polaire de* $Vect\{\varphi_i\}$ *puis d'invoquer la proposition précdente et le théorème du rang.* ■

6.6 Conventions de bidualité (dimension finie)

Le paragraphe précédent permet, en dimension finie donc, permet grâce à ev d'identifier V et son bidual, polaire W_*° de W_* et orthogonal $\mathrm{W}_*^\perp$, W et biorthogonal $\mathrm{W}^{\perp\perp}$. On note alors en général simplement $\mathrm{W}_*^\perp$ pour W_*°. D'une manière générale, en dimension finie, on considère espaces et dual, mais on ne dualise pas le dual grâce à ev et on écrit simplement $\mathrm{W} = \mathrm{W}^{\perp\perp}$ que W soit un sous-espace de V ou de V^*.
A titre d'illustration, donnons le lemme algébrique, facile mais important, qui dans les cas réel est le contenu algébrique du théorème extrema liés en géométrie différentielle (interpréter le résultat en termes d'espaces tangents de sous variétés de $\mathbf{R}^n$ dans l'esprit de l'exemple 6.1.0.1).

Lemme 6.6.0.1. *Soient* φ *et* φ_i*,* $i \in \mathrm{I}$ *des formes linéaires de* V*. Alors,* φ *est compbinaison linéaire des* φ_i *si et seulement si* $\cap_i \mathrm{Ker}(\varphi_i) \subset \mathrm{Ker}(\varphi)$*.*

Démonstration. *Par stricte décoissance de l'orthogonal, la condition*

$$\cap_i \mathrm{Ker}(\varphi_i) = \mathrm{Vect}(\varphi_i)^\perp \subset \mathrm{Ker}(\varphi) = \mathrm{Vect}(\varphi)^\perp$$

est équivalente à l'inclusion

$$\mathrm{Vect}(\varphi) = \mathrm{Vect}(\varphi)^{\perp\perp} \subset \mathrm{Vect}(\varphi_i)^{\perp\perp} = \mathrm{Vect}(\varphi_i).$$

■

Remarque(s) 6.6.0.2 (Lemme de Farkas). *Si* $\mathbf{k} = \mathbf{R}$*, on a un réultat analogue pour les familles finies de demi-espaces* $\mathrm{H}^+, \mathrm{H}_i^+$ *dédinies par les inéquations* $f \geq 0, f_i \geq 0$*. On a en effet* $\cap_i \mathrm{H}_i^+ \subset \mathrm{H}^+$ *si et seulement si* φ *est combinaison linéaire à coefficients positifs des* φ_i*. Voir par exemple David Bart, "A short algebraic proof of the Farkas lemma", Siam PublicationsSIAM journal on optimization, 2008, Vol.19 (1), p.234-239.*

6.7 Contravariance

Soient $\mathrm{V}_i, i = 1, 2, 3$ sont des espaces vectoriels arbitraires,

Définition 6.7.0.1. *Si* $f \in \mathrm{Hom}_{\mathbf{k}}(\mathrm{V}_1, \mathrm{V}_2)$*, on note* ${}^t f \in \mathrm{Hom}_{\mathbf{k}}(\mathrm{V}_2^*, \mathrm{V}_1^*)$ *la transposée de* f *définie par* ${}^t f(\varphi_2) = \varphi_2 \circ f$*, autrement dit,* $\langle v_1, {}^t f(\varphi_2)\rangle = \langle f(v_1), \varphi_2\rangle$ *pour tout* $\varphi_2 \in \mathrm{V}_2^*, v_1 \in \mathrm{V}_1$*.*

On a la proposition (formelle) suivante

Proposition 6.7.0.2. *Soit* $f \in \mathrm{Hom}_{\mathbf{k}}(\mathrm{V}_1, \mathrm{V}_2)$ *et* $\mathcal{B}_i$ *des bases de* V_i*.*

1. *On a* $\mathrm{Mat}_{\mathcal{B}_2^*, \mathcal{B}_1^*}({}^t f) = {}^t\mathrm{Mat}_{\mathcal{B}_1, \mathcal{B}_2}(f)$*.*
2. *L'application* $f \mapsto {}^t f$ *est linéaire injective.*
3. *Si* $f_i \in \mathrm{Hom}_{\mathbf{k}}(\mathrm{V}_i, \mathrm{V}_{i+1})$*, on a (*contravariance de la transposée*)* ${}^t(f_2 \circ f_1) = {}^t f_1 \circ f_2)$*.*
4. *En dimension finie, avec les identifications 6.6, la transposition est involutive.*
5. $\mathrm{Im}({}^t f) = \mathrm{Ker}(f)^\perp$ *et* $\mathrm{Ker}({}^t f) = \mathrm{Im}(f)^\perp$*.*
6. *Si* $\mathrm{V}_1 = \mathrm{V}_2 = \mathrm{V}$*, un sous-espace* W *de* V *est stable par* f *si et seulement si* $\mathrm{W}^\perp$ *est stable par* ${}^t f$*.*

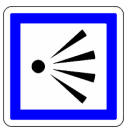
Démonstration. *Donnons un argument pour le 4)(la vérification du reste est laissée en exercice). D'abord il suffit de montrer une des deux formules (changer* f *en* ${}^t f$ *et utiliser l'involutivité de la transposition et de l'orthogonal). Ensuite,* $\mathrm{Im}({}^t f) = \mathrm{Ker}(f)^\perp$ *ayant la même dimension d'après 1) et 6.5.0.1, il suffit de prouver les* $\mathrm{Im}({}^t f) \subset \mathrm{Ker}(f)^\perp$*. Or, si* $f(v_1) = 0$*, on a* $\langle {}^t f(\varphi_2), v_1\rangle = \langle \varphi_2, f(v_1)\rangle = 0$*.*

■

Chapitre 7

Sous-espaces stables

7.1 Point de vue



Comme on va le voir, les espaces stables dépendent fortement du corps de base, contrairement à tout ce qui précède pour l'essentieL C'est ce qui explique que chapitre va en partie déroger à notre volonté de donner des preuves algorithmiques concrètement par la considération de sous-espaces stables particuliers : les sous-espaces caractéristiques. Il sont en effet définis grâce à la décomposition en facteurs irréductibles du polynôme caractéristique d'un endomorphisme, qui dépendent du corps de base et qu'en général on ne sait pas obtenir concrètement.

Même dans le cas complexe, on sait bien qu'il n'est pas possible de« calculer explicitement » les racines d'un polynôme. Ce chapitre n'en reste pas moins important car il existe des cas importants où on a accès aux valeurs propres. Il permet notamment de comprendre la topologie des classes de similitude de matrices dans le cas complexe. Plus généralement, nous discuterons les aspects de continuité des constructions en jeu dans la mesure précisément où l'on ne sait qu'approximer les racines d'un polynôme en général.

7.2 Généralités

On sait que les sous-espaces stables par $a \in \mathrm{End}_{\mathbf{k}}(\mathrm{V})$ sont ses sous-modules. D'après 2.2.3.1, si a est cyclique, ce sont exactement les $\mathrm{P}(a)(\mathrm{V})$ avec P diviseur unitaires de χ. Ils sont en particulier en nombre fini. Il est remarquable que la réciproque soit essentiellement vraie.

Proposition 7.2.0.1. *Si* $\mathbf{k}$ *est infini, un endomorphisme qui n'a qu'un nombre fini de sous-espaces stables est cyclique.*

Démonstration. *Soit a un tel endomorphisme. Il s'agit de trouver un vecteur cyclique. La famille des sous-esapces monogènes stricts de V_a est une famille de sous-espaces stables donc est finie. Comme* $\mathbf{k}$ *est infini, leur réunion n'est pas* V *tout entier. En effet, dans le cas contraire, leur réunion serait* V *tout entier. Choisissons alors pour chacun de ces sous-espaces stricts* W *une forme linéaire non nulle qui s'annule sur* W*. Le produit de ses formes est une fonction polynomiale identiquement nulle. Comme* $\mathbf{k}$ *est infini, l'anneau des fonctions polynomiales sur $\mathrm{V} = \mathbf{k}^n$ est isomorphe à l'anneau des polynômes à n variables, anneau qui est intègre. Donc, une des formes qui est un facteur du produit serait identiquement nulle, une contradiction.* ∎

Évidemment, si $\mathbf{k}$ est fini la proposition est fausse puisqu'il n'y a qu'un nombre fini de sous-espaces de V dans ce ca, stables ou pas.

Remarque(s) 7.2.0.2. *Lorsque $\mathbf{k} = \mathbf{C}$, tout endomorphisme f en dimension > 1 admet des espaces stables non triviaux (prendre des droites propres). Lorsque $\mathbf{k} = \mathbf{R}$, soit il admet des droites stables (valeurs propres réelles) ou des plans stables (prendre par exemple le plan défini par les parties réelles et imaginaires des coordonnées d'un vecteur propre non nul associé à une valeur propre non réelle de la matrice de f dans une base ou, ce qui revient au même, considérer un facteur irréductible de degré 2 polynôme caractéristique). Si $\mathbf{k} = \mathbf{Q}$ et si $\mathrm{P} \in \mathbf{Q}[\mathrm{X}]$ est irréductible de degré n (prendre par exemple $\mathrm{P}(\mathrm{X}) = \mathrm{X}^n - 2$), alors l'endomorphisme de multiplication par* X *sur $\mathbf{Q}[\mathrm{X}]/(\mathrm{P})$ n'a pas de sous-espace stable non trivial puisqu'il est cyclique et que son minimal n'a pas de diviseur strict : les espaces stables d'un endomorphisme dépendent fortement de l'arithmétique du corps de base.*

On pourrait espérer traitre le cas général en se ramenant au cas cyclique grâce à la décomposition de Frobenius 4.13.3. Le problème est que les sous-espaces cycliques qui apparaissent ne sont pas canoniques et aisni se comprotent mal quand on les intersecte avec un sous-espace stable. Ce n'est pas le cas des sous-espaces caractéristiques.

7.3 Sous-espaces caractéristiques

Soit $a \in \mathrm{End}(\mathrm{V})$. On rappelle (4.6.0.1) que μ_a et χ_a ont les mêmes facteurs irréductibles et que le module V_a est annulé par son polynôme caractéristique (Cayley-Hamilton).

Définition 7.3.0.1. *Soit $\pi \in \mathbf{k}[\mathrm{T}]$ un facteur irréductible unitaire de son polynôme caractéristique $\chi_a(\mathrm{T})$. Le sous-espace caractéristique de a associé à π est la composante π-primaire $\mathrm{V}_a[\pi]$ de V_a*

$$\mathrm{V}_a[\pi] = \cup_{i\geq 1} \mathrm{Ann}_{\mathrm{V}_a}[\pi^i] = \cup_{i\geq 1} \mathrm{Ker}(\pi^i(a)).$$

C'est un sous-espace stable par a.

On a alors
D'après le lemme chinois 1.2.0.1 et , on a alors

Proposition 7.3.0.2. *Soit $\chi = \prod \pi^{v_\pi(\chi)}$ la décomposition de $\chi = \chi_a$ en facteurs irréductibles et $\mu_a = \prod \pi^{v_\pi(\mu_a)}$ celle de μ_a.*

1. *On a $v_\pi(\mu_a) \leq v_\pi(\chi)$.*
2. *Il existe $u_{\pi,\chi} \in \mathbf{k}[\mathrm{T}]$ (presque nuls) ne dépendant que de χ tels que $\sum u_{\pi,\chi}(\chi/\pi^{v_\pi(\chi)}) = 1$.*
3. *On a $\mathrm{V}_a = \oplus \mathrm{V}_a[\pi]$ et la projection p_π sur $\mathrm{V}_a[\pi]$ parallèlement à $\oplus \pi' \neq \pi \mathrm{V}_a[\pi']$ est l'homothétie de rapport $\mathrm{e}_\pi(a) =$ avec $\mathrm{e}_\pi = u_{\pi,\chi}(\chi/\pi^{v_\pi(\chi)}) \in \mathbf{k}[\mathrm{T}]$.*
4. *Les p_π forment une famille orthogonale de projecteurs* spectraux *de V_a* i.e.$\sum p_\pi = \mathrm{Id}$ *et $p_\pi p_{\pi'} = \delta_{\pi,\pi'} p_\pi$.*
5. *Chaque $\mathrm{V}_a[\pi]$ est stable par a et $\mathrm{V}_a[\pi] = \mathrm{Ker}(\pi^{v_\pi(\chi)}(a) = \mathrm{Ker}(\pi^{v_\pi(\mu_a)}$.*
6. *Si W est stable par a, on a $\mathrm{W}_a[pi] = \mathrm{V}_a[\pi] \cap \mathrm{W}$: : tout-sous espace stable par a est somme directe de ses intersections avec les sous-espaces caractéristiques.*
7. *On a $\dim_{\mathbf{k}} \mathrm{V}_a[\pi] = \deg(\pi^{v_\pi(\chi)}) = v_\pi(\chi)\deg(\pi)$.*

Démonstration. *Les 6 premiers points sont une réécriture du lemme chinois 2.6.0.1 et de la fonctorialité des composantes primaires. Pour (7), rappelons que chaque sous-espace caractéristique est stable par a. Puisqu'une puissance de π annule $\mathrm{V}_a[\pi]$, le polynôme caractéristique $\chi_{a|\mathrm{V}_a[\pi]}$ de la restriction de a à $\mathrm{V}_a[\pi]$ est une puissance π^{w_π}. Mais comme V_a est la somme directe des $\mathrm{V}_a[\pi]$, on a*

$$\prod\nolimits_{\pi|\chi_a} \pi^{v_\pi(\chi)} = \chi_a = \prod\nolimits_{\pi|\chi_a} \chi_{a|\mathrm{V}_a[\pi]} = \prod\nolimits_{\pi|\chi_a} \pi^{w_\pi}$$

de sorte que $w_\pi = v_\pi(\chi)$. Mais

$$\dim_{\mathbf{k}} \mathrm{V}_a[\pi] = \deg \chi_{a|\mathrm{V}_a[\pi]} = w_\pi \deg(\pi) = v_\pi(\chi)\deg(\pi).$$

■

Observons que les projecteurs spectraux sont définis par $\mathrm{e}_{\pi,\chi} \in \mathbf{k}[\mathrm{T}]$ qui ne dépend que de χ_a. Ceci illustre le fait que $p_{\pi,\chi} = \mathrm{e}_{\pi,\chi}(a)$ « varie continûment » lorsque a varie continûment, *i.e.*quand a est définie par une fonction matricielle continu $\mathrm{A} : \Omega \to \mathrm{M}_n(\mathbf{k})$, ce dès que le polynôme caractéristique $\chi_{\mathrm{A}(\omega)}(\mathrm{T})$ est indépendant de $\omega \in \Omega$. Il en est de même des espaces caractéristiques. En particulier, leur dimension est (localement) constante sur Ω sous cette condition (très forte évidemment). Le lecteur précisera le sens de cet énoncé quand $\mathbf{k} = \mathbf{R}$ ou $\mathbf{k} = \mathbf{C}$ ou, pour le lecteur savant, dans le cas général pour la topologie de Zariski. Ce point est crucial, même s'il sera un peu caché, dans l'étude topologique des classes de similitude (8 et 7.3.1).

Le lemme suivant est important et découle immédiatement du fait qu'un sous-espace stable est somme de ses intersections avec les sous-espaces caractéristiques. C'est ce type résultat qui va nous permettre de ramener l'étude de la topologie des classes de similitude au cas de la topologie des classes de similitude de matrices nilpotentes.

Lemme 7.3.0.3 (Invariance par extension de corps)**.** *Soit* $\mathrm{A} \in \mathrm{M}_n(\mathbf{k})$ *et* $\chi_{\mathrm{A}} = \prod_\pi \pi^{v_\chi}$ *sa décomposition en facteurs irréductibles (unitaires). On note* $\mathrm{A}, \mathrm{A}_{\mathrm{K}}$ *les endomorphismes correspondants de* $\mathbf{k}^n, \mathrm{K}^n$*. On a alors* $\mathrm{Ker}(\pi^{v_\pi(\chi)}(\mathrm{A}_{\mathrm{K}})) = \oplus_{\tilde{\pi}|\pi} \mathrm{V}_{\mathrm{A}_{\mathrm{K}}}[\tilde{\pi}]$ *où* $\tilde{\pi}$ *décrit les diviseurs unitaires irréductibles de* π *dans* $\mathrm{K}[\mathrm{T}]$.

7.3.1 Propriétés topologiques dans le cas complexe

On considère dans ce numéro une suite de matrices $\mathrm{A}_n \in \mathrm{M}_d(\mathbf{C})$ (identifiées à des endomorphismes de $\mathbf{C}^n$) dont on veut étudier la convergence éventuelle vers une matrice notée $\mathrm{A}_\infty \in \mathrm{M}_d(\mathbf{C})$ pour la topologie définie par une norme sur $\mathrm{V} = \mathbf{C}^n$ [1] en fonction de ses projections sur ses espaces spectraux. Comme on s'intéressera (cf. chapitre 8) au cas où les A_n sont toutes dans une même classe de similitude (dont on cherche à étudier l'adhérence), on suppose de plus que le polynôme caractéristique de A_n est un polynôme constant.

Par continuité du polynôme caractéristique en les coefficients de la matrice (ce sont des polynômes en les coefficients), la convergence de A_n vers A_∞ impose $\chi_{\mathrm{A}_\infty}(\mathrm{T}) = \det(\mathrm{T}\,\mathrm{Id} - \mathrm{A}_\infty) = \chi_{\mathrm{A}_n}(\mathrm{T})$, condition que l'on suppose donc réalisée.

Soit donc Λ le spectre de A_∞, ensemble des racines complexes de χ_{A_∞} et v_λ leurs multiplicités correspondantes. Comme dans 7.3.0.2, choisissons $u_\lambda(\mathrm{T}) \in \mathbf{C}[\mathrm{T}]$ tels que

$$\sum_{\lambda \in \Lambda} u_\lambda(\mathrm{T}) \frac{\chi_\lambda(\mathrm{T})}{(\mathrm{X} - \lambda)^{v_\lambda}} = 1$$

de sorte que les polynômes

$$\mathrm{e}_\lambda(\mathrm{T}) = u_\lambda(\mathrm{T}) \frac{\chi_\lambda(\mathrm{T})}{(\mathrm{X} - \lambda)^{v_\lambda}}$$

1. Comme évoqué plus haut, le lecteur savant pourra utilement discuter le cas d'un corps infini quelconque avec l'espace de matrice $\mathrm{M}_n(\mathbf{k})$ muni par exemple de la topologie de Zariski, tous les fermés intervenant de manière essentielle étant définis par des équations polynomiales comme il apparaîtra naturellement.

définissent les projecteur spectraux

(i) $$p_{\lambda,n} = \mathrm{e}_\lambda(\mathrm{A}_n),\ n \in \overline{\mathbf{N}} = \mathbf{N} \cup \{\infty\}$$

associés à A_n . Puisqu'on a

$$\sum_\Lambda p_{\lambda,n} = \mathrm{Id}_\mathrm{V}$$

on en déduit

(ii) $$\mathrm{A}_n = \sum_\Lambda \mathrm{A}_{n,\lambda}\ n \in \overline{\mathrm{N}}$$

où $\mathrm{A}_{n,\lambda} = \mathrm{A}_n p_{\lambda_n,\ n\in\overline{\mathbf{N}}}$ est la restriction de A_n sur l'espace caractéristique associé à λ et 0 sur les autres de sorte que $\mathrm{A}_{n,\lambda} - \lambda\,\mathrm{Id}$ est nilpotent. Une autre manière de dire est que la partie semi-simple de A_n est $\sum \lambda \mathrm{e}_\lambda(\mathrm{A}_n)$.

Proposition 7.3.1.1. *Avec les notations précédentes et l'hypothèse $\chi_{\mathrm{A}_n}(\mathrm{T})$ indépendant de $n \in \overline{\mathbf{N}}$, on a $\lim \mathrm{A}_n = \mathrm{A}_\infty$ si et seulement si pour tout $\lambda \in \Lambda$, $\lim \mathrm{A}_{n,\lambda} = \mathrm{A}_{\infty,\lambda}$.*

Démonstration. *C'est une conséquence immédiate des formules (i) et (i).* ∎

Lemme 7.3.1.2. *Soient $\mathrm{P}_{n,d}|\cdots|\mathrm{P}_{n,1}$ les invariants de similitudes de A_n. Alors, les invariants de similitude $\mathrm{A}_{n,\lambda}$ sont $1,\cdots,1,(\mathrm{X}-\lambda)^{v_\lambda(\mathrm{P}_{n,i})},\ i = d,\cdots,1$ où les 1 sont répétés $d - v_\lambda(\chi)$-fois.*

Démonstration. *C'est une autre façon d'écrire la remarque 4.12.0.3. Rappelons l'argument sans explicitement invoquer la décomposition de Jordan. Pour $\mathrm{A} = \mathrm{A}_n$ d'invariants $\mathrm{P}_i = \mathrm{P}_{n,i}$, le module V_A est isomorphe à $\oplus k[\mathrm{T}]/\mathrm{P}_i(\mathrm{T})$, d'écrire la décomposition $\mathrm{P}_i = \prod_\Lambda (\mathrm{X}-\lambda)^{v_\lambda(\mathrm{P}_i)}$ (car P_i divise χ_A) puis d'invoquer le lemme chinois pour écrire*

$$\mathrm{V}_\mathrm{A} \simeq \oplus_i \oplus_\Lambda k[\mathrm{T}]/(\mathrm{T}-\lambda)^{v_\lambda(\mathrm{P}_i)}.$$

Mais A_λ agit par A sur la composante $(\mathrm{X}-\lambda)$-primaire $\mathrm{V}_{\mathrm{A},\lambda} = \mathrm{e}_\lambda(\mathrm{T})\mathrm{V}_\mathrm{A}$ et par 0 sur les composantes $(\mathrm{X}-\tilde{\lambda})$-primaires $\mathrm{V}_{\mathrm{A},\tilde{\lambda}} = \mathrm{e}_{\tilde{\lambda}}(\mathrm{T})\mathrm{V}_\mathrm{A}$ si $\tilde{\lambda} \neq \lambda$. La composante $(\mathrm{T}-\lambda)$-primaire (de dimension la multiplicité $v_\lambda(\chi)$ de la racine λ de χ d'après (7.3.0.2) s'écrit alors

$$\mathrm{V}_{\mathrm{A},\lambda} = \mathrm{V}_\mathrm{A}[\mathrm{T}-\lambda] \simeq \oplus_i k[\mathrm{T}]/(\mathrm{T}-\lambda)^{v_\lambda(\mathrm{P}_i)}$$

et on conclut grâce à l'unicité des invariants de similitude. ∎

On peut également donner une forme légèrement différente de l'énoncé précédent.

Proposition 7.3.1.3. *Avec les notations précédentes et l'hypothèse $\chi_{\mathrm{A}_n}(\mathrm{T})$ indépendant de $n \in \overline{\mathbf{N}}$, on a $\lim \mathrm{A}_n = \mathrm{A}_\infty$ si et seulement si les parties semi-simples (resp. nilpotentes) de la décomposition de Jordan-Chevalley convergent vers la partie semi-simple (resp. nilpotente) de $\mathrm{A}_{\infty,\lambda}$.*

Démonstration. *C'est une conséquence immédiate du fait qu'il existe $\mathrm{P} \in \mathbf{k}[\mathrm{T}]$ ne dépendant que de χ telles que les parties semi-simples et nilpotentes de $\mathrm{A}_n, n \in \overline{\mathbf{N}}$ soient $\mathrm{P}(\mathrm{A}_n)$ (resp. $\mathrm{A} - \mathrm{P}(\mathrm{A}_n)$) d'après (5.3.2.1).* ∎

7.3.2 Racines d-ièmes dans GL_n

Si $\mathrm{A} \in \mathrm{M}_n(\mathbf{k})$ avec $\chi_{\mathrm{A}}(\mathrm{T}) = \prod(\mathrm{X} - \lambda)^{v_\lambda}$ scindé, on retrouve donc la définition usuelle rencontrée en algèbre linéaire. Si $\mathrm{pr}_\lambda = \mathrm{e}_\lambda(\mathrm{A})$ est comme plus haut, le projecteur spectral sur $\mathrm{V}[\mathrm{T} - \lambda] = \mathrm{Ker}(\mathrm{A} - \lambda)^{v_\lambda(\chi)}$), la décomposition de Jordan-Chevalley $\mathrm{A} = \mathrm{D} + \mathrm{N}$ se calcule simplement par

$$d = \sum \lambda \mathrm{e}_\lambda(\mathrm{A}) \text{ et } \mathrm{N} = \mathrm{A} - \mathrm{D}$$

comme on vient de le voir. Une application immédiate, et utile est l'existence de racines d-ièmes polynomiales dans le cas algébriquement clos.

Proposition 7.3.2.1. *Soit d un entier > 0 et supposons $\mathbf{k}$ algébriquement clos de caractéristique première à d. Soit χ unitaire de degré n. Il existe $\mathrm{P}_{d,\chi} \in \mathbf{k}[\mathrm{T}]$ tel que pour toute matrice $\mathrm{A} \in \mathrm{GL}_n(\mathbf{k})$ tel que $\chi_{\mathrm{A}} = \chi$ on ait $\mathrm{P}_{d,\chi}(\mathrm{A})^d = \mathrm{A}$.*

Démonstration. *Comme $\chi(0) \neq 0$, les polynômes χ et T sont premiers entre eux et on peut écrire une identité de Bézout $\mathrm{UT} + \mathrm{V}\chi = 1$ dans $\mathbf{k}[\mathrm{T}]$. Avec les notations précédentes, comme $\chi_{\mathrm{D}} = \chi_{\mathrm{A}} = \chi$, la matrice D est inversible d'inverse $\mathrm{U}(\mathrm{D})$. Comme D et N commutent,*

$$\mathrm{A} = \mathrm{D}(\mathrm{Id} + \mathrm{D}^{-1}\mathrm{N}) = \mathrm{D}(\mathrm{Id} + \mathrm{U}(\mathrm{D})\mathrm{N})$$

avec $\mathrm{D}^{-1}\mathrm{N}$ est nilpotente. On peut alors écrire une racine d-ième de D sous la forme

$$\mathrm{D}^{1/d} = \sum \lambda^{1/d} \mathrm{e}_\lambda(\mathrm{A})$$

qui est donc un polynôme ne dépendant que de χ et d évalué en A. Par ailleurs, les coefficients de la série entière $(1+z)^{1/d}$ sont les coefficients binomiaux généralisés $\binom{1/d}{i}$, $i \geq 0$ et donc sont dans $\mathbf{Z}[1/d]$. Comme d est inversible dans $\mathbf{k}$ et $(\mathrm{D}^{-1}\mathrm{N})^n = 0$, on a une racine d-ième

$$(\mathrm{D}^{-1}\mathrm{N})^{1/d} = \sum_{i<d} \binom{1/d}{i} (\mathrm{D}^{-1}\mathrm{N})^i$$

qui est bien un polynôme ne dépendant que de χ et d évalué en A comme le sont D^{-1} et N, ce qu'on voulait ∎

On ne peut espérer mieux. D'une part, l'énoncé est clairement faux dans le cas non algébriquement clos en général, déjà dans le cas $n = 1$. D'autre part, une matrice nilpotente N non nulle n'admet pas de racine d-ième. En effet, celle-ci serait nilpotente de sorte que sa puissance n-ième serait nulle mais aussi égale à $n\,!$

Chapitre 8

Topologie des classes de similitude

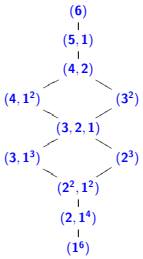
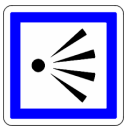


Diagramme de Hasse de GL_6

8.1 Point de vue



On donné ici un aperçu de la géométrie des classes de similitude via leur topologie. Pour éviter le formalisme, on s'est restreint à la topologie usuelle sur les matrices complexes même si la topologie, dite de Zariski, dont les fermés sont définis par des familles d'équations polynomiales, aurait été plus naturelle [1].

8.2 Introduction

Définition 8.2.0.1. *Un n-type est une suite de polynômes unitaires $\underline{\mathrm{P}} = (\mathrm{P}_n|\mathrm{P}_{n-1}|\cdots\mathrm{P}_1)$ suite de polynômes de $\mathbf{k}[\mathrm{T}]$ tels que $\sum \deg(\mathrm{P}_i) = n$. On note $\mathrm{O}(\underline{\mathrm{P}})$ l'ensemble des matrices de $\mathrm{M}_n(\mathbf{k})$ semblables à la matrice compagnon $\mathrm{C}(\underline{\mathrm{P}})$.*

1. Comme plus haut donc, dans le cas d'un corps infini général, il faudrait considérer la topologie de Zariski, ce qui n'ajoute aucune difficulté réelle lorsqu'on en connaît la définition. Il suffit en fait que la topologie soit plus fine que celle de Zariski, que les opérations usuelles sur les matrices soient continues et que les points de $\mathbf{k}$ ne soient pas ouverts, assurant ainsi que l'adhérence de $\mathbf{k}^*$ est $\mathbf{k}$. C'est ici que l'infinitude du corps intervient dans le cas de la topologie de Zariski..

Ainsi, $O(\underline{P})$ est l'orbite de $C(\underline{P})$ sous l'action de $GL_n(\mathbf{k})$ par conjugaison. La théorie des invariants de similitude nous dit que $O(\underline{P})$ est formée des matrices d'invariants de similitude $\underline{P}$ et que $M_d(\mathbf{k})$ est la réunion disjointe des $O(\underline{P})$ lorsque $\underline{P}$ parcourt tous les n-types (4.9.0.2).
Notre but est d'étudier l'adhérence $\overline{O(\underline{P})}$ des orbites $O(\underline{P})$. On supposera donc dans la suite de ce chapitre que $\mathbf{k}$ est le corps des complexes, les espaces de matrices étant munies d'une norme (elles sont toutes équivalentes).
On définit alors une relation (topologique) sur les n-types complexes par

$$\underline{P} \preceq \underline{Q} \text{ si et seulement si } O(\underline{P}) \text{ est contenue dans l'adhérence } \overline{O(\underline{Q})}.$$

C'est visiblement une relation d'ordre, mais partiel comme on le verra. Comme $\overline{O(\underline{Q})}$ est invariante par conjugaison, c'est une réunion d'orbites et on a $\overline{O(\underline{Q})} = \cup_{\underline{P} \preceq \underline{Q}} O(\underline{P})$. Nous allons caractériser cette relation d'ordre de manière combinatoire de la manière suivante.
On définit une relation (combinatoire [2]) sur les n-types complexes par

$$\underline{P} \leq \underline{Q} \text{ si et seulement si et seulement on a la divisibilité } \prod_{j \leq i} P_j | \prod_{j \leq i} Q_j \text{ pour tout } i = 1, \cdots, n.$$

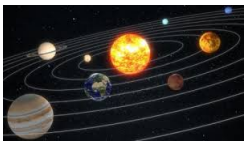
C'est également une relation d'ordre. Observons que nécessairement on a alors $\prod_{i=1}^{n} P_i = \prod_{i=1} Q_i$ pour des raisons de degré.

Théorème 8.2.0.2. *Soit $\underline{P}, \underline{Q}$ deux n-types complexes. Alors, $\underline{P} \preceq \underline{Q}$ si et seulement $\underline{P} \leq \underline{Q}$. Autrement dit, les ordres topologique et combinatoire sur les n-types coïncident.*

Remarque(s) 8.2.0.3. *Ce théorème est une reformulation, plus transparente à mon avis, du théorème 4 de [Ger61]. C'est en effet à ma connaissance Gerstenhaber qui a explicité de manière complète la structure des adhérences d'orbites même si je n'ai pas réussi à y trouver cet énoncé stricto sensu.*

Nous allons procéder par réduction au cas nilpotent grâce aux résultats topologiques de 7.3.1. Commençons donc par le cas crucial.

8.3 Adhérence d'une orbite nilpotente



Les orbites nilpotentes sont classifiées par les partitions $\underline{d}$ de n (4.12.0.2), le dictionnaire entre type et partition étant donné par $\underline{d} \mapsto T^{\underline{d}}$. On notera alors $O(\underline{d})$ l'orbite $O(T^{\underline{d}})$ correspondante.

On a donc encore une relation d'ordre topologique sur les partitions de n définie par

2. Comparer avec cf. 8.3.

$\underline{d} \preceq \underline{\delta}$ si et seulement si $\mathrm{O}(\underline{d})$ est contenue dans l'adhérence $\overline{\mathrm{O}(\underline{\delta})}$

et une relation d'ordre combinatoire

$\underline{d} \leq \underline{\delta}$ si et seulement pour tout $i = 1, \cdots, n$ on a l'inégalité $\sum_{j \leq i} d_j \leq \sum_{j \leq i} \delta_j$.

Le théorème 8.2.0.2 devient alors dans ce cas

Théorème 8.3.0.1 (Cas nilpotent). *Soit $\underline{d}, \underline{\delta}$ deux partitions de n. Alors, $\underline{d} \preceq \underline{\delta}$ si et seulement pour tout $\underline{d} \leq \underline{\delta}$.*

Ainsi, on veut montrer que les ordres topologique $\preceq$ et combinatoire $\leq$ sur les partitions coïncident. On écrit souvent une partition en indiquant le nombre de fois où l'on répète un entier, souvent en ordre croissant. Pour $n = 6$ par exemple, la partition $(3, 1, 1, 1, 0, 0)$ est alors noté $(1^3, 3)$ tandis que la partition $(6, 0, 0, 0, 0, 0)$ est notée (6). Le diagramme décrivant l'odre s'appelle alors diagramme de Hasse. Nous n'utiliserons pas ces notations exceptée dans l'illustration de la section.

8.3.1 Ordre et dualité sur les partitions

On utilise notations et résultats sur les matrices nilpotentes de 4.12.2. Nous allons démontrer que la dualité des partitions est décroissante pour l'ordre combinatoire $\leq$. Pour cela, et la suite, la clef est le classique lemme de dévissage clef suivant dont je reprends la preuve de [Ros20].

On dira que $\underline{d} \underset{\mathrm{e}}{\leq} \underline{\delta}$ ($\underline{d}$ élémentairement inférieur à $\underline{\delta}$) si il existe des indices $i < j$ tels que

$$(\delta_1, \ldots, \delta_n) = (d_1, \ldots, d_{i-1}, d_i + 1, \ldots, d_j - 1, \ldots, d_n).$$

On a évidemment

$$\underline{d} \leq_{\mathrm{e}} \underline{\delta} \Rightarrow \underline{d} \leq \underline{\delta}$$

Lemme 8.3.1.1. *Soit $\underline{d}, \underline{\delta}$ deux partitions de n. Alors, $\underline{d} \leq \underline{\delta}$ si et seulement il existe une suite d'inégalités élémentaires $\underline{d} = \nu_0 \underset{\mathrm{e}}{\leq} \underline{\nu}_1 \underset{\mathrm{e}}{\leq} \ldots \underset{\mathrm{e}}{\leq} \underline{\nu}_{\mathrm{N}-1} \underset{\mathrm{e}}{\leq} \underline{\nu}_{\mathrm{N}} = \underline{\delta}$.*

Démonstration. *Il suffit de prouver l'existence d'une partition $\underline{\nu}$ telle que $\underline{d} \underset{\mathrm{e}}{\leq} \underline{\nu} \leq \underline{\delta}$ lorsque $\underline{d} \neq \underline{\delta}$ et d'itérer le processus (qui s'arrête lorsque $\underline{\nu}_{\mathrm{N}} = \underline{\delta}$.) On cherche donc $i < j$ tel que $\underline{\nu} \leq \underline{\delta}$ avec*

$$\underline{\nu} = (d_1, \ldots, d_{i-1}, d_i + 1, \ldots, d_j - 1, \ldots, d_n).$$

Si $\underline{\nu} = \underline{\delta}$, on a terminé. Sinon, $\underline{\nu} < \underline{\delta}$.

Il existe donc k tel que

$$d_1 + \cdots + d_k < \delta_1 + \cdots + \delta_k \tag{1}$$

Soit i le plus petit entier k vérifiant (1)
Par ailleurs, comme $\sum d_k = \sum \delta_k$, il existe donc $k > i$ tel que

$$d_1 + \cdots + d_k \geq \delta_1 + \cdots + \delta_k \tag{2}.$$

Soit j le plus petit entier $k > i$ vérifiant (2).
On a donc

$$d_1 + \cdots + d_k + 1 \leq \delta_1 + \cdots + \delta_k \text{ pour tout } k \in [i, j-1] \tag{3}$$

et

$$d_1 + \cdots + d_j = \delta_1 + \cdots + \delta_j \tag{4}$$

Avec ces valeurs de i et j, montrons que $\underline{\nu}$ est une partition, i.e. *$d_{i-1} > d_i$ (ou $i = 1$) d'une part et $d_j > d_{j+1}$ d'autre part.*
Par construction, i est le plus petit entier tel que $d_i < \delta_i$ et donc $d_i < \delta_i \leq \delta_{i-1} = d_{i-1}$ (ou $i = 1$).
D'autre part, plus, puisque $d_1 + \cdots + d_{j-1} < \delta_1 + \cdots + \delta_{j-1}$ et $d_1 + \cdots + d_j = \delta_1 + \cdots + \delta_j$ on a $\delta_j < d_j$; comme de plus et $d_1 + \cdots + d_{j+1} \leq \delta_1 + \cdots + \delta_{j+1}$ on a aussi $d_{j+1} \leq \delta_{j+1}$. En combinant les deux, on obtient $d_{j+1} \leq \delta_{j+1} \leq \delta_j < d_j$, ce qu'on voulait.
On observe alors que 'inégalité $\underline{\nu} \leq \underline{\delta}$ est équivalente (3). ■

Corollaire 8.3.1.2. *La dualité des partitions est (strictement) décroissante.*

Démonstration. *Il suffit de montrer la décroissance dans le cas élémentaire $\underline{d} \leq \underline{\delta}$. Pour cela on observe que $\underline{\delta}^*$ vérifie*

$$\delta_k^* = \begin{cases} d_k & si\ k \neq d_i, d_j \\ d_k - 1 & si\ k = d_i \\ d_k + 1 & si\ k = d_j \end{cases}$$

de sorte que $\underline{\delta}^* \leq \underline{d}^*$. Pour le voir, on observe que $d_i > d_j$ et on regarde le tableau suivant

k	$\underline{d}^*$	$\underline{\delta}^*$	*comparaison*	$\mathrm{Card}(\underline{\delta}^*) - \mathrm{Card}(\underline{d}^*)$
[1,i-1]	$d_k \geq \alpha$	$d_k \geq \alpha$	*idem*	*0*
i	$d_k \geq \alpha$	$d_k \geq \alpha + 1$	*idem sauf si* $\alpha = d_i$	*-1*
[i-1,j-1]	$d_k \geq \alpha$	$d_k \geq \alpha$	*idem*	*0*
j	$d_k \geq \alpha$	$d_k \geq \alpha - 1$	*idem sauf si* $\alpha = d_j$	*+1*
[j+1,n]	$d_k \geq \alpha$	$d_k \geq \alpha$	*idem*	*0*

en utilisant la formule de calcul de la partition duale $d_\alpha^* = \mathrm{Card}\{k | d_k \geq \alpha\}$ (4.12.2.2). La preuve donne également la décroissance stricte (même si le caractère strict découle du fait que la dualité est involutive) ■

8.3.2 Rang et orbites nilpotentes

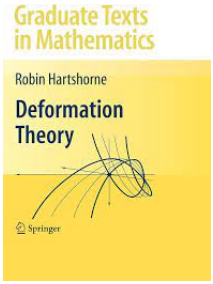
Soit M une matrice nilpotente de partition associée $\underline{d}$. D'après la formule (vi) de 4.12.2, on a pour tout $n - \mathrm{rg}(\mathrm{M}^i) = \sum_{j \leq i} d_j^*$. Mais le rang est semi-continu inférieurement : il existe un voisinage U de M où toutes les matrices sont de rang supérieur ou égal à celui de M. Si M est dans l'adhérence de $\mathrm{O}(\underline{\delta})$, ce voisinage rencontre $\mathrm{O}(\underline{\delta})$: soit donc $\mathrm{N} \in \mathrm{U} \cap \mathrm{O}(\underline{\delta})$. On a alors $n - \mathrm{rg}(\mathrm{N}^i) \leq n - \mathrm{rg}(\mathrm{M}^i)$ pour tout i, autrement dit $\underline{\delta}^* \leq \underline{d}^*$ et donc $\underline{\delta} \leq \underline{d}$.

Corollaire 8.3.2.1. *Soient $\underline{d}, \underline{\delta}$ des partitions de n. Alors,*

$$\underline{d} \preceq \underline{\delta} \Rightarrow \underline{d} \leq \underline{\delta}.$$

Montrons l'implication réciproque.

8.3.3 Une déformation de matrice nilpotente



D'après le lemme de dévissage 8.3.1.1, il s'agit simplement de montrer l'implication dans le cas élémentaire. Soit donc $\underline{d} \underset{\mathrm{e}}{\leq} \underline{\delta}$ et montrons qu'on a $\underline{d} \preceq \underline{\delta}$. Il donc existe des indices $i < j$ tels que

$$(\delta_1, \ldots, \delta_n) = (d_1, \ldots, d_{i-1}, d_i + 1, \ldots, d_j - 1, \ldots, d_n).$$

On considère donc $\mathrm{J}_{\underline{d}}$ dont on veut montrer qu'elle est dans l'adhérence de $\mathrm{O}(\underline{\delta})$, donc on veut montrer que $\mathrm{J}_{\underline{d}}$ est une limite de matrices de $\mathrm{O}(\underline{\delta})$.

Comme $\underline{d}$ et $\underline{\delta}$ ne diffèrent qu'aux indices i et j, on peut supposer sans perte de généralité qu'on n'a que deux indices. On doit donc montrer que $\mathrm{J}_{(d_i,d_j)}$ est dans l'adhérence de $\mathrm{O}((d_i - 1, d_j + 1))$. Posons par exemple alors $\mathrm{N}(x) = \mathrm{J}_{(d_i,d_j)} + x\mathrm{E}_{d_i+d_j,d_i}$. C'est une matrice triangulaire par blocs de taille $d_i + d_j$ et de rang $d_i + d_j - 2$ avec $d_i > d_j$. Son type est caractérisé par son indice de nilpotence qui est $d_i - 1$

(4.12.0.4) pour x non nul de sorte que $\mathrm{N}(x)$ est de type $d_i - 1, d_j + 1$. Ainsi, $\mathrm{N}(0) = \lim_{x\to 0} \mathrm{N}(x) \in \overline{\mathrm{O}(\underline{\delta})}$ et $\underline{d} \preceq \underline{\delta}$. On a donc en se souvenant 8.3.2.1

$$\underline{d} \preceq \underline{\delta} \Longleftrightarrow \underline{d} \leq \underline{\delta}$$

On a bien démontré le théorème 8.3.0.1 de calcul de l'adhérence d'une orbite nilpotente.

Remarque(s) 8.3.3.1. *C'est pour cet argument de suite (et celui du paragraphe suivant) que le lecteur savant voulant généraliser à la topologie de Zariski des corps généraux utilisera l'hypothèse que le corps est infini.*

Passons au cas général.

8.4 Adhérence d'une orbite quelconque

Tout le travail a été fait pour ramener le cas général au cas nilpotent. Expliquons. On considère donc deux n-types $\underline{\mathrm{P}}, \underline{\mathrm{Q}}$ et on étudie l'inclusion $\mathrm{O}(\underline{\mathrm{P}}) \subset \overline{\mathrm{O}(\underline{\mathrm{Q}})}$. Autrement dit, on considère une suite de matrices A_m dans $\mathrm{O}(\underline{\mathrm{Q}})$ qui converge vers $\mathrm{A}_\infty \in \mathrm{O}(\underline{\mathrm{P}})$. On utilise alors librement notations et résultats de 7.3.1.

Par continuité du polynôme caractéristique, assure déjà que $\chi_{\mathrm{A}_m}, m \in \overline{\mathbf{N}}$ est un polynôme constant χ dont on note Λ l'ensemble de ses racines complexes. En particulier, les espaces caractéristiques de A_m ont une dimension constante d_λ : l'ordre de multiplicité de la racine λ de χ.
Ensuite (7.3.1.1), on a

$$\lim \mathrm{A}_m = \mathrm{A}_\infty \textit{ si et seulement si pour tout } \lambda \in \Lambda, \qquad \lim \mathrm{A}_{m,\lambda} = \mathrm{A}_{\infty,\lambda}$$

Mais, pour chaque λ, la matrice $\mathrm{A}_{m,\lambda} - \lambda\,\mathrm{Id} \in \mathrm{M}_{n,\mathbf{C}}$ est nilpotente et son n-type est (7.3.1.2) est

$$\underline{\delta}_\lambda = 1, \cdots, 1, (\mathrm{X} - \lambda)^{v_\lambda(\mathrm{Q})},\ i = d_\lambda, \cdots, 1 \text{ si } n < \infty$$

et

$$\underline{d}_\lambda = 1, \cdots, 1, (\mathrm{X} - \lambda)^{v_\lambda(\mathrm{P})},\ i = d, \cdots, 1 \text{ sinon}$$

où les 1 sont répétés $d_\lambda - v_\lambda(\chi)$-fois dans tous les cas. Mais d'après la caractérisation des orbites nilpotentes -partie condition nécessaire- (8.3.0.1), l'existence de cette suite de matrice entraîne

(i) $$\text{Pour tout } \lambda \in \Lambda,\ \underline{d}_\lambda \leq \underline{\delta}_\lambda$$

Inversement, supposons cette condition vérifiée. On note p_λ les projecteurs spectraux de A_∞ de type $\underline{\mathrm{P}}$. D'après la partie suffisante de la caractérisation de orbites nilpotentes (8.3.0.1, il existe pour tout λ des

matrices nilpotentes $\mathrm{N}_{m,\lambda}$ qui convergent vers $\mathrm{N}_{\infty,\lambda} = \mathrm{A}_{\infty,\lambda} - \lambda p_\lambda$. En posant $\mathrm{A}_m = \sum_\lambda (\mathrm{N}_{m,\lambda} + \lambda p_\lambda)$, on a $\lim \mathrm{A}_m = \mathrm{A}_\infty$. On a donc

$$\underline{\mathrm{P}} \preceq \underline{\mathrm{Q}} \Longleftrightarrow \text{pour tout } \lambda \in \Lambda,\ \underline{d}_\lambda \leq \underline{\delta}_\lambda.$$

Or, pour deux polynômes P, Q dont les racines sont dans Λ, on a

$$\mathrm{P}|\mathrm{Q} \Longleftrightarrow \text{pour tout } v_\lambda(\mathrm{P}) \leq v_\lambda(\mathrm{Q})$$

La condiotion (i équivaut donc à

$$\textit{pour tout } i = 1, \cdots, n, \text{ on a } \prod\nolimits_{j\leq i} \mathrm{P}_j | \prod\nolimits_{j\leq i} \mathrm{Q}_j$$

Ceci achève la preuve du théorème 8.2.0.2 .

8.5 Exercices supplémentaires

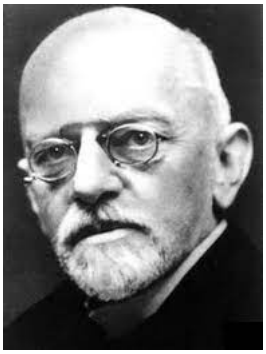
Exercice(s) 8.5.0.1. *Soit $\underline{\mathrm{Q}}$ un n-type et $\chi = \prod \mathrm{Q}_i$ le polynôme caractéristique correspondant.*

1. *Montrer que $\mathrm{O}(\underline{\chi}_{red})$ (cf. 5.3.4.1) est la seule orbite fermée contenue dans $\overline{\mathrm{O}(\underline{\mathrm{Q}})}$. En déduire que les orbites fermées sont les orbites semi-simples et que $\chi_{red}) = (\chi_n, \ldots, \chi_1)$ est un type minimal pour $\preceq$.*
2. *Montrer que l'adhérence de $\mathrm{O}(\underline{\chi}_{red}))$ est l'ensemble des matrices A telles que $\chi_1(\mathrm{A}) = 0$ et $\chi_\mathrm{A} = \chi$.*
3. *Montrez en général que les n-types minimaux sont de la forme $\underline{\chi}_{red}$ pour χ unitaire de dégré n. Pouvez-vous démonter ce résultat directement ?*
4. *Inversement, montrez que les n-types maximaux sont de la forme $(1, \ldots, 1, \chi)$. En déduire que les orbites maximales sont celles des matrices compagnons $\mathrm{C}(\chi)$.*
5. *Montrer que l'adhérence de $\mathrm{O}(\mathrm{C}(\chi))$ est l'ensemble des matrices A telles que $\chi_\mathrm{A} = \chi$.*

Exercice(s) 8.5.0.2. *Soit $\mathbf{k}$ un sous-corps de $\mathbf{C}$. On ne considère ici que des n-type $\mathbf{k}$-rationnels $\underline{d}$,* i.e. *vérifiant $\mathrm{P}_i \in \mathbf{k}[\mathrm{T}], i = 1, \ldots, n$. On note $\mathrm{O}_\mathbf{k}(\underline{d})$ la classe de conjugaison de $\mathrm{C}(\underline{d})$ sous $\mathrm{GL}_n(\mathbf{k})$. Montrer dans ce cas $\mathrm{O}_\mathbf{k}(\underline{\mathrm{P}}) = \mathrm{O}_\mathbf{C}(\underline{\mathrm{P}}) \cap \mathrm{M}_n(\mathbf{k})$. En utilisant 7.3.1.3 et le théorème principal 8.2.0.2, montrer $\overline{\mathrm{O}_\mathbf{k}(\underline{\mathrm{Q}})} = \cup_{\underline{\mathrm{P}} \leq \underline{\mathrm{Q}}} \mathrm{O}_\mathbf{k}(\underline{\mathrm{P}})$.*

Chapitre 9

Propriétés de finitude des modules



David Hilbert

Emmy Noether

9.1 Introduction

La notion d'anneau noethérien renvoie inévitablement au papier fondateur de Hilbert de 1890 [Hil90] avec ses trois grands théorèmes, le premier étant le théorème de transfert 9.3.3.1 (ou de la base de Hilbert) dans le cas des anneaux de polynômes. Toutefois, comme me l'a fait justement fait remarquer une étudiante, ne parler que de ce (formidable) article [1] est injuste. C'est en effet Emmy Noether qui en a dégagé la vision générale dès 1920 ([Noe21]).

9.2 Intégralité

En complément de l'importance des espaces vectoriels de dimension finie, montrons par quelques exemples l'importance des modules de type fini en général. Exemples qui seront largement nourris par la suite.

1. Les deux autres théorèmes de l'article sont rien moins que le Nullstellensatz et le théorème des syzygies !

9.2.1 Principe de prolongement des identités algébriques

Ce principe, extrêmement utile, est basé sur une trivialité. Soit $P \in \mathbf{Z}[T_1, \cdots, T_n]$ et $I_i, 1 = 1, \cdots, n$ des ensembles infinis d'un corps de caractéristique nulle k. Alors, si P est nul sur $\prod I_i$, pour tout anneau R et tout $(r_i) \in R^n$, on a $P(r_1, \cdots, r_n) = 0$. En effet, on observe qu'on a alors $\mathbf{Z}[T_1, \cdots, T_n] \subset k[T_1, \cdots, x_n]$ et on se ramène par récurrence au fait qu'un polynôme à 1 variable non identiquement nul n'a qu'un nombre fini de racines.

Corollaire 9.2.1.1. *Soit* $A \in M_n(R)$ *et* $\chi_A(T) = \det(TId - A)$. *On a alors,* $\chi_A(A) = 0$.

Démonstration. *L'équation matricielle* $\chi_A(A) = 0$ *est un système de* n^2 *équations polynomiales à coefficients entiers (les coefficients* $\chi_A(A) = 0$ *où* A *est la matrice générique* $A = [T_{i,j}]$*). Mais ces polynômes sont nuls sur* $M_n(\mathbf{C})$ *d'après Cayley-Hamilton usuel. On conclut en posant* $I_{i,j} = \mathbf{C}$ *grâce à la discussion précédente.*

9.2.2 Une application de Cayley-Hamilton

Proposition 9.2.2.1 (Ruse du déterminant)**.** *Soit* f *un endomorphisme d'un* R- *module de type fini* M *Il existe* $P \in R[T]$ *unitaire annulant* f. *Si de plus* $f(M) \subset IM$, *on peut supposer que les coefficients de* f *d'indice* $< \deg(P)$ *sont dans* I.

Démonstration. *Soit* m_i; $1 \leq \mathbf{i} \leq n$ *un famille finie de générateurs de* M *et considérons une matrice* $A = [a_{i,j}$ *de* f, i.e.*pour chaque* j*,écrivons (de manière non unique)*

$$f(m_j) = \sum_i a_{i,j} m_i.$$

Notons que si $f(M) \subset IM$, *on peut supposer* $a_{i,j} \in I$. *Il suffit alors de poser* $P = \det(TId - A)$ *et d'invoquer par exemple Cayley-Hamilton (9.2.1.1) pour* $A \in M_n(R)$.

En appliquant la proposition à $f = \mathrm{Id}_M$, on obtient le fameux lemme de Nakayama, très important en algèbre commutative (plutôt en M2).

Corollaire 9.2.2.2 (Nakayama)**.** *Soit* M *un module de type fini et* I *u n idéal tel que* $M = IM$. *Alors, il existe* $i \in I$ *tel que* $(1 + i)M = 0$. *En particulier, si* $1 + i$ *est inversible (eg si* i *est nilpotent),* $M = 0$.

9.2.3 Anneaux des entiers

Soit R′ une R-algèbre (autrement dit on considère un morphisme d'anneaux $R \to R'$). On dit que $r' \in R'$ est entier sur R s'il est annulé par un polynôme unitaire à coefficients dans R.

Théorème 9.2.3.1. *Le sous-ensemble de* R' *des éléments entiers sur* R *est un sous-anneau de* R'.

Démonstration. 0 *et* 1 *sont entiers. On doit donc prouver que la différence et le produits de deux éléments entiers* r' *et* $r"$ *sont entiers. Posons* $\mathrm{M} = \mathrm{R}[r', r"]$ *l'anneau des expressions polynomiales en* r' *et* $r"$ *à coefficients dans* R. *Si* r' *et* $r"$ *sont annulés par des polynômes unitaires de degrés* n' *et* $n"$, *la famille* $r'^i r"^j 1 \leq n', j \leq n"$ *engendre* M *et contient* $r' - r"$ *et* $r'r"$. *Mais si* $\rho \in \mathrm{M}$, *l'homotéthie de rapport* ρ *définit un endomorphisme* h_ρ *de* M *et donc (9.2.2.1) il existe* $\mathrm{P} \in \mathrm{R}[\mathrm{T}]$ *unitaire tel que* $\mathrm{P}(h_\rho) = h_{\mathrm{P}(\rho)} = 0$. *En appliquant à* $1 \in \mathrm{M}$, *on obtient* $\mathrm{P}(\rho) = 0$ *de sorte que tous les éléments de* M *son entiers sur* R.

Corollaire 9.2.3.2. *Soit* k *un sous-corps d'un corps* k'. *Alors le sous-ensemble des éléments de* k' *qui sont algébriques sur* k *est un sous corps de* k'.

Démonstration. *D'après 9.2.3.1 appliqué à* $\mathrm{R} = k$, *il suffit de montrer que l'inverse d'un élément algébrique* $r' \in k'$ *non nul est encore non nul. Soit donc* P *est un annulateur unitaire de* r'. *Mais alors,* $\mathrm{T}^{\deg(\mathrm{P})}\mathrm{P}(1/\mathrm{T})$ *est un annulateur non nul de* $1/r'$.

Exercice(s) 9.2.3.3.

1. *Montrer qu'un nombre rationnel est entier sur* $\mathbf{Z}$ *si et seulement si il est entier.*
2. *Montrer que le polynôme unitaire de degré minimal* $\mathrm{P} \in \mathbf{Q}[\mathrm{T}]$ *annulateur de* $\exp(\frac{2i\pi}{n})$ *est à coefficients entiers.*

On verra en TD la formule $\mathrm{P} = \Phi_k(\mathrm{T}) = \prod_{k \in (\mathbf{Z}/n\mathbf{Z})^\times} (\mathrm{T} - \exp(\frac{2ik\pi}{n}))$.

9.3 Modules noethériens

L'image d'une famille de générateurs d'un modules par un morphisme engendre le module image. Ainsi, *tout quotient d'un module de type fini est encore de type fini.* En revanche, si un sous-module d'un module R module de type fini est encore de type fini lorsque R est un corps, il n'en est rien en général (cf 2.2.4). C'est en revanche le cas dans le cas noethérien.

Lemme 9.3.0.1. *Soit* M *un* R *module. Les propriétés suivantes sont équivalentes.*

1. *Tout sous-module de* M *est de type fini.*
2. *Toute suite croissante de sous-modules stationne.*
3. *Toute famille non vide de sous-modules de* M *admet un élément maximal pour l'inclusion.*

Démonstration. $1 \Rightarrow 2$. *Soit* M_i *une suite croissante de sous-modules. Alors,* $\cup \mathrm{M}_i$ *est un sous-module de* M*, donc de type fini. Choisissons une famille finie de générateurs : pour* n *assez grand, ils appartiennent tous à* M_n *et donc* $\mathrm{M}_i = \mathrm{M}_n$ *si* $i \geq n$*.*
$2 \Rightarrow 3$. *Soit* $\mathcal{F}$ *une famille non-vide de sous-modules* M *sans élément maximal (preuve par contrapposée). On construit une suite strictement croissante d'éléments de* $\mathcal{F} \neq \varnothing$ *par récurrence en choisissant* M_0 *un de ses éléments arbitrairement puis par récurrence, supposant la suite construite pour* $i \leq n$*, on observe que* M_n *n'est pas maximal donc il existe* M_{n+1} *dans* $\mathcal{F}$ *qui le contient strictement.*
$3 \Rightarrow 1$. *Soit donc* N *un sous-module de* M *et soit* $\mathcal{F}$ *la famille de ses sous-modules de type fini. Comme* $\{0\} \in \mathcal{F}$*, cette famille est non vide. Soit* N' *un élément maximal. Il est de type fini contenu dans* N *par construction. Inversement, soit* $n \in \mathrm{N}$*. Le module* $\mathrm{R}n + \mathrm{N}'$ *est dans* $\mathcal{F}$ *et contient l'élément maximal* N' *: il lui est donc égal de sorte que* $n \in \mathrm{N}'$*. On a donc* $\mathrm{N}' = \mathrm{N}$ *et donc* N *de type fini.*

Définition 9.3.0.2. *1. Un module vérifiant les conditions équivalentes précédentes est dit noethérien.*
2. Un anneau qui est noethérien comme module sur lui-même est dit anneau noethérien.

On a donc R noethérien s'il vérifie l'une des trois propositions équivalentes suivantes

1. Tout idéal est de type fini.
2. Toute suite croissante d'idéaux stationne.
3. Toute famille non vide d'idéaux admet un élément maximal pour l'inclusion.

Exemple(s) 9.3.0.3. *Les sous-modules de modules noethériens sont noethériens (tautologique) ainsi que les quotients de modules noethériens (exercice facile). Les corps, les anneaux principaux, les anneaux quotients de d'anneaux noethériens sont noethériens. En revanche, un sous anneau d'un anneau noethérien n'est en général pas de noethérien (par exemple un anneau de polynômes sur un corp à une infinité de variables n'est pas noethérien alors que c'est un sous-anneau de son corps des fractions qui lui l'est !)*

9.3.1 Stabilité par suite exacte

Proposition 9.3.1.1. *Donnons-nous une suite exacte de modules*

$$0 \to \mathrm{M}_1 \xrightarrow{j} \mathrm{M}_2 \xrightarrow{p} \mathrm{M}_3 \to 0.$$

Alors M_2 *est noethérien si et seulement si* M_1 *et* M_3 *le sont.*

Démonstration. *La partoie directe a déjà été observée dans l'exemple précédent. Inversement, supposons* M_1 *et* M_3 *noethérien et soit* M_2' *un sous-module de* M_2*. On a une suite exacte*

$$0 \to j^{-1}(\mathrm{M}_2') \to \mathrm{M}_2' \to p(\mathrm{M}_2') \to 0.$$

Mais $j^{-1}(\mathrm{M}'_2)$ et $p(\mathrm{M}'_2)$ sont de type fini comme sous-modules de M_1 et M_3. On peut donc choisi une famille finie de générateurs de $p(\mathrm{M}'_2)$ de la forme $p'(g_{2,i})$ et une famille finie de générateurs $g_{1,j}$ de $j^{-1}(\mathrm{M}'_2)$. La famille finie $g_{1,j}, g'_{2,i}$ de M'_2 l'engendre.

En particulier, si R est noethérien, R^n est un module noethérien et donc il en de même de tout quotient. On en déduit le corollaire important suivant.

Corollaire 9.3.1.2. *Les modules noethériens sur un anneau noethérien sont les modules de type fini.*

9.3.2 Existence de décomposition en irréductibles

Rappelons que $r \in \mathrm{R}$ est dit irréductible s'il est non nul et non inversible d'une part et si ses seuls diviseurs sont soit inversibles soit lui sont associés. Autrement dit, $r \in \mathrm{R}^*$ est irréductible si l'équation $r = r_1 r_2$ entraîne r_1 ou r_2 inversible.

Lemme 9.3.2.1. *Tout élément non nul et non inversible dans un anneau noethérien* R *est produit d'éléments irréductibles.*

Démonstration. *Remarquons que le fait que r soit irréductible ne dépend que de (r) ie est invariant par multiplication par un inversible. Soit alors $\mathcal{F}$ l'ensemble des idéaux principaux propres et non nuls de* R *dont un des générateurs n'est pas produit d'irréductible. Si $\mathcal{F}$ était non vide, il admettrait un élément maximal (r) pour l'inclusion. Mais r n'est pas irréductible car sinon $(r) \notin \mathcal{F}$ de sorte que r s'écrit $r_1 r_2$ avec r_1 et r_2 non inversibles/ Ainsi $(r) \subsetneq (r_i)$. Par maximalité, $(r_i) \notin \mathbf{F}$ de sorte que chaque r_i est produit d'irréductible, et de même pour leur produit r. Une contradiction.*

Ainsi, l'existence de décomposition en irréductibles est banale. C'est l'unicité à multiplication par inversible près (et ordre près) qui est importante (comme on le verra c'est exactement le contenu du lemme d'Euclide dans les anneaux factoriels). Par exemple, d'après ce qui précède, l'anneau $\mathbf{R}[\mathrm{T}, \mathrm{Y}]/(\mathrm{T}^2 - \mathrm{Y}^3)$ est noethérien, visiblement intègre (exercice). Pourtant, l'élement $\mathrm{T}^2 = \mathrm{Y}^3$ du quotient a bien deux décompositions (non équivalentes) car tant T que Y sont irréductibles dans le quotient et non associés (exercice).

9.3.3 Le théorème de transfert de Hilbert

Théorème 9.3.3.1. *Soit* R *un anneau noethérien.*

1. *L'anneau de polynômes* $\mathrm{R}[\mathrm{T}]$ *est noethérien.*
2. *Toute* R*-algèbre de type fini est un anneau noethérien.*

Démonstration. *Le second point est une conséquence immédiate du premier (par récurrence, tout anneau de polynômes sur* R *à* n *variables est noethérien, et donc itou pour tout quotient). Passons au premier point.*

Soit I *un idéal de* $\mathrm{R}[\mathrm{T}]$ *et* $\mathrm{I}* = \mathrm{I} - \{0\}$. *Si* P *est un polynôme non nul, on note* $\mathrm{dom}(\mathrm{P})$ *son coefficient de plus haut degré non nul. La formule* $\mathrm{dom}(\mathrm{T}^n\mathrm{P}) = \mathrm{dom}(\mathrm{P})$ *assure que* $\{0\} \cup \mathrm{dom}(\mathrm{I}^*)$ *est un idéal de* R *(exercice). Il admet donc un nombre fini de générateurs de la forme* $\mathrm{dom}(\mathrm{P}_i), \mathrm{P}_i \in \mathrm{I}^*$ *qu'on peut supposer de même degré* $d \geq 0$ *d'après la formule précédente. Une récurrence immédiate montre alors* $\mathrm{I} \cap \mathrm{R}_{\geq d}[\mathrm{T}] = \langle \mathrm{P}_i \rangle$. *Mais* $\mathrm{I} \cap \mathrm{R}_{\leq d}[\mathrm{T}]$ *est un sous-*R*-module de* $\mathrm{R}_{<d}[\mathrm{T}] \simeq \mathrm{R}^d$ *: c'est donc un module noethérien comme* R^d *(9.3.1.2). On peut donc en prendre un nombre fini de générateurs* Q_j *(comme* R*-module) et la famille finie* $(\mathrm{P}_i, \mathrm{Q}_j)$ *engendre* I. ∎

On a en fait repris l'argument de division euclidienne utilisé pour montrer que $\mathbf{k}[\mathrm{T}]$ est principal, le problème étant qu'on ne peut diviser dans $\mathbf{k}[\mathrm{T}]$ que si le coefficient dominant du polynôme est un inversible de $\mathrm{R}^\times$. C'est la raison qui pousse à introduire les idéaux de coefficients dominants de I.

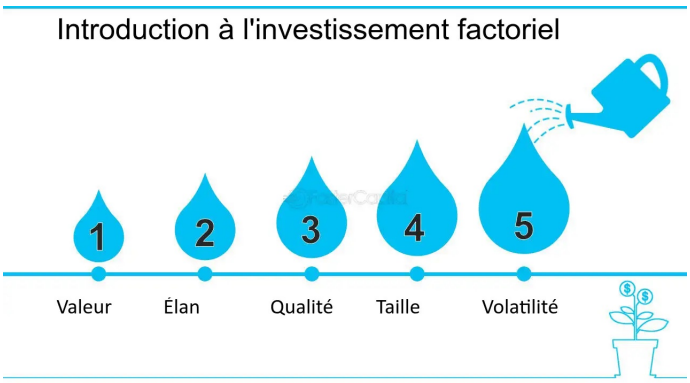
9.4 Exercices

Exercice(s) 9.4.0.1. *Soit* G *un groupe fini opérant (à gauche) sur un anneau* R. *On suppose que le cardinal* n *de* G *est inversible dans* R *et on note* R^G *le sous-anneau de* R *des éléments invariants par* G. *On note* $\pi : \mathrm{R} \to \mathrm{R}$ *l'application* $x \mapsto \frac{1}{n}\sum_{g\in\mathrm{G}} gx$.

1. *Montrer que* p *est un projecteur d'image* R^G.
2. *Montrer* p *est* R^G *linéaire.*
3. *Montrer que si* R *est noethérien,* R^G *est noethérien.*

Chapitre 10

Rappels sur les anneaux factoriels



10.1 Introduction

Dans ce chapitre, R désigne un anneau *intègre* (commutatif unitaire comme toujours). On notera **k** son corps des fractions. On a donc une notion de divisibilité qui définit une relation d'ordre partielle sur $\mathrm{R}^* = \mathrm{R} - \{0\}$. On s'intéresse aux anneaux pour lesquels il existe une décomposition en facteurs premiers existe et est raisonnablement unique. On utilisera sans cesse le fait que deux éléments $a, b \in \mathrm{R}$ sont égaux à un multiple inversible $u \in \mathbf{R}^\times$ près si et seulement si les idéaux qu'ils engendrent sont égaux : on dit alors qu'ils sont *associés* et on écrit $a \sim b$. Ceci définit une relation d'équivalence sur R compatible au produit de sorte que $\mathrm{R}/\sim$ est muni d'une multiplication associative avec une unité, la classe de 1 : c'est ce qu'on appelle un monoïde (commutatif unitaire).

10.2 Caractérisation

Rappelons que $r \in \mathrm{R}$ est dit irréductible s'il est non nul et non inversible d'une part et si ses seuls diviseurs sont soit inversibles soit lui sont associés. Autrement dit, $r \in \mathrm{R}^*$ est irréductible si l'équation $r = r_1 r_2$ entraîne r_1 ou r_2 inversible.

10.2.1 Critère d'unicité

On sait que les entiers positifs p qui sont irréductibles sont les nombres premiers. De manière générale, on a

Lemme 10.2.1.1. *Soit $r \in \mathrm{R}^*$. Alors si l'idéal (r) est premier, r est irréductible.*

Démonstration. *Si $r = r_1 r_2$, le produit $r_1 r_2$ est nul dans $\mathrm{R}/(r)$ qui par définition est intègre. La classe $(r_1 \mod r)$ par exemple est donc nulle de sorte que $r = \rho_1 r$ et $r = \rho_1 r r_2$. En simplifiant par r (intégrité), on a r_2 inversible.* ∎

La réciproque est la propriété d'Euclide, qui elle n'et essentiellement vraie que dans les anneaux factoriels (voir ? ? ? pour un énoncé précis).

Définition 10.2.1.2 (Lemme d'Euclide)**.** *On dit (par abus) que le lemme d'Euclide est vrai dans* R *si l'idéal engendré par un irréductible est premier, autrement dit si tout irréductible divisant un produit divise l'un des facteurs.*

On sait bien que tout anneau principal vérifie le lemme d'Euclide (conséquence immédiate du lemme de Bézout -exercice-).

Définition 10.2.1.3. *On dira qu'une décomposition*

$$r = u \prod_{i=1}^{r} p_i$$

avec $u \in \mathrm{R}^\vee$ et p_i irréductible est unique *si pour toute autre telle décomposition*

$$r = u' \prod_{i=1}^{r'} p'_i,$$

on a $r = r'$ et, à renumérotation près*, $p_i \sim p'_i$ pour tout i. On dit aussi que les (classes des) p_i sont uniques à l'ordre près. Un anneau est dit factoriel (unique factorization domain en anglais) si tout élément non nul admet une unique décomposition en produit d'irréductibles au sens précéden*[1].

1. Un inversible a comme décomposition lui même multiplié par un produit vide d'irréductibles, un produit vide étant égal à 1.

Le lien avec ce qui précède est

Lemme 10.2.1.4 (Lemme d'unicité)**.** *Supposons tout élément non inversible de* R *admette une décomposition en produit d'irréductible. Alors, ces décompositions sont uniques si et seulement si le lemme d'Euclide est vrai dans* R*. C'est le cas pour un anneau principal.*

Démonstration. *Supposons qu'on a l'unicité et soit r irréductible (donc non nul). Si $r = r_1 r_2$, on écrit les décompositions en irréductibles $r_i = u_i \prod_j p_{i,j}$ de sorte que $r = u_1 u_2 \prod_{i,j} p_{1,i} p_{2,j}$. On a donc deux décompositions de r en irréductibles, l'une n'ayant qu'un facteur, lui-même ! Ainsi, par unicité, r est l'un des $p_{i,j}$ (et même le seul) de sorte qu'il divise r_i.*
Inversement, si le lemme d'Euclide est vrai, on procède par une récurrence sans mystère sur la somme des longueurs de deux éventuelles décomposition d'un même élément non nul. ∎

En invoquant l'existence des décompositions dans le cas noethérien (9.3.2.1), on a

Corollaire 10.2.1.5. *Un anneau noethérien intègre est factoriel si et seulement si il vérifie le lemme d'Euclide. C'est le cas pour un anneau principal.*

10.3 Transfert

Nous allons démontrer le théorème de transfert de la factorialité aux anneaux de polynômes

Théorème 10.3.0.1. *Si* R *est factoriel, alors* R[T] *est factoriel.*

On doit donc démontrer l'unicité des décomposition (donc le lemme d'Euclide) et leur unicité. Pour cela on va comparer la notion d'irréductibles dans R[X] et $\mathbf{k}$[X] en utilisant l notion de contenu (due à Gauss). On utilisera l'égalité $(\mathrm{R}[\mathrm{T}])^\vee = \mathrm{R}^\vee$ (exercice, utiliser le degré des polynômes).

10.3.1 Rappels : PGCD, PPCM

Soit (r_i) un famille finie d'éléments de R qu'on supposera non identiquement nulle. Rappelons qu'un élément $r \in \mathrm{R}^*$ est un PGCD des r_i s'il est maximal parmi les diviseurs communs aux r_i. Deux PGCD d'une même famille, lorsqu'ils existent, sont bien entendus associés, raison pour laquelle on parle « du fg PGCD. On peut donc considérer les PGCD, PPCM comme des éléments du monoïde $\mathrm{R}/\sim$. En considérant des multiples communs maximaux, on obtient la notion de PPCM. Comme sur les entiers, on a

Lemme 10.3.1.1. *Si* R *est factoriel, le* PGCD *et le* PPCM *des* (r_i) *existe.*

Démonstration. *Considérons des décompositions en facteurs irréductibles de chacun des* $r_i \neq 0$ *et soit* q_j *une famille d'irréductibles deux à deux non associés tels que tous ces facteurs soient associés à exactement un des* p_i*. On peut alors écrire de manière unique*

$$r_i = u_i \prod_j q_j^{v_{i,j}},\ v_{i,j\geq 0} \text{ et } u_i \in \mathrm{R}^\vee.$$

On pose alors

$$\mathrm{PGCD}(r_i) = \prod_j q_j^{\min_i v_{i,j}} \text{ et } \mathrm{PGCD}(r_i) = \prod_j q_j^{\max_i v_{i,j}}$$

dont on vérifie qu'ils conviennent. ■

Notons que PGCD et PPCM sont homogènes de poids 1 pour la multiplication par R^*.

Exercice(s) 10.3.1.2. *Montrer que si* R *est principal,* $\mathrm{PGCD}(r_i)$ *est une générateur de l'idéal engendré par les* (r_i)*. Donner une caractérisation du* PPCM *en termes d'idéaux.*

10.3.2 Contenu

Dans le reste de ce chapitre section, R désigne un anneau factoriel.

Définition 10.3.2.1. *Soit* $\mathrm{P} \in \mathrm{R}[\mathrm{T}]$ *non nul. Le contenu* $c(\mathrm{P}) \in \mathrm{R}/\sim$ *de* P *le* PGCD *de ses coefficients. Un polynôme de contenu* $c(\mathrm{P}) = 1$ *est dit primitif.*

Par exemple, les polynômes unitaires de R[T] sont primitifs. Le contenu est homogène de poids 1 sous R^* comme le PGCD
þ[Gauss] Soient $\mathrm{P}, \mathrm{Q} \in \mathrm{R}[\mathrm{T}]$ tous deux non nuls. Alors, $c(\mathrm{PQ}) = c(\mathrm{P})c(\mathrm{Q})$.

Démonstration. *Par homogénéité, on peut supposer* P, Q *primitifs et on doit démontrer que* PQ *est primitif. Sinon, soit* p *un irréductible de* R *divisant* $c(\mathrm{PQ})$*. Comme* R *est factoriel, il vérifie le lemme d'Euclide et le quotient* $\overline{\mathrm{R}} = \mathrm{R}/(p)$ *est intègre. Le morphisme* $\mathrm{R} \to \overline{\mathrm{R}}$ *de réduction des coefficients induit un morphisme d'anneaux* $\mathrm{R}[\mathrm{T}] \to \mathrm{R}/(p)$ *de sorte qu'on a* $0 = \overline{\mathrm{PQ}} = \overline{\mathrm{P}} \cdot \overline{\mathrm{Q}}$*. Puisque* $\overline{\mathrm{R}}[\mathrm{T}]$ *est intègre comme* $\overline{\mathrm{R}}$*, on a par exemple* $\overline{\mathrm{P}} = 0$*,* i.e.$p|c(\mathrm{P})$ *(avec un petit abus d'écriture), une contradiction car* $c(\mathrm{P}) = 1$*.* ■

Corollaire 10.3.2.2. *Les irréductibles de* R[T] *sont*

1. *Les irréductibles de* R *;*
2. *Les polynômes primitifs de* R[T] *qui sont irréductibles dans* $\mathbf{k}[\mathrm{X}]$*.*

Démonstration. *Rappelons l'égalité* $(\mathrm{R}[\mathrm{T}])^* = \mathrm{R}^\vee$. *Le premier point en découle immédiatement pour des raisons de degré.*
Si P *est irréductibles dans* R[T] *de degré* > 0, *il est certainement primitif d'après le premier point.*
Supposons donc qu'il soit le produits de deux polynômes $\tilde{\mathrm{P}}_1, \tilde{\mathrm{P}}_2 \in \mathbf{k}[\mathrm{T}]$. *En réduisant à un même dénominateur* $d_i \in \mathrm{R}^*$ *les coefficients de* $\tilde{\mathrm{P}}_i$, *on peut écrire* $\tilde{\mathrm{P}}_i = \mathrm{P}_i/d_i$ *avec* $\mathrm{P}_i \in \mathrm{R}[\mathrm{X}]$. *On a donc*

$$(*) \qquad d_1 d_2 \mathrm{P} = \mathrm{P}_1 \mathrm{P}_2$$

de sorte que $d_1 d_2 = d_1 d_2 c(\mathrm{P}) = c(\mathrm{P}_1)c(\mathrm{P}_2)$ *(homogénéité et multiplicativité du contenu). En remplaçant dans (*), on obtient*

$$\mathrm{P} = \mathrm{P}_1/c(\mathrm{P}_1)\mathrm{P}_2/c(\mathrm{P}_2)$$

avec $\mathrm{P}_i/c(\mathrm{P}_i) \in \mathrm{R}[\mathrm{T}]$ *par définition du contenu. Comme* P *est irréductible dans* R[T], *on en déduit par exemple* $\mathrm{P}_1/c(\mathrm{P}_1)$ *inversible, donc de degré zéro, et donc de même pour* $\tilde{\mathrm{P}}_1$ *qui lui est proportionnel par un scalaire D'où l'irréductibilité dans* $\mathbf{k}[\mathrm{T}]$.
La réciproque est tautologique (qui peut le plus peut le moins) ∎

10.3.3 Le théorème de transfert

Théorème 10.3.3.1. *Si* R *est factoriel, alors* R[T] *est factoriel (par exemple si* R *est principal).*

Démonstration. *Existence de la décomposition. Soit* $\mathrm{P} \in \mathrm{R}[\mathrm{X}]$ *non nul. Si* P *est une constante* $r \in \mathrm{R}^*$, *on écrit la décomposition* $r = \prod p_i$ *en facteurs irréductibles dans* R *et on invoque (10.3.2.2).*
Si P *est de degré* > 0, *quitte à factoriser par un* PGCD *de ses coefficients, on peut supposer* P *primitif. Comme dans la preuve de 10.3.2.2, un argument de dénominateur commun permet alors écrire sa décomposition dans l'anneau principal donc factoriel* $\mathbf{k}[\mathrm{X}]$

$$\mathrm{P} = \prod \mathrm{P}_i/d_i$$

avec $\mathrm{P}_i \in \mathrm{R}[\mathrm{T}]$ *irréductible dans* $\mathbf{k}[\mathrm{T}]$ *et* $d_i \in \mathrm{R}^*$. *En prenant les contenus, on a* $c(\mathrm{P}) = \prod d_i$ *et* $\mathrm{P} = \prod \mathrm{P}_i/c(\mathrm{P}_i)$ *qui est la décomposition cherchée (toujours avec un petit abus d'écriture).*
Unicité de la décomposition dans R[T]. *Montrons que* R[T] *vérifie le lemme d'Euclide (10.2.1.2). Soit donc* P *irréductible divisant le produit de* $\mathrm{P}_1, \mathrm{P}_2 \in \mathrm{R}[\mathrm{T}]$. *Si* P *est de degré* > 0, *il est primitif et irréductible dans* $\mathbf{k}[\mathrm{T}]$ *d'après (10.3.2.2). Comme* $\mathbf{k}[\mathrm{T}]$ *est factoriel puisque principal,* $\mathrm{P}|\mathrm{P}_1$ *par exemple (dans* $\mathbf{k}[\mathrm{T}]$ *)et un argument de dénominateur commun permet d'écrire une fois de plus* $d\mathrm{P}_1 = \mathrm{Q}_1 \cdot \mathrm{P}$ *avec* $d \in \mathrm{R}^*, \mathrm{Q}_1 \in \mathrm{R}[\mathrm{T}]$. *En prenant les contenus on a de nouveau* $dc(\mathrm{P}_1) = c(\mathrm{Q}_1)$ *et donc* $\mathrm{P}_1 = c(\mathrm{P}_1)\mathrm{Q}_1/c(\mathrm{Q}_1)\mathrm{P}$ *et donc* P *divise* P_1 *dans* R[T]. ∎

Par exemple, un anneau de polynômes à n variables sur un corps, un anneau principal plus généralement est factoriel. Mais attention, cette remarquable stabilité de la factorialité ne passe pas aux quotients comme le fait la propriété d'être noethérien. Le lecteur savant la reliera à la notion de non singularité en géométrie.

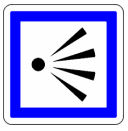
Exercice(s) 10.3.3.2. *Montrer que l'anneau* $\mathbf{R}[\mathrm{X},\mathrm{Y}]/(\mathrm{X}^2-\mathrm{Y}^3)$ *est intègre, noethérien mais pas factoriel.*

Chapitre 11

Formes bilinéaires et sesquilinéaires

Charles Hermite

11.1 Point de vue



On aborde les notions de forme bilinéaire et sesquiliénaire, sans symétrie *a priori*, ce d'un point de vue unifié même si ce sont les cas symétriques ou hermitiens qui nous occuperont rapidement. La raison de ce choix est qu'elles interviennent au-delà de la géométrie euclidienne ou hermitienne complexe, en géométrie projective par exemple et que ceci n'apporte aucune difficulté pour les notions de base.

11.2 Introduction

Comme dans le cas linéaire, deux points de vue équivalents -morphismes et matrices- se complètent de manière fructueuse donnant lieu naturellement à une nouvelle relation d'équivalence sur les matrices carrées, la congruence. Comme on le verra, il n'y a aucun espoir, mis à part le cas alterné (13.4), de décrire les classes de congruence de manière unifiée sur un corps quelconque comme dans le cas de la relation de similitude car c'est impossible déjà en dimension 1 : un forme bilinéaire à congruence près étant un scalaire à multiplication par un carré non nul près !

11.2.1 Notations et rappels

— Dans ce chapitre, E est un espace vectoriel[1] sur $\mathbf{k}$ de dimension finie n (même si la plupart des définitions formelles se généralisent sans cette hypothèse comme le lecteur s'en convaincra aisément).

— On note $\mathrm{E}^\star$ le dual de E et, si $\mathcal{B}$ est une base de E, on notera $\mathcal{B}^\star$ la base duale associée lorsque E est supposé de dimension finie.

— On rappelle que si $\mathcal{B} = (\mathrm{e}_i)$ et $\mathcal{C} = (\mathrm{e}'_i)$ sont des bases de E (de dimension finie), les colonnes de la matrice de passage $\mathrm{P} = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\mathrm{Id}_\mathrm{E})$ sont les coordonnées des vecteurs de $\mathcal{C}$ dans la base $\mathcal{B}$. Si $\mathrm{X} = [x]_\mathcal{B}$ (resp. $\mathrm{X}' = [x]_\mathcal{C}$) sont les coordonnées de $x \in \mathrm{E}$ dans $\mathcal{B}$ (resp. dans $\mathcal{C}$), on a donc $\mathrm{X} = \mathrm{PX}'$.

— La matrice de passage La matrice de passage $\mathrm{P}^\vee = \mathrm{Mat}_{\mathcal{B}^\star,\mathcal{C}^\star}$ vérifie ${}^t\mathrm{P}^\vee \cdot \mathrm{P} = \mathrm{Id}$.

11.3 Formes bilinéaires/sesquilinéaires

11.3.1 Formes bilinéaires

Définition 11.3.1.1. *Une forme bilinéaire est une application $b : \mathrm{E} \times \mathrm{E} \to \mathbf{k}$ telle que, pour tout $y \in \mathrm{X}$, $b_1 : x \mapsto b(x, y)$ est linéaire et pour tout $x \in \mathrm{X}$, $b_2 : y \mapsto b(x, y)$ est linéaire.*
On dit que b est symétrique si $b(x, y) = b(y, x)$, antisymétrique si $b(x, y) = -b(y, x)$, alternée si $b(x, x) = 0$ pour tous x, y et ε-symétrique si elle est symétrique ou antisymétrique.
Si E *est de dimension finie et si $\mathcal{B} = \{\mathrm{e}_i\}_{1\leq i\leq n}$ est une base, on définit* $\mathrm{Mat}(b, \mathcal{B}) = (b(\mathrm{e}_i, \mathrm{e}_j))_{i,j}$. *La base $\mathcal{B}$ est dite orthogonale(resp. orthonormée) si* $\mathrm{Mat}(b, \mathcal{B})$ *est diagonale (resp. l'identité).*

Dire que b est symétrique (resp. alternée), c'est dire que sa matrice est symétrique (resp. alternée, ie antisymétrique à coefficients diagonaux nuls). Si X et Y sont les coordonnées de vecteurs x et y dans la base $\mathcal{B}$, on note X' et Y' leurs coordonnées dans $\mathcal{C}$ une autre base et P la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$. Les formules $\mathrm{X} = \mathrm{PX}'$ et $\mathrm{Y} = \mathrm{PY}'$ donnent alors

(i)
$$\boxed{\begin{array}{rcl} b(x, y) & = & {}^t\mathrm{X} \cdot \mathrm{Mat}(b, \mathcal{B}) \cdot \mathrm{Y} \\ \mathrm{Mat}(b, \mathcal{C}) & = & {}^t\mathrm{P}\,\mathrm{Mat}(b, \mathcal{B})\mathrm{P} \end{array}}$$

Ainsi, grâce aux formules précédentes (i), le choix d'une base permet d'une base d'identifier b avec sa matrice M **dans $\mathcal{B}$, ce qu'on fera librement, les propriétés de ε-symétrie se lisant sur la propriétés matricielles analogues . On remarque que le changement de base ne correspond plus à la similitude comme mais à la congruence des matrices** $\mathrm{M} \mapsto {}^t\mathrm{PMP}$**. En particulier,** b

1. On a choisi E comme Euclide plutôt que V comme vecteur pour distinguer le contexte quadratique/hermitien du contexte vectoriel général.

est invariante par $a \in \mathrm{End}_k(\mathrm{E})$ **de matrice** $\mathrm{A} = \mathrm{Mat}(\mathcal{B}, a)$ **si et seulement si**

$$^t\mathrm{A}\,\mathrm{Mat}(b, \mathcal{B})\mathrm{A} = \mathrm{Mat}(b, \mathcal{B}).$$

Remarque(s) 11.3.1.2. *Pour tout corps,* $\mathbf{k}$*,* b *alternée implique* b *antisymétrique. Réciproquement,* b *antisymétrique implique* b *alternée si* $car(\mathbf{k}) \neq 2$ *; si* $car(\mathbf{k}) = 2$*, les notions de symétrie et d'antisymétrie coïncident.*

On a alors la notion d'espace bilinéaire (espace vectoriel muni d'une forme bilinéaire b, un morphisme entre un espaces bilinéaires étant une application linéaire préservant les applications bilinéaires, un sous-espace vectoriel définissant un (sous)-espace bilinéaire par restriction de $b, \dots$). Matriciellement, ceci signifie que la matrice de b dans une base adaptée à la décomposition est diagonale.

11.3.2 Généralisation sesquilinéaire

On suppose de plus $\mathbf{k}$ muni d'une involution de corps σ.

Définition 11.3.2.1. *On dit que* $b_\sigma : \mathrm{E} \times \mathrm{E} \to \mathbf{k}$ *est* σ*-sesquilinéaire si* b_σ *est linéaire en la première variable, additive en la seconde et si* $b_\sigma(x, \lambda y) = \sigma(\lambda) b_\sigma(x, y)$*. Une forme* σ*-sesquilinéaire est dite hermitienne si si* $b_\sigma(y, x) = \sigma(b_\sigma(x, y))$*, anti-hermitienne*$b_\sigma(x, y) = -b_\sigma(y, x)$ *pour tous* x, y *et* ε*-hermitienne si elle est hermitienne ou antihermitienne.*
Si E *est de dimension finie et si* $\mathcal{B} = \{\mathrm{e}_i\}_{1 \leq i \leq n}$ *est une base, on définit* $\mathrm{Mat}(b_\sigma, \mathcal{B}) = (b_\sigma(\mathrm{e}_i, \mathrm{e}_j))_{i,j}$*. La base* $\mathcal{B}$ *est dite orthogonale (resp. orthonormée) si* $\mathrm{Mat}(b_\sigma, \mathcal{B})$ *est diagonale (resp. l'identité).*

Les formules i deviennent

(ii)
$$\boxed{\begin{array}{rcl} b_\sigma(x, y) & = & {}^t\mathrm{X} \cdot \mathrm{Mat}(b_\sigma, \mathcal{B}) \cdot \sigma(\mathrm{Y}) \\ \mathrm{Mat}(b_\sigma, \mathcal{C}) & = & {}^t\mathrm{P}\,\mathrm{Mat}(b_\sigma, \mathcal{B})\sigma(\mathrm{P}) \end{array}}$$

Ainsi, grâce aux formules précédentes (i), le choix d'une base permet d'une base d'identifier b_σ **avec sa matrice dans** $\mathcal{B}$**, ce qu'on fera librement, les propriétés « ε-hermitiennes » se lisant sur la propriétés matricielles analogues . On remarque que le changement de base ne correspond plus à la similitude comme mais à la $\star$-congruence des matrices** $\mathrm{M} \mapsto {}^t\mathrm{PM}\sigma(\mathrm{P})$.
Le lecteur généralisera la notion d'espace bilinéaire au cas sesquilinéaire
Ainsi, le cas bilinéaire est le cas particulier sesquilinéaire avec $\sigma = \mathrm{Id}$. Il est à noter que le déterminant de la matrice de b_σ est défini à une norme $\mathrm{N}(\lambda) = \lambda\sigma(\lambda)$ non nulle près (un carré non nul dans le cas bilinéaire). Cet élément de l'ensemble quotient $\mathbf{k}/\mathrm{N}(\mathbf{k}^*)$ s'appelle le discriminant de b_σ.

1. Le cas le plus important pour nous étant la conjugaison complexe.

11.3.3 Formes non dégénérées

Le lecteur intéressé uniquement par le cas bilinéaire fera $\sigma = \mathrm{Id}$ et $\mathrm{E} = \mathrm{E}_\sigma$ dans ce qui suit.

On note E_σ le l'espace vectoriel dont le groupe sous-jacent est E dont la multiplication externe $.\sigma$ est tordue par σ, *i.e.* $\lambda._\sigma x = \sigma(\lambda)v$. Une base $\mathcal{B}$ de E est encore est base notée $\mathcal{B}_\sigma$ de E_σ de sorte que E et E_σ ont même dimension. Une application σ-linéaire de E dans E′ s'identifie, au choix -c'est l'avantage d'avoir une involution- comme un élément de $\mathrm{Hom}_{\mathbf{k}}(\mathrm{E}, \mathrm{E}'_\sigma) = \mathrm{Hom}_{\mathbf{k}}(\mathrm{E}_\sigma, \mathrm{E}')$. En particulier, le dual tordu $\mathrm{E}^\star_\sigma = \mathrm{Hom}_{\mathbf{k}}(\mathrm{E}_\sigma, \mathbf{k})$ est l'espace des formes σ-linéaires avec une base σ-duale $\mathcal{B}^\star_\sigma$ qui est la base duale de $\mathcal{B}_\sigma$.

On peut associer à b_σ les applications linéaires

(iii) $$\check{b}_\sigma : \left\{ \begin{array}{ccc} \mathrm{E} & \to & \mathrm{E}^\star_\sigma \\ y & \mapsto & (x \mapsto b_\sigma(x,y)) \end{array} \right. \quad \text{et} \quad \widehat{b}_\sigma : \left\{ \begin{array}{ccc} \mathrm{E} & \to & \mathrm{E}^\star_\sigma \\ x & \mapsto & (y \mapsto \sigma(b_\sigma(x,y))) \end{array} \right.$$

On remarque que $\sigma(\mathrm{Mat}(b_\sigma, \mathcal{B})) = \mathrm{Mat}(\check{b}_\sigma, \mathcal{B}, \mathcal{B}^\star_\sigma)$ et ${}^t\mathrm{Mat}(b_\sigma, \mathcal{B}) = \mathrm{Mat}(\widehat{b}_\sigma, \mathcal{B}, \mathcal{B}^\star_\sigma)$.

Définition 11.3.3.1. *On définit le noyau (à gauche)* $\mathrm{Ker}\, b_\sigma$ *de* b_σ *par*

$$\mathrm{Ker}\, b_\sigma = \mathrm{Ker}\, \widehat{b}_\sigma = \{x \in \mathrm{E},\ \forall\, y \in \mathrm{E},\ b_\sigma(x,y) = 0\}.$$

On dira qu'une forme bilinéaire b_σ *est non dégénérée si son noyau est nul* i.e. *si sa matrice dans une base* $\mathcal{B}$ *est inversible.*

Une forme non dégénérée identifie donc E et son dual $\mathrm{E}^\star_\sigma$ grâce à iii).

Remarque(s) 11.3.3.2. *Du point de vue matriciel, on a* $\mathrm{Ker}(b_\sigma) = \mathrm{Ker}(\sigma(\mathrm{Mat}(b_\sigma, \mathcal{B}))) = \mathrm{Ker}(\mathrm{Mat}(b_\sigma, \mathcal{B}))$.

- *Si on avait échangé les rôles de* x *et* y, i.e. *utilisé* $\widehat{b_\sigma}$ *pour définir le noyau (à droite), celui-ci aurait été celui de la transposée* ${}^t\mathrm{Mat}(b_\sigma, \mathcal{B})$. *La notion de dégénérescence n'aurait donc pas changé.*
- *Si maintenant* b_σ *est de plus supposé* ε*-hermitienne, la notion de noyau à droite et à gauche coïncide est aussi la même puisqu'alors* ${}^t\mathrm{Mat}(b, \mathcal{B}) = \pm\,\mathrm{Mat}(b, \mathcal{B})$.
- *Dans le cas non dégénéré,* ${}^t\widehat{b_\sigma}$ *est un isomorphisme : on définit alors traditionnellement l'*asymétrie

$$\beta = {}^t\widehat{b}_\sigma^{-1} \circ \check{b}_\sigma \in \mathrm{End}_{\mathbf{k}}(\mathrm{E}_\sigma) = \mathrm{End}_{\mathbf{k}}(\mathrm{E})$$

de la forme non dégénérée b_σ. *Sa matrice*[2] *est simplement* ${}^t\mathrm{Mat}(b_\sigma, \mathcal{B})^{-1}\sigma(\mathrm{Mat}(b_\sigma, \mathcal{B}))$. *C'est l'unique isomorphisme vérifiant*

$$b_\sigma(y,x) = b_\sigma(\beta(x), y)$$

pour tout $x, y \in \mathrm{E}$.

- *La notion d'orthogonalité d'une forme sans symétrie hermitienne est délicate. On réserve la notion de somme directe orthogonale de deux-sous espaces* $\mathrm{E}_1, \mathrm{E}_2$ *au cas où les sous-espaces sont*

d'une part en somme directe et orthogonaux à droite et à gauche d'autre part, ie $b_\sigma(\mathrm{E}_1, \mathrm{E}_2) = b_\sigma(\mathrm{E}_2, \mathrm{E}_1) = \{0\}$. *On écrit alors* $\mathrm{E}_1 \overset{\perp}{\oplus} \mathrm{E}_2$. *Matriciellement, dans une base adaptée à la somme directe, ceci siginifie que la matrice de* b_σ *est diagonale par blocs et que la forme est non dégénérée si et seulement les blocs le sont.*

Si b_σ est dégénérée, alors la matrice associée dans une base adaptée à la somme directe du noyau et d'un supplémentaire (arbitraire) a $\dim \operatorname{Ker} b_\sigma$ colonnes nulles. Si de plus elle est ε-hermitienne, les lignes correspondantes sont également nulles de sorte que b_σ passe au quotient en une forme non dégénérée (exercice) :

Toute forme ε-hermitienne b_σ définit une forme sesquilinéaire non dégénérée sur $\mathrm{E}/\operatorname{Ker}(b_\sigma)$.

Si la forme est quelconque, *i.e.* pas ε-hermitienne donc, ce n'est plus le cas (cf. 12).

11.3.4 Adjoint

La proposition suivante est simple mais est importante.

Proposition 11.3.4.1. *Soit* b_σ *une forme sesquilinéaire* ***non dégénérée*** *et* f *un endomorphismes de* E. *Il existe un unique endomorphisme* f^* *de* E *dit* adjoint *de* f *(relativement à* b_σ*) tel que, pour tout* $x, y \in \mathrm{E}$, *on ait*

$$b_\sigma(f(x), y) = b_\sigma(x, f^*(y)).$$

On a

(iv) $$\operatorname{Mat}(\mathcal{B}, f^*) = \sigma(\operatorname{Mat}(\mathcal{B}, b_\sigma)^{-1})\sigma({}^t \operatorname{Mat}(\mathcal{B}, f))\sigma(\operatorname{Mat}(\mathcal{B}, b_\sigma))$$

En particulier, a et a^* *on même rang et si* $\mathcal{B}$ *est* orthonormée, *on a*

$$\operatorname{Mat}(\mathcal{B}, f^*) = \sigma({}^t \operatorname{Mat}(\mathcal{B}, f))$$

Démonstration. *Soit* M *la matrice de* b_σ *dans une une base* $\mathcal{B}$ *arbitraire et* A *celle de* f. *On écrit matriciellement l'idéntité cherchée tenant comte de* $b_\sigma(x, y) = {}^t\mathrm{X}\mathrm{M}\sigma(\mathrm{Y})$ *:*

$${}^t(\mathrm{AX})\mathrm{M}\sigma(\mathrm{Y}) = {}^t\mathrm{X}{}^t\mathrm{A}\mathrm{M}\sigma(\mathrm{Y}) = {}^t\mathrm{X}\mathrm{M}\mathrm{M}^{-1}{}^t\mathrm{A}\mathrm{M}\sigma(\mathrm{Y}) = {}^t\mathrm{X}\mathrm{M}\sigma(\sigma(\mathrm{M}^{-1})\sigma({}^t\mathrm{A})\sigma(\mathrm{M})\mathrm{Y})$$

2. Dans le cas bilinéaire ($\sigma = \mathrm{Id}$), on dit que cette matrice est le *cocarré* de $\operatorname{Mat}(b_\sigma, \mathcal{B})$.

■

Les propositions usuelles de la transposition donnent les formules habituelles (linéarité de l'adjonction, $(f \circ g)^* = g^* \circ f^*$, $\mathrm{Id}^* = \mathrm{Id}$. Notons que dans le cas bilinéaire ($\sigma = \mathrm{Id}$), u et u^* sont semblables (4.5.0.3).

Exercice(s) 11.3.4.2. *Si b_σ est non dégénérée, montrer que l'isomorphisme $\check{b} : \mathrm{E} \to \mathrm{E}^*_\sigma$ défini par b_σ (cf. 11.3.3) identifie l'adjoint f^* de $f \in \mathrm{End}_{\mathbf{k}}(\mathrm{E})$ à sa transposée ${}^t f \in \mathrm{End}_{\mathbf{k}}(\mathrm{E}^\star_\sigma *) = \mathrm{End}_{\mathbf{k}}(\mathrm{E}^\star)$ (cf. 6.7.0.1).*

Définition 11.3.4.3. *Un endormorphisme d'un espace sesquilinéaire non dégénéré est* auto-adjoint *si $f = f^*$, normal s'il commute avec son adjoint, unitaire (ou simplement orthogonal lorsque $\sigma = \mathrm{Id}$) s'il est inversible et si $f^{-1} = f^*$. On dit aussi que f une isométrie.*

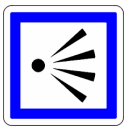
Exercice(s) 11.3.4.4. *Montrer que f laisse b_σ invariante si et seulement si f est unitaire.*

Chapitre 12

Complément : formes bilinéaires générales

Herbert Westren Turnbull

12.1 Point de vue



Nous, avons choisi ce thème, au delà de son importance mathématique, pour montrer un lien un peu inattendu entre congruence des matrices inversibles M et similitudes de leurs cocarrés (11.3.3.2) ${}^t\mathrm{M}^{-1}\mathrm{M}$ qui visiblement est caché dans le cas des formes ε-symétriques. Ce sujet a une longue histoire (voir [DT16]) pour un historique.

12.2 Introduction

On s'intéresse à la congruence $A \mapsto {}^tPAP$ avec $P \in GL_n(\mathbf{k})$ et $A \in M_n(\mathbf{k})$. Par exemple, si J_d est un bloc de Jordan, J_d et tJ_d sont congruentes via une congruence diagonale dès que t est non nul. Si P est une matrice de permutation, ${}^tPAP$ se déduit de A en permutant les lignes puis les colonnes de même indice. De même, si $P = T_{i,j}(\lambda)$ est une matrice de transvection, on déduit ${}^tPAP$ de A en ajoutant λ fois la i-ème colonne à la j-ème colonne puis en ajoutant λ fois la i-ème ligne à la j-ème ligne. On parlera de congruences $\equiv$ permises.

En particulier, si X est un élément non nul du noyau de A, on peut se ramener par des congruences permises à une première colonne nulle (écrire $X = \sum x_i e_i$ avec $\ell = x_i \neq 0$, faire la congruence associée à la transpositions $(1, i)$ si $i > 1$ puis les opérations permises associées aux $T_{1,j}(-a_{1,j}/\ell), j > 1$).

La classification complète [1] dans le cas algébriquement clos est obtenue en 12.6.0.6. On sait qu'on ne peut espérer un résultat sur un corps quelconque, même en dimension 1 ! Comme on le verra, l'existence de racines carrées est la clef de cette classification, comme dans le cas symétrique complexe (13.5.7.1) ou réel (13.5.8.1). Mais dès qu'elles existent, on peut classifier.

Il n'y a aucune difficulté à adapter au cas sesquilinéaire. On laisse cela au lecteur intéressé. Traiter le cas bilinéaire a simplement l'avantage de simplifier les notations.

12.3 Existence d'une décomposition

Le résultat suivant de décomposition des matrices permet de ramener l'étude des formes bilinéaires à celle des formes non dégénérées. Étrangement, il est assez récent, dû à P. Gabriel [Gab74]. Nous en donnons une version simplifiée essentiellement issue de [Ikr18] (pour la partie existence) et de [DS04] (pour la partie unicité). Ce lemme est évident dans le cas $\pm$-symétrique (ou hermitien) du fait de la coïncidence des noyaux de M et de ${}^tM$. L'intersection $\mathrm{Ker}(M) \cap \mathrm{Ker}({}^tM)$ apparaît clairement dans la preuve du résultat suivant.

Lemme 12.3.0.1. *Toute matrice est (algorithmiquement) congruente à une matrice diagonale par blocs* $\mathrm{diag}(A, J_{\underline{d}})$ *où* $A \in GL_r(\mathbf{k})$, $r = \mathrm{rang}(A)$ *et* $J_{\underline{d}} = \mathrm{diag}(J_{d_i})$ *est la matrice diagonale par blocs de Jordan de taille* d_i *associée à une partition* $\underline{d} = (d_i)$ *de* $n - r$.

Démonstration. *Il suffit de montrer que toute matrice est (algorithmiquement) congruente à une matrice diagonale par blocs* $\mathrm{diag}(A, J(\underline{\delta}))$ *où* $A \in GL_r(\mathbf{k})$, $r = \mathrm{rang}(A), \underline{\delta} \in (\mathbf{k} - \{0\})^{n-r}$, *et* $J(\underline{\delta}) = \mathrm{diag}(\delta_i J_{d_i})$ *est la matrice diagonale par blocs de Jordan de taille* d_i *associée à une partition* $\underline{d} = (d_i)$ *de* $n - r$. *En effet, une congruence diagonale ramène à la forme cherchée.*

On procède par récurrence sur $r = \mathrm{rang}(M)$. *On peut supposer* $0 < r < n$ *et le théorème prouvé en rang* $r - 1$. *D'après la remarque précédente, on peut donc supposer après des opérations permises que la*

1. Avec le système de représentants de Turnbull et Aitken

première colonne de M *est nulle (*$\mathrm{e}_1 \in \mathrm{Ker}(\mathrm{M})$*) :*

$$\mathrm{M} \equiv \begin{pmatrix} 0 & \mathrm{L}_{1,n-1} \\ 0 & \mathrm{M}'_{n-1,n-1} \end{pmatrix}.$$

Si $\mathrm{e}_1 \in \mathrm{Ker}({}^t\mathrm{M})$ i.e. *si* $\mathrm{L}_{1,n-1} = 0$*, on a* $\mathrm{M} = \mathrm{diag}(0, \mathrm{M}')$ *avec* M' *de rang* $r-1$ *et on conclut par récurrence. Sinon, un coefficient* $\delta_1 = \mathrm{L}_{1,j}, j > 1$ *est non nul. Comme la première colonne est nulle donc invariante par opérations de lignes, on se ramène par congruences permises à* $j = 2$*. En utilisant* δ_1 *comme pivot, on se ramène par congruences permises d'abord à*

$$\mathrm{M} \equiv \begin{pmatrix} 0 & \delta_1 & 0_{1,n-2} \\ 0 & \gamma & \mathrm{L}'_{1,n-2} \\ 0 & \mathrm{C}_{n-2,1} & \mathrm{M}'_{n-2,n-2} \end{pmatrix}.$$

La congruence permise soustrayant à la 2*-ième ligne* γ/δ_1 *fois la première ne change pas la seconde colonne puisque la première colonne est nulle !.De même pour les lignes d'indice* $> n-1$*. Ainsi, on a une congruence permise*

$$\mathrm{M} \equiv \begin{pmatrix} 0 & \delta_1 & 0_{1,n-2} \\ 0 & 0 & \mathrm{L}_{1,n-2} \\ 0 & 0_{n-2,1} & \mathrm{M}''_{n-2,n-2} \end{pmatrix}$$

Si $\mathrm{L}_{1,n-2} = 0$*, on conclut par récurrence appliquée à* M'' *qui est de rang* $r-1$*.*
Si $\mathrm{L}_{1,n-2} \neq 0$*, un des coefficients* δ_2 *de* $\mathrm{L}_{1,n-2} = 0$ *est non nul : il se trouve dans une colonne de* $\tilde{\mathrm{M}}$ *d'indice* $j \geq 3$*. Par opération permise, peut donc supposer* $j = 3$ *sans changer deux premières colonnes de* $\tilde{\mathrm{M}}$ *de sorte qu'on a*

$$\mathrm{M} \equiv \begin{pmatrix} 0 & \delta_1 & 0 & 0_{1,n-3} \\ 0 & 0 & \delta_2 & \mathrm{L}'_{1,n-3} \\ 0 & 0 & \mathrm{C}_{n-3,1} & \mathrm{M}_{n-3,n-3} \end{pmatrix}$$

Comme précédemment, par opérations permises sur les lignes d'indices > 2*, on peut supposer que le coefficient d'indice* $(2,3)$ *est le seul coefficient non nul dans sa colonne et donc*

$$\mathrm{M} \equiv \begin{pmatrix} 0 & \delta_1 & 0 & 0_{1,n-3} \\ 0 & 0 & \delta_2 & \mathrm{L}_{1,n-3} \\ 0 & 0 & 0_{n-3,1} & \mathrm{M}''_{n-3,n-3} \end{pmatrix}.$$

Si $\mathrm{L}_{1,n-3} = 0$*, on applique l'hypothèse de récurrence à* $\mathrm{M}''_{n-3,n-3}$*. Sinon, on itère le processus qui est clairement fini ce qui achève l'existence cherchée. Quitte à multiplier par une matrice diagonale à droite et à gauche, on peut modifier les* δ *par multiplication par des carrés non nuls d'où le secodn point.*

■

Remarque(s) 12.3.0.2. *Notons que dans le cas symétrique usuel (ou alternée, ou hermitien), la considération d'une base obtenue par complétion d'une base du noyau de* M *donne immédiatement le lemme,* A *étant congruente à la matrice de la forme* Ψ *induite par* M *sur* $\mathbf{k}^n/\mathrm{Ker}(\Psi)$. *On obtient alors sans difficulté un énoncé d'unicité dans ce cas. Même en remplaçant par exemple noyau par le noyau à gauche* $\mathrm{L}(\mathrm{V})$ *de* $\mathrm{V} = (\mathbf{k}^n, \Psi)$, *le problème dans le cas général, comme on va le voir, est que* Ψ *ne se factorise que sur un certain sous-espace* $\mathrm{L}^2(\mathrm{V})/\mathrm{L}(\mathrm{V})$ *du quotient* $\mathrm{V}/\mathrm{L}(\mathrm{V})$.

12.4 L'espace bilinéaire type

Notons qu'à ce stade, il n'est pas clair que ni $\underline{d}$ ni la classe de congruence de A soient déterminés par la classe de congruence de M.

Notons simplement $\mathrm{V_M}$ l'espace bilinéaire $\mathbf{k}^n$ muni de la forme bilinéaire $(\mathrm{X}, \mathrm{Y}) \mapsto {}^t\mathrm{XMY}$ (de matrice M dans la base canonique)

On note simplement V_d l'espace $\mathrm{V_{J_d}} = \mathbf{k}[\mathrm{T}]/(\mathrm{T}^d)$ associé au bloc de Jordan standard de taille $d \geq 1$ vu comme la matrice de la multiplication par T dans la base des monômes comme d'habitude.On a donc $(\mathrm{T}^i, \mathrm{T}^j) = \delta_{j+1,i}$. On a $\mathrm{V}_d = \{0\}$ si $d \leq 0$ car J_d est la matrice vide.

On dispose (par exemple) du noyau à droite

$$\mathrm{R}(\mathrm{V}) = \{v \in \mathrm{V} | \Psi(\mathrm{V}, v) = \{0\}$$

et à gauche

$$\mathrm{L}(\mathrm{V}) = \{v \in \mathrm{V} | \Psi(v, \mathrm{V}) = \{0\}$$

. qui est un invariant par isomorphisme d'espaces bilinéaires. On a alors immédiatement,

$$\mathrm{R}(\mathrm{V_M}) = \mathrm{Ker}(\mathrm{M}) \text{ et } \mathrm{R}(\mathrm{V_M}) = \mathrm{Ker}({}^t\mathrm{M}).$$

On définit également

$$\mathrm{R}^2(\mathrm{V}) = \{v \in \mathrm{V} | \Psi(\mathrm{R}(\mathrm{V}), v) = \{0\}.$$

C'est un sous-espace qui contient $\mathrm{R}(\mathrm{V})$ et Ψ induit une forme sur $\mathrm{R}^2(\mathrm{V})/\mathrm{R}(\mathrm{V})$ en faisant ainsi un espace bilinéaire.

Exercice(s) 12.4.0.1. *Montrer que* $\mathrm{R}^2(\mathrm{V})/\mathrm{R}(\mathrm{V})$ *est le plus grand sous-espace de* $\mathrm{V}/\mathrm{R}(\mathrm{V})$ *sur lequel* Ψ *passe au quotient.*

Pour pour $\mathrm{M} = \mathrm{J}_d$ et tout $d \geq 1$, on a

- $\mathrm{R}(\mathrm{V}_d) = \langle \mathrm{T}^{d-1} \rangle \overset{\times \mathrm{T}^{1-d}}{\simeq} \mathrm{V}_1$ et $\mathrm{L}(\mathrm{V}_d) = \langle 1 \rangle \simeq \mathrm{V}_1$
- $\dim \mathrm{R}(\mathrm{V}_d) \cap \mathrm{L}(\mathrm{V}_d) = \delta_{1,d}$
- $\mathrm{R}^2(\mathrm{V}_d) = \langle 1, \mathrm{T}, \ldots, \mathrm{T}^{d-3}, \mathrm{T}^{d-1} \rangle$
- $\mathrm{R}^2(\mathrm{V}_d)/\mathrm{R}(\mathrm{V}_d) = \mathrm{T}^2\mathbf{k}[\mathrm{T}]/(\mathrm{T}^{d-1}) \overset{\times \mathrm{T}^{-2}}{\simeq} \mathbf{k}[\mathrm{T}]/(\mathrm{T}^{d-3}) = \mathrm{V}_{d-2}$.

Pour A inversible, on a

— $\mathrm{R}(\mathrm{V_A}) = \{0\}$

— $\mathrm{R}^2(\mathrm{V_A}) = \mathrm{V_A}$

— $\mathrm{R}^2(\mathrm{V_A})/\mathrm{R}(\mathrm{V_A}) = \mathrm{V_A}$.

On dira que deux sous-espaces $\mathrm{V}_1, \mathrm{V}_2$ d'un espace bilinéaire V sont en somme directe orthogonale s'ils sont en somme directe et si $(\mathrm{V}_1, \mathrm{V}_2) = (\mathrm{V}_2, \mathrm{V}_1) = \{0\}$, et on écrit alors $\mathrm{V} = \mathrm{V}_1 \overset{\perp}{\oplus} \mathrm{V}_2$. On alors

$$\mathrm{R}(\mathrm{V}_1 \overset{\perp}{\oplus} \mathrm{V}_2) = \mathrm{R}(\mathrm{V}_1) \overset{\perp}{\oplus} \mathrm{R}(\mathrm{V}_2) \text{ et } \mathrm{R}^2(\mathrm{V}_1 \overset{\perp}{\oplus} \mathrm{V}_2) = \mathrm{R}^2(\mathrm{V}_1) \overset{\perp}{\oplus} \mathrm{R}^2(\mathrm{V}_2)$$

12.5 Unicité

Avec les notations précédentes, le lemme entraîne l'existence d'un isomorphisme d'espaces bilinéaires

$$\mathrm{V_M} \simeq \mathrm{V_A} \overset{\perp}{\oplus}\overset{\perp}{\oplus}_{d\in\underline{d}} \mathrm{V}_d.$$

On souhaite *in fine* récupérer les classe d'isomorphismes de $\mathrm{V_A}$ (donc de congruence de A) et de $\mathrm{V}_{\underline{\delta}}$ et même la partition $\underline{d}$. Précisément, avec des notations évidentes, on veut montrer

Lemme 12.5.0.1. *Si on a un isomorphismes d'espaces bilinéaires*

$$\mathrm{V_A} \overset{\perp}{\oplus}\overset{\perp}{\oplus}_{d\in\underline{d}} \mathrm{V}_d \simeq \mathrm{V_{A'}} \overset{\perp}{\oplus}\overset{\perp}{\oplus}_{d'\in\underline{d}'} \mathrm{V}'_d$$

avec A *et* A′ *inversibles, on a* A *et* A′ *congruentes et* $\underline{d} = \underline{d}'$.

Démonstration. *Pour cela, on va procéder par récurrence sur* $n + \max(\underline{d})$ *simplement en calculant les espaces bilinéaires* $\mathrm{R}(\mathrm{V_M})$ *et* $\mathrm{R}^2(\mathrm{V_M})/\mathrm{R}(\mathrm{V_M})$ *(qui sont des invariants par isomorphismes d'espaces bilinéaires rappelons le). On applique les formules de la section précédente qui donnent*

1. $\mathrm{R}(\mathrm{V_M}) \simeq \overset{\perp}{\oplus}_{\substack{d\in\underline{d}\\ d\geq 1}} \mathrm{V}_1$
2. $\mathrm{R}^2(\mathrm{V_M})/\mathrm{R}(\mathrm{V_M}) = \mathrm{V_A} \overset{\perp}{\oplus}\overset{\perp}{\oplus}_{\substack{d\in\underline{d}\\ d\geq 3}} \mathrm{V}_{d-2}$

On a donc $\mathrm{R} \simeq \mathrm{R}'$ *et* $\mathrm{R}^2/\mathrm{R} \simeq \mathrm{R}^{2\prime}/\mathrm{R}'$.

— *Si* $\dim(\mathrm{R}^2/\mathrm{R}) < n$, *l'hypothèse de récurrence assure* A *et* A′ *sont congruentes et* $\underline{d}, \underline{d}'$ *coïncident sur leurs éléments* $d, d' \geq 3$. *Mais on a alors* $\dim \mathrm{R} \cap \mathrm{L} = \mathrm{Card}\{d \in \underline{d} | d = 1\}$ *et* $\dim \mathrm{R}(\mathrm{V}_d) = \mathrm{Card}\{d \in \underline{d} | d \geq 1\}$ *montrent que les valeurs* 1 *et* 2 *apparaissent avec le même poids dans* $\underline{d}$ *et* $\underline{d}'$ *de sorte qu'*in fine $\underline{d} = \underline{d}'$.

— *Supposons* $\dim(\mathrm{R}^2/\mathrm{R}) = n$.

- *Si* $\max(\underline{d}) \geq 3$, *on peut également appliquer l'hypothèse de récurrence et conclure comme précédemment.*
- *Sinon, on a* $\max(\underline{d}) \leq 2$ *on a* $\mathrm{R}^2/\mathrm{R} = \mathrm{V_A} = \mathrm{V}'_{\mathrm{A}}$ *ce qui montre que* A *et* A′ *sont congruentes dans ce cas. Mais dans ce cas également, la considération de* R *et* $\mathrm{R} \cap \mathrm{L}$ *montre comme précédemment* $\underline{d} = \underline{d}'$ *puisque qu'il n'y a pas d'élément* $d \in \underline{d} \cup \underline{d}'$ *tel que* $d \geq 2$.

■

Des lemmes 12.3.0.1 et 12.5.0.1 on déduit immédiatement

Théorème 12.5.0.2. *Toute matrice* $M \in M_n(\mathbf{k})$ *est (algorithmiquement) congruente à une matrice diagonale par blocs* $\mathrm{diag}(A, J_{\underline{d}})$ *où* $A \in GL_r(\mathbf{k})$, $r = \mathrm{rang}(A)$ *et* $J_{\underline{d}} = \mathrm{diag}(J_{d_i})$ *est la matrice diagonale par blocs de Jordan de taille* d_i *associée à une partition* $\underline{d} = (d_i)$ *de* $n-r$. *De plus la classe de congruence de* A *et* $\underline{d}$ *sont uniquement déterminées par la classe de congruence de* M.

Remarque(s) 12.5.0.3. *Le lecteur adaptera ces résultats sans aucune difficulté au cas sesquilinéaire.*

12.6 Classification : cas algébriquement clos

On suppose ici $\mathbf{k}$ algébriquement clos (ou si on veut, on suppose seulement que tout élément est un carré). Le résultat suivant est ancien, sans doute dû à Turnbull dès 1936 ([TA61]). Je reprends ici l'approche de [HP94].

Proposition 12.6.0.1. *Deux matrices inversibles* A, B *sont congruentes si et seulement si les asymétries*[2] ${}^tA^{-1}A$ *et* ${}^tB^{-1}B$ *des formes associées sont semblables,* i.e. *aient les mêmes invariants de similitude.*

Notons que ceci est consistant avec la classifications des formes quadratiques, alternées (ou hermitiennes, cf. remarque supra) dans le cas algébriquement clos.

Démonstration. *Si* M *est inversible, on pose* $M' = {}^tM^{-1}M$ *(appelée le cocarré dans la littérature). Si* ${}^tPAP = B$ *avec* P *inversible, on a* $P^{-1}A'AP$ *et* $B'B$, *d'où le sens direct.*
Inversement, supposons que $A'A$ *et* $B'B$ *sont semblables. Alors en considérant les* pinceaux $A + {}^tAT$ *et* $B + {}^tBT$, *il existe d'après 4.5.0.4,* P, Q *inversibles tels que*

$$PAQ = B \text{ et } P{}^tAQ = {}^tB \text{ et donc } {}^tQA{}^tP = B.$$

On déduit

$$\begin{aligned} PAQ &= {}^tQA{}^tP \\ XA &= B{}^tX \text{ avec } X = Q'P \text{ et donc} \\ \Xi A &= B{}^t\Xi \text{ pour tout } \Xi \in \mathbf{k}[X] \end{aligned}$$

Comme X *est inversible et* $\mathbf{k}$ *algébriquement clos, on choisit alors* $\Xi \in \mathbf{k}[X]$ *tel que* $\Xi^2 = X$ *(7.3.2.1). On a alors*

2. voir 11.3.3.2.

$$\begin{aligned} \mathrm{B} &= {}^t\mathrm{Q}\mathrm{A}{}^t\mathrm{P} \\ &= {}^t\mathrm{Q}\mathrm{A}{}^t\mathrm{X}\mathrm{Q} \\ &= {}^t\mathrm{Q}\mathrm{A}({}^t\Xi)^2\mathrm{Q} \\ &= {}^t\mathrm{Q}(\mathrm{A}{}^t\Xi){}^t\Xi\mathrm{Q} \\ &= {}^t\mathrm{Q}(\Xi\mathrm{A}){}^t\Xi\mathrm{Q} \\ &= {}^t({}^t\Xi\mathrm{Q})\mathrm{A}({}^t\Xi\mathrm{Q}) \end{aligned}$$

■

Exercice(s) 12.6.0.2. *Retrouver le résultat 15.1.13.1.*

Reste à trouver une famille de matrices non congruentes deux à deux qui décrit les invariants de similitude possible des cocarrés. Il y a forcément des restrictions puisque le déterminant d'un cocarré est 1. Si P est un polynôme unitaire tel que $\mathrm{P}(0) \neq 0$, on note

$$\mathrm{P}^*(\mathrm{T}) = \frac{\mathrm{T}^{\deg(\mathrm{P})}\mathrm{P}(1/\mathrm{T})}{\mathrm{P}(0)}$$

son polynôme (unitaire) aux inverses. Si A est inversible d'invariants de similitude $\underline{\mathrm{P}} = (\mathrm{P}_i)$, on note $\underline{\mathrm{P}}^* = (\mathrm{P}_i^*)$.

Lemme 12.6.0.3. *Soit* $\underline{\mathrm{P}}$ *les invariants de* A *inversible.*

1. *La famille des invariants de similitude de* A^{-1} *est* $\underline{\mathrm{P}}^*$.
2. *Si* A *est un cocarré, on a* $\underline{\mathrm{P}} = \underline{\mathrm{P}}^*$.
3. *Les invariants*

Démonstration. *Soit* P *un invariant de similitude de* A. *Comme* A *inversible,* T *est inversible dans* $\mathrm{V} = \mathbf{k}[\mathrm{T}]/(\mathrm{P})$ *(d'inverse* $(\mathrm{P}(\mathrm{T}) - \mathrm{P}(0)/(\mathrm{TP}(0))$. *Mais* $\mathrm{V}_{\mathrm{T}^{-1}}$ *est un* $\mathbf{k}[\mathrm{T}^{-1}]$ *module cyclique (engendré par n'importe quel monôme) annulé par* P^* *qui est donc son minimal pour des raisons de dimensions, d'où le (1). Comme on l'a déjà remarqué,* $^t\mathrm{BT} - \mathrm{B}$ *est équivalente à* $\mathrm{Id} - {}^t\mathrm{B}^{-1}\mathrm{B}$ *de sorte que les invariants de similitude de* $\mathrm{A} = {}^t\mathrm{B}^{-1}\mathrm{B}$ *sont les facteurs invariants de* $^t\mathrm{BT} - \mathrm{B}$, *donc coïncident avec ceux de sa transposée* $\mathrm{TB} - {}^t\mathrm{B}$ *donc avec les invariants de similitude de* $\mathrm{B}^{-1}{}^t\mathrm{B} = \mathrm{A}^{-1}$. ■

Désignons par $\Lambda_\pm$ l'ensemble des paires (non ordonnées) $\overline{\lambda} = \{\lambda, \lambda^{-1}\}$ -donc avec $\lambda \neq \pm 1$. Les invariants de similitude d'un cocarré de spectre s'écrivent donc

$$\mathrm{P}_i(\mathrm{T}) = \prod\nolimits_{\Lambda_\pm} [(\mathrm{T} - \lambda)(\mathrm{T} - \lambda^{-1})]^{v_{\overline{\lambda},i}} (\mathrm{T} - 1)^{v_{i,+}} (\mathrm{T} + 1)^{v_i,-}$$

de sorte que les blocs associés de la réduite de Jordan sont (vec un micro-abus d'écriture)

$$\mathrm{diag}(\lambda\,\mathrm{Id} + \mathrm{J}_{v_{\overline{\lambda},i}}, \lambda^{-1}\,\mathrm{Id} + \mathrm{J}_{v_{\overline{\lambda},i}}, \mathrm{Id}_{v_{i,+}}, -\mathrm{Id}_{v_{i,-}})$$

Autrement dit, la réduite de Jordan d'un cocarré a deux blocs $\pm\mathrm{Id}$ de taille éventuellement différentes [3] et des blocs de type

(i) $$\mathrm{diag}(\lambda\,\mathrm{Id}+\mathrm{J}_d, \lambda^{-1}\,\mathrm{Id}+\mathrm{J}_d) \text{ avec } \lambda\neq\pm 1$$

Remarque(s) 12.6.0.4. *Si β est l'asymétrie de l'espace bilinéaire V et λ,μ des valeur propres, on vérifie sans difficulté que les (sommes d')espaces caractéristiques $\mathrm{V}[\mathrm{T}-\lambda]+\mathrm{V}[\mathrm{T}-\lambda^{-1}]$ et $\mathrm{V}[\mathrm{T}-\mu]$ sont orthogonaux dès que $\lambda'\notin\{\lambda,\lambda^{-1}\}$. C'est un phénomène général, même si le corps n'est pas algébriquement en remplaçant les espaces caractéristiques par les composantes primaires associées à des irréductibles unitaires P,P^* d'une part et $\mathrm{Q}\notin\{\mathrm{P},\mathrm{P}^*\}$ de l'autre. Ainsi, la décomposition de Jordan, avec des blocs convenablement regroupés, correspond à une décomposition orthogonale !*

Reste à exhiber une famille de matrices dont les cocarrés sont comme dans i. On va regarder les représentants de [HP94] (d'autres choixi classiques existent, voir par exemple [HS08]).

Lemme 12.6.0.5. *Soit $d>0, \mathrm{V}=\mathrm{T}^{-d}\mathbf{k}_{\leq 2d}[\mathrm{T}]/\mathbf{k}.1$ et $\lambda\neq\pm 1$. Soit $a\in\mathrm{End}_{\mathbf{k}}(\mathrm{V})$ défini par*

$$\begin{aligned} a(\mathrm{T}^i) &= \mathrm{T}^{-i} \text{ si } i<0 \\ &= \lambda\mathrm{T}^{-i}+\mathrm{T}^{-i+1} \text{ si } i>0. \end{aligned}$$

Alors, la réduite de Jordan du cocarré de a est $\mathrm{diag}(\lambda\,\mathrm{Id}+\mathrm{J}_d,\lambda^{-1}\,\mathrm{Id}+\mathrm{J}_d)$.

Démonstration. *Soit A la matrice de a dans la base $\mathrm{T}^i, 0<|i|\leq d$ et P la matrice de permutation $i\mapsto -i$. On remarque alors que la matrice de $(\mathrm{T}^t\mathrm{A}-\mathrm{A})\mathrm{P}$ est égale à*

$$\mathrm{diag}((\mathrm{T}-\lambda)\,\mathrm{Id}+\mathrm{J}_d,(\lambda\mathrm{T}-1)\,\mathrm{Id}+\mathrm{T}\mathrm{J}_d)\approx\mathrm{diag}((\mathrm{T}-\lambda)\,\mathrm{Id}+\mathrm{J}_d,(\mathrm{T}-\lambda^{-1})\,\mathrm{Id}+\mathrm{T}\mathrm{J}_d).$$

Plaçons nous dans l'anneau principal $\mathbf{k}[\mathrm{T},\mathrm{T}^{-1}]$. Comme la sous diagonale de chacun des deux blocs est 1 ou T, le premier mineur principal d'ordre $d-1$ obtenu en rayant les premières ligne et colonne vaut 1 ou T donc est inversible. On déduit que les facteurs invariants de ces matrices dans $\mathbf{k}[\mathrm{T}^{-1},\mathrm{T}$, définis à un inversible près,] est $(1,\ldots,(\mathrm{T})\lambda)^d)$ où $(1,\ldots,(\mathrm{T}-\lambda^{-1})^d)$. Les facteurs invariants de ces matrices dans $\mathbf{k}[\mathrm{T}]$ sont les mêmes que dans $\mathbf{k}[\mathrm{T}^{-1},\mathrm{T}]$, à un inversible de $\mathbf{k}[\mathrm{T}^{-1},\mathrm{T}]$ de la forme T^i près. Mais comme ils sont premiers avec T par hypothèse, ce sont donc bien $(1,\ldots,(\mathrm{T}^\lambda)^d)$ où $(1,\ldots,(\mathrm{T}-\lambda^{-1})^d)$. ■

Bien entendu, la matrice de la forme bilinéaire alternée non dégénérée standard (13.4) de rang $2d$ a pour cocarré $-\mathrm{Id}_{2n}$. Pour distinguer 1 de -1, on se place en caractéristique différente de 2.

3. La taille du bloc $-\mathrm{Id}$ est nécessairement paire en caractéristique différente de 2 : c'est le cas alterné non dégénéré !

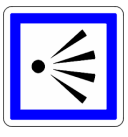
Théorème 12.6.0.6. *Soit* $\mathbf{k}$ *algébriquement clos de caractéristique différente de* 2 *et* V *un espace bilinéaire quelconque,* r *le rang de la forme bilinéaire. Alors, il existe une unique partition* $\underline{d}$ *de* $\dim(\mathrm{V}) - r$*, il existe une décomposition orthogonale* $\mathrm{V} = \mathrm{W} \overset{\perp}{\oplus} \mathrm{V}_{\underline{d}}$ *avec* W *non dégénéré, unique à isomorphisme bilinéaire près. De plus,* W *se décompose en somme directe orthogonale de sa partie symétrique, de sa partie et antisymétrique et d'espaces non dégénérés de matrices* $\mathrm{diag}(\lambda\,\mathrm{Id} + \mathrm{J}_d, \lambda^{-1}\,\mathrm{Id} + \mathrm{J}_d), \lambda \neq \pm 1$*. Cette décomposition est unique à isomorphisme bilinéaire près.*

Chapitre 13

Formes ε-symétriques, Formes quadratiques

Taj Mahal

13.1 Point de vue



Les formes quadratiques en dimension finie peuvent être vues comme des polynômes homogènes de degré 2 à n variables ou [1]comme des formes bilinéaires symétriques sur $\mathbf{k}^n$ -ou leur matrice associée-. Le groupe linéaire agit sur les premiers par changement de variable sur les secondes par congruence. Nous allons utiliser les deux points de vue pour leur étude, étude qui dépend fortement du corps $\mathbf{k}$ contrairement à l'algèbre linéaire ordinaire.

Comme nous allons le voir, le premier outil fondamental dans la classification des formes ε-symétriques est l'orthogonalité qui permet de ramener nombre de problèmes à des formes sur des droites ou des plans.

1. En caractéristique différente de 2 du moins.

13.2 Introduction

La théorie des formes formes sesquilinéaires sesquilinéaires servent de fondement à la géométrie : les formes bilinéaires symétriques réelles conduisent à la géométrie riemannienne, sont omniprésentes en théorie des nombres, et les formes hermitiennes complexes à la géométrie complexe et holomorphe. Par ailleurs, la partie imaginaire d'une forme hermitienne complexe définit une forme alternée, formes alternées qui sont au cœur de la de la géométrie symplectique. Ces géométries sont à la base de la physique classique (géométrie euclidienne), relativiste (géométrie Lorentzienne pour la relativité restreinte, riemannienne générale pour la relativité générale), de la mécanique (géométrie symplectique) et de la mécanique quantique (géométrie hermitienne).

On garde les notations 11.2.1. On rappelle (11.3.3) qu'une forme sesquilinéaire b_σ qui est ε-hermitienne définit une forme non dégénérée sur $\mathrm{E}/\operatorname{Ker}(b_\sigma)$, ou, ce qui revient au même, sur n'importe quel supplémentaire du noyau ramenant ainsi leur étude au cas non dégénéré. On dit parfois que ce noyau est le radical de b_σ ou de (E, b_σ) voire le radical $\operatorname{rad}(\mathrm{E})$ de E quand aucune confusion n'est à craindre.

13.3 Orthogonalité

Dans ce seul numéro du chapitre, considérons b_σ une forme sesquilinéaire ε-hermitienne. On a donc

$$b_\sigma(x, y) = 0 \Rightarrow b_\sigma(y, x) = 0.$$

Définition 13.3.0.1. *Si* $\mathrm{F} \subset \mathrm{E}$ *est un* sous-ensemble*, l'orthogonal (à gauche) de* F *est le* sous-espace vectoriel *de* E*.*

$$\mathrm{F}^\perp = \{x \in \mathrm{E},\ \forall\, y \in \mathrm{F},\ b_\sigma(x, y) = 0\}.$$

On a donc $b(\mathrm{F}^\perp, \mathrm{F}) = \{0\}$ et donc $b(\mathrm{F}, \mathrm{F}^\perp) = \{0\}$ de sorte que $\mathrm{F} \subset (\mathrm{F}^\perp)^\perp$.

Proposition 13.3.0.2. *Supposons* b_σ *non dégénérée et soit* F, G *des sous-espaces de* E*.*

1. $\dim \mathrm{F} + \dim \mathrm{F}^\perp = \dim \mathrm{E}$*,*
2. $\mathrm{F} = (\mathrm{F}^\perp)^\perp$*,*
3. $(\mathrm{F} + \mathrm{G})^\perp = \mathrm{F}^\perp \cap \mathrm{G}^\perp$ *et* $(\mathrm{F} \cap \mathrm{G})^\perp = \mathrm{F}^\perp + \mathrm{G}^\perp$*.*
4. *L'orthogonalité est décroissante.*
5. *Les conditions suivantes sont équivalentes*
   - *(a)* $\mathrm{F} \cap \mathrm{F}^\perp = \{0\}$
   - *(b)* b_σ *non dégénérée sur* F
   - *(c)* b_σ *non dégénérée sur* $\mathrm{F}^\perp$

Démonstration. *Par définition,* $\widehat{b}_\sigma(\mathrm{F})$ *est l'orthogonal en dualité de* F *dans* $\mathrm{E}^\star_\sigma$*. Les deux premières formules découlent du calcule de dimension de l'orthogonal en dualité (11.3.4.1). Les deux dernières sont formelles (mais utiles). Pour 5), l'équivalence de a) et b) est tautologique. Mais en invoquant cette équivalence pour* $\mathrm{F}^\perp$*, on conclut l'équivalence de a) et b) grâce à 2).* ∎

La preuve justifie *a posteriori* l'abus de notation $\perp$ pour l'orthogonal qui en général ne prête pas à confusion.

Définition 13.3.0.3. *On dit qu'un vecteur* x *est* isotrope *si* $b_\sigma(x,x)=0$*. On note par* $\mathcal{C}(b_\sigma)$ *le* cône *des vecteurs isotropes ; il contient* $\mathrm{Ker}(b_\sigma)$*.*
Un espace vectoriel F *est* isotrope *si* $\mathrm{F}\cap\mathrm{F}^\perp\neq\{0\}$*, totalement isotrope si* $\mathrm{F}\subset\mathrm{F}^\perp$*. Il est* anistrope *si si* $\mathrm{F}\cap\mathrm{F}^\perp=\{0\}$
Enfin, on dit que F *est* totalement isotrope *si* $\mathrm{F}\subset\mathrm{F}^\perp$*, autrement dit, si* $b_\sigma|_{\mathrm{F}\times\mathrm{F}}\equiv 0$*.*

Exemple. Si $\mathrm{E}=\mathbf{R}^n$ et $b(\mathrm{X},\mathrm{Y})=x_1y_2+y_1x_2$, alors tous les vecteurs de la base canonique sont isotropes.

Remarque(s) 13.3.0.4. *Si* F *est isotrope, alors* $\mathrm{F}\cap\mathrm{F}^\perp$ *est totalement isotrope.*
Si b_σ *est non dégénérée et* F *totalement isotrope, alors* $\dim\mathrm{F}\leq\dim\mathrm{E}/2$ *(13.3.0.2) ; si en revanche* F *est anisotrope, alors on a* $\mathrm{E}=\mathrm{F}\overset{\perp}{\oplus}\mathrm{F}^\perp$*.*
Notons qu'en général on a évidemment $\mathrm{rad}(\mathrm{E}_1\overset{\perp}{\oplus}\mathrm{E}_2)=\mathrm{rad}(\mathrm{E}_1)\overset{\perp}{\oplus}\mathrm{rad}(\mathrm{E}_2)=$ *ce qui permettra de systématiquement se ramener au cas non dégénéré.*

Proposition 13.3.0.5. *Supposons* b_σ *non dégénérée et soit* $a\in\mathrm{End}(\mathrm{E})$ *et* a^* *son adjoint. Alors on a*

1. $\mathrm{Ker}(a^*)=\mathrm{Im}(a)^\perp$.
2. $\mathrm{Im}(a^*)=\mathrm{Ker}(a)^\perp$.

Démonstration. *On peut comme dans 11.3.4.2 le déduire formellement de la propriété analogue pour la transposée. Donnons en une preuve directe (qui est une copie de la preuve par dualité). D'abord il suffit de montrer une des deux formules (changer* a *en* a^* *et utiliser l'involutivité de l'adjonction et de l'orthogonal). Ensuite,* $\mathrm{Im}(a^*)=\mathrm{Ker}(a)^\perp$ *ayant la même dimension d'après 11.3.4.1 et 13.3.0.2, il suffit de prouver les* $\mathrm{Im}(a^*)\subset\mathrm{Ker}(a)^\perp$*. Or, si* $a(x)=0$*, on a* $b_\sigma(x,a^*(y)=b_\sigma(a(x),y)=0$ *et donc* $a^*(y)\in\mathrm{Ker}(a)^\perp$*.* ∎

13.4 Formes alternées

Nous allons donner dans le cas alterné une première illustration du principe de l'introduction : une grande partie de la géométrie est contenue en dimension ≤ 2, la dimension plus grande s'obtenant par orthogonal.

13.4.1 Classification

Soit donc b une forme bilinéaire alternée sur V de dimension finie d.

Théorème 13.4.1.1. *Les formes alternées sont classifiées par leur rang. Une forme non dégénérée est de rang pair $2n$ et sa matrice dans une base convenable est*

$$\mathrm{W}_n = \mathrm{diag}(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix})$$

Démonstration. *Quitte à se restreindre à un supplémentaire du noyau, on peut supposerb non dégénérée et* $\dim(\mathrm{V}) > 0$. *Soit* $x_1 \in \mathrm{V}$ *non nul. Comme le noyau de b est nul, on peut choisir* y_1 *tel que* $b(x_1, y_1)$ *de sorte que* $\Pi = \langle x_1, y_1 \rangle$ *est un plan (si* y_1 *était colinéaire à* x_1 *on aurait* $b(x, y_1) = 0$*). La dimension de l'orthogonal de* Π *est* $n-2$. *Si* $z = \alpha x_1 + \beta y_1 \in \Pi^\perp$, *on a* $0 = b(z, x_1) = -\beta$ *et* $0 = b(z, y_1) = -\alpha$ *de sorte que* $z = 0$ *et on a une décomposition orthogonale* $\mathrm{V} = \Pi \overset{\perp}{\oplus} \Pi^\perp$ *où la restriction de b à* $\Pi^\perp$ *est non dégénérée (11.3.3.2). On conclut par une récurrence sans mystère.* ∎

En particulier, les formes alternées non dégénérées n'existent qu'en dimension paire.

Exercice(s) 13.4.1.2. *En utilisant le principe de prolongement des identités algébriques, montrer que le déterminant de toute matrice de taille impaire, alternée et à coefficients dans un anneau est nul.*

Exercice(s) 13.4.1.3. *Énoncer et prouver un résultat analogue à 13.4.1.1 dans le cas anti-hermitien.*

13.4.2 Pfaffien

Le théorème 13.4.1.1 prouve qu'une matrice alternée inversible est de taille $2n$ et est congruente à W_n. Comme son déterminant est 1, le déterminant de toute matrice alternée est un carré de $\mathbf{k}$. Cela est vrai de manière universelle. Soit $\mathrm{R} = \mathbf{Z}[\mathrm{T}_{i,j}], 1 \leq i < j \leq 2n$ l'anneau de polynômes à coefficients entiers à $(n(2n-1))$ indéterminées. C'est un anneau factoriel dont on note K le corps des fractions (10.3.3.1). Soit $\mathrm{M} \in \mathrm{M}_n(\mathrm{K})$ la matrice de coefficients $\mathrm{sign}(i-j))\mathrm{T}_{i,j}$. C'est une matrice alternée polynomiale qui définit donc une fonction matricielle sur les matrices alternées de $\mathrm{M}_n(\mathbf{k})$.

Proposition 13.4.2.1. *Il existe un unique polynôme* $\mathrm{Pf}(\mathrm{M}) \in \mathrm{R}$ *de carré* $\det(\mathrm{M})$ *et qui vaut* 1 *lorsque* $\mathrm{M} = \mathrm{W}_n$.

Démonstration. *Si on avait un second polynôme* Q *vérifiant la proposition, on aurait* $\mathrm{Q}^2 = \mathrm{Pf}^2$ *et donc* $\mathrm{Q} = \pm\mathrm{Pf}$ *par l'intégrité de* R. *Mais en regardant la valeur sur* W_n, *on conclut* $\mathrm{Q} = \mathrm{Pf}$. *Pour l'existence, observons que* $\det(\mathrm{M}) \in \mathrm{K}$ *est non nul (car cela est vrai lorsque* $\mathrm{M} = \mathrm{W}_n$*). Soit donc* $\mathrm{P} \in \mathrm{GL}_{2n}(\mathrm{K})$ *tel que* ${}^t\mathrm{PW}_n\mathrm{P} = \mathrm{M}$ *(13.4.1.1). On a donc* $\det(\mathrm{M}) = (\det(\mathrm{P}))^2$ *avec* $\det(\mathrm{M}) \in$ *et* $\det(\mathrm{P}) \in \mathrm{K}^*$. *Écrivons la décomposition* $\det(\mathrm{M}) = \prod p_i^{v_i}$ *en facteurs irréductibles dans l'anneau factoriel* R *et de même, en écrivant celles des numérateurs et dénominateurs de* $\det(\mathrm{P})$, *écrivons* $\det(\mathrm{P}) = u\prod p_i^{w_i}$ *avec* u *inversible de* R. *On a donc* $v_i \geq 0,\ w_i \in \mathbf{Z}$ *et, par unicité de la décomposition,* $2w_i = v_i \geq 0,\ u^2 = 1$. *On pose alors* $\mathrm{P}f = \pm u\prod p_i^{w_i} \in \mathrm{R}$ *le signe étant choisi pour avoir* $\mathrm{Pf}(\mathrm{W}_n) = 1$. ∎

La géométrie symplectique est l'étude des propriétés qui sont invariantes par le groupe symplectique $\mathrm{Sp}_{2n}(\mathbf{k})$ des matrices P préservant W_n, *i.e.* telles que ${}^t\mathrm{PW}_n\mathrm{P} = \mathrm{W}_n$. Elle est très riche, pleine de questions ouvertes mais dépasse le cadre choisi.

13.5 Formes quadratiques

Nous donnons une définition générale. Si la théorie des formes quadratiques en caractéristique 2 est utile et intéressante, elle diffère notablement du cas de caractéristique $\neq 2$. Dès lors,

sauf mention expresse du contraire, k désigne à partir de 13.5.1 un corps de caractéristique différente de 2.

Définition 13.5.0.1. *Une application* $q : \mathrm{E} \to \mathbf{k}$ *est une forme quadratique si*

1. q *est homogène de poids* 2, i.e.*pour tout* $x \in, \lambda \in \mathbf{k},\ q(\lambda x) = \lambda^2 q(x)$;
2. *l'application* ${}^2b : \begin{cases} \mathrm{E}\times\mathrm{E} & \to & \mathbf{k} \\ (x,y) & \mapsto & q(x+y)-q(x)-q(y) \end{cases}$ *est bilinéaire (symétrique).*

Exemple(s) 13.5.0.2. *Soit* $\mathrm{E} = \mathbf{k}^n$ *et* $\mathrm{P} \in \mathbf{k}[\mathrm{X}, \cdots, \mathrm{X}_n]$ *un polynôme homogène de degré* 2 *à* n *variables. Alors,* P *définit une forme quadratique* $(x_i) \mapsto \mathrm{P}(x_i)$ *et ce réciproquement une fois une base* $\mathcal{B}$ *de* E *étant donnée.*

Notons qu'on a ${}^2b(x,x) = 2q(x)$. Ainsi,

1. Si $\mathrm{car}(\mathbf{k}) = 2$, on a b_2 à la fois symétrique et alternée. Observons, que $(x_i) \mapsto x_1^2$ sur $\mathbf{k}^n$ est une forme quadratique, comme l'est tout polynôme homogène de degré 2 à n variables. Pourtant, ${}^2b(x,y) = (x_1+y_1)^2 - x_1^2 - y_1^2 = 0$ de sorte que q et 0 ont la même forme bilinéaire associée !

2. Si $\mathrm{car}(\mathbf{k}) \neq 2$, on a $q(x) = \frac{1}{2}{}^2 b(x,x)$ se donner q équivaut à se donner une forme bilinéaire symétrique ou encore une matrice symétrique.

13.5.1 Forme polaire

Définition 13.5.1.1. *La forme polaire de q est la forme bilinéaire symétrique sur $\mathrm{E} \times \mathrm{E}$ définie par $b(x,y) = \frac{1}{2}{}^2 b(x,y) = \frac{1}{2}(q(x+y) - q(x) - q(y))$.*

Les notions définies pour une forme bilinéaire se propagent aux formes quadratiques. Ainsi, on dira espace quadratique, morphisme d'espace quadratique, matrice d'une forme quadratique... pour espace bilinéaire, matrice d'une forme bilinéaire.... Un isomorphisme d'espaces quadratiques est souvent appelé une isométrie [2] par analogie avec le cas euclidien. Cette terminologie est cohérente avec 11.3.4.3. Si q est la forme quadratique sur $\mathbf{k}^n$ définie par $q(x_i) = \sum_{i \leq j} a_{i,j} x_i x_j,\ a_{i,j} \in \mathbf{k}$, sa matrice $\mathrm{M}(\mathcal{B}, q) = \mathrm{S}$ dans la base canonique $\mathcal{B}$ est définie par $\mathrm{S}_{i,j} = a_{i,j}/2$ si $\mathrm{S}_{i,i} = a_{i,i}$ avec la formule

$$b(\mathrm{X}, \mathrm{Y}) = {}^t\mathrm{XSY}$$

une fois identifié E à $\mathbf{k}^n$ grâce à $\mathcal{B}$. Comme précédemment, peut donc voir une forme quadratique q de matrice $\mathrm{M}(\mathcal{B}, q) = \mathrm{S}$ comme

1. une forme bilinéaire symétrique b ;
2. une matrice symétrique S ;
3. un polynôme homogène de poids 2,

points de vue équivalents qu'on utilisera librement. Rappelons (11.3.1), que q est invariante par a si et seulement si ${}^t\mathrm{ASA} = \mathrm{S}$ avec $\mathrm{A} = \mathrm{Mat}(\mathcal{B}, a)$.

Exemple(s) 13.5.1.2. — *Si (X, μ) est un espace mesuré, alors la formule*

$$q(f) = \int f^2 d\mu$$

définit une forme quadratique (en dimension infinie en général) sur $\mathrm{L}^2(\mathrm{X}, \mu; \mathbf{R}) \to \mathbf{R}$ de forme polaire

$$b(f,g) = \int f g d\mu.$$

— *Si $\varphi_1, \ldots, \varphi_r \in \mathrm{E}^*$ et si $(\lambda_i) \in \mathbf{k}^r$, alors la formule*

$$q(x) = \sum_{1 \leq j \leq r} \lambda_j (\varphi_j(x))^2$$

2. On trouve parfois dans la littérature le terme isométrie pour morphisme d'espaces quadratiques. On ne l'utilisera pas en ce sens car dans le cas dégénéré cela peut prêter à confusion, un tel morphisme n'étant pas nécessairement un isomorphisme contrairement au cas euclidien usuel duquel est issu la terminologie.

définit une forme quadratique sur E *de forme polaire*

$$b(x,y) = \sum_{1\leq j\leq r} \lambda_j \varphi_j(x)\varphi_j(y)\,.$$

Le lecteur vérifiera (exercice) que son rang est r *dès que les* λ_i *sont non nuls et les* φ_i *indépendantes.*

— *Si* $\mathrm{M}\in \mathrm{M}_n(\mathbf{k})$*, alors* $q(\mathrm{M}) = \mathrm{tr}({}^t\mathrm{MM})$ *définit une forme quadratique de forme polaire* $b(\mathrm{M},\mathrm{N}) = \mathrm{tr}({}^t\mathrm{MN})$

— *On définit sur* $\mathrm{M}_2(\mathbf{k})$ *la forme* $q(\mathrm{M}) = \det \mathrm{M}$*. On remarque que* q *est un polynôme quadratique homogène en les coefficients de* M*. De plus, (vérification directe ou théorème de Cayley-Hamilton), on a*

$$\mathrm{M}^2 - (\mathrm{tr}\,\mathrm{M})\cdot \mathrm{M} + (\det \mathrm{M})\cdot \mathrm{I}_2 = 0\,.$$

En prenant la trace, on trouve

$$q(\mathrm{M}) = \frac{(\mathrm{tr}\,\mathrm{M})^2 - \mathrm{tr}\,\mathrm{M}^2}{2}\,.$$

Du coup, la forme polaire associée est

$$b(\mathrm{M},\mathrm{N}) = \frac{(\mathrm{tr}\,\mathrm{M})(\mathrm{tr}\,\mathrm{N}) - \mathrm{tr}(\mathrm{MN})}{2}\,.$$

— *Si* $\mathbf{k}$ *est une extension finie de* $\mathbf{Q}$*, la multiplication par* $x\in\mathbf{k}$ *définit un endomorphisme* $\mathbf{Q}$ *linéaire et donc possède une trace notée* $\mathrm{tr}_{\mathbf{k}/\mathbf{Q}}(x)$*. L'application* $x\mapsto \mathrm{tr}_{\mathbf{k}/\mathbf{Q}}(x^2)$ *est une forme bilinéaire sur le* $\mathbf{Q}$*-espace vectoriel* $\mathbf{k}$ *dont on peut montrer sans trop de peine qu'elle est non dégénérée.*

13.5.2 Bases orthogonales

Rappelons (11.3.1) qu'une base $\mathcal{B}$ est orthogonale pour q si et seulement si sa matrice est diagonale ou encore si dans les les coordonnées associées on a $q(x) = \sum \lambda_i x_i^2$. Cette forme diagonale est notée traditionnellement $\langle \lambda_1,\dots,\lambda_n\rangle$. La considération d'une congruence diagonale $\mathrm{diag}(t_i)$ montre

$$\langle \lambda_1,\dots,\lambda_n\rangle \approxeq \langle t_1^2\lambda_1,\dots,t_n^2\lambda_n\rangle$$

que et donc $\langle \lambda_1,\dots,\lambda_n\rangle$ ne dépend que de la classe des λ_i dans $\mathbf{k}^*/(\mathbf{k}^*)^2$.

Théorème 13.5.2.1 (Gauss)**.** *Tout espace quadratique*[3] (E,q) *admet une base orthogonale* (e_i)*. Le rang de* q *est alors le nombre d'indices* i *tels que* $q(\mathrm{e}_i)\neq 0$*. De plus, une telle base peut-être obtenue par l'algorithme* infra*, dit algorithme de Gauss.*

Démonstration. *On procède par récurrence sur* n*. On peut supposer* $n>0$ *et* q *non nulle. Si* q *est dégénérée, la restriction on prend un vecteur non nul* e_n *du noyau et une base orthogonale (récurrence)*

3. La preuve se recopie dans le cas hermitien

de la restriction de q à un supplémentaire (nécessairement orthogonal) qui forment une base orthogonale. Si q est non dégénérée, on choisit alors $\mathrm{e}_{n+1} \in \mathrm{E}$ tel que $q(\mathrm{e}_{n+1}) \neq 0$ de sorte que l'orthogonal $\mathrm{H} = \mathrm{e}_n^{\perp}$ est un hyperplan ne contenant pas e_n puisqu'il est non isotrope (13.3.0.4). Il suffit de compléter e_n par une base orthogonale de (H, q) (récurrence). ■

Algorithme de Gauss. On part de $q(\mathrm{X}) = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j$ qu'on va récursivement transformer par changements de variables en $\sum \lambda_i \varphi_i^2$ où les φ forment une base de $\mathrm{E}^\star$.

A chaque étape, l'algorithme fait apparaître un carré et fait disparaître une coordonnée dans le reste de la forme.

Initialisation. On peut supposer $n \geq 2$ (en dimension ≤ 1 on n'a pas vraiment de choix...) et q non identiquement nulle (dans ce cas on prend pour φ_i les coordonnées associées à la base canonique par exemple et $\lambda_i = 0$).

Récursivité.

— Si il existe i tel que $a_{i,i} \neq 0$, quitte à permuter les variables, on peut supposer $a_{11} \neq 0$, on met x_1 en facteurs dans tous les monômes possibles :

$$\begin{aligned} q(\mathrm{X}) &= a_{11}\left(x_1^2 + \sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_1 x_j\right) + \sum_{2 \leq i \leq j \leq n} a_{ij} x_i x_j \\ &= a_{11}\left(x_1 + \frac{1}{2}\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j\right)^2 - \frac{a_{11}}{4}\left(\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j\right)^2 + \sum_{2 \leq i \leq j \leq n} a_{ij} x_i x_j \\ &= a_{11}\varphi_1(x)^2 + \sum_{2 \leq i \leq j \leq n} \alpha_{ij} x_i x_j \end{aligned}$$

avec $\varphi_1(x) = x_1 + \frac{1}{2}\sum_{j=2}^{n} \frac{a_{1j}}{a_{11}} x_j$ indépendant des $n-1$ formes $x_j, j \geq 2$ et on applique (récursivement) l'algorithme à $\sum_{2 \leq i \leq j \leq n} \alpha_{ij} x_i x_j$ qui n'a plus que $n-2$ variables.

— Si tous les termes carrés sont nuls ($a_{jj} = 0$), alors, quitte à permuter les variables, on peut supposer $a_{1,2} \neq 0$. On écrit

$$x_1 x_2 = \frac{(x_1 + x_2)^2 - (x_1 - x_2)^2}{4}\,.$$

On pose alors

$$\begin{cases} \varphi_1(x) = \dfrac{x_1 + x_2}{2} \\ \\ \varphi_2(x) = \dfrac{x_1 - x_2}{2} \end{cases}$$

de sorte que $q(x) = a_{1,2}\varphi_1(x)^2 - a_{1,2}\varphi_2(x)^2 + \sum_{2 \leq i \leq j \leq n} a_{ij} x_i x_j$ avec $\varphi_1, \varphi_2, x_i,\ i \geq 3$ indépendantes et on applique (récursivement) l'algorithme à $\tilde{q}$ qui n'a plus que $n-2$ variables.

Une fois qu'on a les φ_i, sa base ante-duale (11.3.4.1) est la base orthogonale cherchée.

Exercice(s) 13.5.2.2. *Programmer (en SAGE par exemple) l'algorithme de Gauss. Discuter sa stabilité numérique. Donner une version matricielle de l'algorithme en utilisant uniquement des congruences par des matrices de permutation et de transvections.*

Exemple(s) 13.5.2.3. *On considère sur* $\mathbf{R}^3$ *la forme*

$$q(x,y,z) = xy + yz + xz\,.$$

Dans la base canonique $\mathcal{B}$*, on a donc*

$$\mathrm{Mat}(q,\mathcal{B}) = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

On pose alors

$$\begin{cases} u = \dfrac{x+y}{2} \\ \\ v = \dfrac{x-y}{2} \end{cases}$$

Il vient

$$q(x,y,z) = u^2 - v^2 + (u-v)z + (u+v)z = u^2 + 2uz - v^2\,;$$

du coup,

$$q(x,y,z) = (u+z)^2 - v^2 - z^2\,.$$

Notons $\mathcal{B}$ *la base canonique, et*

$$\mathrm{P}^* = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

la matrice de passage de $\mathcal{B}^\star$ *à une base* $\mathcal{C}^\star$ *base duale d'une base orthogonale* $\mathcal{C}$ *que l'on veut déterminer. On calcule donc* $\mathrm{P} = ({}^t\mathrm{P}^\star)^{-1}$ *et on obtient*

$$\mathrm{P} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

Dans cette base, $q(x,y,z) = x^2 - y^2 - z^2$*.*

13.5.3 Plans quadratiques

Conformément à notre stratégie, étudions les plans quadratiques.

Proposition 13.5.3.1. *Soit* (E, q) *un espace quadratique de dimension* 2.
De quatre choses l'une :

1. *q est anisotrope ($-\operatorname{disc}(q) \notin (k^*)^2$) ;*
2. *il existe exactement une seule droite isotrope ($\operatorname{disc}(q) = 0$) , q est dégénérée et dans une base appropriée, $q(x, y) = ax^2$, pour $a \neq 0$;*
3. *il existe exactement deux droites isotropes ($-\operatorname{disc}(q) \in (k^*)^2$), dans une base appropriée $q(x, y) = xy$ et dans ce cas, on dit que* E *est hyperbolique.*
4. *il existe au moins trois droites isotropes ($\operatorname{disc}(q) = 0$) et q est identiquement nulle.*

Démonstration. *On peut supposer q non identiquement nulle.*

Quitte à multiplier q par un scalaire non nulle, q s'écrit alors dans une base orthogonale convenable[4] *$q(x, y) = x^2 - \lambda y^2$.*

Si λ n'est pas un carré ($\operatorname{disc} = (q)$ non carré) et q est anisotrope.

Sinon, on écrit $\lambda = \mu^2$ et $q(x, y) = x^2 - \mu^2 y^2 = (x - \lambda y)(x + \lambda y)$.

SI λ est nul, on est dans le cas 2) -l'axe vertical $x = 0$ est la seule droite isotrope- si λ est non nul on est dans le cas 3) -les droites d'équations $x - \lambda y = 0$ et $x - \lambda y = 0$ étant les seules droites isotropes. ■

Avec le vocabulaire de 13.3.0.3, on déduit la caractérisation suivante des plans hyperboliques.

Proposition 13.5.3.2. *Un plan quadratique est hyperbolique si et seulement si il est non dégénéré et isotrope ou encore si son discriminant est un carré non nul.*

13.5.4 Plans anisotropes

On peut donner les modèles des plans anisotropes de la manière suivante. Soit donc $\alpha \in \mathbf{k}^*$ tel que $-\alpha$ n'est pas un carré et soit $\mathrm{K} = k[\sqrt{-\alpha}] = \mathbf{k}[\mathrm{T}]/(\mathrm{T}^2 + \alpha)$. C'est un corps de dimension 2 sur $\mathbf{k}$ de base $\mathcal{B} = (1, \sqrt{-\alpha})$ et l'endomorphisme σ de matrice $\operatorname{diag}(1, -1)$ est un morphisme de corps (de même que la conjugaison complexe est un morphisme de corps...). Bien entendu, les éléments fixés par σ sont exactement les éléments de $\mathbf{k}$. Posons alors $\mathrm{N}(z =) = z\sigma(z) = x^2 - \alpha y^2$ pour tout $z = x + \sqrt{-\alpha} \in \mathrm{K}$: c'est une forme quadratique sur K à valeurs dans $\mathbf{k}$ dont la matrice dans $\mathcal{B}$ est $\operatorname{diag}(1, -\alpha)$. On notera simplement Π_α ce plan quadratique. D'après 13.5.3.1, tout plan anisotrope est équivalent à $\Pi_{-\operatorname{disc}(q)}$.

4. L'algorithme de Gauss se réduit alors à la factorisation canonique d'un polynôme de degré 2 ! À vrai dire l'énoncé revient à dire qu'un trinôme su second degré non nul admet au plus 2 racines...

13.5.5 Invariants de formes quadratiques

Définition 13.5.5.1. *Deux espaces* (E, q) *et* (E', q') *sont équivalents* ($\approxeq$) *s'il existe une isométrie* $u : (\mathrm{E}, q) \to (\mathrm{E}', q')$, i.e. *un isomorphisme* $u : \mathrm{E} \to \mathrm{E}'$ *tel que,pour tout* $x \in \mathrm{E}$, *on a* $q'(u(x)) = q(x)$, *autrement dit si les matrices de* q *et* q' *sont congruentes. Un invariant est une application sur le quotient correspondant* $\{(\mathrm{E}, q)\}/_{\approxeq}$.

On dispose pour le moment de deux invariants par congruences d'une forme quadratique q (ou d'une matrice symétrique) : le rang $\mathrm{rg}(\mathrm{M})$ et le discriminant $\mathrm{disc}(q) = \det(\mathrm{M}) \in k/k^{*2}$ (11.3.2) d'une matrice de q. On va voir que si $\mathbf{k}$ algébriquement clos, le rang classifie les formes quadratiques (13.5.7.1). Dans ce cas, on a d'ailleurs $k/k^{*2} = \{0, 1\}$ et $\mathrm{disc}(q) = \mathrm{sign}(r)$.

Nous verrons que dans le cas des corps finis, rang et discriminant classifient les formes quadratiques (13.5.9.2).

En général, ces deux invariants ne sont pas suffisants. Par exemple, les formes réelles à quatre variables $x^2-y^2-z^2-t^2$ et $x^2+y^2-z^2-t^2$ ont même rang mais ne sont pas équivalentes car leurs discriminants sont -1 et 1 respectivement qui sont différents dans $\mathbf{R}/\mathbf{R}^{*2}\{-1, 0, 1\}$. On va voir (13.5.8.1) qu'un troisième invariant est toutefois nécessaire, l'indice (13.5.6.1), ces trois invariants étant résumés dans la signature de la forme quadratique réelle.

Dans tous ces cas, il n'y a qu'un nombre fini de classes d'équivalence. Ce n'est pas vrai en général.

Donnons un exemple. On définit, pour tout nombre premier p la forme quadratique

$$q_p(x) = \sum_{1 \leq j < n} x_j^2 + p x_n^2$$

sur $\mathbf{Q}^n$. Elles sont deux à deux non équivalentes dès que $n > 0$ et se différencient par leur discriminant qui est $p \mod (\mathbf{Q}^{*2})$ (cf. l'exercice 13.6.0.3). On sait classifier en général sur $\mathbf{Q}$ mais il faut de nouveaux invariants reliés à la classification sur les corps finis, les symboles de Hilbert (cf. le magnifique ouvrage [Ser77]).

De manière générale, la classification est un problème extrêmement difficile. On peut voir cela de manière inverse : les formes quadratiques permettent de définir des invariants de corps subtiles (cf. l'exercice 13.6.0.2).

13.5.6 Isotropie et indice

Nous allons définir un troisième invariant : l'indice.

Définition 13.5.6.1. *L'*indice ν *d'une forme quadratique* q *la dimension maximale des espaces totalement isotropes. Si* $\nu = 0$ i.e., *si* $q(x) = 0 \Rightarrow x = 0$, *on dit que* q *est anisotrope ou définie.*

Par exemple, si $\mathbf{k} = \mathbf{R}$, un argument de continuité assure si q est définie, alors, ou bien pour tout $x \neq 0$ on a $q(x) > 0$, ou bien, pour tout $x \neq 0$ on a $q(x) < 0$.

On va en déduire la décomposition générale d'un espace (E, q). On commence par un lemme.

Lemme 13.5.6.2. *Soit (E, q) un espace quadratique non dégénéré.*

1. *Si x est isotrope, il existe un plan hyperbolique contenant x.*
2. *L'indice ν de $\overset{\perp}{\oplus}_{1\leq j\leq r} \mathrm{P}_j$ est r.*
3. *Il existe r plans hyperboliques P_j et (F, q) est anisotrope tel que*
$$\mathrm{E} = (\overset{\perp}{\oplus}_{1\leq j\leq r} \mathrm{P}_j) \overset{\perp}{\oplus} \mathrm{F}.$$
4. *Dans une telle décomposition existe, on a $r = \nu$.*

Si q est quelconque, on a une décomposition
$$\mathrm{E} = \mathrm{rad}(\mathrm{E}) \overset{\perp}{\oplus} (\overset{\perp}{\oplus}_{1\leq j\leq r} \mathrm{P}_j) \overset{\perp}{\oplus} \mathrm{F}$$
avec $r + \dim(\mathrm{rad}(\mathrm{E})) = \nu$.

Démonstration. *Il existe y tel que $b(x, y) \neq 0$. Du coup, x et y engendrent bien un plan quadratique dont le discriminant $-b(x, y)^2 \neq 0$. Comme il a une droite isotrope, c'est un plan hyperbolique (13.5.3.1) d'où 1).*

Pour 2), on peut et supposer q isotrope et $n \geq 3$ (d'après 13.5.3.1). Soit alors P un plan hyperbolique contenu dans E. Montrons $\mathrm{E} = \mathrm{P} \overset{\perp}{\oplus} \mathrm{P}^\perp$. Comme q est non dégénérée, les dimensions sont les bonnes. Si $v \in \mathrm{P} \cap \mathrm{P}^\perp$ alors $b(v, \mathrm{P}) = \{0\}$ contredisant la non dégénérescence des plans isotropes. Ceci contredit que q est non dégénérée. Donc $\mathrm{E} = \mathrm{P} \overset{\perp}{\oplus} \mathrm{P}^\perp$ et on applique l'hypothèse de récurrence à $\mathrm{P}^\perp$.

Pour 3), notons $\mathrm{e}_i, \mathrm{e}'_i$ une base des deux droites isotropes de P_i. On a évidemment $\nu \geq r$ car $\mathrm{Vect}(\mathrm{e}_i)$ est totalement isotrope de dimension r. Comme $\mathrm{Vect}(\mathrm{e}_i + \mathrm{e}'i)$ est anisotrope de codmension r, on a également $\nu \geq r - r'$.

Pour le 4), soit G isotrope de dimension ν et notons p la projection orthogonale sur F (parallèlement à $\overset{\perp}{\oplus}_{1\leq j\leq r} \mathrm{P}_j$ donc). L'espace $p(\mathrm{G})$ est isotrope par construction donc nul puisque F est anisotrope. On déduit que G est inclus dans $\overset{\perp}{\oplus}_{1\leq j\leq r} \mathrm{P}_j$ et donc $\nu \leq r$ d'après 3). Inversement, comme $\mathrm{Vect}(\mathrm{e}_i)$ est isotrope, on a $r \leq \nu$.

Le dernier point découle des précédents de l'additivité des radicaux par somme directe orthogonale. ∎

Théorème 13.5.6.3 (Simplification de Witt)**.** *Soit (E, q) un espace quadratique et*
$$\mathrm{E} = \mathrm{rad}(\mathrm{E}) \oplus (\overset{\perp}{\oplus}_{1\leq j\leq \nu} \mathrm{P}_j) \overset{\perp}{\oplus} \mathrm{F}$$
comme dans le lemme 13.5.6.2.

1. *La classe d'isomorphisme quadratique de* F *est bien déterminée et* $r = \nu$.
2. *Si on a un isomorphisme d'espaces quadratiques*

$$\mathrm{E} \overset{\perp}{\oplus} \mathrm{F} \approxeq \mathrm{E} \overset{\perp}{\oplus} \mathrm{F}'$$

alors

$$\mathrm{F} \approxeq \mathrm{F}'.$$

3. *Si deux familles de scalaires non nulles vérifient*

$$\langle a, a_1, \ldots, a_n\rangle \approxeq \langle a, b_1, \ldots, b_n\rangle$$

alors

$$\langle a_1, \ldots, a_n\rangle \approxeq \langle b_1, \ldots, b_n\rangle$$

Démonstration. *On va montrer 3) puis les deux implications* $3) \Rightarrow 2)$ *et* $2) \Rightarrow 1)$.
Preuve de 3). On note $x_i,\ 0 \leq i \leq n$ *les formes linéaires de coordonnées sur* $\mathbf{k}^{n+1}$, *autrement dit la base duale de la base canonique* (e_i). *Dire*

$$\langle a, a_1, \ldots, a_n\rangle \approxeq \langle a, b_1, \ldots, b_n\rangle$$

c'est dire l'existence de formes linéaires $\varphi_i \in (\mathbf{k}^{n+1})^\star$ *indépendantes(les lignes de la matrice de passage définissant la congruence) telles que*

$$(*) \qquad ax_0^2 + \sum_{i\geq 1} a_i x_i^2 = a\varphi_0(x)^2 + \sum_{i\geq 1} b_i \varphi_i^2$$

Au moins une des deux formes linéaires $x_0 \pm \varphi(x_0)$ *n'annule pas* e_0 *: soit* $\varepsilon = \pm 1$ *tel que* $\langle x_0 + \varepsilon\varphi_0(x), \mathrm{e}_0\rangle = \lambda \neq 0$ *de sorte que* $x_0 + \varepsilon\varphi_0(x)$ *s'écrit sous la forme*

$$x_0 + \varepsilon\varphi_0(x) = \lambda x_0 - \lambda\psi(x_1, \cdots, x_n)$$

avec ψ *forme linéaire sur* $\mathbf{k}^n$. *Si on fait la substitution* $x_0 = \psi(x_1, \cdots, x_n)$ *dans* $(*)$, *on a donc*

$$\sum_{i\geq 1} a_i x_i^2 = \sum_{i\geq 1} b_i \varphi_i(\psi(x_1, \ldots, x_n), x_1, \ldots, x_n)^2 = \sum_{i\geq 1} b_i \tilde{\varphi}_i(x_1, \ldots, x_n)^2$$

où $\tilde{\varphi}$ *sont des forme linéaires sur* $\mathbf{k}^n$. *Si* $\Psi \in \mathrm{M}_n(\mathbf{k})$ *est la matrice qu'elles définissent, on a donc* $\mathrm{diag}(a_i) = {}^t\Psi\,\mathrm{diag}(b_i)\Psi$ *ce qui assure l'inversibilité de* Ψ *en prenant les déterminants et donc*

$$\langle a_1, \ldots, a_n\rangle \approxeq \langle b_1, \ldots, b_n\rangle$$

d'où 3).
$3) \Rightarrow 2)$ *On note* q, q' *les formes quadratiques associées,* r *le rang de la restriction de* q *à* E, *et* ρ, ρ' *ceux des restrictions de* q, q' *à* F, F'. *Comme les sommes sont orthogonales, on* $\mathrm{rang}(q) = r + \rho = r + \rho'$ *de sorte que* $\rho = \rho'$ *et bien entendu* $\dim(\mathrm{F}) = \dim(\mathrm{F}')$. *Le radical de* q *et* q' *étant engendrés par les vecteurs*

des bases orthogonales correspondantes d'indices tels $a_i = b_j = 0$ et $a_i = b'_j = 0$, on peut en passant au quotient par les radicaux supposer que les formes sont non dégénérées. On a donc

$$\langle a_1, \dots, a_r, b_1, \dots, b_\rho \rangle \approx\!\!\!\approx \langle a_1, \dots, a_r, b'_1, \dots, b'_\rho \rangle$$

de sorte que

$$\langle b_1, \dots, b_\rho \rangle \approx\!\!\!\approx \langle b'_1, \dots, b'_\rho \rangle$$

d'après 3) ce qui prouve 2).

2) $\Rightarrow$ 1) *Comme plus haut, on se ramène au cas non dégénéré par passage au quotient. Supposons que*

$$(\overset{\perp}{\oplus}_{1\le j\le r} \mathrm{P}_j) \overset{\perp}{\oplus} \mathrm{F} \approx\!\!\!\approx (\overset{\perp}{\oplus}_{1\le j\le r'} \mathrm{P}'_j) \overset{\perp}{\oplus} \mathrm{F}'$$

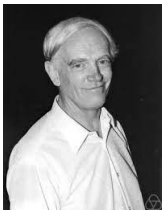
avec F, F' anisotropes, les $\mathrm{P}_i, \mathrm{P}'_j$ hyperboliques et par exemple $r' \le r$. On a déjà vu dans le lemme 13.5.6.2 $r = \nu = r'$. Grâce à 2), on déduit alors $\mathrm{F} \approx\!\!\!\approx \mathrm{F}'$. ∎

La clef de cet important résultat est donc le point 3), le reste étant assez formel. Mais finalement c'est juste un calcul très simple mais qui à ma connaissance n'est apparu que récemment ([CMM17]). La preuve classique est proposée en exercice plus bas (13.6.0.1). Bien entendu, d'après la discussion précdente, tout espace quadratique (non dégénéré ou pas) se décompose en

$$\mathrm{E} = \mathrm{Ker}(q) \overset{\perp}{\oplus} (\overset{\perp}{\oplus}_{1\le j\le \nu} \mathrm{P}_j) \overset{\perp}{\oplus} \mathrm{F}$$

avec F bien défini à isométrie près (et anisotrope).

On a donc défini un quatrième invariant : la partie anisotrope $(\mathrm{E}, q)^{\text{☺}} = (\mathrm{F}, q) \mod \approx\!\!\!\approx$ de (E, q).



Ernst Witt

La personnalité de Witt est controversée de par, *a minima*, sa collaboration active avec le régime nazi (membre du parti nazi dès 1933 puis SA). La personnalité complexe de Witt semble toutefois un peu relativiser son action. Voir `https://mathshistory.st-andrews.ac.uk/Biographies/Witt/` pour une biographie.

13.5.7 Classification sur un corps algébriquement clos

Une base $\mathcal{B}$ de E de dimension n étant donnée, comme toujours $(x_j) = [x]_{\mathcal{B}}$ désigne les coordonnées de x dans $\mathcal{B}$.

Théorème 13.5.7.1. *Si* **k** *est algébriquement clos, alors, pour tout q, il existe une base $\mathcal{B}$ telle que*

$$q(x) = \sum_{1\le j\le \mathrm{rg}(q)} x_j^2.$$

On a exactement $n+1$ classes d'équivalences, qui se différencient à l'aide du rang de q.

Démonstration. *Dans une base orthogonale* $\mathcal{B} = (e_1, \dots, e_n)$*, on a*

$$q(x) = \sum_{1 \leq j \leq \mathrm{rg}(q)} \lambda_j x_j^2,$$

avec $\lambda_j \in \mathbf{k}^*$*. Comme* $\mathbf{k}$ *est algébriquement clos, il existe* $\mu_j \in \mathbf{k}$ *tel que* $\mu_j^2 = \lambda_j$*. En posant* $f_j = e_j/\mu_j$ *si* $i \leq \mathrm{rg}(q)$ *et* $f_j = e_j$ *sinon , on obtient*

$$q(x) = \sum \lambda_j x_j^2 = \sum (\mu_j x_j)^2 = \sum \xi_j^2$$

avec $(\xi_j = \mu_j x_j)$ *sont les coordonnées de* x *dans la base* (f_j) ∎

Exercice(s) 13.5.7.2. *Vérifier dans ce cas les formules* $\nu = \dim(\mathrm{rad}(q)) + [\mathrm{rg}(q)/2]$ *et* $E^{☺} = \{0\}$ *ou* $E^{☺} \approxeq \langle 1 \rangle$ *suivant la parité du rang.*

13.5.8 Classification sur R

Théorème 13.5.8.1 (Inertie de Sylvester)**.** *Si* $\mathbf{k} = \mathbf{R}$*, alors, pour tout* q*, il existe un unique couple d'entiers naturels* (s,t) *appelé la signature de* q *tel qu'il existe une base dans laquelle*

$$q(x) = \sum_{1 \leq j \leq s} x_j^2 - \sum_{s+1 \leq j \leq \mathrm{rg}(q)} x_j^2.$$

On a alors

1. $s + t = \mathrm{rg}(q)$.
2. $s + \dim(\mathrm{Ker}(q) = \max\{\dim(F) | q_{|F} \geq 0\}$
3. $t + \dim(\mathrm{Ker}(q) = \max\{\dim(F) | q_{|F} \leq 0\}$

On a exactement $(\dim E+1)(\dim E+2)/2$ *classes d'équivalence, qui se distinguent à l'aide de la signature de* q*.*

Démonstration. *En séparant les ans une base orthogonale convenable* $\mathcal{B} = (e_1, \dots, e_n)$*, on a dans une base orthogonale* $\mathcal{B} = (e_1, \dots, e_n)$*, on a*

$$q(x) = \sum_{1 \leq j \leq \mathrm{rg}(q)} \lambda_j x_j^2,$$

avec $\lambda_j \in \mathbf{R}^*$*. On peur suppose que les* s *premiers scalaires sont positifs et les* t *derniers négatifs. On pose alors* $\mu_j = \sqrt{|\lambda_j|}$ *et* $f_j = e_j$ *sinon , on obtient*

$$q(x) = \sum_{1 \leq j \leq s} \xi_j^2 - \sum_{s+1 \leq j \leq \mathrm{rg}(q)} \xi_j^2$$

avec $(\xi_j = \mu_j x_j)$ *sont les coordonnées de* x *dans la base* (f_j)*. Il reste à montrer les points 1), 2) et 3), le reste suivant immédiatement. Le 1) est clair et le 3) découle de 2) en changeant* q *en* $-q$*.*
Soit alors, $F' = \mathrm{Vect}\{e_1, \dots, e_s\} \overset{\perp}{\oplus} \mathrm{Ker}(q)$ *et* $G = \mathrm{Vect}\{e_{s+1}, \dots, e_n\}$*. Comme* $n = \dim(\mathrm{Ker}(q)) + \mathrm{rg}(q)$ *on* $\dim(G) = t$*. Comme* q *est* ≥ 0 *sur* F'*, on a* $s + \dim(\mathrm{Ker}(q)) \leq \max\{\dim(F) | q_{|F} \geq 0\}$*. Inversement, si*

il existe F *de dimension* $> s + \dim(\mathrm{Ker}(q)) = n - t$, *on aurait* $\dim(\mathrm{F} \cap \mathrm{G}) > 0$ *et donc un vecteur* x *tel que* $q(x) \geq 0$ *car* $x \in \mathrm{F}$ *et* $q(x) < 0$ *car* $x \in \mathrm{G} - \{0\}$. ■

Exercice(s) 13.5.8.2. *Vérifier dans ce cas les formules* $\nu = \dim(\mathrm{rad}(q)) + inf(p,q)$ *et* $\dim(\mathrm{E}^{\text{☺}}) = |s-t|$ *avec* $\mathrm{E}^{\text{☺}} = \langle \mathrm{signe}(s-t), \ldots, \mathrm{signe}(s-t) \rangle$ *si* $s \neq t$.

13.5.9 Classification sur les corps finis.

Si $\mathbf{k}$ est est un corps fini (de caractéristique différente de 2), on sait que son cardinal est de la forme p^d où $p \neq 2$ est sa caractéristique -et donc est premier- (simplement car c'est un $\mathbf{F}_p$-vectoriel). Conformément à la stratégie générale, regardons ce qui se passe dans le cas crucial de la dimension ≤ 2 en commençant par la dimension 1, à savoir comprendre le discriminant.

Lemme 13.5.9.1. *Le groupe multiplicatif* $\mathbf{F}_{p^d}^*/(\mathbf{k}^*)^2$ *a deux éléments* $1 \neq \alpha$ *de sorte que le discriminant d'une forme quadratique sur* $\mathbf{k}$ *est à valeurs dans l'ensemble* $\{0, 1, \alpha\}$. *Il existe exactement deux plans quadratiques non dégénérés équivalents à* $\langle 1, 1\rangle$ *et* $\langle 1, \alpha\rangle$ *distingués par le discriminant.*

Démonstration. *On a une suite exacte* $\{1\} \to \mathrm{Ker}(sq) \to \mathbf{k}^* \xrightarrow{sq} (\mathbf{k}^*)^2 \to \{1\}$ *de groupe multiplicatifs où sq est le morphisme d'élévation au carré. Comme* $\mathbf{k}$ *est intègre, l'équation* $x^2 = 1$ *a pour solutions* ± 1. *Comme* $p \neq 2$, *on a* $1 \neq -1$. *On déduit* $\mathrm{Card}((\mathbf{k}*)^2) = \frac{q-1}{2}$ *et* $\mathrm{Card}(\mathbf{k}^*/(\mathbf{k}^*)^2) = 2$ *ce qui prouve le premier point.*

Soit alors un plan quadratique non dégénéré de forme q. *On peut supposer* $q(x_1, x_2) = ax_1^2 + bx_2^2$ *avec* $ab \neq 0$. *Or, il y a* $1 + \frac{q-1}{2} = \frac{q+1}{2}$ *carrés dans* $\mathbf{k}$ *de sorte que les cardinaux de* $\{at^2,\ x \in \mathbf{k}|\}$ *et* $\{1 - bt^2,\ y \in \mathbf{k}|\}$ *valent* $\frac{q+1}{2}$ *et donc ont au moins un point d'intersection qui définit* $\mathrm{e}_1 = (x_1, x_2)$ *tel que* $q(x_1, x_2) = 1$. *Soit alors* e_2 *base de l'orthogonal de* e_1. *En écrivant* q *dans cette base, on a* $q \approx\!\!\approx \langle 1, \mathrm{disc}(q)\rangle$ *(cf. 13.5.2 pour l'invariance par congruence de l'écriture).* ■

On a alors.

Théorème 13.5.9.2. *Si* $\mathbf{k}$ *est fini de caractéristique* $p \neq 2$, *alors, toute forme non dégénérée est (uniquement) équivalente à* $\langle 1, \ldots, 1\rangle$ *ou* $\langle 1, \ldots, 1, \alpha\rangle$. *Ces classes sont distinguées par leurs discriminants.*

Démonstration. *On procède par récurrence sur* n. *On peut supposer* $n \geq 3$ *et* (e_i) *base orthogonale. Comme le plan quadratique* $\mathrm{Vect}(\mathrm{e}_1, \mathrm{e}_2)$ *est non dégénéré, on peut choisir* ε_1 *dans ce plan* $q(\varepsilon_1) = 1$. *On a alors* $\mathrm{E} = \mathbf{k}\varepsilon_1 \overset{\perp}{\oplus} \varepsilon_1^{\perp}$ *de sorte que* $\varepsilon_1^{\perp}$ *est un hyperplan quadratique à qui on peut appliquer l'hypothèse de récurrence. Le second point est clair.* ■

Exercice(s) 13.5.9.3. *Soit* $\mathbf{k}$ *est fini de caractéristique* $p \neq 2$*et* q *non dégénérée sur* E.

1. *En utilisant* $\mathrm{Card}(\mathbf{k}^*) = \frac{p^d-1}{2}$, *montrer que* -1 *est un carré si et seulement si* $p \equiv 1 \mod (4)$.
2. *Calculer* $\nu(\mathrm{E})$ *et* $\mathrm{E}^{\text{☺}}$ *en fonction de* $n, \mathrm{disc}(q)$ *et* $p \mod (4)$.

13.5.10 Le théorème de prolongement de Witt

Théorème 13.5.10.1 (de Witt). *Soit $u : \mathrm{F} \to \mathrm{E}$ un morphisme injectif d'espaces quadratiques avec E non dégénéré. Alors, il existe $\tilde{u} \in \mathrm{O}(q)$ tel que $\tilde{u}_{|\mathrm{F}} = u$.*

Démonstration. *Observons que l'injectivité du morphisme u assure $\mathrm{F} \approxeq u(\mathrm{F})$. On procède par récurrence sur $\mathrm{codim}(\mathrm{F})$. On peut supposer $\mathrm{codim}(\mathrm{F}) > 0$.*
- Si $q_{|\mathrm{F}}$ non dégénérée, i.e.$\mathrm{K} = \mathrm{F} \cap \mathrm{F}^\perp = \{0\}$, *alors $\mathrm{E} = \mathrm{F} \overset{\perp}{\oplus} \mathrm{F}^\perp$. Si $u(f) \in u(\mathrm{F}) \cap u(\mathrm{F})^\perp$, on a $0 = b(u(f), u(\mathrm{F}^\perp)) = b(f, \mathrm{F}^\perp) = \{0\}$ donc $f = 0$ et ainsi $u(\mathrm{F}) \cap u(\mathrm{F})^\perp = \{0\}$ de sorte que*

$$\mathrm{F} \overset{\perp}{\oplus} \mathrm{F}^\perp = u(\mathrm{F}) \overset{\perp}{\oplus} u(\mathrm{F})^\perp \approxeq \mathrm{F} \overset{\perp}{\oplus} u(\mathrm{F})^\perp$$

de sorte que $\mathrm{F}^\perp \approxeq u(\mathrm{F})^\perp$ d'après le théorème de simplification de Witt 13.5.6.3 . Un peut donc choisir un isomorphismes d'espaces quadratiques $u^\perp : \mathrm{F}^\perp \to u(\mathrm{F})^\perp$ et $\tilde{u} \overset{\perp}{\oplus} u^\perp \in \mathrm{O}(q)$ convient.
- Si $q_{|\mathrm{F}}$ est dégénérée, choisissons x non nul dans $\mathrm{F} \cap \mathrm{F}^\perp$ et un plan isotrope P qui le contient (13.5.6.2). Si y dirige la seconde droite isotrope de P, on a $b(x, y) = 1$ de sorte que $y \notin \mathrm{F}$ et

$$\mathrm{codim}(\mathrm{G} = \mathrm{F} \oplus \mathbf{k}y) = \mathrm{codim}(\mathrm{F}) - 1$$

Il suffit de prolonger u à G (en préservant b). On cherche donc $z \in \mathrm{E}$ tel que

1. $b(u(f), z) = b(f, y)$ pour tout $f \in \mathrm{F}$
2. $b(u(x), z) = 1$
3. $b(z, z) = 0$

qui assure que le prolongement de u défini par $\tilde{u}(y) = z$ est une isométrie (sur son image). Notons que $1) \Rightarrow 2)$ et que changer z en $z + \lambda u(x)$ ne change pas 1) car $x \in \mathrm{F}^\perp$. Or,

$$b(z + \lambda u(x), z + \lambda u(x)) = b(z, z) + 2\lambda b(u(x), z) = b(z, z) + 2\lambda b(x, y) = b(z, z) + 2\lambda$$

de sorte qu'en changeant z en $z - \frac{b(z,z)}{2}x$, on aura le prolongement cherché. Il suffit donc de trouver z vérifiant 1).
Soit alors S un supplémentaire quelconque de $u(\mathrm{F})$ dans E et $\varphi \in \mathrm{E}^*$ la forme linéaire nulle sur S et valant $u^{-1}(t), y$ pour tout $t \in u(\mathrm{F})$ -rappelons que u est supposée injective donc bijective de F sur $u(\mathrm{F})$-. Comme b est non dégénérée, il existe $z \in \mathrm{E}$ tel que $\varphi = b(., z)$ et z vérifie 1).

13.6 Exercices

Exercice(s) 13.6.0.1. *TBD*

Exercice(s) 13.6.0.2. *TBD*

Exercice(s) 13.6.0.3. *Soit $\mathcal{P}$ l'ensemble des nombres premiers, $v_p(x)$ l'exposant de p dans la décomposition en facteurs irréductibles de $x \in \mathbf{Q}^*$ et $v_\infty(x) \in \mathbf{Z}/2\mathbf{Z}$ défini par $sign(x) = (-1)^{v_\infty(x)}$. Montrer que l'application $x \mapsto\to (v_i(x))$ défini un isomorphisme de groupes $\mathbf{Q}^*/(\mathbf{Q}^*)^2 \simeq (\mathbf{Z}/2\mathbf{Z})^{(\mathcal{P} \cup \{\infty\})}$.*

Chapitre 14

Le groupe orthogonal d'une forme quadratique non dégénérée

Dans ce chapitre (E, q) désigne un espace quadratique non dégénéré de dimension $n > 0$ et S la matrice (inversible) de q dans une base $\mathcal{B}$ donnée.

14.1 Définition

Notons qu'un endomorphisme u de matrice $\mathrm{M} = \mathrm{Mat}(u, \mathcal{B})$ préserve q si et seulement si ${}^t\mathrm{MSM} = \mathrm{S}$ de sorte que $\det(\mathrm{M}) \neq 0$. Ainsi, u est automatiquement une isométrie (ou un endomorphisme orthogonal comme on préfère, 11.3.4.3). L'ensemble des isométries de (E, q) est un sous-groupe $\mathrm{O}(q)$ de $\mathrm{GL}(\mathrm{E})$. On remarque que les groupes orthogonaux de deux formes quadratiques équivalentes sont isomorphes (exercice).

Remarque(s) 14.1.0.1. *Si b est hermitienne ou alternée, on peut encore parler d'isométries. Dans le cas hermitien, on parle du groupe unitaire* $\mathrm{U}(b)$ *et dans le cas alterné du groupe symplectique* $\mathrm{Sp}(b)$

.

Puisque $u \in \mathrm{O}(q)$ si et seulement $u^* = u^{-1}$ c'est-à-dire si

$$ {}^t\mathrm{Mat}(u, \mathcal{B}) \cdot \mathrm{Mat}(q, \mathcal{B}) \cdot \mathrm{Mat}(u, \mathcal{B}) = \mathrm{Mat}(q, \mathcal{B}). $$

on a $(\det u)^2 = 1$. On définit alors le groupe spécial orthogonal $\mathrm{SO}(q)$ comme le sous-groupe normal des isométries de déterminant 1 de sorte qu'on a une suite exacte

$$ \{1\} \to \mathrm{SO}(q) \to \mathrm{O}(q) \to \{\pm 1\} \to \{1\} $$

14.2 Le cas de la dimension 2

Un plan quadratique non dégénéré est soit hyperbolique soit anisotrope (13.5.3.1) suivant que $-\operatorname{disc}(q)$ est un carré ou non. Étudions $\mathrm{SO}(q)$ dans chacun de ces cas.

Proposition 14.2.0.1. *Le groupe spécial orthogonal d'un plan hyperbolique est commutatif et isomorphe à* $\mathbf{k}^*$. *Explicitement* $q(x,y) = xy$, *on a* $\mathrm{SO}(q) = \{\mathrm{diag}(a, a^{-1}),\ a \in \mathbf{k}^*\} \simeq \mathbf{k}^*$.

Démonstration. *On choisit des coordonnées de telle sorte que* $q(x,y) = xy$ *de sorte que* $\mathrm{SO}(q)$ *s'identifie aux matrices* $\mathrm{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *de déterminant* 1 *telles que* ${}^t\mathrm{MSM} = \mathrm{S}$ *avec* $\mathrm{S} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. *Les équations correspondantes sont alors*

$$\begin{aligned} 2ac &= 0 \\ bc + ad &= 1 \\ 2bd &= 0 \\ ad - bc &= 1 \end{aligned}$$

Des seconde et quatrième on tire $ad = 1$ *donc* a *et* d *inversibles. De la première et la seconde on tire alors* $b = c = 0$. ∎

Soit $-\alpha$ non carré dans $\mathbf{k}^*$ et q_α de discriminant α. Rappelons (13.5.4) qu'un plan anisotrope est isométrique au corps $\mathrm{K} = \mathbf{k}[\sqrt{-\alpha}]$ (isomorphe $\mathbf{k}^2$ comme $\mathbf{k}$-espace vectoriel) muni de la forme $\mathrm{N}(z) = z\sigma(z)$ où $\sigma(x + \sqrt{-\alpha}y) = x - \sqrt{-\alpha}y$.

Proposition 14.2.0.2. *Le groupe spécial orthogonal d'un plan hyperbolique est commutatif et isomorphe au sous-groupe* $\{z \in \mathrm{K}^* |\ \mathrm{N}(z) = 1\}$ *de* K^* *qui agit par multiplication sur* K *muni de la forme quadratique* N.

Démonstration. *La matrice de la multiplication par* $z = a + \alpha c$ *dans* K *dans sa base naturelle* $(1, \sqrt{-\alpha})$ *est* $\mathrm{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *et le déterminant de* M *est* $\mathrm{N}(z)$. *Dans cette base,* $q(x,y) = x^2 + \alpha y^2$ *de sorte qu'on a bien défini un morphisme injectif de* $\{z \in \mathrm{K}^* |\ \mathrm{N}(z) = 1\}$ *dans* $\mathrm{SO}(q_\alpha)$. *Pour la surjectivité, c'est un calcul analoque au précédent avec* $q(x,y) = x^2 + \alpha y^2$ *de sorte que* $\mathrm{SO}(q)$ *s'identifie aux matrices* $\mathrm{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *de déterminant* 1 *telles que* ${}^t\mathrm{MSM} = \mathrm{S}$ *avec* $\mathrm{S} = \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix}$.
Comme $\det(\mathrm{M}) = 1$, *on a*

$$\begin{pmatrix} a & \alpha c \\ \alpha^{-1} b & d \end{pmatrix} = \mathrm{S}^{-1\,t}\mathrm{MS} = \mathrm{M}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

et donc $a = d,\ c = -b/\alpha$ *avec* $\det(\mathrm{M}) = a^2 + \alpha c^2 = 1$ *ce qui prouve la surjectivité.* ■

Le lecteur reconnaîtra pour $\mathbf{k} = \mathbf{R}$ et $\alpha = 1$ l'isomorphisme du groupe des rotations planes avec le cercle unité du plan complexe.

14.3 Symétries orthogonales

Une symétrie est un endomorphisme u tel que $u \circ u = \mathrm{Id}$. En particulier, elles sont inversibles. De plus, les valeurs propres sont ± 1, et il existe une décomposition de E en somme directe $\mathrm{E} = \mathrm{E}_+ \oplus \mathrm{E}_-$, où E_+ est l'espace propre associé à la valeur propre 1, et E_- à la valeur propre -1.

Proposition 14.3.0.1. *Une symétrie est orthogonale si et seulement si les* E_+ *et* E_- *sont orthogonaux. Dans ce cas, ces espaces sont non isotropes.*
Réciproquement, si F *est un sous-espace non isotrope, alors il existe une unique symétrie orthogonale telle que* F *soit exactement l'espace propre associé à la valeur propre* 1.

Démonstration. *Si* u *est orthogonale, alors, pour* $x \in \mathrm{E}_+$ *et* $y \in \mathrm{E}_-$, *on a*

$$b(x, y) = b(u(x), u(y)) = -b(x, y)$$

donc $b(x, y) = 0$ *car car* $\mathbf{k} \neq 2$.
Réciproquement, si ces espaces sont orthogonaux, alors, soient $x, y \in \mathrm{E}$. *On écrit*

$$\begin{cases} x = x_+ + x_-, & (x_+, x_-) \in \mathrm{E}_+ \times \mathrm{E}_- \\ y = y_+ + y_-, & (y_+, y_-) \in \mathrm{E}_+ \times \mathrm{E}_- \end{cases}$$

Il vient

$$b(u(x), u(y)) = b(x_+, y_+) + b(x_-, y_-) = b(x, y)\,.$$

Soit F *non isotrope. On note* $\mathrm{H} = \mathrm{F}^\perp$. *On a donc* $\mathrm{E} = \mathrm{F} \overset{\perp}{\oplus} \mathrm{H}$, *et on peut définir* $u \in \mathrm{O}(q)$ *par* $u|_{\mathrm{F}} = \mathrm{Id}$ *et* $u|_{\mathrm{H}} = -\mathrm{Id}$. ■

Définition 14.3.0.2. *Lorsque* $\dim \mathrm{E}_- = 1$, *on dit que* u *est une* réflexion, *et quand* $\dim \mathrm{E}_- = 2$, *on parle de* renversement *(demi-tour d'axe* E_+*).*

Exemple(s) 14.3.0.3. *Si* y *est anisotrope, l'endomorphisme* s_y *défini par* $s_y(x) = x - 2\frac{b(x,y)}{b(y,y)}y$ *est la réflexion avec* $\mathrm{E}_+ = \{y\}^\perp$ *et* $\mathrm{E}_- = \mathbf{k}y$. *En particulier, si* $q(x) = q(y)$ *et* $x - y$ *anisotrope, on a* $s_{x-y}(x) = y$. *Notons au passage que la formule précédente prouve que la restriction d'une réflexion* s_x *à un espace stable contenant* x *est encore une réflexion et que l'opposé d'une réflexion en dimension* 3 *est un retournement.*

14.4 Similitudes

De manière générale, il s'agit des endomorphismes u tels qu'il existe un scalaire $\lambda \in \mathbf{k}^*$ tel que $b(u(x), u(y)) = \lambda \cdot b(x, y)$. Ils forment un groupe $\mathrm{GO}(q)$ et on a la suite exacte

$$1 \to \mathrm{O}(q) \to \mathrm{GO}(q) \to \mathbf{k}^*$$

où la dernière flèche est donnée par le scalaire λ.
Matriciellement, on obtient l'identité suivante :

$$^t\mathrm{Mat}(u, \mathcal{B}) \cdot \mathrm{Mat}(q, \mathcal{B}) \cdot \mathrm{Mat}(u, \mathcal{B}) = \lambda \cdot \mathrm{Mat}(q, \mathcal{B})\,.$$

Du coup, $\det^2 u = \lambda^n$.

Lorsque $\mathbf{k}$ est algébriquement clos, ou lorsque $(\mathbf{k}^*)^2 = \mathbf{k}^*$, on a la suite exacte courte :

$$1 \to \mathrm{O}(q) \to \mathrm{GO}(q) \to \mathbf{k}^* \to 1.$$

En effet, si $\mu \in \mathbf{k}^*$, on considère $\lambda \in \mathbf{k}$ tel que $\lambda^2 = \mu$, et $u = \lambda \mathrm{I}$. Il vient $b(u(x), u(y)) = \lambda^2 \cdot b(x, y) = \mu \cdot b(x, y)$. En général, il n'est pas évident de trouver une section qui donne la racine carrée d'un scalaire.

On a la caractérisation suivante des similitudes.

Proposition 14.4.0.1. *Soit* E *un* $\mathbf{k}$*-espace vectoriel de dimension finie muni d'une forme quadratique* q *non dégénérée. Soit* $u \in \mathrm{GL(E)}$*. Alors,* u *est une similitude si et seulement si* u *préserve l'orthogonalité, soit*

$$\forall\, x, y \in \mathrm{E},\ x \perp y \iff u(x) \perp u(y)\,.$$

Démonstration. *Il est aisé de vérifier qu'une similitude préserve l'orthogonalité. Inversement, on considère une base orthogonale* $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ *de* E*. On considère* $\varepsilon_i = u(\mathrm{e}_i)$*,* $i = 1, \ldots, n$*, qui forment aussi une base orthogonale par hypothèse.*
Comme q *est non dégénérée,* $q(\mathrm{e}_i)$*,* $q(\varepsilon_i)$ *sont non nuls, donc il existe* $\lambda_i \in \mathbf{k}^*$ *tel que* $q(\varepsilon_i) = \lambda_i q(\mathrm{e}_i)$*. Il suffit de montrer que* λ_i *est indépendant de* i *pour conclure que* u *est une similitude. On se donne deux indices* $i \neq j$ *et on pose* $\lambda = -q(\mathrm{e}_i)/q(\mathrm{e}_j)$*. Il vient*

$$b(\mathrm{e}_i + \mathrm{e}_j, \mathrm{e}_i + \lambda \mathrm{e}_j) = q(\mathrm{e}_i) + \lambda q(\mathrm{e}_j) = 0$$

donc ces vecteurs sont orthogonaux. Du coup, $u(\mathrm{e}_i + \mathrm{e}_j) = \varepsilon_i + \varepsilon_j$ *et* $u(\mathrm{e}_i + \lambda \mathrm{e}_j) = \varepsilon_i + \lambda \varepsilon_j$ *sont aussi orthogonaux et on en déduit*

$$\lambda = -\frac{q(\varepsilon_i)}{q(\varepsilon_j)} = -\frac{\lambda_i}{\lambda_j} \cdot \frac{q(\mathrm{e}_i)}{q(\mathrm{e}_j)} = \lambda \cdot \frac{\lambda_i}{\lambda_j}\,.$$

Ceci montre bien que λ_i *est une fonction constante de* i*.* ∎

14.5 Générateurs du groupe orthogonal

Nous allons démontrer que les réflexions (resp. retournements) engendrent $\mathrm{O}(q)$ (resp $\mathrm{SO}(q)$.) Commençons par un cas simple.

Lemme 14.5.0.1. *Si q est anisotrope, toute isométrie u est produit d'au plus $\mathrm{rg}(u-\mathrm{Id})$-réflexions.*

Démonstration. *On fait une récurrence sur*

$$d = \mathrm{rg}(u-\mathrm{Id}) = n - \dim \mathrm{Ker}(u-\mathrm{Id}) \leq n$$

Si $d=0$, l'identité est bien produit de 0 symétries. Supposons $0<d\leq n$ et le théorème prouvé pour tout isométrie v telle que $\mathrm{rg}(v-\mathrm{Id})\leq d-1$.
Comme $d>0$, on peut choisir x n'appartenant pas à $\mathrm{Ker}(u-\mathrm{Id})$, c'est-à-dire non fixé par u. Posons $y=u(x)\neq x$. On a $q(x)=q(y)$ et $y-x$ anisotrope puisque non nul. On a $s_{x-y}(x)=y$ d'après 14.3.0.3 et donc $v(x)=x$ avec $v=s_{x-y}\circ u$. Mais si $z\in\mathrm{Ker}(u-\mathrm{Id})$, on a

$$b(z,x-y)=b(z,x)-b(z,u(x))=b(z,x)-b(u(z),u(x))=b(z,x)-b(z,x)=0$$

de sorte que $\mathrm{Ker}(u-\mathrm{Id})\subset\{x-y\}^\perp=\mathrm{Ker}(s_{x-y}-\mathrm{Id})$. Ainsi, $\mathbf{k}x\oplus\mathrm{Ker}(u-\mathrm{Id})\subset\mathrm{Ker}(v-\mathrm{Id})$ et donc la codimension de $\mathrm{rg}(v-\mathrm{Id})\leq d-1$. On conclut en appliquant l'hypothèse de récurrence à v. ∎

Proposition 14.5.0.2. *Si q est non dégénérée, toute isométrie est produit d'au plus $2n$ symétries hyperplanes.*

Démonstration. *On procède par récurrence sur n et soit $u\in\mathrm{O}(q)$. Soit x non isotrope de sorte $u=u(x)$ l'est également. L'un des vecteurs $x-y$ ou $x+y$ est donc non isotrope d'après la formule de polarisation.*
Si $x-y$ est non isotrope, $v=s_{x-y}\circ u$ fixe x. Comme x est non isotrope, l'hyperplan $\mathrm{H}=x^\perp$ est un supplémentaire de $\mathbf{k}x$ qui est stable par v. On applique alors l'hypothèse de récurrence à la restriction de v à H en remarquant que (H,q) est non dégénéré.
Si $x+y$ est non isotrope, on pose $v=s_y\circ s_{x+y}\circ u$ et on conclut comme plus haut. ∎

Remarque(s) 14.5.0.3. *On peut prouver que n symétries suffisent (Cartan-Dieudonné).*

Proposition 14.5.0.4. *Si $n\geq 3$, tout élément de $\mathrm{SO}(q)$ est produit d'au plus n renversements.*

Démonstration. *Tout élément de* $\mathrm{SO}(q)$ *est produit d'un nombre pair de réflexions de sorte qu'il s'agit de démontrer que le produit* $u = s_x \circ s_y$ *de deux réflexions est un composé de retournements.*
On peut supposer que x, y *sont non colinéaires de sorte qu'il engendrent un plan* P*. Comme* $q_{|\mathrm{P}}$ *est non nulle (*x *et* y *anisotropes par définition), son noyau* $\mathrm{P} \cap \mathrm{P}^\perp$ *est au plus de dimension* 1*. Notons que la restrictions de* v *à* $\mathrm{P}^\perp = x^\perp \cap y^\perp$ *est l'identité et que* P *est stable par* v*.*
Si $\mathrm{P} \cap \mathrm{P}^\perp = \{0\}$*, on a* $\mathrm{E} = \mathrm{P} \overset{\perp}{\oplus} \mathrm{P}^\perp$ *et* $\mathrm{P}^\perp$ *non dégénéré comme* E *de sorte qu'on peut prendre* $z \in \mathrm{P}^\perp$ *non isotrope. Alors,* $\mathrm{W} = \mathrm{Vect}(x, y, z)$ *est non dégénéré de sorte qu'on a une décomposition* $\mathrm{E} = \mathrm{W} \overset{\perp}{\oplus} \mathrm{W}^\perp$ *qui est stable par* v *et on conclut par récurrence.*
Si $\mathrm{P} \cap \mathrm{P}^\perp$ *est un droite* D*, soit* $z \notin \mathrm{D}^\perp$ *et* $\mathrm{W} = \mathrm{Vect}(x, y, z)$*. Comme* $\mathrm{D} \subset \mathrm{P}^\perp$*, on a* $\mathrm{P} \subset \mathrm{D}^\perp$ *de sorte que* $z \notin \mathrm{P}$ *et* W *est de dimension* 3*. Montrons que* W *est non dégénéré. Soit donc* $w \in \mathrm{W} \cap \mathrm{W}^\perp$ *non nul. Comme* $w \in z^\perp$*, on a* $w \in \mathrm{D} \subset \mathrm{P}$*. Mais* $\mathrm{W}^\perp \subset \mathrm{P}^\perp$ *donc* $w \in \mathrm{P} \cap \mathrm{P}^\perp = \mathrm{D}$ *et* $\mathrm{D} = \mathbf{k}w$*. Mais* $b(w, z) = 0$*, donc* $z \in \mathrm{D}^\perp$*, une contradiction. On a alors* $\mathrm{E} = \mathrm{W} \overset{\perp}{\oplus} \mathrm{W}^\perp$*. Comme* u *agit sur l'identité sur* $\mathrm{P}^\perp \subset \mathrm{W}^\perp$*, il agit aussi sur l'identité sur* $\mathrm{W}^\perp$ *et donc laisse stable* W*. Mais les restrictions de* $-s_x$ *et* $-s_y$ *à* W *sont des retournements (14.3.0.3) ainsi que leur prolongements* r_x, r_y *par l'identité sur* $\mathrm{W}^\perp$*. Et on a* $y = r_x \circ r_y$*.* ∎

Exercice(s) 14.5.0.5. *Montrer que le conjugué d'une réflexion par une isométrie est une isométrie dont on précisera les éléments caractéristiques. En déduire le centre des groupes orthogonaux et spéciaux orthogonaux.*

Chapitre 15

Géométrie euclidienne

15.1 Généralités

La plupart de cette section est constituée de rappels (orthogonalité, espaces orthogonaux et supplémentaires, Gram-Schmidt et existence de l'adjoint d'un endomorphisme d'un espace euclidien). On a remis des preuves pour mémoire mais elles ne sont pas forcément traitées à l'oral : à réviser donc !

Euclide par Raphael

15.1.1 Norme euclidienne

Définition 15.1.1.1. *Un espace euclidien est un espace vectoriel de dimension finie n sur $\mathbf{R}$ muni d'une forme quadratique q définie positive,* i.e.*de signature* $(n, 0)$.

Dans cette situation, on appelle la forme polaire *un produit scalaire* et on le note communément $\langle \cdot, \cdot \rangle$. On remarque, et c'est aussi trivial qu'important, que tout sous-espace d'une espace euclidien est euclidien. Dans ce chapitre, (E, q) désigne un espace euclidien. Rappelons que dans un espace euclidien, il existe toujours une base orthonormée. Dans une telle base, la matrice de q est l'identité.

Exemple(s) 15.1.1.2. — *Si* (X, μ) *est un espace mesuré (de mesure* > 0*), alors on définit* $q : \mathrm{L}^2(\mathrm{X}, \mu; \mathbf{R}) \to \mathbf{R}$ *par*

$$q(f) = \int f^2 d\mu.$$

On a bien $q(f) > 0$ *pour tout* $f \neq 0$ *dans* L^2*. Toute restriction à un sous-espace de dimension finie confère une structure d'espace euclidien.*

— *Si* $\mathrm{M} \in \mathrm{M}_{p,q}(\mathbf{R})$*, alors* $q(\mathrm{M}) = \mathrm{tr}({}^t\mathrm{MM})$ *détermine une forme quadratique définie positive. Pour cela, on note* $\mathrm{M} = (a_{ij})$ *et on calcule les termes diagonaux de* ${}^t\mathrm{MM} = (b_{ij})$ *:*

$$b_{jj} = \sum_k a_{kj} a_{kj} = \sum_k a_{kj}^2$$

et

$$\mathrm{tr}({}^t\mathrm{MM}) = \sum_{i,j} a_{ij}^2$$

donc $q(\mathrm{M}) \geq 0$ *et* $q(\mathrm{M}) = 0$ *implique* $\mathrm{M} = 0$*.*

Théorème 15.1.1.3. *L'application* $x \mapsto \|x\| = \sqrt{q(x)}$ *est une norme dite euclidienne.*

Démonstration. *On pose, pour* $x \in \mathrm{E}$*,* $\|x\| = \sqrt{q(x)}$*. Rappelons l'inégalité de Cauchy-Schwarz : pour tout* $x, y \in \mathrm{E}$*,*

$$\langle x, y \rangle \leq q(x) q(y)$$

avec égalité si et seulement si x, y *sont positivement liés (qui est une inégalité en dimension* ≤ 2*, donc de géométrie plane usuelle.). Comme* q *est définie, pour montrer que* $\|\cdot\|$ *est une norme, il suffit de vérifier l'inégalité triangulaire.*

$$\begin{aligned}(\|x\| + \|y\|) - \|x + y\| &= \frac{q(x) + q(y) + 2\sqrt{q(x)q(y)} - q(x+y)}{\|x+y\| + \|x\| + \|y\|} \\ &\geq \frac{2\sqrt{q(x)q(y)} - 2\langle x, y \rangle}{\|x+y\| + \|x\| + \|y\|} \geq 0\,.\end{aligned}$$

■

Rappel : démonstration l'inégalité de Cauchy-Schwarz. — On peut supposer x et y non nuls. Pour $t \in \mathbf{R}$, on a $q(xt + y) \geq 0$. Or $q(tx + y) = q(x)t^2 + 2t\langle x, y \rangle + q(y)$. Comme ce polynôme quadratique ne prend que des valeurs positives, il ne peut s'annuler au plus qu'une fois, donc son discriminant doit être négatif :

$$\langle x, y \rangle^2 - q(x) q(y) \leq 0$$

d'où le premier point. En cas d'égalité, le trinôme admet une racine réelle double t comme q est définie, on a $x + ty = 0$. Mais comme $\langle x, y \rangle$ est alors ≥ 0, ceci force t à être négatif. ■

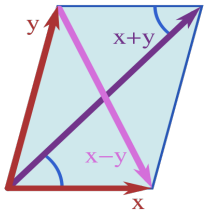
Démonstration. *On applique le théorème de classification des formes quadratiques sur* $\mathbf{R}$. ■

La norme euclidienne est caractérisée par l'identité de la médiane :

Exercice(s) 15.1.1.4. *Montrer que* E *normé de dimension finie est euclidien si et seulement si pour tout* $x, y \in \mathrm{E}$*, on a*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Expliquer pourquoi cette identité s'appelle identité de la médiane.



L'identification entre E et son dual dans le cas euclidien définie par le produit scalaire permet de définir le gradient $\nabla_x(f) \in \mathbf{R}^n$ d'une fonction $f : \mathbf{R}^n \to \mathrm{R}$ différentiable en un point x par la relation

$$\forall v \in \mathbf{R}^n,\ (v, \nabla_x(f)) = df(x).v.$$

15.1.2 Orthogonalisation, projection

On aurait pu précédemment utiliser l'algorithme suivant

Proposition 15.1.2.1 (Algorithme de Gram-Schmidt.)**.** *Soit* $\mathrm{e}_1, \cdots \mathrm{e}_d$ *une famille libre dans* E *euclidien. Alors, il existe une unique famille orthonormée* $\varepsilon_1, \cdots, \varepsilon_d$ *telle que*

- $\mathrm{Vect}\,\mathrm{e}_1, \cdots \mathrm{e}_i = \mathrm{Vect}\,\varepsilon_1, \cdots, \varepsilon_i$ *pour* $i = 1, \cdots, d$.
- $(\mathrm{e}_i, \varepsilon_i) > 0$ *pour* $i = 1, \cdots, d$.

Démonstration. *On fait une récurrence sur* i*. On pose* $\varepsilon_1 = \mathrm{e}_1/\|\mathrm{e}_1\|$*, supposant* $\varepsilon_1, \cdots, \varepsilon_i$ *construits, on doit chercher* ε_{i+1} *sous le forme* $\varepsilon_{i+1} = \lambda_{i+1}\mathrm{e}_{i+1} + \sum_{k \le i} \lambda_k \varepsilon_k$*. Ceci force l'égalité* $0 = \lambda_{i+1}(\mathrm{e}_{i+1}, \varepsilon_k) + \lambda_k$ *pour* $k \le i$ *et impose la valeur de* λ_k *en fonction de* λ_{i+1}*. D'autre part, on a* $1 = \lambda_{i+1}(\mathrm{e}_{i+1}, \varepsilon_{i+1}) = \lambda_{i+1}^2(\mathrm{e}_{i+1}, \mathrm{e}_{i+1}) + 0$ *ce qui force la valeur de* λ_{i+1} *(qui doit être* > 0*). Inversement, ces valeurs conviennent.* ■

Exercice(s) 15.1.2.2. *Programmer l'algorithme précédent sur SAGE.*

L'orthogonalisation de Gram-Schmidt implique aussi que, pour tout sous-espace F de E, il existe une base orthonormée dont les premiers $\dim \mathrm{F}$ vecteurs forment une base de F.

Corollaire 15.1.2.3 (Décomposition d'Iwasawa)**.** *Toute matrice réelle* $\mathrm{M} \in \mathrm{GL}_n(\mathbf{R})$ *se décompose de manière unique en un produit* $\mathrm{M} = \mathrm{QR}$ *avec* Q *orthogonale et* R *triangulaire supérieure à coefficients diagonaux* > 0*.*

Démonstration. *Si* ε_i *désigne les colonnes de* Q*, dire que* $\mathrm{M} = \mathrm{QR}$ *comme dans la décompostion d'Iwasawa, c'est dire que epsilon*$_i$ *est orthonormée,* $\mathrm{R}_{i,j} = \langle \mathrm{e}_i, \varepsilon_j \rangle$ *avec* e_j *la* j*-ième colonne de* M*, autrement dit que* e_i *est l'orthonormalisée de Schmidt des colonnes de* M*. D'où existence et unicité.* ■

Corollaire 15.1.2.4 (Inégalité de Hadamard)**.** *Le volume euclidien d'un parallépipède est inférieur au produit des lonqueurs des côtés avec égalité si et seulement si il est rectangle.*

Démonstration. *Soit donc* $\mathrm{e}_1, \cdots, \mathrm{e}_n$ *une famille de* n *vecteurs libres de* $\mathbf{R}^n$ *et* M *la matrice* $[\mathrm{e}_1, \cdots, \mathrm{e}_n]$*. On veut montrer* $|\det(\mathrm{M})| \leq \prod \|\mathrm{e}_i\|$*. En gardant les notations précédentes, on a*

$$|\det(\mathrm{M})| = |\det(\mathrm{R})| = |\prod (\mathrm{e}_i, \varepsilon_i)| \overset{Cauchy-Schwarz}{\leq} \prod \|\mathrm{e}_i\| \|\varepsilon_i\| = \prod \|\mathrm{e}_i\|.$$

En cas d'égalité, on a égalité dans Cauchy-Schwartz de sorte que chaque e_i *est (positivement) lié à* ε_i*.*■

Exercice(s) 15.1.2.5. *Généraliser l'existence de la décomposition* QR *aux matrices rectangles quelconques. A-t-on unicité ?*

Proposition 15.1.2.6 (Projection orthogonale)**.** *Soit* F *un sous-espace vectoriel de* E *euclidien et* $x \in \mathrm{E}$*.*

1. *Il existe un unique* $p_{\mathrm{F}}(x) \in \mathrm{F}$ *tel que* $\|x - p_{\mathrm{F}}(x)\| = \inf_{y \in \mathrm{F}} \|x - y\|$*.*
2. *L'application* $x \mapsto p_{\mathrm{F}}(x)$ *est la projection linéaire orthogonale sur* F *parallèlement à* $\mathbf{F}^{\perp}$*.*

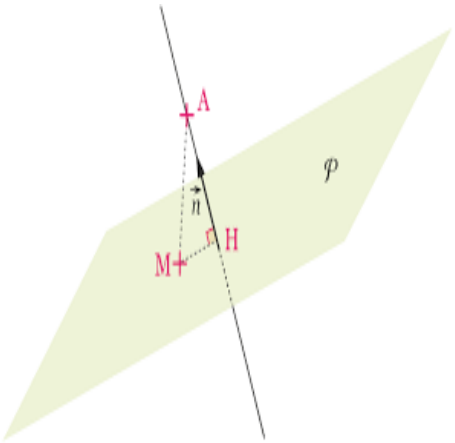
Remarque(s) 15.1.2.7. *Cette proposition se généralise (en perdant la linéarité) au cas où* F *est simplement un convexe fermé, et ce en supposant seulement* E *avec une forme définie et complet pour la norme associée (Hilbert réel).*

Démonstration. *Exercice.*

Exercice(s) 15.1.2.8. *Soit* H *un hyperplan de* $\mathbf{R}^n$ *d'équation* $\sum a_i x_i = 0$*. Montrer l'égalité*

$$d(x, \mathrm{H}) = \frac{|\sum a_i x_i|}{\sqrt{\sum a_i^2}}.$$

Pour toute famille finie x_i de vecteurs de E euclidien, on note $\mathrm{Gram}(x_i) = \det(b(x_i, x_j)_{i,j})$ (déterminant de Gram). Comme b est définie positive, Gram ≥ 0 et n'est nul que si les x_i sont liés.



projection

Proposition 15.1.2.9 (Distance d'un point à un sous-espace)**.** *Soit* $\mathrm{e}_1 \cdots, \mathrm{e}_d$ *une bas de* F *un sous-espace vectoriel de* E *euclidien* $x \in \mathrm{E}$*. On a*

$$d(x, \mathrm{H})^2 = \|x - p_{\mathrm{F}}(x)\|^2 = \frac{\mathrm{Gram}(x, \mathrm{e}_1, \cdots, \mathrm{e}_d)}{\mathrm{Gram}(\mathrm{e}_1, \cdots, \mathrm{e}_d)}.$$

Démonstration. *Comme* $p_{\mathrm{F}}(x) \in \mathrm{F}$*, la famille* $p_{\mathrm{F}}(x), \mathrm{e}_1, \cdots, \mathrm{e}_d$ *est liée et donc son déterminant de Gram est nul. La multilinéarité du déterminant assure alors*

$$\mathrm{Gram}(x, \mathrm{e}_1, \cdots, \mathrm{e}_d) = \mathrm{Gram}(x - p_{\mathrm{F}}(x), \mathrm{e}_1, \cdots, \mathrm{e}_d).$$

Comme $x - p_{\mathrm{F}}(x) \in \mathrm{F}^{\perp}$*, on a* $(x - p_{\mathrm{F}}(x), \mathrm{e}_i) = 0$ *pour tout* i *et donc*

$$\mathrm{Gram}(x - p_{\mathrm{F}}(x), \mathrm{e}_1, \cdots, \mathrm{e}_d) = \|x - p_{\mathrm{F}}(x)\|^2 \,\mathrm{Gram}(\mathrm{e}_1, \cdots, \mathrm{e}_d)$$

ce qui montre la proposition. ∎

15.1.3 Isométries en petite dimension, rappels

Même si tout ceci se déduit du cas général (§15.1.5), il est bon d'avoir une vision claire des isométries (endomorphismes qui préservent la norme euclidienne ou le produits scalaire, c'estla même chose par polarisation) en dimension 2 et 3. Les preuves sont laissées au lecteur à titre d'exercice. **Cas de la dimension 2.** Considérons espace euclidien de dimension 2 orienté. Une isométrie $u \in \mathrm{O}_2(\mathbf{R})$ du plan euclidien, identifié à $\mathbf{C}$ muni de la valeur absolue par le choix d'une quelconque base orthonormée directe $\mathcal{B}$, est

— Une rotation d'angle $\theta \in \mathrm{U}(1) = \mathbf{R}/2\pi\mathbf{Z}$ si $u \in \mathrm{SO}_2(\mathbf{R})$ (isométrie directe) de matrice [1]

$$\mathrm{R}_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

1. Quelle que soit la base orthonormée directe, reflet de la commutativité de $\mathrm{SO}_2(\mathbf{R}) \simeq \mathrm{U}(1)$

qui s'écrit $z \mapsto \exp(\mathbf{i}\theta)z$ en complexe ;

— Une symétrie par rapport à une droite $D_{\theta/2}$ avec $\theta \in \mathbf{R}/\pi\mathbf{Z}$ sinon (isométrie indirecte) qui a pour matrice lorsque le premier vecteur de $\mathcal{B}$ dirige l'axe de symétrie

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

et s'écrit dans l'identification au plan complexe correspondante $z \mapsto \exp(\mathbf{i}\theta)\overline{z}$ en complexe.

Si on change l'orientation, θ est changé en son opposé et l'angle d'une rotation n'est défini qu'au signe près, caractérisé par $\mathrm{tr}(R_\theta) = 2\cos(\theta)$.

Le composé de deux symétries orthogonales est alors une rotation d'angle le double entre les axes de symétries. On déduit qu'elles engendrent $O_2(\mathbf{R})$.

Cas de la dimension 3. Regardons le cas d'ue isométrie u d'un espace euclidien orienté de dimension 3. Du fait que le polynôme caractéristique est réel degré 3, on déduit que ± 1 est valeur propre. Quitte à changer u en $-u$, on peut supposer que 1 est valeur propre. En considérant le plan orthogonal à un vecteur unitaire invariant, on déduit du cas de la dimension 2 que u

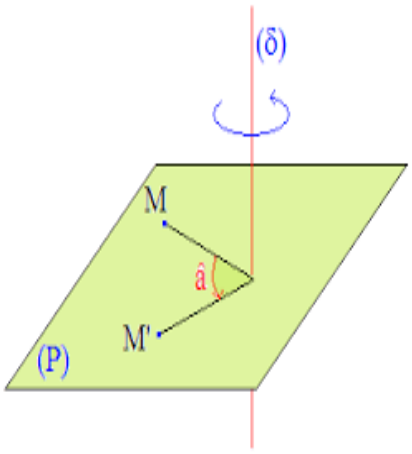
— une isométrie directe d'axe Δ (ponctuellement invariant) et d'angle $\theta \mod 2\pi$ (dans le plan orthogonal à Δ) défini au signe près si l'axe n'est pas orienté ou $2\pi\mathbf{Z}$ si Δ est orientés de matrice dans une base orthonormée directe adaptée

$$\begin{pmatrix} 1 & 0 \\ 0 & R_\theta \end{pmatrix} ;$$

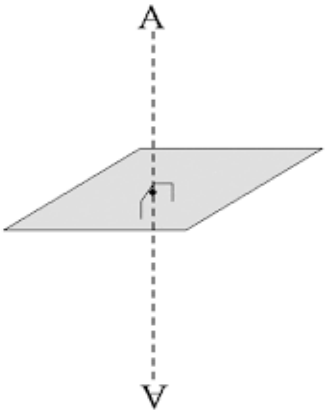
L'angle est caractérisé par $\mathrm{tr}(u) = 1 + 2\cos(\theta)$.

— l'opposé d'une rotation.

L'opposé des rotations d'angle π sont les symétries orthogonales par rapport à des plans (retournements), qui, comme dans le cas plan, engendrent $O_2(\mathbf{R})$ [le composé de deux retournements est une rotation d'axe l'intersection des plans fixes et d'angle le double de l'angle entre ces deux plans -penser à deux feuilles d'un livre entre-ouvert-].

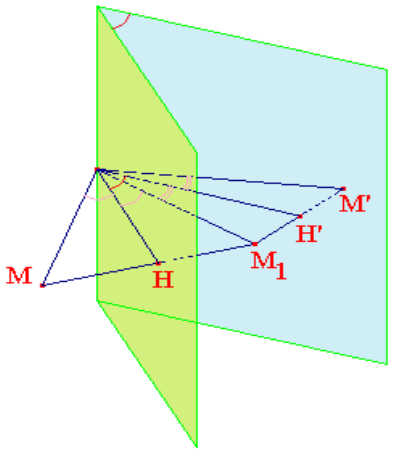


rotation



symétrie

Exercice(s) 15.1.3.1. *Caractériser géométriquement le composé de deux réflexions. En déduire que* $O_3()$ *est engendré par les réflexions. En déduire que* $SO_3()$ *est engendré par les retournements (rotations d'angle* π*).*



composé de réflexions

15.1.4 Endomorphismes normaux réels

Rappelons (11.3.4.1) que la matrice dans une base orthonormée de l'adjoint d'un endomorphisme est sa transposée[2]. Dès lors, il est normal (s'il commute avec son adjoint) si et seulement si cette matrice commute avec sa transposée et sa restriction à un sous espace stable d'un tel endomorphisme normal est encore normal.. Les principaux exemples proviennent, dans une base orthonormée, des matrices réelles M symétriques ${}^t\mathrm{M} = \mathrm{M}$, antisymétriques ${}^t\mathrm{M} = -\mathrm{M}$ et orthogonales ${}^t\mathrm{M} = \mathrm{M}^{-1}$. Nous allons voir, qu'une fois de plus, tout se ramène à la géométrie plane !

L'outil pour se ramener en dimension ≤ 2 est le comportement des orthogonaux vis à vis des endomorphismes normaux :

Proposition 15.1.4.1. *Soit* $\mathrm{M} = \begin{pmatrix} \mathrm{A} & \mathrm{C} \\ 0 & \mathrm{B} \end{pmatrix}$ *une matrice carrée réelle par blocs qui commute avec sa transposée. Alors,* $\mathrm{C} = 0$*. Autrement dit, l'orthogonal d'un espace stable par un endomorphisme normal est stable.*

Démonstration. *Le second point est une conséquence immédiate du premier en écrivant la matrice (symétrique) de u dans une base orthonormée union d'une base orthonormée de l'espace stable en question et de son orthogonal. Pour le premier point, le bloc* $(1,1)$ *de* $\mathrm{M}{}^t\mathrm{M} - {}^t\mathrm{M}\mathrm{M}$ *s'écrit* $\mathrm{C}{}^t\mathrm{C} + \mathrm{A}{}^t\mathrm{A} - {}^t\mathrm{A}\mathrm{A}$ *et est nul. En prenant sa trace, on a* $\mathrm{tr}(\mathrm{C}{}^t\mathrm{C}) = \sum c_{i,j}^2 = 0$ *et donc* $\mathrm{C} = 0$ *(c'est ici que sert le caractère réel des coefficients).* ∎

2. Attention, c'est faux si la base n'est pas supposée orthonormée

Exercice(s) 15.1.4.2. *Soit u un endomorphisme normal d'un espace euclidien. Montrer les points suivants*

1. *On a l'égalité* $\operatorname{Ker} u = \operatorname{Ker} u^*$.
2. *Plus généralement, u et u^* ont mêmes valeurs propres et espaces propres. Ceux-ci sont orthogonaux.*
3. *Les endomorphismes normaux diagonalisables sont diagonalisables dans une base orthonormée.*
4. *Plus généralement encore, u et u^* ont même espaces stables.*

Proposition 15.1.4.3. *Soit u un endomorphisme du plan euclidien et $\mathrm{M} \in \mathrm{M}_2(\mathbf{R})$.*

1. M *commute avec sa transposée si et seulement si elle est soit symétrique soit de la forme* $\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \rho \mathrm{R}_\theta$ *avec* $a + ib = \rho \exp(i\theta), r > 0$.
2. M *commute avec sa transposée et est diagonalisable si et seulement si* M *est symétrique. Dans ce cas, elle est orthogonalement diagonalisable.*
3. *u est soit auto-adjoint soit une similitude directe. Les similitudes directes du plan euclidien qui sont auto-adointes sont les homothéthies.*
4. *u est normal et diagonalisable si et seulement si il est auto-adjoint ; il est alors diagonalisable dans une base orthonormée.*

Démonstration. *Montrons le premier point avec* $\mathrm{M} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. *On traduit sur ses coefficients le fait d'être normal : on obtient le système*

$$\begin{cases} (c-b)(c+b) = 0 \\ (a-d)(b-c) = 0 \end{cases}$$

Soit M *est symétrique et $b = c$, soit* M *ne l'est pas de sorte que $b \neq c$ et donc $b = -c$, $a = d$. On a alors*

$$\mathrm{M} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \rho \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

en posant $a + ib = \rho \exp(i\theta), r > 0$. Pour le second point, vérifions d'abord qu'une matrice symétrique réelle $\mathrm{M} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ *est diagonalisable.*
D'après le lemme clef 15.1.4.1 (ou l'exercice 15.1.4.2 comme on veut), il s'agit simplement de montrer que $\chi_{\mathrm{M}}(\mathrm{T})$ est scindé. Or son discriminant est

$$(a+d)^2 - 4(ad - b^2) = (a-d)^2 + b^2 \geq 0.$$

Il faut pour finir montrer que la similitude directe $r\mathrm{R}_\theta$ n'est diagonalisable qu si elle est symétrique. Or, ses valeurs propres sont $r\exp(\pm i\theta)$ donc ne sont réelles que si $\theta \mod \pi = 0$ et donc $\mathrm{M} = \pm r\,\mathrm{Id}$ qui est bien symétrique ! Les deux derniers points en découlent immédiatement en considérant la matrice de u dans une base orthonormée. ∎

Remarque(s) 15.1.4.4. *L'angle d'une matrice de similitude n'est défini qu'au signe près (ou à π près, comme on veut), exactement comme pour les rotations planes. En effet, si l'angle est θ dans une base orthonormée (e,e_2), il devient θ dans la base (d'orientation opposée) (e_2,e_1).*

Une récurrence immédiate jointe aux propositions 15.1.4.1 et 15.1.4.3 donne

Théorème 15.1.4.5 (Réduction des normaux)**.** *Les endomorphismes normaux u sont les endomorphismes tels qu'il existe une base orthonormée, des réels $\alpha_1,\dots\alpha_p$, $\rho_1>0,\dots,\rho_q>0$ et $\theta_1,\dots,\theta_q\in\mathbf{R}-\pi\mathbf{Z}$ tels que*

$$\mathrm{Mat}(u)=\begin{pmatrix}\alpha_1 & & & & & & \\ & \ddots & & & 0 & & \\ & & \alpha_p & & & & \\ & & & \rho_1\cdot \mathrm{R}_{\theta_1} & & & \\ & & & & \ddots & & \\ & & & & & \rho_q\cdot \mathrm{R}_{\theta_q}\end{pmatrix}$$

Les $\alpha_i,\rho_j,(\theta_j \mod \pi)$ sont uniques à l'ordre près.

Démonstration. *On raisonne par récurrence sur la dimension. Les cas de la dimension 1 et 2 sont faits. On suppose que le théorème est vrai jusqu'à la dimension n et on suppose que E est de dimension $n+1$. D'après la remarque 7.2.0.2, u fixe une droite ou un plan F. Dans cet espace, on sait (15.1.4.3) trouver une base orthonormée qui donnera la forme voulue pour $u|_{\mathrm{F}}$. Ensuite, on sait (15.1.4.1)que u préserve $\mathrm{F}^\perp$ avec de plus $u|_{\mathrm{F}^\perp}$ normal. On peut appliquer l'hypothèse de récurrence.* ∎

D'un point de vue matriciel, ceci se réécrit en disant si $\mathrm{M}\in\mathrm{M}_n(\mathbf{R})$ commute avec sa transposée, il existe O telle que

$${}^t\mathrm{OO}=\mathrm{Id}\ \text{ et }\ \mathrm{O}^{-1}\mathrm{MO}={}^t\mathrm{OMO}=\begin{pmatrix}\alpha_1 & & & & & & \\ & \ddots & & & 0 & & \\ & & \alpha_p & & & & \\ & & & \rho_1\cdot \mathrm{R}_{\theta_1} & & & \\ & 0 & & & \ddots & & \\ & & & & & \rho_q\cdot \mathrm{R}_{\theta_q}\end{pmatrix}.$$

15.1.5 Application aux endomorphismes orthogonaux

On suppose que E est un espace euclidien muni d'une base orthonormée $\mathcal{B}$. Rappelons

Proposition 15.1.5.1. *Soit u un endomorphisme de* E. *Les propriétés suivantes sont équivalentes.*

- $u \in \mathrm{O(E)}$;
- *pour tout* $x \in \mathrm{E}$, *on a* $\|u(x)\| = \|x\|$;
- $u \circ u^* = \mathrm{Id}$;
- $u^* \circ u = \mathrm{Id}$;
- $\mathrm{Mat}(u, \mathcal{B}) \cdot {}^t\mathrm{Mat}(u, \mathcal{B}) = \mathrm{I}$;
- ${}^t\mathrm{Mat}(u, \mathcal{B}) \cdot \mathrm{Mat}(u, \mathcal{B}) = \mathrm{I}$;
- *les colonnes de* $\mathrm{Mat}(u, \mathcal{B})$ *forment une base orthonormée ;*
- *les lignes de* $\mathrm{Mat}(u, \mathcal{B})$ *forment une base orthonormée ;*
- u *transforme une base orthonormée en une base orthonormée.*

La démonstration est laissée en exercice.

Corollaire 15.1.5.2. *Si* $u \in \mathrm{O(E)}$, *alors* $\det u = \pm 1$ *et donc* $\mathrm{SO(E)} = \det^{-1}\{1\} \cap \mathrm{O(E)}$ *est distingué d'indice* 2.

Lorsqu'on munit $\mathrm{E} = \mathbf{R}^n$ de sa base canonique et de son produit scalaire usuel (pour laquelle elle est orthonormée), les endomorphismes orthogonaux s'identifient donc aux matrices orthogonales O usuelles (qui vérifient ${}^t\mathrm{OO} = \mathrm{Id}$) qui sont les endomorphismes préservant la norme euclidienne (isométries). On note alors $\mathrm{O(E)} = \mathrm{O}_n(\mathbf{R})$ le groupe orthogonal, sous-groupe de $\mathrm{GL}_n(\mathbf{R})$ des orthogonales et $\mathrm{SO}_n(\mathbf{R})$ son sous-groupe d'indice 2 des matrices de déterminant 1.

Théorème 15.1.5.3 (Réduction des isométries)**.** *Les isométries d'un espace euclidien sont les endomorphismes dont une matrice dans une base orthonormée convenable est diagonale par blocs de la forme*

$$\begin{pmatrix} \mathrm{Id}_p & & & & \\ & -\mathrm{Id}_m & & & \\ & & \mathrm{R}_{\theta_1} & & \\ & & & \ddots & \\ & & & & \mathrm{R}_{\theta_q} \end{pmatrix}.$$

avec avec $p, m \in \mathbf{N}$ *et* $\theta_1, \ldots, \theta_q \in \mathbf{R} - \pi\mathbf{Z}$

Démonstration. *Vace les notations du théorème 15.1.4.5 de réduction des normaux, il suffit d'observer que la conditions de préserver la norme équivaut à* $\alpha_i = \pm 1$ *et* $\rho_i = 1$. ∎

Remarque(s) 15.1.5.4. *Si* E *est de dimension impaire, les matrices orthogonales comme plus haut vérifient* m *pair de sorte que* $-\mathrm{Id}_m = \mathrm{diag}_{m/2}(\mathrm{R}_\pi)$ *et* n *impair de sorte que* $\mathrm{Id}_p = \mathrm{diag}(1, \mathrm{diag}_{(p-1)/2}(\mathrm{R}_0))$. *Ainsi, la matrice d'une isométrie positive en dimension impaire est de la forme* $\mathrm{diag}(1, \mathrm{diag}\,\mathrm{R}_{\tilde{\theta}_i})$ *avec* $\tilde{\theta}_i$ *réels, généralisant la forme usuelle dans l'espace euclidien à trois dimensions. De même, en dimension paire elle est de la forme* $\mathrm{diag}\,\mathrm{R}_{\tilde{\theta}_i}$, *donc sans forcément d'« axe » invariant comme en dimension impaire.*

15.1.6 Application aux auto-adjoints

Théorème 15.1.6.1 (Réduction des autoadjoints)**.** *Les endomorphismes auto-adjoints d'une espace euclidien sont les endomorphismes dont une matrice dans une base orthonormée convenable est diagonale. Matriciellement, le matrices symétriques réelles sont les matrices orthogonalement diagonalisables.*

Démonstration. *Comme* u *est normal, on peut appliquer le théorème 15.1.4.5 et se ramener à montrer qu'une matrice symétrique réelle de la fome*

$$\begin{pmatrix} \alpha_1 & & & & & & \\ & \ddots & & & 0 & & \\ & & \alpha_p & & & & \\ & & & \rho_1 \cdot \mathrm{R}_{\theta_1} & & & \\ & 0 & & & \ddots & & \\ & & & & & \rho_q \cdot \mathrm{R}_{\theta_q} & \end{pmatrix}$$

avec $\alpha_1, \dots \alpha_p$, $\rho_1 > 0, \dots, \rho_q > 0$ *et* $\theta_1, \dots, \theta_q \in \mathbf{R} - \pi\mathbf{Z}$ *est diagonale. Mais dans ce cas, on a* $m = 0$ *car les matrices de similitude* $\rho_i \mathrm{R}_{\theta_i}$ *ne sont pas symétriques puisque* $\theta_i \in \mathbf{R} - \pi\mathbf{Z}$. ∎

Remarque(s) 15.1.6.2. *Attention, les matrices symétriques complexes ne sont en général pas diagonalisables. Par exemple,* $\begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$ *est symétrique et nilpotente !*

Corollaire 15.1.6.3 (Réduction simultanée)**.** *Soient* $\mathrm{S}, \mathrm{S}' \in \mathrm{M}_n(\mathbf{R})$ *deux matrices symétriques réelles avec* S *définie positive. Alors, il existe une matrice inversible* $\Pi \in\in \mathrm{GL}_n(\mathbf{R})$ *et une matrice diagonale* $\Delta \in \mathrm{Diag}_n(\mathbf{R})$ *telles que*

$$^t\Pi \mathrm{S} \Pi = \mathrm{Id} \quad \textit{et} \quad ^t\Pi \mathrm{S}' \Pi = \Delta.$$

Démonstration. *Choisissons une base orthonormée de* $\mathbf{R}^n$ *pour la forme euclidienne* q_S. *Si* P *est la matrice de changement de ses coordonnées, on a donc* ${}^t\mathrm{PSP} = \mathrm{Id}$. *D'après le théorème de réduction 15.1.6.1 appliquée à la matrice symétrique* ${}^t\mathrm{PS'P}$, *donc définissant un endomorphisme normal pour le produits scalaire canonique de* $\mathbf{R}^n$, *il existe* $\mathrm{O} \in \mathrm{O}_n(\mathbf{R})$ *telle que*

$$ {}^t\mathrm{O}{}^t\mathrm{PSPO} = \Delta \text{ et donc } {}^t\mathrm{OO} = \mathrm{Id}\,. $$

Il suffit de poser $\Pi = \mathrm{PO}$. ∎

Remarque(s) 15.1.6.4. *Attention,* Δ *n'a rien à voir avec les valeurs propres de* S' *!.*

15.1.7 Application aux anti auto-adjoints

Théorème 15.1.7.1 (Réduction des anti-autoadjoints)**.** *Les endomorphismes auto-adjoints d'une espace euclidien sont les endomorphismes dont une matrice dans une base orthonormée convenable est diagonale par blocs où le blocs non nuls sont de la forme* $r \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $r \in \mathbf{R}^*$.

Démonstration. *On procède comme dans le cas auto-adjoint en observant que la matrice*

$$ \begin{pmatrix} \alpha_1 & & & & & \\ & \ddots & & & 0 & \\ & & \alpha_p & & & \\ & & & \rho_1 \cdot \mathrm{R}_{\theta_1} & & \\ & 0 & & & \ddots & \\ & & & & & \rho_q \cdot \mathrm{R}_{\theta_q} \end{pmatrix} $$

est anti-symétrique si et seulement si $\alpha_i = 0$ *et* $\theta_j = \pm\pi/2$. ∎

Corollaire 15.1.7.2. *L'application* $\exp : \mathcal{A}_n(\mathbf{R}) \to \mathrm{SO(E)}$ *est surjective, où* $\mathcal{A}_n(\mathbf{R})$ *désigne l'ensemble des matrices antisymétriques.*

Démonstration. *Si* A *est antisymétrique, alors* ${}^t\mathrm{A} = -\mathrm{A}$. *Pour toute matrice* A, $\exp(-\mathrm{A}) = (\exp \mathrm{A})^{-1}$ *et* $\exp({}^t\mathrm{A}) = {}^t(\exp \mathrm{A})$. *Du coup, si* A *est antisymétrique, alors* ${}^t\mathrm{A} = -\mathrm{A}$ *et* $(\exp \mathrm{A})^{-1} = {}^t(\exp \mathrm{A})$, *donc* $\exp(\mathcal{A}_n) \subset \mathrm{O(E)}$. *Or* $\det \exp \mathrm{A} = \exp \operatorname{tr} \mathrm{A}$ *donc, pour* $\mathrm{A} \in \mathcal{A}_n$, *on a* $\det \exp \mathrm{A} = 1$ *:* $\exp \mathrm{A} \in \mathrm{SO(E)}$.

Soit

$$\mathrm{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

On a $\mathrm{J}^2 = -\mathrm{I}$*, et, pour* $\theta \in \mathbf{R}$*, on obtient*

$$\exp(\theta \mathrm{J}) = \cos\theta \cdot \mathrm{I} + \sin\theta \cdot \mathrm{J} = \mathrm{R}_\theta\,.$$

Par conséquent, si

$$\mathrm{A} = \begin{pmatrix} 0 & & & & & & \\ & \ddots & & & 0 & & \\ & & 0 & & & & \\ & & & \theta_1 \cdot \mathrm{J} & & & \\ & 0 & & & \ddots & & \\ & & & & & \theta_q \cdot \mathrm{J} & \end{pmatrix}$$

$$\exp \mathrm{A} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & 0 & & \\ & & 1 & & & & \\ & & & \mathrm{R}_{\theta_1} & & & \\ & 0 & & & \ddots & & \\ & & & & & \mathrm{R}_{\theta_q} & \end{pmatrix}$$

Ceci permet de montrer la surjectivité. ∎

15.1.8 Appendice : étude de $\mathcal{S}_n^{++}$

On note $\mathcal{S}_n$ (resp. $\mathcal{S}_n^{++}$) l'ensemble des matrices réelles symétriques (resp. définies positives). Ce sont des convexes.

Corollaire 15.1.8.1. *L'application* $\exp : \mathcal{S}_n \to \mathcal{S}_n^{++}$ *est un homéomorphisme.*

Démonstration. *L'application* exp *est continue, et si* $\mathrm{M} \in \mathcal{S}_n(\mathbf{R})$*, alors il existe* $\mathrm{O} \in \mathrm{O}(n)$ *telle que* $\mathrm{O}^{-1}\mathrm{MO}$ *soit diagonale. On en déduit que* $\exp \mathrm{M} = \mathrm{O}(\exp \mathrm{O}^{-1}\mathrm{MO})\mathrm{O}^{-1} \in \mathcal{S}_n^{++}(\mathbf{R})$*. Réciproquement, si* $\mathrm{M} \in \mathcal{S}_n^{++}(\mathbf{R})$*, il existe* $\mathrm{O} \in \mathrm{O}(n)$ *telle que* $\mathrm{O}^{-1}\mathrm{MO}$ *soit diagonale avec des valeurs propres strictement positives. On peut alors considérer la matrice* N *diagonale formée des logarithmes des valeurs propres de* M*. On a* $\mathrm{ONO}^{-1} \in \mathcal{S}_n(\mathbf{R})$ *et* $\exp(\mathrm{ONO}^{-1}) = \mathrm{M}$*.*

Il reste à voir que l'application est injective (d'inverse continu). Tout d'abord, le théorème de réduction nous permet de diagonaliser une matrice $\mathrm{M} \in \mathcal{S}_n$*. Sous cette forme,* $\exp \mathrm{M}$ *est aussi diagonale et* M *et* $\exp \mathrm{M}$ *ont la même décomposition en espaces propres, et les valeurs propres sont liées via l'exponentielle (numérique). Du coup, si* $\exp \mathrm{M} = \exp \mathrm{N}$*, alors la décomposition en sous-espaces propres nous permet de conclure que* $\mathrm{M} = \mathrm{N}$*.*

Enfin, pour voir que l'application réciproque est continue, il suffit de montrer que exp *est propre. Pour cela, on munit* $\mathcal{S}_n$ *de la norme associée à la forme* $q(\mathrm{M}) = \mathrm{tr}^t\mathrm{MM}$. *Restreinte à* $\mathcal{S}_n$, *elle prend la forme* $q(\mathrm{M}) = \mathrm{tr}\,\mathrm{M}^2$, *qui se traduit par la somme des carrés des valeurs propres de* M. *Par suite, si* $\exp \mathrm{M}$ *reste dans un compact de* $\mathcal{S}_n^{++}$, *les valeurs propres de* $\exp \mathrm{M}$ *restent dans un compact de* $\mathbf{R}_+^*$, *donc les valeurs propres restent dans un compact de* $\mathbf{R}$, *et il s'ensuit que* M *aussi reste dans un compact de* $\mathcal{S}_n$. *Du coup, on en déduit que* exp *est continue, propre et injective, donc un homéomorphisme sur son image.* ■

15.1.9 Coniques et quadriques de $\mathbf{R}^2$ et $\mathbf{R}^3$, ellipsoïde

Une cônique est donnée par une équation de la forme $q(x, y) = 1$, où q est un polynôme homogène de degré 2. Autrement dit, q est une forme quadratique. D'après le théorème 15.1.6.1, il existe une base orthonormée de $\mathbf{R}^2$ telle que P ait une forme canonique, qui nous donne la notion d'ellipse, d'hyperbole...

On définit une quadrique de $\mathbf{R}^3$ comme le lieu

$$\mathcal{Q} = \{(x, y, z) \in \mathbf{R}^3,\ q(x, y, z) = 1\}$$

où q est une forme quadratique non dégénérée. On discute selon la signature de q de la forme de la quadrique.

sig(q)=(0,3) Dans une base adaptée de $\mathbf{R}^3$, on a $q(x, y, z) = -x^2 - y^2 - z^2$, donc $\mathcal{Q} = \varnothing$.

sig(q)=(1,2) Dans une base adaptée de $\mathbf{R}^3$, on a $q(x, y, z) = x^2 - y^2 - z^2$. Donc $\mathcal{Q}$ a deux composantes connexes selon que $x \geq 1$ ou $x \leq -1$. La quadrique coupe le plan $\{x = cste\}$, pour $|x| \geq 1$, en un cercle de rayon $x^2 - 1$. On dit que $\mathcal{Q}$ est un hyperboloïde à deux nappes.

sig(q)=(2,1) Dans une base adaptée de $\mathbf{R}^3$, on a $q(x, y, z) = x^2 + y^2 - z^2$. Donc $\mathcal{Q}$ est connexe. La quadrique coupe le plan $\{z = 0\}$ en un cercle. On dit que $\mathcal{Q}$ est un hyperboloïde à une nappe.

Une propriété importante de cette quadrique est qu'elle est très exactement doublement réglées. Un point appartient à $\mathcal{Q}$ si $(y - z)(y + z) = (1 - x)(1 + x)$. On devine l'équation de deux familles de droites incluses dans $\mathcal{Q}$.

$$\Delta_a \begin{cases} y - z = a(1 - x) \\ (y + z)a = 1 + x \end{cases} a \in \mathbf{R} \quad \text{et } \Delta_\infty \begin{cases} y = -z \\ x = 1 \end{cases}$$

ainsi que

$$\mathrm{D}_b \begin{cases} y + z = b(1 - x) \\ (y - z)b = 1 + x \end{cases} b \in \mathbf{R} \quad \text{et } \mathrm{D}_\infty \begin{cases} y = z \\ x = 1 \end{cases}$$

Un simple calcul montre que ces familles sont transverses, et que seule une droite par famille passe par un point donné de $\mathcal{Q}$.

De plus, si L est une droite incluse dans $\mathcal{Q}$ passant par un point p, et si v est un vecteur directeur, on a, pour tout $t \in \mathbf{R}$,

$$1 = q(p + tv) = q(p) + 2tb(p, v) + t^2 q(v)\,,$$

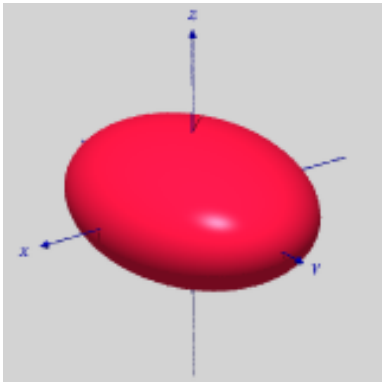
donc $b(p, v) = 0$ et $q(v) = 0$. Ceci implique que $v \in (\mathbf{R}p)^\perp \cap \mathcal{C}(q)$ (il s'agit d'une équivalence). Or, puisque $q(p) \neq 0$, on a $\mathbf{R}^3 = (\mathbf{R}p) \oplus (\mathbf{R}p)^\perp$; de plus $sig(q|_{\mathbf{R}p}) = (1, 0)$ donc $sig(q|_{(\mathbf{R}p)^\perp}) = (1, 1)$.

Or on a vu qu'une forme définie sur un plan avec cette signature avait exactement deux droites isotropes.

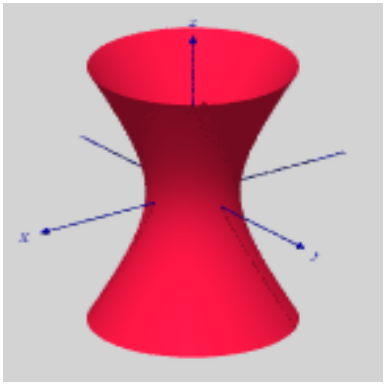
sig(q)=(3,0) Dans une base adaptée de $\mathbf{R}^3$, on a $q(x,y,z) = x^2+y^2+z^2$. Donc $\mathcal{Q}$ est connexe. On dit que $\mathcal{Q}$ est un ellipsoïde. Dans une base orthonormée de $\mathbf{R}^3$, l'équation prend la forme plus générale

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1\,.$$

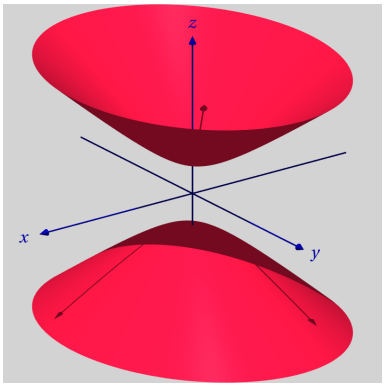
On remarque qu'un ellipsoïde est la sphère unité pour une structure euclidienne appropriée.



ellipsoïde



hyperboloïde à 1 nappe



hyperboloïde à 2 nappes

Exercice(s) 15.1.9.1. *Soit* $\mathrm{S} \in \mathrm{M}_n(\mathbf{R})$ *définie positive,* q_S *la forme quadratique associées. Montrer que le volume* $vol(\mathcal{E}_\mathrm{S})$ *de l'ellipsoïde* $\mathcal{E}_\mathrm{S}$ *défini par l'inégalité* $q_\mathrm{A}(x) \leq 1$ *est* $\mathrm{B}_n/\sqrt{\det(\mathrm{A})}$ *où* B_n *est le volume de la boule unité. En utilisant la stricte concavité du logarithme et le théorème de réduction 15.1.6.1, en déduire que*

$$vol : \begin{cases} \mathcal{S}_n{+}{+} & \to &]0,\infty[\\ \mathrm{S} & \mapsto vol(\mathcal{E}_\mathrm{S}) & \end{cases}$$

est strictement convexe.

15.1.10 Appendice : ellipsoïde de Loewner

Théorème 15.1.10.1 (Théorème de Loewner)**.** *Si* K *est un compact de* $\mathbf{R}^n$ *dont l'intérieur contient l'origine, alors il existe un unique ellipsoïde de volume minimal contenant* K.

En conséquence, si G est un sous-groupe compact de $\mathrm{GL}_n(\mathbf{R})$, il est conjugué à un sous-groupe de $\mathrm{O}(n,\mathbf{R})$ (voir *infra* le théorème 15.1.11.7.)

Remarque(s) 15.1.10.2. *On peut en déduire par polarité le théorème de John qui affirme l'existence d'un ellipsoïde de volume maximal contenu dans* K.

Avant de démontrer le théorème de John, nous établissons quelques lemmes.

Lemme 15.1.10.3. *Soit* $\mathcal{E}_{\mathrm{S}} = \{\mathrm{X} \in \mathbf{R}^n,\ {}^t\mathrm{XSX} \leq 1\}$ *où* $\mathrm{S} \in \mathcal{S}_n^{++}(\mathbf{R})$. *Alors* $\operatorname{vol}\mathcal{E}_{\mathrm{S}} = \mu(\mathrm{S})\operatorname{vol}\mathcal{E}_{\mathrm{I}}$ *où* $\mu : \mathrm{S} \in \mathcal{S}_n^{++}(\mathbf{R}) \mapsto (\det \mathrm{S})^{-1/2}$.

Démonstration. *D'après le théorème de réduction, il existe* $\mathrm{O} \in \mathrm{O}_n(\mathbf{R})$ *telle que* ${}^t\mathrm{OSO} = \mathrm{D}$ *soit une matrice diagonale dont les coefficients (diagonaux)* $\lambda_1, \ldots, \lambda_n$ *sont tous strictement positifs. On considère* D' *la matrice diagonale dont les termes diagonaux sont* $1/\sqrt{\lambda_j}$, *et on note* $\mathrm{R} = \mathrm{OD}'\mathrm{O}^{-1}$, *qui est inversible et symétrique. On a donc* $\mathrm{RSR} = \mathrm{I}_n$. *De plus*

$$\mathcal{E}_{\mathrm{S}} = \{\mathrm{X} \in \mathbf{R}^n,\ {}^t\mathrm{XR}^{-2}\mathrm{X} \leq 1\} = \{\mathrm{X} \in \mathbf{R}^n,\ {}^t(\mathrm{R}^{-1}\mathrm{X})(\mathrm{R}^{-1}\mathrm{X}) \leq 1\} = \{\mathrm{X} \in \mathbf{R}^n,\ \mathrm{R}^{-1}(\mathrm{X}) \in \mathcal{E}_{\mathrm{I}}\} = \mathrm{R}(\mathcal{E}_{\mathrm{I}})\,.$$

Par la formule de changement de variables, on obtient le résultat recherché :

$$\operatorname{vol}\mathcal{E}_{\mathrm{S}} = \det \mathrm{R}\operatorname{vol}\mathcal{E}_{\mathrm{I}} = \mu(\mathrm{S})\operatorname{vol}\mathcal{E}_{\mathrm{I}}\,.$$

■

Lemme 15.1.10.4. *Les espaces* $\mathcal{S}_n^{++}(\mathbf{R})$ *et* $\mathcal{S}_n^{+}(\mathbf{R})$ *sont convexes.*

Démonstration. *On traite le cas de* $\mathcal{S}_n^{++}(\mathbf{R})$. *Soient* $\mathrm{S}_0, \mathrm{S}_1 \in \mathcal{S}_n^{++}(\mathbf{R})$. *On note, pour* $s \in [0,1]$, $\mathrm{S}_s = (1-s)\mathrm{S}_0 + s\mathrm{S}_1$. *On a* $\mathrm{S}_s \in \mathcal{S}_n(\mathbf{R})$, *et, pour tout* $\mathrm{X} \in \mathbf{R}^n \setminus \{0\}$,

$${}^t\mathrm{XS}_s\mathrm{X} = (1-s)({}^t\mathrm{XS}_0\mathrm{X}) + s({}^t\mathrm{XS}_1\mathrm{X}) > 0\,.$$

■

Lemme 15.1.10.5. *L'application* $\mu : \mathrm{S} \in \mathcal{S}_n^{++}(\mathbf{R}) \mapsto (\det \mathrm{S})^{-1/2}$ *est strictement convexe.*

Démonstration. *Soient* $\mathrm{S}_0, \mathrm{S}_1 \in \mathcal{S}_n^{++}(\mathbf{R})$ *distinctes. On note, pour* $s \in [0,1]$, $\mathrm{S}_s = (1-s)\mathrm{S}_0 + s\mathrm{S}_1$. *D'après le théorème de réduction, il existe* $\mathrm{P} \in \mathrm{GL}_n(\mathbf{R})$ *telle que* ${}^t\mathrm{PS}_0\mathrm{P} = \mathrm{I}_n$ *et* ${}^t\mathrm{PS}_1\mathrm{P} = \mathrm{D}$ *soit une matrice diagonale de coefficients (diagonaux)* $\lambda_1, \ldots, \lambda_n$. *Si* $\mathrm{D} = \mathrm{I}_n$ *alors on aurait* ${}^t\mathrm{PS}_0\mathrm{P} = {}^t\mathrm{PS}_1\mathrm{P}$ *et* $\mathrm{S}_0 = \mathrm{S}_1$, *ce qui est contraire à l'hypothèse. Donc* $\mathrm{D} \neq \mathrm{I}_n$, *et on peut donc supposer que* $\lambda_1 \neq 1$.
Du coup,

$$\det \mathrm{S}_s = \frac{1}{\det^2 \mathrm{P}} \det({}^t\mathrm{PS}_s\mathrm{P}) = \frac{1}{\det^2 \mathrm{P}} \det((1-s)\mathrm{I}_n + s\mathrm{D}) = \frac{1}{\det^2 \mathrm{P}} \prod[(1-s) + s\lambda_j]\,.$$

Posons $\mathrm{A}_j(s) = (1-s) + s\lambda_j$ *et* $u(s) = \mu(\mathrm{S}_s)/|\det \mathrm{P}|$. *Cette application est différentiable, et*

$$u'(s) = \sum_{j=1}^{n} \frac{-1}{2} \frac{\lambda_j - 1}{\mathrm{A}_j(s)^{3/2}} \frac{1}{\prod_{i \neq j} \mathrm{A}_i(s)^{1/2}} = \frac{-1}{2} u(s) \sum_{j=1}^{n} \frac{\lambda_j - 1}{\mathrm{A}_j(s)}$$

et

$$u''(s) = \frac{1}{4}u(s)\left(\sum_{j=1}^{n}\frac{\lambda_j - 1}{\mathrm{A}_j(s)}\right)^2 + \frac{1}{2}u(s)\sum_{j=1}^{n}\left(\frac{\lambda_j - 1}{\mathrm{A}_j(s)}\right)^2 \geq \frac{1}{2}u(s)\left(\frac{\lambda_1 - 1}{\mathrm{A}_1(s)}\right)^2 > 0\,.$$

■

Nous pouvons maintenant nous atteler à la démonstration du théorème de John.

Démonstration. *Par hypothèse, il existe* $\rho_1, \rho_2 > 0$ *tel que* $\mathrm{B}(0,\rho_1) \subset \mathrm{K} \subset \mathrm{B}(0,\rho_2)$.

Assertion. *Considérons*

$$\mathcal{C} = \{\mathrm{S} \in \mathcal{S}_n^{++}(\mathbf{R}), \mathrm{K} \subset \mathcal{E}_\mathrm{S} \text{ et } vol(\mathcal{E}_\mathrm{S}) \leq vol(\mathrm{B}(0,\rho_2))\}.$$

est convexe (car l'application voume est convexe (cf. exercice 15.1.9.1)), non vide (car $\rho_2^{-1}\,\mathrm{Id} \in \mathcal{C}$*). Montrons que* $\mathcal{C}$ *est compact. La fermeture de* $\mathcal{C}$ *dans* $\mathcal{S}_n$ *est claire puisque : en effet, dire*

$$\mathrm{S} \in \mathcal{S}_n^{++} \text{ et } vol(\mathcal{E}_\mathrm{S}) \leq vol(\mathrm{B}(0,\rho_2)),$$

c'est dire

$$\mathrm{S} \in \mathcal{S}_n^{+} \text{ et } \sqrt{\det(\mathrm{S})} \geq \rho_2^{-n}$$

qui sont des conditions fermées dans $\mathcal{S}_n$*. De plus, si* x *de norme* 1, $\rho_1 \mathrm{S}x \in \mathcal{E}_\mathrm{S}$ *ie* $q_\mathrm{S}(x) \leq \rho_1^{-2}$ *de sorte* S *est bornée ce qui donne la compacité. Par continuité l'application volume y atteint un minimum en au moins un point. Par stricte convexité,ce point est unique.* ■

15.1.11 Propriétés topologiques du groupe orthogonal

Lemme 15.1.11.1. $\mathrm{O}_n(\mathbf{R})$ *est compact.*

Démonstration. *On considère l'application* $f : \mathrm{M} \in \mathrm{M}_n(\mathbf{R}) \mapsto {}^t\mathrm{MM}$*. On a* $\mathrm{O}_n(\mathbf{R}) = f^{-1}(\mathrm{I}_n)$*, donc* $\mathrm{O}_n(\mathbf{R})$ *est fermé. Pour montrer que* $\mathrm{O}(n)$ *est borné, on considère la norme induite par* $\mathrm{tr}({}^t\mathrm{MM})$*. On a* $\mathrm{O}_n(\mathbf{R}) \subset \mathrm{B}(0,\sqrt{n})$. ■

Proposition 15.1.11.2. *Le groupe* $\mathrm{O(E)}$ *a exactement deux composantes connexes qui sont homéormorphes :* $\mathrm{SO(E)}$ *et* $\mathrm{O(E)}^- = \mathrm{O(E)} \setminus \mathrm{SO(E)}$.

Démonstration. *On montre d'abord que* $\mathrm{SO(E)}$ *est connexe par arcs. Puisque* q *est pair, on peut remplacer* $-\mathrm{I}_q$ *par une matrice par blocs de rotations d'angle* π*. Du coup,*

$$\mathrm{Mat}(u) = \begin{pmatrix} \mathrm{I}_p & & 0 & \\ & \mathrm{R}_{\theta_1} & & \\ & & \ddots & \\ 0 & & & \mathrm{R}_{\theta_q} \end{pmatrix}$$

Pour $t \in [0,1]$, on pose

$$\mathrm{Mat}(u_t) = \begin{pmatrix} \mathrm{I}_p & & 0 & \\ & \mathrm{R}_{t\theta_1} & & \\ & 0 & \ddots & \\ & & & \mathrm{R}_{t\theta_q} \end{pmatrix}$$

On a $u_0 = \mathrm{Id}$ et $u_1 = u$. On vérifie que chaque $u_t \in \mathrm{SO(E)}$.

On considère l'application ι définie en base orthonormée par

$$\mathrm{Mat}(\iota) = \begin{pmatrix} -1 & & 0 & \\ 0 & 1 & & \\ & & \ddots & \\ & 0 & & 1 \end{pmatrix}$$

La multiplication à gauche par cette matrice définit un homéomorphisme entre $\mathrm{SO(E)}$ *et* $\mathrm{O}^-(\mathrm{E})$. *Enfin, l'application* $\det : \mathrm{O(E)} \to \{\pm 1\}$ *montre que* $\mathrm{O(E)}$ *n'est pas connexe.* ∎

Décomposition polaire. L'application $\Phi : \mathrm{O}(n) \times \mathcal{S}_n^{++}(\mathbf{R}) \to \mathrm{GL}_n(\mathbf{R})$ définie par $\Phi(\mathrm{O}, \mathrm{S}) = \mathrm{OS}$ est un homéomorphisme.

Lemme 15.1.11.3. *Si $\mathrm{M} \in \mathrm{GL}_n(\mathbf{R})$, alors ${}^t\mathrm{M} \cdot \mathrm{M}$ est définie positive.*

Démonstration. *Si $\mathrm{X} \in \mathbf{R}^n - 0$, on a*

$${}^t\mathrm{X}({}^t\mathrm{M} \cdot \mathrm{M})\mathrm{X} = {}^t(\mathrm{MX})(\mathrm{MX}) = \|\mathrm{MX}\|^2 > 0.$$

∎

Idée Si $\mathrm{M} = \mathrm{OS}$, alors ${}^t\mathrm{M} = \mathrm{SO}^{-1}$ donc ${}^t\mathrm{M} \cdot \mathrm{M} = \mathrm{S}^2$, avec S définie positive.

Démonstration de la décomposition : existence. — D'après le lemme, il existe $\widehat{\mathrm{O}} \in \mathrm{O}(n)$ telle que

$$\widehat{\mathrm{O}}^{-1} \cdot ({}^t\mathrm{M} \cdot \mathrm{M}) \cdot \widehat{\mathrm{O}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

soit diagonale avec des valeurs propres strictement positives. On note

$$\mathrm{S} = \widehat{\mathrm{O}} \cdot \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \cdot \widehat{\mathrm{O}}^{-1}$$

et $\mathrm{O} = \mathrm{MS}^{-1}$. Comme S est symétrique, S^{-1} aussi, donc

$${}^t\mathrm{O} \cdot \mathrm{O} = {}^t(\mathrm{S}^{-1}) \cdot ({}^t\mathrm{M} \cdot \mathrm{M}) \cdot \mathrm{S}^{-1} = \mathrm{S}^{-1}\mathrm{S}^2\mathrm{S}^{-1} = \mathrm{I}\,.$$

■

Démonstration de la décomposition : unicité. — Soit $P \in \mathbf{R}[X]$ tel que $P(\lambda_j) = \sqrt{\lambda_j}$. On a

$$P({}^t M \cdot M) = \sum a_k \widehat{O} \cdot \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} \cdot \widehat{O}^{-1} = \widehat{O} \cdot \begin{pmatrix} P(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & P(\lambda_n) \end{pmatrix} \cdot \widehat{O}^{-1} = S\,.$$

Donc, si M admet une seconde décomposition polaire $M = O'S'$ alors ${}^t M \cdot M = (S')^2$ donc S' commute avec ${}^t M \cdot M$, donc avec S puisque S est un polynôme en ${}^t M \cdot M$. Du coup, S et S' peuvent être diagonalisé simultanément (en effet, $E_{\lambda_i}(S) = \oplus(E_{\lambda_i}(S) \cap E_{\mu_j}(S'))$). Ceci oblige les valeurs propres à coïncider, donc ces matrices sont les mêmes. ■

Démonstration de la décomposition : homéomorphie. — L'application Φ est clairement continue. Réciproquement, si $M_n = \Phi(O_n, S_n)$ converge vers une matrice $M = OS$, montrons que O_n tend vers O et S_n vers S. Comme $O(E)$ est compact, quitte à extraire une sous-suite, on peut supposer que O_n tend vers une matrice orthogonale $\widehat{O}$. Du coup, S_n tend vers une matrice (symétrique) $\widehat{O}^{-1}M$. Du coup, on a $OS = \widehat{O}(\widehat{O}^{-1}M)$. Par l'unicité de la décomposition polaire, on obtient $O = \widehat{O}$ et S_n tend vers S. ■

On en déduit quelques résultats.

Proposition 15.1.11.4. $GL_n(\mathbf{R})$ *a exactement deux composantes connexes.*

Démonstration. *On a*

$$GL_n(\mathbf{R}) \approx O_n(\mathbf{R}) \times \mathcal{S}_n^{++}(\mathbf{R}) = (SO_n(\mathbf{R}) \times \mathcal{S}_n^{++}(\mathbf{R})) \cup (O_n^-(\mathbf{R}) \times \mathcal{S}_n^{++}(\mathbf{R}))$$

par la décomposition polaire. ■

Exercice(s) 15.1.11.5. *Soit* $D(t)$ *la dilatation* $\mathrm{Id} + (t-1)E_{1,1}$. *Montrer que* $(t, M) \mapsto D(t)M$ *est un homéomorphisme de* $\mathbf{R}^* \times SL_n(\mathbf{R})$ *sur* $GL_n(\mathbf{R})$. *En utilisant le pivot de Gauss, montrer que* $SL_n(\mathbf{R})$ *est engendré par les produits d'au plus* n^2 *transvections. En déduire que* $SL_n(\mathbf{R})$ *est connexe puis que* $GL_n(\mathbf{R})$ *a deux composantes connexes. Que se passe-t-il sur* $\mathbf{C}$ *?*

Proposition 15.1.11.6. $SO_n(\mathbf{R})$ *(resp.* $O_n(\mathbf{R})$*) est un sous-groupe compact maximal de* $SL_n(\mathbf{R})$ *(resp. de* $GL_n(\mathbf{R})$*).*

Démonstration. *Soit* $G \subset SL_n(\mathbf{R})$ *un sous-groupe compact qui contient* $SO_n(\mathbf{R})$*, et soit* $g \in G \setminus SO_n(\mathbf{R})$. *On a* $g = OS$ *avec* $O \in SO_n(\mathbf{R})$ *car* $\det g > 0$ *et* $S \neq I$ *car* $g \notin SO_n(\mathbf{R})$. *Donc* $S \in G$. *Or si* v *est un vecteur propre associé à une valeur propre* λ *de* S *différente de* 1, *alors* $\mathrm{Log}\,\|S^n x\|$ *tend vers l'infini, ce qui contredit la compacité de* G. ■

On montre que l'on peut améliorer ce résultat de la manière suivante.

Théorème 15.1.11.7. *Un sous-groupe compact* G *de* $\mathrm{GL}_n(\mathbf{R})$ *est conjugué à un sous-groupe de* $\mathrm{O}_n(\mathbf{R})$.

Démonstration. *On remarque tout d'abord que si* $\mathcal{E}$ *est un ellipsoïde et si* $\mathrm{M} \in \mathrm{GL}_n(\mathbf{R})$, *alors* $\mathrm{M}\mathcal{E}$ *est aussi un ellipsoïde. En effet si* $\mathcal{E} = \{{}^t\mathrm{XSX} = 1\}$ *avec* $\mathrm{S} \in \mathcal{S}_n^{++}(\mathbf{R})$, *alors*

$$\mathrm{M}\mathcal{E} = \{{}^t(\mathrm{M}^{-1}\mathrm{X})\mathrm{SM}^{-1}\mathrm{X} = 1\} = \{{}^t\mathrm{X}({}^t\mathrm{M}^{-1}\mathrm{SM}^{-1})\mathrm{X} = 1\}$$

Or ${}^t\mathrm{M}^{-1}\mathrm{SM}^{-1}$ *est aussi définie positive car il ne s'agit que d'un changement de variables, donc* $\mathrm{M}\mathcal{E}$ *est aussi un ellipsoïde.*

Soit B *la boule unité fermée de* $\mathbf{R}^n$. *On note* $\mathrm{K} = \cup_{g\in\mathrm{G}} g(\mathrm{B})$. *Alors* K *est compact car* G *et* B *le sont,* K *est invariant par définition, et* 0 *est un point intérieur de* K *car* K *contient* $\mathrm{I}(\mathrm{B}) = \mathrm{B}$. *D'après le théorème de Loewner, il existe un unique ellipsoïde* $\mathcal{E}_\mathrm{S}$ *qui contient* K *de volume minimal.*

Puisque G *est compact, on a, pour tout* $g \in \mathrm{G}$, $|\det g| = 1$. *Donc* $\operatorname{vol} g(\mathcal{E}_\mathrm{S}) = \operatorname{vol}\mathcal{E}_\mathrm{S}$, *et comme* $\mathrm{K} = g(\mathrm{K}) \subset g(\mathcal{E}_\mathrm{S})$, *on obtient* $g(\mathcal{E}_\mathrm{S}) = \mathcal{E}_\mathrm{S}$. *Par suite,* $\mathcal{E}_\mathrm{S}$ *est invariant par* G. *Soit* T *une racine carrée de* S^{-1}. *Alors* $\mathcal{E}_\mathrm{S} = \mathrm{T}(\mathrm{B})$, *et* $\mathrm{TGT}^{-1} \subset \mathrm{O}_n(\mathbf{R})$. ∎

Voici un autre argument. Puisque G est un groupe compact métrique, il existe une mesure de Haar *i.e.*, une mesure de probabilité μ borélienne sur G invariante par multiplication à gauche et à droite dans G. Autrement dit, si $\varphi : \mathrm{G} \to \mathbf{R}$ est une fonction intégrable, alors

$$\int_\mathrm{G} \varphi(g)d\mu(g) = \int_\mathrm{G} \varphi(hg)d\mu(g) = \int_\mathrm{G} \varphi(gh)d\mu(g)$$

pour tout $h \in \mathrm{G}$. On définit sur $\mathbf{R}^n$ la forme

$$\langle x, y\rangle_\mathrm{G} = \int_\mathrm{G} \langle gx, gy\rangle d\mu(g)\,.$$

Il s'agit clairement d'une forme bilinéaire symétrique positive. Si $\langle x, x\rangle_\mathrm{G} = 0$ alors il existe $g \in \mathrm{G}$ tel que $\|g(x)\| = 0$, donc $x = 0$. Du coup, $\langle\cdot,\cdot\rangle_\mathrm{G}$ est un produit scalaire sur $\mathbf{R}^n$ représenté par une matrice symétrique M.

Or, si $h \in \mathrm{G}$, l'invariance de μ implique que

$$\langle h(x), h(y)\rangle_\mathrm{G} = \int_\mathrm{G} \langle ghx, ghy\rangle d\mu(g) = \int_\mathrm{G} \langle gx, gy\rangle d\mu(g) = \langle x, y\rangle_\mathrm{G}$$

donc h est une isométrie pour cette structure euclidienne.

Par le théorème de Sylvester, il existe une matrice $\mathrm{P} \in \mathrm{GL}_n(\mathbf{R})$ telle que ${}^t\mathrm{PMP} = \mathrm{I}_n$. Du coup, $\mathrm{PGP}^{-1} \subset \mathrm{O}_n(\mathbf{R})$.

Remarque(s) 15.1.11.8. *Il est facile de voir que* $\mathrm{O}_2(\mathbf{C})$ *n'est pas compact, donc* $\mathrm{O}_n(\mathbf{C})$ *non plus.*

Proposition 15.1.11.9. *Les espaces* $\mathrm{GL}_n(\mathbf{R})$ *et* $\mathrm{SL}_n(\mathbf{R})$ *sont respectivement homéomorphes à* $\mathrm{O}_n(\mathbf{R}) \times \mathbf{R}^{\frac{n(n+1)}{2}}$ *et* $\mathrm{SO}_n(\mathbf{R}) \times \mathbf{R}^{\frac{n(n-1)}{2}}$.

Démonstration. *L'application* $\exp : \mathcal{S}_n(\mathbf{R}) \to \mathcal{S}_n^{++}(\mathbf{R})$ *est un homéomorphisme et* $\mathcal{S}_n(\mathbf{R}) \approx \mathbf{R}^{\frac{n(n+1)}{2}}$, *donc*

$$\mathrm{GL}_n(\mathbf{R}) \approx \mathrm{O}_n(\mathbf{R}) \times \mathcal{S}_n^{++}(\mathbf{R}) \approx \mathrm{O}_n(\mathbf{R}) \times \mathbf{R}^{\frac{n(n+1)}{2}}.$$

De même, $\mathrm{SL}_n(\mathbf{R}) \approx \mathrm{SO}_n(\mathbf{R}) \times (\mathcal{S}_n^{++}(\mathbf{R}) \cap \mathrm{SL}_n(\mathbf{R}))$ *et l'application* $\exp : \mathcal{S}_n(\mathbf{R}) \cap \mathrm{tr}^{-1}\{0\} \to \mathcal{S}_n^{++}(\mathbf{R}) \cap \mathrm{SL}_n(\mathbf{R})$ *est un homéomorphisme.* ∎

Propriétés algébriques.

Théorème 15.1.11.10. $\mathrm{O(E)}$ *est engendré par des réflexions. Plus précisément, si* $u \in \mathrm{O(E)}$, *alors* u *est produit d'au plus* $\dim \mathrm{E} - \dim \mathrm{Ker}(\mathrm{Id} - u)$ *réflexions.*

Démonstration. *On constate tout d'abord que le produit de* 2 *réflexions dans* $\mathbf{R}^2$ *par rapport à des droites* e_1 *et* e_2 *est une rotation d'angle 2 fois l'angle entre* e_1 *et* e_2. *Donc, si on écrit la forme réduite de* u, *chaque bloc* R_θ *compte pour deux réflexions, alors que chaque* (-1) *compte pour une seule.* ∎

Exercice. Montrer que SO(E) est engendré par des renversements.

Théorème 15.1.11.11. $\mathrm{Z(O(E))} = \{\pm \mathrm{Id}\}$ *et,* $\mathrm{Z(SO(E))} = \{\mathrm{Id}\}$ *si* $\dim \mathrm{E}$ *est impaire,* $\mathrm{Z(SO(E))} = \{\pm \mathrm{Id}\}$ *si* $\dim \mathrm{E}$ *est paire et* $\dim \mathrm{E} \geq 4$, *et* $\mathrm{Z(SO(E))} = \mathrm{SO(E)}$ *si* $\dim \mathrm{E} = 2$.

Démonstration. *Soit* x *de norme* 1. *On complète en une base orthonormée. La symétrie par rapport à* x *s'écrit*

$$\mathrm{Mat}(s_x, \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 0 & -\mathrm{I} \end{pmatrix}.$$

Si $u \in \mathrm{O(E)}$, *alors* $us_x u^{-1} = s_{u(x)}$, *donc si* $u \in \mathrm{Z(O(E))}$, *alors* $s_x = s_{u(x)}$, *donc il existe* $\lambda_x \in \mathbf{R}$ *telle que* $u(x) = \lambda_x x$. *Comme* $u \in \mathrm{O(E)}$, *on a* $\lambda_x = \pm 1$. *Ceci implique que* u *est une homothétie de rapport* $\lambda = \pm 1$. *En effet, on a, pour* x, y *indépendants,*

$$u(x+y) = \lambda_{x+y} x + \lambda_{x+y} y = \lambda_x x + \lambda_y y.$$

Quant au centre de $\mathrm{SO(E)}$, *on raisonne de la même manière.* ∎

Théorème 15.1.11.12. $\mathrm{D}(\mathrm{O}(\mathrm{E})) = \mathrm{SO}(\mathrm{E})$ *et,* $\mathrm{D}(\mathrm{SO}(\mathrm{E})) = \mathrm{SO}(\mathrm{E})$ *si* $\dim \mathrm{E} \geq 3$ *et* $\mathrm{D}(\mathrm{SO}(\mathrm{E})) = \{id\}$ *si* $\dim \mathrm{E} = 2$.

Démonstration. *Si* $u, v \in \mathrm{O}(\mathrm{E})$ *alors* $\det uvu^{-1}v^{-1} = 1$, *donc* $\mathrm{D}(\mathrm{O}(\mathrm{E})) \subset \mathrm{SO}(\mathrm{E})$. *Or les produits pairs de réflexions engendrent* $\mathrm{SO}(\mathrm{E})$. *Montrons que les produits de deux réflexions sont des commutateurs : soient* x, y *unitaires. Il existe* $u \in \mathrm{O}(\mathrm{E})$ *tel que* $u(x) = y$. *On a* $s_y = s_{u(x)} = u \circ s_x \circ u^{-1}$, *donc*

$$s_x \circ s_y = s_x \circ u \circ s_x \circ u^{-1} = s_x \circ u \circ s_x^{-1} \circ u^{-1} .$$

■

Théorème 15.1.11.13. $\mathrm{SO}(\mathrm{E})$ *est simple si* $\dim \mathrm{E} = 3$.

Démonstration. *Soit* G *un sous-groupe distingué de* $\mathrm{SO}(\mathrm{E})$ *non trivial. Pour montrer que* $\mathrm{G} = \mathrm{SO}(\mathrm{E})$, *il suffit de montrer que* G *contient un demi-tour. A ce moment-là, on saura qu'il les contient tous par conjugaison, et donc que* $\mathrm{G} = \mathrm{SO}(\mathrm{E})$. *Soit* $g \in \mathrm{G}$ *non trivial. Comme* $g \in \mathrm{SO}(\mathrm{E})$, *il s'agit d'une rotation d'axe* x *et d'angle* θ. *Si* $\theta = \pi$, *alors on a gagné.*
Si $\theta \neq \pi$, *on remarque que, pour* $v \in \mathrm{SO}(\mathrm{E})$, *on a* $vgv^{-1}g^{-1} \in \mathrm{G}$. *En particulier, si* $v = s_y$, *où* $y \in \mathrm{E}\backslash\{0\}$, *alors* $s_y g s_y g^{-1} = s_y \circ s_{g(y)} \in \mathrm{G}$.
Si y *et* $g(y)$ *sont orthogonaux, alors* $s_y \circ s_{g(y)}$ *serait un demi-tour. Pour voir cela, il suffit de considérer une base orthonormée contenant* y *et* $g(y)$.
Pour conclure, on cherche donc $y \neq 0$ *tel que* $g(y) \perp y$. *Soit* $(x, \mathrm{e}_2, \mathrm{e}_3)$ *une base orthonormée. On a*

$$\mathrm{Mat}(g) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

donc si $y = y_1 x + y_2 \mathrm{e}_2 + y_3 \mathrm{e}_3$, *on cherche à résoudre*

$$y_1^2 + y_2(y_2 \cos\theta - y_3 \sin\theta) + y_3(y_2 \sin\theta + y_3 \cos\theta) = 0 ,$$

soit

$$y_1^2 + (y_2^2 + y_3^2)\cos\theta = 0 .$$

Quitte à itérer g, *on peut supposer que* $\cos\theta \leq 0$. *Du coup, une solution existe.* ■

15.1.12 Appendice : Sous-groupes finis d'isométries en petite dimension

On s'intéresse à la classification des sous-goupes finis de $\mathrm{O}_2(\mathbf{R})$ et de $\mathrm{SO}_3(\mathbf{R})$ et à leurs relations avec la géométrie.

Sous-groupes de $\mathrm{O}_2(\mathbf{R})$**.** On s'intéresse d'abord à $\mathrm{O}_2(\mathbf{R})$ en vue de $\mathrm{SO}_3(\mathbf{R})$.

Proposition 15.1.12.1. *Soit* G *un sous-groupe fini de* $\mathrm{O}_2(\mathbf{R})$ *non trivial. Alors* G *préserve un* n*-gône régulier. Si* G *n'est constitué que de rotations, alors* G *est isomorphe à* $\mathbf{Z}/n\mathbf{Z}$*. Sinon,* G *est isomorphe à un groupe diédral* D_n*.*

Démonstration. *Si* G *n'est constitué que de rotations, alors* G *s'identifie à un sous-groupe fini de* $\mathbf{R}/\mathbf{Z}$*. On note* $p : \mathbf{R} \to \mathbf{R}/\mathbf{Z}$ *la projection canonique. Alors* $p^{-1}(\mathrm{G})$ *est un sous-groupe de* $\mathbf{R}$ *qui contient les entiers. Puisque* G *est fini,* $p^{-1}(\mathrm{G})$ *est discret, donc il s'agit de* $(1/n)\mathbf{Z}$ *pour un entier* $n \geq 1$*. Du coup* G *est isomorphe à* $\mathbf{Z}/n\mathbf{Z}$*. On en déduit que* G *préserve un* n*-gône régulier inscrit dans le disque unité.*

Sinon, notons $\mathrm{SG} = \mathrm{G} \cap \mathrm{SO}_2(\mathbf{R})$*. D'après ci-dessous,* SG *est isomorphe à* $\mathbf{Z}/n\mathbf{Z}$*. De plus, on a la sous-suite exacte courte*

$$1 \to \mathrm{SG} \to \mathrm{G} \to \mathbf{Z}/2\mathbf{Z} \to 1\,.$$

Donc G *est d'ordre* $2n$*.*

Soit σ *une symétrie de* G*. Elle fixe deux poins opposés* x *et* $-x$ *du cercle unité* $\mathbf{S}^1$*. On écrit alors* $\sigma = \sigma_x = \sigma_{-x}$*. On note* X *l'ensemble des points fixes de toutes les symétries de* $\mathrm{G} \setminus \mathrm{SG}$*. On a* n *symétries donc* $2n$ *points fixes.*

D'autre part, on a, pour $g \in \mathrm{G}$*,*

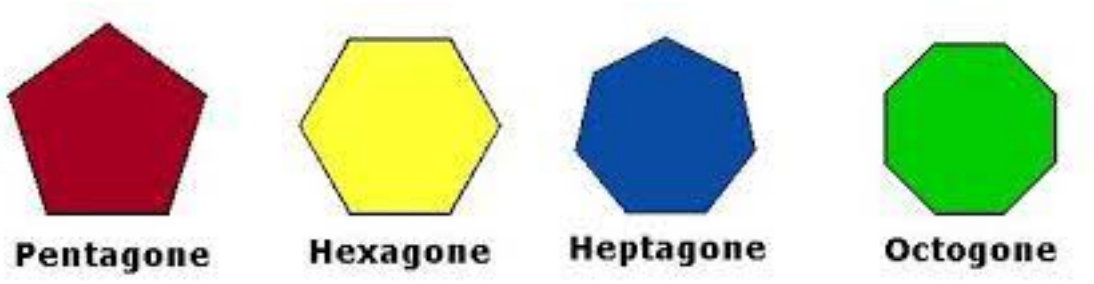
$$g \circ \sigma_x \circ g^{-1}(g(x)) = g(x)$$

donc

$$g \circ \sigma_x \circ g^{-1} = \sigma_{g(x)}$$

et G *opère sur* X*. Le stabilisateur de chaque point de* X *est d'ordre deux, donc chaque orbite est de cardinal* n*. Du coup, on a deux orbites sous l'action de* SG*.*

On en déduit que G *est isomorphe à* D_n*.* ∎



Sous-groupes de $\mathrm{SO}_3(\mathbf{R})$**.** Soit G un sous-groupe de $\mathrm{SO}_3(\mathbf{R})$ d'ordre N. Chaque élément non trivial est une rotation, donc fixe deux points opposés sur la sphère $\mathbf{S}^2$. On note $\mathrm{X} = \mathrm{X}_\mathrm{G}$ l'ensemble de ces points. Comme ci-dessus, G opère sur X.

Le stabilisateur de chaque point x de X fixe le plan orthogonal $x^\perp$. Sa restriction est un sous-groupe fini de $\mathrm{SO}_2(\mathbf{R})$, donc isomorphe à $\mathbf{Z}/r_x\mathbf{Z}$. Son orbite est donc d'ordre $n_x = \mathrm{N}/r_x$. Chaque point x est

le point fixe de $(r_x - 1)$ rotations, étant différentes pour tous les autres points excepté son opposé. Du coup, on a

$$2\mathrm{N} - 2 = \sum_{x\in \mathrm{X}} (r_x - 1) = \sum_{j\in \mathrm{X/G}} n_j (r_j - 1)\,.$$

On en déduit que

$$2 - 2/\mathrm{N} = \sum_{j\in \mathrm{X/G}} (1 - 1/r_j)\,.$$

Or $r_j \geq 2$ par définition, et donc

$$2 > 2 - 2/\mathrm{N} = \sum_{j\in \mathrm{X/G}} (1 - 1/r_j) \geq |\mathrm{X/G}|/2\,.$$

Par conséquent, on a au plus trois orbites. D'autre part, le groupe G n'opère pas transitivement sur X. En effet, on aurait $2 - 2/\mathrm{N} = 1 - 1/r$, soit $1 = 2/\mathrm{N} - 1/r \leq 1/\mathrm{N}$, puisque $r \leq \mathrm{N}$!!

Proposition 15.1.12.2. *Si on a deux orbites, alors* G *est un groupe de rotations du plan, isomorphe à* $\mathbf{Z}/\mathrm{N}\mathbf{Z}$.

Démonstration. *On a* $2 - 2/\mathrm{N} = 2 - (1/r_1 + 1/r_2)$ *soit* $2/\mathrm{N} = 1/r_1 + 1/r_2$. *Si* $\mathrm{N} = r_1$ *alors* $r_2 = \mathrm{N}$. *Du coup,* X *a deux éléments, et* G *fixe leur orthogonal. Par conséquent, il opère comme un sous-groupe de* $\mathrm{SO}_2(\mathbf{R})$ *et il est isomorphe à* $\mathbf{Z}/\mathrm{N}\mathbf{Z}$.
Sinon, on a $2r_1 \leq \mathrm{N}$, *soit* $(1/r_1) \geq 2/\mathrm{N}$ *donc* $r_2 \leq 0$ *!!* ∎

Le cas de trois orbites comporte plusieurs cas. Notre équation s'écrit

$$\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} = 1 + \frac{2}{\mathrm{N}}\,,$$

où on choisit $r_1 \leq r_2 \leq r_3$.
Si $r_1 \geq 3$, alors le terme de gauche est plus petit que 1 alors que le second est strictement plus grand. Donc $r_1 = 2$.

Si $r_2 = 2$, alors $1/r_3 = 2/\mathrm{N}$, soit $\mathrm{N} = 2r_3$. On a donc $r_j = (2, 2, \mathrm{N}/2)$ et $n_j = (\mathrm{N}/2, \mathrm{N}/2, 2)$.

Si $r_2 \geq 3$ alors $1/r_3 = 1/2 + 2/\mathrm{N} - 1/r_2 > 1/2 - 1/3$, donc $r_3 < 6$.

- Si $r_j = (2, 3, 3)$ alors $\mathrm{N} = 12$ et $n_j = (6, 4, 4)$.
- Si $r_j = (2, 3, 4)$ alors $\mathrm{N} = 24$ et $n_j = (12, 8, 6)$.
- Si $r_j = (2, 3, 5)$ alors $\mathrm{N} = 60$ et $n_j = (30, 20, 12)$.

Si $r_2 \geq 4$, alors

$$\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} \leq 1 < 1 + \frac{2}{\mathrm{N}}\,,$$

donc on a la liste complète.

Proposition 15.1.12.3. G *est un groupe diédral* D_r *dans le cas*

$$\mathrm{N} = 2r \quad r_j = (2, 2, r) \quad n_j = (r, r, 2)\,.$$

Démonstration. *Les points opposés ont les mêmes comportements donc la troisième orbite est constitué de deux points opposés. Par suite,* G *fixe leur orthogonale, et les isométries qui ne fixent pas ces points sont des symétries sur ce plan. Donc il s'agit du groupe diédral.* ∎

Si on n'est pas dans un des cas précédents, alors aucune orbite n'est planaire. En effet, la restriction au plan nous ramenerait aux cas déjà traités.

Proposition 15.1.12.4. *Si* $\mathrm{N} = 12$*,* $r_j = (2, 3, 3)$*, et* $n_j = (6, 4, 4)$ *alors* G *est le groupe d'isométries d'un tétraèdre, et est isomorphe à* $\mathfrak{a}_4$*.*

Démonstration. *Soit* $x \in o(2)$*. Son stabilisateur est un groupe de rotations qui opère sur son orbite, donc sur trois points. Ces points forment un triangle équilatéral. Il s'agit donc d'un tétraèdre. De plus,* G *opère sur ces sommets : il s'identifie à un sous-groupe de permutation à* 4 *éléments* $\mathfrak{S}_4$*. Or le seul endomorphisme qui fixe ces quatre points non coplanaires est l'identité. Donc* G *est un sous-groupe d'indice* 2 *: il s'agit de* $\mathfrak{a}_4$*.* ∎

Proposition 15.1.12.5. *Si* $\mathrm{N} = 24$*,* $r_j = (2, 3, 4)$*, et* $n_j = (12, 8, 6)$ *alors* G *est le groupe d'isométries d'un cube et d'un octaèdre, et est isomorphe à* $\mathfrak{S}_4$*.*

Démonstration. *Le stabilisateur d'un point* $x \in o(3)$ *opère sur* $o(2)$*, en deux orbites. Chacune forme un carré, et ces deux carrés ne peuvent pas être coplanaires. En changeant de point de* $o(3)$*, on voit que* $o(2)$ *sont les sommets d'un cube dont les faces sont dans la direction des points de* $o(3)$ *et les arêtes de* $o(1)$*.*
Or G *opère sur les paires de sommets opposés, donc on a un morphisme* $\varphi : \mathrm{G} \to \mathfrak{S}_4$*. Si* $\varphi(g)$ *est l'identité, et* g *échange deux sommets, alors, puisqu'il n'y a que deux points fixes,* g *échange au moins deux autres paires. Par suite,* $g = -\mathrm{Id}$*, mais* g *est une rotation, donc c'est impossible, et* φ *est injective. Par dualité, l'envoloppe convexe de* $o(3)$ *est un octaèdre de faces centrées sur* $o(2)$*.* ∎

On remarque que le cube contient deux tétraèdres "opposés", en considérant pour arêtes des diagonales des faces. Le groupe préserve ces tétraèdres ou les échange. On peut ainsi en déduire que $\mathfrak{a}_4$ est un sous-groupe d'indice 2 de G, donc distingué.

Proposition 15.1.12.6. *Si* $\mathrm{N} = 60$, $r_j = (2, 3, 5)$, *et* $n_j = (30, 20, 12)$ *alors* G *est le groupe d'isométries d'un dodécaèdre et d'un icosaèdre, et est isomorphe à* $\mathfrak{a}_5$.

Démonstration. *Les trois orbites s'organisent par paires.*

Le stabilisateur d'un point $x \in o(3)$ *opère sur* $o(2)$, *en quatre orbites de cinq éléments. Les points de* $o(2)$ *les plus proches de* x *forment un pentagone (on ne peut avoir deux orbites sur un même plan en considérant un autre point de* $o(3)$*). En opérant sur toute l'orbite de* x, *on devine un dodécaèdre à* 20 *sommets et* 30 *arêtes. Par dualité, on obtient un icosaèdre.*

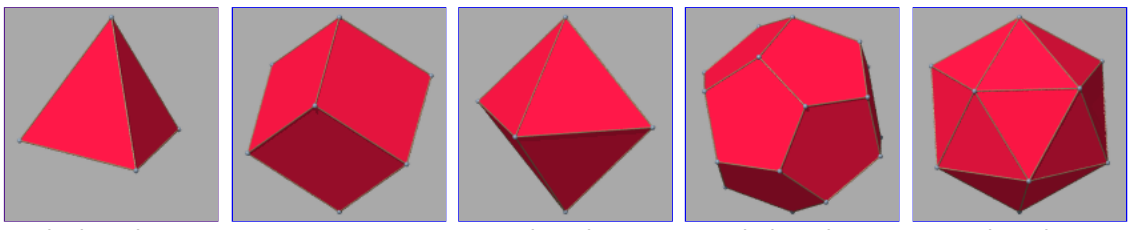
On se fixe un sommet $x \in o(2)$. *Il correspond à un sommet commun de trois pentagones. Le stabilisateur de* x *les permute, ainsi que les arêtes communes. Il opère aussi sur les autres sommets de ces pentagones, en deux orbites. Les trois segments qui joignent* x *à une de ces orbites se coupent à angle droit.*

On devine ainsi un cube, si on considère $(-x)$. *Les centres des faces correspondent à des points de* $o(1)$, *soit des centres des arêtes. On peut voir que chaque arête est de même longueur, et que l'on a bien* 8 *sommets.*

A chaque paire de sommets correspond deux cubes, et chaque cube admet 8 *sommets, soit* 4 *diagonales. Du coup, on obtient* $10 \times 2/4 = 5$ *cubes. De manière équivalente, chaque arête correspond à une face de cube, qui en comporte* 6. *Ce qui nous fait* $30/6 = 5$ *cubes.*

Notre groupe G *opère sur ces cubes par permutation. Supposons qu'un élément* $g \in \mathrm{G}$ *fixe tous ces cubes globalement. S'il est non trivial, il fixe donc son axe de rotation. Si son ordre est* 2, *il fixe une arête du dodécaèdre, donc une face de cube. Il ne peut fixer les autres cubes alors. Si son ordre est* 3, *alors son axe passe par les sommets de deux cubes. Les cubes induits par les sommets contigüs doivent aussi être préservés par* g, *ce qui est impossible. L'ordre ne peut être* 5 *puisqu'aucun élément qui préserve un cube n'est d'ordre multiple de* 5.

Donc $g = \mathrm{Id}$. *Notre groupe est donc un sous-groupe de* $\mathfrak{S}_5$ *d'indice* 2 *(pour une question d'ordre), il s'agit de* $\mathfrak{a}_5$. ∎



solides platoniciens

Sous-groupes distingués. Si G est un sous-groupe fini de $\mathrm{SO}_3(\mathbf{R})$, et si H est un sous-groupe de G, alors $\mathrm{X}_\mathrm{H} \subset \mathrm{X}_\mathrm{G}$. Si $x \in \mathrm{X}_\mathrm{H}$, $g \in \mathrm{G}$, alors il existe $h \in \mathrm{H}$ tel que $h(x) = x$, et $ghg^{-1}(gx) = gx$. Donc l'action de G induite par automorphismes intérieurs sur H est en correspondance avec l'action de G sur les images de X_H.

Si de plus H est distingué, alors X_H est préservé. Donc G opère sur X_H, et chaque point de X_H a une orbite par H qui est une sous-orbite pour G. Autrement dit, $o_H(x)$ divise $o_G(x)$, et X_H est une réunion d'orbites de G. Du coup, les orbites dans X_H de G se décomposent en orbites de H. On montre ainsi facilement que $\mathfrak{a}_5$ est simple.

Remarque(s) 15.1.12.7. *Un polytope convexe est une intersection de demi-espaces affines d'intérieur non vide. Si* P *est un polytope convexe, on peut supposer que l'origine est dans son intérieur. Si on projette le bord de* P *sur* $\mathbf{S}^2$*, alors on obtient une triangulation de la sphère en sommets, arêtes et faces. Si on note* s *le nombre de sommets,* a *le nombre d'arêtes et* f *le nombre de faces, alors* $s - a + f = 2$*. En effet, chaque fois que l'on supprime un sommet avec les arêtes qui le contient, on enlève autant de faces que d'arêtes, sauf que le sommet se transforme en face : $s - a + f$ reste constant lorsque l'on diminue le nombre de sommets. Lorsqu'il ne reste plus que* 4 *sommets, alors on peut vérifier la formule.*

15.1.13 Appendice : Pinceaux quadratiques

La classification des formes quadratiques permet d'étudier les quadriques comme on l'a vu dans le cas réel (ou dans le cas complexe -cf. TD-). Elle est le premier pas pour aborder leur arithmétique. On se propose ici de s'intéresser aux intersections de deux quadriques, l'une étant non dégénérée (hypothèse assez faible en fait -pourquoi ?-). Pour une autre preuve, voir 12..

Soit donc q_1, q_2 deux formes quadratiques sur un $\mathbf{k}$-espace vectoriel V de dimension finie n avec q_1 non dégénérée. On note b_1, b_2 les formes bilinéaires associées et $S_\ell = (b_\ell(e_i, e_j))_{i,j}, \ell = 1, 2$ les matrices dans une base $\mathcal{B} = (e_i)$ choisie arbitrairement. On a donc $S_1 \in GL_n(\mathbf{k})$.

Si $k = \mathbf{R}$ et que q_1 est définie, le théorème de réduction 15.1.6.1 assure qu'on peut trouver une base telle que la matrice de q_1 soit l'identité et celle de q_2 soit diagonale, de sorte qu'on se ramène à un « pinceau diagonal ») :

$$q_1(x) = \sum x_i^2 \text{ et } q_2(x) = \sum \lambda_i x_i^2.$$

Qu'en est-il dans le cas général ? D'une manière générale, on parlera d'un pinceau diagonal pour

$$q_1(x) = \sum \mu_i x_i^2 \text{ et } q_2(x) = \sum \lambda_i x_i^2.$$

Ainsi (q_1, q_2) est un pinceau diagonal dans une base convenable si et seulement si il existe une base de co-orthogonalisation.

Théorème 15.1.13.1. *Avec les notations précédentes, q_1, q_2 est un pinceau diagonal dans une base convenable si et seulement si $\mathrm{S}_1^{-1}\mathrm{S}_2$ est diagonalisable sur $\mathbf{k}$. C'est en particulier le cas si*

$$\mathrm{Card}\{\lambda \in \mathbf{k} | q_2 + tq_1 \ \textit{dégénérée})\} = n.$$

Comme on l'a vu, en général, $\mathrm{S}_1^{-1}\mathrm{S}_2$ n'est pas diagonalisable, même avec $\mathrm{S}_1 = \mathrm{Id}$ (cf. 15.1.6.2). Notons également que si $\mathbf{k}$ est algébriquement clos, on peut prendre $\mu_i = 1$ (changer e_i en $\mathrm{e}_i/\sqrt{\mu_i}$), les valeurs propres de $\mathrm{S}_1^{-1}\mathrm{S}_2$ sont les quotients λ_i/μ_i.

Démonstration. *Soit u l'endomorphisme défini par $\mathrm{S}_1^{-1}\mathrm{S}_2$: il est autoadjoint pour q_1 (cf. 11.3.4.2). Supposons u diagonalisable. On procède par récurrence sur n, le cas $n = 1$ étant tautologique. Supposons $n > 1$ et la proposition vraie en dimension $< n$.*

Si u est une homothétie $\mu\,\mathrm{Id}$, alors $q_2 = \mu q_1$ et (q_1, q_2) est un pinceau diagonal dans n'importe quelle base orthogonale de q_1

Supposons donc que u n'est pas une homothétie. Les espaces propres de u sont deux à deux orthogonaux pour q_1 : si $x \in \mathrm{Ker}(u - \lambda\,\mathrm{Id})$, $x' \in \mathrm{Ker}(y - \lambda'\,\mathrm{Id})$ avec $\lambda \neq \lambda'$, on a

$$\lambda'(x, x')_1 = (x, u(x'))_1 = (u^*(x), x')_1 = (u(x), x')_1 = \lambda(x, x')_1$$

et donc $(x, x')_1 = 0$. En prenant pour chaque espace propre une base q_1-orthogonale, on obtient donc une base q_1-orthogonale $\overline{\mathcal{B}} = (\overline{\mathrm{e}}_i)$ de vecteurs propres de u. Soit P la matrice de passage de $\mathcal{B}$ à $\overline{\mathcal{B}}$ à (les colonnes de P sont les coordonnées des vecteurs de $\overline{\mathcal{B}}$ relativement à $\mathcal{B}$). On a

$$\mathrm{Mat}_{\overline{\mathcal{B}}}(u) = \mathrm{P}^{-1}\mathrm{S}_1^{-1}\mathrm{S}_2\mathrm{P} = (\mathrm{P}^{-1}\mathrm{S}_1^{-1}{}^t\mathrm{P}^{-1})({}^t\mathrm{P}\mathrm{S}_2\mathrm{P}) = ({}^t\mathrm{P}\mathrm{S}_1\mathrm{P})^{-1}({}^t\mathrm{P}\mathrm{S}_2\mathrm{P}) = \overline{\mathrm{S}}_1^{-1}\overline{\mathrm{S}}_2.$$

avec $\overline{\mathrm{S}}_\ell$ matrices de q_ℓ dans $\mathcal{B}'$. Par construction, tant $\overline{\mathrm{S}}_1$ que $\mathrm{Mat}_{\overline{\mathcal{B}}}(u) = \overline{\mathrm{S}}_1^{-1}\overline{\mathrm{S}}_2$ sont diagonales et donc il en de même de $\overline{\mathrm{S}}_2$ qui est leur produit.

La réciproque est claire car si le pinceau est diagonal en considérant $\overline{\mathcal{B}}$ une base $\mathcal{B}$ de diagonalisation, tant $\overline{\mathrm{S}}_1$ que $\overline{\mathrm{S}}_2$ sont diagonalise et donc de même pour $\overline{\mathrm{S}}_1^{-1}\overline{\mathrm{S}}_2$. qui est semblable à $\mathrm{S}_1^{-1}\mathrm{S}_2$ comme on vient de le voir.

Le dernier point signifie que u admet n valeurs propres distinctes : u est bien diagonalisable. ∎

15.1.14 Appendice : Extrema locaux et position d'une hypersurface par rapport à son plan tangent

Si $f : \mathbf{R}^n \to \mathbf{R}$ est une application de classe $\mathcal{C}^2$, alors le théorème de Schwarz implique que la matrice des dérivées partielles secondes est symétrique. Cela nous permet d'appliquer ce qui précède au calcul des variations.

Lemme de Morse. — *Soit $f : \mathbf{R}^n \to \mathbf{R}$ une application de classe $\mathcal{C}^k$, $k \geq 2$ telle que $f(0) = \mathrm{D}_0 f = 0$ et telle que $\mathrm{D}_0^2 f$ soit inversible. Alors il existe un voisinage V de l'origine et des applications de classe*

$\mathcal{C}^{k-2}$ $y_1, \dots, y_r$ *et* $z_1, \dots, z_s$ *définies sur* V *telles que* $r+s=n$ *et, pour* $x \in \mathrm{V}$, *on ait*

$$f(x) = \sum_{1 \leq j \leq r} y_j^2 - \sum_{1 \leq j \leq s} z_j^2 \,.$$

Démonstration. — On considère le développement avec reste intégrale de f au voisinage de l'origine. On a

$$f(x) = \int_0^1 (1-t)\mathrm{D}^2_{tx} f(x,x)dt = {}^t\mathrm{X} \cdot \int_0^1 (1-t)\mathrm{D}^2_{tx} f dt \cdot \mathrm{X} \,.$$

On note

$$\mathrm{A}(x) = \left(\int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(tx)dt \right)_{i,j} \qquad \text{et} \qquad \mathrm{A}_0 = \mathrm{A}(0) \,.$$

On utilise alors le lemme suivant.

Lemme 15.1.14.1. *Si* $\mathrm{A}_0 \in \mathcal{S}_n(\mathbf{R}) \cap \mathrm{GL}_n(\mathbf{R})$, *il existe un voisinage* U *de* A_0 *dans* $\mathcal{S}_n(\mathbf{R})$ *et une application infiniement différentiable* $\psi : \mathrm{U} \to \mathrm{M}_n(\mathbf{R})$ *tels que* $\psi(\mathrm{A}_0) = \mathrm{I}$ *et, pour tout* $\mathrm{A} \in \mathrm{U}$,

$${}^t\psi(\mathrm{A}) \cdot \mathrm{A}_0 \cdot \psi(\mathrm{A}) = \mathrm{A} \,.$$

Du coup, si x est assez proche de l'origine, alors $f(x) = {}^t x {}^t\psi(\mathrm{A}(x))\mathrm{A}_0\psi(\mathrm{A}(x))x$. On pose $\psi_1(x) = \psi(\mathrm{A}(x))x$, et on obtient

$$f(x) = {}^t\psi_1(x)\mathrm{A}_0\psi_1(x) \,.$$

Or il existe une base P dans laquelle A_0 soit diagonale avec r valeurs sur la diagonale égales à 1 et s égales à -1. On note J cette matrice et on pose $\psi_2 = \mathrm{P} \cdot \psi_1$. Il vient $f(x) = {}^t\psi_2(x)\mathrm{J}\psi_2(x)$. Si on appelle $y_1, \dots, y_r$ et $z_1, \dots, z_s$ les coordonnées de ψ_2, on obtient la forme recherchée.
Ceci établit le lemme de Morse modulo le Lemme 15.1.14.1. ∎

Démonstration du Lemme 15.1.14.1. — On considère l'application $h : \mathrm{M}_n(\mathbf{R}) \to \mathcal{S}_n(\mathbf{R})$ définie par $h(\mathrm{M}) = {}^t\mathrm{M}\mathrm{A}_0\mathrm{M}$. On calcule la différentielle à l'identité de h.

$$h(\mathrm{I}+\mathrm{M}) = {}^t(\mathrm{I}+\mathrm{M})\mathrm{A}_0(\mathrm{I}+\mathrm{M}) = \mathrm{A}_0 + ({}^t\mathrm{M}\mathrm{A}_0 + \mathrm{A}_0\mathrm{M}) + {}^t\mathrm{M}\mathrm{A}_0\mathrm{M} = \mathrm{A}_0 + ({}^t\mathrm{M}\mathrm{A}_0 + \mathrm{A}_0\mathrm{M}) + \mathrm{O}(\|\mathrm{M}\|^2) \,.$$

Donc $\mathrm{D}_\mathrm{I}h(\mathrm{M}) = {}^t\mathrm{M}\mathrm{A}_0 + \mathrm{A}_0\mathrm{M}$. Cette application n'est pas inversible puisque $\dim \mathcal{S}_n(\mathbf{R}) < \dim \mathrm{M}_n(\mathbf{R})$. En revanche, le noyau consiste en les matrices M telles que $\mathrm{A}_0\mathrm{M}$ soit antisymétrique. On considère l'espace E des matrices M telles que $\mathrm{A}_0\mathrm{M}$ soit symétrique. Cet espace est supplémentaire à $\operatorname{Ker} \mathrm{D}_\mathrm{I}h$, et la restriction de $\mathrm{D}_\mathrm{I}h$ à E devient maintenant inversible (injective car A_0 est inversible et espaces source et but de même dimension).
Le théorème d'inversion locale appliqué à $h|_\mathrm{E}$ montre qu'il existe des voisinages U de A_0 et V de I et un difféomorphisme infiniment différentiable $\psi : \mathrm{U} \to \mathrm{V}$ qui inverse $h|_\mathrm{E}$. ∎

Corollaire 15.1.14.2. *Sous ces hypothèses* 0 *est un maximum local strict de* f *si et seulement si* $\mathrm{D}^2 f$ *est définie négative, et est un minimum local strict de* f *si et seulement si* $\mathrm{D}^2 f$ *est définie positive.*

Corollaire 15.1.14.3. *On étudie dans* $\mathbf{R}^{n+1}$ *l'hypersurface* $\mathcal{S}$ *définie par* $x_{n+1} = \mathrm{F}(x_1, \ldots, x_n)$ *où* F *est de classe* $\mathcal{C}^k(\mathbf{R}^n)$, $k \geq 2$, *et* $\mathrm{D}^2\mathrm{F}$ *est non dégénérée. Le plan tangent au point* $p = (x_0, \mathrm{F}(x_0))$ *sépare localement* $\mathcal{S}$ *d'un demi-espace si et seulement si* $\mathrm{D}^2_{x_0}\mathrm{F}$ *est définie.*

Démonstration. *On considère l'application* $g(x) = \mathrm{F}(x) - (\mathrm{F}(x_0) + \mathrm{D}_{x_0}\mathrm{F}(x - x_0))$. *Cette application vérifie les hypothèses du lemme de Morse. On en déduit que*

$$\mathrm{F}(x) = \mathrm{F}(x_0) + \mathrm{D}_{x_0}\mathrm{F}(x - x_0) + \sum_{1 \leq j \leq r} y_j^2 - \sum_{1 \leq j \leq s} z_j^2 \,.$$

■

Chapitre 16

Géométrie hermitienne complexe

16.1 Généralités

Définition 16.1.0.1. *Un espace hermitien complexe est un espace vectoriel de dimension finie sur* **C** *muni d'une forme quadratique hermitienne q définie positive.*

Théorème 16.1.0.2. *La partie réelle d'un produit scalaire hermitien est une produit scalaire sur le* **R***-espace vectoriel sous-jacent. C'est en particulier un espace vectoriel normé.*

C'est clair. On a donc l'inégalité de Cacuchy-Scwarz complexe :

Inégalité de Cauchy-Schwarz. — *Soit (E, q) un* **C***-espace vectoriel de dimension finie muni d'une forme hermitienne positive (pour tout $x \in \mathrm{E}$, on a $q(x) \geq 0$). On a, pour tout $x, y \in \mathrm{E}$,*

$$(\operatorname{Re} b(x,y)) \leq \sqrt{q(x)}\sqrt{q(y)}$$

avec égalité si et seulement si x, y sont positivement liés.

Définition 16.1.0.3. *Une base orthonormée est une base $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_n)$ telle que $b(\mathrm{e}_i, \mathrm{e}_j) = \delta_{i,j}$, où $\delta_{i,j}$ est le symbole de Kronecker.*

Théorème 16.1.0.4. *Dans un espace hermitien, il existe toujours une base orthonormée.*

Dans une telle base, la matrice de q est l'identité. Du coup, les éléments unitaires sont représentés par des matrices M telles que ${}^t\overline{\mathrm{M}} \cdot \mathrm{M} = \mathrm{I}$. Dans cette situation, on appelle la forme polaire *un produit scalaire hermitien* et on le note communément $\langle \cdot, \cdot \rangle$. Dans une base orthonormée, on a $\langle \mathrm{X}, y \rangle = {}^t\mathrm{X} \cdot \overline{\mathrm{Y}}$.

Démonstration. *C'est un corollaire de l'énoncé euclidien.*

L'orthogonalisation de Gram-Schmidt implique aussi que, pour tout sous-espace F de E, il existe une base orthonormée dont les premiers $\dim \mathrm{F}$ vecteurs forment une base de F.

Proposition 16.1.0.5. *Soit $b : \mathrm{E}^2 \to \mathbf{C}$ une forme linéaire en x et anti-linéaire en y sur un espace hermitien. Il existe des endomorphismes u et v tels que, pour tout $x, y \in \mathrm{E}$, on ait*

$$b(x,y) = \langle x, u(y)\rangle = \langle v(x), y\rangle .$$

Ces endomorphismes sont uniques.

Démonstration. *Soit $\mathcal{B} = (\mathrm{e}_1, \dots, \mathrm{e}_n)$ une base orthonormée ; on note $\mathrm{M} = (b(\mathrm{e}_i, \mathrm{e}_j))_{i,j}$ la matrice de b. On note* X, Y *les vecteurs coordonnées de $x, y \in \mathrm{E}$. On a*

$$b(x,y) = \sum_{i,j} x_i \overline{y_j} b(\mathrm{e}_i, \mathrm{e}_j) = {}^t\mathrm{X} \cdot \mathrm{M}\overline{\mathrm{Y}} = \langle {}^t\mathrm{M}\mathrm{X}, \mathrm{Y}\rangle = \langle \mathrm{X}, \overline{\mathrm{M}}\mathrm{Y}\rangle .$$

On définit donc, pour tout $j \in \{1, \dots, n\}$,

$$\begin{cases} u(\mathrm{e}_j) = \displaystyle\sum_{i=1}^{n} b(\mathrm{e}_i, \mathrm{e}_j)\mathrm{e}_i , \\ \\ v(\mathrm{e}_i) = \displaystyle\sum_{j=1}^{n} b(\mathrm{e}_i, \mathrm{e}_j)\mathrm{e}_j . \end{cases}$$

On a bien

$$\begin{cases} \langle \mathrm{e}_i, u(\mathrm{e}_j)\rangle = \displaystyle\sum_{k=1}^{n} b(\mathrm{e}_k, \mathrm{e}_j)\langle \mathrm{e}_k, \mathrm{e}_i\rangle = b(\mathrm{e}_i, \mathrm{e}_j) , \\ \\ \langle v(\mathrm{e}_i), \mathrm{e}_j\rangle = \displaystyle\sum_{k=1}^{n} b(\mathrm{e}_i, \mathrm{e}_k)\langle \mathrm{e}_j, \mathrm{e}_k\rangle = b(\mathrm{e}_i, \mathrm{e}_j) . \end{cases}$$

L'unicité provient aussi de ces relations. ∎

Définition 16.1.0.6. *Si $u \in \mathrm{End}(\mathrm{E})$, on appelle* adjoint de u, *que l'on note u^*, l'endomorphisme tel que, pour tout $x, y \in \mathrm{E}$,*

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle .$$

L'adjoint existe toujours d'après la proposition précédente puisque $b : (x,y) \mapsto \langle x, u(y)\rangle$ est sesquilinéaire.
De plus, $u^{**} = u$.
Si M est la matrice de u dans une base orthonormée, alors

$$\langle x, u^*(y)\rangle = \langle u(x), y\rangle = {}^t(\mathrm{M} \cdot \mathrm{X}) \cdot \overline{\mathrm{Y}} = {}^t\mathrm{X} \cdot {}^t\mathrm{M} \cdot \overline{\mathrm{Y}}$$

donc la matrice de u^* est ${}^t\overline{\mathrm{M}}$.

On dit que u est hermitien si $u = u^*$, antihermitien si $u = -u^*$ et unitaire si $uu^* = \mathrm{Id}$.

Remarque(s) 16.1.0.7. *Les polynômes caractéristiques de u et u^* sont conjugués.*

16.2 Endomorphismes normaux complexes

Rappelons qu'un endomorphisme est normal s'il commute avec son adjoint. Le résultat principal de ce paragraphe est le suivant. Les résultats et les preuves sont en tout point analogues au cas réel, en plus simple car on a toujour ds droites stables.

Théorème 16.2.0.1. *Pour tout endomorphisme normal u d'un espace hermitien, il existe une base orthonormée de E qui diagonalise u. Si u est hermitien, alors les valeurs propres sont réelles, si u est antihermitien, elles sont imaginaires pures et si u est unitaire, alors elles sont de module 1.*

Proposition 16.2.0.2. *Soit u un endomorphisme normal.*

1. *On a l'égalité $\operatorname{Ker} u = \operatorname{Ker} u^*$.*
2. *u et u^* ont leurs valeurs propres conjuguées et partagent les mêmes espaces propres. Ceux-ci sont orthogonaux deux à deux.*
3. *u fixe au moins une droite.*
4. *Si F est stable par u, alors F est aussi stable par u^*, et $\mathrm{F}^\perp$ est stable par u et u^*.*

Démonstration. 1. *Si u est normal, alors, pour tout $x \in \mathrm{E}$, on a $\|u(x)\| = \|u^*(x)\|$. En effet,*

$$\|u(x)\|^2 = \langle u(x), u(x)\rangle = \langle x, u^*(u(x))\rangle = \langle x, u(u^*(x))\rangle = \langle u^*(x), u^*(x)\rangle = \|u^*(x)\|^2 .$$

Donc $u(x)$ et $u^*(x)$ s'annulent simultanément.

2. *provient du fait que $u+\lambda\mathrm{I}$ est normal si u est normal, donc $\operatorname{Ker}(u+\lambda\,\mathrm{Id}) = \operatorname{Ker}(u^*+\overline{\lambda}\,\mathrm{Id})$ d'après ci-dessus. Soient λ, μ deux valeurs propres distinctes, et x, y deux vecteurs propres associés : $u(x) = \lambda x$ et $u(y) = \mu y$. Il vient*

$$\lambda\langle x, y\rangle = \langle u(x), y\rangle = \langle x, u^*(y)\rangle = \mu\langle x, y\rangle .$$

Du coup $(\lambda - \mu)\langle x, y\rangle = 0$ et $\langle x, y\rangle = 0$.

3. *Puisque $\mathbf{C}$ est algébriquement clos, il existe toujours une valeur propre et un vecteur propre associé.*

4. *On considère une base orthonormée de* E *telle que les premiers* $p = \dim \mathrm{F}$ *vecteurs forment une base de* F *et les derniers une base de* $\mathrm{F}^\perp$. *Dire que* F *est stable par* u *signifie qu'il existe des matrices* $\mathrm{A} \in \mathrm{M}_p(\mathbf{C})$, $\mathrm{B} \in \mathrm{M}_{p,n-p}(\mathbf{C})$ *et* $\mathrm{C} \in \mathrm{M}_{n-p}(\mathbf{C})$ *telles que la matrice de* u *s'écrive*

$$\mathrm{Mat}(u) = \begin{pmatrix} \mathrm{A} & \mathrm{B} \\ 0 & \mathrm{C} \end{pmatrix}$$

Dire que u *est normal signifie que* $\mathrm{A} \cdot {}^t\overline{\mathrm{A}} + \mathrm{B} \cdot {}^t\overline{\mathrm{B}} = {}^t\overline{\mathrm{A}} \cdot \mathrm{A}$. *Or, un simple calcul montre que* $\mathrm{tr}(\mathrm{A} \cdot {}^t\overline{\mathrm{A}}) = \mathrm{tr}({}^t\overline{\mathrm{A}} \cdot \mathrm{A})$, *donc* $\mathrm{tr}(\mathrm{B} \cdot {}^t\overline{\mathrm{B}}) = 0$. *Mais, si* $\mathrm{B} = (b_{ij})_{1 \leq i \leq p, 1 \leq j \leq n-p}$, *alors*

$$\mathrm{tr}(\mathrm{B} \cdot {}^t\overline{\mathrm{B}}) = \sum_{j=1}^{p} \left(\sum_{k=1}^{n-p} b_{jk} \overline{b_{jk}} \right) = \sum_{j=1}^{p} \sum_{k=1}^{n-p} |b_{jk}|^2 .$$

Donc $b_{jk} = 0$ *pour tout* j, k *et* $\mathrm{B} = 0$. *Du coup,*

$$\mathrm{Mat}(u) = \begin{pmatrix} \mathrm{A} & 0 \\ 0 & \mathrm{C} \end{pmatrix}$$

avec A *et* C *normaux. Ceci montre que* F *et* $\mathrm{F}^\perp$ *sont stables par* u *et* u^*.

■

Démonstration du Théorème de réduction. — On procède par récurrence sur la dimension de E. Pour $n = 1$, il n'y a rien à dire. Supposons que ce soit vrai pour tout espace de dimension n, et que E est de dimension $n+1$. D'après la proposition précédente, il existe une valeur propre λ, et un vecteur propre (normé) e_{n+1}. On note $\mathrm{F} = \mathbf{C}\mathrm{e}_{n+1}$. Cet espace est stable par u donc $\mathrm{F}^\perp$ aussi, qui est un supplémentaire de F. L'hypothèse de récurrence s'applique à $u|_{\mathrm{F}^\perp}$. Cette base se complète en une base orthonormée de E, et la matrice de u dans cette base est diagonale. ■

Corollaire 16.2.0.3. *L'application* $\exp : \mathcal{H}_n(\mathbf{C}) \to \mathcal{H}_n^{++}(\mathbf{C})$ *est un homéomorphisme.*

Démonstration. *L'application est* exp *est continue, et si* $\mathrm{M} \in \mathcal{H}_n(\mathbf{C})$, *alors il existe* $\mathrm{U} \in \mathrm{U}(n)$ *telle que* $\mathrm{U}^{-1}\mathrm{MU}$ *soit diagonale. On en déduit que* $\exp \mathrm{M} = \mathrm{U}(\exp \mathrm{U}^{-1}\mathrm{MU})\mathrm{U}^{-1} \in \mathcal{H}_n^{++}(\mathbf{C})$. *Réciproquement, si* $\mathrm{M} \in\in \mathcal{H}_n^{++}(\mathbf{C})$, *il existe* $\mathrm{U} \in \mathrm{U}(n)$ *telle que* $\mathrm{U}^{-1}\mathrm{MU}$ *soit diagonale avec des valeurs propres strictement positives. On peut alors considérer la matrice* N *diagonale formée des logarithmes des valeurs propres de* M. *On a* $\mathrm{UNU}^{-1} \in \mathcal{H}_n(\mathbf{C})$ *et* $\exp(\mathrm{UNU}^{-1}) = \mathrm{M}$.

Il reste à voir que l'application est injective. On suppose donc que $\exp \mathrm{M} = \exp \mathrm{N}$. *Quitte à changer de base, on peut supposer que* M *est diagonale. Par suite,* $\exp \mathrm{M}$ *est diagonale aussi. Ceci montre que les valeurs propres de* M *et* $\exp \mathrm{M}$ *sont liées, ainsi que les espaces propres associés. Il en est donc de même pour* N. *On en déduit que* $\mathrm{M} = \mathrm{N}$.

Enfin, pour voir que l'application réciproque est continue, on observe que $\mathcal{H}_n(\mathbf{C})$ *admet une exhaustion par des compacts en considérant celles dont le spectre est contenu dans un intervalle compact. Du coup, chaque restriction est un homéomorphisme sur son image.* ■

16.3 Le groupe unitaire

On suppose que E est un espace hermitien muni d'une base orthonormée $\mathcal{B}$.

Théorème 16.3.0.1. *Soit u un endomorphisme de* E. *Les propriétés suivantes sont équivalentes.*

- $u \in \mathrm{U(E)}$;
- *pour tout $x \in \mathrm{E}$, on a $\|u(x)\| = \|x\|$;*
- $u \circ u^* = \mathrm{Id}$;
- $u^* \circ u = \mathrm{Id}$;
- $\mathrm{Mat}(u, \mathcal{B}) \cdot {}^t\overline{\mathrm{Mat}(u, \mathcal{B})} = \mathrm{I}$;
- ${}^t\overline{\mathrm{Mat}(u, \mathcal{B})} \cdot \mathrm{Mat}(u, \mathcal{B}) = \mathrm{I}$;
- *les colonnes de* $\mathrm{Mat}(u, \mathcal{B})$ *forment une base orthonormée ;*
- *les lignes de* $\mathrm{Mat}(u, \mathcal{B})$ *forment une base orthonormée ;*
- *u transforme une base orthonormée en une base orthonormée.*

La démonstration est laissée en exercice.

Corollaire 16.3.0.2. *Si $u \in \mathrm{U(E)}$, alors $|\det u| = 1$ et donc $\mathrm{SU(E)} = \det^{-1}\{1\} \cap \mathrm{U(E)}$ est distingué.*

Propriétés topologiques. On s'intéresse aux propriétés topologiques du groupe unitaire.

Lemme 16.3.0.3. $\mathrm{U}_n(\mathbf{C})$ *est compact.*

Démonstration. *On considère l'application $f : \mathrm{M} \in \mathrm{M}_n(\mathbf{C}) \mapsto {}^t\overline{\mathrm{M}}\mathrm{M}$. On a $\mathrm{U}_n(\mathbf{C}) = f^{-1}(\mathrm{I}_n)$, donc $\mathrm{U}_n(\mathbf{C})$ est fermé. Pour montrer que $\mathrm{U}(n)$ est borné, on considère la norme induite par $\mathrm{tr}\,{}^t\overline{\mathrm{M}}\mathrm{M}$. On a $\mathrm{U}_n(\mathbf{C}) \subset \mathrm{B}(0, \sqrt{n})$.* ∎

Proposition 16.3.0.4. *Les groupes* U(E) *et* SU(E) *sont connexes.*

Démonstration. *On montre d'abord que* U(E) *est connexe par arcs. Le þde réduction nous permet de diagonaliser en base orthonormée tout éléments de* U(E) *: étant donné $\mathrm{U} \in \mathrm{U(E)}$, il existe $\widehat{\mathrm{U}} \in \mathrm{U(E)}$ et $\theta_1, \ldots, \theta_n \in \mathbf{R}$ tels que*

$$\mathrm{U} = \widehat{\mathrm{U}} \begin{pmatrix} \mathrm{e}^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{i\theta_n} \end{pmatrix} \widehat{\mathrm{U}}^{-1}.$$

On note, pour $s \in [0,1]$,

$$\mathrm{U}_s = \widehat{\mathrm{U}} \begin{pmatrix} \mathrm{e}^{is\theta_1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{is\theta_n} \end{pmatrix} \widehat{\mathrm{U}}^{-1}\,.$$

Pour SU(E)*, il faut s'y prendre un peu différemment. On sait que $\sum \theta_j \in 2\pi\mathbf{Z}$. On fait une récurrence sur la dimension pour ramener chaque valeur propre à 1 sans changer le déterminant. Pour cela, on définit $\theta_1(s) = \theta_1 - s$, $\theta_2(s) = \theta_1 + s$ et $\theta_j(s) = \theta_j$, pour $s \in [0, \theta_1]$.* ∎

Décomposition polaire. L'application $\Phi : \mathrm{U}(n) \times \mathcal{H}_n^{++}(\mathbf{R}) \to \mathrm{GL}_n(\mathbf{C})$ définie par $\Phi(\mathrm{U}, \mathrm{H}) = \mathrm{UH}$ est un homéomorphisme.

Lemme 16.3.0.5. *Si $\mathrm{M} \in \mathrm{GL}_n(\mathbf{C})$, alors ${}^t\overline{\mathrm{M}} \cdot \mathrm{M}$ est définie positive.*

Démonstration. *On considère une base orthonormée de $\mathbf{C}^n$ et on considère $\overline{\mathrm{M}}$ comme la matrice d'un endomorphisme u (inversible) dans cette base. On a, pour $x \neq 0$,*

$$\langle x, u^* \circ u(x)\rangle = \langle u(x), u(x)\rangle > 0\,..$$

∎

Démonstration de la décomposition : existence. — D'après le lemme, il existe $\widehat{\mathrm{U}} \in \mathrm{U}(n)$ telle que

$$\widehat{\mathrm{U}}^{-1} \cdot ({}^t\overline{\mathrm{M}} \cdot \mathrm{M}) \cdot \widehat{\mathrm{U}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

soit diagonale avec des valeurs propres strictement positives. On note

$$\mathrm{H} = \widehat{\mathrm{U}} \cdot \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} \cdot \widehat{\mathrm{U}}^{-1}$$

et $\mathrm{U} = \mathrm{MH}^{-1}$. Comme H est hermitienne, H^{-1} aussi, donc

$${}^t\overline{\mathrm{U}} \cdot \mathrm{U} = {}^t(\overline{\mathrm{H}^{-1}}) \cdot ({}^t\overline{\mathrm{M}} \cdot \mathrm{M}) \cdot \mathrm{H}^{-1} = \mathrm{H}^{-1}\mathrm{H}^2\mathrm{H}^{-1} = \mathrm{I}\,.$$

∎

Démonstration de la décomposition : unicité. — Soit $\mathrm{P} \in \mathbf{C}[\mathrm{X}]$ tel que $\mathrm{P}(\lambda_j) = \sqrt{\lambda_j}$. On a

$$\mathrm{P}({}^t\overline{\mathrm{M}} \cdot \mathrm{M}) = \sum a_k \widehat{\mathrm{U}} \cdot \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} \cdot \widehat{\mathrm{U}}^{-1} = \widehat{\mathrm{U}} \cdot \begin{pmatrix} \mathrm{P}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & \mathrm{P}(\lambda_n) \end{pmatrix} \cdot \widehat{\mathrm{U}}^{-1} = \mathrm{H}\,.$$

Donc, si M admet une seconde décomposition polaire $M = U'H'$ alors ${}^t\overline{M} \cdot M = (H')^2$ donc H' commute avec ${}^t\overline{M} \cdot M$, donc avec H puisque H est un polynôme en ${}^t\overline{M} \cdot M$. Du coup, H et H' peuvent être diagonalisées simultanément (en effet, $E_{\lambda_i}(H) = \oplus(E_{\lambda_i}(H) \cap E_{\mu_j}(H'))$). Ceci oblige les valeurs propres à coïncider, donc ces matrices sont les mêmes. ∎

Démonstration de la décomposition : homéomorphie. — L'application Φ est clairement continue. Réciproquement, si $M_n = \Phi(U_n, H_n)$ converge vers une matrice $M = UH$, montrons que U_n tend vers U et H_n vers H. Comme U(E) est compact, quitte à extraire une sous-suite, on peut supposer que U_n tend vers une matrice unitaire $\widehat{U}$. Du coup, H_n tend vers une matrice (hermitienne) $\widehat{U}^{-1}M$. Du coup, on a $UH = \widehat{U}(\widehat{U}^{-1}M)$. Par l'unicité de la décomposition polaire, on obtient $U = \widehat{U}$ et H_n tend vers H. ∎

On en déduit quelques résultats.

Proposition 16.3.0.6. *Deux matrices réelles* A *et* B *unitairement semblables sont orthogonalement semblables.*

Démonstration. *On suppose qu'il existe* $U \in U_n(\mathbf{C})$ *telle que* $AU = UB$. *On écrit* $U = U_1 + iU_2$, *où* U_1 *et* U_2 *sont des matrices réelles. Puisque* ${}^t\overline{U} = U^{-1}$, *on a* $U^{-1} = {}^tU_1 - i{}^tU_2$.
De plus, la similitude de A *et* B *implique* $AU_1 = U_1B$ *et* $AU_2 = U_2B$. *Par suite, si* $\lambda \in \mathbf{R}$, *alors* $A(U_1 + \lambda U_2) = (U_1 + \lambda U_2)B$. *On écrit* $Q(\lambda) = \det(U_1 + \lambda U_2) \in \mathbf{R}[\lambda]$. *Comme* $Q(i) = \det U \neq 0$, *on en déduit que* Q *est un polynôme non nul, donc il existe* $\lambda \in \mathbf{R}$ *tel que* $Q(\lambda) \neq 0$ *et* $P := U_1 + \lambda U_2 \in GL_n(\mathbf{R})$.
D'autre part, ${}^tA = {}^tU^{-1}\,{}^tB\,{}^tU = \overline{U}\,{}^tB\overline{U^{-1}}$ *donc* ${}^tA = U\,{}^tBU^{-1}$ *et* ${}^tAP = P\,{}^tB$. *On a*

$$P\,{}^tBP^{-1} = A = {}^t({}^tA) = {}^tP^{-1}B\,{}^tP$$

donc ${}^tPPB = B\,{}^tPP$. *Or, la décomposition polaire nous donne* $P = OS$, *où* S *est un polynôme en* ${}^tPP$. *Donc* $BS = SB$ *et on obtient*

$$A = PBP^{-1} = OSBS^{-1}O^{-1} = OBO^{-1}\,.$$

∎

Proposition 16.3.0.7. $GL_n(\mathbf{C})$ *est connexe.*

Démonstration. *On a*

$$GL_n(\mathbf{C}) \approx U_n(\mathbf{C}) \times \mathcal{H}_n^{++}(\mathbf{C})$$

par la décomposition polaire. ∎

Remarque.— On a une démonstration plus simple : soient $P, Q \in GL_n(\mathbf{C})$, et notons $R(\lambda) = \det(P + \lambda Q)$ qui est un polynôme complexe non nul puisque $R(0) \neq 0$. Donc il existe un chemin dans $\mathbf{C} \setminus R^{-1}\{0\}$ qui relie $\lambda = 0$ à $\lambda = 1$.

Proposition 16.3.0.8. *$\mathrm{SU}_n(\mathbf{C})$ (resp. $\mathrm{U}_n(\mathbf{C})$) est un sous-groupe compact maximal de $\mathrm{SL}_n(\mathbf{C})$ (resp. $\mathrm{GL}_n(\mathbf{C})$).*

Démonstration. *Soit $\mathrm{G} \subset \mathrm{SL}_n(\mathbf{C})$ un sous-groupe compact qui contient $\mathrm{SU}_n(\mathbf{C})$, et soit $g \in \mathrm{G} \setminus \mathrm{SU}_n(\mathbf{C})$. On a $g = \mathrm{UH}$ avec $\mathrm{U} \in \mathrm{SU}_n(\mathbf{C})$ car $\det g > 0$ et $\mathrm{H} \neq \mathrm{I}$ car $g \notin \mathrm{SU}_n(\mathbf{C})$. Donc $\mathrm{H} \in \mathrm{G}$. Or si v est un vecteur propre associé à une valeur propre λ de H de moule différent de 1, alors $\mathrm{Log}\,\|\mathrm{H}^n x\|$ tend vers l'infini, ce qui contredit la compacité de G. Donc $\mathrm{H} \in \mathrm{SU}_n(\mathbf{C})$.* ∎

Remarque(s) 16.3.0.9. *Il est facile de voir que $\mathrm{O}_2(\mathbf{C})$ n'est pas compact, donc $\mathrm{O}_n(\mathbf{C})$ non plus.*

Proposition 16.3.0.10. *Les espaces $\mathrm{GL}_n(\mathbf{C})$ et $\mathrm{SL}_n(\mathbf{C})$ sont respectivement homéomorphes à $\mathrm{U}_n(\mathbf{C}) \times \mathbf{C}^{n^2}$ et $\mathrm{SU}_n(\mathbf{C}) \times \mathbf{C}^{n^2-1}$.*

Démonstration. *L'application $\exp : \mathcal{H}_n(\mathbf{C}) \to \mathcal{H}_n^{++}(\mathbf{C})$ est un homéomorphisme et $\mathcal{H}_n(\mathbf{C}) \approx \mathbf{C}^{n^2}$, donc*

$$\mathrm{GL}_n(\mathbf{C}) \approx \mathrm{U}_n(\mathbf{C}) \times \mathcal{H}_n^{++}(\mathbf{C}) \approx \mathrm{U}_n(\mathbf{C}) \times \mathbf{C}^{n^2}.$$

De même, $\mathrm{SL}_n(\mathbf{C}) \approx \mathrm{SU}_n(\mathbf{C}) \times (\mathcal{H}_n^{++}(\mathbf{C}) \cap \mathrm{SL}_n(\mathbf{C}))$ et l'application $\exp : \mathcal{H}_n(\mathbf{C}) \cap \mathrm{tr}^{-1}\{0\} \to \mathcal{H}_n^{++}(\mathbf{C}) \cap \mathrm{SL}_n(\mathbf{C})$ est un homéomorphisme. ∎

Proposition 16.3.0.11. *Une matrice $\mathrm{A} \in \mathrm{M}_n(\mathbf{C})$ appartient à $\mathrm{O}_n(\mathbf{C})$ si et seulement si il existe $\mathrm{O} \in \mathrm{O}_n(\mathbf{R})$ et $\Theta \in \mathrm{A}_n(\mathbf{R})$ telles que $\mathrm{A} = \mathrm{O} \exp i\Theta$.*

Démonstration. *Si $\mathrm{A} = \mathrm{O} \exp i\Theta$ alors*

$${}^t\mathrm{AA} = (\exp i\,{}^t\Theta)\,{}^t\mathrm{OO} \exp i\Theta = \exp i({}^t\Theta + \Theta) = \mathrm{I}_n.$$

Réciproquement, on utilise la décomposition polaire complexe : il existe $\mathrm{U} \in \mathrm{U}_n(\mathbf{C})$ et P hermitienne définie positive telle que $\mathrm{A} = \mathrm{UP}$. Or ${}^t\mathrm{AA} = \mathrm{I}$ implique que ${}^t\mathrm{P}\,{}^t\mathrm{UUP} = \mathrm{I}$, soit ${}^t\mathrm{UUP} = {}^t\mathrm{P}^{-1}$. Donc l'unicité de la décomposition nous montre que ${}^t\mathrm{UU} = \mathrm{I}$ et $\mathrm{P} = {}^t\mathrm{P}^{-1}$. La première condition nous conduit à $\mathrm{U} = \overline{\mathrm{U}}$ car U est unitaire, donc $\mathrm{U} \in \mathrm{O}_n(\mathbf{R})$. Quant à la seconde, on écrit $\mathrm{P} = \exp \mathrm{L}$, où L est hermitienne. On obtient $\mathrm{L} = -{}^t\mathrm{L}$, donc L est antisymétrique. De plus, comme L est hermitienne, on a ${}^t\mathrm{L} = \overline{\mathrm{L}} = -\mathrm{L}$ donc $\mathrm{L} = i\Theta$ avec $\Theta \in \mathrm{A}_n(\mathbf{R})$. ∎

Comme corollaire, on en déduit que $\mathrm{O}_n(\mathbf{C})$ est homéomorphe à $\mathrm{O}_n(\mathbf{R}) \times \mathbf{R}^{\frac{n(n-1)}{2}}$ et que $\mathrm{O}_n(\mathbf{C})$ a exactement deux composantes connexes.

16.4 Appendice : Le cas de $\mathrm{SU}_2(\mathbf{C})$

Le cas de $\mathrm{SU}_2(\mathbf{C})$ est crucial pour la classification des représentations groupes compacts d'une part, et en physique car ses représentations sont le cœur du spin des particules en mécanique quantique. Donnons quelques éléments sur sa géométrie.

Théorème 16.4.0.1. *On a $\mathrm{SU}_2(\mathbf{C})/\{\pm \mathrm{I}_2\}$ est isomorphe à $\mathrm{SO}_3(\mathbf{R})$.*

Lemme 16.4.0.2. *On considère le corps des quaternion s sous la forme*

$$\mathbf{H} = \left\{ \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix},\ a, b \in \mathbf{C} \right\}$$

que l'on munit de $q(\mathrm{M}) = \det(\mathrm{M})$. Alors $\mathbf{H}$ est un espace euclidien de dimension 4 et $\mathrm{SU}_2(\mathbf{C})$ s'identifie à la sphère unité de $\mathbf{H}$.

Démonstration. *Il est aisé de voir que $\mathbf{H}$ est un $\mathbf{R}$-espace vectoriel de dimension 4. Le déterminant est bien une forme quadratique par le théorème de Cayley-Hamilton : on constate que*

$$\det \mathrm{M} = \frac{1}{2}((\operatorname{tr} \mathrm{M})^2 - \operatorname{tr} \mathrm{M}^2)$$

et $\det \mathrm{M} \geq 0$ pour toute $\mathrm{M} \in \mathbf{H}$. De plus $\det \mathrm{M} = 0$ si et seulement si $\mathrm{M} = 0$. Du coup, q est positive sans vecteur isotrope. On a bien une structure euclidienne.
Si $\mathrm{M} \in \mathrm{SU}_2$, alors les vecteurs colonnes sont orthogonaux, donc $a\overline{b} + c\overline{d} = 0$. Si $d = 0$ alors $bc = 1$ et $|b| = 1$ donc $c = -\overline{b}$ et $a = 0$. Sinon, $ad\overline{b} + c|d|^2 = 0$ et $(|b|^2 + |d|^2)c + \overline{b} = 0$ car $\det \mathrm{M} = 1$, soit $c + \overline{b} = 0$ car les vecteurs colonnes sont unitaires. En reprenant la première équation, on obtient $a - \overline{d} = 0$ si $b \neq 0$. Sinon, un argument similaire montre aussi que $d = \overline{a}$. Donc $\mathrm{M} \in \mathbf{H}$. ∎

On passe maintenant à la démonstration du théorème.

Démonstration. *On note $\mathbf{H}_0 = \mathbf{H} \cap \operatorname{tr}^{-1}\{0\}$ l'ensemble des quaternions purs (on a $\mathrm{I}^\perp = \mathbf{H}_0$). Cet espace $\mathbf{H}_0$ est un $\mathbf{R}$-espace vectoriel de dimension 3. L'application*

$$\mathcal{P} : \mathrm{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} ix_1 & -x_2 + ix_3 \\ x_2 + ix_3 & -ix_1 \end{pmatrix}$$

définit d'ailleurs un isomorphisme entre $\mathbf{R}^3$ et $\mathbf{H}_0$, et on observe que

$$\det \mathcal{P}(\mathrm{X}) = \|\mathrm{X}\|$$

donc $\mathcal{P}$ réalise en fait une isométrie (euclidienne) si on munit $\mathbf{H}_0$ de la forme "$\det$".
On remarque que les éléments H de $\mathbf{H}_0$ vérifient $\mathrm{H} = -\overline{{}^t\mathrm{H}}$ et $\operatorname{tr} \mathrm{H} = 0$.

Si $\mathrm{H} \in \mathbf{H}_0$ *et si* $\mathrm{U} \in \mathrm{SU}_2(\mathbf{C})$*, alors*

$$-{}^t(\overline{\mathrm{UHU}^{-1}}) = {}^t\,\overline{\mathrm{U}^{-1}}(-{}^t\overline{\mathrm{H}}){}^t\overline{\mathrm{U}} = \mathrm{UHU}^{-1}$$

donc l'application φ_{U} *définie par conjugaison par* U *est un endomorphisme de* $\mathbf{H}_0$*.*

De plus, on a $\det\varphi_{\mathrm{U}}(\mathrm{H}) = \det\mathrm{H}$ *donc* $\varphi_{\mathrm{U}} \in \mathrm{O}(\mathbf{H}_0)$*. Or* $\varphi : \mathrm{SU}_2(\mathbf{C}) \to \mathrm{O}(\mathbf{H}_0)$ *ainsi induite est un morphisme de groupes continu, et comme* $\mathrm{SU}_2(\mathbf{C})$ *est connexe, on a* $\varphi\mathrm{SU}_2(\mathbf{C}) \subset \mathrm{SO}(\mathbf{H}_0)$*.*

Une matrice U *est dans le noyau de* φ *si, pour tout* $\mathrm{H} \in \mathbf{H}_0$*, on a* $\mathrm{UH} = \mathrm{HU}$*. En considérant les matrices*

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad et \quad \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

on trouve $\operatorname{Ker}\varphi = \{\pm\mathrm{Id}\}$*.*

Il reste à voir que φ *est surjective. Pour cela, nous allons montré que* $\operatorname{Im}\varphi$ *contient tous les demi-tours. Soit* $\mathrm{A} \in \mathbf{H}_0$ *telle que* $\det\mathrm{A} = 1$*. Par le théorème de Cayley-Hamilton, on a* $\mathrm{A}^2 = -\mathrm{I}$*. Or, puisque* $\mathrm{A} \in \mathbf{H}_0$*, on a aussi* $\mathrm{A} = -{}^t\overline{\mathrm{A}}$*. Donc* $\mathrm{A}{}^t\overline{\mathrm{A}} = -\mathrm{A}^2 = \mathrm{I}$ *et* $\mathrm{A} \in \mathrm{SU}_2(\mathbf{C})$*. Par suite,* φ_{A} *n'est pas l'identité car* $\mathrm{A} \neq \pm\mathrm{I}$*, mais* $\varphi_{\mathrm{A}} \circ \varphi_{\mathrm{A}}(\mathrm{H}) = \varphi_{\mathrm{A}^2}(\mathrm{H}) = \mathrm{H}$ *car* $\mathrm{A}^2 = -\mathrm{I}$ *et* $\varphi_{\mathrm{A}}(\mathrm{A}) = \mathrm{A}$*. Donc* φ_{A} *est le demi-tour d'axe* $\mathbf{R}\mathrm{A}$*.* ∎

Remarque(s) 16.4.0.3. *Le lecteur connaissant la notion de sous-variété pourra trouver une preuve plus naturelle en regardant la conjuqaison de* SU_2 *sur son espace tangent à l'origine. Ceci équivaut à regarder l'action par conjugaison sur l'espace des matrices antihermitiennes muni de la forme quadratique positive invariante* $q(\mathrm{M}) = -\operatorname{tr}(\mathrm{M}^2)$ *dite forme de Killing.*

Théorème 16.4.0.4. *On a*

$$\mathrm{SU}_2 \times \mathrm{SU}_2/\{\pm(\mathrm{I},\mathrm{I})\} \approx \mathrm{SO}_4\,.$$

Démonstration. *On fait agir* $\mathrm{SU}_2 \times \mathrm{SU}_2$ *sur* $\mathbf{H}$ *par* $(\mathrm{U},\mathrm{V})\cdot\mathrm{H} \mapsto \mathrm{UHV}^{-1}$*. Cette application continue envoie* $\mathrm{SU}_2 \times \mathrm{SU}_2$ *dans* O_4*, mais comme* $\mathrm{SU}_2 \times \mathrm{SU}_2$ *est connexe, on définit ainsi un morphisme* $\Phi : \mathrm{SU}_2 \times \mathrm{SU}_2 \to \mathrm{SO}_4$*.*

Si $(\mathrm{U},\mathrm{V}) \in \operatorname{Ker}\Phi$*, alors, pour tout* $\mathrm{A} \in \mathbf{H}$*, on a* $\mathrm{UAV}^{-1} = \mathrm{A}$*. En particulier, pour* $\mathrm{A} = \mathrm{I}$*, on obtient* $\mathrm{U} = \mathrm{V}$*. Ceci nous ramène à un cas connu... On en déduit que* $\operatorname{Ker}\Phi = \{\pm\mathrm{Id}\}$*.*

Montrons la surjectivité. Il suffit de montrer que tout renversement est dans l'image de Φ*. On se fixe donc un plan de* $\mathbf{H}$ *engendré par deux éléments orthogonaux et unitaires* M_1 *et* M_2*. Il existe* $\mathrm{U} \in \mathrm{SU}_2(\mathbf{C})$ *telle que* $\mathrm{UM}_1 \in \mathbf{H}_0^{\perp} = \mathbf{R}\mathrm{I}$ *(prendre* $\mathrm{U} = \mathrm{M}_1^{-1}$*, en remarquant que* $\mathrm{M}_1 \in \mathrm{SU}_2$*). Du coup,* $\mathrm{UM}_2 \in \mathbf{H}_0$ *: on s'est ramené à la situation précédente : on cherche* $(\mathrm{U},\mathrm{V}) \in \mathrm{SU}_2$ *tel que* $\Phi(\mathrm{U},\mathrm{V})$ *fixe* $\mathbf{H}_0^{\perp}$ *et un élément* $\mathrm{H} \in \mathbf{H}_0$*. Il suffit de prendre* $\mathrm{U} = \mathrm{V} = \mathrm{H}$*. Pour conclure, on conjugue cette application par* $\mathrm{M} \mapsto \mathrm{U}^{-1}\mathrm{M}$ *qui est une isométrie, donc conjugue notre renversement à celui que nous cherchions.* ∎

Exercice(s) 16.4.0.5. *Montrer que les deux isomorphismes précédents sont des homéomorphismes.*

Chapitre 17

Index et bibliographie

Index

Bibliographie

[Bev16] Anthony J. Bevelacqua. A family of non-Euclidean PIDs. *Amer. Math. Monthly*, 123(9) :936–939, 2016.

[Bis67] Errett Bishop. *Foundations of constructive analysis.* McGraw-Hill Book Co., New York-Toronto-London, 1967.

[BMS67] H. Bass, J. Milnor, and J.-P. Serre. Solution of the congruence subgroup problem for SL_n ($n \geq 3$) and Sp_{2n} ($n \geq 2$). *Inst. Hautes Études Sci. Publ. Math.*, (33) :59–137, 1967.

[Bor91] Armand Borel. *Linear algebraic groups*, volume 126 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1991.

[Bou70] N. Bourbaki. *Éléments de mathématique. Algèbre. Chapitres 1 à 3.* Hermann, Paris, 1970.

[Bou07] N. Bourbaki. *Éléments de mathématique. Algèbre. Chapitre 4 –7. Algèbre.* Springer-Verlag, Berlin, 2007.

[CMM17] Sunil K. Chebolu, Dan McQuillan, and Ján Mináč. Witt's cancellation theorem seen as a cancellation. *Expo. Math.*, 35(3) :300–314, 2017.

[DS04] Dragomir Z. Doković and Fernando Szechtman. An elementary proof of Gabriel's theorem on degenerate bilinear forms and its generalization. *J. Algebra*, 279(1) :121–125, 2004.

[DT16] Fernando De Terán. Canonical forms for congruence of matrices and T-palindromic matrix pencils : a tribute to H. W. Turnbull and A. C. Aitken. *SeMA J.*, 73(1) :7–16, 2016.

[Gab74] Peter Gabriel. Appendix : degenerate bilinear forms. *J. Algebra*, 31 :67–72, 1974.

[Ger61] Murray Gerstenhaber. On dominance and varieties of commuting matrices. *Ann. of Math. (2)*, 73 :324–348, 1961.

[Gra81] Daniel R. Grayson. Sk1 of an interesting principal ideal domain. *Journal of Pure and Applied Algebra*, 20 :157–163, 1981.

[Hil90] David Hilbert. Ueber die Theorie der algebraischen Formen. *Math. Ann.*, 36(4) :473–534, 1890.

[HP94] W. V. D. Hodge and D. Pedoe. *Methods of algebraic geometry. Vol. II.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1994. Book III : General theory of algebraic varieties in projective space, Book IV : Quadrics and Grassmann varieties, Reprint of the 1952 original.

[HS08] Roger A. Horn and Vladimir V. Sergeichuk. Canonical matrices of bilinear and sesquilinear forms. *Linear Algebra Appl.*, 428(1) :193–223, 2008.

[Ikr18] Khakim Ikramov. On the congruent selection of Jordan blocks from a singular square matrix. *Numerical Analysis and Applications*, 11 :204–207, 07 2018.

[Jac85] Nathan Jacobson. *Basic algebra. I.* W. H. Freeman and Company, New York, second edition, 1985.

[Kle74] Felix Klein. *Le programme d'Erlangen.* Collection "Discours de la Méthode". Gauthier-Villars Éditeur, Paris-Brussels-Montreal, Que., 1974. Considérations comparatives sur les recherches géométriques modernes, Traduit de l'allemand par H. Padé, Préface de J. Dieudonné, Postface de Françcois Russo.

[Mil66] J. Milnor. Whitehead torsion. *Bull. Amer. Math. Soc.*, 72 :358–426, 1966.

[Noe21] Emmy Noether. Idealtheorie in Ringbereichen. *Math. Ann.*, 83(1-2) :24–66, 1921.

[Per88] Daniel Perrin. *Cours d'algèbre.* Ellipses, Paris, 1988.

[Ros94] Jonathan Rosenberg. *Algebraic* K*-theory and its applications*, volume 147 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994.

[Ros20] Hjalmar Rosengren. Proof of the duality of the dominance order on partitions. `https://math.stackexchange.com/q/3429855`, 2020.

[Ser77] Jean-Pierre Serre. *Cours d'arithmétique*, volume No. 2 of *Le Mathématicien [The Mathematician]*. Presses Universitaires de France, Paris, 1977. Deuxième édition revue et corrigée.

[TA61] H. W. Turnbull and A. C. Aitken. *An introduction to the theory of canonical matrices.* Dover Publications, Inc., New York, 1961.

[vdW50] B. L. van der Waerden. *Modern Algebra. Vol. II.* Frederick Ungar Publishing Co., New York, N. Y., 1950.