Algèbre linéaire

Thomas Dedieu

Rentrée 2019, compilé le 13 avril 2020

Avertissement. Ceci n'est pas un polycopié de cours. Ce texte est constitué de mes notes personnelles pour un cours de préparation à l'agrégation, grossièrement mises en forme. Néanmoins, si vous trouvez des erreurs il m'intéresse que vous me les communiquiez.

Prologue **3**

1 Produits de matrices ; systèmes linéaires **3**

2 Déterminants **8**

3 Polynôme caractéristique **12**

4 Matrices par blocs et sous-espaces stables ; quotients **14**

4.1 Matrices par blocs . . . 14

4.2 Rappels et compléments sur les quotients . . . 15

4.3 Sous-espaces stables et matrices triangulaires par blocs . . . 17

5 Dualité **18**

I Réduction des endomorphismes : analyse **21**

1 Polynômes d'endomorphismes, polynômes annulateurs **21**

2 Trigonalisation **21**

3 Diagonalisation **22**

4 Étude des nilpotents **24**

5 Endomorphismes à polynôme caractéristique scindé **27**

6 Réduction dans le cas général **30**

6.1 Théorème de Cayley–Hamilton . . . 30

6.2 Remarques sur les endomorphismes cycliques . . . 31

II Réduction des endomorphismes : synthèse **33**

1 Simplicité et semi-simplicité **33**

2 Structure de l'algèbre $\mathbf{k}[u]$ **35**

2.1 Liens entre réduction d'un endomorphisme u et structure de l'algèbre $\mathbf{k}[u]$ . . . 35

2.2 Applications . . . 36

3 Classes d'équivalence de matrices à coefficients dans un anneau principal **38**

3.1 Algorithme de Gauss . . . 39

4 Classes de similitude dans $\mathcal{M}_n(\mathbf{k})$ **42**

5 Approche pédestre des invariants de similitude **44**

Appendices **47**

A Formule de Laplace **47**

B Exponentielle **50**

C Exercices d'hiver **55**

Références **56**

Prologue

Quelques principes suivant lesquels aborder ce cours (en espérant qu'ils permettent de ne pas trouver le préambule *bizarre*) :

- révisions et compléments de théorie de la réduction des endomorphismes (ce sera la partie linéaire et conventionnelle de ce cours) ;
- parenthèses digressives souvent de taille conséquente ;
- tâcher de s'entraîner à rendre concrets les énoncés rencontrés (aptitude attendue à l'oral du concours, et peu développée dans les années d'études précédentes) ;
- on s'adresse à de futurs enseignants (y compris ceux qui poursuivront vers la recherche), en conséquence de quoi y a-t-il certains points dont il faut être sûr qu'ils sont bien clairs.

J'ajoute le principe universel :

- tâcher d'être toujours acteur du cours, autrement dit de ne pas rester spectateur, et passer le cerveau du mode téléspectateur au mode actif. (Pas si facile !)

Dans tout ce cours $\mathbf{k}$ est un corps arbitraire ; dans corps il y a écrit commutatif, sinon on dit corps gauche ou mieux, algèbre à division. Avec une algèbre à division il y a des morceaux importants de la théorie qui se mettent à déconner, par exemple un polynôme peut avoir une infinité de racines, ce qui est fâcheux. Ainsi, le polynôme X^2+1 a une infinité de racines distinctes dans le pas-corps des quaternions $\mathbf{H}$ (par exemple i, j, k), toutes conjuguées à i.

Notre $\mathbf{k}$-espace vectoriel de travail s'appellera toujours E sauf quand il ne s'appellera pas E, et il sera supposé de dimension finie. Souvent F est un sous-espace vectoriel de E, mais peut-être pas toujours. Je mélange sans vergogne applications linéaires et matrices : c'est mal.

1 – Produits de matrices ; systèmes linéaires

(1.1) On multiplie à droite par des colonnes, et à gauche par des lignes (la seconde opération se déduisant de la première par transposition).

(1.2) On en déduit comment multiplier par des transvections/dilatations/permutations sans se tromper.

(1.2.1) Exercice. Pour tout $i \neq j$, $\lambda \in (\mathbf{k}, +) \mapsto T_{ij}(\lambda) \in (\mathrm{GL}, \times)$ est un morphisme de groupes.

(1.2.2) Application. Générateurs des groupes (général) linéaire et spécial linéaire :

- GL est engendré par les transvections et les dilatations ;
- SL est engendré par les transvections.

Au moins pour GL, ça se démontre avec le pivot de Gauss, en utilisant le fait astucieux que toute transposition est produit de transvections et dilatations :

$$\begin{pmatrix} L_1 \\ L_2 \end{pmatrix} \to \begin{pmatrix} L_1+L_2 \\ L_2 \end{pmatrix} \to \begin{pmatrix} L_1+L_2 \\ L_2-(L_1+L_2) \end{pmatrix} = \begin{pmatrix} L_1+L_2 \\ -L_1 \end{pmatrix} \to \begin{pmatrix} L_2 \\ -L_1 \end{pmatrix} \to \begin{pmatrix} L_2 \\ L_1 \end{pmatrix}.$$

(1.2.3) Application. Réinterprétation d'une méthode pour inverser les matrices : effectuer les mêmes opérations élémentaires sur les lignes de A et de Id, jusqu'à avoir transformé A en l'identité. Ceci revient à multiplier à gauche A et Id par la même matrice inversible P, produit de matrices de transvections, dilatations, et permutation. On ne connaît pas P mais elle même,

mais bien la liste de tous ces facteurs de transvections, dilatations, et permutation. À la fin : $PA = \mathrm{Id}$, donc l'inverse de A est P qui est $P\mathrm{Id}$, autrement dit on connaît P en effectuant toutes les opérations élémentaires sur les lignes de Id.

Il y a une façon tristement plus élémentaire de comprendre cette méthode. On la donne en (1.3.3), puisqu'il faut au préalable voir le paragraphe (1.3) ci-dessous.

(1.3) Inversibilité et systèmes linéaires. *Soit $A \in \mathcal{M}_n(\mathbf{k})$. Les propositions suivantes sont équivalentes :*

(i) *A est inversible ;*
(ii) $\forall Y \in \mathbf{k}^n,\ \exists! X \in \mathbf{k}^n : AX = Y$;
(iii) $\exists B \in \mathcal{M}_n(\mathbf{k}) : \forall X, Y \in \mathbf{k}^n,\ AX = Y \Leftrightarrow X = BY$.

Avant d'attaquer la preuve, rappellons que par définition A est inversible s'il existe $B \in \mathcal{M}_n(\mathbf{k})$ telle que $AB = BA = \mathrm{Id}_n$.

(1.3.1) Preuve. Manifestement (i) implique (ii) et (iii), et (iii) implique (ii). On va démontrer (ii) $\Rightarrow$ (iii) puis (iii) implique (i), ce qui suffit pour conclure.

Commençons par montrer (ii) $\Rightarrow$ (iii). Il s'agit de démontrer que l'unique solution du système $AX = Y$ s'exprime linéairement en fonction du second membre. Notons $C_1, \ldots, C_n \in \mathbf{k}^n$ les colonnes de la base canonique, et considérons $X_1, \ldots, X_n \in \mathbf{k}^n$ les solutions respectives des équations $AX = C_i$, $i = 1, \ldots, n$, d'inconnue X. Soit $Y = {}^{\mathsf{T}}(y_1, \ldots, y_n) \in \mathbf{k}^n$. On a

$$A \times (y_1 X_1 + \cdots + y_n X_n) = y_1 AX_1 + \cdots + y_n AX_n = y_1 C_1 + \cdots + y_n C_n = Y,$$

donc l'unique solution de $AX = Y$ est

$$X = y_1 X_1 + \cdots + y_n X_n = \left(\begin{array}{c|c|c} X_1 & \cdots & X_n \end{array}\right) \times \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

ce qui prouve que (iii) vaut avec B la matrice obtenue en concaténant les colonnes $X_1, \ldots, X_n$.

À présent montrons que (iii) implique (i). À nouveau, on note $C_1, \ldots, C_n \in \mathbf{k}^n$ les colonnes de la base canonique. Les colonnes $B_1, \ldots, B_n$ de B sont $BC_1, \ldots, BC_n$, et on a donc $AB_i = A(BC_i) = C_i$ pour tout $i = 1, \ldots, n$. Ainsi

$$AB = A \times \left(\begin{array}{c|c|c} B_1 & \cdots & B_n \end{array}\right) = \left(\begin{array}{c|c|c} C_1 & \cdots & C_n \end{array}\right) = \mathrm{Id}_n,$$

et donc $AB = \mathrm{Id}$. Il reste à démontrer que $BA = \mathrm{Id}$, ce qui n'est pas tautologique. Pour tout $Z \in \mathbf{k}^n$, $B(AZ)$ est l'unique X solution de l'équation $AX = AZ$. Manifestement $X = Z$ est solution de cette équation, donc par unicité de la solution on a $B(AZ) = Z$. Les colonnes de la matrice BA sont les $BA \times C_i$, et d'après ce qu'on vient de dire $BAC_i = C_i$, donc $BA = \mathrm{Id}$ comme il fallait démontrer. □

(1.3.2) Remarque. Les équivalences (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) s'obtiennent uniquement par du calcul matriciel élémentaire, en particulier elles valent encore lorqu'on remplace $\mathbf{k}$ par un anneau commutatif.

(1.3.3) Application au calcul de l'inverse. D'après (1.3), pour l'inverse de A est la matrice B telle que $AX = Y \Leftrightarrow X = BY$ pour tout $X, Y \in \mathbf{k}^n$. On obtient donc l'inverse en résolvant le système

$$\text{(1.3.3.i)} \qquad \begin{cases} a_{11}x_1 & +\cdots+ & a_{1n}x_n & = & y_1 & & \\ \vdots & & \vdots & \vdots & & \ddots & \\ a_{n1}x_1 & +\cdots+ & a_{nn}x_n & = & & & y_n \end{cases}$$

où les x_i sont les inconnues, les a_{ii} sont des scalaires, et les y_j sont des variables. La solution de ce système donnera chaque x_i en fonction des variables y_j, sous la forme suivante :

$$\text{(1.3.3.ii)} \qquad \begin{cases} x_1 & & = & b_{11}y_1 & +\cdots+ & b_{1n}y_n \\ & \ddots & \vdots & \vdots & & \vdots \\ & x_n & = & b_{n1}y_1 & +\cdots+ & b_{nn}y_n. \end{cases}$$

et la matrice (b_{ij}) est l'inverse de A.

Ceci donne une explication élémentaire à la méthode magique (1.2.3). On présente le système (1.3.3.i) sous forme d'un tableau de nombres, en n'écrivant plus ni les inconnues ni les variables et en réservant une colonne pour chacune d'entre elles. Ainsi, (1.3.3.i) est représenté par le tableau

$$\text{(1.3.3.i')} \qquad \left(\begin{array}{ccc||ccc} a_{11} & \cdots & a_{1n} & 1 & \cdots & 0 \\ \vdots & & \vdots & & \ddots & \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 1 \end{array} \right).$$

On résout le système par le pivot de Gauss, en opérant sur les lignes du système (1.3.3.i), ce qui se traduit par des opérations sur les lignes du tableau (1.3.3.i'). À la fin de l'algorithme on arrive à la solution (1.3.3.ii), dont la représentation sous forme d'un tableau de nombres est

$$\text{(1.3.3.ii')} \qquad \left(\begin{array}{ccc||ccc} 1 & \cdots & 0 & b_{11} & \cdots & b_{1n} \\ & \ddots & & \vdots & & \vdots \\ 0 & \cdots & 1 & b_{n1} & \cdots & b_{nn} \end{array} \right).$$

On constate alors qu'il n'y a rien de magique dans la méthode (1.2.3).

(1.4) Proposition. *Soit $A, B \in \mathcal{M}_n(\mathbf{k})$. Les propositions suivantes sont équivalentes :*
(i) $\forall X, Y \in \mathbf{k}^n,\ AX = Y \Leftrightarrow X = BY$;
(ii) $\forall X, Y \in \mathbf{k}^n,\ AX = Y \Leftarrow X = BY$;
(iii) $\forall X, Y \in \mathbf{k}^n,\ AX = Y \Rightarrow X = BY$.

On va démontrer dans un premier temps cet énoncé en utilisant la théorie de la dimension. On verra en (2.9.1) qu'en fait cet ingrédient n'est pas nécessaire. Il est utile de pouvoir s'en passer si on travaille sur un anneau commutatif au lieu de notre corps $\mathbf{k}$ (commutatif lui aussi, par définition).

Preuve. Manifestement (i) implique (ii) et (iii). Il suffit de démontrer que réciproquement (ii) $\Rightarrow$ (i) et (iii) $\Rightarrow$ (i).

Commençons par (iii) $\Rightarrow$ (i). (iii) dit que pour tout $Y \in \mathbf{k}^n$, l'équation $AX = Y$ possède au plus une solution, nécessairement $X = BY$. Ainsi l'application linéaire $X \in \mathbf{k}^n \mapsto AX \in \mathbf{k}^n$ est

injective. Elle est donc aussi surjective, puisque c'est une application linéaire entre espaces de même dimension. Autrement dit l'équation $AX = Y$ possède toujours une solution, dont on a vu qu'elle est forcément $X = BY$. Ainsi (i) vaut.

La preuve de (ii) $\Rightarrow$ (i) est analogue. (ii) dit que l'équation $AX = Y$ possède toujours la solution $X = BY$. Ainsi $X \in \mathbf{k}^n \mapsto AX \in \mathbf{k}^n$ est surjective, et donc aussi injective par la théorie de la dimension, et (i) vaut. □

(1.5) Corollaire. *Soit $A, B \in \mathcal{M}_n(\mathbf{k})$. Les propositions suivantes sont équivalentes :*
(i) $AB = BA = \mathrm{Id}$ *;*
(ii) $AB = \mathrm{Id}$ *;*
(iii) $BA = \mathrm{Id}$.

Preuve. Le corollaire s'obtient à partir de la Proposition (1.4) en constatant que (i), (ii) et (iii) du corollaire équivalent respectivement à (i), (ii) et (iii) de la proposition. Pour (i), on l'a déjà vu en démontrant (i) $\Leftrightarrow$(iii) dans la preuve de (1.3). Je le fais pour (ii), et laisse (iii) au lecteur.

Si $AB = \mathrm{Id}$, alors le système $AX = Y$ possède toujours la solution $X = BY$. Réciproquement, si pour tout Y l'équation $AX = Y$ possède la solution $X = BY$, alors pour les colonnes $C_1, \ldots, C_n$ de la base canonique de $\mathbf{k}^n$ on a $A(BC_i) = C_i$. Or $A(BC_i) = (AB)C_i$ est la i-ème colonne de AB. On a donc $AB = \mathrm{Id}$ comme il fallait. □

(1.6) Allure de l'ensemble des solutions d'un système linéaire. L'ensemble des solutions de l'équation $AX = Y$ est non-vide si et seulement si $Y \in \mathrm{im}(A)$; dans ce cas, c'est un espace affine dont l'espace vectoriel sous-jacent est $\ker(A)$.

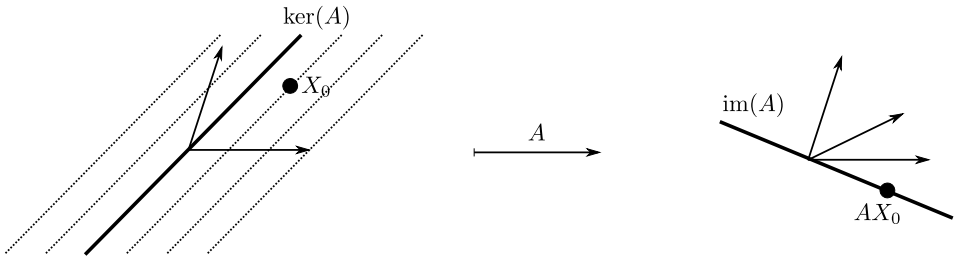


Figure 1: Lignes de niveau d'une application linéaire $\mathbf{k}^2 \to \mathbf{k}^3$ de rang 1. Si $AX_0 = Y_0$, l'ensemble des solutions de $AX = Y_0$ est la ligne de niveau passant par X_0.

Réciproquement, tout sous-espace affine de $\mathbf{k}^n$ peut être décrit comme l'ensemble des solutions d'une équation de la forme $AX = B$. On retrouve les équations de droites affines dans le plan de quand on était petit.

(1.7) Exercice. On considère un système linéaire $AX = B$, $A \in \mathcal{M}_{mn}$, $B \in \mathcal{M}_m$. Soit $I \subseteq [\![1, m]\!]$ tel que les lignes A_i, $i \in I$ de A forment une base de l'espace engendré par les lignes. On considère A_I et B_I les matrices obtenues en extrayant les lignes indicées par I de A et B respectivement. Démontrer que les deux propositions suivantes sont équivalentes :
(i) $AX = B$ possède au moins une solution ;
(ii) toute équation/ligne du système $AX = B$ est combinaison linéaire d'une équation/ligne du système $A_I X = B_I$.

(1.8) Lemme. *Un sev de E défini par une famille de rang r d'équations linéaires a codimension m dans E.*

(1.9) Théorème du rang. *Soit $A \in \mathcal{M}_{m,n}(\mathbf{k})$. Le rang c de la famille de vecteurs de $\mathbf{k}^n$ constituée des colonnes de A est*

$$c = \dim(E) - \dim\big(\ker(A)\big).$$

(1.10) Rang et rang de la transposée. Ici on démontre que le rang d'une matrice est égal au rang de sa transposée. La première preuve fait de cet énoncé un corollaire direct des deux résultats (1.8) et (1.9) ci-dessus. La seconde est un peu plus astucieuse et me plaît moins ; on peut y déceler l'influence du pivot de Gauss.

(1.10.1) Preuve 1. On considère $A \in \mathcal{M}_{m,n}(\mathbf{k})$, et on note c le rang de ses n colonnes dans $\mathbf{k}^m$, r le rang de ses m lignes dans ${}^{\mathsf{T}}\mathbf{k}^n$. [1] Par définition, c est le rang de A et r est celui de sa transposée.

Le théorème du rang nous dit que c, le rang des colonnes, est la codimension du noyau dans E. Mais par définition, la codimension du noyau est égale au rang des lignes, r, donc $c = r$. □

(1.10.2) Preuve 2. On commence par démontrer (savez-vous vraiment le faire ?) que $A \in \mathcal{M}_{m,n}(\mathbf{k})$ est de rang r si et seulement si il existe $P \in \mathrm{GL}_m$ et $Q \in \mathrm{GL}_n$ telles que

$$P^{-1}AQ = J^r_{m,n}$$

où $J^r_{m,n}$ est la matrice $m \times n$ avec des 1 sur les r premiers coefficients de la diagonale, et des 0 partout ailleurs.

Ensuite, si A est de rang r on trouve P et Q comme ci-dessus, et en transposant l'identité on obtient

$${}^{\mathsf{T}}Q\,{}^{\mathsf{T}}A\,{}^{\mathsf{T}}P^{-1} = {}^{\mathsf{T}}J^r_{m,n} = J^r_{n,m},$$

qui prouve que ${}^{\mathsf{T}}A$ est de rang r elle aussi. □

(1.11) Systèmes d'équations linéaires à coefficients dans un anneau commutatif. On se contente de quelques observations. C'est le moment d'étudier le Théorème des facteurs invariants (3.2) ?

(1.12) Exercice. Approche pédestre des Grassmanniennes vues comme des espaces homogènes.
1) Soit $n, a \in \mathbf{N}^*$, et considérons $M \in \mathrm{GL}_n$. Montrer que pour tout $M' \in \mathcal{M}_{n,a}$ il existe une unique matrice H, dont on déterminera la taille, telle que $M' = MH$.

2) Soit $n \in \mathbf{N}^*$, $r \in [\![1, n]\!]$, et $M, M' \in \mathcal{M}_{n,r}$. On suppose que M et M' sont toutes les deux de rang r. On note $M_1, \ldots, M_r \in \mathbf{k}^n$ (resp. $M'_1, \ldots, M'_r \in \mathbf{k}^n$) les colonnes de M (resp. M'). Démontrer l'équivalence des deux propositions suivantes :
(i) il existe $H \in \mathrm{GL}_r$ tel que $M' = M \times H$;
(ii) $\mathrm{Vect}(M_1, \ldots, M_r) = \mathrm{Vect}(M'_1, \ldots, M'_r)$.

3) Soit $n \in \mathbf{N}^*$ et $p \in [\![1, n]\!]$. On considère l'ensemble P des matrices de GL_{n+1} triangulaires supérieures par blocs

$$\left(\begin{array}{c|c} A & B' \\ \hline 0 & B \end{array}\right)$$

avec $A \in \mathrm{GL}_p$ et $B \in \mathrm{GL}_{n+1-p}$.
a) Montrer que P est un sous-groupe de GL_{n+1}.

1. r comme *riga* ou *row* (ligne en italien et anglais), et c comme cornichon.

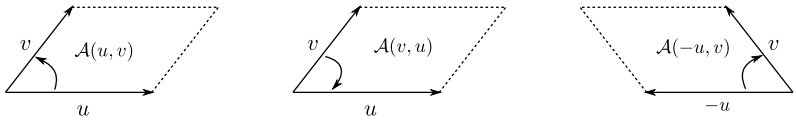
b) Soit $M, M' \in \mathrm{GL}_{n+1}$. On note $M_1, \ldots, M_{n+1} \in \mathbf{k}^{n+1}$ (resp. $M'_1, \ldots, M'_{n+1} \in \mathbf{k}^{n+1}$) les colonnes de M (resp. M'). Montrer que les deux propositions suivantes sont équivalentes :
(i) il existe $H \in P$ tel que $M' = M \times H$;
(ii) $\mathrm{Vect}(M_1, \ldots, M_p) = \mathrm{Vect}(M'_1, \ldots, M'_p)$.

(*Remarque.* Ceci permet de démontrer que l'ensemble quotient $\mathrm{GL}_{n+1}(\mathbf{k})/P$ s'identifie à l'ensemble des sous-espaces vectoriels de dimension p de $\mathbf{k}^{n+1}$.)

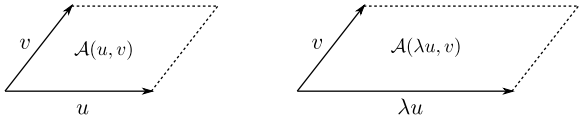
2 – Déterminants

(2.1) Théorème : les formes n-linéaires alternées sur E constituent un $\mathbf{k}$-ev de dimension 1. Soit $\mathcal{B}$ une base. Définition : $\det_{\mathcal{B}}$ est l'unique forme n-linéaire alternée f telle que $f(\mathcal{B}) = 1$.

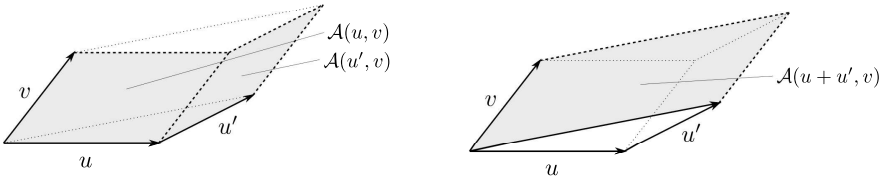
(2.2) Interprétation géométrique. Slogan : « un déterminant c'est un volume (orienté) ». Aire *orientée* : $\mathcal{A}(v,u) = \mathcal{A}(-u,v) = -\mathcal{A}(u,v)$.



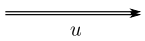
Axiomes d'un volume orienté :
— homogénéité par rapport à chaque facteur : $\mathcal{A}(\lambda u, v) = \lambda \mathcal{A}(u,v)$;



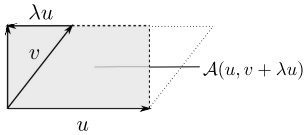
— additivité par rapport à chaque facteur : $\mathcal{A}(u+u',v) = \mathcal{A}(u,v) + \mathcal{A}(u',v)$;



— caractère alterné : $\mathcal{A}(u,u) = 0$.



(2.2.1) Exemple. Aire d'un parallèlogramme = base × hauteur.



(2.2.2) En profiter pour distinguer déterminants d'une famille de vecteurs, d'un endomorphisme, d'une matrice.

(2.2.3) Cas euclidien. Si E est un espace euclidien, on a un choix canonique pour l'unité de volume, à savoir le volume orienté défini par une base orthnormée directe. Le déterminant dans une telle base ne dépend pas du choix de ladite base, et s'appelle le *produit mixte*.

(2.3) Polynôme déterminant.
(2.3.1) Définition. On considère l'anneau $\mathbf{Z}[X_{ij}]$ des polynômes à coefficients entiers en les n^2 indéterminées X_{ij}, $1 \leqslant i, j \leqslant n$. Le *polynôme déterminant* est l'élément

$$\det := \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) X_{\sigma(1),1} \cdots X_{\sigma(n),n} \in \mathbf{Z}[X_{ij}].$$

(2.3.2) Soit A un anneau commutatif unitaire. Le morphisme naturel $\mathbf{Z} \to A$ est défini par

$$n \in \mathbf{Z} \mapsto \underbrace{1_A + \cdots + 1_A}_{n \text{ fois}} \in A$$

(avec l'adaptation habituelle si $n < 0$). Ce morphisme s'étend en un morphisme $\mathbf{Z}[X_{ij}] \to A[X_{ij}]$.
(2.3.3) Définition. Soit A un anneau commutatif unitaire, et $M = (a_{ij}) \in \mathcal{M}_n(A)$ une matrice carrée de taille n à coefficients dans A. Le *déterminant de M* est l'élément $\det(M) \in A$ obtenu en évaluant le polynôme déterminant en les coefficients de M, autrement dit

$$\det(M) := \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \in A.$$

(2.3.4) Proposition. Soit $M = (a_{ij}) \in \mathcal{M}_n(\mathbf{k})$ une matrice carrée de taille n à coefficients dans un corps $\mathbf{k}$. Notons $M_1, \ldots, M_n \in \mathbf{k}^n$ les colonnes de M. Le déterminant de M est le déterminant de la famille de vecteurs $(M_1, \ldots, M_n)$ dans la base canonique de $\mathbf{k}^n$.

(2.4) La bonne façon de calculer un déterminant c'est d'utiliser le pivot de Gauss. Explication par le caractère n-linéaire alterné de la formule pour le déterminant d'une matrice triangulaire : c'est le même calcul qu'en (2.2.1) !
(2.4.1) Exercice. Évaluer la complexité d'un tel calcul.

(2.5) Il y a plusieurs façons de démontrer le fait que toutes les formes n-linéaires alternées sur E sont proportionnelles.

Une première voie est de montrer que toute forme n-linéaire alternée est proportionnelle au polynôme déterminant. Ça se fait par un calcul assez naturel : soit $(e_1, \ldots, e_n)$ une base, f n-linéaire alternée. On a

$$\begin{aligned} f(\sum a_{i1} e_i, \ldots, \sum a_{in} e_i) &= \sum_{i_1, \ldots, i_n} a_{i_1,1} \cdots a_{i_n,n} \cdot f(e_{i_1}, \ldots, e_{i_n}) \\ &= \sum_{\sigma \in \mathfrak{S}_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \cdot f(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in \mathfrak{S}_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \cdot \varepsilon(\sigma) \cdot f(e_1, \ldots, e_n) \end{aligned}$$

(la première égalité est un développement par n-linéarité ; pour la seconde, on observe que par le caractère alterné seuls les n-uplets $(i_1, \ldots, i_n)$ avec les i_s deux à deux distincts contribuent non trivialement, et ces n-uplets sont en bijection avec $\mathfrak{S}_n$; pour la troisième on utilise à nouveau le caractère alterné pour réordonner les arguments de f).

Une autre façon de faire est de procéder par récurrence sur n, en démontrant le caractère nécessaire de la formule de développement par rapport à une colonne. C'est techniquement moins agréable que la première approche.

(2.6) Formule de Laplace. $A = (a_{ij})_{1\leqslant i,j\leqslant n}$, $J = \{j_1 < \cdots < j_p\}$ fixé :

$$\det(A) = \sum_{I=\{i_1<\cdots<i_p\}} (-1)^{|I|+|J|} \det(A_{IJ}) \det(A_{\bar{I}\bar{J}}),$$

où $|I| = i_1 + \cdots + i_p$.

La formule de Laplace généralise la formule de développement par rapport à une colonne (celle-ci est le cas $p = 1$). Cette dernière peut se démontrer soit par un calcul direct sur le polynôme déterminant (c'est le plus agréable) ; sinon on peut la montrer directement, comme dans la preuve par récurrence sur n du théorème énoncé dans (2.1).

Je repousse la preuve de la formule de Laplace en Annexe A. À nouveau on peut procéder par un calcul direct sur le polynôme déterminant, la principale difficulté par rapport au cas $p = 1$ consistant alors en un calcul (un peu fin) de signature de permutation. Sinon on peut procéder par récurrence sur p, le cas de base étant alors le développement par rapport à une colonne.

(2.7) Application 1 : déterminant des matrices triangulaires par blocs.

(2.7.1) Mise en garde dans le cas général = non-triangulaire (si on y réfléchit deux secondes, un tel résultat n'a pas lieu d'être car les blocs n'ont a priori aucune raison d'être carrés tous les quatre).

(2.8) Application 2 : caractérisation du rang par les mineurs. *Le rang d'une matrice est la taille de son plus grand mineur non-nul.*

Il devrait être évident que c'est une très mauvaise façon de calculer le rang en pratique ; comme d'habitude la bonne attitude est d'utiliser le pivot de Gauss. En revanche cet énoncé est très utile d'un point de vue théorique, voir (2.8.2) pour une première utilisation.

(2.8.1) Preuve. On va démontrer l'énoncé équivalent :

$$\forall r \in \mathbf{N} : \quad \text{rang} < r \Leftrightarrow \text{tous les mineurs de taille } r \text{ sont nuls.}$$

Montrons $\Rightarrow$. Soit $A \in \mathcal{M}_{m,n}$ une matrice de rang $< r$. Pour tout $1 \leqslant j_1 < \cdots < j_r \leqslant n$, il existe une relation de dépendance linéaire entre les colonnes de A d'indices $j_1, \ldots, j_r$. Celle-ci induit pour tout $1 \leqslant i_1 < \cdots < i_r \leqslant n$ une relation de dépendance linéaire entre les colonnes de la matrice extraite A_{IJ}, et on a donc $\det(A_{IJ}) = 0$, comme il fallait.

Montrons maintenant $\Leftarrow$ par contraposée. Soit donc $A \in \mathcal{M}_{m,n}$ une matrice de rang $\geqslant r$, et montrons qu'elle possède un mineur de taille r non-nul. D'après le théorème de la base extraite il existe r colonnes de A qui sont linéairement indépendantes. Je choisis r telles colonnes, et je les complète en une base de $\mathbf{k}^m$, en invoquant le théorème de la base incomplète. J'ai alors sous les yeux m colonnes dans $\mathbf{k}^m$, et je considère la matrice $\tilde{A} \in \mathcal{M}_m(\mathbf{k})$ obtenue en les concaténant. Par construction, on a $\det(A) \neq 0$. D'après la formule de Laplace, ce déterminant est une combinaison linéaire de déterminants de taille r extraits des r premières colonnes de $\tilde{A}$, et ainsi une combinaison linéaire de mineurs de taille r de A. Puisque $\det(A) \neq 0$, il est nécessaire que ces mineurs ne soient pas tous nuls. □

(2.8.2) Application à l'invariance du rang par extension de corps. Soit $A \in \mathcal{M}_{mn}(\mathbf{k})$. Pour toute extension de corps $\mathbf{k}'/\mathbf{k}$, on peut voir A comme une matrice à coefficients dans $\mathbf{k}'$; notons ici A' cette matérialisation de A. Le rang de A est la dimension du sev de $\mathbf{k}^m$ engendré par les colonnes de A, celui de A' la dimension du sev de $(\mathbf{k}')^m$ engendré par ces mêmes colonnes. *A priori* il n'y a rien d'évident à ce que ces deux rangs coïncident.

Il est cependant bien vrai que $\text{rg}(A) = \text{rg}(A')$, et ceci se voit bien avec l'énoncé (2.8).

(2.8.3) Preuve de (2.8) sans formule de Laplace. On peut préférer une preuve moins technique de la caractérisation du rang par les mineurs. En voici une naturelle, utilisant de manière essentielle le fait qu'une matrice et sa transposée ont le même rang, voir (1.10).

On utilise le même argument qu'en (2.8.1) pour démontrer que si $\mathrm{rg}(A) = r$, alors tous les mineurs de taille $> r$ de A sont nuls (cet argument est élémentaire, et ne met pas en jeu la formule de Laplace).

Il reste donc à démontrer que si $\mathrm{rg}(A) = r$, alors il existe un mineur de taille r non nul. On commence par extraire r colonnes linéirement indépendantes $A_{j_1}, \ldots, A_{j_r}$ de A. Alors la matrice $(A_{j_1}, \ldots, A_{j_r}) \in \mathcal{M}_{mr}$ est de rang r, donc par (1.10) ses lignes forment une famille rang r. On peut donc extraire r lignes indépendantes de $(A_{j_1}, \ldots, A_{j_r})$; ceci donne une matrice inversible de taille r, dont le déterminant est un mineur de taille r non-nul de A. □

(2.8.4) Remarque. La preuve de la caractérisation du rang par les mineurs avec la formule de Laplace n'utilise pas l'invariance du rang par transposition. Puisque la caractérisation par les mineurs implique que le rang est égal au rang de la transposée, nous avons une troisième preuve de ce dernier énoncé.

(2.9) Formules de Cramer. *Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbf{k})$. On considère sa comatrice* $\mathrm{Com}(A) = \big((-1)^{i+j} A_{ij}\big)$. *On a*

$$A \times {}^{\mathsf{T}}\mathrm{Com}(A) = {}^{\mathsf{T}}\mathrm{Com}(A) \times A = \det(A).\mathrm{Id}.$$

Preuve. On utilise la formule pour développer un déterminant par rapport à une ligne (resp. colonne) pour calculer $A \times {}^{\mathsf{T}}\mathrm{Com}(A)$ (resp. ${}^{\mathsf{T}}\mathrm{Com}(A) \times A$).

Le coefficients d'indice i, j de $A \times {}^{\mathsf{T}}\mathrm{Com}(A)$ est

$$\sum_{k=1}^{n} a_{ik}(-1)^{k+j} A_{jk} = \det \begin{pmatrix} L_1 \\ \hline \vdots \\ \hline L_{j-1} \\ \hline L_i \\ \hline L_{j+1} \\ \hline \vdots \\ L_n \end{pmatrix} = \delta_{ij} \det(A)$$

(ici L_i est la i-ème ligne de A, et la matrice ci-dessus est obtenue à partir de A en mettant la i-ème ligne à la place de la j-ème). De la même façon, Le coefficients d'indice i, j de ${}^{\mathsf{T}}\mathrm{Com}(A) \times A$ est

$$\sum_{k=1}^{n} (-1)^{k+i} A_{ki} a_{kj} = \det \left(\begin{array}{c|c|c|c|c|c|c} & & & & & & \\ C_1 & \cdots & C_{i-1} & C_j & C_{i+1} & \cdots & C_n \\ & & & & & & \end{array} \right) = \delta_{ij}.\det(A).$$

□

(2.9.1) Corollaire : preuve de (1.5) sans théorie de la dimension. Supposons $AB = \mathrm{Id}$. Alors $\det(A) \cdot \det(B) = 1$, donc en posant $A' = \det(B).\,\mathrm{Com}(A)$, on a

$$A \times A' = A' \times A = \det(B)\det(A).\mathrm{Id} = \mathrm{Id}$$

par les formules de Cramer. Ainsi A est inversible. Ensuite $AB = \mathrm{Id}$ implique (en multipliant à gauche par A^{-1}) que $B = A^{-1}$. Ceci prouve (ii) ⇒ (i). L'implication (iii) ⇒ (i) se démontre de manière analogue. □

(2.9.2) Dans le même ordre d'idée, ce sont les formules de Cramer qui permettent de démontrer II.(3.1)

(2.10) Exercice. Déterminer le rang de la comatrice $\mathrm{Com}(A)$ en fonction du rang de A.

3 – Polynôme caractéristique

(3.1) Convention : $\chi_A = \det(X \cdot \mathrm{Id} - A)$. Définition*s* des valeurs propres.

(3.2) Application de la formule de Laplace : calcul des coefficients de χ_M.

$$\det(X \cdot e_1 - C_1, \ldots, X \cdot e_n - C_n) = \sum_{k=0}^{n} X^k \left(\sum_{i_1 < \cdots < i_k} \det(-C_1, \ldots, e_{i_1}, \ldots, e_{i_k}, \ldots, -C_n) \right)$$

et les déterminants dans la somme entre parenthèses à droite se calculent en un coup avec la formule de Laplace.
(3.2.1) Exercice. Étudier $\det(A + X \cdot M)$.
(3.2.2) Corollaire. $M \mapsto \chi_M$ est continue (on verra que ce n'est pas le cas de $M \mapsto \mu_M$). Si $\mathbf{k} = \mathbf{R}$ ou $\mathbf{C}$, sinon topologie de Zariski.

(3.3) Développement. *Soit $A, B \in \mathcal{M}_n(\mathbf{k})$. On a $\chi_{AB} = \chi_{BA}$.*

Je présente ce résultat essentiellement motivé par la méthode de preuve. Toutefois, cet énoncé a son intérêt. C'est une généralisation de la formule bien connue $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. En fait, pour qui connaît un peu d'algèbre extérieure ce n'est pas vraiment une généralisation, puisqu'au signe près les coefficients du polynôme caractéristique χ_M sont les traces des matrices $M^{\wedge i}$, $i = 0, \ldots, n$.
(3.3.1) Cas A inversible. Dans ce cas le résultat s'obtient par un calcul qui ne dévoile pas vulgairement la raison profonde de notre énoncé. On calcule des déterminants de matrices à coefficients dans l'anneau commutatif $\mathbf{k}[X]$.

$$\begin{aligned} \chi_{AB} = \det(X.\mathrm{id} - AB) &= \det\big[A(X.A^{-1} - B)\big] \\ &= \det\big[(X.A^{-1} - B)A\big] = \det(X.\mathrm{id} - BA) = \chi_{BA} \end{aligned}$$

et c'est gagné.

Voyons maintenant comment déduire le résultat sans condition sur A par un argument de passage à la limite qui peut se formuler de plusieurs façons.
(3.3.2) Version 1. Si $\mathbf{k}$ est infini, il existe une infinité de $\lambda \in \mathbf{k}$ tels que $A - \lambda.\mathrm{Id}$ est inversible. Pour tous ces λ, on a

$$\chi_{(A-\lambda \mathrm{Id})B} = \chi_{B(A-\lambda \mathrm{Id})}.$$

Les coefficients du polynôme en X $\chi_{(A-\lambda \mathrm{Id})B} - \chi_{B(A-\lambda \mathrm{Id})}$ sont des polynômes en λ, nuls pour une infinité de valeurs de λ, et donc uniformément nuls. Autrement dit, l'annulation de $\chi_{(A-\lambda \mathrm{Id})B} - \chi_{B(A-\lambda \mathrm{Id})} \in \mathbf{k}[X]$ pour une infinité de λ implique son annulation tout court. Pour $\lambda = 0$, ceci nous donne le résultat voulu.

Si $\mathbf{k}$ n'est pas infini, on trouve une extension $\mathbf{k}'$ de $\mathbf{k}$ qui l'est, par exemple $\mathbf{k}(T)$ (fraction rationnelles en T). Par le même argument on obtient l'identité $\chi_{AB} = \chi_{BA}$ qui se fiche pas mal de savoir si on habite dans $\mathcal{M}_n(\mathbf{k})$ ou $\mathcal{M}_n(\mathbf{k}')$.

(3.3.3) Version 2. Si $\mathbf{k} = \mathbf{R}$ ou $\mathbf{C}$, ou encore mieux si on connaît la topologie de Zariski, alors on peut arguer que

$$(A, B) \in \mathcal{M}_n(\mathbf{k}) \times \mathcal{M}_n(\mathbf{k}) \mapsto \chi_{AB} - \chi_{BA} \in \mathbf{k}[X]_n$$

est continue et nulle sur l'ouvert dense des (A, B) tels que A inversible, et donc uniformément nulle.

(3.3.4) Version 3. On considère les deux matrices $A = (a_{ij})$ et $B = (b_{ij})$ à coefficients indéterminés. Il s'agit de démontrer l'identité $\chi_{AB} = \chi_{BA}$ entre polynômes à coefficients dans l'anneau de polynômes $\mathbf{Z}[a_{ij}, b_{ij}]$ (polynômes en $2n^2$ indéterminées, à coefficients entiers) : $\chi_{AB}, \chi_{BA} \in \mathbf{Z}[a_{ij}, b_{ij}][X]$.

On considère le corps des fractions de cet anneau intègre, qui est le corps de fractions rationnelles $\mathbf{K} := \mathbf{Q}(a_{ij}, b_{ij})$. Puisque le déterminant est un polynôme non-nul, A est inversible dans $\mathcal{M}_n(\mathbf{K})$. On en déduit par notre calcul basique l'identité $\chi_{AB} = \chi_{BA}$ dans $\mathbf{K}[X]$. Bien sûr les deux polynômes sont en fait dans le sous-anneau $\mathbf{Z}[a_{ij}, b_{ij}][X]$ de $\mathbf{K}[X]$, et l'identité $\chi_{AB} = \chi_{BA}$ est une égalité entre polynômes à coefficients dans $\mathbf{Z}[a_{ij}, b_{ij}]$: notre résultat est démontré !

Remarque. Le couple (A, B) comme ci-dessus est le point générique au sens de la géométrie algébrique de l'espace affine $\mathcal{M}_n(\mathbf{Z}) \times \mathcal{M}_n(\mathbf{Z})$.

(3.4) Preuve directe de Cayley–Hamilton. En préambule, on propose de se demander pourquoi la preuve consistant à dire « j'évalue $\chi_A(X) = \det(X.\mathrm{Id} - A)$ en A, ainsi $\chi_A(A) = \det(A \times \mathrm{Id} - A) = \det(0) = 0$ » est irrémédiablement privée de sens. [2]

Preuve (de Cayley–Hamilton). Avant de se lancer pour de bon dans la preuve, on rappelle l'identité

$$X^k.\mathrm{Id} - A^k = (X.\mathrm{Id} - A) \underbrace{\sum_{0 \leqslant j \leqslant k-1} X^{k-1-j}.A^j}_{=:Q_{A,k}(X)} \tag{3.4.1}$$

valable puisque A et $X.\mathrm{Id}$ commutent.

Le point de départ est la formule de Cramer pour la matrice $X.\mathrm{Id} - A \in \mathcal{M}_n(\mathbf{k}[X])$,

$$(X.\mathrm{Id} - A) \times \mathrm{Com}(X.\mathrm{Id} - A) = \det(X.\mathrm{Id} - A).\mathrm{Id} = \chi_A(X).\mathrm{Id}. \tag{3.4.2}$$

D'autre part, notant $\chi_A(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0$ (on a $a_n = 1$, mais OSEF), on obtient grâce à (3.4.1)

$$\chi_A(X).\mathrm{Id} - \chi_A(A) = \sum\nolimits_k a_k (X^k.\mathrm{Id} - A^k) = (X.\mathrm{Id} - A) \times \sum\nolimits_k a_k Q_{A,k}(X) \tag{3.4.3}$$

Finalement,

$$\begin{aligned} \chi_A(A) &= \chi_A(X).\mathrm{Id} - \big(\chi_A(X).\mathrm{Id} - \chi_A(A)\big) \\ &= (X.\mathrm{Id} - A) \times \Big[\mathrm{Com}(X.\mathrm{Id} - A) - \sum\nolimits_k a_k Q_{A,k}(X)\Big]. \end{aligned}$$

On a envie de dire que $X.\mathrm{Id} - A$, polynôme de degré 1 (à coefficients dans $\mathcal{M}_n(\mathbf{k})$), divise $\chi_A(A)$ qui est de degré $\leqslant 0$, donc nécessairement $\chi_A(A) = 0$; l'anneau $\mathcal{M}_n(\mathbf{k})$ n'est pas intègre donc il faut être prudent, mais ici tout va aller car le coefficient dominant de $X.\mathrm{Id} - A$ est inversible.

2. Indice sur votre écran : '.' et '×' ce n'est pas la même chose.

Pour éviter tout ressentiment, écrivons l'argument explicitement. Notant $X^N.B_N + \cdots + X.B_1 + B_0$ le terme entre crochets, on a

$$\begin{aligned}\chi_A(A) &= (X.\mathrm{Id} - A) \times \left[X^N.B_N + \cdots + X.B_1 + B_0\right] \\ &= X^{N+1}.B_N + X^N.(B_{N-1} - A \times B_N) + \cdots + X.(B_0 - A \times B_1) - A \times B_0,\end{aligned}$$

dont on déduit par identification des coefficients

$$B_N = B_{N-1} - A \times B_N = B_0 - A \times B_1 = -A \times B_0 - \chi_A(A) = 0,$$

ce qui de proche en proche donne $\chi_A(A) = 0$ comme il fallait démontrer. □

4 – Matrices par blocs et sous-espaces stables ; quotients

Avertissement. Il y a deux définition concurrentes de projection : (i) $\pi : E \to E/F$ et ses avatars après choix d'un supplémentaire ; celle-ci n'est pas un endomorphisme en général ; (ii) $p \in \mathcal{L}(E)$ projection sur F_1 dans la direction de F_2 ; ce p est un endomorphisme de E, et satisfait à la relation $p^2 = p$. Dans ce paragraphe, c'est plutôt la version (i) qui est utilisée.

4.1 – Matrices par blocs

(4.1) Sommes directes. On dit que $E_1 \oplus \cdots \oplus E_r = E$ si pour tout $x \in E$ il existe un unique $(x_1, \ldots, x_r) \in E_1 \times \cdots \times E_r$ tel que $x = x_1 + \cdots + x_r$.

(4.2) Décomposition d'une application linéaire selon une somme directe. On suppose $E = E_1 \oplus \cdots \oplus E_r$ et $F = F_1 \oplus \cdots \oplus F_s$. Soit $f \in \mathcal{L}(E, F)$. Il existe une unique famille d'applications linéaires $f_{ij} \in \mathcal{L}(E_j, F_i)$ telle que

$$\forall (x_1, \ldots, x_r) \in E_1 \times \cdots \times E_r,\ f(x_1 + \cdots + x_r) = \sum\nolimits_{ij} f_{ij}(x_j). \tag{4.2.1}$$

Réciproquement, pour toute famille d'applications linéaires $f_{ij} \in \mathcal{L}(E_j, F_i)$, il existe une unique $f \in \mathcal{L}(E, F)$ telle que (4.2.1) soit vérifiée.

(4.2.2) Exemple. Pour $E = E_1 + E_2$, définition de l'endomorphisme de E projecteur sur E_1 dans la direction de E_2, ainsi que de la projection de E sur E_1 dans la direction de E_2.

(4.2.3) Remarque. On a $f_{ij} = \pi_{F,F_i} \circ f \circ \iota_{E_j,E} \in \mathcal{L}(E_j, F_i)$.

(4.3) Bases compatibles à une somme directe. Si $\mathcal{B}_1, \ldots, \mathcal{B}_r$ sont des bases de $E_1, \ldots, E_r$ respectivement et $E = E_1 \oplus \cdots \oplus E_r$, alors $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$ est une base de E.

Dans les conditions du (4.2), si $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_r)$ et $\mathcal{D} = (\mathcal{D}_1, \ldots, \mathcal{D}_s)$ sont des bases de E et F respectivement, compatibles aux sommes directes, alors $\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(f)$ se décompose par blocs $\mathrm{Mat}_{\mathcal{B}_j,\mathcal{D}_i}(f_{ij})$.

(4.4) Composition/produit par blocs. On considère trois espaces vectoriels décomposés en sommes directes $E = E_1 \oplus \cdots \oplus E_r$, $F = F_1 \oplus \cdots \oplus F_s$, $G = G_1 \oplus \cdots \oplus G_t$, et deux applications linéaires $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$ décomposées respectivement en f_{ij} et g_{ij} comme en (4.2.1). On a pour tout $j = 1, \ldots, r$ et $i = 1, \ldots, t$

$$(g \circ f)_{ij} = \sum\nolimits_{1 \leqslant k \leqslant s} g_{ik} \circ f_{kj}.$$

Le calcul à faire pour démontrer cette formule est exactement le même que celui qui établit la formule

$$\mathrm{Mat}(g \circ f) = \mathrm{Mat}(g) \times \mathrm{Mat}(f).$$

C'est ce qui explique les formules pour un produit de matrices par blocs, qui se résument en disant "qu'on peut prétendre que les blocs sont des scalaires". Attention juste au fait que les produits de blocs ne sont pas commutatifs.

4.2 – Rappels et compléments sur les quotients

(4.5) Définition. etc.

(4.6) Lemme. *Considérons une filtration*

$$\{0\} = F_0 \subseteq F_1 \subseteq \cdots \subseteq F_p = E.$$

Soit une famille de vecteurs

$$\varepsilon_1, \ldots, \varepsilon_{r_1}, \varepsilon_{r_1+1}, \ldots, \varepsilon_{r_2}, \varepsilon_{r_2+1}, \ldots\ldots, \varepsilon_{r_{p-1}+1}, \ldots, \varepsilon_{r_p} \in E. \tag{4.6.1}$$

Si pour tout $i = 1, \ldots, p$, (i) les vecteurs $\varepsilon_{r_{i-1}+1}, \ldots, \varepsilon_{r_i}$ sont dans F_i, et (ii) leurs classes constituent une base de F_i/F_{i-1}, alors la famille (4.6.1) *est une base de E.*

Preuve. Par récurrence sur p, il suffit de savoir le faire pour $p = 2$. Soit F un sous-espace vectoriel de E, $(e_1, \ldots, e_p)$ base de F, et $e_{p+1}, \ldots, e_{p+q} \in E$ tels que $(\bar{e}_{p+1}, \ldots, \bar{e}_{p+q})$ soit une base du quotient E/F. Montrons que $(e_1, \ldots, e_p, e_{p+1}, \ldots, e_{p+q})$ est une base de E.

Soit $\lambda_1, \ldots, \lambda_{p+q}$ tels que

$$\lambda_1.e_1 + \cdots + \lambda_p.e_p + \lambda_{p+1}e_{p+1} + \cdots + \lambda_{p+q}.e_{p+q} = 0. \tag{4.6.2}$$

Après projection dans E/F, reste

$$\lambda_{p+1}\bar{e}_{p+1} + \cdots + \lambda_{p+q}.\bar{e}_{p+q} = 0,$$

qui implique $\lambda_{p+1} = \cdots = \lambda_{p+q} = 0$. (4.6.2) devient alors une combinaison linéaire nulle de $e_1, \ldots, e_p$, qui à son tour donne $\lambda_1 = \cdots = \lambda_p = 0$. Ceci prouve la liberté.

D'autre part, soit $x \in E$. Sa projection $\bar{x}$ dans le quotient E/F s'écrit comme une combinaison linéaire

$$\bar{x} = \lambda_{p+1}.\bar{e}_{p+1} + \cdots + \lambda_{p+q}.\bar{e}_{p+q}.$$

Considérons $\tilde{x} = \lambda_{p+1}.e_{p+1} + \cdots + \lambda_{p+q}.e_{p+q} \in E$. Le vecteur $x - \tilde{x}$ appartient à F, puisque sa classe modulo F est nulle, en conséquence de quoi il s'écrit comme combinaison linéaire des $e_1, \ldots, e_p$. □

(4.6.3) Remarque au passage. Pour montrer l'égalité

$$\dim\big(E/F\big) = \dim E - \dim F,$$

on fait comme dans la preuve ci-dessus.

(4.7) Propriété universelle du quotient. *Le quotient E/F jouit des deux propriétés suivantes :*
(a) il existe $\pi : E \to E/F$ surjective et de noyau F ;
(b) pour tout G, l'application linéaire (c'en est une)

$$u \in \mathrm{Hom}(E/F, G) \mapsto u \circ \pi \in \mathrm{Hom}(E, G)$$

induit un isomorphisme (fonctoriel...)

$$\Phi : \mathrm{Hom}(E/F, G) \cong \ker\big(\mathrm{restr} : \mathrm{Hom}(E, G) \to \mathrm{Hom}(F, G)\big).$$

S'il existe E' jouissant lui aussi de l'une ou l'autre de ces deux propriétés, alors il existe un unique isomorphisme $\phi : E/F \cong E'$ tel que $\pi' = \phi \circ \pi$.

NB : écrire plutôt $\mathcal{L}(E, F)$ que $\mathrm{Hom}(E, F)$.

(4.7.1) Preuves. On laisse au lecteur la preuve de la propriété (a) pour E/F et sa projection canonique. On va montrer que les propriétés (a) et (b) sont équivalentes, puis qu'un **k**-ev satisfaisant à ces propriétés s'identifie canoniquement à E/F.

Soit E' vérifiant (a) (*i.e.*, on suppose E' **k**-ev muni de $\pi' : E \to E'$ linéaire, surjective et de noyau F), et montrons que (b) vaut pour E' (cela montrera au passage que (b) vaut bien pour le quotient). On commence par constater que $u \mapsto u \circ \pi'$ donne bien une application linéaire

$$\Phi_{\pi'} : \mathcal{L}(E', G) \to \ker\big(\mathrm{restr} : \mathcal{L}(E, G) \to \mathcal{L}(F, G)\big)$$

pour tout **k**-ev G. Reste à voir que c'est effectivement un isomorphisme. Si $u \circ \pi' = 0$, alors $u = 0$ car π' est surjective, d'où l'injectivité de $\Phi_{\pi'}$.

Pour la surjectivité, on construit à la main un antécédent pour toute application linéaire $v : E \to G$ s'annulant sur F. Pour tout $y \in E'$ il existe $x \in E$ tel que $y = \pi'(x)$, et on pose $u(y) := v(x)$; ça ne dépend pas du choix de x car $\ker(\pi') = F$ et $v|_F = 0$. Ceci définit une application linéaire $u : E' \to G$ qui visiblement factorise comme il faut (*i.e.*, $v = u \circ \pi'$). □

Réciproquement, soit E' vérifiant (b), et montrons directement que (a) vaut pour E' (à nouveau, cela montrera au passage que le quotient catégoriel est bien le quotient défini par la relation de congruence). Précisément, on suppose qu'il existe une application linéaire $\pi' : E \to E'$ tel que $\Phi_{\pi'}$ soit un isomorphisme pour tout G, et il s'agit de montrer que π' est surjective et de noyau F.

On regarde le $\Phi_{\pi'}$ pour $G = E'$: puisque $\pi' = \Phi_{\pi'}(\mathrm{id}_{E'})$, π' est dans le noyau de la restriction à F, donc $F \subseteq \ker(\pi')$. D'autre part, on regarde le $\Phi_{\pi'}$ pour $G = E/F$: puisque π est dans le noyau de la restriction à F, il existe un $u : E' \to E/F$ tel que $\pi = u \circ \pi'$. Ceci implique $\ker(\pi') \subseteq \ker(\pi) = F$, donc on conclut que $\ker(\pi') = F$.

D'autre part, si π' n'était pas surjective on pourrait contredire l'injectivité de $\Phi_{\pi'}$ pour n'importe quel $G \neq \{0\}$ de la manière suivante. Soit $x \in E'$ non atteint par π'. On considère G un **k**-ev non nul et $u' : E' \to G$. Alors pour g n'importe quel vecteur non nul de G et ℓ une forme linéaire de noyau un hyperplan transverse à x, on a

$$(u' + \ell.g) \circ \pi' = u' \circ \pi',$$

ce qui contredit l'injectivité de $\Phi_{\pi'}$. □

Soit E' vérifiant (a). Alors $\pi' \in \mathcal{L}(E, E')$ est dans le noyau de la restriction à F, donc il existe un unique $\phi : E/F \to E'$ tel que $\pi' = \phi \circ \pi$. Reste à voir que ce ϕ est un isomorphisme. Puisque π est surjective, $\ker \phi \neq \{0\} \Rightarrow \ker \pi' \supsetneq F$ et donc nécessairement ϕ est injective. D'autre part ϕ est surjective car π' l'est. □

(4.8) Quotient et supplémentaire. Choisissons un supplémentaire F' de F dans E. On dispose de la projection $\tilde{p} : E \to F'$ dans la direction de F, qui est surjective et de noyau F. D'après la propriété universelle du quotient, il existe un isomorphisme canonique entre F' et E/F. Cette réalisation du quotient dépend d'un choix et n'est donc pas canonique.

Le choix d'un supplémentaire de F dans E équivaut au choix d'une injection linéaire $j : E/F \hookrightarrow E$ telle que $\mathrm{im}(j) \cap F = \{0\}$.

4.3 – Sous-espaces stables et matrices triangulaires par blocs

(4.9) F est stable par f si pour tout $x \in F$, $(x) \in F$. Dans ce cas on a un *endomorphisme* induit $f_F \in \mathcal{L}(F)$.

(On note que $u_F = \pi_F \circ u|_F \in \mathcal{L}(F)$ ne dépend que de F, et pas du choix du supplémentaire choisi pour définir la projection $\pi_F \in \mathcal{L}(E, F)$).

(4.10) Proposition. *Si f a une matrice triangulaire supérieure par blocs (taille p) dans la base $(e_1, \ldots, e_p, \ldots, e_n)$, alors $F := \mathrm{Vect}(e_1, \ldots, e_p)$ est stable par f.*

Réciproquement, si F est un sev stable par f, alors pour toute base $\mathcal{B}$ de E obtenue en complétant une base $\mathcal{B}_F$ de F, $\mathrm{Mat}_{\mathcal{B}}(f)$ est triangulaire supérieure par blocs, et le bloc supérieur est $\mathrm{Mat}_{\mathcal{B}_F}(f_F)$.

On va interpréter l'autre bloc diagonal en termes du quotient E/F. On en profite pour faire quelques rappels et compléments sur les quotients.

(4.11) Endomorphisme induit sur le quotient par un sous-espaces stables. Si F est stable par u, on a aussi canoniquement défini un $\bar{u}_F \in \mathcal{L}(E/F)$.

(4.11.1) Lemme. Pour $\mathcal{B} = (\mathcal{B}_F, \mathcal{B}')$ comme ci-dessus, $\bar{\mathcal{B}}'$ est une base de E/F.

(4.11.2) Le second bloc diagonal de $\mathrm{Mat}_{\mathcal{B}}(f)$ est $\mathrm{Mat}_{\bar{\mathcal{B}}'}(\bar{u_F})$.

(4.11.3) Pour un autre choix de base $\mathcal{D} = (\mathcal{D}_F, \mathcal{D}')$ comme ci-dessus, on a une autre matrice triangulaire supérieure par blocs. Les blocs diagonaux sont deux à deux semblables.

Preuve. Les blocs diagonaux sont respectivement $\mathrm{Mat}_{\mathcal{B}_F}(u_F)$ et $\mathrm{Mat}_{\mathcal{D}_F}(u_F)$ d'une part, et $\mathrm{Mat}_{\bar{\mathcal{B}}'}(\bar{u_F})$ et $\mathrm{Mat}_{\bar{\mathcal{D}}'}(\bar{u_F})$ d'autre part. □

(4.12) Attention ! Pas de supplémentaire stable en général. Même si u définit canoniquement un endomorphisme de E/F, pour F' supplémentaire de F dans E, F' n'est en général pas stable par u ; pire, en général aucun supplémentaire F' n'est stable par u. Ainsi $\bar{u}_F$ n'est pas $u_{F'}$ puisque ce dernier *n'existe pas*.

Les endomorphismes dont tous les stables ont un supplémentaire stable portent un nom, ce sont les semi-simples. Sur un corps algébriquement clos, ce sont les diagonalisables.

Ceci traduit le fait qu'un sous-module n'admet en général pas de supplémentaire, voir partie II.

(4.13) Remarque. La donnée de u_F et $\bar{u}_F$ ne suffit pas à reconstruire u. Exemple par le contre-exemple universel de la réduction des endomorphismes :

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \textit{vs.} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Ceci traduit le fait que les suites exactes de modules ne sont pas toutes scindées.

(4.14) Exercice. Si F est stable par $u, v \in \mathcal{L}(E)$, alors

$$(u \circ v)_F = u_F \circ v_F \quad \text{et} \quad (\overline{u \circ v})_F = \bar{u}_F \circ \bar{v}_F.$$

On retrouve ainsi les formules pour les produits de matrices triangulaires par blocs.

(4.15) Proposition. *On suppose toujours F stable par u. Alors,*
(i) $\det(u) = \det(u_F) \cdot \det(\bar{u}_F)$ *;*
(ii) $\chi_u = \chi_{u_F} \cdot \chi_{\bar{u}_F}$.
Preuve. Utiliser les formules pour le déterminant des matrices triangulaires par blocs. □

(4.15.1) Remarque. En lien avec (4.13), on remarque que la donnée de u_F et $\bar{u}_F$ suffit pour déterminer χ_u, mais c'est bien trop peu pour prétendre comprendre u. Il manque d'autre invariants, les invariants de similitude.

5 – Dualité

(5.1) Dual, bidual. $E^\vee$ (ou E^*) est le $\mathbf{k}$-ev des formes linéaires sur E.
On a donc $E^\vee = \mathcal{L}(E, \mathbf{k})$, et $E^\vee$ est de dimension n.
(5.1.1) Exercice important. Montrer que

$$\llcorner : e \in E \mapsto (\llcorner e : \ell \in E^\vee \mapsto \ell \llcorner e := \ell(e) \in \mathbf{k})$$

établit un isomorphisme canonique $E \cong (E^\vee)^\vee$. On peut préférer à la notation 'produit intérieur' la notation 'évaluation', $\mathrm{ev}_x : \ell \mapsto \ell(x)$.

(Ici on utilise de manière essentielle le fait que E est de dimension finie : on montre que $\llcorner$ est injective, et on conclut par un argument de dimension). [3]

(5.2) Base duale. Voici une formulation qui n'est pas intrinsèque. Étant donnée une base $(e_1, \ldots, e_n)$ de E, on a un système de coordonnées $(x_1, \ldots, x_n)$ sur E. L'espace $E^\vee$ des formes linéaires sur E est le $\mathbf{k}$-ev des polynômes *homogènes* de degré 1 en $x_1, \ldots, x_n$; la famille $(x_1, \ldots, x_n)$ en est une base.

Définition de la base duale. Remarquer que les $e_1^\vee, \ldots, e_n^\vee$ sont les $dx_1, \ldots, dx_n$ du calcul intégral.
(5.2.1) Attention. Si on se donne juste un vecteur $e \in E$ (tout seul), $e^\vee$ n'est pas défini (il faut tout une base).
(5.2.2) Si L est la colonne (!) des coordonnées de ℓ dans $\mathcal{B}^\vee$, la matrice de ℓ dans $\mathcal{B}$ et la base canonique de $\mathbf{k}$ est la ligne ${}^{\mathsf{T}}L$.

La base canonique de $\mathbf{k}$ étant plus que canonique, j'allège la notation $\mathrm{Mat}_{\mathcal{B},\mathcal{B}_{can}}$ en $\mathrm{Mat}_{\mathcal{B}}$.
(5.2.3) Lemme. $\ell = \ell(e_1).e_1^\vee + \cdots + \ell(e_n).e_n^\vee$.
(5.2.4) Changement de base. $\mathrm{Mat}(\mathcal{B}_F^\vee, \mathcal{B}_E^\vee) = {}^{\mathsf{T}}(\mathrm{Mat}(\mathcal{B}_E, \mathcal{B}_F))$ est la formule naturelle ; si on insiste pour avoir $\mathrm{Mat}(\mathcal{B}_E^\vee, \mathcal{B}_F^\vee)$, c'est ${}^{\mathsf{T}}(\mathrm{Mat}(\mathcal{B}_E, \mathcal{B}_F))^{-1}$.
Exemple. Soit E une droite engendrée par $e \neq 0$. Pour $\lambda \in \mathbf{k}^*$, $\lambda.e$ est une autre base de E. On a :

$$(\lambda e)^\vee = \frac{1}{\lambda} e^\vee.$$

3. sinon il faut rajouter des hypothèses de complétude, et se restreindre aux formes linéaires continues. Soit H un espace de Hilbert. L'inégalité de Cauchy–Schwartz assure que pour tout $x \in H$, la forme linéaire $\langle x, \cdot \rangle$ est continue. Le théorème de représentation de Riesz assure que la réciproque est vraie. (Ne pas confondre avec le théorème de convexité de Riesz).

En effet :

$$(\lambda e)^\vee(e) = \frac{1}{\lambda}(\lambda e)^\vee(\lambda e) = \frac{1}{\lambda} = \frac{1}{\lambda}e^\vee(e).$$

(5.2.5) Exercice. $(\mathcal{B}^\vee)^\vee \cong \mathcal{B}$ par l'isomorphisme de bidualité.

(5.3) Transposition. Pour $u \in \mathcal{L}(E,F)$, ${}^\mathsf{T}u \in \mathcal{L}(F^\vee, E^\vee)$ est

$$\ell \mapsto \ell \circ u.$$

(5.3.1) $\mathrm{Mat}_{\mathcal{B}_F^\vee,\mathcal{B}_E^\vee}({}^\mathsf{T}u) = {}^\mathsf{T}\big(\mathrm{Mat}_{\mathcal{B}_F,\mathcal{B}_E}(u)\big)$.
Preuve. Je conserve le système de notation de (5.2.2) : L est la colonne de ℓ dans $\mathcal{B}_F^\vee$. La matrice ligne de ${}^\mathsf{T}u(\ell)$ est

$$\mathrm{Mat}_{\mathcal{B}_E}\big({}^\mathsf{T}u(\ell)\big) = {}^\mathsf{T}L \times A,$$

donc ${}^\mathsf{T}u$ s'écrit matriciellement $L \mapsto {}^\mathsf{T}A \times L$. □
Exercice. Retrouver la formule de changement de base (5.2.4) à partir de (5.3.1).
(5.3.2) Exercice (abscons). ${}^\mathsf{T}({}^\mathsf{T}u) = u$ *via* les isomorphismes canoniques $(E^\vee)^\vee \cong E$.

(5.4) Orthogonalité. Définition. L'orthogonalité renverse les inclusions. $(A,B)^\perp = A^\perp \cap B^\perp$; $A^\perp + B^\perp \subseteq (A \cap B)^\perp$. $F \subseteq (F^\perp)^\perp$ *via* l'isomorphisme de bidualité.
(5.4.1) $\ker({}^\mathsf{T}u) = (\mathrm{im}(u))^\perp$; $\mathrm{im}({}^\mathsf{T}u) = (\ker u)^\perp$.
(5.4.2) $(E/F)^\vee \cong F^\perp$. ${}^\mathsf{T}\big(\pi : E \twoheadrightarrow E/F\big) = \big(\iota : F^\perp \hookrightarrow E^\vee\big)$.

(5.5) Dualité géométrique. On a une correspondance biunivoque entre les sev de dim r de $E^\vee$ et les sev de dim $c := n - r$ de E, donnée par $\Lambda \mapsto \Lambda^\perp$. Voir Lemme (1.8).

Cas particulier important : l'ensemble des hyperplans (vectoriels) de E s'identifie à l'ensemble des droites (vectorielles) de $E^\vee$. Voici un contexte dans lequel la géométrie projective apparaît de manière naturelle. [4]

(5.5.1) Preuve. L'argument de dimension est inévitable, dans la mesure où cet énoncé ne fonctionne pas en dimension infinie (voir exemple (5.5.3)). Le truc à comprendre est que $\Lambda^\perp$ est un sev de E défini par un système d'équations linéaires. Ensuite, une bonne façon de comprendre la formule $\dim \Lambda^\perp = \dim E - \dim \Lambda$ est d'utiliser le pivot de Gauss. Une autre possibilité est de prendre $(\ell_1, \ldots, \ell_r)$ base de Λ, de la compléter en une base de $E^\vee$, et de constater en utilisant la base duale que $\dim(\ell_1, \ldots, \ell_r)^\perp = n - r$.

(5.5.2) Remarque. Caché là-dedans se trouve le fait que "rang = rang de la transposée". En effet, considérons une matrice A. Le calcul de $\dim \Lambda^\perp$ nous dit que la dimension du noyau de A est n (nombre de colonnes) moins le rang des lignes de A. D'autre part, le théorème du rang nous dit que la dimension du noyau est aussi n moins le rang des colonnes de A.

(5.5.3) Application. En dimension finie, on a $A^\perp + B^\perp = (A \cap B)^\perp$ et $F = (F^\perp)^\perp$.
Exemple. En dimension infinie, prenant un peu d'avance sur la dualité par rapport à une forme quadratique, on peut considérer la forme bilinéaire non-dégénérée sur $\mathbf{k}[X]$ standard

$$\langle a_d X^d + \cdots + a_0, b_d X^d + \cdots + b_0 \rangle = \sum a_i b_i,$$

et $F = \{P : P(1) = 0\}$. On vérifie que $F^\perp = \{0\}$ et donc $(F^\perp)^\perp = \mathbf{k}[X]$.

4. je ne résiste pas ! Ce contexte naturel est un cas particulier de celui fondamental des systèmes linéaires : les hyperplans de $\mathbf{k}^{n+1}$ -ou de $\mathbf{P}^n$!) forment un espace projectif.

(5.5.4) Application. Calculer un système d'équation de $\mathrm{Vect}(u_1, \ldots, u_r)$ connaissant les coordonnées de chaque u_i dans $\mathcal{B}$, sous l'hypothèse que les u_i sont linéairement indépendants. i) en trouvant un mineur non-nul dans la matrice des u_i ; ii) il est bon de parler du pivot dans ce contexte.

Une fois qu'on a $A_r \in \mathrm{GL}_r(\mathbf{k})$,

$$(5.5.5) \qquad \left(\begin{array}{c|c} & x_1 \\ A_r & \vdots \\ & x_r \\ \hline L_{r+1} & x_{r+1} \\ \vdots & \vdots \\ L_n & x_n \end{array}\right) \iff \forall i = 1, \ldots, n-r,\ \det \left(\begin{array}{c|c} & x_1 \\ A_r & \vdots \\ & x_r \\ \hline L_{r+i} & x_{r+i} \end{array}\right) = 0.$$

En effet, on cherche les équations d'un espace de dimension r. Les équations du côté droit de (5.5.5) sont $n - r$ équations linéaires indépendantes vérifiées par les vecteurs de notre espace. Elles suffisent pour définir notre sous-espace.

On retrouve essentiellement la formule pour l'équation d'un plan dans $\mathbf{k}^3$ avec le produit vectoriel de deux vecteurs générateurs.

Un autre bon point pour la géométrie projective : si on la connaît, la méthode ci-dessus permet aussi d'écrire les équations des sous-espaces affines.

(5.5.6) Application. Savoir calculer la dimension d'une intersection d'hyperplans (rapport du jury).

I – Réduction des endomorphismes : analyse

1 – Polynômes d'endomorphismes, polynômes annulateurs

(1.1) Définition. Le morphisme d'algèbres $\theta_f : P \in \mathbf{k}[X] \mapsto P(f) \in \mathcal{L}(E)$, dit morphisme d'évaluation en f.

On peut commenter que la donnée de $f \in \mathcal{L}(E)$ produit une représentation de l'algèbre $\mathbf{k}[X]$, et que ce point de vue sur f s'avère très fructueux, c'est le fondement de la réduction classique des endomorphismes.

On note $\mathbf{k}[f]$ l'image du morphisme d'évaluation ; c'est une sous-algèbre *commutative* de $\mathcal{L}_{\mathbf{k}}(E)$. Deux polynômes en f commutent toujours.

(1.2) Exemples. L'inverse de f (s'il existe), l'exponentielle de f, sont des polynômes en f.

On verra plus tard des exemples de projecteurs intrinsèquement liés à f (à savoir, ceux sur ses sous-espaces caractéristiques) qui sont des polynômes en f.

(1.3) Pour tout polynôme P, $\ker P(f)$ est stable par f (ceci contient le cas des sous-espaces propres de f).

Pour P bien choisi, on pourra trouver un supplémentaire stable à $\ker\big(P(f)\big)$, mais ça n'existe pas pour n'importe quel P. *(Ne pas faire cette remarque, trop de risque de confusion).*

(1.4) Polynôme minimal. Définition comme unique générateur unitaire de l'idéal $\ker(\theta_f)$. On en déduit l'isomorphisme canonique

$$\mathbf{k}[u] \cong \mathbf{k}[X]/(\mu_u). \tag{1.4.1}$$

La dimension de cette $\mathbf{k}$-algèbre est $\deg(\mu)$.

Lien avec la formulation classique : $P \in \mathbf{k}[X]$ annule f si et seulement si $\mu | P$.

(1.4.2) Remarque. Le polynôme minimal est un invariant de similitude.

(1.4.3) Exercice.

$$\deg(\mu_u) = \min\big\{p : (\mathrm{id}, u, \dots, u^p) \text{ est liée}\big\}.$$

(1.4.4) Application. Invariance de μ par extension de corps, *via* l'invariance du rang. (La bonne preuve est que μ est un invariant de similitude, et que ceux-ci se calculent par un pivot de Gauss).

Soit $\mathbf{k} \to \mathbf{k}'$ une extension de corps. Certainement $\mu_{u^{\mathbf{k}'}}(u^{\mathbf{k}}) = 0$, donc $\mu_{u^{\mathbf{k}}} | \mu_{u^{\mathbf{k}'}}$. Ces deux polynômes ont le même degré par (1.4.3) et l'invariance du rang par extension de corps (voir (2.8.2) du Prologue), donc diffèrent d'une constante multiplicative. □

(1.5) Proposition. *Soit F sous-espace stable par $f \in \mathcal{L}(E)$. Alors μ_{f_F} et $\mu_{\bar{f}_F}$ divisent tous les deux μ_f (de manière équivalente,* $\mathrm{ppcm}(\mu_{f_F}, \mu_{\bar{f}_F})$ *divise μ_f).*

2 – Trigonalisation

(2.1) Définition.

(2.1.1) Remarque.
f trigonalisable $\Leftrightarrow$ $\exists$ une base $(e_1, \ldots, e_n)$ t.q. chaque $\mathrm{Vect}(e_1, \ldots, e_p)$ est stable par f
$\Leftrightarrow$ $\exists$ une filtration $\{0\} = F_0 \subsetneq F_1 \subsetneq \ldots \subsetneq F_n = E$ t.q. chaque F_p est stable par f.
(2.1.2) Exercice. Les deux notions de trigonalisabilité supérieure et inférieure sont elles équivalentes ?

(2.2) Théorème. *f trigonalisable $\Leftrightarrow$ χ_f scindé $\Leftrightarrow$ μ_f scindé.*

(2.2.1) Remarque. On obtient ainsi l'équivalence χ scindé $\Leftrightarrow$ μ scindé qui n'a rien d'évident.

(2.2.2) Remarque. En faisant un tout petit peu attention dans la preuve de f trigonalisable $\Rightarrow \mu_f$ scindé, on voit qu'on a démontré le théorème de Cayley–Hamilton pour les endomorphismes trigonalisables. On peut en déduire la version générale par un argument d'extension des scalaires.

(2.3) Lemme. Si $fg = gf$, alors les sous-espaces $\ker P(f)$ sont stables par g pour tout $P \in \mathbf{k}[X]$.
(2.3.1) Attention ! Même sous l'hypothèse de commutativité, il peut y avoir des stables par f qui ne sont pas stables par g. (Exemple : tout-le-monde est stable par id !)

(2.4) Proposition. Si f et g sont trigonalisables et $fg = gf$, alors ils sont simultanément trigonalisables.
(2.4.1) Remarque. La réciproque est fausse. Il suffit de se baisser pour ramasser un exemple.

(2.5) Si χ_f est scindé, on a la décomposition de Dunford qui nous dit que f se réduit à une partie diagonalisable et une partie nilpotente ; [5] en particulier f est diagonalisable si et seulement si la partie nilpotente est triviale.

Ceci devrait suffire à motiver l'étude des diagonalisables à la Section 3 et des nilpotents à la Section 4. On sera capable de donner un critère de similitude pour deux endomorphismes à polynôme caractéristique scindé (*i.e.*, , un critère de similitude pour les endomorphismes trigonalisables).

3 – Diagonalisation

(3.1) Lemme des noyaux. Pour $P_1, \ldots, P_r$ *deux à deux* premiers entre eux.
Les projecteurs associés sont des polynômes en f, et on sait les calculer.

(3.1.1) Corollaire. Si P et Q sont premiers entre eux, alors $P(f)_{\ker Q(f)}$ est injectif.

(3.1.2) NB. Si $P_1, \ldots, P_r$ sont deux à deux premiers entre eux (ce qui est une condition plus forte qu'être premiers entre eux dans leur ensemble ; en êtes-vous bien convaincus ?), alors les $Q_i := \prod_{j \neq i} P_j$ sont globalement premiers entre eux, [6] autrement dit $(Q_1, \ldots, Q_r) = (1)$ et on a l'identité qu'il nous faut

$$1 = U_1 Q_1 + \cdots + U_r Q_r.$$

Pour le calcul pratique des U_i par récurrence sur r avec l'algorithme d'Euclide, voir II.(2.10).

5. pour obtenir cette décomposition, il est souhaitable de bien comprendre aussi bien les diagonalisables que les nilpotents, on l'énoncera donc seulement quand ce sera le cas.

6. soit par l'absurde H facteur irréductible commun à tous les Q_i. H divise Q_1, donc il doit diviser un P_i, $i > 1$, disons P_2. H divise aussi Q_2, donc nécessairement aussi un P_j avec $j \neq 2$. H serait alors diviseur irréuctible commun à P_2 et P_j $\lightning$

(3.1.3) Le cas typique d'application du lemme des noyaux est quand $P_i = H_i^{\alpha_i}$ avec les H_i irréductibles deux à deux distincts.

Le lemme suivant sera utile à l'occasion.

(3.2) Lemme. *On suppose $E = \bigoplus F_i$ où les F_i sont stables par f. Si g laisse lui aussi les F_i stables (par exemple si $g = Q(f)$, ou si les F_i sont des* $\ker((P_i(f))$ *et g commute avec f), alors*

$$\ker(g) = \bigoplus(\ker(g) \cap F_i).$$

(3.3) Proposition. f diagonalisable $\Leftrightarrow$ μ_f scindé à racines simples.
(3.3.1) Corollaire. f diagonalisable $\Leftrightarrow \exists\, P$ annulateur de f scindé à racines simples.
(3.3.2) Mise en garde. χ_f scindé à racines simples $\Rightarrow$ f diagonalisable (et $\chi_f = \mu_f$), mais la réciproque est fausse en général.
f diagonalisable $\Rightarrow$ χ_f scindé, mais pas nécessairement à racines simples.

(3.4) (i) Si f diagonalisable, alors tout stable par f se décompose selon les sous-espaces propres de f.
(ii) Un diagonalisable est semi-simple.

(3.5) Diagonalisation simultanée.

4 – Étude des nilpotents

(4.1) Proposition. *i) La suite des* $K_i := \ker(f^i)$ *est d'abord strictement croissante, puis constante.*
ii) La suite $(\dim K_{i+1} - \dim K_i)_{i\geqslant 0}$ *est décroissante.*

Notation : $k_i = \dim(K_i)$, $k_i(\lambda) = \dim\big(K_i(f-\lambda.\mathrm{id})\big)$, etc.

Preuve. On a $f(K_{i+1}) \subseteq K_i$, donc en composant f avec la projection $K_i \to K_i/K_{i-1}$, on obtient une application linéaire ${}^\sharp\bar{f}_{i+1} : K_{i+1} \to K_i/K_{i-1}$. Son noyau est K_i :

$$ {}^\sharp\bar{f}_{i+1}(x) = 0 \Leftrightarrow f(x) \in K_{i-1} \Leftrightarrow f^i(x) = 0. $$

Ainsi, ${}^\sharp\bar{f}_{i+1}$ induit une application linéaire injective

$$ \bar{f}_{i+1} : K_{i+1}/K_i \to K_i/K_{i-1}, \tag{4.1.1} $$

et $d_{i+1} - d_i \leqslant d_i - d_{i-1}$. □

(4.1.2) Remarque. On a des résultats analogues avec la suite des images des itérés de f.

(4.2) Corollaire. Si f nilpotent, alors son indice de nilpotence est $\leqslant \dim E$.

(4.3) Théorème. *Soit* $f, g \in \mathcal{L}(E)$ *deux endomorphismes nilpotents. Les deux propositions suivantes sont équivalentes :*
(i) f *et* g *sont semblables ;*
(ii) pour tout i, $\dim\big(\ker(f^i)\big) = \dim\big(\ker(g^i)\big)$.

(4.3.1) Remarque. Pour que l'énoncé ci-dessus contienne le théorème de Jordan, il manque l'unicité de la forme normale à permutation des blocs près. Ceci est très bien expliqué dans les notes [Lyon I]. Une bonne façon de démontrer l'unicité est de montrer l'injectivité de l'application qui à une forme réduite de Jordan associe la suite des dimensions des noyaux de ses itérés, voir (4.4).

Avant de se lancer dans la preuve, il est bon de se souvenir qu'elle s'écrit de droite à gauche.

Preuve. $(i) \Rightarrow (ii)$ est immédiat. La stratégie pour prouver la réciproque et d'écrire tout endomorphisme nilpotent sous une forme normale ne dépendant que des dimensions des noyaux de ses itérés ; ainsi si f et g satisfont à la propriété (ii), ils ont une forme normale commune et sont donc semblables.

Soit donc f nilpotent d'indice de nilpotence p, et notons pour tout i, $d_i = \dim K_i = \dim\big(\ker(f^i)\big)$ et $s_i = d_i - d_{i-1}$. On commence par choisir une base $(\bar{e}_1, \ldots, \bar{e}_{s_p})$ de $K_p/K_{p-1} = E/K_{p-1}$. Puisque l'application linéaire

$$ \bar{f}_p : \bar{x} \in K_p/K_{p-1} \mapsto \overline{f(x)} \in K_{p-1}/K_{p-2} $$

de (4.1.1) est injective, $\big(\overline{f(e_1)}, \ldots, \overline{f(e_{s_p})}\big)$ est une famille libre de vecteurs de K_{p-1}/K_{p-2}. On la complète en une base

$$ \big(\overline{f(e_1)}, \ldots, \overline{f(e_{s_p})}, \bar{e}_{s_p+1}, \ldots, \bar{e}_{s_{p-1}}\big), $$

remarquant au passage que la notation est consistante puisque $s_p \leqslant s_{p-1}$.

On obtient ainsi par récurrence des vecteurs

$$ e_1, \ldots, e_{s_p}, e_{s_p+1}, \ldots\ldots, e_{s_2}, e_{s_2+1}, \ldots, e_{s_1} \in E $$

(à nouveau : $s_p \leqslant \cdots \leqslant s_2 \leqslant s_1$) tels que pour tout $i = 1, \ldots, p$, (i) les vecteurs $e_{s_{i+1}+1}, \ldots, e_{s_i}$ sont dans K_i, et (ii) les classes des vecteurs

$$f^{p-i}(e_1), \ldots, f^{p-i}(e_{s_p}), \ldots, f(e_{s_{i+2}+1}), \ldots, f(e_{s_{i+1}}), e_{s_{i+1}+1}, \ldots, e_{s_i}$$

constituent une base de K_i/K_{i-1}.

Le Lemme (4.6) du Prologue nous assure alors que la famille

$$\begin{array}{cccccc} e_1, \ldots, e_{s_p}, & & & & & \\ f(e_1), \ldots, f(e_{s_p}), & & e_{s_p+1}, \ldots, e_{s_{p-1}}, & & & \\ \vdots & & \vdots & & \ddots & \\ f^{p-1}(e_1), \ldots, f^{p-1}(e_{s_p}), & f^{p-2}(e_{s_p+1}), \ldots, f^{p-2}(e_{s_{p-1}}), & \ldots\ldots, & e_{s_{1+1}+1}, \ldots, e_{s_1} \end{array}$$

est une base de E. On obtient une forme normale de Jordan dans la base obtenue en renumérotant la base ci-dessus en lisant colonne par colonne de gauche à droite et de haut en bas. [7] Le nombre de blocs de Jordan de taille i est $s_i - s_{i+1}$ (notons que $s_i = 0$ pour $i > p$), entièrement déterminé par la suite $(d_i)_{i \geqslant 0}$. □

(4.4) Unicité de la forme normale de Jordan. (Examen 2019 ; inspiré par [Lyon I]).

1) Soit $n \geqslant 1$. On considère la matrice $A_n = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbf{Q})$ définie par

$$\forall i, j \in [\![1, n]\!] \quad a_{ij} = \begin{cases} j & \text{si } j \leqslant i \\ i & \text{si } j > i. \end{cases}$$

a) Montrer que A_3 est inversible et calculer son inverse.

b) Montrer que la matrice

$$P_n = \begin{pmatrix} 2 & -1 & & & \\ -1 & 1 & & & \\ -1 & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \\ -1 & 0 & & & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbf{Q})$$

est inversible, puis calculer $P_n \times A_n$.

c) Montrer que A_n est inversible pour tout $n \geqslant 1$.

2) On considère $F_1, \ldots, F_r$ sous-espaces vectoriels de E tels que $E = \bigoplus_{1 \leqslant i \leqslant r} F_i$.

a) Soit $g \in \mathcal{L}(E)$ tel que pour tout $i = 1, \ldots, r$, F_i est stable par g. Montrer que

$$\ker(g) = \bigoplus_{1 \leqslant i \leqslant r} \big(\ker(g) \cap F_i\big).$$

b) Soit $f \in \mathcal{L}(E)$ tel que pour tout $i = 1, \ldots, r$, F_i est stable par f. Montrer que pour tout $P \in \mathbf{k}[X]$,

$$\ker\big(P(f)\big) = \bigoplus_{1 \leqslant i \leqslant r} \big(\ker(P(f)) \cap F_i\big).$$

3) On considère l'endomorphisme f de $\mathbf{k}^r$ défini par la multiplication à gauche par la matrice

$$J_r = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}.$$

7. ceci fournit des blocs de Jordan sous forme de matrice compagnon, avec les '1' sous la diagonale ; pour les avoir au dessus, il faut lire les colonnes de bas en haut.

Calculer $\dim\big(\ker(f^i)\big)$ pour tout $i \in \mathbf{N}$.
4) Soit $f \in \mathcal{L}(E)$ et $b_1, \ldots, b_n$ ($n = \dim(E)$) des entiers. On suppose qu'il existe une décomposition

$$E = \bigoplus_{1 \leqslant r \leqslant n} \bigoplus_{1 \leqslant a \leqslant b_r} F_{r,a}$$

telle que chaque $F_{r,a}$ est de dimension r et stable par f, et $f_{F_{r,a}}$ est semblable à l'endomorphisme de $\mathbf{k}^r$ de la question 3.
a) En utilisant les questions 2b et 3, calculer $k_i := \dim\big(\ker(f^i)\big)$ pour tout $i \in \mathbf{N}$. En déduire que

$$\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = A_n \times \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

où A_n est la matrice de la question 1.
b) Conclure que s'il existe une autre décomposition

$$E = \bigoplus_{1 \leqslant r \leqslant n} \bigoplus_{1 \leqslant a \leqslant b'_r} F'_{r,a}$$

telle que chaque $F'_{r,a}$ est de dimension r et stable par f, et $f_{F'_{r,a}}$ est semblable à l'endomorphisme de $\mathbf{k}^r$ de la question 3, alors $b_i = b'_i$ pour tout $i = 1, \ldots, n$. □

(4.5) Curiosité. Quelles valeurs la suite des k_i peut-elle prendre ? Les seules contraintes sont celles imposées par les inégalités du (4.1), qui se réduisent à

$$0 \leqslant k_n - k_{n-1} \leqslant \ldots \leqslant k_2 - k_1 \leqslant k_1 - k_0 = k_1.$$

En effet, ces inégalités donnent une condition nécessaire et suffisante pour que

$$A^{-1} \times \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & 1 \\ & & & -1 & 1 \end{pmatrix} \times \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$$

ait toutes ses entrées $\geqslant 0$.

On conclut par le traditionnel énoncé de simultanée propriété-ité.

(4.6) Proposition. *Si n et n' sont nilpotents et commutent, alors $n + n'$ est nilpotent.*

5 – Endomorphismes à polynôme caractéristique scindé

En vertu du Théorème (2.2), cette section pourrait s'appeler *Endomorphismes trigonalisables, II.* Dans toute cette section, on suppose que χ_f est scindé, ou de manière équivalente μ_f est scindé.

(5.1) Lemme. *Les polynômes χ et μ ont les mêmes racines.*

Remarque. On démontre plus loin (Prop. (6.1)) qu'en fait, χ et μ ont toujours les mêmes facteurs irréductibles, même s'ils ne sont pas scindés.

Preuve. Soit λ racine de χ. C'est une valeur propre, donc il existe $x \neq 0$ tel que $f(x) = \lambda.x$. On a $\mu(f)(x) = \mu(\lambda).x = 0$, donc $\mu(\lambda) = 0$.

Soit λ racine de μ, et écrivons $\mu = (X - \lambda)Q$. Par minimalité de μ, il existe $x \in E$ tel que $Q(f)(x) \neq 0$. Alors nécessairement $f(x) = \lambda.x$ donc $f - \lambda.\mathrm{id}$ n'est pas inversible et $\chi(\lambda) = 0$. □

(5.2) Notation. Écrivons

$$\mu = \prod\nolimits_{\lambda\in \mathrm{Sp}}(X-\lambda)^{a_\lambda} \quad \text{et} \quad \chi = \prod\nolimits_{\lambda\in \mathrm{Sp}}(X-\lambda)^{b_\lambda}.$$

(5.3) Définition. Les *sous-espaces caractéristiques* de f sont les $C_\lambda = \ker\big((f-\lambda.\mathrm{id})^{a_\lambda}\big)$ pour $\lambda \in \mathrm{Sp}(f)$. On a la décomposition

$$E = \bigoplus\nolimits_{\lambda\in \mathrm{Sp}} C_\lambda \tag{5.3.1}$$

en somme de sous-espaces stables par f (et par tout g commutant avec f). Les *projecteurs spectraux* $p_\lambda \in \mathcal{L}(E)$, qui s'écrivent

$$p_{\lambda_i}(x_{\lambda_1}, \ldots, x_{\lambda_r}) = (0, \ldots, x_{\lambda_i}, \ldots, 0)$$

dans la décomposition (6.2.1), sont des polynômes en f.

(5.4) Proposition. *Soit $\lambda \in \mathrm{Sp}(f)$.*
i) L'entier a_λ est l'indice de stagnation de la suite des $K_i(\lambda) = \ker\big((f-\lambda.\mathrm{id})^i\big)$, et l'indice de nilpotence de $f_{C_\lambda} - \lambda.\mathrm{id}_{C_\lambda} \in \mathcal{L}(C_\lambda)$.
ii) L'entier b_λ est la dimension du sous-espace caractéristique C_λ.

On verra plus loin une version un peu plus générale de ce résultat (Proposition (6.3)). Les preuves sont strictement identiques, et en fait la formulation pour (6.3) me semble plus lisible car moins polluée par les notations.

Preuve. i) D'après le lemme des noyaux, $f-\lambda.\mathrm{id}$ est inversible sur $\bigoplus_{\lambda'\neq\lambda} C_{\lambda'}$, donc $K_i(\lambda) = C_\lambda \cap \ker\big((f-\lambda.\mathrm{id})^i\big)$ par le Lemme (3.2), et ce noyau s'identifie canoniquement à $\ker\big((f_{C_\lambda}-\lambda.\mathrm{id}_{C_\lambda})^i\big)$.

Par définition de C_λ, la suite $K_i(\lambda)$ est constante égale à C_λ lui-même pour $i \geqslant a_\lambda$. Par minimalité de μ, $K_i(\lambda) \subsetneq C_\lambda$ si $i < a_\lambda$.

ii) La décomposition $\bigoplus C_\lambda$ est une somme de sous-espaces stables, donc $\chi_f = \prod_\lambda \chi_{f_{C_\lambda}}$. Or $\chi_{f_{C_\lambda}} = (X-\lambda)^{\dim C_\lambda}$, puisque d'une part $\mu_{f_{C_\lambda}} = (X-\lambda)^{a_\lambda}$, donc $\chi_{f_{C_\lambda}}$ a comme seul facteur irréductible $X - \lambda$, et d'autre part $\deg(\chi_{f_{C_\lambda}}) = \dim(C_\lambda)$. □

(5.5) Théorème (Décomposition de Jordan–Chevalley, aussi dite de Dunford).

(5.5.1) Remarque. La condition de commutativité est aussi importante que le reste ; elle assure qu'on peut réduire d et n simultanément, donc utiliser cette décomposition pour comprendre f.

La diagonalisabilité et la nilpotence sont deux choix de vie incompatibles : un endomorphisme simultanémente diagonalisable et nilpotent est nécessairement nul.

(5.5.2) Corollaire. Il existe une base dans laquelle d a une matrice diagonale et n est sous forme réduite de Jordan. Une telle base est en particulier trigonalisante pour $f = d + n$.

Preuve. On note $p_\lambda \in \mathcal{L}(E)$ les projecteurs spectraux ; ce sont des polynômes en u.[8] On a

$$u = u \circ \sum\nolimits_\lambda p_\lambda = \underbrace{\sum\nolimits_\lambda \lambda.p_\lambda}_{=:d} + \underbrace{\sum\nolimits_\lambda (u - \lambda.\mathrm{id}) \circ p_\lambda}_{=:n}.$$

À ce point là, il vaut mieux affirmer froidement « [qu']on a toutes les propriétés cherchées » plutôt que de s'embarquer dans des explications lourdingues : on a déjà tout démontré dans les résultats précédents.

Reste quand même à voir l'unicité. Soit $d' + n'$ une autre décomposition. d' et n' commutent à u, donc à tout polynôme en u, et en particulier aux d et n construits ci-dessus. On en déduit que $d' - d$ est diagonalisable, et $n - n'$ nilpotent, et donc que $d = d'$ et $n = n'$ puisque

$$d' - d = n - n'.$$

□

(5.6) Exemple. Décomposition de

$$\begin{pmatrix} -3 & -2 & -2 \\ -2 & 0 & -1 \\ 10 & 5 & 6 \end{pmatrix}$$

($\chi = (X-1)^3$, voir (4.6)).

(5.7) Exemple. Décomposition de

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

Remarque : ceci donne de façon amusante un exemple de deux matrices triangulaires supérieures qui ne commutent pas.

(5.8) Théorème (Caractérisation des classes de similitude). *Soit f et g deux endomorphismes trigonalisables. Les deux conditions suivantes sont équivalentes :*
(i) f et g sont semblables ;
(ii) pour tout $\lambda \in \mathbf{k}$ et tout $s \in \mathbf{N}$,

$$\dim\big(\ker(f - \lambda.\mathrm{id})^s\big) = \dim\big(\ker(g - \lambda.\mathrm{id})^s\big).$$

Preuve. Il suffit de savoir le faire sous-espace caractéristique par sous-espace caractéristique. Sur un sous-espace caractéristique on peut appliquer le critère (4.3) pour les endomorphismes nilpotents, en se souvenant de l'identification

$$\ker(f - \lambda\mathrm{id})^i = C_\lambda \cap \ker(f - \lambda\mathrm{id})^i \cong \ker(f_{C_\lambda} - \lambda\mathrm{id}C_\lambda)^i.$$

□

8. et donc bien des éléments de $\mathcal{L}(E)$ et non de $\mathcal{L}(E, C_\lambda)$.

(5.9) Exercice. Soit $\lambda, \mu \in \mathbf{k}$ distincts. Donner un représentant de toutes les classes de similitude de matrices de $\mathcal{M}_5(\mathbf{k})$ dont le polynôme minimal est $(X-\lambda)^2(X-\mu)$. Bien justifier que les matrices que vous donnerez sont deux à deux non semblables.

Voir aussi Exercice (4.7) (partie II) pour des questions du même goût.

Avec la Proposition (5.4), le théorème de Cayley–Hamilton est à portée de main.

(5.10) Théorème (Théorème de Cayley–Hamilton). $\chi(f) = 0$.

Preuve. Quitte à étendre les scalaires, on peut supposer que χ_f est scindé sur $\mathbf{k}$. On dispose ainsi de la décomposition en sous-espaces caractéristiques comme précédemment. Pour chaque $\lambda \in \mathrm{Sp}(f)$, a_λ est l'indice de nilpotence de $f_{C_\lambda} - \lambda.\mathrm{id} \in \mathcal{L}(C_\lambda)$, donc $a_\lambda \leqslant \dim(C_\lambda)$. Puisque d'autre part $b_\lambda = \dim(C_\lambda)$, on a donc $a_\lambda \leqslant b_\lambda$ pour tout $\lambda \in \mathrm{Sp}(f)$, et ainsi $\mu_f | \chi_f$. □

6 – Réduction dans le cas général

(6.1) Proposition. *Les polynômes* μ *et* χ *ont les mêmes facteurs irréductibles.*

Rappel. $\mathbf{k}[X]$ est factoriel.

Preuve. L'argument clef est le lemme (5.1) combiné au fait que tout polynôme possède un corps de rupture.

Soit P facteur irréductible de χ. On considère $\mathbf{k}' = \mathbf{k}[X]/(P)$, c'est un corps étendant $\mathbf{k}$. P possède une racine dans $\mathbf{k}'$, qui est aussi racine de μ d'après le lemme (5.1).[9] On a donc $\operatorname{pgcd}(P, \mu) \neq 1$,[10] et ainsi $P|\mu$ car P irréductible.

Le même raisonnement prouve que tout facteur irréductible de μ divise χ. □

(6.2) Sous-espaces caractéristiques. On appelle $\mathrm{Irr} \subset \mathbf{k}[x]$ l'ensemble des facteurs irréductibles de P et/ou μ. On note

$$\mu = \prod\nolimits_{P\in\mathrm{Irr}} P^{a_P} \quad \text{et} \quad \chi = \prod\nolimits_{P\in\mathrm{Irr}} P^{b_P}.$$

Les *sous-espaces caractéristiques* de f sont les $C_i = \ker\big(P_i^{a_i}(f)\big)$ pour $i = 1, \dots, r$, et on a la décomposition

$$E = \bigoplus\nolimits_{P\in\mathrm{Irr}} C_P. \tag{6.2.1}$$

(6.3) Proposition. *Soit* P *facteur irréductible de* χ *et/ou* μ.
i) L'entier a_P *est l'indice de stagnation de la suite des* $K_i(P) = \ker\big(P(f)^i\big)$*, et l'indice de nilpotence de* $P(f)_{C_P} \in \mathcal{L}(C_P)$.
ii) L'entier b_P *est* $\dim(C_P)/\deg(P)$.

Preuve. i) L'endomorphisme $P(f)$ est injectif sur $\bigoplus_{Q\neq P} C_Q$, donc les $K_i(P)$ s'identifient canoniquement aux noyaux des $P(f_{C_P})^i$. La suite des $K_i(P)$ est constante pour $i \geqslant a_P$ par définition de C_P, et pas avant par minimalité de μ_f.

ii) On a $\chi_f = \prod_{P\in\mathrm{Irr}} \chi_{f_{C_P}}$. Chaque f_{C_P} est annulé par P^{a_P}, donc puisque P est irréductible son polynôme caractéristique est une puissance de P. Comme $\deg(\chi_{f_{C_P}}) = \dim(C_P)$, on a donc nécessairement $\chi_{f_{C_P}} = P^{\dim(C_P)/\deg(P)}$. □

6.1 – Théorème de Cayley–Hamilton

On donne ici une preuve du théorème de Cayley–Hamilton ne passant pas par un argument d'extension des scalaires pour se ramener au cas trigonalisable. L'argument donné nous amènera naturellement à l'étude des endomorphismes cycliques, qui elle-même nous mènera à la décomposition de Frobenius, puis à l'étude des invariants de similitude.

(6.4) Preuve du théorème de Cayley–Hamilton. On considère $x \in E$ quelconque, et on introduit F_x le plus petit sev stable par f contenant x, et μ_x le générateur unitaire de l'idéal $\{P \in \mathbf{k}[X] \mid P(f)(x) = 0\}$. On écrit

$$\mu_x = X^p + a_{p-1}X^{p-1} + \cdots + a_0.$$

9. Ici il faut faire attention aux petites subtilités, mais « tout va bien puisque χ et μ sont tous les deux invariants par extension de corps. »

10. Puisque $P, \mu \in \mathbf{k}[X]$, leur pgcd est le même calculé dans $\mathbf{k}[X]$ ou $\mathbf{k}'[X]$, l'algorithme d'Euclide ne faisant pas sortir de $\mathbf{k}[X]$.

(6.4.1) Lemme. La famille $(x, f(x), \dots, f^{p-1}(x))$ est une base de F_x.
Étapes de la preuve. $F_x = \mathrm{Vect}\big(f^k(x), k \in \mathbf{N}\big) = \mathrm{Vect}\big(x, f(x), \dots, f^{p-1}(x)\big)$. La famille $(x, f(x), \dots, f^{p-1}(x))$ est libre. □

Ensuite on écrit la matrice C_x de $f_{F_x} \in \mathcal{L}(F_x)$ dans cette base,

$$C_x = \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -a_{p-2} \\ & & 1 & -a_{p-1} \end{pmatrix}.$$

(6.4.2) Lemme. $\chi_{C_x} = \mu_x$.
Preuve. Par récurrence sur p (développer par rapport à la première ligne). On peut aussi faire la preuve en un seul coup, en développant par rapport à la dernière colonne, mais ça nécessite d'être sûr de soi pour écrire tous les p mineurs. □

(6.4.3) Conclusion. $\chi_{C_x} = \chi_{f_{F_x}} | \chi_f$ car F_x est stable, donc $\chi_f \in (\mu_x)$ et $\chi_f(f)(x) = 0$. □

(6.5) Exercice. Soit $P \in \mathbf{k}[X]$. On considère le $\mathbf{k}$-espace vectoriel $E_P = \mathbf{k}[X]/(P)$ (il se trouve que c'est aussi une $\mathbf{k}[X]$-algèbre), et l'endomorphisme $m_P \in \mathcal{L}(E_P)$ de multiplication par X :

$$\forall H \in \mathbf{k}[X], \quad m_P(\bar{H}) = \overline{X \cdot H}.$$

1) On note $n = \deg(P)$. Montrer que E_P est de dimension n, muni de la base canonique $(\bar{1}, \bar{X}, \dots, \bar{X}^{n-1})$.

2) Démontrer que $\mu_{m_P} = P$.

3) Écrire la matrice de m_P dans la base canonique. En conclure que le polynôme *minimal* de la matrice compagnon associée à P est P lui-même.

6.2 – Remarques sur les endomorphismes cycliques

[C'est hors-programme 2018, donc je le mentionne juste en passant].

(6.6) Définitions.
(i) F_x est le sous-espace cyclique associé à x pour f.
(ii) f est dit cyclique s'il existe x tel que $F_x = E$.
Notation. On note C_P la matrice compagnon du polynôme P.

(6.7) Exercice. Montrer que $\mu_{C_P} = P$. En déduire que pour tout $x \in E$, $\mu_x | \mu_f$.

(6.8) Lemme. Il existe toujours un $x \in E$ tel que $\mu_x = \mu_f$.

Preuve. On écrit $\mu_f = P_1^{a_1} \cdots P_r^{a_r}$ et on regarde la décomposition en sous-espaces caractéristiques

$$E = C_1 \oplus \cdots \oplus C_r.$$

Pour chaque i, il existe un $x_i \in C_i$ tel que $\mu_{x_i} = P_i^{a_i}$. En effet, pour tout $x \in C_i$, $\mu_x | P_i^{a_i}$ car $P_i^{a_i}(f)(x) = 0$, donc $\mu_x = P_i^{a_x}$. Si pour tout $x \in C_i$ on a $a_x < a_i$, alors $P_i^{a_i - 1}(u_{C_i}) = 0$ et ceci contredit la minimalité de μ. Donc notre x_i existe bel et bien.

Ensuite,

$$x = x_1 + \cdots + x_r$$

convient. □

(6.9) Corollaire. $\deg(\mu) \leqslant \dim(E)$.
(Cette inégalité est essentiellement équivalente au théorème de Cayley–Hamilton).

(6.10) Proposition. *f cyclique* $\Leftrightarrow \chi_f = \mu_f$.

Preuve. $\Rightarrow$) Il existe une base de E dans laquelle la matrice de f est une C_P. Or $\chi_{C_P} = \mu_{C_P}$ par l'exercice (6.7).
$\Leftarrow$) Je trouve x tel que $\mu_x = \mu_f$ par le Lemme (6.8) ; on a donc $\mu_x = \chi_f$. Du coup, $\dim F_x = \deg \mu_x = \deg \chi_f = \dim E$ et $E = F_x$. □

(6.11) Décomposition de Frobenius. *Tout endomorphisme se décompose de manière unique en somme directe d'endomorphismes cycliques $f_1, \ldots, f_r$ tels que $\chi_r | \cdots | \chi_1$.*

La théorie des invariants de similitude fournit ceci de manière définitive. On peut procéder à la main de la manière suivante, *cf.* [H2G2, vol. 1, III]. Je prends x tel que $\mu_x = \mu_f$. Son F_x possède un supplémentaire stable E' (c'est un résultat ; la preuve n'est pas méchante, mais je n'ai toujours pas compris la raison profonde). Je recommence sur $f_{E'}$, tenant compte du fait que $\mu_f(f_{E'}) = 0$, donc $\mu_{f_{E'}} | \mu_f$. Ça donne l'existence d'une décomposition sans trop de dégât. L'unicité est nettement plus délicate.

II – Réduction des endomorphismes : synthèse

(0.1) Remarques sur les modules. Définition d'un module. Ce qui change par rapport aux espaces vectoriels : il n'est pas toujours possible de résoudre les systèmes d'équations linéaires, en particulier ciao la théorie de la dimension. Dans la catégorie des espaces vectoriels, (i) E de dimension n possède des sev de toutes les dimensions $\leqslant n$, (ii) F sev possède toujours un supplémentaire, et (iii) toute suite exacte est scindée. Dans la catégorie des modules ces trois énoncés sont faux (on verra qu'ils le sont aussi dans la catégorie des groupes).

1 – Simplicité et semi-simplicité

(1.1) Le $\mathbf{k}[X]$-module $M_{E,f}$. Ses sous modules sont les M_{F,f_F}, F stable par f. Ceci dicte les définitions :

(i) f simple s'il n'a pas de sev stable non trivial ;
(ii) f semi-simple si tout stable possède un supplémentaire stable.

(1.2) Proposition. *f simple $\Leftrightarrow$ χ_f irréductible.*

(1.2.1) Exemples. Il n'y a que les homothéties d'une droite si $\mathbf{k}$ algébriquement clos. Rotations du plan euclidien.

(1.2.2) Preuve de la proposition. Supposons f simple, et montre que ceci impose à χ_f d'être irréductible. Soit $P, Q \in \mathbf{k}[x]$ tels que $\chi = PQ$. On a $P(f)Q(f) = 0$, donc $P(f)$ ou $Q(f)$ a un noyau non-nul. Disons que c'est P, et soit $x \in \ker(P(f))$ non nul. Par simplicité de f, F_x le plus petit stable contenant x est E tout entier, donc $\deg(\mu_x) = n$ (on utilise les notations de la preuve de Cayley–Hamilton). Ceci implique $\deg P \geqslant n$, et on conclut que χ est irréductible.

Si au contraire f possède un stable non-trivial F, alors χ_{f_F} est diviseur strict de χ_f, puisque $\deg(\chi_{f_F}) < \deg(\chi_f)$, donc χ n'est pas irréductible. □

(1.3) Proposition. *f semi-simple $\Leftrightarrow$ aucun carré non constant ne divise μ_f.*

(1.3.1) Preuve. '$\Rightarrow$'. Soit P facteur irréductible de μ, et considérons $C_P^\circ := \ker(P(f))$ (attention, en général ce n'est pas le sous-espace caractéristique associé à P ; justement ça l'est ssi P^2 ne divise pas μ). Si f est semi-simple, C_P° possède un supplémentaire stable F. L'endomorphisme $P(f)_F \in \mathcal{L}(F)$ est injectif donc inversible, en conséquence de quoi P et μ_F sont premiers entre eux. Puisque $P \cdot \mu_{f_F}$ annule f, on en déduit que P^2 ne divise pas μ_f.

'$\Leftarrow$'. Écrivons $\mu = P_1 \cdots P_r$, produit d'irréductibles deux à deux distincts, et considérons la décomposition en sous-espaces caractéristiques $E = C_1 \oplus \cdots \oplus C_r$ et les projecteurs spectraux $p_1, \ldots, p_r$. Soit F stable. Puisque les p_i sont des polynômes en f, chacun laisse F stable. On en déduit la décomposition

$$F = (F \cap C_1) \oplus \cdots \oplus (F \cap C_r). \tag{1.3.i}$$

On va montrer que chaque $F \cap C_i$ possède un supplémentaire G_i dans C_i stable par f, ou plutôt par f_{C_i}. Alors $G := \bigoplus G_i$ sera un supplémentaire de F dans E stable par f, et on aura gagné.

Pour montrer l'existence d'un supplémentaire stable à $F \cap C_i$ dans C_i, on utilise un argument un peu baroque. [11] Soit $\mathbf{k}_i := \mathbf{k}[X]/(P_i)$; puisque P_i est irréductible, $\mathbf{k}_i$ est un corps. L'opération de composition externe

$$(\bar{Q}, x) \in \mathbf{k}_i \times C_i \mapsto \bar{Q}.x := Q(f)(x)$$

munit le $\mathbf{k}$-ev C_i d'une structure de $\mathbf{k}_i$-ev. On observe que les $\mathbf{k}_i$-sev de C_i correspondent ensemblistement aux $\mathbf{k}$-sev de C_i qui sont stables par f. Ainsi $F \cap C_i$ est un $\mathbf{k}_i$-sev de C_i. On lui choisit un $\mathbf{k}_i$-supplémentaire : c'est un $\mathbf{k}$-sev de C_i supplémentaire à $F \cap C_i$ qui est stable par f, et nous avons donc trouvé notre G_i. □

(1.3.2) Exemples. Les endomorphismes qui ne sont typiquement pas semi-simples sont les nilpotents. Pour ceux-là, $X^2|\mu$ et le noyau est un stable sans supplémentaire stable (sauf si $f = 0$).

A contrario le prototype de l'endomorphisme semi-simple est l'endomorphisme diagonalisable. D'ailleurs, la preuve donnée ci-dessus du fait que si μ est sans facteur carré alors f est semi-simple est parallèle à celle du fait que les diagonalisables sont semi-simples ; la différence est que dans le cas diagonalisable, on a $\mathbf{k}_i \cong \mathbf{k}$ et donc l'argument baroque est invisible (mais bien là tapis dans l'ombre).

(1.3.3) Remarque. Il est toujours vrai, sans condition sur f, que tout sous-espace stable F se décompose selon les sous-espaces caractéristiques comme en (1.3.i).

En revanche, il est grossièrement faux que pour une décomposition arbitraire $E = \bigoplus H_i$ on a $F = \bigoplus(F \cap H_i)$.

(1.4) Corollaire. *semi-simple ⇔ somme directe de simples.*

Preuve. '⇒' s'obtient par abstract nonsense, mais '⇐' nécessite les caractérisations (1.1) et (1.2). C'est la décomposition selon les sous-espaces caractéristiques qui donne la décomposition en somme directe de simples. □

(1.5) Corollaire. *(i) Si* $\mathbf{k}$ *alg. clos, semi-simple ⇔ diagonalisable.*
(ii) Si $\mathbf{k}$ *parfait, alors* $M \in \mathcal{M}_n(\mathbf{k})$ *est semi-simple ssi elle est diagonalisable dans une extension finie de* $\mathbf{k}$*.*

(1.5.1) Rappel. Un corps $\mathbf{k}$ est parfait si tout polynôme irréductible de $\mathbf{k}[X]$ est scindé à racines simples dans une extension de décomposition. Les corps parfait sont les corps de caractéristique 0, et les corps de caractéristique $p > 0$ tels que $\mathbf{k}^p = \mathbf{k}$. Typiquement, s'il existe $\alpha \in \mathbf{k} - \mathbf{k}^p$, le polynôme $X^p - \alpha$ est irréductible dans $\mathbf{k}[X]$ mais possède une unique racine de multiplicité p dans tout corps de décomposition : si $\beta \in \mathbf{k}'$ est racine p-ème de $\alpha \in \mathbf{k}$, alors

$$X^p - \alpha = X^p - \beta^p = (X - \beta)^p.$$

Un exemple de corps non-parfait : $\mathbf{F}_p(T)$. Un exemple de semi-simple sur $\mathbf{F}_p(T)$ qui n'est diagonalisable dans aucune extension : *cf.* [H2G2, II].

11. On peut éviter cet argument et n'utiliser que des techniques élémentaires de réduction, voir [Gourdon]. Attention ce n'est pas si facile. Dans l'esprit de cette partie, la preuve naturelle est celle donnée ici.

2 – Structure de l'algèbre $\mathbf{k}[u]$

(2.1) L'algèbre $\mathbf{k}[u]$. On a déjà vu en (1.4) qu'elle est isomorphe à $\mathbf{k}[X]/(\mu_u)$, donc de dimension finie $\deg \mu_u$ sur $\mathbf{k}$.

Parmi ses « propriétés globales », outre sa dimension il faut citer son commutant dans $\mathcal{L}(E)$, *i.e.*, l'ensemble des endomorphismes de E qui commutent avec u. Ceci s'identifie naturellement à l'espace des endomorphismes de E comme $\mathbf{k}[X]$-module :

$$\mathrm{Com}_{\mathcal{L}(E)}(u) \cong \mathrm{End}_{\mathbf{k}[X]}\big(M_{E,u}\big)$$

(source : [H2G2, Vol. 2, II.2]).

2.1 – Liens entre réduction d'un endomorphisme u et structure de l'algèbre $\mathbf{k}[u]$

(2.2) La première chose à dire c'est que, notant $\mu_u = \prod P_i^{a_i}$ la décomposition en produit de facteurs irréductible, on a par le lemme chinois

$$\mathbf{k}[u] \cong \bigoplus_i \mathbf{k}[X]/\big(P_i^{a_i}\big), \tag{2.2.1}$$

qui présente un parallèle évident avec la décomposition en sous-espaces caractéristiques.

On peut d'ores et déjà prouver l'énoncé suivant.

(2.3) Semi-simplicité et nilpotents. *L'endomorphisme u est semi-simple ssi l'algèbre $\mathbf{k}[u]$ n'a pas d'élément nilpotent non-nul.*

Preuve. Le point clef est que semi-simple $\Leftrightarrow$ aucun carré non constant ne divise μ. Ainsi, vu la décomposition (2.2.1), il s'agit de se convaincre que $\mathbf{k}[X]/\big(P^a\big)$, P irréductible, possède des nilpotents non-triviaux ssi $a > 1$.

Si $a > 1$, $\bar{P}$ est un nilpotent non-nul. Si $a = 1$, soit $\bar{F}$ un nilpotent : il existe $s \geqslant 1$ tel que $\bar{F}^s = 0$, *i.e.*, $P|F^s$. Comme P irréductible, par le lemme de Gauss ceci implique $P|F$, *i.e.*, $\bar{F} = 0$. □

Pour la suite de notre étude, on aura besoin d'expliciter l'inverse de l'isomorphisme chinois. C'est l'objet des deux paragraphes suivants.

(2.4) Lemme chinois. *Soit $P, Q \in \mathbf{k}[X]$ premiers entre eux. L'application*

$$\phi : (H \bmod PQ) \in \frac{\mathbf{k}[X]}{(PQ)} \mapsto (H \bmod P, H \bmod Q) \in \frac{\mathbf{k}[X]}{(P)} \oplus \frac{\mathbf{k}[X]}{(Q)} \tag{2.4.1}$$

est un isomorphisme d'algèbres.

Preuve. On commence par constater que ϕ est bien définie, et que c'est un morphisme de $\mathbf{k}$-algèbres. Puisqu'on a l'égalité des dimensions

$$\dim\left(\frac{\mathbf{k}[X]}{(PQ)}\right) = \dim\left(\frac{\mathbf{k}[X]}{(P)} \oplus \frac{\mathbf{k}[X]}{(Q)}\right) = \deg(P) + \deg(Q),$$

il suffit de montrer l'injectivité de ϕ pour prouver que c'est un isomorphisme.

Soit donc $\bar{H}$ une classe modulo PQ tel que H est congru à 0 modulo P et modulo Q. Le polynôme H est divisible par P et par Q, donc par PQ puisque P et Q sont premiers entre eux, donc $\bar{H} = 0$, ce qui prouve l'injectivité de ϕ. □

(2.5) Réciproque de l'isomorphisme chinois. *Sous les hypthèses de (2.4), considérons la relation de Bezout*

$$UP + VQ = 1. \tag{2.5.1}$$

L'application

$$\psi : (H_1 \bmod P, H_2 \bmod Q) \mapsto H_1VQ + H_2UP \bmod PQ$$

est l'inverse de l'isomorphisme ϕ de (2.4.1).

Preuve. À nouveau, on commence par se convaincre que ψ est bien définie et est un morphisme de $\mathbf{k}$-algèbres. Ensuite, soit $\bar{H}_1, \bar{H}_2$ des classes modulo P et Q respectivement. On a

$$\begin{aligned} H_1VQ + H_2UP &\equiv H_1VQ \mod P \\ &\equiv H_1 \mod P \end{aligned}$$

puisque $VQ \equiv 1 \mod P$, et de la même manière $H_1VQ + H_2UP \equiv H_2 \mod Q$. Ceci prouve que $\phi \circ \psi = \mathrm{id}$. □

(2.6) Idempotents et projecteurs sur les sous-espaces caractéristiques. *Les idempotents de l'algèbre $\mathbf{k}[u]$ sont les (sommes de) projecteurs sur les sous-espaces caractéristiques.*

Preuve. On commence par constater que les projecteurs spectraux sont les éléments de $\mathbf{k}[u]$ qui dans la décomposition (2.2.1) s'écrivent $(0, \ldots, 1, \ldots, 0)$. Pour cela, on relit l'expression des projecteurs spectraux comme polynômes en u fournie par la preuve du lemme des noyaux (3.1), puis celle de l'inverse de l'isomorphisme chinois (2.5).

Ceci fait, il s'agit de montrer que les idempotents de $\mathbf{k}[X]/\left(P^a\right)$ sont 0 et 1. $\bar{F}$ est idempotent ssi

$$\begin{aligned} F^2 \equiv F \mod P^a &\Leftrightarrow P^a | F(F-1) \\ &\Leftrightarrow P^a | F \text{ ou } P^a | (F-1) \end{aligned}$$

car P irréductible et F et $F-1$ premiers entre eux. □

2.2 – Applications

(2.7) Remarque. Soit $u \in \mathcal{L}(E)$, $P \in \mathbf{k}[X]$. Les deux propositions suivantes sont équivalentes :
(i) $P(u)$ est inversible ;
(ii) P et μ_u sont premiers entre eux.

(2.8) Calcul de puissance. La remarque obligatoire à ce sujet est que pour calculer A^k, il suffit de calculer $P(A)$, où $P \in \mathbf{k}[X]$ est le reste de X^k modulo μ_A. Ou modulo n'importe quel polynôme annulateur qui serait plus facile à trouver que μ_A, d'ailleurs.

D'autre part, les techniques du paragraphe suivant s'appliquent aussi.

(2.9) Calcul d'exponentielle. On le fait sur un exemple simple et typique. Soit $A \in \mathcal{M}_n(\mathbf{k})$ telle que

$$(A - \mathrm{Id})(A - 2\mathrm{Id}) = 0,$$

et calculons e^A.

Soit p et q les projecteurs de $\mathbf{k}^n$ sur les deux sous-espaces caractéristiques éventuellement non triviaux de A (en l'occurence, ce sont des sous-espaces propres). On note bien que p et q s'expriment facilement et explicitement comme des polynômes en A. On a

$$\begin{aligned} e^A \circ p = e^{\mathrm{Id}} \circ e^{A-\mathrm{Id}} \circ p &= e. \sum \frac{1}{n!}(A-\mathrm{Id})^n \circ p \\ &= e.p \end{aligned}$$

puisque $(A-\mathrm{Id}) \circ p = 0$, et de la même manière $e^A \circ q = e^2.q$. Conclusion : puisque $p+q = \mathrm{id}$, on a

$$e^A = e^A \circ (p+q) = e.p + e^2.q.$$

Insistons sur le fait qu'on n'a jamais eu besoin de diagonaliser A où quoi que ce soit de ce type (même si au fond...).

Voici une autre façon de voir ce calcul. On a

$$A = p + 2q, \quad pq = qp = 0.$$

Ainsi p et q commutent, et on a

$$\begin{aligned} e^A = e^p e^{2q} &= \left(\sum_{n\geqslant 0} \frac{1}{n!} p^n\right) \left(\sum_{n\geqslant 0} \frac{1}{n!}(2q)^n\right) \\ &= \left(\mathrm{id} + \left(\sum_{n>0} \frac{1}{n!}\right) p\right) \circ \left(\mathrm{id} + \left(\sum_{n>0} \frac{2^n}{n!}\right) q\right) \qquad \text{puisque } p^2 = p \text{ et } q^2 = q \\ &= \big(\mathrm{id} + (e-1).p\big) \circ \big(\mathrm{id} + (e^2-1).q\big) \\ &= \mathrm{id} + (e-1).p + (e^2-1).q \\ &= e.p + e^2.q \qquad \text{puisque } \mathrm{id} = p+q. \end{aligned}$$

(2.10) Reprise du calcul précédent. Supposons avoir un polynôme annulateur P pour notre matrice M avec trois facteurs deux à deux premiers entre eux, $P = ABC$. Pour avoir les projecteurs correspondants, il nous faut une identité

$$1 = ABU + ACV + BCW$$

(on voit que A, B, C deux à deux premiers entre eux implique que AB, AC, BC sont globalement premiers entre eux, qui est la propriété dont on a vraiment besoin).

En pratique, on commence par écrire $1 = AU + BV$, qui donne

$$C = ACU + BCV.$$

Ensuite, puisque C et AB sont premiers entre eux, on peut trouver une écriture

$$1 = CS + ABW.$$

En remplaçant C, on arrive à

$$1 = AC(US) + BC(VS) + ABW$$

qui est l'écriture qu'on voulait.

(2.10.1) Remarque. En pratique, pour calculer $ACUS(M)$ on commencera par effectuer la division euclidienne de $ACUS$ par $P = ABC$ (ce dernier étant un polynôme annulateur), ce qui revient à diviser US par B.

3 – Classes d'équivalence de matrices à coefficients dans un anneau principal

(3.1) Inversibilité dans $\mathcal{M}_n(A)$.

(3.2) Théorème des facteurs invariants. *Soit A un anneau principal, $M \in \mathcal{M}_{m,n}(A)$ ($m \leqslant n$ disons). La matrice M est équivalente dans $\mathcal{M}_{m,n}(A)$ à une matrice*

$$\text{(3.2.i)} \qquad \begin{pmatrix} a_1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & a_m & 0 & \cdots & 0 \end{pmatrix}$$

où $a_1 | \cdots | a_m$. Les a_i sont uniquement déterminés (à multiplication par un inversible de A près).

Dans l'énoncé ci-dessus, les derniers a_i sont autorisés à être nuls. Les a_i sont appelés les *facteurs invariants* de la matrice M.

On appellera par abus de langage matrice *diagonale* une matrice comme en (3.2.i) (sans condition sur les facteurs diagonaux), et on la notera $\mathrm{diag}(a_1, \ldots, a_m)$ — ou éventuellement $\mathrm{diag}_{mn}(a_1, \ldots, a_m)$.

Deux mots sur la démonstration. L'existence d'une forme normale (3.2.i) s'obtient par un avatar du pivot de Gauss : c'est le pivot de Gauss authentique si A est euclidien, sinon il faut rajouter les opérations élémentaires de type Bezout (c'est inutile dans le cas euclidien, car celles-ci se décomposent en opérations élémentaires usuelles par un algorithme d'Euclide ; en pratique, on préférera néanmoins utiliser aussi des opérations de type Bezout).

Le pivot de Gauss ne donne aucunement l'unicité. Un moyen élémentaire et efficace de l'obtenir est de considérer les

$$\delta_k(M) = \mathrm{pgcd}\,\big(\text{mineurs } k \times k \text{ de } M\big).$$

Il faut se convaincre que la multiplication à gauche ou à droite par une matrice inversible maintient tous ces δ_k constants.

Les preuves qu'on trouve en général dans la littérature utilisent des techniques plus sophistiquées de théorie des modules, notamment la décomposition des modules de torsion selon les composantes primaires. Les lignes de preuves indiquées ci-dessus sont celles suivies dans [Debarre].

(3.2.1) Opération de type Bezout. Soit $a, b \in A$, $\mathrm{pgcd}(a, b) = \delta$, et considérons une relation de Bezout

$$au + bv = \delta.\,^{12}$$

On constate que la matrice

$$\begin{pmatrix} u & v \\ -b/\delta & a/\delta \end{pmatrix}$$

est dans $\mathrm{SL}_2(A)$ donc inversible, puis que

$$\begin{pmatrix} u & v \\ -b/\delta & a/\delta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \delta \\ 0 \end{pmatrix}.$$

12. notation a' réservée pour $a = a'\delta$

(3.2.2) Résolution pratique d'un système à coefficients dans un anneau. Plus haut je prétends qu'on arrive à la forme normale (3.2.i) par un pivot de Gauss. C'est vrai, mais il faut bien préciser qu'il faut autoriser les opérations élémentaires aussi bien sur les lignes (multiplication à gauche par $P \in \mathrm{GL}_m$) que sur les colonnes (multiplication à droite par $Q \in \mathrm{GL}_n$).

Si on se trouve concrètement en face d'un système d'équations linéaires, les opérations sur les lignes sont parfaitement banales, mais celles sur les colonnes reviennent à opérer des changements de variables sur les inconnues. En théorie ces changements de variables ne sont en aucun cas une obstruction à la résolution du système, mais en pratique leur accumulation s'avère rapidement douloureuse. Il existe d'autres méthodes de résolution plus habiles en pratique, mais je n'en parle pas ici.

(3.2.3) Anneau principal non-euclidien. Les exemples ne sont pas évidents à établir. Pour information, des exemples concrets sont les

$$\mathbf{Z}\big[\frac{1+\sqrt{-d}}{2}\big], \quad d = 19, 43, 67, 163.$$

(3.3) Structure des modules sur un anneau principal. Le théorème des facteurs invariants donne l'existence pour tout A-module M d'une décomposition

$$M = A^{\oplus r} \oplus A/(a_1) \oplus \cdots \oplus A/(a_{m-r}) \tag{3.3.1}$$

avec $\{0\} \subsetneq (a_{m-r}) \subseteq \cdots \subseteq (a_1)$, mais il ne suffit pas à garantir l'unicité de cette décomposition.

Pour obtenir l'unicité, il faut en plus garantir que deux présentations de M correspondent à des matrices équivalentes, ce qui n'a rien d'évident (même s'il est facile de se ramener à deux présentations avec les mêmes m et n, en rajoutant des générateurs et relations triviales de part ou d'autre).

À nouveau, on peut invoquer la décomposition des modules de torsion pour conclure à l'unicité, après avoir identifié r comme la dimension de l'espace vectoriel $M \otimes \mathrm{Frac}(A)$. Une approche élémentaire est possible, voir mes notes [groupes] d'après [duCloux].

3.1 – Algorithme de Gauss

Je donne ici une preuve de la partie existence du Théorème (3.2), basée sur l'algorithme de Gauss.

(3.4) Opérations élémentaires. En effet, les opérations élémentaires sur les lignes et les colonnes d'une matrice reviennent [...]

Dans un premier temps je me concentre sur le cas euclidien, dans lequel la preuve sera algorithmique — si la division euclidienne peut se faire de façon algorithmique.

À partir d'ici et jusqu'à mention explicite du contraire, l'anneau A est supposé euclidien.

(3.5) Opération de Bezout. Lorsque A est euclidien, il est commode de considérer des opérations sur les lignes et colonnes d'un type différent, qui ne font pas sortir de la classe d'équivalence.

Soit $a, b \in A$, $d = \mathrm{pgcd}(a, b)$, et considérons une relation de Bezout : soit $u, v \in A$ tels que

$$au + bv = d \iff a'u + b'v = 1$$

où l'on a écrit $a = a'd$ et $b = b'd$, avec $a', b' \in A$. La matrice

$$\begin{pmatrix} u & v \\ -b' & a' \end{pmatrix}$$

est inversible dans $\mathcal{M}_2(A)$ car de déterminant 1.

On en déduit qu'effectuer simultanément les deux opérations

$$\begin{cases} L_i \leftarrow uL_i + vL_j \\ L_j \leftarrow -b'L_i + a'L_j \end{cases} \quad \left(\text{resp.} \quad \begin{cases} C_i \leftarrow uC_i + vC_j \\ C_j \leftarrow -b'C_i + a'C_j \end{cases}\right)$$

(avec $i \neq j$) ne fait pas sortir de la classe d'équivalence : en effet cela revient à multiplier la matrice modifiée à gauche (resp. à droite) par une matrice inversible dans $\mathcal{M}_m(A)$ (resp. $\mathcal{M}_n(A)$) qu'on laisse au lecteur le soin d'écrire. Ces opérations transforment les lignes

$$\begin{matrix} L_i = \begin{pmatrix} a & * & \cdots & * \end{pmatrix} \\ L_j = \begin{pmatrix} b & * & \cdots & * \end{pmatrix} \end{matrix} \quad \text{en} \quad \begin{matrix} L_i' = \begin{pmatrix} d & *' & \cdots & *' \end{pmatrix} \\ L_j' = \begin{pmatrix} 0 & *' & \cdots & *' \end{pmatrix}. \end{matrix}$$

(3.6) Stathme euclidien. On note note

$$\varphi : A - \{0\} \to \mathbf{N}$$

le stathme euclidien, qu'on étend à A tout entier avec la convention $\varphi(0) = -1$. L'existence d'une division euclidienne assure que pour $a, b \in A - \{0\}$ il existe $q, r \in A$ tels que

$$a = bq + r \quad \text{et} \quad \varphi(r) < \varphi(a).$$

D'autre part, on peut supposer librement que

$$\forall a \in A - \{0\}, \quad \varphi(a) = 0 \Leftrightarrow a \text{ inversible}$$

(quitte à décaler la valeur de φ par une constante additive).

(3.6.1) Exemples. Les conventions ci-dessus peuvent paraître saugrenues dans le cas emblématique où $A = \mathbf{Z}$. Dans ce cas, une définition naturelle du stathme en accord avec ces conventions est

$$\forall a \in A, \quad \varphi(a) = |a| - 1.$$

Un autre cas important est celui où $A = \mathbf{k}[x]$ avec $\mathbf{k}$ un corps. Dans ce cas on peut poser simplement

$$\forall P \in \mathbf{k}[X], \quad \varphi(P) = \deg(P).$$

(3.7) Lemme. *Soit $m, n \in \mathbf{Z}_{\geqslant 2}$ et $M \in \mathcal{M}_{mn}(A)$. Il existe une matrice M' équivalente à M de la forme*

$$\begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{pmatrix}. \tag{3.7.1}$$

Preuve. Pour $M = (a_{ij})$, on pose

$$\varphi(M) := \max\Big(\min_{a_{ij} \neq 0} \varphi(a_{ij}), -1\Big).$$

On a donc $\varphi(M) = -1$ si et seulement si $M = 0$, et sinon seuls les coefficients non nuls de M entrent en compte dans le calcul de $\varphi(M)$.

Montrons alors le résultat par récurrence sur $\varphi(M) \in \mathbf{Z}_{\geqslant -1}$. Si $\varphi(M) = -1$ alors $M = 0$ et le résultat est vrai. Supposons donc $\varphi(M) > -1$. Dans ce cas $\varphi(M)$ est calculé par un coefficient non nul de M, et quitte à permuter les lignes et les colonnes on peut supposer que $\varphi(M) = a_{11}$.

Pour tout $i = 2, \ldots, m$ on écrit une division euclidienne $a_{i1} = a_{11}q_i + r_i$ de a_{i1} par a_{11}, et on effectue l'opération élémentaire $L_i \leftarrow L_i - q_i L_1$. De même pour tout $j = 2, \ldots, n$ on écrit une division euclidienne $a_{1j} = a_{11}p_i + s_i$ de a_{1i} par a_{11}, et on effectue l'opération élémentaire $C_j \leftarrow C_j - p_j C_1$. On obtient ainsi une matrice M' équivalente à M de la forme

$$M' = \begin{pmatrix} a_{11} & s_2 & \cdots & s_n \\ r_2 & & & \\ \vdots & & N & \\ r_m & & & \end{pmatrix}.$$

Si tous les r_i et s_j sont nuls, c'est une matrice M' de la forme voulue. Sinon on a

$$\varphi(M') \leqslant \min\Big(\min_{r_i \neq 0} \varphi(r_i), \min_{s_i \neq 0} \varphi(s_i)\Big) < \varphi(a_{11}) = \varphi(M)$$

et on conclut en appliquant l'hypothèse de récurrence à la matrice M'. □

(3.8) Lemme. *Soit $a, b \in A$ euclidien, $d = \mathrm{pgcd}(a, b)$ et $m = \mathrm{ppcm}(a, b)$. Les deux matrices*

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \textit{et} \quad \begin{pmatrix} d & 0 \\ 0 & m \end{pmatrix}$$

sont équivalentes dans $\mathcal{M}_2(A)$.

Preuve. On commence par effectuer l'opération $C_1 \leftarrow C_1 + C_2$, qui donne

$$\begin{pmatrix} a & 0 \\ b & b \end{pmatrix}.$$

Ensuite on effectue une opération de Bezout comme en (3.5) (dont on reprend les notations), on obtient

$$\begin{pmatrix} d & bv \\ 0 & a'b \end{pmatrix}.$$

Il reste alors à faire $C_2 \leftarrow C_2 - b'vC_1$ pour arriver à

$$\begin{pmatrix} d & 0 \\ 0 & a'b \end{pmatrix}.$$

qui est la matrice cherchée puisque $a'b = da'b' = \mathrm{ppcm}(a, b)$. □

(3.9) Preuve de l'existence des facteurs invariants dans le cas euclidien. Montrons par récurrence sur $m+n \in \mathbf{N}$ que toute matrice $M \in \mathcal{M}_{mn}(A)$ possède une matrice de la forme ((3.2)) dans sa classe d'équivalence. Quitte à raisonner sur la transposée, on peut supposer $m \leqslant n$. Si $m = 1$, on obtient directement le résultat en raisonnant comme dans la preuve du Lemme (3.7).

Supposons donc $m \geqslant 2$. Alors il existe une matrice diagonale par blocs comme en (3.7.1) équivalente à M. Par hypothèse de récurrence appliquée au bloc N, on obtient une matrice $\mathrm{diag}(a, a_2, \ldots, a_m)$ équivalente à M avec $a_2 | \cdots | a_m$.

On se ramène à une matrice du type voulu par application répétée du Lemme (3.8). On commence par appliquer le Lemme aux deux premières lignes : on obtient une matrice équivalente $\mathrm{diag}(d, \tilde{a}_2, a_3, \ldots, a_m)$ avec $d = \mathrm{pgcd}(a, a_2)$ et $\tilde{a}_2 = \mathrm{ppcm}(a, a_2)$. Le coefficient d divise tous les autres, puisqu'il divise a_2 donc aussi tous les autres a_i ; en particulier $\mathrm{pgcd}(d, a_i) = d$ pour tout i.

On répète l'opération pour les lignes numéros 1 et 3, ce qui donne la matrice équivalente $\mathrm{diag}(d, \tilde{a}_2, \tilde{a}_3, a_4 \ldots, a_m)$ puisque $\mathrm{pgcd}(d, a_3) = d$. D'autre part on observe que $\tilde{a}_3$ est multiple de d et a_3, donc *a fortiori* de d et a_2 et par suite de $\tilde{a}_2 = \mathrm{ppcm}(d, a_2)$. Répétant l'opération pour toutes les lignes restantes, on arrive à la matrice $\mathrm{diag}(d, \tilde{a}_2, \ldots, \tilde{a}_m)$, où $\tilde{a}_i = \mathrm{ppcm}(d, a_i)$ pour tout i. Cette dernière matrice est du type voulu puisque $d|\tilde{a}_2|\cdots|\tilde{a}_m$. □

4 – Classes de similitude dans $\mathcal{M}_n(\mathbf{k})$

(4.1) Définition des invariants de similitude. On prend une base, on fabrique $X \cdot \mathrm{Id}_n - M$, et on la met sous forme réduite.

(4.2) Théorème. *Deux matrices sont semblables ssi elles ont les mêmes invariants de similitude.*

Remarque. Je sais déjà décider si deux endomorphismes à polynôme caractéristique scindé sont semblables, d'abord en comparant leurs polynômes caractéristiques, puis les noyaux des itérés de $f - \lambda\mathrm{id}$ pour toute valeur propre λ. C'est algorithmique !

L'utilisation des invariants de similitude, outre le fait qu'elle est conceptuellement plus satisfaisante, est encore plus algorithmique, on fait seulement un pivot de Gauss ! En plus, elle ne nécessite pas de factoriser χ ou quelque autre polynôme que ce soit, opération désagréable et coûteuse.

(4.2.1) Les $\mathbf{k}[X]$-modules isomorphes à $M_{E,f}$ sont exactement les $M_{E',f'}$ pour lesquels qu'il existe $\varphi : E \cong E'$ tel que

$$f = \varphi^{-1} \circ f' \circ \varphi.$$

Preuve. Soit $\varphi : M_{E,f} \cong M$ isomorphisme de $\mathbf{k}[X]$-modules. Je définis $E' := M$ muni de la structure de $\mathbf{k}$-ev sous-jacente, et $f' := m_X \in \mathcal{L}(E')$. La relation $f = \varphi^{-1} \circ f' \circ \varphi$ est offerte par le fait que φ est un isomorphisme de $\mathbf{k}[X]$-modules, donc commute à la multiplication par $X \in \mathbf{k}[X]$. □

(4.2.2) Le $\mathbf{k}[X]$-module $M_{E,f}$ est isomorphe au quotient $E[X]/\mathrm{im}(X \cdot \mathrm{id} - f)$ via

$$\pi : \sum e_i X^i \mapsto \sum f^i(e_i).$$

Preuve. on vérifie que $\pi\big((X\mathrm{id} - f)(\sum e_i X^i)\big) = 0$, et réciproquement si $\sum e_i X^i \in \ker \pi$ alors

$$\begin{aligned} \sum e_i X^i &= \sum e_i X^i - 0 \\ &= \sum e_i X^i - \sum f^i(e_i) \\ &= \sum (X^i\mathrm{id} - f^i)(e_i), \end{aligned}$$

et chaque $X^i\mathrm{id} - f^i = (X\mathrm{id})^i - f^i$ se factorise par $X\mathrm{id} - f$. □

(4.2.3) Conclusion.

$$\frac{E[X]}{\operatorname{im}(X\cdot\mathrm{id}-f)} \;\cong\; \frac{\mathbf{k}[X]}{(P_1)}\oplus\cdots\oplus\frac{\mathbf{k}[X]}{(P_n)}.$$

□

(4.3) Décomposition de Frobenius. Chaque $E_i = \mathbf{k}[X]/(P_i)$ est un sev stable par f et $f_i := f_{E_i}$ est cyclique avec $\chi_{f_i} = \mu_{f_i} = P_i$. La décomposition

$$M_{E,f} = E_1 \oplus \cdots \oplus E_n$$

est une somme directe de sous-espaces cycliques pour f.

On en déduit $\chi_f = \chi_{f_1}\cdots\chi_{f_n} = P_1\cdots P_n$, et $\mu_f = P_n$. Ceci dévoile le Théorème de Cayley–Hamilton (je ne dirais pas que ça le trivialise...), y compris sa version améliorée "χ et μ ont les mêmes facteurs irréductibles".

(4.3.1) Exercice. f cyclique $\Leftrightarrow$ tous les P_i sauf le dernier sont égaux à 1.

[En particulier, on fait remarquer qu'en général il y a un certain nombre d'invariants égaux à 1, et que ceux-ci donnent des facteurs triviaux dans la décomposition cyclique.]

(4.4) Décomposition de Dunford–Jordan d'un bloc cyclique.

$$\frac{\mathbf{k}[X]}{\big((X-\lambda_1)^{a_1}\cdots(X-\lambda_r)^{a_r}\big)} \;\cong\; \frac{\mathbf{k}[X]}{(X-\lambda_1)^{a_1}}\oplus\cdots\oplus\frac{\mathbf{k}[X]}{(X-\lambda_r)^{a_r}}$$

par le lemme chinois ; pour chaque morceau de la décomposition de droite, la multiplication par X s'écrit

(4.4.1)
$$\begin{pmatrix}\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1\\ & & & \lambda\end{pmatrix}$$

dans la base $\big((X-\lambda)^{a-1},\dots,X-\lambda,1\big)$, puisque

$$X\cdot(X-\lambda)^i = (X-\lambda)^{i+1} + \lambda(X-\lambda)^i.$$

(4.4.1) est un bloc de Dunford–Jordan de taille a.

(4.5) Proposition. *Les invariants de similitude sont invariants par extension de corps.*

Preuve. Ils sont obtenus par un pivot de Gauss. □

(4.5.1) Exemple. Regarder une rotation d'angle $\pi/2$, non diagonalisable sur $\mathbf{R}$, mais diagonalisable après extension des scalaires. Écrire ses invariants de similitude sur $\mathbf{R}$ et sur $\mathbf{C}$ (!).

(4.6) Traiter l'exemple de

$$\begin{pmatrix}-3 & -2 & -2\\ -2 & 0 & -1\\ 10 & 5 & 6\end{pmatrix}.$$

(4.6.1) Remarque pratique. Pour s'épargner des opérations dans le pivot de Gauss, ne pas hésiter à utiliser directement le fait que pour P et Q quelconques

$$\mathbf{k}[X]/(P) \oplus \mathbf{k}[X]/(Q) \cong \mathbf{k}[X](\text{pgcd}) \oplus \mathbf{k}[X]/(\text{ppcm}),$$

en conséquence de quoi les matrices

$$\begin{pmatrix} P & \\ & Q \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} \text{pgcd} & \\ & \text{ppcm} \end{pmatrix}$$

sont équivalentes dans $\mathcal{M}(\mathbf{k}[X])$.

(4.7) Exercice.

1) Soit $\alpha \in \mathbf{C}$. Donner un représentant de chacune des classes de similitude dans $\mathcal{M}_4(\mathbf{C})$ constituées de matrices ayant α comme unique valeur propre.

2) Pour chaque classe $\overline{A}$ comme dans la question précédente, calculer :
a) les polynômes caractéristique et minimal χ_A et μ_A ;
b) pour tout $\beta \in \mathbf{C}$ et $i \in \mathbf{N}$, la dimension du noyau de $(\beta \mathrm{Id}_4 - A)^i$;
c) les invariants de similitude de A.

3) Soit $\alpha \in \mathbf{R}$. Déterminer deux matrices $A', A'' \in \mathcal{M}_4(\mathbf{R})$ qui ne sont pas semblables, mais qui sont telles que pour $A = A', A''$:

$$\text{(i)}\ \dim\big(\ker(\alpha \mathrm{Id}_4 - A)^i\big) = \begin{cases} i & \text{si } i \leqslant 1 \\ 2 & \text{sinon} \end{cases} ; \qquad \text{et (ii)}\ \dim\big(\ker(\beta \mathrm{Id}_4 - A)^i\big) = 0$$

pour tout $\beta \in \mathbf{R} - \{\alpha\}$ et tout $i \in \mathbf{N}$.

5 – Approche pédestre des invariants de similitude

Cette section contient une tentative de présentation des résultats de la section précédente sous forme plus digeste, ou en tout cas sans prononcer le mot 'module'.

(5.1) On considère E un $\mathbf{k}$-espace vectoriel. La donnée d'un endomorphisme $f \in \mathcal{L}(E)$ induit une représentation de $\mathbf{k}$-algèbres $\mathbf{k}[X] \circlearrowleft E$. Une telle représentation est simplement la donnée du morphisme de $\mathbf{k}$-algèbres

$$\theta_f : P \in \mathbf{k}[X] \mapsto P(f) \in \mathcal{L}(E).$$

(5.2) Représentations de $\mathbf{k}[X]$. Commençons par remarquer qu'une représentation $\mathbf{k}[X] \circlearrowleft E$ est entièrement déterminée par l'action de X. Ainsi, toute représentation de dimension finie de $\mathbf{k}[X]$ est du type décrit en (5.1).

(5.2.1) Représentation tautologique. C'est la représentation $\mathbf{k}[X] \circlearrowleft \mathbf{k}[X]$ donnée par la multiplication à gauche :

$$\forall P, Q \in \mathbf{k}[X] \quad P.Q = P \times Q.$$

L'image de X dans $\mathcal{L}(\mathbf{k}[x])$ est l'endomorphisme m_X de multiplication par X, qui lui même n'est autre que $X \cdot \mathrm{id}_{\mathbf{k}[X]}$:

$$\forall P \in \mathbf{k}[X] \quad m_X(P) = X \times P = X\mathrm{id}(P).$$

(5.2.2) Représentation cyclique. On peut construire des représentations de dimension finie à partir de la représentation tautologique de la manière suivante.

Soit $P \in \mathbf{k}[X]$ unitaire. La multiplication par X induit une action $\mathbf{k}[X] \circlearrowleft \mathbf{k}[X]/(P)$. Si on veut, c'est parce que le sev (ou plutôt l'idéal) $(P) = \mathbf{k}[X] \cdot P$ est stable par m_X. Plus simplement, c'est $\overline{X} \cdot \mathrm{id}$.

Sa matrice dans la base $(\overline{1}, \overline{X}, \ldots, \overline{X}^{d-1})$, $d = \deg P$, est la matrice compagnon de P. Ce point de vue sur les endomorphismes cycliques se prête particulièrement bien à la démonstration du fait que $\chi = \mu = P$ pour l'endomorphisme cyclique relatif au polynôme P.

Exemple : $P = (X - \lambda)^d$. Dans la base $(\overline{1}, \overline{X - \lambda}, \ldots, \overline{X - \lambda}^{d-1})$, la matrice de $X\mathrm{id}$ est un bloc de Jordan de taille d

$$\begin{pmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix},$$

puisque

$$X \cdot (X - \lambda^k) = (X - \lambda)(X - \lambda)^k + \lambda(X - \lambda)^k.$$

Exemple : $P = (X - \lambda)(X - \mu)$, $\lambda \neq \mu$. Dans la base $(\overline{X - \mu}, \overline{X - \lambda})$, la matrice de $X\mathrm{id}$ est

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

puisque

$$X \cdot \overline{X - \mu} = \overline{(X - \lambda)(X - \mu)} + \lambda \overline{X - \mu} \quad \text{et} \quad \circlearrowleft .$$

(5.3) Théorème. *Soit $\rho_f : \mathbf{k}[X] \circlearrowleft E$ une représentation de dimension finie, correspondant à $f \in \mathcal{L}(E)$. Il existe une unique représentation isomorphe à ρ_f de la forme*

$$\mathbf{k}[X] \circlearrowleft \big(\mathbf{k}[X]/(P_1) \oplus \cdots \mathbf{k}[X]/(P_r)\big)$$

où les P_i sont des polynômes unitaires tels que $P_1 | P_2 | \cdots | P_r$.

(5.4) Corollaire. *Les polynômes P_i caractérisent l'endomorphisme f à similitude près. On les appelle les invariants de similitude de f.*

La démonstration du corollaire est relativement immédiate. Avant de se lancer dans la preuve du théorème, on remarque tout d'abord que $\mu_f = P_r$ et $\chi_f = P_1 \cdots P_r$. On verra que le calcul des P_i se fait par un pivot de Gauss, donc en particulier celui de μ aussi. Dans le cas où P_r est scindé (et ainsi tous les P_i avec lui), la donnée des P_i détermine les dimensions des noyaux de tous les $(f - \lambda\mathrm{id})^k$.

Démonstration. En choisissant une base de E, on identifie f à un endomorphisme

$$\mathbf{c} \in \mathbf{k}^n \mapsto A \times \mathbf{c} \in \mathbf{k}^n,$$

où A est une matrice à coefficients dans $\mathbf{k}$. Ceci induit un endomorphisme de $\mathbf{k}[X]^n$, à savoir

$$\mathbf{P} \in \mathbf{k}[X]^n \mapsto A \times \mathbf{P} \in \mathbf{k}[X]^n.$$

On notera $A \in \mathcal{L}(\mathbf{k}[X]^n)$ cet endomorphisme.

D'autre part, la multiplication par X donne un autre endomorphisme de $\mathbf{k}[X]^n$, à savoir $X \cdot \mathrm{id} \in \mathcal{L}(\mathbf{k}[X]^n)$. On va trouver un quotient de $\mathbf{k}[X]^n$ isomorphe à $\mathbf{k}^n$ sur lequel les deux endomorphismes A et $X \cdot \mathrm{id}$ s'identifient naturellement : ce quotient est (évidemment !)

$$\mathbf{k}[X]^n / \operatorname{im}(X \cdot \mathrm{id} - A).$$

Nos deux endomorphismes $X \cdot \mathrm{id}$ et A de $\mathbf{k}[X]^n$ induisent des endomorphismes de ce quotient, parce que $\operatorname{im}(X \cdot \mathrm{id} - A)$ est stable par $X \cdot \mathrm{id}$ et A, puisque $X \cdot \mathrm{id}$ et A commutent avec $X \cdot \mathrm{id} - A$. Ce sont respectivement la multiplication par $\overline{X}$ coordonnée par coordonnée et la multiplication par la matrice A à coefficients dans $\mathbf{k}$, et ils s'identifient tautologiquement sur le quotient : pour tout $\mathbf{P} \in \mathbf{k}[X]^n$,

$$\begin{aligned} X \cdot \mathbf{P} - A \times \mathbf{P} \in \operatorname{im}(X \cdot \mathrm{id} - A) \quad &\Longleftrightarrow \quad \overline{X \cdot \mathbf{P}} - \overline{A \times \mathbf{P}} = 0 \\ &\Longleftrightarrow \quad \overline{X} \cdot \overline{\mathbf{P}} = A \times \overline{\mathbf{P}}. \end{aligned} \tag{5.4.1}$$

L'identification de la représentation $\mathbf{k}[X] \circlearrowleft \mathbf{k}[X]^n / \operatorname{im}(X \cdot \mathrm{id} - A)$ induite par $X \cdot \mathrm{id}$ avec une unique représentation comme dans l'énoncé fonctionne comme l'identitication de tout groupe abélien fini avec une somme de groupes cycliques, en utilisant le Théorème (3.2) pour l'existence. Son identification avec ρ_f se déduit du Lemme ci-dessous. □

(5.5) Lemme. *L'application linéaire*

$$\theta_A : \mathbf{P} = X^d \mathbf{a}_d + \cdot + X\mathbf{a}_1 + \mathbf{a}_0 \in \mathbf{k}[X]^n \longmapsto A^d \times \mathbf{a}_d + \cdot + A \times \mathbf{a}_1 + \mathbf{a}_0 \in \mathbf{k}^n$$

induit un isomorphisme $\Theta_A : \mathbf{k}[X]^n / \operatorname{im}(X \cdot \mathrm{id} - A) \cong \mathbf{k}^n$*, tel que le diagramme ci-dessous est commutatif.*

$$\begin{array}{ccc} \mathbf{k}[X]^n / \operatorname{im}(X \cdot \mathrm{id} - A) & \xrightarrow[\Theta_A]{\cong} & \mathbf{k}^n \\ \Big\downarrow{\scriptstyle \overline{X} \cdot \mathrm{id}} & & \Big\downarrow{\scriptstyle A} \\ \mathbf{k}[X]^n / \operatorname{im}(X \cdot \mathrm{id} - A) & \xrightarrow[\Theta_A]{\cong} & \mathbf{k}^n \end{array}$$

Preuve. Soit $\mathbf{P} = (P_1, \ldots, P_n) \in \mathbf{k}[X]^n$. En regardant les coefficients de chacun des P_i, on constate qu'il existe une écriture $\mathbf{P} = X^d \mathbf{a}_d + \cdot + X\mathbf{a}_1 + \mathbf{a}_0$, avec $\mathbf{a}_d, \ldots, \mathbf{a}_0 \in \mathbf{k}^n$ et $d = \max(\deg P_i)$. D'après (5.4.1), on a

$$\overline{\mathbf{P}} = \overline{X^d \mathbf{a}_d} + \cdot + \overline{X\mathbf{a}_1} + \overline{\mathbf{a}_0} = \overline{X^d \mathbf{a}_d} + \cdot + \overline{X\mathbf{a}_1} + \overline{\mathbf{a}_0} = A^d \overline{\mathbf{a}_d} + \cdot + A\overline{\mathbf{a}_1} + \overline{\mathbf{a}_0}.$$

Autrement dit, $\theta_A(\mathbf{P})$ est un vecteur de $\mathbf{k}^n \subset \mathbf{k}[X]^n$ qui représente la classe $\overline{\mathbf{P}} \in \mathbf{k}[X]^n / \operatorname{im}(X \cdot \mathrm{id} - A)$.

Montrons qu'il existe un unique vecteur $\mathbf{c} \in \mathbf{k}^n$ représentant $\overline{\mathbf{P}}$. Il suffit de voir que 0 est l'unique vecteur de $\mathbf{k}^n \cap \operatorname{im}(X \cdot \mathrm{id} - A)$. Or si $\mathbf{c} \in \mathbf{k}^n$ est dans $\operatorname{im}(X \cdot \mathrm{id} - A)$, il existe un $\mathbf{P} \in \mathbf{k}[X]^n$ tel que $\mathbf{c} = X\mathbf{P} + A\mathbf{P}$, et en regardant les coefficients dominants coordonnée par coordonnée, on voit que nécessairement $\mathbf{P} = 0$ et donc $\mathbf{c} = 0$ comme on voulait.

Ainsi, θ_A induit une application linéaire $\mathbf{k}[X]^n / \operatorname{im}(X \cdot \mathrm{id} - A) \to \mathbf{k}^n$, qui à toute classe $\overline{\mathbf{P}}$ associe son unique représentant $\mathbf{c} \in \mathbf{k}^n$. Cette application est évidemment injective : si deux classes $\overline{\mathbf{P}}$ et $\overline{\mathbf{P}}'$ ont un représentant en commun, on a $\overline{\mathbf{P}} = \overline{\mathbf{P}}'$. Elle est tout aussi évidemment surjective, puisque tout $\mathbf{c} \in \mathbf{k}^n$ est un représentant de $\overline{\mathbf{c}}$. La commutativité du diagramme est une conséquence de (5.4.1). □

Appendices

A – Formule de Laplace

(A.1) Proposition. *Soit $A = (a_{ij})_{1\leqslant i,j\leqslant n}$, matrice carrée à coefficients dans* **k** *anneau commutatif (oui !). On fixe un ensemble de colonnes $J = \{j_1 < \cdots < j_p\}$. On a*

$$\det(A) = \sum_{I=\{i_1<\cdots<i_p\}} (-1)^{|I|+|J|} \det(A_{IJ}) \det(A_{\bar{I}\bar{J}}),$$

où $|I| = i_1 + \cdots + i_p$.

La preuve par calcul sur le polynôme déterminant est essentiellement semblable à celle du développement par rapport à une seule colonne, sauf qu'il y a un lemme sur la signature qui cesse d'être facile dans notre situation.

(A.2) Notations. Les notations générales de théorie des groupes sont celles de [groupes].

Pour tout $p \in [\![1, n]\!]$, on note $\mathcal{P}_p$ l'ensemble des parties à p éléments de $\{1, \ldots, n\}$. Pour $I = \{i_1 < \cdots < i_p\} \in \mathcal{P}_p$, on note $\bar{I}$ le complémentaire de I dans $[\![1, n]\!]$, $\bar{\imath}_1 < \cdots < \bar{\imath}_{n-p}$ les $n-p$ entiers tels que

$$\{i_1, \ldots, i_p, \bar{\imath}_1, \ldots, \bar{\imath}_{n-p}\} = \{1, \ldots, n\}.$$

On pose $\omega_J^I = \big\{\sigma \in \mathfrak{S}_n : \sigma(J) = I\big\}$.

(A.3) Preuve par calcul sur le polynôme déterminant. Le point de départ est la partition $\mathfrak{S}_n = \coprod \omega_J^I$, qui donne

$$\text{(A.3.1)} \qquad \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = \sum_{I\in\mathcal{P}_p} \sum_{\sigma\in\omega_J^I} \varepsilon(\sigma) a_{\sigma(j_1)j_1} \cdots a_{\sigma(j_p)j_p} a_{\sigma(\bar{\jmath}_1)\bar{\jmath}_1} \cdots a_{\sigma(\bar{\jmath}_p)\bar{\jmath}_p}.$$

Ceci suggère de décomposer $\sigma \in \omega_J^I$ en $(\sigma', \sigma'') \in \mathfrak{S}_I \times \mathfrak{S}_{\bar{I}}$. Précisément, on définit

$$\Phi_J^I : \mathfrak{S}_p \times \mathfrak{S}_{n-p} \to \omega_J^I$$
$$\text{par} \quad \Phi_J^I(\sigma', \sigma'')(j_s) = i_{\sigma'(s)}, \quad \Phi_J^I(\sigma', \sigma'')(\bar{\jmath}_s) = \bar{\imath}_{\sigma''(s)}.$$

Ceci est une bonne définition, et manifestement Φ_J^I est injective. Il est à peu près aussi manifeste qu'elle est surjective, mais de toute façon, nous allons devoir exhiber une réciproque à Φ_J^I plus loin. Attention toutefois : Φ_J^I n'est pas un morphisme de groupes, d'ailleurs ω_J^I n'est même pas un groupe.

Au point où on en est, on a pour tout $I \in \mathcal{P}_p$

$$\text{(A.3.2)} \qquad \sum_{\sigma\in\omega_J^I} \varepsilon(\sigma) a_{\sigma(j_1)j_1} \cdots a_{\sigma(j_p)j_p} a_{\sigma(\bar{\jmath}_1)\bar{\jmath}_1} \cdots a_{\sigma(\bar{\jmath}_p)\bar{\jmath}_p}$$
$$= \sum_{\sigma'\in\mathfrak{S}_p} \sum_{\sigma''\in\mathfrak{S}_{n-p}} \varepsilon\big(\Phi_J^I(\sigma', \sigma'')\big) a_{i_{\sigma'(1)}j_1} \cdots a_{i_{\sigma'(p)}j_p} a_{\bar{\imath}_{\sigma''(1)}\bar{\jmath}_1} \cdots a_{\bar{\imath}_{\sigma''(n-p)}\bar{\jmath}_{n-p}}.$$

On calcule $\varepsilon\big(\Phi_J^I(\sigma',\sigma'')\big)$ dans le Lemme (A.4), et on en déduit que le second membre de (A.3.2) s'écrit

$$(-1)^{|I|+|J|}\left(\sum_{\sigma'\in\mathfrak{S}_p}\varepsilon(\sigma')a_{i_{\sigma'(1)}j_1}\cdots a_{i_{\sigma'(p)}j_p}\right)\left(\sum_{\sigma''\in\mathfrak{S}_{n-p}}\varepsilon(\sigma'')a_{\bar{i}_{\sigma''(1)}\bar{j}_1}\cdots a_{\bar{i}_{\sigma''(n-p)}\bar{j}_{n-p}}\right),$$

ce qui compte tenu de (A.3.1) achève la preuve. □

(A.4) Lemme. *On a* $\varepsilon\big(\Phi_J^I(\sigma',\sigma'')\big)=(-1)^{|I|+|J|}\varepsilon(\sigma')\varepsilon(\sigma'')$.

Preuve. A tout $(\sigma',\sigma'')\in\mathfrak{S}_p\times\mathfrak{S}_{n-p}$, on associe une permutation $\Psi(\sigma',\sigma'')\in\mathfrak{S}_n$ définie par

$$\Psi(\sigma',\sigma'')=\begin{pmatrix}1&\cdots&p&p+1&\cdots&p+(n-p)\\ \sigma'(1)&\cdots&\sigma'(p)&p+\sigma''(1)&\cdots&p+\sigma''(n-p)\end{pmatrix}.$$

Pour tout $I\in\mathcal{P}_p$, on définit une permutation $\rho_I\in\mathfrak{S}_n$ par

$$\rho_I=\begin{pmatrix}1&\cdots&p&p+1&\cdots&p+(n-p)\\ i_1&\cdots&i_p&\bar{\imath}_1&\cdots&\bar{\imath}_{n-p}\end{pmatrix}.$$

Je prétends que

$$\Phi(\sigma',\sigma'')=\rho_I\circ\Psi(\sigma',\sigma'')\circ\rho_J^{-1}, \tag{A.4.1}$$

et je laisse au lecteur le soin de le démontrer.

On va maintenant pouvoir montrer l'identité voulue sur les signatures en calculant des nombres d'inversions. D'une part, $\varepsilon(\Psi(\sigma',\sigma''))=\varepsilon(\sigma')\varepsilon(\sigma'')$ car

$$\mathcal{I}(\Psi(\sigma',\sigma''))=\big(\mathcal{I}(\sigma')\big)\cup\big((p,p)+\mathcal{I}(\sigma'')\big)$$

($\mathcal{I}$ désigne l'ensemble des inversions, dont il faut calculer le cardinal pour définir le nombre d'inversions). D'autre part,

$$\begin{aligned}\mathcal{I}(\rho_I)&=\big\{(r,s')\in[\![1,p]\!]\times[\![p+1,n]\!]:i_r>\bar{\imath}_{s'-p}\big\}\\ &\simeq\coprod_{1\leqslant r\leqslant p}\big\{s\in[\![1,n-p]\!]:\bar{\imath}_s<i_r\big\}\\ &\simeq\coprod_{1\leqslant r\leqslant p}\bar{I}\cap[\![1,i_r-1]\!]\\ &\simeq\coprod_{1\leqslant r\leqslant p}[\![1,i_r-1]\!]\setminus\{i_1,\dots,i_{r-1}\}\end{aligned}$$

donc

$$I(\rho_I)=\sum_{1\leqslant r\leqslant p}\big(i_r-1-(r-1)\big)=\sum_{1\leqslant r\leqslant p}i_r-\sum_{1\leqslant r\leqslant p}r=|I|-\frac{p(p+1)}{2}.$$

□

(A.5) Preuve par récurrence sur p. On laisse l'initialisation de la récurrence au lecteur. Supposons par récurrence la formule connue pour $p'\leqslant p-1$ colonnes. On commence par développer par rapport à la colonne j_1 :

$$\det(A)=\sum_{1\leqslant i\leqslant n}(-1)^{i+j_1}a_{i,j_1}\det(A_{\hat{\imath},\hat{\jmath}_1}).$$

Ensuite on développe chacun des déterminants de taille $n-1$ par rapports aux colonnes correspondant aux colonnes $j_2, \dots, j_p$ de A. Attention, il y a un décalage de numérotation, elles sont devenues les colonnes numéros $j_2-1, \dots, j_p-1$ de la matrice $A_{\hat{\imath},\hat{\jmath}_1}$. On a un décalage du même type dans la numérotation des lignes ; pour $I' \in [\![1,n]\!]^{p-1}$ ne contenant pas i, on note $s(i,I')$ le nombre de $i' \in I$ qui sont $> i$ (c'est le nombre de lignes dont l'indice va être décalé par la suppression de la ligne i). On obtient

$$\det(A_{\hat{\imath},\hat{\jmath}_1}) = \sum_{I' \not\ni i} (-1)^{|I'|-s(i,I')+|J'|-(p-1)} \det(A^{\hat{\imath},\hat{\jmath}_1}_{I'J'}) \det(A_{\bar{I}\bar{J}}),$$

où on a posé $J' = \{j_2 < \cdots j_p\}$, et $A^{\hat{\imath},\hat{\jmath}_1}_{I'J'}$ est la matrice de taille $p-1$ obtenue en extrayant de $A_{\hat{\imath},\hat{\jmath}_1}$ les lignes I' et les colonnes J' de A.

Reste à recombiner ces déterminants $p-1$ entre eux, à nouveau avec la formule de développement par rapport à une colonne, utilisée à l'envers. Le point sur lequel il faut se fixer est que chaque $I = \{i_1 < \cdots < i_p\}$ apparaît sous les p formes (i_r, I'_r) avec $I'_r = I - \{i_r\}$, $r = 1, \dots, p$; notons au passage que $s(i_r, I'_r) = p - r$.

$$\begin{aligned}\det(A) &= \sum_i \sum_{I' \not\ni i} (-1)^{i+j_1+|I'|-s(i,I')+|J'|-(p-1)} a_{i,j_1} \det(A^{\hat{\imath},\hat{\jmath}_1}_{I'J'}) \det(A_{\bar{I}\bar{J}}) \\ &= \sum_I (-1)^{|I|+|J|} \left(\sum_{r=1}^{p} (-1)^{-s(i_r,I'_r)-p+1} a_{i_1,j_1} \det(A^{\hat{\imath}_1,\hat{\jmath}_1}_{I'J'}) \right) \det(A_{\bar{I}\bar{J}}).\end{aligned}$$

Le terme entre parenthèses est le développement de $\det(A_{IJ})$ par rapport à sa première ligne, puisque $(-1)^{-s(i_r,I'_r)-p+1} = (-1)^{r+1}$ (voir ci-dessous). □

(A.5.1) Remarque. On a utilisé la formule diabolique

$$\forall \varepsilon_1, \dots, \varepsilon_r = \pm 1 : \quad (-1)^{\varepsilon_1 a_1 + \cdots + \varepsilon_r a_r} = (-1)^{a_1 + \cdots + a_r},$$

qui vaut en vertu de la non moins diabolique identité $(-1)^{-1} = -1$.

B – Exponentielle

Ici $\mathbf{k}$ doit être un sous-corps de $\mathbf{C}$.

(B.1) Décomposition de Dunford multiplicative. *Toute matrice $A \in \mathrm{GL}_n(\mathbf{k})$ à polynôme caractéristique scindé s'écrit de manière unique*

$$A = DU$$

où D est diagonalisable, U unipotente, et D et U commutent. De plus, D et U sont des polynômes en A.
(U unipotente $\underset{\text{déf.}}{\Leftrightarrow}$ $U - \mathrm{Id}$ nilpotente).

Il est bien écrit que ça vaut pour les matrices inversibles, c'est une décomposition sur $\mathrm{GL}_n(\mathbf{k})$, pas sur $\mathcal{M}_n(\mathbf{k})$, ce qui est parfaitement moral.

Cette décomposition est essentielle en pratique pour étudier les propriétés de surjectivité et/ou injectivité de l'exponentielle, voir (B.3) et suivants.

Preuve. « Identique à celle de la décomposition additive. » (Mais il faut l'avoir fait au moins une fois correctement dans sa vie, sans oublier l'unicité ni la polynomialité). □

(B.1.1) Passage d'une décomposition à l'autre. On a

$$A = DU = D + D(U - \mathrm{Id})$$

où D est diagonalisable, et $D(U - \mathrm{Id})$ nilpotente (on utilise que D et U commutent) ; la commutation est claire.

Réciproquement, si D est inversible (CNS pour que $D + N$ soit inversible), on a

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

où tout est comme il faut.

(B.2) Décomposition de Dunford réelle. Pour appliquer utilement la décomposition de Dunford à l'étude des propriétés d'injectivité et de surjectivité de l'exponentielle, il est nécessaire d'avoir une version valable sur $\mathbf{R}$. Une telle décomposition existe inconditionnellement sur tout corps parfait, en conséquence du Corollaire (1.5) qui dit que sur un corps parfait les semi-simples sont diagonalisables sur un corps de décomposition du polynôme caractéristique.

Soit $\mathbf{k}$ un corps parfait. Toute matrice $A \in \mathcal{M}_n(\mathbf{k})$ se décompose de manière unique en $A = D+N$ avec $D, N \in \mathcal{M}_n(\mathbf{k})$ respectivement semi-simple et nilpotente, telles que $DN = ND$. De plus, D et N sont des polynômes (à coefficients dans $\mathbf{k}$) en A.

La preuve est un peu triste : on passe à $\mathbf{k}'$ une extension de décomposition de χ_A, on applique la décomposition de Dunford habituelle à A qui est trigonalisable sur $\mathbf{k}'$, et on vérifie comme on peut que la décomposition obtenue vit en fait sur $\mathbf{k}$. Dans le cas réel, on constate que $D+N$ et $\overline{D} + \overline{N}$ sont deux décompositions de $A = \overline{A}$, et on conclut par unicité que $D = \overline{D}$ et $N = \overline{N}$.

En général on peut effectuer le même raisonnement avec un peu de théorie de Galois (on utilise là aussi que $\mathbf{k}$ est parfait). Il y a une jolie preuve évitant cet outil délicat dans [Debarre].

(B.3) Étant donné une matrice ($\mathcal{M}_n(\mathbf{k})$) décomposée additivement, on a de manière naturelle la décomposition multiplicative de son exponentielle (qui est inversible).

En effet :

$$e^{D+N} = e^D e^N$$

puisque D et N commutent, la commutation vaut puisque D et N sont toutes les deux des polynômes en $M = D + N$, e^D est diagonalisable, et e^N est unipotente.

(B.3.1) Recherche de pré-image par l'exponentielle. Soit $M \in \mathrm{GL}_n(\mathbf{k})$, décomposée multiplicativement en $M = DU$. Par unicité des décompositions de Dunford, pour toute $A \in \mathcal{M}_n(\mathbf{k})$ telle que $e^A = M$ décomposée en $S + N$, on a nécessairement

$$e^S = D \quad \text{et} \quad e^N = U.$$

Si $M \in \mathrm{GL}_n(\mathbf{R})$, on a une décomposition de Dunford DU où D est semi-simple et U unipotente, réelles toutes les deux. Si on cherche une pré-image A réelle elle aussi, les deux composantes de sa décomposition de Dunford seront réelles elles aussi, et on pourra suivre la stratégie habituelle.

(B.4) Proposition. *L'exponentielle réalise un homémomorphisme de l'espace des matrices nilpotentes dans celui des matrices unipotentes (aussi bien sur* $\mathbf{R}$ *que sur* $\mathbf{C}$*).*

Preuve. Voir [H2G2, VI.B.12 et B.11]. On peut utiliser le logarithme,

$$L(N) = N - \frac{1}{2}N^2 + \cdots \quad \text{pour } N \text{ nilpotente},$$

qui vérifie $e^{L(N)} = \mathrm{Id} + N$ (et ne fait pas sortir des matrices réelles, le cas échéant). □

(B.5) Proposition. *Soit $A \in \mathcal{M}_n(\mathbf{C})$. La matrice A est diagonalisable si et seulement si e^A l'est. (♲ pour A réelle en remplaçant diagonalisable par semi-simple ?)*

Preuve. Cf. [H2G2, VI.B.15]. On regarde les décompositions de Dunford : $A = D + N$,

$$e^A = e^D + e^D(e^N - \mathrm{Id}).$$

On voit alors

$$e^A \text{ diagonalisable} \Leftrightarrow e^N = \mathrm{Id}.$$

Cette dernière condition équivaut à $N = 0$ (c'est un petit résultat pas difficile sur les matrices nilpotentes, complément de la Proposition (B.4)). Le même argument fonctionne pour une matrice réelle, en utilisant la version adéquate de la décomposition de Dunford. □

(B.5.1) Exercice. Déterminer toutes les matrices X telles que $e^X = \mathrm{Id}$. (Ce sont toutes les matrices diagonales avec des coefficients dans $2i\pi\mathbf{Z}$).

(B.6) Théorème. *L'exponentielle réalise une surjection $\mathcal{M}_n(\mathbf{C}) \to \mathrm{GL}_n(\mathbf{C})$. L'exponentielle réelle n'est pas surjective, son image est l'ensemble des matrices qui s'écrivent B^2, $B \in \mathrm{GL}_n(\mathbf{R})$.*

En fait pour $A \in \mathrm{GL}_n(C)$, on peut carrément trouver un polynôme P (dépendant de A bien sûr) tel que $e^{P(A)} = A$ (*cf.* [H2G2, exo B.13]).

Une preuve qui fonctionne simplement exploite la décomposition de Jordan : pour chaque bloc, on trouve une pré-image en bricolant un tout petit peu à partir de la Proposition (B.4).

Le critère pour les matrices réelles n'est pas super praticable en pratique, comme on le verra avec les exemples. Jacques Sauloy me dit qu'il y a un critère en termes de décomposition de Jordan, dont je n'ai pas retrouvé la trace.

(B.7) Remarque. L'exponentielle réelle est injective (en dimension 1) mais surjective, l'exponentielle complexe est surjective mais pas injective.
L'exponentielle réelle reste-t-elle injective en dimension supérieure ? Évidemment non, en prenant θ et $\theta + 2k\pi$ dans (B.8).

(B.8) Exemple.

$$\exp\begin{pmatrix}0 & -\theta\\ \theta & 0\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

(faire explicitement le calcul de la série entière). C'est $e^{i\theta}$! On voit explicitement que toutes les matrices de rotations (réelles) sont dans l'image de l'exponentielle ; il n'est pas trop difficile de voir que ce sont des carrés.

En particulier, $-\mathrm{Id}_2$ est dans l'image de l'exponentielle, attention à ne pas parler trop vite !

(B.9) Un exemple plus simple que celui du jury. Avant de se lancer, deux remarques. Tout d'abord finalement il m'apparaît que le plus simple est d'utiliser le critère "A est un carré ou non", surtout si on veut éviter de parler de semi-simples. Ensuite, on note pour la recherche de racine carrée l'identité de matrices par blocs

$$\begin{pmatrix}0 & -B\\ B & 0\end{pmatrix}^2 = \begin{pmatrix}-B^2 & 0\\ 0 & B^2\end{pmatrix}, \tag{B.9.1}$$

qui généralise ce qu'on utilise dans (B.8), qui n'est autre que $(ib)^2 = -b^2$!

(B.9.2) On sait bien que -1 n'est pas dans $\exp(\mathbf{R})$, mais on a vu que $-\mathrm{Id}_2$ est $\exp\begin{pmatrix}0 & -\pi\\ \pi & 0\end{pmatrix}$.

(B.9.3) La matrice $\begin{pmatrix}-1 & 0\\ 0 & -2\end{pmatrix}$ n'est pas dans l'image de l'exponentielle réelle. En effet, un antécédent complexe a nécessairement deux valeurs propres non-réelles et non-conjuguée, donc ne peut pas être réel.

En rechanche, pour la matrice

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

cette obstruction disparaît, et en effet on voit facilement que cette matrice est un carré avec (B.9.1). On peut tout aussi facilement exhiber une pré-image réelle par l'exponentielle.

(B.10) Exemple du rapport 2017.

(B.10.1) La matrice

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

n'est pas dans l'image de l'exponentielle.

a) L'argument rapide. Si $A = B^2$, B réelle, alors B a ses valeurs propres parmi i et $-i$, et celles-ci doivent être deux à deux conjuguées. Donc B a i et $-i$ comme valeurs propres simples, et est diagonalisable sur $\mathbf{C}$. Alors A est tout aussi diagonalisable sur $\mathbf{C}$, ce qui n'est manifestement pas le cas. $\lightning$

b) Argument élémentaire bourrin. En 2×2, il est possible de voir à la main en écrivant plus ou moins habilement les équations sur les coefficients que A n'est pas le carré d'une matrice réelle.

c) Stratégie reproductible. On décompose selon Dunford

$$A = \begin{pmatrix}-1 & 1\\ 0 & -1\end{pmatrix} = -\mathrm{Id}\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix},$$

et dans cette décomposition

$$U := \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \exp(N), \quad N = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}.$$

On cherche donc D_B réelle semi-simple (*i.e.*, réelle et diagonalisable sur $\mathbf{C}$), qui commute à N, et telle que $\exp(D_B) = -\mathrm{Id}$. On calcule que la condition de commuter à N impose

$$D_B = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}.$$

On aboutit rapidement à une contradiction.
(B.10.2) La matrice 4×4 par blocs

$$A' = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}$$

est exponentielle d'une matrice réelle.

On écrit la décomposition de Dunford $A' = -\mathrm{Id}U'$, et comme précédemment $U' = \exp(N')$. Reste à trouver D' diagonalisable réelle ($-\mathrm{Id}$ l'est !), commutant à N', et telle que $\exp(D') = -\mathrm{Id}_4$. À force de bricolage, on trouve que

$$D' = \pi \begin{pmatrix} 0 & \mathrm{Id}_2 \\ -\mathrm{Id}_2 & 0 \end{pmatrix}$$

convient.

Comme application, on peut trouver une racine carrée réelle B' à A' :

$$B' = \begin{pmatrix} 0 & V \\ -V & 0 \end{pmatrix}, \quad V = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{pmatrix}$$

fonctionne.

Finalement je prétends qu'on peut trouver relativement naturellement cette racine carrée, et que c'est le moyen de le plus simple de se convaincre que notre matrice est dans l'image de l'exponentielle réelle. On cherche une racine carrée de la forme (B.9.1), ce qui revient à chercher B telle que $B^2 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. On trouve $B = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{pmatrix}$ parce qu'on connaît la formule

$$\begin{pmatrix} \mathrm{Id}_n & \mathbf{e} \\ & 1 \end{pmatrix} \times \begin{pmatrix} \mathrm{Id}_n & \mathbf{e}' \\ & 1 \end{pmatrix} = \begin{pmatrix} \mathrm{Id}_n & \mathbf{e}+\mathbf{e}' \\ & 1 \end{pmatrix}$$

qui permet de voir le groupe des translations dans les projectivités.

(B.11) Lien avec la décomposition polaire. La décomposition polaire dit qu'on a des homéomorphismes (et nul doute que ce sont en fait des C^∞-difféomorphismes)

$$\mathrm{GL}_n(\mathbf{R}) \cong \mathrm{O}_n(\mathbf{R}) \times \mathcal{S}_n^{++}(\mathbf{R}) \cong \mathrm{O}_n(\mathbf{R}) \times \mathrm{GL}_n(\mathbf{R})/\mathrm{O}_n(\mathbf{R})$$
$$\mathrm{GL}_n(\mathbf{C}) \cong \mathrm{U}_n(\mathbf{C}) \times \mathcal{H}_n^{++}(\mathbf{C}) \cong \mathrm{U}_n(\mathbf{C}) \times \mathrm{GL}_n(\mathbf{C})/\mathrm{U}_n(\mathbf{C}).$$

(Les identifications $\mathcal{S}_n^{++}(\mathbf{R}) \cong \mathrm{GL}_n(\mathbf{R})/\mathrm{O}_n(\mathbf{R})$ et $\circlearrowleft$ sont données par le fait que $\mathcal{S}_n^{++}(\mathbf{R})$ est l'orbite de Id_n sous l'action de GL_n sur lui-même par congruence). Ces identités ne sont évidemment pas des isomorphismes de groupes, dans la mesure où $\mathcal{S}_n^{++}(\mathbf{R})$ et $\circlearrowleft$ ne sont même pas des groupes !

L'exponentielle apporte des homémorphismes

$$\mathcal{S}_n(\mathbf{R}) \to \mathcal{S}_n^{++}(\mathbf{R})$$
$$\mathcal{H}_n(\mathbf{C}) \to \mathcal{H}_n^{++}(\mathbf{C}).$$

En combinant avec la décomposition polaire, on obtient

$$\mathrm{GL}_n(\mathbf{R}) \cong \mathrm{O}_n(\mathbf{R}) \times \mathbf{R}^{\frac{n(n+1)}{2}} \quad \text{et} \quad \mathrm{GL}_n(\mathbf{C}) \cong \mathrm{U}_n(\mathbf{C}) \times \mathbf{R}^{n^2}.$$

(B.11.1) Mises en garde. Comme dit ci-dessus, $\mathcal{S}_n^{++}(\mathbf{R})$ n'est pas un groupe. On pourrait trop hâtivement l'envisager, partant du fait que c'est l'image par l'exponentielle d'un espace vectoriel (ce serait même un groupe abélien !). Oui mais $e^{A+B} = e^A e^B$ ne vaut que si A et B commutent, et il n'y a aucune raison pour que deux matrices, fussent-elles symétriques, commutent. Exemple : $\begin{pmatrix} 1 & \\ & \lambda \end{pmatrix}$ et $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ ne commutent pas en général. Pire, leur produit n'est même pas symétrique en général, donc même la symétrie n'est pas stable par produit.

Dès lors il n'y aucune raison de croire que si $M = OS$ et $M' = O'S'$, alors la décomposition polaire de MM' est $OO'SS'$.

Une autre mauvaise raison d'imaginer que $\mathcal{S}_n^{++}(\mathbf{R})$ est un groupe consiste à invoquer le théorème spectral. Il ne permet pas de dire que deux matrices symétriques réelles définies positives sont simultanément diagonalisables.

(B.11.2) Une autre identification : $\exp : \mathcal{A}_n(\mathbf{R}) \twoheadrightarrow \mathrm{SO}_n(\mathbf{R})$.

La preuve est basée sur la réduction des endomorphismes orthogonaux (ils sont semi-simples, avec des stables de dimension $\leqslant 2$), cf. [FGN, Algèbre 3, 1.35], et sur le petit calcul (B.8) pour chacun des blocs, cf. [FGN, *ibid.*, 1.36].

C – Exercices d'hiver

(C.1) Caractérisation des matrices diagonales complexes par la topologie de leur classe de similitude. (Examen 2014 ; pour une référence voir [FGN, II, 4.16]).

1) Rappeler comment on peut munir $\mathcal{M}_n(\mathbf{C})$ d'une topologie naturelle.

2) Pour tout $M \in \mathcal{M}_n(\mathbf{C})$, montrer que l'adhérence de la classe de similitude de M contient une matrice diagonale.

3) On considère les applications

$$\begin{array}{rcl} \chi : \mathcal{M}_n(\mathbf{C}) & \to & \mathbf{C}[X]_{\leq n} \\ A & \mapsto & \chi_A \end{array} \qquad \begin{array}{rcl} \mu : \mathcal{M}_n(\mathbf{C}) & \to & \mathbf{C}[X]_{\leq n} \\ A & \mapsto & \mu_A, \end{array}$$

où $\mathbf{C}[X]_{\leq n}$ est l'espace des polynômes de degré au plus n.
a) L'application χ est-elle continue (justifier par un court argument) ?
b) L'application μ est-elle continue (idem) ?

4) Soit D une matrice diagonalisable. Montrer que la classe de similitude de D est égale à $\chi^{-1}(\chi_D) \cap \mu_D^{-1}(0)$.

5) En déduire une caractérisation des matrices diagonalisables en terme de la topologie de leur classe de similitude.

(C.1.1) Complément. Exhiber une matrice réelle non-diagonalisable dont la classe de similitude est fermée.

(C.2) Matrices de rang 1. (Examen 2013).

1) Soit n et m deux entiers $\geqslant 1$. On considère $V \in \mathbf{k}^n$ et $L \in {}^{\mathsf{T}}(\mathbf{k}^m) := \mathcal{M}_{1,m}(\mathbf{k})$. Préciser la taille de la matrice $V \times L$, et montrer qu'elle est de rang 1.

2) Soit E et F deux $\mathbf{k}$–espaces vectoriels, et considérons deux applications linéaires $u_1, u_2 \in \mathcal{L}(E, F)$, toutes les deux de rang 1. Montrer que si

$$\ker u_1 = \ker u_2 \quad \text{et} \quad \operatorname{im} u_1 = \operatorname{im} u_2,$$

alors il existe deux homothéties $h_E \in \mathcal{L}(E)$ et $h_F \in \mathcal{L}(F)$ telles que

$$u_1 = u_2 \circ h_E \quad \text{et} \quad u_1 = h_F \circ u_2.$$

3) Identifier le noyau et l'image de $V \times L$.

4) Montrer que toute matrice de rang 1 est le produit d'une matrice colonne avec une matrice ligne.

5) Soit $u \in \mathcal{L}(E, F)$, et r un entier. Montrer l'équivalence des deux propriétés suivantes :
(i) $\operatorname{rg} u \leqslant r$;
(ii) $\exists u_1, \ldots, u_r \in \mathcal{L}(E, F)$ toutes de rang 1, telles que $u = u_1 + \cdots + u_r$.

6) Soit $A \in \mathcal{M}_n(\mathbf{k})$. Écrire explicitement

$$A = V_1 \times L_1 + \cdots + V_n \times L_n$$

où les V_i (resp. L_i) sont des matrices colonnes (resp. lignes), $1 \leqslant i \leqslant n$.

(C.2.1) Complément. Démontrer que $\{\operatorname{rg} \leqslant r\}$ est l'adhérence de $\{\operatorname{rg} = r\}$ (sur $\mathbf{k} = \mathbf{R}$ ou $\mathbf{C}$, ou mieux en topologie de Zariski).

Références

[CEZ] D. Couty, J. Esterle, R. Zarouf, *Décomposition effective de Jordan-Chevalley*, Gazette des Mathématiciens **129** (2011), 29–49

[Debarre] Olivier Debarre, *Réduction des endomorphismes*, lien.

[duCloux] Fokko du Cloux, *Groupes abéliens*, chapitre du cours d'algèbre de maîtrise, lien.

[FGN] S. Francinou, H. Gianella, S. Nicolas, Exercices de mathématiques — Oraux X-ENS, Cassini, 2008.

[Gourdon] X. Gourdon, *Les maths en tête — Algèbre*, Ellipses, 2009.

[Grifone] Joseph Grifone, *Algèbre linéaire*, Cépaduès Éditions, 4^{e} édition, 2011.

[groupes] Th. Dedieu, Notes sur les groupes

[H2G2] Philippe Caldero et Jérôme Germoni, *(Nouvelles) Histoires hédonistes de groupes et de géométries*, Calvage & Mounet, (2017) 2015.

[Laszlo] Yves Laszlo, *Introduction à l'algèbre commutative et homologique*, Cours de Maîtrise 2003–2004.

[Lyon I] Prépa Agreg de l'UCB—Lyon 1, *Endomorphismes trigonalisables et nilpotents*, 2009–2010. trigo_nilpotent.pdf.

[repr.] Th. Dedieu, Notes *Représentations linéaires (complexes de dimension finie des groupes finis)*

[Sernesi] Edoardo Sernesi, *Geometria* 1 e 2, Bollati Boringhieri editore, 1989.