



TECHNIQUES
DE L'INGÉNIEUR

Réf. : **AF87 V1**

Date de publication :
10 avril 1999

Réduction des endomorphismes

Cet article est issu de : **Sciences fondamentales | Mathématiques**

par **Rached MNEIMNÉ**

Pour toute question :
Service Relation clientèle
Techniques de l'Ingénieur
Immeuble Pleyad 1
39, boulevard Ornano
93288 Saint-Denis Cedex

Par mail :
infos.clients@teching.com
Par téléphone :
00 33 (0)1 53 35 20 20

Document téléchargé le : **23/12/2022**

Pour le compte : **7200035676 - insa rouen normandie // 195.220.135.37**

© Techniques de l'Ingénieur | tous droits réservés

Réduction des endomorphismes

par **Rached MNEIMNÉ**
Maître de conférences à l'Université Paris VII, Denis-Diderot
Agrégré en mathématiques
Ancien élève de l'École normale supérieure de Saint-Cloud

1. Manipulations premières sur la relation de similitude	AF 87 - 3
2. Valeurs propres. Polynôme caractéristique. Polynôme minimal	— 4
3. La partition de $M(n, \mathbb{C})$ par classes de similitude	— 9
4. Suite des noyaux itérés. Tableaux de Young.....	— 10
5. Matrices nilpotentes. Cône nilpotent	— 12
6. La jordanisation pour elle-même	— 15
7. Familles particulières de matrices. Les matrices de la classe δ	— 15
8. Calcul de la dimension du commutant.....	— 16
9. Réduction simultanée	— 17
10. Autre point de vue sur la jordanisation. La version $\mathbb{K}[X]$ - modules	— 18
11. Matrices de Hessenberg.....	— 19
12. Le cas réel.....	— 20
13. Similitude et congruence. Matrices symétriques réelles.....	— 22
14. Quelques exemples récapitulatifs.....	— 25
Références bibliographiques	— 26

L'algèbre linéaire naît historiquement du besoin de fonder sur des bases solides l'étude des systèmes d'équations linéaires, mais, également, de celui de saisir ce qui survit à la géométrie d'Euclide, une fois gommé l'effet des translations, et, éventuellement, oubliée l'idée de distance. La réduction des endomorphismes n'apparaît que plus tard, et c'est lors de l'examen des équations différentielles à singularités régulières (théorie de Fuchs) que C. Jordan aborde la réduction qui portera son nom.

L'algèbre linéaire se développe petit à petit en une spécialité digne d'intérêt en elle-même, et devient, au sens élémentaire du terme, la « science » qui s'occupe de matrices ou encore d'espaces vectoriels et d'applications linéaires entre ces espaces vectoriels. Les objectifs de base se réduisent, grosso modo, à l'examen de quatre, voire cinq, principales relations d'équivalence définies entre matrices. Il s'agit en fait :

- de la **r-équivalence** ($A = PBQ$) ;
- de la **PG-équivalence** ($A = PB$), qui fonde la première des sources historiques évoquées ci-dessus (PG comme pivot de Gauss) ;
- de la **similitude** ($A = BPB^{-1}$), qui est l'objet de notre étude ;
- de la **congruence** ($A = PB^tP$).

Une autre relation établit enfin certains liens entre similitude et congruence ; elle est donnée par la similitude orthogonale

$$A = OBO^{-1} = OB^tO.$$

Il va s'agir dès lors de chercher à dégager des critères d'appartenance ou de non-appartenance à une classe d'équivalence donnée, à défaut de pouvoir toujours donner une description explicite de ces classes. La présentation adoptée ici fait libre usage du langage des groupes opérant, chaque classe étant une orbite sous l'action du groupe adéquat à la situation.

■ Pour la **similitude**, deux aspects sont à prendre en compte.

Un **aspect classique** consiste, une fois choisie une matrice A d'ordre n à coefficients dans le corps \mathbb{K} , à trouver dans sa classe de similitude une matrice ayant une forme simple (diagonale, quand c'est possible, ou, à défaut, tridiagonale ou triangulaire, etc.), et l'on dit alors la **réduire**, puis trouver un élément du groupe linéaire $GL(n, \mathbb{K})$ qui « transporte » A vers sa forme simple considérée, et l'on parle alors de **matrice de passage**. Cela correspond, pour l'endomorphisme de \mathbb{K}^n canoniquement associé à A , à un changement de base.

Le **deuxième aspect**, qui se développe actuellement aux côtés du premier, consiste en l'examen, pour une matrice donnée A , de la géométrie de sa classe de similitude regardée comme un tout, mais aussi de la géométrie de l'ensemble de toutes les classes de similitude, c'est-à-dire l'espace des orbites.

■ **L'étude de la réduction** soulève de nombreux problèmes d'algorithmique ou d'approximation, dus essentiellement au fait que le calcul des valeurs propres passe, dans un premier temps, par le calcul d'un déterminant à coefficients polynomiaux (le polynôme caractéristique) et dans un second temps par le « calcul » de ses racines. Des résolutions de systèmes linéaires et des inversions de matrices sont également à prendre en considération. C'est la réduction des endomorphismes « effective ».

Enfin, le chapitre de la réduction s'articule sur le chapitre de la réduction des formes quadratiques (la relation de congruence pour les matrices symétriques). C'est le problème de la réduction des opérateurs symétriques dans les espaces euclidiens ou, plus généralement, des opérateurs normaux dans les espaces hermitiens. Similitude et congruence dépendent différemment de la nature du corps de base. La réduction des endomorphismes fait peu intervenir la nature du corps (polynôme caractéristique scindé ou pas) alors que la congruence et les résultats qui s'y rattachent dépendent énormément de l'arithmétique du corps. On se contentera, sauf exception, de regarder la similitude dans les cas de \mathbb{R} et de \mathbb{C} .

■ Quelques applications classiques, en physique ou ailleurs, de la similitude devraient être ici évoquées. Les axes d'inertie d'un solide ou les états propres d'un système de masses avec ressorts illustrent les idées subtiles de la théorie mais ne sont pas des exemples fondamentaux d'application ; on se limitera en fait à l'exemple des ressorts de Trubowitz. On laissera également de côté l'intervention de la réduction dans la théorie de Fuchs. Enfin, on se doit d'indiquer que l'étude des systèmes dynamiques et de la nature de leurs points d'équilibre (pendule, circuit RLC, ressort avec frottements, etc.), étude qui se fait au niveau du système linéaire associé, dépend largement de la réduction des endomorphismes et, notamment, des signes des parties réelles des valeurs propres de la matrice associée.

Nota : on pourra se reporter en bibliographie aux références [7], [17] et [19].

1. Manipulations premières sur la relation de similitude

Définition 1. Deux matrices A et B à coefficients dans le corps \mathbb{K} sont dites semblables, et l'on écrit $A \sim B$, s'il existe une matrice inversible $P \in \text{GL}(n, \mathbb{K})$ telle que $A = PBP^{-1}$.

La définition met déjà en évidence quelques propriétés essentielles de la similitude, et une certaine prudence s'impose quand il s'agit d'appliquer à la similitude des propriétés propres à la congruence ou à la PG-équivalence.

1.1 Similitude et rang

On remarque en premier lieu que deux matrices semblables sont en particulier r -équivalentes, ou, ce qui revient au même, ont même rang (le **rang** classifie les classes d'équivalences pour la r -équivalence des matrices) :

$$A \sim B \Rightarrow A \stackrel{r}{\equiv} B.$$

Par ailleurs, si A et B sont semblables, il en est de même de leurs puissances :

$$\forall k \in \mathbb{N}, \quad A \sim B \Rightarrow A^k \sim B^k,$$

et également, pour tout λ dans \mathbb{K} , des matrices $A - \lambda I$ et $B - \lambda I$, où I désigne la matrice identité :

$$\forall \lambda \in \mathbb{K}, \quad A \sim B \Rightarrow (A - \lambda I) \sim (B - \lambda I).$$

On peut dès lors énoncer le théorème suivant.

Théorème 1 (Weyr). Soit \mathbb{K} un corps algébriquement clos. Deux matrices A et B à coefficients dans \mathbb{K} sont semblables si, et seulement si, pour tout scalaire $\lambda \in \mathbb{K}$ et pour tout entier $k \geq 0$, on a :

$$\text{rg}(A - \lambda I)^k = \text{rg}(B - \lambda I)^k.$$

Corollaire. Toute matrice A est semblable à sa transposée.

a) La démonstration de ce théorème découle de la jordanisation, qui sera traitée dans le paragraphe 4.

b) Il s'avère que l'on peut conjuguer A et sa transposée au moyen d'une matrice de passage symétrique, soit

$${}^tA = SAS^{-1},$$

et déduire de là par exemple que toute matrice complexe est produit de deux matrices symétriques, soit

$$A = (AS^{-1})S.$$

1.2 Similitude et PG-équivalence

Si A est une matrice d'ordre n à coefficients dans \mathbb{K} , on note encore A l'endomorphisme de \mathbb{K}^n , rapporté à sa base canonique

(e_1, \dots, e_n) , canoniquement associé à A . Si donc $A = PBP^{-1}$, on a alors :

$$\text{Ker } A = P(\text{Ker } B) \quad \text{et} \quad \text{Im } A = P(\text{Im } B).$$

Quand A et B sont PG-équivalentes, soit $A = PB$, on a

$$\text{Ker } A = \text{Ker } B;$$

d'ailleurs, la réciproque est vraie : les classes d'équivalence sous l'action de $\text{GL}(n, \mathbb{K})$ par translations à gauche sont paramétrées par les sous-espaces vectoriels de \mathbb{K}^n .

1.3 Similitude et congruence

Remarquons, par ailleurs, que le caractère symétrique ou anti-symétrique d'une matrice ne se conserve pas en général par similitude, comme c'est le cas avec la congruence, notée \approx :

$$(A \approx B \text{ et } {}^tA = \pm A) \Rightarrow {}^tB = \pm B.$$

Il est facile par exemple de prouver (cf § 2.4) que la matrice

$\begin{bmatrix} 1 & x \\ 0 & 2 \end{bmatrix}$ est semblable, pour tout x , à une matrice (diagonale) symétrique.

On démontre, plus généralement, que toute classe de similitude, dans $M(n, \mathbb{C})$, contient une matrice symétrique ; par contre, sur \mathbb{R} , on établit que pour qu'une classe de similitude contienne des matrices symétriques, il faut et il suffit qu'elle contienne une matrice diagonale (cf. § 13.7).

Il est de même aisé de prouver que la matrice $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ est

semblable à la matrice symétrique $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & i \\ 0 & i & 0 \end{bmatrix}$ en même temps qu'à

la matrice antisymétrique $\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -i \\ 0 & i & 0 \end{bmatrix}$ (cf. § 4.4).

1.4 Fonctions polynomiales invariantes

La fonction trace $M \mapsto \text{Tr } M$ et la fonction déterminant $M \mapsto \det M$ sont invariantes par similitude ; il en est de même, par exemple, de $M \mapsto \det M \cdot \text{Tr } M^2$. Les fonctions invariantes, par définition, sont constantes sur les orbites de $\text{GL}(n, \mathbb{K})$ opérant par similitude sur $M(n, \mathbb{K})$. On dispose ainsi d'un moyen pour établir que deux matrices ne sont pas semblables. L'algèbre des fonctions polynomiales sur $M(n, \mathbb{R})$ ou $M(n, \mathbb{C})$ (c'est-à-dire des fonctions qui dépendent polynomialement des coefficients des matrices considérées), qui sont invariantes par similitude, est une algèbre de type fini et même une algèbre de polynômes (en n indéterminées).

Théorème 2. L'algèbre des fonctions polynomiales sur $M(n, \mathbb{C})$ invariantes sous l'action par similitude de $\text{GL}(n, \mathbb{C})$ est engendrée par les n fonctions polynomiales :

$$M \mapsto \text{Tr}(M^i), \quad \text{où } i \in [1, n].$$

Ces fonctions sont algébriquement indépendantes.

a) Notons que les fonctions polynomiales invariantes ne séparent pas dans ce cas les orbites (s'il est vrai que deux matrices semblables ont même polynôme caractéristique, la réciproque est en général fautive - cf. § 3.3).

b) Le théorème implique que le déterminant s'exprime polynomialement en fonction des traces des puissances k -ièmes. Pour une matrice 2×2 , on a :

$$2 \det A = (\operatorname{Tr} A)^2 - \operatorname{Tr} A^2.$$

Pour une matrice 3×3 , on a :

$$6 \det A = (\operatorname{Tr} A)^3 - 3(\operatorname{Tr} A)(\operatorname{Tr} A^2) + 2\operatorname{Tr} A^3.$$

1.5 Commutant d'une matrice et matrices de changement de base

Supposons les matrices carrées A et B , d'ordre n , semblables. Les matrices qui conjuguent A et B s'obtiennent toutes en fonction de l'une d'elles et du groupe des matrices (inversibles) qui commutent avec l'une d'elles. De façon plus précise, soit :

$$Z_M = \{X \in \operatorname{GL}(n, \mathbb{K}), MX = XM\},$$

le groupe Z_M est alors égal au stabilisateur de M sous l'action de $\operatorname{GL}(n, \mathbb{K})$ opérant par similitude (on dit également parfois «par conjugaison») sur $M(n, \mathbb{K})$.

Il est facile de voir que si A et B sont semblables, soit $A = PBP^{-1}$, les sous-groupes Z_A et Z_B sont conjugués, c'est-à-dire que

$$Z_A = P Z_B P^{-1}.$$

Si, de plus, la matrice inversible P_0 conjugue A en B , alors toutes celles qui le font aussi sont obtenues en multipliant P_0 à droite par une matrice de Z_A (ou bien à gauche par une matrice de Z_B).

a) Le groupe $Z(M)$ apparaît également comme le groupe des éléments inversibles de l'algèbre $\beta(M)$ (de dimension finie) des matrices qui commutent avec M .

Le groupe $Z(M)$ est désigné souvent comme le **centralisateur** de M et l'algèbre $\beta(M)$ comme le **commutant** de M .

b) Les matrices dont la classe de similitude est réduite à un point sont exactement les matrices scalaires.

2. Valeurs propres. Polynôme caractéristique. Polynôme minimal

2.1 Généralités

L'existence d'une famille finie de scalaires (il s'agit ici des valeurs propres) associés à une matrice, ou plus précisément à sa classe de similitude, est un phénomène remarquable. On l'abordera mathématiquement de plusieurs points de vue. Ce phénomène se manifeste ou se laisse voir également dans certaines situations physiques. Un exemple facile à traiter est donné par les **ressorts de Trubowitz** : il s'agit de l'étude du mouvement de n masses identiques, mobiles sur un rail circulaire et reliées par des ressorts égale-

ment identiques. L'existence et la détermination de ce que l'on appelle les états purs du système (cas où les masses oscillent toutes avec la même fréquence, et qui sont en nombre fini), sont en fait intimement liées à l'existence et à l'étude des valeurs propres d'une matrice naturellement associée au système (cf. § 14.4).

■ Commençons par un premier constat. Pour une matrice A d'ordre n , le rang de $A - \lambda I$ est un entier qui ne dépend que de la classe de similitude de A et qui a la propriété remarquable d'être constamment égal à n sauf pour un nombre fini de scalaires.

■ Dans la classe de similitude d'une matrice complexe A , il y a toujours des matrices triangulaires. Le fait que toutes ces matrices triangulaires ont mêmes coefficients diagonaux (à une permutation près) ne peut passer inaperçu.

Si deux matrices triangulaires A et B sont semblables, le produit respectif de leurs termes diagonaux n'est autre que l'invariant donné par le déterminant commun à toutes les matrices semblables à A ou B , mais ces coefficients se lisent également dans le (polynôme) déterminant commun aux deux matrices $A - XI$ et $B - XI$ qui sont encore semblables, comme racines de ce polynôme.

■ Étant donné une matrice A , la considération de sous-espaces vectoriels (maximaux et non nuls) de \mathbb{K}^n sur lesquels l'opérateur linéaire associé à A agit par une homothétie nous fait découvrir la notion de sous-espace propre. Le fait remarquable est que ces sous-espaces (qui existent dès lors que l'on se place dans \mathbb{C}) sont en situation de somme directe. Ils sont donc en nombre fini. Les sommes directes ainsi associées aux deux matrices semblables A et $B = PAP^{-1}$ s'appliqueront l'une sur l'autre par l'action de l'opérateur P .

■ La série entière $\sum_k (\operatorname{Tr} A^k) t^k$ a un rayon de convergence R_A non

nul et se prolonge en une fonction méromorphe sur \mathbb{C} dont les pôles, en nombre fini, sont donnés par les inverses des valeurs propres. Le nombre $\rho(A) = 1/R_A$ n'est autre que le plus grand des modules de ces valeurs propres.

Le calcul des résidus en ces pôles apporte de l'information sur la classe de similitude de A [4].

2.2 Valeur, vecteur et sous-espace propres

On donne dans ce paragraphe les définitions usuelles de valeur propre et de vecteur propre, et l'on s'étendra plus longuement sur l'idée de **sous-espace propre**.

Définition 2. Un sous-espace F de $E = \mathbb{K}^n$ est dit stable sous l'action de $f \in \operatorname{End}(E)$ si $f(F) \subset F$. L'application f définit par restriction un endomorphisme de F appelé l'endomorphisme induit par f sur le sous-espace stable F .

La trigonalisation d'un endomorphisme, comme son éventuelle diagonalisation, est intimement liée à l'examen de ses sous-espaces stables : un endomorphisme f d'un espace E de dimension n est **trigonalisable** s'il existe une base dans laquelle sa matrice est triangulaire. Cela s'exprime aussi par l'existence d'une **filtration**, ou **drapeau**, (c'est-à-dire d'une famille de sous-espaces)

$$\{0\} \subset F_1 \subset F_2 \subset \dots \subset F_k$$

maximale (c'est-à-dire formée de $k = n$ sous-espaces distincts) stable par f :

$$\text{pour tout } i, f(F_i) \subset F_i.$$

De même, les endomorphismes réels, diagonalisables sur \mathbb{C} (on parle alors d'**endomorphismes semi-simples**), sont caractérisés par le fait que tout sous-espace stable par un tel endomorphisme admet un supplémentaire stable.

Définition 3. Soit $f \in \text{End}(E)$. Quand l'endomorphisme f admet une droite stable (ou droite propre), il y induit, par restriction, une homothétie de rapport $\lambda \in \mathbb{K}$.

Le scalaire λ ainsi obtenu est appelé une **valeur propre** de f . L'ensemble des valeurs propres de f dans \mathbb{K} s'appelle le **spectre** de f .

a) L'étude de l'ensemble des droites stables par un endomorphisme f est liée à la notion de sous-espaces en somme directe (cf. proposition 1 dans le paragraphe 2.3)

b) Génériquement, en $f \in \text{End}(E)$, où $\dim_{\mathbb{C}} E = n$, il existe un nombre fini (égal à n) de droites propres. Si un endomorphisme en a strictement plus que n , il en a une infinité. En dimension 2, l'existence de trois droites stables distinctes implique que les valeurs propres qui leur sont associées sont égales, auquel cas, toute droite de ce plan est stable. Dans le cas général, la réunion de toutes les droites stables où f agit avec le même scalaire λ est un sous-espace vectoriel, qui n'est autre d'ailleurs que le noyau de $f - \lambda \text{Id}$. Ce sous-espace, non nul lorsque λ est valeur propre, est noté $E_f(\lambda)$ et appelé **sous-espace propre** associé à la valeur propre λ .

Définition 4. On dit que v est un **vecteur propre** de f s'il est non nul et si la droite $D = \mathbb{K}v$ qu'il engendre est stable par f .

On dit que v est un vecteur propre associé à λ , si v est un vecteur non nul de $\text{Ker } f - \lambda \text{Id}$: il est alors vecteur propre de f . De plus, un scalaire λ est une valeur propre si, et seulement si, il existe un vecteur propre v tel que $f(v) = \lambda v$.

2.3 Sous-espaces en somme directe et sous-espaces propres

Définition 5.

a) On dit que **deux** sous-espaces F_1 et F_2 sont en **somme directe** (ou encore **transverses**) si leur intersection est réduite à $\{0\}$. Le sous-espace somme, autrement dit le sous-espace qu'ils engendrent, est noté alors

$$F_1 \oplus F_2.$$

b) On dit que l'espace E est **somme directe** des sous-espaces F_1 et F_2 si

$$E = F_1 \oplus F_2.$$

c) On dit que **trois** sous-espaces sont en **somme directe** si F_1 est transverse à F_2 et que F_3 est transverse à la somme directe $F_1 \oplus F_2$. Cette propriété ne dépend pas de l'ordre avec lequel on procède. On note alors $F_1 \oplus F_2 \oplus F_3$ le sous-espace $F_1 + F_2 + F_3$ qu'ils engendrent. On dit que E est somme directe des sous-espaces F_i si

$$E = F_1 \oplus F_2 \oplus F_3.$$

Ces définitions se généralisent aisément au cas de n sous-espaces vectoriels.

Proposition 1.

Soit n sous-espaces vectoriels F_1, \dots, F_n . Les propriétés suivantes sont équivalentes.

a) Les sous-espaces F_i sont en somme directe.

b) La dimension de $F_1 + \dots + F_n$ est égale à la somme des dimensions des F_i .

c) Si $\sum_i x_i = 0$, où x_i est dans F_i pour tout i , alors les x_i sont tous nuls.

d) Il existe un gonflement des sous-espaces F_i en des sous-espaces G_i (c'est-à-dire que, pour tout i , $F_i \subset G_i$) tel que E soit somme directe des sous-espaces G_i .

e) Les sous-espaces F_i sont les sous-espaces propres d'un endomorphisme de E .

a) Les valeurs propres sont en nombre fini.

b) Les droites propres dessinent ensemble la réunion des sous-espaces propres ; cela précise la géométrie des droites propres.

c) Un hyperplan stable par f est appelé hyperplan propre. Pour décrire la géométrie des hyperplans propres, on regarde dans le dual E^* les sous-espaces propres de ${}^t f$. Les hyperplans cherchés sont les hyperplans qui passent par les orthogonaux dans E de ces sous-espaces propres.

2.4 Matrices diagonalisables

Définition 6.

a) Si la classe de similitude d'une matrice A contient au moins une matrice diagonale, on dit que la matrice A est **diagonalisable**. Le nombre des matrices diagonales d'une même classe de similitude est fini.

b) Un endomorphisme f est dit diagonalisable lorsque sa matrice dans une base l'est. Pour qu'il en soit ainsi, il faut et il suffit que l'espace E soit somme directe des sous-espaces propres de f .

a) Une matrice $A \in M(n, \mathbb{K})$ qui a, dans \mathbb{K} , n valeurs propres distinctes est diagonalisable.

b) Une fois acquise l'existence de valeurs propres (cf. § 2.5), il est facile d'établir qu'un endomorphisme u de $E = \mathbb{C}^n$ est diagonalisable si, et seulement si, tout sous-espace vectoriel stable par u admet un supplémentaire stable (on considérera un supplémentaire stable de la somme de tous les sous-espaces propres).

Quand un endomorphisme $f \in \text{End}(E)$, où E est de dimension finie, est donné sur un corps algébriquement clos, il existe un gonflement privilégié de ses sous-espaces propres en une somme directe de l'espace E ; il s'agit des sous-espaces caractéristiques.

Définition 7. Pour $\lambda \in \mathbb{K}$, valeur propre de l'endomorphisme f , le sous-espace vectoriel formé des vecteurs annulés par l'une ou l'autre des puissances de $f - \lambda \text{Id}$ est appelé le **sous-espace caractéristique** associé à la valeur propre λ et est noté $F_f(\lambda)$.

La restriction de f à $F = F_f(\lambda)$ s'écrit $\lambda \text{Id}_F + z$, où une puissance de l'endomorphisme z est nulle (invoker la dimension finie), c'est-à-dire que z est **nilpotent**.

On a ainsi (**théorème spectral**) :

$$E = F_f(\lambda_1) \oplus \dots \oplus F_f(\lambda_k),$$

et par suite aussi un système **complet** privilégié d'**idempotents orthogonaux** donné par les projections p_i sur chaque facteur parallèlement à la somme des autres :

$$p_i p_j = \delta_{ij} p_i \text{ et } 1 = p_1 + \dots + p_k.$$

On établira que ces idempotents, notés $p_i(f)$, appartiennent à l'algèbre $\mathbb{K}[f]$ des polynômes en f .

L'endomorphisme s de E dont les sous-espaces propres coïncident avec les sous-espaces caractéristiques de f , et qui vaut λ_i sur $F_f(\lambda_i)$ pour tout i , s'appelle la composante semi-simple de f .

Il s'avère que $s = \sum \lambda_i p_i$ commute avec f et que $n = f - s$ est nilpotent, c'est-à-dire qu'une puissance de n est nulle.

L'écriture $f = s + n$ s'appelle la **décomposition de Jordan** de f ou encore la **décomposition de Dunford** de f .

Le fait que s soit diagonalisable dans \mathbb{C} , que n soit nilpotent et que $sn = ns$ caractérise les composantes semi-simple et nilpotente de f .

a) Les vecteurs de $F_f(\lambda)$ sont appelés parfois **vecteurs propres généralisés** de f associés à la valeur propre λ .

b) Le sous-espace caractéristique $F_f(\lambda)$ apparaît ici comme la réunion croissante des noyaux itérés $\text{Ker}(f - \lambda I)^k$, avec $k \in \mathbb{N}$.

2.5 Polynôme caractéristique

L'existence de valeurs propres et de vecteurs propres n'est pas en général garantie : il s'avère en effet que pour que tout endomorphisme de tout espace $E \neq \{0\}$ sur \mathbb{K} ait une droite stable, il faut et il suffit que le corps de base \mathbb{K} soit algébriquement clos.

2.5.1 Généralités

Définition 8. Si A est une matrice d'ordre n à coefficients dans \mathbb{K} , les valeurs propres et les vecteurs propres de l'endomorphisme canoniquement associé à A sont appelés les valeurs propres et les vecteurs propres de la matrice A . Le polynôme unitaire de degré n défini par

$$\chi_A(X) = (-1)^n \det(A - XI_n)$$

est appelé le **polynôme caractéristique** de A . Le $(-1)^n$ est parfois omis ; il vise ici à rendre le polynôme χ_A normalisé.

Les racines dans \mathbb{K} du polynôme caractéristique sont les valeurs propres de A . De plus, le polynôme (unitaire) :

$$P = X^n - a_{n-1}X^{n-1} - \dots - a_0$$

(de degré n) est polynôme caractéristique de sa **matrice compagnon** C_P définie comme la matrice ayant des 1 sur sa sous-diagonale, les a_i sur sa dernière colonne et des zéros partout ailleurs, soit :

$$C_P = \begin{pmatrix} 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & \dots & 0 & a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_{n-2} \\ 0 & \dots & 0 & 1 & a_{n-1} \end{pmatrix}$$

a) Lorsque \mathbb{K} est algébriquement clos, le déterminant d'une matrice est bien le produit des valeurs propres.

b) Le polynôme caractéristique $\chi_A(X)$ ne dépend que de la classe de similitude de A . Plus généralement, les matrices AB et BA ont même polynôme caractéristique et donc mêmes valeurs propres.

c) Le **théorème de Sourour** : aucun lien ne semble par contre exister entre les polynômes caractéristiques de deux matrices et celui de leur produit, autre que celui donné par la multiplicité du déterminant, qui y apparaît comme le terme constant.

En effet, si $\alpha_1, \dots, \alpha_n$ sont les valeurs propres de A , et si β_1, \dots, β_n et $\gamma_1, \dots, \gamma_n$ sont deux familles quelconques de scalaires, l'existence de deux matrices B et C ayant les β_i et les γ_i respectivement pour valeurs propres et telles que $A = BC$ équivaut à l'égalité du produit des β_i et de celui des γ_i .

2.5.2 Matrices trigonalisables

Théorème 3. Sur un corps algébriquement clos, comme \mathbb{C} , toute matrice est trigonalisable.

Preuve. \diamond Soit λ une valeur propre de l'endomorphisme f canoniquement associé à la matrice considérée. Tout hyperplan contenant l'image de $f - \lambda \text{Id}$ est stable par $f - \lambda \text{Id}$, donc par f .

En recommençant avec un tel hyperplan et une valeur propre de la restriction de f à cet hyperplan, et ainsi de suite, on exhibe un **drapeau maximal** stable par f . \diamond

Si l'on dispose d'un drapeau stable par un endomorphisme f , la matrice de f dans une base adaptée à ce drapeau est une matrice triangulaire en blocs. Le polynôme caractéristique de f est, d'après une propriété bien classique du déterminant, le produit des polynômes caractéristiques des blocs diagonaux. Il est facile à partir de là de démontrer que pour qu'une matrice à coefficients dans un corps \mathbb{K} soit trigonalisable sur \mathbb{K} , il faut et il suffit que son polynôme caractéristique soit scindé sur \mathbb{K} .

a) Si $P = p_m X^m + \dots + p_1 X + p_0 \in \mathbb{C}[X]$, on note $P(A)$ la matrice $p_m A^m + \dots + p_1 A + p_0 I$.

Les valeurs propres de $P(A)$ sont exactement les $P(\lambda)$, pour λ parcourant le spectre de A (trigonaliser). En particulier, pour $P = \chi_A(X)$, on constate que $\chi_A(A)$ est nilpotente ; en fait, le théorème de Cayley-Hamilton (cf. § 2.6) dit que $\chi_A(A)$ est nulle.

b) Une matrice nilpotente N a ses valeurs propres nulles. Si maintenant $N \in M(n, \mathbb{K})$ a toutes ses valeurs propres nulles, elle est semblable à une matrice triangulaire supérieure stricte T ; on a clairement $T^n = 0$, il en est donc de même de N .

2.5.3 Coefficients du polynôme caractéristique

Le calcul effectif du polynôme caractéristique donne lieu à différentes méthodes en calcul numérique, que nous ne traitons pas ici.

Pour une matrice 2×2 , on a :

$$\chi_A(X) = \det \begin{pmatrix} a-X & b \\ c & d-X \end{pmatrix} = X^2 - \text{Tr}(A)X + \det A.$$

Pour une matrice $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$, on a :

$$\chi_A(X) = X^3 - \text{Tr}(A)X^2 + [(ae - bd) + (ak - cg) + (ek - fh)]X - \det A.$$

Plus généralement, le coefficient de X^k est égal, au signe $(-1)^{n-k}$ près, à la trace de la puissance extérieure $\Lambda^k(A)$ de A , ou encore à la somme des mineurs principaux d'ordre k . C'est aussi la fonction symétrique $(n-k)$ -ième des racines de $\chi_A(X)$, c'est-à-dire des valeurs propres. Ainsi, la trace est la somme des valeurs propres et le déterminant leur produit.

2.5.4 Multiplicités d'une valeur propre

Définition 9. La multiplicité d'une valeur propre λ en tant que racine du polynôme caractéristique de A est appelée sa **multiplicité algébrique** et est notée $m_a(\lambda)$.

La dimension du sous-espace propre $E_A(\lambda)$ est désignée comme la **multiplicité géométrique** et est notée $m_g(\lambda)$.

Théorème 4. On a

$$m_a(\lambda) = \dim F_A(\lambda) = \text{Tr}(p_i(A)).$$

On a donc toujours

$$m_g(\lambda) \leq m_a(\lambda).$$

Si la matrice A a toutes ses racines dans \mathbb{K} , alors elle est diagonalisable si, et seulement si, pour tout λ dans son spectre, on a :

$$m_g(\lambda) = m_a(\lambda).$$

Preuve. \diamond Le polynôme caractéristique de la restriction de A au sous-espace stable E_λ divise celui de A ; on en déduit l'inégalité :

$$m_g(\lambda) \leq m_a(\lambda).$$

Si A est diagonalisable, l'espace est somme directe des sous-espaces propres et l'on a donc :

$$n = \sum_i m_g(\lambda_i).$$

Mais :

$$n = \deg \chi_A(X) = \sum_i m_a(\lambda_i).$$

L'inégalité implique la condition nécessaire. La condition suffisante est facile.

Reste la première assertion. La deuxième égalité résulte de ce que le rang d'un projecteur est égal à sa trace (le diagonaliser). Par ailleurs, le théorème spectral s'interprète en disant que la matrice A est semblable à une matrice B diagonale en blocs (les sous-espaces caractéristiques sont stables), et telle que chaque bloc est de la forme $\lambda I + N$ où λ est la valeur propre correspondante et où N nilpotent. Le calcul du polynôme caractéristique de B est alors facile et permet de conclure. \diamond

Si N est une matrice nilpotente d'ordre n , on a :

$$\chi_{\lambda I + N}(X) = (X - \lambda)^n.$$

2.6 Théorème de Cayley-Hamilton

Théorème 5 (Cayley-Hamilton). Le polynôme caractéristique de A annule A .

Preuve. \diamond Si λ_1 est une valeur propre de A et H un hyperplan contenant l'image de $A - \lambda_1 I$, la restriction de A à H a les mêmes valeurs propres que A , la valeur propre λ_1 en moins.

En répétant l'opération, il est facile d'établir que le rang de

$$(A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_i I)$$

est inférieur ou égal à $n - i$. \diamond

Si toutes les valeurs propres de $A \in M(n, \mathbb{K})$ sont nulles, le polynôme caractéristique est donc égal à X^n . Le théorème de Cayley-Hamilton nous permet donc de retrouver le fait que la matrice A est nilpotente et vérifie $A^n = 0$.

Définition 10. L'ensemble des polynômes qui annulent A est un idéal de l'anneau principal $\mathbb{K}[X]$. Le polynôme unitaire qui engendre cet idéal s'appelle le **polynôme minimal** de A et est noté $\mu_A(X)$. Il divise le polynôme caractéristique et son degré qui est aussi la dimension de l'algèbre $\mathbb{K}[A]$ des polynômes en A est donc $\leq n$.

a) Soit $A \in \text{GL}(n, \mathbb{K})$. L'inverse A^{-1} de A est un polynôme en A . On peut en outre exprimer A^{-1} à l'aide de $\mu_A(X)$ ou encore de $\chi_A(X)$, car alors le terme constant dans chacun de ces polynômes est non nul.

b) Soit $A = \text{Diag}(a_1, \dots, a_n)$. Supposons les a_i distincts. L'algèbre $\mathbb{K}[A]$ coïncide avec l'algèbre des matrices diagonales (on pensera aux polynômes d'interpolation de Lagrange), et l'on a :

$$\chi_A(X) = \mu_A(X) = \prod_i (X - a_i).$$

c) Le polynôme minimal d'une matrice diagonale en blocs est égal au PPCM des polynômes minimaux des blocs diagonaux.

2.7 Théorème spectral et autre point de vue sur le théorème de Cayley-Hamilton

Le **théorème spectral** peut s'établir de diverses manières. Nous en indiquons deux. La première passe par l'examen des noyaux des polynômes en A et l'autre par des considérations élémentaires sur la structure de l'algèbre des polynômes en A et de ses idempotents. On utilisera dans l'une de façon franche l'identité de Bezout et on pourra éviter de le faire dans la suivante. Cependant, l'identité de Bezout ou le théorème chinois, et leur pratique effective sont au cœur de cette approche. On s'en sert pour calculer par exemple les idempotents du théorème spectral mais aussi la décomposition de Jordan explicite d'une matrice ou pour calculer une exponentielle, etc.

Une fois acquis le théorème spectral, son interprétation matricielle implique aisément le **théorème de Cayley-Hamilton**.

La matrice A est, en effet, semblable à une matrice diagonale en blocs, avec au niveau du i -ième bloc, une matrice de la forme $\lambda_i I + N_i$, où N_i est une matrice nilpotente dont l'indice de nilpotence

est la multiplicité de λ_i dans le polynôme minimal ; on la note $m_m(\lambda_i)$.

Le calcul du polynôme caractéristique indique que l'ordre des matrices figurant dans le i -ième bloc est exactement la multiplicité $m_a(\lambda_i)$ de la valeur propre λ_i dans le polynôme caractéristique. Le théorème de Cayley-Hamilton exprime simplement que la multiplicité $m_m(\lambda)$, qui est en fait ≥ 1 , est $\leq m_a(\lambda)$. Cela résulte de ce que l'indice de nilpotence d'une matrice nilpotente quelconque N est plus petit que son ordre, soit m : cela a été déjà constaté par une trigonalisation, mais on peut aussi l'obtenir en remarquant que pour une matrice quelconque M d'ordre m opérant dans \mathbb{K}^m , la suite des noyaux itérés

$$\{0\} \subset \text{Ker } M \subset \text{Ker } M^2 \subset \dots \subset \text{Ker } M^i \subset \dots$$

est strictement croissante avant d'être stationnaire.

Si $\text{Ker } M^m \neq \text{Ker } M^{m+1}$, toutes les inclusions précédentes sont strictes et la dimension de $\text{Ker } M^i$ est supérieure ou égale à i pour tout i , y compris $i = m + 1$, d'où contradiction.

L'application de ce résultat dans le cas de N prouve que la suite des noyaux itérés de N stationne avant l'étape m ; or, elle stationne en \mathbb{K}^m à l'étape donnée par son indice de nilpotence.

Soit λ une valeur propre de M . On retient donc que la suite des noyaux itérés

$$\{0\} \subset E_M(\lambda) = \text{Ker } (M - \lambda I) \subset \text{Ker } (M - \lambda I)^2 \subset \dots \subset \text{Ker } (M - \lambda I)^i \subset \dots$$

est strictement croissante jusqu'à l'étape donnée par la multiplicité $m_m(\lambda)$ de λ dans le polynôme minimal, et qu'elle stationne au delà. L'espace auquel aboutit la suite est le sous-espace caractéristique $F_M(\lambda)$, qui lui est de dimension $m_a(\lambda)$.

2.8 Théorème spectral : deux démonstrations

2.8.1 Noyaux des polynômes en une matrice

On note $\Delta(X)$ et $\Gamma(X)$ le PGCD et le PPCM des deux polynômes P et Q de l'anneau principal $\mathbb{K}[X]$. Ils peuvent se définir comme les polynômes unitaires qui engendrent les idéaux somme et produit des idéaux engendrés par P et Q , soit

$$(P) + (Q) = (\Delta)$$

$$(P) \cap (Q) = (\Gamma).$$

Si A est une matrice, on a par contre [2] :

$$\text{Ker } P(A) + \text{Ker } Q(A) = \text{Ker } \Gamma(A)$$

$$\text{Ker } P(A) \cap \text{Ker } Q(A) = \text{Ker } \Delta(A).$$

L'identité de Bezout est ici omniprésente.

Si maintenant

$$\mu_A(X) = (X - \lambda_1)^{\alpha_1} (X - \lambda_2)^{\alpha_2} \dots (X - \lambda_k)^{\alpha_k}$$

est la décomposition du polynôme minimal (ou de n'importe quel polynôme annulateur) $\mu_A(X)$ de A en facteurs irréductibles, on en déduit que

$$E = \bigoplus_i \text{Ker } (A - \lambda_i I)^{\alpha_i}.$$

Par ailleurs, les projecteurs sur les facteurs de cette somme directe sont encore des polynômes en A , comme le montre

l'exemple suivant, où $k = 3$. L'endomorphisme $P(A)$, où P vérifie le système de congruences

$$P \equiv 0 \pmod{(X - \lambda_1)^{\alpha_1}}$$

$$P \equiv 0 \pmod{(X - \lambda_2)^{\alpha_2}}$$

$$P \equiv 0 \pmod{(X - \lambda_3)^{\alpha_3}}$$

s'annule sur $\text{Ker } (A - \lambda_1 I)^{\alpha_1} \oplus \text{Ker } (A - \lambda_3 I)^{\alpha_3}$ et vaut l'identité sur $\text{Ker } (A - \lambda_2 I)^{\alpha_2}$; c'est donc le projecteur sur le deuxième facteur de la somme directe

$$E = \text{Ker } (A - \lambda_1 I)^{\alpha_1} \oplus \text{Ker } (A - \lambda_2 I)^{\alpha_2} \oplus \text{Ker } (A - \lambda_3 I)^{\alpha_3},$$

parallèlement à la somme des deux autres. Le théorème chinois assure l'existence de P . Le calcul effectif de $P(A)$ découle de la résolution effective du théorème chinois.

Applications

a) Calcul de la partie semi-simple d'une matrice

On considère la matrice

$$M = \begin{bmatrix} a & x & z \\ 0 & a & y \\ 0 & 0 & b \end{bmatrix},$$

où l'on a supposé $a \neq b$.

La partie semi-simple de M est donnée par

$$S = \begin{bmatrix} a & 0 & \frac{xy}{b-a} + z \\ 0 & a & y \\ 0 & 0 & b \end{bmatrix}.$$

En effet, le polynôme $p(X) = a + \frac{1}{b-a}(X-a)^2$ vérifie

$$p(X) \equiv a \pmod{(X-a)^2}$$

et

$$p(X) \equiv b \pmod{(X-b)}.$$

On a donc

$$S = p(M).$$

b) Calcul de l'exponentielle

Soit à calculer l'exponentielle d'une matrice A vérifiant

$$(A - I)^2 (A - 2I) = 0.$$

On considère pour cela l'écriture suivante de E en somme directe

$$E = \text{Ker } (A - I)^2 \oplus \text{Ker } (A - 2I),$$

qui répond à l'identité de Bezout $1 = a(x)(x-1)^2 + b(x)(x-2)$, et l'on note p la projection sur $\text{Ker}(A - 2I)$ parallèlement à $\text{Ker}(A - I)^2$ et $q = I - p$. On a alors :

$$e^A = e^A p + e^A q$$

et

$$e^A p = e^2 e^{A-2I} p = e^2 \left[\sum \frac{(A-2I)^n}{n!} \right] p = e^2 p$$

et de même

$$e^A q = e e^{A-I} q = e(I + A - I) q.$$

On peut prendre $a(x) = 1$ et $b(x) = x$. On en déduit :

$$e^A = e^2(A - I)^2 + eAA(A - 2I).$$

2.8.2 L'algèbre $\mathbb{K}[A]$ et ses idempotents

L'anneau $\mathbb{K}[X]/\mu_A(X)$, quotient de l'anneau des polynômes $\mathbb{K}[X]$ par l'idéal engendré par le polynôme

$$\mu_A(X) = (X - \lambda_1)^{\alpha_1}(X - \lambda_2)^{\alpha_2}\dots(X - \lambda_k)^{\alpha_k},$$

s'identifie à l'algèbre $\mathbb{K}[A]$ des polynômes en A . Cet anneau est produit des anneaux $\mathbb{K}[X]/(X - \lambda_i)^{\alpha_i}$. On a donc :

$$\mathbb{K}[X]/\mu_A(X) = \mathbb{K}[X]/(X - \lambda_1)^{\alpha_1} \oplus \dots \oplus \mathbb{K}[X]/(X - \lambda_k)^{\alpha_k}.$$

Preuve. \diamond Cela résulte du fait que $P = P_1P_2$, où P_1 et P_2 sont premiers entre eux, l'application naturelle

$$\mathbb{K}[X]/P(X) \rightarrow \mathbb{K}[X]/P_1(X) \times \mathbb{K}[X]/P_2(X)$$

est un homomorphisme injectif d'algèbres de même dimension. C'est donc un isomorphisme. \diamond

Lemme 1. Si, pour $i = 1, \dots, k$, on dispose d'une sous-algèbre unitaire \mathcal{A}_i de $\text{End}(E_i)$, alors l'algèbre somme $\mathcal{A} = \bigoplus_i \mathcal{A}_i$ est naturellement une sous-algèbre unitaire de $\text{End}(E)$, où $E = \bigoplus_i E_i$.

Inversement, à une réalisation d'une sous-algèbre unitaire $\mathcal{A} \subset \text{End}(E)$ comme produit d'algèbres unitaires \mathcal{A}_i correspond une décomposition de E en une somme directe de sous-espaces vectoriels E_i , telle que chaque \mathcal{A}_i apparait comme une sous-algèbre unitaire de $\text{End}(E_i)$ et que \mathcal{A} s'identifie à la sous-algèbre $\bigoplus_i \mathcal{A}_i$. On a donc :

$$\mathcal{A} \simeq \bigoplus_i \mathcal{A}_i \subset \bigoplus_i \text{End}(E_i) \subset \text{End}(E) \stackrel{\text{e.v.}}{\simeq} \bigoplus_{i,j} \text{Hom}(E_i, E_j).$$

Preuve. \diamond La première assertion est facile, en considérant des matrices diagonales en blocs. Il s'agit cependant pour le reste de bien comprendre comment l'unité de chaque \mathcal{A}_i devient un idempotent p_i de l'algèbre produit, et comment, partant d'une décomposition de l'unité $I = p_1 + \dots + p_k$ dans $\mathcal{A} \subset \text{End}(E)$ en une somme d'idempotents orthogonaux, c'est-à-dire en des projecteurs de E vérifiant $p_i p_j = \delta_{ij}$, on obtient E comme somme directe des images des p_i . \diamond

Le théorème spectral s'obtient maintenant grâce au lemme 1 et à la remarque suivante : l'endomorphisme A s'identifie à la classe de X dans le quotient $\mathbb{K}[X]/\mu_A(X)$ laquelle s'identifie au k -uplet formé des classes de X dans chaque quotient $\mathbb{K}[X]/(X - \lambda_i)^{\alpha_i}$.

Mais, dans un tel quotient, on a $X = \lambda + (X - \lambda)$, où λ est inversible et $(X - \lambda)$ est nilpotent d'indice α .

Il s'ensuit que la matrice A est semblable à une matrice $B = PAP^{-1}$ diagonale en blocs, avec au niveau du i -ième bloc, une matrice de la forme $\lambda_i I + N_i$, où N_i est une matrice nilpotente dont l'indice de nilpotence est la multiplicité $m_{\lambda_i}(\lambda)$ dans le polynôme minimal.

La vérification de l'assertion à examiner est directe pour la matrice B . Mais les polynômes caractéristiques de A et B sont égaux et leurs sous-espaces caractéristiques sont **conjugués** par P et donc de même dimension. Enfin, les sous-espaces caractéristiques de B sont les sous-espaces apparaissant dans sa décomposition en

blocs. Ils sont donc trivialement en somme directe, il en est de même des sous-espaces caractéristiques de A .

2.9 Résultant du polynôme caractéristique

L'ouvert (algébrique) des matrices ayant des valeurs propres distinctes (on parlera de matrices **régulières**) est, comme tout complémentaire des zéros d'un polynôme, un ouvert connexe dense pour la topologie usuelle de $M(n, \mathbb{C})$. Le polynôme dont il s'agit est le résultant du polynôme caractéristique de la matrice générale (X_{ij}) .

Une matrice régulière vérifie à la fois qu'elle est diagonalisable et que son polynôme caractéristique est égal à son polynôme minimal (matrice **monogène**). Ces deux propriétés caractérisent les matrices régulières.

La matrice générale (X_{ij}) a ses valeurs propres distinctes, car le résultant de son polynôme caractéristique, qui est un polynôme dans $\mathbb{Z}[X_{11}, \dots, X_{nn}]$, est non nul, car sinon, par spécialisation, toute matrice aurait au moins une valeur propre double, ce qui n'est évidemment pas le cas. Le théorème de Cayley-Hamilton, facile à établir dans le cas diagonalisable, est donc vrai pour la matrice générale. Par spécialisation, il est donc valable pour toute matrice.

Remarque sur la **terminologie**. Certains auteurs réservent le vocable « matrices régulières » aux matrices appelées ici « monogènes ». Les matrices que nous avons appelées régulières deviennent chez quelques-uns « matrices génériques » et judicieusement chez d'autres « matrices semi-simples régulières ». Les notions concernées trouvent des répondants dans le cadre des algèbres de Lie. Une discussion sérieuse sur la valeur respective de tel ou tel choix terminologique doit tenir compte de son adaptabilité à ce cadre. Les choix effectués ici ont été dictés par des considérations d'usage ou de simplicité.

3. La partition de $M(n, \mathbb{C})$ par classes de similitude

3.1 Généralités

Alors que la r -équivalence donne des classes d'équivalence en nombre fini (paramétrées par le rang), et que la congruence donne, dans le cas des matrices symétriques réelles, un nombre également fini d'orbites (paramétrées par la signature), le nombre d'orbites sous l'action à gauche de $\text{GL}(n, \mathbb{K})$ est infini dès que \mathbb{K} l'est. En effet, les classes d'équivalence sont alors paramétrées par les matrices échelonnées ou encore par les sous-espaces vectoriels de \mathbb{K}^n .

Le nombre d'orbites est également infini dans l'action par conjugaison de $\text{GL}(n, \mathbb{K})$ dès que \mathbb{K} l'est aussi. On a cependant le fait remarquable suivant : pour la relation d'équivalence

$$(A \mathcal{P} B) \Leftrightarrow A \text{ et } B \text{ ont même polynôme caractéristique,}$$

laquelle relation est moins fine que la similitude, on est en présence d'une partition de $M(n, \mathbb{C})$ en une infinité (bien contrôlable) de classes d'équivalence (car clairement paramétrées par les polynômes de degré n ou ce qui revient au même par les orbites du groupe symétrique \mathfrak{S}_n sur \mathbb{C}^n) qui **chacune** est réunion **finie** de classes de similitude! On se représentera cela dans la suite comme une infinité d'**armoires** comportant chacune un nombre fini de **tiroirs** (figure 1).

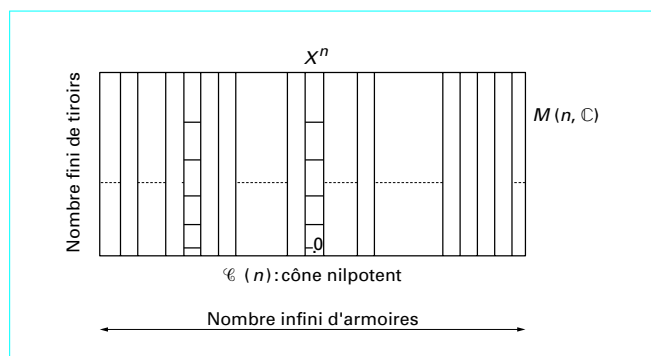


Figure 1 – La partition de $M(n, \mathbb{C})$ par classes de similitude

3.2 La partition donnée par l'égalité des polynômes caractéristiques

La relation \mathcal{P} définie paragraphe 3.1 répartit les matrices en familles ayant le même polynôme caractéristique. Une classe d'équivalence est attachée à chaque polynôme unitaire de degré n . L'espace quotient est donc l'ensemble $\mathbb{C}_U^n[X]$ de ces polynômes. Mais un polynôme unitaire

$$X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0$$

de degré n dépend des n scalaires a_0, \dots, a_{n-1} .

Leur ensemble s'identifie donc à \mathbb{C}^n . Par ailleurs, l'application qui au n -uplet $(\lambda_1, \dots, \lambda_n)$ associe le polynôme $\prod_i (X - \lambda_i)$ passe au quotient par l'action de \mathfrak{S}_n sur \mathbb{C}^n et identifie l'espace des orbites $\mathfrak{S}_n / \mathbb{C}^n$ avec l'ensemble $\mathbb{C}_U^n[X]$. On a donc

$$M(n, \mathbb{C}) / \mathcal{P} \simeq \mathbb{C}_U^n[X] \simeq \mathbb{C}^n$$

et

$$\mathfrak{S}_n \backslash \mathbb{C}^n \simeq \mathbb{C}_U^n[X].$$

Dans l'identification

$$\mathfrak{S}_n \backslash \mathbb{C}^n \simeq \mathbb{C}^n$$

$$\varphi : (\lambda_1, \dots, \lambda_n) \mapsto \left(\sigma_1 = \sum_i \lambda_i ; \sigma_2 = \sum_{i < j} \lambda_i \lambda_j ; \dots ; \sigma_n = \prod_i \lambda_i \right)$$

qui en découle, le sous-ensemble de $\mathfrak{S}_n \backslash \mathbb{C}^n$ formé des orbites de n -uplets distincts (on parle alors de **point régulier**) correspond au sous-ensemble de \mathbb{C}^n formé des points en dehors de la variété algébrique affine définie par le résultant du polynôme $\prod_i (X - z_i)$. La

plupart des classes d'équivalence modulo \mathcal{P} sont donc formées de matrices régulières. Il reste qu'une classe modulo \mathcal{P} sera appelée par les valeurs propres communes aux matrices qu'elle contient ou par le polynôme unitaire de degré n qui la paramètre.

A ce stade, deux remarques s'imposent :

— la première est que dans une classe modulo \mathcal{P} formée de matrices régulières (associée donc à un point régulier de \mathbb{C}^n), toutes les matrices sont semblables (autrement dit, une armoire régulière ne contient qu'un seul tiroir, c'est donc une **penderie**) et c'est le

cas de toutes les armoires ou peu s'en faut (on dit que cela est le cas génériquement) ;

— la deuxième est que, parmi toutes les classes modulo \mathcal{P} , il en est une qui occupe une place très particulière aussi bien au niveau géométrique qu'au niveau de l'importance, c'est celle appelée par l'orbite de l'origine sous l'action de \mathfrak{S}_n ou encore par l'élément nul de \mathbb{C}^n correspondant donc au polynôme unitaire X^n . Cette classe n'est autre que le **cône nilpotent** $\mathcal{C}(n)$.

3.3 Description des classes de similitude d'une même classe modulo \mathcal{P}

Soit $P(X)$ un polynôme unitaire de degré n . On le décompose en produit de facteurs irréductibles, soit

$$P(X) = (X - \lambda_1)^{\alpha_1} \dots (X - \lambda_k)^{\alpha_k}.$$

L'ensemble des matrices \mathcal{A}_P ayant $P(X)$ comme polynôme caractéristique (qui est donc $\mu^{-1}[\varphi(\lambda_1, \dots, \lambda_n)]$) est réunion de classes de similitude en nombre fini, égal exactement au produit $p(\alpha_1) \dots p(\alpha_k)$, où $p(m)$ est le nombre de partitions de l'entier $m \in \mathbb{N}$, c'est-à-dire le nombre de façons d'écrire l'entier

$$m = \sum_i m_i$$

comme somme d'entiers $m_1 \geq m_2 \geq \dots \geq m_h$ allant en décroissant.

Le cas particulier où les α_i sont tous égaux à 1 a déjà été évoqué paragraphe 3.2.

Un autre cas particulier est celui où $k = 1$, duquel relèvent et le cône nilpotent $\mathcal{C}(n) = \mathcal{A}_{X^n}$ et, ce qui revient essentiellement au même, l'ensemble des matrices A telles que $A - \lambda I$ est nilpotent. Le cône nilpotent $\mathcal{C}(n)$ contient ainsi $p(n)$ classes de similitude.

Par ailleurs, deux classes de similitude (éventuellement confondues) occupent dans « l'armoire » \mathcal{A}_P une place spéciale :

— la première, de taille nettement plus grande que les autres (quand il y en a), est formée des matrices dont le polynôme minimal est P ;

— la deuxième, plus petite que les autres (quand il y en a), est formée des matrices de \mathcal{A}_P qui sont diagonalisables.

On a déjà constaté que ces deux classes sont les mêmes précisément dans le cas régulier. Dans le cas nilpotent, la première est formée des matrices nilpotentes telles que A^{n-1} est non nulle et l'autre est réduite à la matrice nulle. Les autres classes de similitude qui sont contenues dans le cône nilpotent sont particulièrement bien saisies grâce aux tableaux de Young qui en donnent un paramétrage (cf. § 4.4).

Le cas général s'obtient à partir de là en associant à chaque valeur propre λ_i un tableau de Young comportant α_i cases correspondantes qui résume l'information sur les rangs des puissances successives de $A - \lambda_i I$. Les classes de similitude de l'armoire \mathcal{A}_P sont ainsi paramétrées par les systèmes de k tableaux de Young, ayant chacun α_k cases.

4. Suite des noyaux itérés. Tableaux de Young

Soit λ une valeur propre de A . En remplaçant dans ce qui suit A par $A - \lambda I$, on pourra supposer que λ est nulle. On va donc considérer pour la valeur propre $\lambda = 0$, une suite croissante de sous-espaces vectoriels, et de là on associera à $\lambda = 0$ un tableau de Young, noté $\text{TY}(A, \lambda)$.

4.1 Une suite qui s'essouffle

La suite croissante

$$\{0\} \subset \text{Ker } A \subset \text{Ker } A^2 \subset \dots \subset \text{Ker } A^k \dots$$

est une suite qui « s'essouffle », en ce sens que les sauts de dimension vont en diminuant ; ce résultat clé découle des **injections de Frobenius** induites par l'endomorphisme A :

$$\text{Ker } A^{k+1} / \text{Ker } A^k \hookrightarrow \text{Ker } A^k / \text{Ker } A^{k-1}.$$

On peut également constater cela à partir des **suites exactes courtes de Frobenius** :

$$\{0\} \rightarrow \text{Ker } A^m \rightarrow \text{Ker } A^{m+1} \xrightarrow{A^m} \text{Ker } A \cap \text{Im } A^m \rightarrow \{0\}.$$

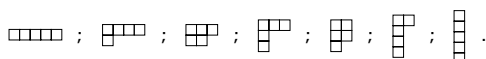
Du fait de la dimension finie n , la suite des noyaux itérés est strictement croissante jusqu'au moment où elle devient stationnaire et l'on a sûrement

$$\text{Ker } A^n = \text{Ker } A^{n+1},$$

car dans le cas contraire, toutes les inclusions précédentes seraient strictes et $\text{Ker } A^{n+1}$ serait au moins de dimension $n+1$.

4.2 Tableaux de Young

Un **tableau de Young** est un tableau formé de cases disposées en colonnes (ou si l'on préfère par lignes) et dont les longueurs vont en décroissant (on notera que le tableau transposé d'un tableau de Young en est encore un). Les tableaux suivants sont tous les tableaux comportant 5 cases :



Le nombre de tableaux de Young ayant un nombre total de ℓ cases est en fait égal à $p(\ell)$, le nombre de **partitions** de l'entier ℓ .

4.3 Tableau de Young associé à une valeur propre

A partir de la suite des noyaux itérés, on constitue le **tableau de Young associé à la valeur propre 0** à partir de ses colonnes et ce en plaçant :

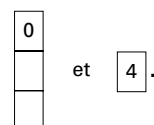
- à la première colonne, un nombre de cases égal à la dimension de $\text{Ker } A$;
- à la deuxième colonne, le saut de dimension entre $\text{Ker } A$ et $\text{Ker } A^2$ (égal aussi à la dimension de $\text{Ker } A \cap \text{Im } A$) ;
- ainsi de suite, en plaçant donc à la k -ième colonne le saut de dimension entre $\text{Ker } A^{k-1}$ et $\text{Ker } A^k$ (égal aussi à la dimension de $\text{Ker } A \cap \text{Im } A^{k-1}$)...

Le nombre de cases dans le tableau de Young correspond à la dimension du sous-espace caractéristique associé à la valeur propre $\lambda = 0$, c'est-à-dire la multiplicité de $\lambda = 0$ dans le polynôme caractéristique. Quant à la longueur de la première ligne, elle indique la multiplicité de la valeur propre $\lambda = 0$ dans le polynôme minimal.

Le **théorème de Weyr** s'exprime, dans ce contexte, comme suit : pour que deux matrices A et B soient semblables, il faut et il suffit que leurs polynômes caractéristiques soient égaux et que, pour chaque valeur propre λ , on ait

$$\text{TY}(A, \lambda) = \text{TY}(B, \lambda).$$

En présence de plusieurs valeurs propres, on pourra inscrire la valeur propre dans la première case du tableau de Young correspondant. Ainsi, la classe de similitude de la matrice réelle d'ordre 4 ayant des 1 partout est paramétrée par les deux tableaux de Young :



4.4 Pratique de la réduction de Jordan pour une matrice nilpotente

Un **bloc** de Jordan J_n , ou encore **cellule** de Jordan, est une matrice triangulaire supérieure stricte dont tous les coefficients sont nuls sauf ceux de la surdiagonale qui valent 1. Ainsi le bloc J_4 est donné par :

$$J_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$


La forme de Jordan canonique d'une matrice nilpotente consiste en une matrice diagonale en blocs, avec, sur la diagonale, des cellules de Jordan de tailles décroissantes. De façon plus précise, comme les sauts de dimension dans la suite croissante des noyaux itérés

$$\{0\} \subset \text{Ker } A \subset \text{Ker } A^2 \subset \dots \subset \text{Ker } A^k \subset \dots$$

vont en diminuant, on dispose ces sauts de dimension en colonnes dans un tableau de Young et l'on obtient la forme de Jordan en considérant les lignes. Les tailles des cellules de Jordan dans la forme de Jordan sont les longueurs des lignes du tableau de Young associé

Ainsi, par **exemple**, pour une **matrice A d'ordre 8** telle que

$$\begin{aligned} \dim \text{Ker } A &= 3, \\ \dim \text{Ker } A^2 &= 5, \\ \dim \text{Ker } A^3 &= 7, \\ \text{et} \quad A^4 &= 0, \end{aligned}$$

le tableau des sauts de dimension est donné par  ; la réduite de Jordan canonique de A est alors :

$$\begin{bmatrix} \boxed{\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}} & & & \\ & \boxed{\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}} & & \\ & & \boxed{\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}} & \\ & & & \boxed{0} \end{bmatrix}$$

La détermination effective d'une base de jordanisation, ou ce qui revient au même, d'une matrice de passage, sera traitée dans le paragraphe 6.1.

4.5 Pratique de la réduction de Jordan pour une matrice quelconque

Le cas d'une matrice quelconque M se déduit aisément du cas nilpotent (§ 4.4). On constitue les différents tableaux de Young associés aux différentes valeurs propres de M . Sur chacun des sous-espaces caractéristiques, la matrice M est, à une homothétie près, nilpotente. Pour une valeur propre λ non nécessairement nulle, la forme de Jordan canonique de la restriction de M au sous-espace caractéristique $F_M(\lambda)$ est donnée par un tableau diagonal de cellules de Jordan $J_k(\lambda)$ associées à λ et définies par

$$J_k(\lambda) = \lambda I_k + J_k$$

Les tailles décroissantes de ces cellules correspondent aux longueurs des lignes du tableau de Young $TY(M, \lambda)$ de M relatif à la valeur propre λ .

Ainsi, si A est une matrice d'ordre 5 ayant 1 puis 2 pour valeurs propres (et que l'on ordonne comme indiqué), et pour tableaux de Young respectifs associés \square et $\square\square$, la forme de Jordan canonique de A est donnée par :

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

On peut dès lors énoncer le théorème suivant.

Théorème 6. Une matrice $M \in M(n, \mathbb{C})$ est diagonalisable si, et seulement si, tous les tableaux de Young associés à ses valeurs propres sont unicolonnés, et M est monogène (c'est-à-dire ayant un polynôme minimal égal à son polynôme caractéristique) si, et seulement si, les tableaux de Young associés à toutes ses valeurs propres sont unilignes.

5. Matrices nilpotentes. Cône nilpotent

L'ensemble des matrices nilpotentes est le cône nilpotent, noté $\mathcal{C}(n)$.

5.1 Matrices nilpotentes

On obtient facilement les caractérisations suivantes. Une matrice $A \in M(n, \mathbb{C})$ est dite nilpotente si elle satisfait l'une des conditions équivalentes suivantes, où le crochet $[X, Y]$ des matrices X et Y désigne la matrice $XY - YX$.

- La matrice A a une puissance nulle.
- La puissance n -ième de A est nulle.
- La suite des noyaux itérés aboutit à $E = \mathbb{C}^n$.
- Le tableau de Young associé à la valeur propre $\lambda = 0$ comporte n cases.
- Le polynôme caractéristique de A est égal à X^n , ou, ce qui revient au même, toutes les sommes des mineurs principaux de A sont nulles.
- Les valeurs propres de A sont toutes nulles.

- Les traces des puissances k -ièmes de A sont nulles pour tout $k = 1, \dots, n$.
- La matrice A est semblable à sa moitié.
- La classe de similitude de A contient la matrice nulle dans son adhérence.
- Il existe H telle que $[H, X] = 2X$.
- La matrice A est semblable à une matrice triangulaire ayant une diagonale nulle.
- La matrice A est semblable à une matrice diagonale en blocs de cellules de Jordan (associées au scalaire 0) de tailles décroissantes.
- La matrice A est limite de matrices semblables à la cellule de Jordan $J_n(0)$.
- La restriction de A à chaque sous-espace stable est non inversible.
- Il existe X tel que $A = [A, [A, X]]$.
- Une puissance de l'endomorphisme $\text{ad } A : X \mapsto [A, X]$ est nulle.

- Si A commute avec $AB - BA$, un calcul facile et la condition g) ci-dessus montrent que la matrice $AB - BA$ est alors nilpotente.
- Les matrices triangulaires supérieures de $\mathcal{C}(n)$ forment un sous-espace vectoriel maximal de dimension $n(n-1)/2$. Tout autre sous-espace vectoriel est de dimension inférieure ou égale à ce plafond [3].

5.2 Filtration et gradué associés à un endomorphisme nilpotent

Si les sommes directes sont l'outil de choix dans l'étude des endomorphismes diagonalisables ou, plus généralement, semi-simples, les filtrations sont l'outil adapté au cas nilpotent : cela se justifie par la présence de la suite des noyaux itérés.

Une filtration de E est la même chose qu'un drapeau. Le gradué associé à une filtration

$$\mathcal{F} : \{0\} \subset E_1 \subset E_2 \subset \dots \subset E_k = E$$

est l'espace vectoriel (gradué) :

$$\text{Gr}(E, \mathcal{F}) = E_1 \oplus (E_2/E_1) \oplus \dots \oplus (E_k/E_{k-1}).$$

Une filtration est stable par un endomorphisme f si chacun des sous-espaces E_i l'est : un tel endomorphisme induit un endomorphisme noté $\text{gr}(f)$ dans le gradué associé. Un calcul facile de déterminant d'une matrice triangulaire en blocs montre que le polynôme caractéristique de f est le même que celui de $\text{gr}(f)$.

Quand A est un endomorphisme nilpotent, on note $\text{Gr}(E, A)$ le gradué associé à la filtration donnée par la suite des noyaux itérés ; on a donc pour un endomorphisme nilpotent A de E , d'indice de nilpotence m :

$$\text{Gr}(E, A) = (\text{Ker } A) \oplus (\text{Ker } A^2 / \text{Ker } A) \oplus \dots \oplus (\text{Ker } A^m / \text{Ker } A^{m-1})$$

Comme l'endomorphisme A applique $\text{Ker } A^{i+1}$ dans $\text{Ker } A^i$, il induit une application linéaire de $\text{Gr}(E, A)$ dans $\text{Gr}(E, A)[-1]$, où la graduation à l'arrivée a été décalée vers la gauche de telle sorte que la i -ième composante de $\text{Gr}(E, A)[-1]$ soit la $(i-1)$ -ième de $\text{Gr}(E, A)$. Le résultat suivant est une illustration du bon usage que l'on peut faire des filtrations associées à un endomorphisme nilpotent.

Proposition 2 (Kostant).

Si A nilpotent commute avec $AB - BA$, alors AB est nilpotent.

Preuve. \diamond On établit d'abord que BA laisse invariante la filtration associée canoniquement à A : si l'on note D la dérivation d'algèbre associative donnée par le crochet avec B , alors si $D(A)$ commute avec A , on a

$$D(A^k) = kD(A)A^{k-1},$$

ce qui s'écrit encore

$$A^k B - BA^k = kABA^{k-1} - kBA^k;$$

la stabilité de $\text{Ker } A^k$ par BA en découle après simple multiplication par A à droite. On écrit ensuite le **diagramme commutatif** suivant :

$$\begin{array}{ccc} E \oplus E \cdots \oplus E & \xleftarrow{h} & \text{Gr}(E, A) \\ \downarrow v & & \downarrow \text{gr}(BA) \\ E \oplus E \cdots \oplus E & \xleftarrow{h} & \text{Gr}(E, A) \end{array}$$

L'application h est construite sur $\text{Gr}(E, A)$ à partir de la somme de toutes les applications

$$h_k : \text{Ker } A^{k+1} / \text{Ker } A^k \rightarrow E_k = E$$

avec

$$h_k(x) = A^k(x)$$

et l'application v est définie sur le k -ième facteur du produit par

$$v(x_k) = k(AB - BA)(x_k).$$

L'application horizontale h est injective et permet donc de plonger le gradué $\text{Gr}(E, A)$ dans le produit $E \times \cdots \times E$; l'endomorphisme $\text{gr}(BA)$ apparaît ainsi comme une restriction du morphisme vertical v de gauche. Son spectre est donc contenu dans le spectre de v , spectre réduit à $\{0\}$ (voir encadré dans § 5.1) ; il en est de même du spectre de BA . \diamond

Proposition 3.

Si $A(AB - BA) = 0$, alors $AB - BA$ est nilpotent.

Preuve. \diamond L'hypothèse $ABA = A^2B$ implique que l'on peut passer A à gauche de B chaque fois que B est entourée par A ; ainsi on a

$$ABA^k = A^{k+1}B$$

et de là on prouve que

$$(AB)^k = A^k B^k.$$

Enfin, pour tout k , la trace de

$$\begin{aligned} (AB - BA)^{k+1} &= (AB - BA) \cdots (AB - BA)(AB - BA) \\ &= AB \cdots AB(AB - BA) \\ &= A^{k+1}B^{k+1} - A^k B^k BA \end{aligned}$$

est clairement nulle. \diamond

5.3 Cône nilpotent.

Exemple des matrices d'ordre 8

On trouvera, sur la figure 2 le graphe de l'ensemble des classes de similitude (tiroirs) dans le cône nilpotent (l'armoire indexée par le polynôme X^n).

Les **sommets** de ce graphe représentent les différentes classes de similitude ; le **nombre** qui figure à côté d'un sommet indique la dimension de la classe de similitude en tant que sous-variété ; ce nombre ajouté à la dimension du commutant d'une matrice de la classe donne n^2 .

Les **tableaux de Young** que l'on voit indiquent le paramétrage par tableaux de Young des classes de similitude ; lu par ses colonnes, il donne la dimension des noyaux itérés des matrices d'une même classe, et lu en lignes, il donne la forme de Jordan privilégiée des matrices de la même classe.

La suite des polynômes qui apparaît ici ou là donne ce que l'on appelle la **liste des invariants de similitude** du $\mathbb{K}[X]$ -module correspondant (cf. § 10.2) ; elle se lit dans les lignes du tableau de Young ou ce qui est évidemment équivalent dans la forme de Jordan.

De manière plus précise et sur l'exemple de la classe de similitude de dimension 30, on a que le $\mathbb{K}[X]$ -module E_μ est isomorphe à

$$\mathbb{K}[X]/X \oplus \mathbb{K}[X]/X \oplus \mathbb{K}[X]/X^2 \oplus \mathbb{K}[X]/X^2 \oplus \mathbb{K}[X]/X^2.$$

On aura constaté que les dimensions des classes sont paires. Cela est un phénomène général et témoigne, d'une certaine manière, du fait plus subtil que les classes de similitude sont des variétés symplectiques. L'espace tangent en une matrice M à sa classe de similitude est isomorphe à $M(n, \mathbb{K})/\beta(M)$, où $\beta(M)$, est le noyau d'une forme bilinéaire alternée, à savoir

$$(X, Y) \mapsto \text{Tr}(M[X, Y]),$$

il est donc de codimension paire.

Deux classes de similitude (c'est-à-dire deux sommets du graphe) reliées par une arête du graphe sont telles que l'une (celle du bas) est dans l'adhérence de l'autre. C'est le cas quand le tableau de l'une s'obtient par dégradations successives (d'une case du tableau de Young) à partir de l'autre. Cela indique d'ailleurs comment se construit le graphe, en partant de la classe des éléments réguliers qui, elle, est dense dans le cône nilpotent. La chute d'une case d'un niveau au suivant correspond à la perte de deux dimensions pour la classe (on pensera à la manière avec laquelle se calcule la dimension du commutant, cf. § 8). L'adhérence d'une classe de similitude se lit comme la réunion de toutes les classes qui sont à un niveau qui lui est inférieur (ou égal) et qui lui sont connectées.

Il est facile de comprendre maintenant le résultat suivant en rapportant une matrice quelconque M dans son armoire : si $M = s + n$ est la décomposition de Dunford-Jordan (en semi-simple + nilpotent) de M , la matrice s est dans l'adhérence de la classe de similitude de M .

5.4 Cône nilpotent et classes de similitude en dimension 2

Soit $A \in M(2, \mathbb{R})$. La classe de similitude de A est à une translation près la même que celle de $M = A - \frac{1}{2} \text{Tr}(A)I$, qui, elle, est de trace nulle. Une représentation géométrique est dès lors envisageable puisque l'ensemble $\mathfrak{sl}(2, \mathbb{R}) = \left\{ M = \begin{bmatrix} x & y+z \\ y-z & -x \end{bmatrix} \right\}$ des matrices de trace nulle est un espace vectoriel de dimension 3. Le cône nilpotent est alors donné par le cône usuel d'équation $x^2 + y^2 - z^2$. Son sommet représente l'orbite de la matrice nulle, et

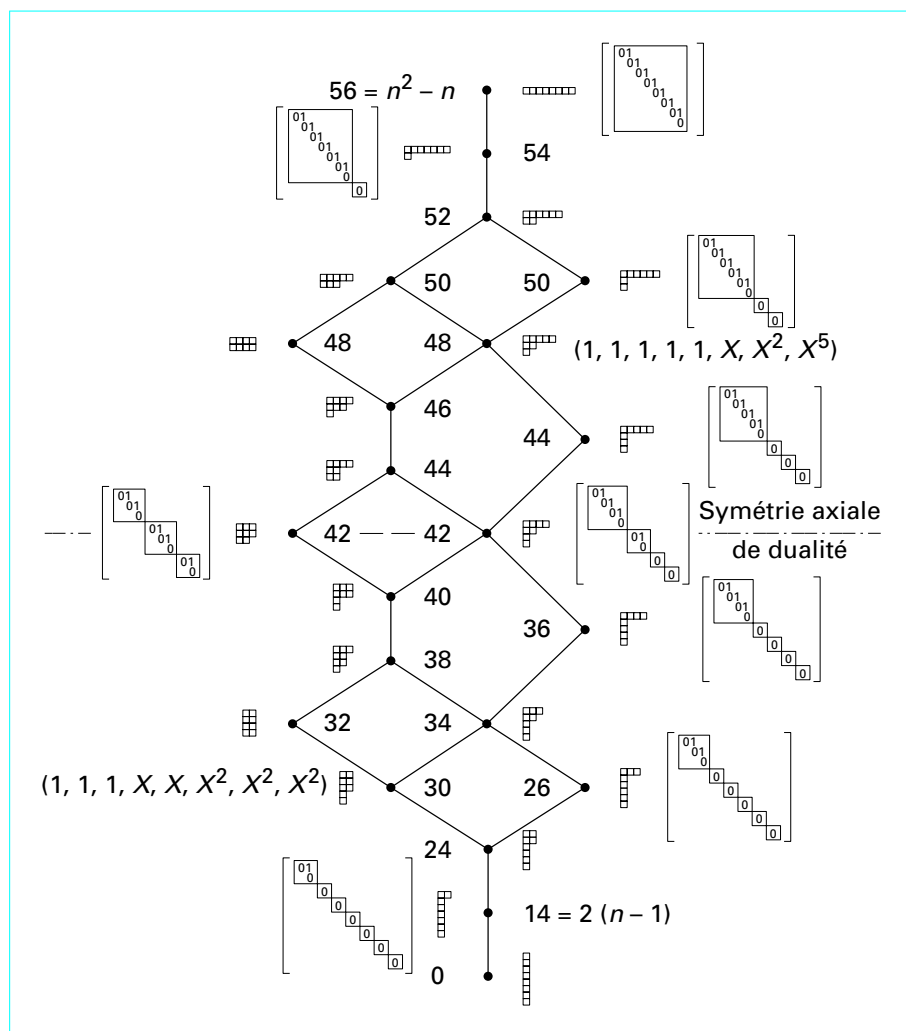


Figure 2 – Graphe des orbites nilpotentes dans $M(8, \mathbb{C})$

le cône épointé est à lui seul une classe de similitude, celle de

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ par exemple.}$$

Les points situés en dehors du cône (il s'agit des matrices dont les valeurs propres sont réelles non nulles) se répartissent en classes de similitudes qui sont des hyperboloïdes à une nappe (asymptotes au cône) ; elles sont indexées par les matrices **symétriques**

$$\begin{bmatrix} 0 & a \\ a & 0 \end{bmatrix}, \text{ avec } a > 0.$$

Quant aux points situés à l'intérieur du cône, leurs classes de similitude sont des hyperboloïdes à deux nappes (toujours asymptotes au cône) ; elles sont indexées par les matrices **antisymétriques**

$$\begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}, \text{ avec } a < 0.$$

Quelques classes représentatives sont visibles dans la figure 3.

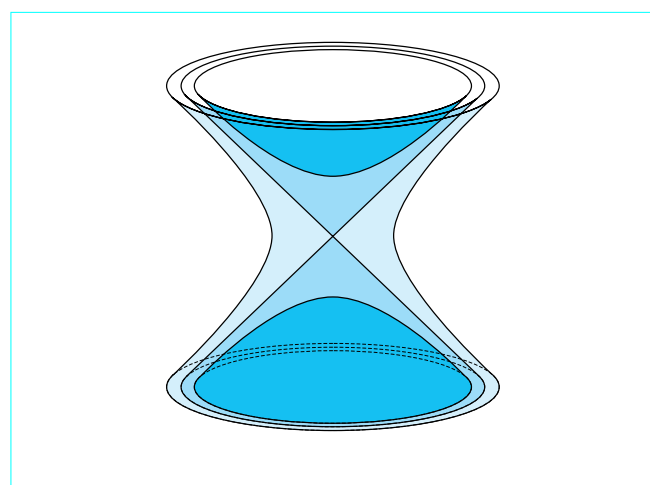


Figure 3 – Classes de similitude en dimension 2

6. La jordanisation pour elle-même

6.1 Pratique et preuve

La démonstration du théorème de Jordan pour une matrice nilpotente A d'ordre n est facile si on prend le soin de faire usage du tableau de Young correspondant qui comporte n cases. On désigne par m l'indice de nilpotence de A .

On commence par choisir des vecteurs v_1, \dots, v_p dans E qui relèvent une base du quotient $E/\text{Ker } A^{m-1}$, c'est-à-dire

$$p = \dim E - \dim \text{Ker } A^{m-1}$$

vecteurs de $\text{Ker } A^m$ qui sont indépendants modulo $\text{Ker } A^{m-1}$.

On place ces vecteurs dans les cases de la dernière colonne du tableau de Young de A . À chacun de ces vecteurs, on applique successivement les puissances de A de façon à remplir, en allant de droite à gauche, la ligne de longueur m qui y aboutit.

Une fois ces cases remplies, on procède de même avec la restriction de A à $\text{Ker } A^{m-1}$, dont le tableau est celui de A auquel on a ôté la dernière colonne. Simplement ici, les cases de la dernière colonne (c'est-à-dire de l'avant-dernière colonne du tableau de A) sont déjà en partie occupées par la famille **libre** des vecteurs $A(v_1), \dots, A(v_p)$; on fait donc appel au nombre (éventuellement nul) de vecteurs w_i qu'il faut pour remplir cette colonne avec des vecteurs qui s'appliquent sur une base du quotient $\text{Ker } A^{m-1}/\text{Ker } A^{m-2}$. On applique les puissances de A à ces nouveaux vecteurs et l'on remplit les lignes qui y aboutissent.

On répète le procédé jusqu'à remplir ainsi l'ensemble des cases du tableau de Young.

$A^3 v_1$	$A^2 v_1$	$A v_1$	v_1
$A^3 v_2$	$A^2 v_2$	$A v_2$	v_2
$A^2 w_1$	$A w_1$	w_1	
u_1			
u_2			

La famille des n vecteurs figurant dans les cases du tableau de Young de A forme une base de E , car ces vecteurs sont linéairement indépendants : en effet, en appliquant A^{m-1} à une combinaison linéaire de ces vecteurs, il est aisé de prouver que les coefficients des vecteurs de la dernière colonne sont nuls ; une fois débarrassés dans la combinaison de ces vecteurs, on applique à la combinaison restante l'endomorphisme $\text{Ker } A^{m-2}$ et l'on recommence.

6.2 Arrangements possibles d'une base de Jordan

La jordanisation classique consiste à numéroter les vecteurs, figurant dans le tableau, de gauche à droite et de haut en bas.

1 ^{er}	2 ^e	3 ^e	4
5	6	7	8
9	10	11	
12			
13			

Mais on peut faire autrement. Avant ce faire, remarquons que ce qui est extraordinaire avec les vecteurs trouvés, c'est qu'ils donnent une base de $\text{Ker } A^i$ pour tout i , mais aussi une base de $\text{Im } A^i$ et même une base de $\text{Ker } A^i \cap \text{Im } A^i$. Ainsi par exemple, les vecteurs qui restent, une fois mis de côté tous les vecteurs figurant en bout des lignes du tableau de Young, fournissent une base de $\text{Im } A$ et l'on comprend mieux pourquoi les sauts de dimension sont donnés par la dimension de $\text{Ker } A \cap \text{Im } A^i$ (cf. § 4.1).

Un arrangement utile est donné en ordonnant les vecteurs de base (inscrits dans le tableau de Young) en colonnes et de bas en haut.

5	8	11	13
4	7	10	12
3	6 ^e	9	
2 ^e			
1 ^{er}			

Cette pseudo-jordanisation donne une réduction qui s'avère entre autre efficace dans l'étude du centralisateur $Z(A)$ [11]. Tous les blocs de cette réduction qui figurent en dehors de la surdiagonale sont nuls. Ceux sur la surdiagonale sont des blocs rectangulaires de la forme $[0, I_k]$ (à lire en lignes), où k est l'ordre du bloc diagonal figurant juste en dessous.

Exemple. Soit la classe de similitude définie par le tableau de Young $\begin{smallmatrix} \square & \square & \square \\ \square & \square & \square \end{smallmatrix}$. Si l'on ordonne la base comme indiqué juste avant, soit :

v_2	v_3	v_4	v_5
v_1			

la matrice est y donnée alors par

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

. Nous allons

prouver que la classe correspondant au tableau $\begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix}$ (obtenu au moyen d'une dégradation) est dans l'adhérence de la précédente; il s'agit de faire en sorte que le vecteur v_5 descende à la deuxième ligne et ait une image égale à v_1 . Pour cela, on considère les matrices

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \varepsilon \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

qui sont dans la classe initiale et qui, lorsque ε tend

vers 0, convergent vers une matrice qui est dans la seconde.

Cette méthode expliquée sur l'exemple précédent donne une démonstration du principe de dégradations successives.

7. Familles particulières de matrices. Les matrices de la classe δ

Le sous-espace $\text{Ker } M^2$ a une dimension qui varie entre deux bornes : la dimension de $\text{Ker } A$ et son double. Les matrices telles que

$$\text{Ker } A = \text{Ker } A^2$$

et les matrices telles que

$$\dim(\text{Ker } A^2) = 2 \dim(\text{Ker } A)$$

forment deux familles qui méritent une attention particulière. Les propositions suivantes donnent quelques propriétés caractéristiques de l'une ou l'autre de ces deux familles. Il est bon de garder à l'esprit la suite exacte

$$\{0\} \rightarrow \text{Ker } A \rightarrow \text{Ker } A^2 \xrightarrow{A} \text{Ker } A \cap \text{Im } A \rightarrow \{0\}.$$

7.1 Une première famille

Proposition 4.

Soit A une matrice complexe non inversible. Les propriétés suivantes sont équivalentes.

- $\text{Ker } A^2 = \text{Ker } A$.
- Le tableau de Young $\text{TY}(A; 0)$ de A relatif à la valeur propre nulle comporte une seule colonne.
- 0 est racine simple du polynôme minimal.
- La suite des noyaux itérés s'arrête à $\text{Ker } A$.
- La multiplicité géométrique de la valeur propre 0 est égale à sa multiplicité algébrique.
- La décomposition de Jordan de A dans sa partie associée à la valeur propre 0 ne comporte que des cellules de taille 1.
- $\text{rg}(A) = \text{rg}(A^2)$.
- L'espace E est somme directe de $\text{Ker } A$ et de $\text{Im } A$.
- La matrice A est diagonalisable par rapport à sa valeur propre 0 .
- La matrice A s'écrit comme produit de deux matrices singulières qui sont des polynômes en A .
- La matrice A admet une racine qui est un polynôme en A .
- La matrice A est annulée par un polynôme qui admet 0 comme racine simple.
- La matrice A est semblable à une matrice $\begin{bmatrix} 0 & 0 \\ 0 & P \end{bmatrix}$, où P est inversible.

Il est clair qu'une matrice est diagonalisable si, et seulement si, pour toute valeur propre λ , la matrice $A - \lambda I$ vérifie les propriétés précédentes. Par ailleurs, il est intéressant de noter la proposition suivante.

Proposition 5.

Soit A une matrice complexe. Les propriétés suivantes sont équivalentes.

- L'endomorphisme $\text{ad}(A) : X \mapsto AX - XA$ vérifie les conditions précédentes.
- La matrice A est diagonalisable.
- L'endomorphisme $\text{ad}(A)$ est diagonalisable.
- Tout élément de déterminant nul dans $\mathbb{C}[A]$ est produit de deux éléments de déterminant nul de $\mathbb{C}[A]$.

7.2 Une seconde famille

L'autre famille est décrite dans la proposition suivante.

Proposition 6.

Soit A une matrice complexe non inversible d'ordre n . Les propriétés suivantes sont équivalentes, et caractérisent les matrices dites de **type** δ :

- $\dim \text{Ker } A^2 = 2 \dim \text{Ker } A$.

b) Les deux premières colonnes du tableau de Young $\text{TY}(A; 0)$ sont de même longueur.

c) Les cellules de Jordan de A associées à la valeur propre 0 sont toutes de taille supérieure ou égale à 2.

d) La matrice A n'est pas semblable à une matrice de la forme

$$\begin{bmatrix} 0 & 0 \\ 0 & A' \end{bmatrix}, \text{ où } A' \text{ est d'ordre } n-1.$$

e) $\text{Ker } A \subset \text{Im } A$.

f) Si $A = A_1 \cdots A_{s+1}$, où $s = \dim(\text{Ker } A)$ et les A_i commutent avec A , alors l'une au moins des matrices A_i est inversible.

g) Aucune droite du noyau n'admet de supplémentaire stable dans E .

8. Calcul de la dimension du commutant

Ce calcul sert pour le calcul de la dimension des orbites. La méthode consiste à trouver pour une matrice nilpotente la dimension de son commutant car cela suffit vu le lemme suivant.

Lemme 2. Si $M = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ est une matrice diagonale en blocs

vérifiant que les polynômes caractéristiques de A et de B sont premiers entre eux, alors les matrices qui commutent avec M

sont diagonales en blocs, soit $\begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix}$ avec

$$AX = XA \text{ et } BY = YB.$$

Preuve. \diamond L'équation $AY = YB$ implique que

$$P(A)Y = YP(B)$$

pour tout polynôme P , en particulier pour $P = \chi_A(X)$; on conclut, grâce au théorème de Cayley-Hamilton et au fait que $\chi_A(B)$ est inversible, car ses valeurs propres sont toutes non nulles. \diamond

Proposition 7.

Pour une matrice nilpotente N , la dimension du commutant est égale à la somme des carrés des longueurs des colonnes de son tableau de Young.

Preuve. \diamond Il s'agit de déterminer le nombre de degrés de liberté ℓ dans le choix d'un opérateur M qui commute avec A . Dans une base de jordanisation placée dans le tableau de Young, les images des vecteurs situés dans la dernière colonne sont totalement libres; par contre, ceux qui en découlent par itération par A et qui se situent sur leurs lignes sont complètement déterminés par la condition

$$M A^k(v) = A^k M(v).$$

Si k est le nombre des lignes les plus longues, on a ainsi kn degrés de liberté, qu'on répartit en plaçant k dans chaque case du tableau :

2	2	2	$k = 2$
2	2	2	2
2	2		
2			

L'étape suivante consiste à compter les degrés de liberté dans le choix des vecteurs situés à l'extrémité droite des ℓ lignes de

longueur m suivante. Ces vecteurs sont dans $\text{Ker } A^m$ qui doit être laissé stable par M , on a donc $\dim(\text{Ker } A^m) \ell$ degrés de liberté nouveaux qu'on répartit en plaçant ℓ dans chaque case des m premières colonnes :

2 + 1	2 + 1	2	2
2 + 1	2 + 1	2	2
2 + 1	2 + 1		
2 + 1			

et ainsi de suite,

2 + 1 + 1	2 + 1	2	2
2 + 1 + 1	2 + 1	2	2
2 + 1 + 1	2 + 1		
2 + 1 + 1			

en se plaçant dans les cases des tableaux de Young.

◇

a) Il est facile de voir que la dimension du commutant de la matrice $M \in M(n, \mathbb{K})$ est toujours supérieure ou égale à n , de codimension paire et dont les deux plus hautes valeurs sont :
 n^2 (cas des matrices scalaires) ;
 $(n-1)^2 + 1$ (cas des matrices semblables à

$$\lambda I_{n-1} \oplus \mu I_1, \text{ avec } \lambda \neq \mu$$

et cas de $\lambda I_n + N$ avec N nilpotente de rang 1).

Enfin, il est facile de voir que la dimension du commutant est n si, et seulement si, la matrice est monogène (polynôme minimal égal au polynôme caractéristique).

b) Le bicommutant de M est réduit à l'ensemble des polynômes en M (cf. § 10.1).

9. Réduction simultanée

On entend par cela tout résultat qui réduit deux ou plusieurs endomorphismes avec une même matrice de passage.

9.1 Cas de deux matrices

Le plus connu est le fait que si A et B sont deux matrices complexes qui commutent, alors elles ont un même vecteur propre en commun et par suite, il existe une **même** base où elles ont toutes les deux la forme triangulaire. Toute matrice de la forme

$$P(A, B)(AB - BA),$$

où P est un polynôme (non commutatif) en deux variables, est alors nilpotente.

Le théorème suivant éclaire mieux les choses.

Théorème 7. Si $P(A, B)(AB - BA)$ est nilpotent pour tout polynôme P , alors A et B sont simultanément trigonalisables (cf. référence dans [16]).

Un cas particulier de ce résultat et qui se démontre directement et aisément est que si $AB = 0$, alors A et B sont simultanément trigonalisables. Il en est de même si A et B commutent avec $AB - BA$. On a encore une trigonalisation simultanée pour tout couple de matrices telles que $\text{rg}(AB - BA) \leq 1$ [16].

Le problème de la réduction simultanée de deux matrices quelconques est difficile et a suscité beaucoup d'articles. D'ailleurs, on démontre qu'il n'est pas possible d'obtenir de classification, pour les classes de similitude simultanée de couples de matrices (A, B) , comme celle dont on dispose dans le cas d'une seule matrice.

Un problème apparenté, plus simple, fut étudié et résolu par Kronecker : il s'agit de classer les classes de r -équivalence simultanée. Il est clair que si $A = PBQ$ et $A' = PB'Q$, alors on a $A'A^{-1}$ semblable à $B'B^{-1}$. De plus (A, I) est simultanément r -équivalent à (B, I) si, et seulement si, A et B sont semblables. La solution du problème de Kronecker contient la solution du problème de la réduction d'une matrice qui nous a occupé dans cet article.

9.2 Théorèmes de Engel et de Lie

Théorème 8 (Engel). Soit \mathfrak{g} une sous-algèbre de Lie de $M(n, \mathbb{K})$, c'est-à-dire un sous-espace vectoriel de matrices stable par crochet. On suppose que tous les éléments de \mathfrak{g} sont nilpotents. Il existe alors une base où tous les éléments de \mathfrak{g} sont triangulaires supérieures avec des 0 sur la diagonale.

Le résultat suivant en est une généralisation, sur un corps algébriquement clos de caractéristique nulle. On définit pour une sous-algèbre de Lie \mathfrak{g} de $M(n, \mathbb{K})$ la sous-algèbre dérivée $[\mathfrak{g}, \mathfrak{g}]$ comme le sous-espace vectoriel engendré par les crochets des éléments de \mathfrak{g} . On notera $\mathfrak{g}^{(m)}$ la sous-algèbre $[\mathfrak{g}^{(m-1)}, \mathfrak{g}^{(m-1)}]$, avec la convention

$$\mathfrak{g}^{(1)} = [\mathfrak{g}, \mathfrak{g}].$$

Théorème 9 (Lie). Les éléments d'une même sous-algèbre de Lie résoluble \mathfrak{g} de $M(n, \mathbb{C})$ (c'est-à-dire telle que $\mathfrak{g}^{(k)} = \{0\}$ pour un certain k) sont simultanément trigonalisables.

9.3 Théorème de Kolchin

On dispose d'une version groupe du théorème de Engel.

Théorème 10 (Kolchin). Les éléments d'un sous-groupe G de $\text{GL}(n, \mathbb{C})$ qui est formé d'éléments unipotents sont simultanément trigonalisables. Autrement dit, le groupe G est conjugué à un sous-groupe du groupe des matrices triangulaires supérieures ayant des 1 sur la diagonale.

9.4 Diagonalisation simultanée

La diagonalisation simultanée d'une famille commutative d'endomorphismes diagonalisables est encore possible.

De ce fait, si G est un sous-groupe commutatif fini de $\text{GL}(n, \mathbb{C})$, les éléments de G sont simultanément diagonalisables. Le groupe G est conjugué à un sous-groupe du groupe des matrices diagonales inversibles.

On aura noté qu'un élément d'ordre fini dans $GL(n, \mathbb{C})$ est diagonalisable car il annule un polynôme du type $X^m - 1$ qui est à racines distinctes. On peut invoquer aussi le fait que ses valeurs propres étant des racines de l'unité, sa partie semi-simple est comme lui d'ordre fini : il en est donc également de sa partie unipotente, qui est donc égale à l'identité. La décomposition de Dunford-Jordan $M = S + N$ a une version multiplicative lorsque M est inversible :

$$M = S(I + S^{-1}N) = SU,$$

où U est unipotente et commute à S . La matrice U s'appelle alors la partie unipotente de M .

Si maintenant \mathfrak{g} est une sous-algèbre de Lie de $M(n, \mathbb{C})$ formée d'éléments diagonalisables, il existe une base qui les réduit simultanément ; en particulier, elle est commutative.

9.5 Réduction simultanée et théorème de Sylow

On termine ce paragraphe par un exemple de réduction simultanée déduit du théorème de Sylow. Le corps ici est exceptionnellement un corps fini.

Proposition 8.

Soit G un sous-groupe de $GL(n, \mathbb{F}_p)$ de cardinal $p^{n(n-1)/2}$. Alors les éléments de G sont unipotents et simultanément trigonalisables.

10. Autre point de vue sur la jordanisation. La version $\mathbb{K}[X]$ - modules

Soit E un espace vectoriel de dimension finie n sur \mathbb{K} et soit $M \in \text{End}(E)$. L'espace E , muni de M , a une structure naturelle de $\mathbb{K}[X]$ - module induite par la multiplication externe

$$P(X) \cdot v = P(M)(v).$$

L'espace E muni de cette structure de $\mathbb{K}[X]$ - module est noté usuellement E_M . Il est facile de voir que ce module est un module de torsion de type fini sur l'anneau principal $\mathbb{K}[X]$. De tels modules sur les anneaux principaux ont une structure bien connue. Par ailleurs, on a la proposition suivante.

Proposition 9.

Les deux $\mathbb{K}[X]$ - modules E_M et E_N sont isomorphes si, et seulement si, M et N sont semblables.

10.1 Retour sur le commutant. Application au bicommutant

Le commutant d'une matrice M se voit de ce point de vue comme l'algèbre des endomorphismes du $\mathbb{K}[X]$ - module E_M . Le bicommutant de M n'est autre que le centre de cette algèbre.

Proposition 10.

Le bicommutant de M est réduit aux polynômes en M .

Preuve. \diamond Le $\mathbb{K}[X]$ - module E_M s'écrit

$$\mathbb{K}[X]/(P_1) \oplus \dots \oplus \mathbb{K}[X]/(P_n),$$

où P_i divise P_{i+1} pour $i \in [1, n-1]$.

Un endomorphisme du $\mathbb{K}[X]$ - module E_M est caractérisé par ses blocs opérant entre les différents facteurs de la somme directe. Maintenant, un homomorphisme entre deux blocs est entièrement déterminé par le choix de l'image de la classe du polynôme constant 1, laquelle image est nécessairement annulée par le polynôme engendrant l'idéal apparaissant dans le facteur de départ. Cela nous laisse sans contrainte si le polynôme de départ est multiple de celui de l'arrivée et, dans le cas inverse, l'image considérée doit se trouver dans le sous-module engendré par la classe du quotient de ces deux polynômes.

Un élément du centre de $\text{End}_{\mathbb{K}[X]}(E_M)$ doit commuter en particulier avec les éléments qui sont nuls sur tous les blocs sauf le dernier et qui sont quelconques sur le dernier.

Un calcul similaire à celui de l'ensemble des matrices qui commutent sur un corps avec les matrices qui sont nulles sur l'hyper-plan des $n-1$ premières coordonnées nous convainc que les éléments du centre sont les endomorphismes scalaires de notre $\mathbb{K}[X]$ - module. \diamond

10.2 Facteurs invariants et forme normale de Smith

Théorème 11. Deux matrices de $M(n, \mathbb{K})$ sont semblables si, et seulement si, il existe deux matrices P et Q à coefficients polynomiaux, de déterminants constants non nuls, et tels que

$$A - XI = P(B - XI)Q.$$

L'idée dans cet énoncé est que la base canonique (e_1, \dots, e_n) de $E = \mathbb{K}^n$ en tant que \mathbb{K} - espace vectoriel, n'est plus qu'un système de générateurs de E en tant que $\mathbb{K}[X]$ - module. Les relations entre ces générateurs sont données comme on l'imagine par les écritures $Xe_j - \sum_i a_{ij} e_i$ qui correspondent à l'endomorphisme de E défini par

la matrice A . Ces relations sont des vecteurs de $\mathbb{K}[X]^n$, et y engendrent un sous-module (noyau) identique à l'image de l'homomorphisme de $\mathbb{K}[X]^n$ défini par la matrice $XI - A$.

Le théorème résulte du lemme suivant, qui s'établit par une chasse de diagramme élémentaire.

Lemme 3. Soit R un anneau commutatif. Deux homomorphismes de R^n dans R^m sont entrelacés par des automorphismes de R^m et R^n si, et seulement si, leurs conoyaux sont isomorphes.

Les matrices qui, comme P ou Q , ont un déterminant constant non nul sont les matrices inversibles dans l'anneau des matrices à coefficients dans $\mathbb{K}[X]$. Deux matrices, à coefficients dans l'anneau, entrelacées par des matrices inversibles seront dites φ -équivalentes.

Mais sur un anneau principal R tel que $\mathbb{K}[X]$, on dispose pour les matrices à coefficients dans R de la **forme normale de Smith**, qui exprime qu'une matrice M , y est équivalente à une matrice diagonale

$$\text{Diag}(s_1, \dots, s_r, 0, \dots),$$

où chaque s_i divise le suivant. Les s_i s'appellent les facteurs invariants de M et $s_1 \dots s_i$ est égal au PGCD des mineurs d'ordre i de M .

On peut énoncer le théorème suivant.

Théorème 12. Les matrices A et B de $M(n, \mathbb{K})$ sont semblables si, et seulement si, les matrices $A - XI$ et $B - XI$ ont les mêmes facteurs invariants, P_1, P_2, \dots, P_n . Ces facteurs invariants vérifient

$$P_1(X) \mid P_2(X) \mid \dots \mid P_n(X).$$

Le polynôme $P_n(X)$ est le polynôme minimal commun à A et B et le produit des P_i est égal à leur polynôme caractéristique.

a) L'énoncé sur les s_i s'établit facilement d'abord pour s_1 et s'obtient pour les autres en passant par les puissances extérieures.

b) Le polynôme minimal est donné, vu le théorème 12, comme le quotient du polynôme caractéristique par le PGCD des coefficients de la comatrice de $A - XI$.

c) Sur un anneau euclidien tel que $\mathbb{K}[X]$, les facteurs invariants dans la forme de Smith s'obtiennent de façon effective par des opérations sur lignes et colonnes combinées avec l'algorithme d'Euclide.

d) Les polynômes $P_i(X)$ apparaissant dans la preuve de la proposition 10 et liés à la structure du $\mathbb{K}[X]$ -module de type fini E_M sont les mêmes que les facteurs invariants issus de la forme normale de Smith.

11. Matrices de Hessenberg

11.1 Généralités

Une matrice de Hessenberg est une matrice telle que tous les termes au dessous de sa sous-diagonale sont nuls. Elle est dite H-régulière si tous les termes de sa sous-diagonale sont non nuls. Elle est évidemment triangulaire (supérieure) si tous les coefficients de sa sous-diagonale sont nuls.

L'intérêt pratique des matrices de Hessenberg est double :

- d'abord, leurs déterminants, comme leurs polynômes caractéristiques, se calculent facilement ;
- surtout, toute matrice est (de façon effective et quel que soit le corps de base) semblable à une matrice de Hessenberg.

Ce dernier point s'établit grâce à des opérations élémentaires simultanées (c'est-à-dire de façon à rester dans la même classe de similitude) sur les lignes et les colonnes : pour peu que la première colonne soit non nulle (sinon on passe à la suivante), on peut, par une permutation éventuelle (simultanée) des lignes et colonnes, supposer que le terme b_{21} est non nul. A partir de là, on annule, par des opérations élémentaires sur les lignes, tous les termes de la première colonne qui se trouvent en dessous et on effectue au fur et à mesure les opérations (simultanées) sur les colonnes.

Comme, durant ce processus, la première ligne n'est pas modifiée mais surtout n'intervient pas, les opérations simultanées correspondantes sur les colonnes n'affectent pas la première colonne et l'on se retrouve avec une matrice semblable à celle de départ mais ayant une première colonne comme il faut (on notera que c'est pour cela que l'on ne peut obtenir la trigonalisation – qui suppose d'ailleurs le corps de base algébriquement clos – en usant d'un processus analogue).

Le processus poursuivi établit notre propos. Cela donne, par exemple, avec des matrices d'ordre 3 les écritures suivantes, où l'on a supposé $d \neq 0$:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} \rightarrow \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & h - e(g/d) & k - f(g/d) \end{bmatrix} \\ \rightarrow \begin{bmatrix} a & b + c(g/d) & c \\ d & e + f(g/d) & f \\ 0 & h - eg/d + (k - fg/d)/(g/d) & k - fg/d \end{bmatrix}$$

Une matrice est semblable à une matrice de Hessenberg H-régulière si, et seulement si, son polynôme minimal est égal à son polynôme caractéristique (c'est-à-dire si elle est monogène) et que toute matrice complexe est semblable à une matrice de Hessenberg dont le premier terme (ou, si l'on veut, le dernier) de sa sous-diagonale est nul (on considérera une droite ou un hyperplan stable).

11.2 Calcul du déterminant

Le calcul du déterminant est fondé sur une formule simple de calcul du déterminant d'une matrice dont les coefficients de la dernière ligne sont nuls sauf peut-être les deux derniers. On aura besoin pour l'exprimer d'introduire une application bilinéaire alternée naturelle de

$$\mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}^{n-1}.$$

Pour un vecteur colonne \vec{V} de $\mathbb{R}^n = \mathbb{R}^{n-1} \oplus \mathbb{R}$, on écrit $\vec{V} = \vec{V}_1 + u_n$

et l'on définit $|\vec{V}, \vec{W}|$ comme le vecteur $w_n \vec{V}_1 - u_n \vec{W}_1$.

Lemme 4. Soit M une matrice carrée d'ordre n dont les coefficients de la dernière ligne sont nuls sauf, éventuellement, les deux derniers.

Si l'on note A la matrice rectangulaire $(n-1) \times (n-2)$ obtenue en rayant la dernière ligne et les deux dernières colonnes \vec{V} et \vec{W} de M , alors le déterminant de M est égal au déterminant de la matrice d'ordre $n-1$ obtenue en adjoignant à la matrice A une dernière colonne donnée par $|\vec{V}, \vec{W}|$.

Ainsi, par exemple, le polynôme caractéristique de la matrice

$$B = \begin{bmatrix} x & y & z & t \\ a & u & v & w \\ 0 & b & r & s \\ 0 & 0 & c & m \end{bmatrix}$$

est donné par

$$\chi_B(X) = \det \begin{bmatrix} x-X & y & z(m-X) - ct \\ a & u-X & v(m-X) - cw \\ 0 & b & (r-X)(m-X) - cs \end{bmatrix} \\ = \det \begin{bmatrix} x-X & y[(r-X)(m-X) - cs] - b[z(m-X) - ct] \\ a & (u-X)[(r-X)(m-X) - cs] - b[v(m-X) - cw] \end{bmatrix}$$

Les matrices de Hessenberg sont également utiles dans d'autres situations. Ainsi, elles permettent par exemple d'établir

facilement que la matrice $\begin{bmatrix} x & 0 & 0 & 0 \\ 1 & y & 0 & 0 \\ 0 & 2 & z & 0 \\ 0 & 0 & 3 & t \end{bmatrix}$ est diagonalisable si, et

seulement si, x, y, z et t sont distincts deux à deux.

De même, il est immédiat de voir que les matrices (dépen-

dants de quatre paramètres indépendants) $\begin{bmatrix} x & xy & 0 & 0 \\ 1 & y & 0 & 0 \\ 0 & 1 & z & zt \\ 0 & 0 & 1 & t \end{bmatrix}$

ne sont jamais diagonalisables (car ce sont des matrices de Hessenberg H-régulières qui ont 0 comme valeur propre au moins d'ordre deux).

12. Le cas réel

12.1 Généralités

L'examen de la place occupée par les matrices réelles dans la partition de $M(n, \mathbb{C})$ en armoires prouve que seules les armoires associées à des polynômes caractéristiques réels contiennent des matrices réelles. Les valeurs propres de telles matrices sont réelles ou complexes conjuguées et pour λ , valeur propre non réelle, les multiplicités algébriques de λ et de $\bar{\lambda}$ sont égales. De plus, pour A complexe

$$\chi_{\bar{A}}(X) = \overline{\chi_A(X)}$$

et le système de tableaux de Young qui paramètre la classe de similitude de \bar{A} se déduit de celui de A en prenant pour tableau associé à $\bar{\lambda}$ celui de λ .

Si maintenant A est semblable à une matrice réelle, elle est nécessairement semblable à sa conjuguée ; son tiroir contiendra donc des matrices réelles si, et seulement si, le système de tableaux de Young associé est tel que pour toute valeur propre λ non réelle,

$$\text{TY}(A; \lambda) = \text{TY}(A; \bar{\lambda}).$$

Enfin, deux matrices réelles d'un même tiroir complexe sont dans un même tiroir réel (autrement dit, si deux matrices réelles sont semblables sur \mathbb{C} , alors elles sont semblables sur \mathbb{R}).

Exemple : la matrice

$$\begin{bmatrix} i & 1 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \end{bmatrix}$$

a un polynôme caractéristique réel mais n'est semblable à aucune matrice réelle (on explicitera le polynôme minimal ou bien comparer les rangs de $A - iI$ et de $A + iI$).

12.2 De certaines matrices monogènes

Soit M est une matrice réelle que l'on a réduite, sur \mathbb{C} , sous forme de Jordan. La réduire sur \mathbb{R} revient essentiellement à examiner les matrices réelles qui sont dans la classe de similitude de la matrice d'ordre $2n$ donnée par

$$J_n(\lambda) \oplus J_n(\bar{\lambda}),$$

où λ est non réel. Ces matrices sont monogènes et admettent

$$X^2 - (2\text{Re}(\lambda)X + |\lambda|^2)^n$$

pour polynôme caractéristique ou polynôme minimal. La matrice compagnon associée est déjà une matrice réelle de leur classe de similitude. Toute matrice réelle monogène M ayant le même polynôme caractéristique convient également.

On se propose de rechercher de telles matrices qui soient de la forme

$$M = \begin{bmatrix} A & B & 0 & \dots & 0 \\ 0 & A & B & \dots & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & A & B \\ 0 & 0 & \dots & 0 & A \end{bmatrix},$$

où A est une matrice réelle d'ordre deux ayant

$$\lambda = a + ib \text{ et } \bar{\lambda} = a - ib$$

comme valeurs propres.

Remarquons, d'abord, que toutes les matrices A soumises à ces conditions forment une seule classe de similitude, dans laquelle on

trouve par exemple les matrices $\begin{bmatrix} a-b & \\ b & a \end{bmatrix}$ et $\begin{bmatrix} 0 & -a^2-b^2 \\ 1 & 2a \end{bmatrix}$, mais

où ne figurent ni matrice triangulaire supérieure (ou inférieure), ni matrice symétrique. Cette classe de similitude, qui est géométriquement un hyperboloïde à deux nappes, sera notée $\mathcal{H}_2(a+ib)$.

Par ailleurs, le polynôme minimal de M est de degré pair, car il est réel et divise le polynôme caractéristique. Aussi, la condition sur M sera assurée dès lors que B est choisie de telle sorte que la matrice

$$M^2 - \text{Tr}(A)M + \det(A)I_{2n}$$

ait une puissance $(n-1)$ -ième non nulle ou, ce qui revient au même, que la matrice

$$AB + BA - \text{Tr}(A)B$$

soit non nilpotente!

Il est facile alors, en regardant les valeurs propres ou en invoquant la trace, de voir que B peut être choisie égale à I_2 , à A elle-même, à

la matrice $X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, à $Y = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ ou enfin à $X - Y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

De plus, si

$$A = A(a, b) = \begin{bmatrix} a-b & \\ b & a \end{bmatrix},$$

on peut prendre B de trace nulle non symétrique quelconque.

12.3 Jordanisation réelle

Théorème 13. Les classes de similitude réelles d'une même armoire à polynôme caractéristique réel sont paramétrées par les systèmes de tableaux de Young où pour toute racine non réelle λ , on a

$$TY(A; \lambda) = TY(A; \bar{\lambda}).$$

Pour un tel système, la réduite de Jordan réelle peut être choisie de telle sorte qu'elle soit une somme de blocs de Jordan habituels associés aux valeurs propres réelles auxquels on ajoute, pour les valeurs propres complexes $\lambda = a + ib$ non réelles, des blocs de taille paire n'ayant que des matrices $A \in \mathcal{H}_2(a+ib)$ sur la diagonale **et** sur la surdiagonale et des zéros partout ailleurs.

D'autres choix sont possibles. On peut, en effet, remplacer les matrices A figurant sur la surdiagonale par des matrices B **toutes** égales à l'identité, ou bien à la matrice X , ou Y ou $X - Y$.

En fait, on a un tableau de Young associé à chaque facteur irréductible du polynôme caractéristique. Dans la jordanisation réelle, le tableau de Young associé à un tel facteur de degré deux générera, pour chacune de ses lignes, un bloc de Jordan de taille double, ayant (dans le choix usuel) les mêmes matrices $A = A(a, b)$ sur la diagonale, la matrice I_2 le long de la surdiagonale et des zéros partout ailleurs.

a) Le tiroir de haut d'une armoire associée à un polynôme unitaire de degré n réel contient toujours un sous-tiroir réel (considérer la matrice compagnon de ce polynôme) et le tiroir du bas également. Le sous-tiroir du bas est constitué des matrices semi-simples réelles de l'armoire en question.

b) Les tiroirs d'une même armoire contiennent tous des sous-tiroirs réels si, et seulement si, les racines complexes non réelles du polynôme caractéristique commun à toutes les matrices sont, quand elles existent, simples.

12.4 Graphe d'une armoire à polynôme caractéristique réel et dans $M(8, \mathbb{C})$

Considérons l'armoire \mathcal{A} correspondant dans $M(8, \mathbb{C})$ au polynôme unitaire réel $(X-1)^2(X^2+1)^3$.

Le nombre de classes de similitudes (ou tiroirs) contenues dans \mathcal{A} est égal à $p(2)p(3) = 2 \times 3^2 = 18$. Chaque classe est étiquetée par un système de trois tableaux de Young, deux d'entre eux ayant trois cases et correspondant aux valeurs propres $\pm i$ et un troisième tableau à deux cases correspondant à la valeur propre 1 (figure 4).

La symétrie verticale du graphe reflète la conjugaison $M \mapsto \bar{M}$ qui laisse stable l'armoire considérée. Les tiroirs centrés sur l'axe de symétrie sont les tiroirs réels : ils ne sont plus que 6. On peut les paramétrer par des couples de tableaux de Young associés aux facteurs réels irréductibles de

$$(X-1)^2(X^2+1)^3.$$

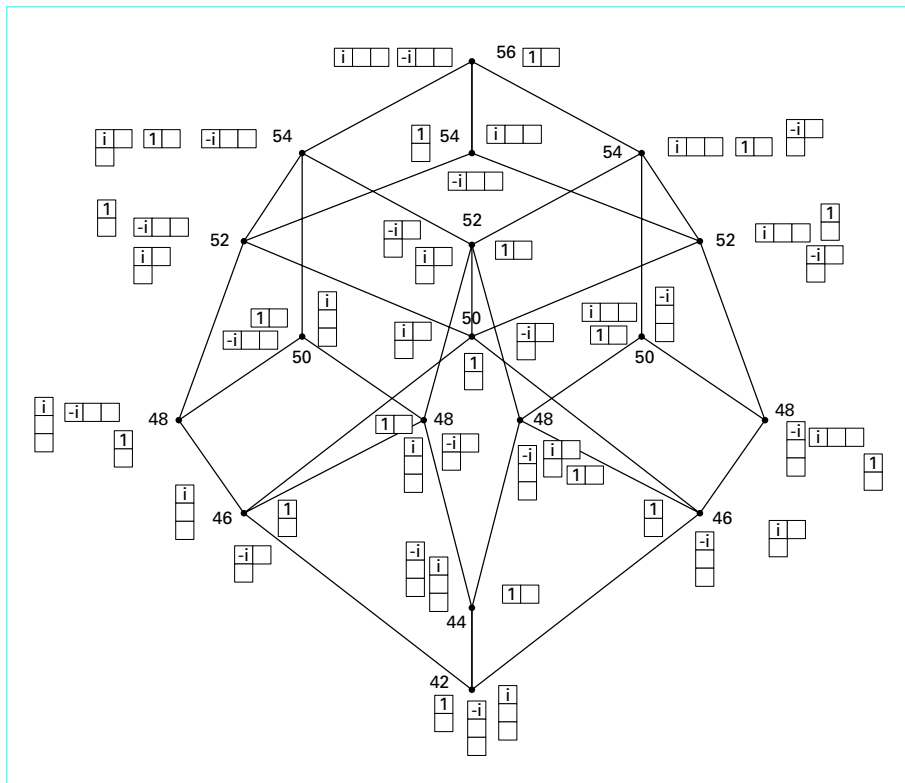


Figure 4 – Graphe d'une armoire complexe à polynôme caractéristique réel

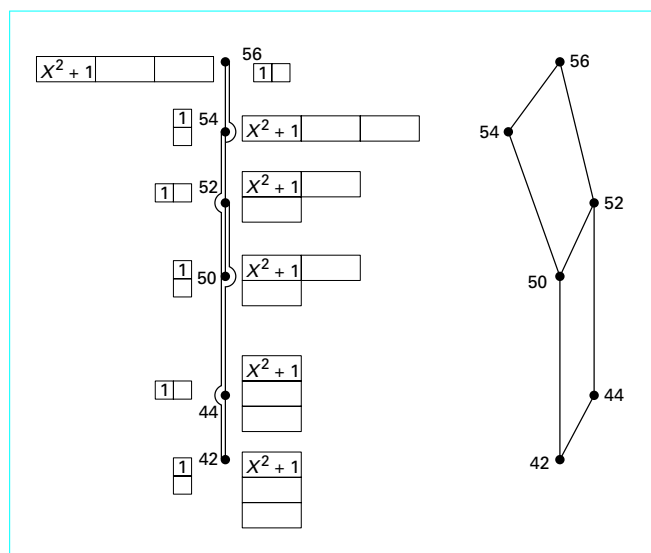


Figure 5 – Graphe d'une armoire réelle

La classe de similitude de dimension 54 correspondra ainsi au couple

$$\begin{pmatrix} X-1 & \\ & \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} X^2+1 & & \\ & & \\ & & \end{pmatrix}$$

La matrice

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}$$

est dans la classe de dimension 52. La dégradation d'une case correspondant à un tableau associé à $X^2 + 1$ entraîne une chute de 4 dimensions par niveau (figure 5).

13. Similitude et congruence. Matrices symétriques réelles

13.1 Généralités

On examine ici le lien entre similitude et congruence. Le groupe orthogonal est défini par

$$O(n) = \{M \in GL(n, \mathbb{R}) \mid {}^tMM = I_n\}$$

Deux matrices réelles A et B conjuguées sous l'action de $O(n)$ sont à la fois semblables et congruentes : les propriétés invariantes par la similitude et la congruence sont conservées pour A et B .

L'introduction dans ce cadre des espaces hermitiens et du groupe unitaire est subtile et simplifie considérablement l'approche : elle est du même ordre que le passage de \mathbb{R} à \mathbb{C} , passage fort utile dans certaines situations. Il s'agit d'étudier l'action par conjugaison du groupe unitaire

$$U(n) = \{M \in GL(n, \mathbb{C}) \mid MM^* = I_n\}, \text{ où } M^* = {}^t\overline{M}.$$

Il s'avère que l'on a la proposition suivante.

Proposition 11.

Si deux matrices réelles sont unitairement semblables alors elles sont orthogonalement semblables.

13.2 Action du groupe unitaire

Toute orbite sous l'action par similitude du groupe unitaire contient une matrice triangulaire (Schur) ; celles qui contiennent des matrices diagonales sont les orbites des matrices normales, c'est-à-dire des matrices qui commutent avec

$$A^* = {}^t\overline{A}.$$

Quand l'espace $E = \mathbb{C}^n$ est muni de la forme hermitienne définie positive canonique

$$\langle v, w \rangle = \sum_i v_i \overline{w_i},$$

ces matrices sont aussi caractérisées par le fait que si F en est un sous-espace stable, son orthogonal F^\perp est stable.

13.3 Matrices normales. Premières propriétés

Les résultats précédents résultent de considérations élémentaires.

Si A est quelconque, $\text{Ker } A = \text{Ker } A^*A$.

Si A est normale alors $\text{Ker } A = \text{Ker } A^*$; par suite, si \vec{v} est un vecteur propre de A associé à la valeur propre λ , il est vecteur propre de A^* associé à la valeur propre $\overline{\lambda}$. Les sous-espaces propres associés à des valeurs propres distinctes sont orthogonaux pour le produit hermitien :

$$\lambda \langle v, w \rangle = \langle A(v), w \rangle = \langle v, A^*w \rangle = \langle v, \overline{\mu}w \rangle = \overline{\mu} \langle v, w \rangle.$$

Et si l'on prend la somme des sous-espaces propres, elle est exhaustive car sinon, son orthogonal serait stable par l'adjoint A^* qui y aurait un vecteur propre, lequel est également vecteur propre de A .

$$\text{Si } A \text{ est normale, } \text{Ker } A = (\text{Im } A)^\perp.$$

13.4 Matrices normales. Autres caractérisations

Une matrice est normale si, et seulement si

$$\sum |\lambda_i|^2 = \text{Tr}(A^*A),$$

comme il résulte du fait que l'application $M \mapsto \text{Tr}(M^*M)$ est $U(n)$ -invariante et du fait du théorème de Schur. On voit ainsi qu'une

matrice triangulaire est normale si, et seulement si, elle est diagonale.

Une matrice A est normale si, et seulement si, A^* est un polynôme en A .

On constate aussi que la matrice A est normale si, et seulement si, tout vecteur propre de A est vecteur propre de A^* . En effet, si \vec{v} est un vecteur propre pour A , il l'est également pour A^* et donc A laisse stable l'hyperplan orthogonal ; une récurrence s'engage alors.

Exemple. Quand A et B sont normales, l'opérateur $\Phi : X \mapsto AX - XB$ de $M(n, \mathbb{C})$ est normal une fois que $M(n, \mathbb{C})$ a été muni du produit hermitien

$$\langle A, B \rangle = \text{Tr}(AB^*)$$

pour lequel d'ailleurs la base canonique (E_{ij}) , $i, j = 1, \dots, n$ est orthonormée. En comparant les noyaux de Φ et de Φ^* , on en déduit que si A et B sont normales et $AT = TB$, alors :

$$A^*T = TB^*.$$

Cela s'obtient également en remarquant d'abord que l'hypothèse $AT = TB$ implique que $P(A)T = TP(B)$ pour n'importe quel polynôme $P \in \mathbb{C}[X]$, le résultat en découle sachant que

$$A^* = P(A) \text{ et } B^* = P(B)$$

pour un même polynôme P (on considérera la matrice normale

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}).$$

13.5 Cas $n = 2$

Une matrice réelle 2×2 est normale si, et seulement si, elle est

symétrique ou bien est de la forme $\begin{bmatrix} a-b & \\ b & a \end{bmatrix}$ (par le calcul direct, ou

bien en utilisant la réduction sur \mathbb{C} puis sur \mathbb{R}).

Toute matrice réelle normale est orthogonalement semblable à une matrice bloc diagonale avec sur la diagonale des blocs de taille 1 ou 2, ces derniers étant des similitudes.

Cela s'exprime aussi par le fait qu'un endomorphisme normal u dans un espace euclidien E (réel) vérifie la chose suivante : ses sous-espaces propres (réels) sont orthogonaux deux à deux (il peut ne pas en exister) et leur somme est de codimension paire. L'orthogonal de cette somme est stable par u et la restriction de u à cette somme est somme directe orthogonale de similitudes planes. Comme

$$\begin{bmatrix} a-b & \\ b & a \end{bmatrix} = - \begin{bmatrix} a-b & \\ b & a \end{bmatrix} + 2a I_2,$$

on voit que la transposée tA d'une matrice normale réelle A est un polynôme (réel) en A .

13.6 Théorème de Specht

En fait, les fonctions $A \mapsto \text{Tr}m(A, A^*)$ où $m(X, Y)$ est un mot quelconque en deux matrices X et Y , sont invariantes sous l'action du groupe $U(n)$. Et, un théorème du à Specht affirme que ces fonctions séparent les orbites ! On notera que ce théorème s'énonce plus simplement dans le cas des matrices normales, puisque l'orbite

d'une matrice normale est entièrement déterminée par les valeurs propres, lesquelles sont déterminées par leurs seules sommes de Newton : deux matrices normales A et B sont semblables si, et seulement si,

$$\text{Tr}(A^k) = \text{Tr}(B^k)$$

pour tout k .

Ce critère n'est pas effectif ; cependant une amélioration due à Percy indique que l'on peut se contenter de mots de degré $\leq 2n^2$.

Exemples.

a) On montre que deux matrices A et B d'ordre deux sont unitairement semblables, si et seulement si,

$$\text{Tr}(A) = \text{Tr}(B),$$

$$\text{Tr}(A^2) = \text{Tr}(B^2)$$

et

$$\text{Tr}(AA^*) = \text{Tr}(BB^*).$$

(On notera que le scalaire $\text{Tr}(AA^*) = \sum_i |\lambda_i|^2$ est un invariant.)

b) Les deux matrices d'ordre $2n$ suivantes

$$\begin{bmatrix} A & A \\ A & A \end{bmatrix} \text{ et } \begin{bmatrix} 2A & 0 \\ 0 & 0 \end{bmatrix}$$

sont unitairement semblables.

13.7 Matrices symétriques et antisymétriques réelles

■ Une **matrice symétrique** réelle $S \in M(n, \mathbb{R})$ est, en particulier, normale. Ses valeurs propres sont réelles (comme toute matrice hermitienne, c'est-à-dire toute matrice H complexe vérifiant $H^* = H$) et elle est diagonalisable dans une base orthonormée pour le produit scalaire. On peut écrire $S = ODO^{-1}$, avec D diagonale et O orthogonale. Les matrices S et D sont à la fois semblables et congruentes. En particulier, si (p, q) est la signature de la forme quadratique sur \mathbb{R}^n définie par S , l'entier p (resp. q) coïncide avec le nombre de valeurs propres > 0 (resp. < 0) de S .

■ Il est également facile d'établir qu'une **matrice antisymétrique** réelle a toutes ses valeurs propres imaginaires pures et qu'elle est orthogonalement semblable à une matrice diagonale en blocs, nuls

ou de la forme $\begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}$, où $\alpha \in \mathbb{R}$. Son rang est en particulier pair.

Exemple : le coefficient q_A de X^{n-2} dans le polynôme caractéristique $\chi_A(X)$ de

$$A = (\alpha_{ij}) \in M(n, \mathbb{R}), \text{ avec } n \geq 2,$$

est une forme quadratique en (les coefficients de) A , qui est non dégénérée et de signature

$$[n(n-1)/2 + 1 ; n(n+1)/2 - 1].$$

En effet, les termes non diagonaux de la matrice A interviennent pour $i < j$ par $-a_{ij}a_{ji}$; chacune de ces $n(n-1)/2$ expressions contribue par $(1;1)$ dans la signature.

Quant aux termes diagonaux, ils interviennent par l'expression $\sum_{i \leq j} a_{ii}a_{jj}$.

Cette forme quadratique a même signature que la forme quadratique définie par la matrice $2J - I$, où I est la matrice identité et J la matrice qui a des 1 partout. La matrice J de rang 1 est semblable à $\text{Diag}(n, 0, \dots, 0)$. La matrice symétrique $2J - I$ a donc pour signature $(1; n-1)$. Le résultat cherché est immédiat.

13.8 Théorème de Horn

Le vecteur $(h_{ii})_{i=1, \dots, n}$ de \mathbb{R}^n correspondant à la diagonale d'une matrice hermitienne quelconque appartenant à une orbite sous l'action de $U(n)$ se trouve évidemment dans le sous-espace affine d'équation

$$\sum_i x_i = c$$

où c désigne la trace commune à toutes les matrices de l'orbite.

En fait, un théorème dû à **Horn** précise que les vecteurs qui correspondent ainsi aux diagonales des matrices hermitiennes, d'une même orbite sous l'action de $U(n)$ décrivent exactement l'enveloppe convexe des vecteurs de coordonnées $(\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(n)})$, où les λ_i sont les valeurs propres et σ une permutation quelconque dans \mathfrak{S}_n .

Ce résultat est à comparer avec le fait que les diagonales des matrices semblables (sous l'action de $GL(n, \mathbb{C})$) au bloc de Jordan J_n décrivent quant à elles l'ensemble des points de l'hyperplan $\sum_i x_i = 0$. Le fait que J_n et λJ_n soient semblables pour tout λ non nul, prouve alors que l'ensemble cherché est un cône ; il suffit alors d'établir que le vecteur nul en est un point intérieur, et pour cela, on procède par un argument de géométrie différentielle.

On considère à cet effet l'application $P \mapsto PJ_n P^{-1}$ qui admet pour différentielle au point I_n l'application

$$\text{ad } J_n : X \mapsto [J_n, X] = J_n X - X J_n.$$

La manière avec laquelle opère $\text{ad } J_n$ est bien connue : elle consiste à pousser une pardiagonale (c'est-à-dire l'ensemble des coefficients x_{ij} tels que $i - j$ est constante) vers la pardiagonale juste plus haut. Seule la sous-diagonale compte donc pour voir ce qui se passe sur la diagonale, et, là, les sous-diagonales se surjectent sur les diagonales dont la somme des coefficients est nulle. L'application, qui à P associe la diagonale de $PJ_n P^{-1}$ est une submersion en I_n , elle est donc ouverte au voisinage de I_n .

13.9 Théorème de Lyapounov

Théorème 14 (Lyapounov). Une matrice A a son spectre dans $\text{Re}(z) > 0$ si, et seulement si, l'équation linéaire

$$AX + XA^* = I$$

admet une solution X hermitienne définie positive. Si tel est le cas, la matrice X est unique et est donnée par

$$\int_0^\infty e^{-tA} e^{-tA^*} dt.$$

L'idée de départ est que le groupe $GL(n, \mathbb{C})$ opère naturellement sur l'ensemble \mathcal{T} des triplets (A, X, H) de matrices (où les deux dernières matrices sont hermitiennes définies positives) vérifiant l'équation de Lyapounov $AX + XA^* = H$:

$$P \cdot (A, X, H) = (PAP^{-1}, PXP^*, PHP^*).$$

Il s'avère que l'application $(A, X, H) \mapsto A$ applique l'ensemble \mathcal{T} sur l'ouvert des **matrices positivement stables** (c'est-à-dire dont le spectre est contenu dans le demi-plan ouvert de droite) et la fibre au dessus d'une telle matrice A (c'est-à-dire l'ensemble des triplets de \mathcal{T} dont la première composante est précisément A) est donnée par le graphe d'un difféomorphisme

$$H \mapsto X = \int_0^\infty \exp(-tA) H \exp(-tA^*) dt$$

entre le cône HDP (des matrices hermitiennes définies positives) et un cône qui y est contenu.

On se contentera à cet effet de **trois remarques**.

La **première** est que si $(A, X, H) \in \mathcal{T}$, alors la matrice A est positivement stable car, du fait de l'action sur les triplets, une matrice B qui lui est semblable vérifie que $B + B^*$ est définie positive et donc si (λ, x) est un couple B -propre,

$$2 \text{Re}(\lambda) \langle x, x \rangle = \langle \lambda x, x \rangle + \langle x, \lambda x \rangle = \langle (B + B^*)(x), x \rangle.$$

La **deuxième** est que

$$\begin{aligned} H &= [\exp(-tA) H \exp(-tA^*)]_0^\infty \\ &= \int_0^\infty A [\exp(-tA) H \exp(-tA^*) + \exp(-tA) H \exp(-tA^*) A^*] dt \\ &= AX + XA^*. \end{aligned}$$

Les intégrales ci-dessus convergent, car si $A = S + N$ est la décomposition de Jordan-Dunford de A , $\mu = \inf \text{Re}(\lambda)$ et c_A une constante dépendant de A bien choisie,

$$\begin{aligned} \|\exp(-tA)\| &= \|\exp(-t(S+N))\| \leq \|\exp(-tS)\| \|\exp(-tN)\| \\ &\leq c_A e^{-\mu t} \|\exp(-tN)\| = O(t^n) e^{-t\mu}, \end{aligned}$$

quand $t \mapsto \infty$, puisque l'exponentielle en N nilpotente est un polynôme en N de degré au plus n .

Enfin, la **dernière** est que si A est positivement stable, l'application linéaire

$$\Psi : X \mapsto AX + XA^*$$

est bijective, car son spectre est donné par les $\lambda_i + \bar{\lambda}_j$, où les λ_i sont les valeurs propres de A (on notera que les deux opérateurs $X \mapsto AX$ et $X \mapsto XA^*$ commutent et sont donc simultanément trigonalisables).

Il s'ensuit que la composante X d'un élément (A, X, H) de \mathcal{T} est entièrement déterminée par les deux autres et s'obtient en inversant l'opérateur Ψ qui est à spectre dans le demi-plan $\{z, \text{Re}(z) > 0\}$. Or, pour un tel opérateur, l'analyse fonctionnelle élémentaire nous donne

$$\Psi^{-1} = \int_0^\infty e^{-t\Psi} dt,$$

et l'exponentielle de l'endomorphisme $\varphi : X \mapsto AX + XA^*$ est donnée par

$$Z \mapsto e^A Z e^B;$$

(le groupe à un paramètre associé à φ coïncide avec

$$\Phi : t \mapsto [Z \mapsto e^{tA} Z e^{tB}],$$

qui vérifie bien :

$$\Phi(t_1 + t_2) = \Phi(t_1)\Phi(t_2);$$

il suffit en effet de dériver ce dernier sous-groupe à un paramètre en $t = 0$.

14. Quelques exemples récapitulatifs

14.1 Une diagonalisation explicite

$$\text{Soit } A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix} \in M(3, \mathbb{R}).$$

On obtient

$$\chi_A(X) = X(X-2)(X-3).$$

La matrice A est donc diagonalisable, car à valeurs propres distinctes. La résolution des équations

$$Ax_1 = 0; \quad Ax_2 = 2x_2; \quad Ax_3 = 3x_3$$

conduit à choisir par exemple

$$x_1 = (1, 1, 0); \quad x_2 = (1, -1, 0) \text{ et } x_3 = (-1, 2, 3).$$

$$\text{Si on pose } P = [x_1 x_2 x_3] = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 2 \\ 0 & 0 & 3 \end{bmatrix}, \text{ on a}$$

$$P^{-1}AP = \text{Diag}(0, 2, 3).$$

Le calcul des puissances A^k de A , pour k grand, devient aisé. Il s'agit cependant de calculer l'inverse de P .

14.2 Une jordanisation effective

$$\text{Soit } A = \begin{bmatrix} 4 & 1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 1 \end{bmatrix} \in M(3, \mathbb{R}). \text{ On a}$$

$$\chi_A(X) = (X-1)^2(X-2).$$

Le vecteur $x_3 = (1, -2, 0)$ vérifie l'équation $Ax_3 = 2x_3$ et fournit donc une base du sous-espace propre $E_A(2)$.

Le sous-espace propre $E_A(1)$ s'avère être de dimension 1 et engendré par $(1, -2, 1)$. La matrice A n'est donc pas diagonalisable.

Par ailleurs, la matrice $(A-I)^2$ de rang 1 est donnée par

$$\begin{bmatrix} 1 & 0 & -1 \\ -2 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Le vecteur $v_2 = (0, 1, 0)$ est dans $\text{Ker}(A-I)^2$ mais non dans $\text{Ker}(A-I)$.

La base (v_1, v_2, v_3) , où $v_1 = (A-I)(v_2) = (1, -2, 1)$, est donc une base de jordanisation de A .

$$\text{En posant } P = [v_1 v_2 v_3] = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & -2 \\ 1 & 0 & 0 \end{bmatrix}, \text{ on a}$$

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

Le recours à une réduction effective peut s'avérer très utile si l'on a à calculer par exemple une puissance A^k de A pour k assez grand. Il est alors nécessaire de calculer P^{-1} .

14.3 Autre exemple

Soit à réduire la matrice

$$M(a, b) = \begin{bmatrix} a & a & a \\ b & a & a \\ b & b & a \end{bmatrix} \in M(3, \mathbb{C}).$$

Le cas $b = 0$ et $a \neq 0$ est facile : la matrice $M - aI$ est de rang 2 et la matrice M est semblable au bloc de Jordan $J_3(a)$ associé à la valeur propre a , soit

$$\begin{bmatrix} a & a & a \\ 0 & a & a \\ 0 & 0 & a \end{bmatrix} \sim \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}.$$

Dans le cas où b est non nul, on peut en considérant $\frac{1}{b}M$ supposer que $b = 1$. Un calcul par Maple donne, quand $n = 3$, une

valeur propre de $A = \begin{bmatrix} a & a & a \\ 1 & a & a \\ 1 & 1 & a \end{bmatrix}$ qui vaut $a + a^{1/3} + a^{2/3}$ et un poly-

nôme caractéristique dont les racines ne sautent donc pas aux yeux. Cela laisse à penser que A est en fait un polynôme en une matrice plus facile ! On a en effet

$$A = B^3 + B^2 + B, \text{ où } B = \begin{bmatrix} 0 & 0 & a \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Cette dernière matrice est une matrice compagnon et son polynôme caractéristique est donné par $X^3 - a$. La réduction de B est facile, celle de A s'en déduit tout aussi facilement.

14.4 Ressorts de Trubowitz

Dans des coordonnées naturelles, l'équation du système formé par les ressorts oscillant sur un rail circulaire sans frottement et reliés par des ressorts identiques, est de la forme

$$(d/dt)^2 x = Ax,$$

où x est le vecteur des positions des n masses et A la matrice

$$I_n - \frac{1}{2}(C + C^{-1}), \text{ avec pour } C \text{ la matrice compagnon du polynôme}$$

X^{n-1} . En particulier, A est circulaire, c'est-à-dire combinaison linéaire de puissances de la matrice C . Comme toute matrice circulaire, la matrice A se diagonalise facilement sur le corps des nombres complexes (les vecteurs $(1, z, z^2, \dots, z^{n-1}) \in \mathbb{C}^n$, où z parcourt les racines n -ièmes de l'unité, forment une base de vecteurs propres). On réduit ainsi le système ci-dessus à un système semblable de la forme

$$(d/dt)^2 u + Du = 0,$$

où la matrice D est diagonale à coefficients positifs ou nuls. Ce système est équivalent à n équations du type

$$(d/dt)^2 u_i + d_i^2 u_i = 0$$

dont la solution est immédiate :

$$\begin{aligned} & \text{— } u_i = A_i \cos(d_i t) + B_i \sin(d_i t) \text{ si } d_i \text{ est non nul;} \\ & \text{— } u_i = A_i t + B_i \text{ si } d_i \text{ est nul.} \end{aligned}$$

Les solutions du système obtenues en prenant $u_i = 0$ sauf pour un seul i sont appelés des « états purs » du système. Quand le système se trouve dans un état pur, toutes les masses oscillent avec la même fréquence. On trouvera dans [10], une liste de liens sur les dessins (animés) de quelques états purs et d'un état qui ne l'est pas.

Références bibliographiques

- [1] ADKINS (A.) et WEINTRAUB (S.). – *Algebra. An Introduction via module Theory*, Springer 1992. ISBN 0-387-97839-9.
- [2] CHAMBADAL (L.) et OVAERT (J.-L.). – *Algèbre linéaire et algèbre tensorielle*, Dunod 1968.
- [3] FRENEL (J.). – *Algèbre des matrices*, Hermann 1997. ISBN 2-7056-1439-7.
- [4] GLAZMAN (I.) et LIUBITCH (Y.). – *Analyse linéaire dans les espaces de dimensions finies*, Éditions Mir 1974.
- [5] GOBLOT (R.). – *Algèbre linéaire*, Scientifika 1994. ISBN 2-909894-49-5.
- [6] HORN (R.) et JOHNSON (C.). – *Topics in matrix analysis*, Cambridge University Press 1991. ISBN 0-521-46713-6.
- [7] HUBBARD (J.), WEST (B.) et GAUTHERON (V.). – *Équations différentielles et systèmes dynamiques*, Cassini 1999. ISBN 2-84225-015-X.
- [8] JACOBSON (N.). – *Basic Algebra I*, Freeman and co 1974. ISBN 0-7167-0453-6.
- [9] KAPLANSKY (I.). – *Linear Algebra and Geometry. A second course*, Chelsea publishing 1974.
- [10] KELLER (B.). – *Les ressorts de Trubowitz, une application de la diagonalisation*, <http://www.math.jussieu.fr/~mneimne/diag/>, 1998.
- [11] MNEIMNÉ (R.). – *Éléments de Géométrie. Actions de groupes*, Cassini 1997. ISBN 2-84225-003-6.
- [12] MNEIMNÉ (R.). – *Réduction des endomorphismes. Cône nilpotent et tableaux de Young. Le poly vert*, Université Paris VII, Denis-Diderot. 1998.
- [13] MNEIMNÉ (R.) et TESTARD (F.). – *Introduction à la théorie des groupes de Lie classiques*, Hermann 1986. ISBN 2-7056-6040-2.
- [14] MNEIMNÉ (R.) et TESTARD (F.). – *On products of singular elements, Séminaire de théorie des nombres*, Série 2, vol. 3 (1991), 337-350.
- [15] PHAM (F.) et DILLINGER (H.). – *Algèbre linéaire*, Diderot 1996. ISBN 2-84134-066-x.
- [16] PRASOLOV (V.). – *Problems and Theorems in Linear Algebra*, AMS Translations, vol. 134 – 1994. ISBN 0-8218-0236-4.
- [17] ROUVIÈRE (F.). – *Petit guide de calcul différentiel*, Cassini 1999. ISBN 2-84225-008-7.
- [18] TAUVEL (P.). – *Mathématiques générales pour l'agrégation*, Masson 1993. ISBN 2-225-82733-8.
- [19] VALIRON (G.). – *Cours d'analyse mathématique II – Équations fonctionnelles, applications*, Masson 1950.

Gagnez du temps et sécurisez vos projets en utilisant une source actualisée et fiable



RÉDIGÉE ET VALIDÉE
PAR DES EXPERTS




MISE À JOUR
PERMANENTE



100 % COMPATIBLE
SUR TOUS SUPPORTS
NUMÉRIQUES



SERVICES INCLUS
DANS CHAQUE OFFRE

- + de 340 000 utilisateurs chaque mois
- + de 10 000 articles de référence et fiches pratiques
- Des Quiz interactifs pour valider la compréhension 

SERVICES ET OUTILS PRATIQUES



Questions aux experts*

Les meilleurs experts techniques et scientifiques vous répondent



Articles Découverte

La possibilité de consulter des articles en dehors de votre offre



Dictionnaire technique multilingue

45 000 termes en français, anglais, espagnol et allemand



Archives

Technologies anciennes et versions antérieures des articles



Info parution

Recevez par email toutes les nouveautés de vos ressources documentaires

*Questions aux experts est un service réservé aux entreprises, non proposé dans les offres écoles, universités ou pour tout autre organisme de formation.

Les offres Techniques de l'Ingénieur



INNOVATION

- Éco-conception et innovation responsable
- Nanosciences et nanotechnologies
- Innovations technologiques
- Management et ingénierie de l'innovation
- Smart city – Ville intelligente



MATÉRIAUX

- Bois et papiers
- Verres et céramiques
- Textiles
- Corrosion – Vieillessement
- Études et propriétés des métaux
- Mise en forme des métaux et fonderie
- Matériaux fonctionnels. Matériaux biosourcés
- Traitements des métaux
- Élaboration et recyclage des métaux
- Plastiques et composites



MÉCANIQUE

- Frottement, usure et lubrification
- Fonctions et composants mécaniques
- Travail des matériaux – Assemblage
- Machines hydrauliques, aérodynamiques et thermiques
- Fabrication additive – Impression 3D



ENVIRONNEMENT – SÉCURITÉ

- Sécurité et gestion des risques
- Environnement
- Génie écologique
- Technologies de l'eau
- Bruit et vibrations
- Métier : Responsable risque chimique
- Métier : Responsable environnement



ÉNERGIES

- Hydrogène
- Ressources énergétiques et stockage
- Froid industriel
- Physique énergétique
- Thermique industrielle
- Génie nucléaire
- Conversion de l'énergie électrique
- Réseaux électriques et applications



GÉNIE INDUSTRIEL

- Industrie du futur
- Management industriel
- Conception et production
- Logistique
- Métier : Responsable qualité
- Emballages
- Maintenance
- Traçabilité
- Métier : Responsable bureau d'étude / conception



ÉLECTRONIQUE – PHOTONIQUE

- Électronique
- Technologies radars et applications
- Optique – Photonique



TECHNOLOGIES DE L'INFORMATION

- Sécurité des systèmes d'information
- Réseaux Télécommunications
- Le traitement du signal et ses applications
- Technologies logicielles – Architectures des systèmes
- Sécurité des systèmes d'information



AUTOMATIQUE – ROBOTIQUE

- Automatique et ingénierie système
- Robotique



INGÉNIERIE DES TRANSPORTS

- Véhicule et mobilité du futur
- Systèmes aéronautiques et spatiaux
- Systèmes ferroviaires
- Transport fluvial et maritime



MESURES – ANALYSES

- Instrumentation et méthodes de mesure
- Mesures et tests électroniques
- Mesures mécaniques et dimensionnelles
- Qualité et sécurité au laboratoire
- Mesures physiques
- Techniques d'analyse
- Contrôle non destructif



PROCÉDÉS CHIMIE – BIO – AGRO

- Formulation
- Bioprocédés et bioproductions
- Chimie verte
- Opérations unitaires. Génie de la réaction chimique
- Agroalimentaire



SCIENCES FONDAMENTALES

- Mathématiques
- Physique Chimie
- Constantes physico-chimiques
- Caractérisation et propriétés de la matière



BIOMÉDICAL – PHARMA

- Technologies biomédicales
- Médicaments et produits pharmaceutiques



CONSTRUCTION ET TRAVAUX PUBLICS

- Droit et organisation générale de la construction
- La construction responsable
- Les superstructures du bâtiment
- Le second œuvre et l'équipement du bâtiment
- Vieillessement, pathologies et réhabilitation du bâtiment
- Travaux publics et infrastructures
- Mécanique des sols et géotechnique
- Préparer la construction
- L'enveloppe du bâtiment
- Le second œuvre et les lots techniques