ALGEBRE LINEAIRE - DIAGONALISATION

Plan

I : Somme de sous-espaces vectoriels

1- Produit d'espaces vectoriels

Soient F_1, ..., F_p des espaces vectoriels. On note $\prod_{i=1}^{p} F_i$ l'ensemble $\{(x_1, ..., x_p), \forall\, i, x_i \in F_i\}$. Il est facile de vérifier qu'il s'agit d'un espace vectoriel lorsqu'on le munit des opérations :

$(x_1, ..., x_p) + (y_1, ..., y_p) = (x_1 + y_1, ..., x_p + y_p)$

$\lambda(x_1, ..., x_p) = (\lambda x_1, ..., \lambda x_p)$

PROPOSITIONS :

Si les F_i *sont de dimension finie, il en est de même de* $\prod_{i=1}^{p} F_i$ *et* $\dim(\prod_{i=1}^{p} F_i) = \sum_{i=1}^{p} \dim(F_i)$.

Démonstration :

Si, pour tout i, on note $(e_{ij}), 1 \leq j \leq \dim(F_i)$, une base de F_i, montrons qu'une base de $\prod_{i=1}^{p} F_i$ est donnée par l'ensemble des $(0, 0, ..., \underset{\substack{\uparrow \\ i \text{ ème rang}}}{e_{ij}}, 0, ..., 0)$. Cette famille comporte en effet $\sum_{i=1}^{p} \dim(F_i)$ vecteurs.

❑ La famille des $(0, 0, ..., e_{ij}, 0, ..., 0)$ est libre. Si $\sum_{i=1}^{p} \sum_{j=1}^{\dim(Ei)} \lambda_{ij} (0, 0, ..., e_{ij}, 0, ..., 0) = (0, ..., 0)$, alors :

$$(0, 0, ..., \sum_{j=1}^{\dim(Ei)} \lambda_{ij} e_{ij}, 0, ..., 0) = (0, ..., 0)$$

donc $\left(\sum_{j=1}^{\dim(E1)} \lambda_{1j} e_{1j}, ..., \sum_{j=1}^{\dim(Ei)} \lambda_{ij} e_{ij}, ..., \sum_{j=1}^{\dim(Ep)} \lambda_{ij} e_{ij} \right) = (0, ..., 0)$

donc $\forall\, i, \sum_{j=1}^{\dim(Ei)} \lambda_{ij} e_{ij} = 0$

et comme pour tout i, (e_{ij}) est une base de F_i, les λ_{ij} sont nuls.

❑ La famille est génératrice. Si $(x_1, ..., x_p)$ appartient à $\prod_{i=1}^{p} F_i$, il suffit de décomposer chaque x_i sous la forme $\sum_{j=1}^{\dim(Ei)} \mu_{ij} e_{ij}$ pour obtenir que $(x_1, ..., x_p) = \sum_{i=1}^{p} \sum_{j=1}^{\dim(Ei)} \mu_{ij} (0, 0, ..., e_{ij}, 0, ..., 0)$

Dans le cas particulier où $F_1 = ... = F_p = \mathbb{K}$, on retrouve $\dim(\mathbb{K}^p) = p$.

2- Somme de *p* sous–espaces vectoriels

Si les F_i sont tous des sous-espaces vectoriels d'un même espace vectoriel E, on peut considérer l'application Φ suivante :

$$\Phi : \begin{array}{l} \prod_{i=1}^{p} F_i \to E \\ (x_1, ..., x_p) \to x_1 + ... + x_p \end{array}$$

$\mathrm{Im}(\Phi)$ s'appelle somme des F_i, notée $F_1 + ... + F_p$ ou $\sum_{i=1}^{p} F_i$. On a donc :

DEFINITION :

Soit E *un espace vectoriel,* $F_1, F_2, ..., F_p$ *p sous–espaces vectoriels de* E. *On appelle somme des* F_i *l'ensemble défini par :*

$$F_1 + F_2 + ... + F_p = \{ z \mid \exists\, x_i \in F_i, z = x_1 + x_2 + ... + x_p \}.$$

$F_1 + F_2 + ... + F_p$ est le sous–espace vectoriel engendré par la réunion des F_i. En effet, il contient chaque F_i, et tout sous-espace contenant chaque F_i contient leur somme. C'est le plus petit sous–espace vectoriel contenant les F_i.

Comme $\mathrm{rg}(\Phi) \le \dim(\prod_{i=1}^{p} F_i)$, on a :

$$\dim(\sum_{i=1}^{p} F_i) \le \sum_{i=1}^{p} \dim(F_i)$$

Si Φ est injective, on dit que la somme est directe. Le fait d'être injective signifie que chaque élément z de l'image n'a qu'un seul antécédent $(x_1, ..., x_p)$. C'est équivalent à dire que $\mathrm{Ker}(\Phi)$ est réduit à l'élément nul.

On a donc :
PROPOSITION :
On dit que la somme des F_i *est directe si l'une des propriétés équivalentes suivantes est vérifiée* :

i) *Tout* z *de* $F_1 + ... + F_p$ *s'écrit de manière unique* $x_1 + ... + x_p$, $x_i \in F_i$.

ii) $[x_1 + x_2 + ... + x_p = 0$ *et* $\forall\, i, x_i \in F_i] \Rightarrow \forall\, i, x_i = 0$

On note $F_1 \oplus F_2 \oplus ... \oplus F_p$ cette somme directe.

Dans le cas de la dimension finie, on dispose du théorème du rang qui permet d'écrire :

$$\dim(\prod_{i=1}^{p} F_i) = \dim(\mathrm{Im}(\Phi)) + \dim(\mathrm{Ker}(\Phi))$$

et donc :

$$\sum_{i=1}^{p} \dim(F_i) = \dim(\sum_{i=1}^{p} F_i) + \dim(\mathrm{Ker}(\Phi))$$

Comme la somme est directe si et seulement si Φ est injective, on a dispose de l'équivalence :

i-ii) *La somme des* F_i *est directe*.

iii) $\dim(\sum_{i=1}^{p} F_i) = \sum_{i=1}^{p} \dim(F_i)$

Une base de la somme directe s'obtient en réunissant des bases de tous les sous–espaces vectoriels F_i. Cela résulte de l'isomorphisme que Φ établit entre $\prod_{i=1}^{p} F_i$ et $\bigoplus_{i=1}^{p} F_i$ et en prenant comme base de $\bigoplus_{i=1}^{p} F_i$ l'image par Φ de la base de $\prod_{i=1}^{p} F_i$ donnée dans le I-1).

Inversement, si on scinde une base de E en p systèmes disjoints, ces systèmes engendrent des sous-espaces vectoriels en somme directe. En particulier, si $(v_1, ..., v_n)$ est une base de E, on a la somme directe : $E = \mathbb{K}v_1 \oplus ... \oplus \mathbb{K}v_n$ où l'on note $\mathbb{K}v$ le sous–espace vectoriel engendré par v.

Si $E = F_1 \oplus ... \oplus F_p$, pour définir une application linéaire u de E dans un espace vectoriel G, il suffit de définir des applications linéaires u_i de E_i dans G. On pose alors :

$u(x_1 + ... + x_p) = u_1(x_1) + ... + u_p(x_p)$.

Réciproquement, u_i est la restriction de u à E_i.

REMARQUE :
Dans le cas de deux sous-espaces vectoriels, la condition de somme directe :

$$x_1 + x_2 = 0,\ x_1 \in E_1,\ x_2 \in E_2 \Rightarrow x_1 = x_2 = 0$$

est équivalente à celle vue en première année $E_1 \cap E_2 = \{0\}$ en écrivant l'égalité sous la forme $x_1 = -x_2$ élément de $E_1 \cap E_2$.

II : Applications linéaires

1- Trace

On appelle trace d'une matrice carrée la somme de ses éléments diagonaux :

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}$$

Tr est une application linéaire de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$. On dispose également de la propriété suivante :

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

En effet :

$$\mathrm{Tr}(AB) = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}b_{ji}$$

alors que :

$$\mathrm{Tr}(BA) = \sum_{i=1}^{n} (BA)_{ii} = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}a_{ji}$$

Il suffit d'intervertir les notations $i \leftrightarrow j$ pour obtenir la même expression. On en déduit en particulier que, pour toute matrice P inversible $n \times n$ et toute matrice M de $\mathcal{M}_n(\mathbb{K})$:

$$\mathrm{Tr}(P^{-1}MP) = \mathrm{Tr}(MPP^{-1}) = \mathrm{Tr}(M)$$

Or si M représente la matrice d'un endomorphisme u dans une base donnée et si P représente la matrice de passage de cette base à une autre base, $P^{-1}MP$ représente la matrice de u dans la nouvelle base. Les deux matrices, bien que différentes, ont la même trace. Elles ont d'ailleurs aussi le même rang et le même déterminant. Cette trace est appelée trace de u et est indépendante de la base choisie.

EXEMPLE :
La trace d'un projecteur est égale à son rang.
La trace d'une rotation d'angle θ en dimension 3 est égale à $1 + 2\cos\theta$.

2- Sous-espaces stables

a) Définition

Soit u un endomorphisme d'un espace vectoriel E. Un sous-espace vectoriel F de E est dit stable par u si $u(F) \subset F$. La restriction $u|_F$ de l'endomorphisme u à F définit alors un endomorphisme de F, appelé endomorphisme sur F induit par u.

On rencontre usuellement les cas suivants en dimension finie :

❑ F stable : si on choisit une base de F que l'on complète en une base de E, la matrice de u a la forme suivante $\begin{pmatrix} A & B \\ O & C \end{pmatrix}$ où A est une matrice carrée de dimension la dimension de F, O une matrice nulle, B et C des matrices quelconques.

❑ Plus généralement, $F_1 \subset F_2 \subset ... \subset F_p = E$ suite croissante de sous-espaces vectoriels stables, donne, dans une base adaptée, une matrice de la forme

$$\begin{pmatrix} A_1 & * & ... & * \\ O & A_2 & ... & * \\ ... & & ... & \\ O & O & ... & A_p \end{pmatrix}$$

La matrice est dite triangulaire par blocs.

❑ En particulier si dim $F_i = i$, la matrice est triangulaire :

$$\begin{pmatrix} a_1 & * & ... & * \\ 0 & a_2 & ... & * \\ ... & & ... & \\ 0 & 0 & ... & a_p \end{pmatrix}$$

On dit que l'endomorphisme est trigonalisable.

❑ $E = F \oplus G$ avec F et G stables :
La matrice de u dans une base de E obtenue en réunissant une base de F et une base de G est alors de la forme $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$. L'intérêt est de scinder l'étude d'une application linéaire en deux études sur deux sous-espaces plus petits, et peut-être plus maniables.

❑ Plus généralement, si $E = E_1 \oplus E_2 \oplus ... \oplus E_p$, avec tous les E_i stables, alors la matrice de u dans une base adaptée est de la forme

$$\begin{pmatrix} A_1 & O & ... & O \\ O & A_2 & ... & O \\ ... & & ... & \\ O & O & ... & A_p \end{pmatrix}.$$

Inversement, un endomorphisme ayant cette matrice stabilise chacun des E_i. La matrice est dite diagonale par blocs.

❑ L'idéal est d'obtenir une matrice diagonale, pour laquelle $u_{|E_i}$ est une homothétie. On dit que l'endomorphisme est diagonalisable, de matrice

$$\begin{pmatrix} a_1 & 0 & ... & 0 \\ 0 & a_2 & ... & 0 \\ ... & & ... & \\ 0 & 0 & ... & a_p \end{pmatrix}.$$

(les a_i étant distincts ou non).

b) Matrices par blocs

On a vu apparaître ci-dessus des matrices par blocs. Nous donnons ci-dessous un exemple de produit de matrices par blocs : il suffit pour cela de réfléchir à ce qu'est la définition du produit de deux matrices, l'ordre dans lequel se font les choses, les dimensions de chaque bloc pour rendre possible les produits :

A est une matrice $p \times q$ E est une matrice $q \times t$

B est une matrice $p \times r$ F est une matrice $q \times u$
C est une matrice $s \times q$ G est une matrice $r \times t$
D est une matrice $s \times r$ H est une matrice $r \times u$

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} E & F \\ G & H \end{pmatrix} = \begin{pmatrix} AE+BG & AF+BH \\ CE+DG & CF+DH \end{pmatrix}$$

avec AE+BG matrice $p \times t$
AF+BH matrice $p \times u$
CE+DG matrice $s \times t$
CF+DH matrice $s \times u$

Signalons le calcul des déterminants triangulaires (ou diagonaux par blocs) :
Soit D = $\begin{vmatrix} A & C \\ O_{n-p,p} & B \end{vmatrix}$ où A est une matrice carrée $p \times p$, B une matrice carrée $(n-p) \times (n-p)$, C une matrice $p \times (n-p)$ et $O_{n-p,p}$ la matrice nulle à $n-p$ lignes et p colonnes. Montrons par récurrence sur p que D = det(A) $\times$ det(B).
Si $p = 1$, le développement par rapport à la première colonne donne le résultat. Supposons le résultat vrai au rang $p-1$ et montrons-le au rang p. Le développement par rapport à la première colonne donne :

$$D = \sum_{i=1}^{p} (-1)^{i-1}\, a_{i1}\, M_{i1}$$

où M_{i1} est le déterminant obtenu en supprimant la ligne i et la colonne 1 du déterminant complet. Il n'est autre que $\begin{vmatrix} A_{i1} & C_i \\ O_{n-p,p-1} & B \end{vmatrix}$ où A_{i1} est la matrice A dont on a supprimé la ligne i et la colonne 1 et C_i est la matrice C dont on a supprimé la ligne i. L'hypothèse de récurrence donne :

$$M_{i1} = \det(A_{i1}) \det(B)$$

$$\Rightarrow \quad D = \sum_{i=1}^{p} (-1)^{i-1}\, a_{i1} \det(A_{i1}) \det(B) = \det(A) \det(B)$$

puisque $\sum_{i=1}^{p} (-1)^{i-1}\, a_{i1} \det(A_{i1})$ n'est autre que le développement de det(A) par rapport à sa première colonne.

Plus généralement, on aura :

$$\begin{vmatrix} A_1 & * & \dots & * \\ O & A_2 & \dots & * \\ \dots & & \dots & \\ O & O & \dots & A_n \end{vmatrix} = \det(A_1)\dots\det(A_n)$$

où les A_i sont des blocs carrés sur la diagonale, les $*$ désignent des blocs de termes quelconques et les O désignent des blocs nuls. Cela se montre par récurrence sur n.

Donnons également un dernier complément sur les déterminants. On appelle déterminant de Vandermonde (1735-1796) un déterminant de la forme $\begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ & \dots & & \dots & \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix}$. Montrons qu'il est égal à $\prod_{i<j}(a_j - a_i)$. S'il existe $i \neq j$ tel que $a_i = a_j$ alors le déterminant est nul puisque deux lignes sont identiques, de même que le produit $\prod_{i<j}(a_j - a_i)$. On peut donc supposer les a_i distincts. Procédons par récurrence. Le cas $n = 2$ est facile à vérifier. Pour le cas n, considérons-le comme un polynôme en a_n de degré $n - 1$, obtenu en le développant par rapport à la dernière ligne. Ce polynôme s'annule en les $n - 1$ racines distinctes $a_n = a_1$, ..., $a_n = a_{n-1}$ et se factorise donc par $(a_n - a_1)...(a_n - a_{n-1})$. Le coefficient dominant est le mineur du terme (n, n), qui n'est autre que le Vandermonde de rang $n - 1$, à savoir, d'après l'hypothèse de récurrence $\prod_{i<j<n}(a_j - a_i)$. Le déterminant cherché vaut donc :

$$(a_n - a_1)...(a_n - a_{n-1}) \prod_{i<j<n}(a_j - a_i) = \prod_{i<j}(a_j - a_i)$$

c) Obtention de sous-espaces vectoriels stables

On dispose de la proposition suivante :

PROPOSITION

Soit v un endomorphisme commutant avec l'endomorphisme u. Alors $\mathrm{Ker}(v)$ *et* $\mathrm{Im}(v)$ *sont stables par u.*

En effet :

$x \in \mathrm{Ker}(v) \Rightarrow v(x) = 0 \Rightarrow u(v(x)) = 0 = v(u(x)) \Rightarrow u(x) \in \mathrm{Ker}(v)$

$x \in \mathrm{Im}(v) \Rightarrow \exists\, z, x = v(z) \Rightarrow \exists\, z, u(x) = u(v(z)) = v(u(z)) \Rightarrow u(x) \in \mathrm{Im}(v)$

Pour trouver des sous-espaces stables par u, il suffit donc de trouver des endomorphismes v commutant avec u. Il y a par exemple u lui-même, mais aussi u^2, u^3, ... et leurs combinaisons linéaires $\sum_{i \geq 0} a_i u^i$. On est ainsi amené à introduire les polynômes d'endomorphisme.

3- Polynômes d'endomorphisme

Soit P un polynôme quelconque de $\mathbb{K}[X]$ et u un endomorphisme. Si $P = \sum_i a_i X^i$, on pose $P(u) = \sum_i a_i u^i$ avec la convention $u^0 = \mathrm{Id}_E$. On définit de même $P(M) = \sum_i a_i M^i$ pour une matrice carrée M, élément de $\mathcal{M}_{np}(\mathbb{R})$ avec $M^0 = I_n$.

On vérifie que les applications $\mathbb{K}[X] \to \mathcal{M}_n(\mathbb{K})$ et $\mathbb{K}[X] \to L(E)$, qui à P associe P(M) ou P(u), sont compatibles avec les opérations dans chaque espace. Ainsi, en notant $+$, $.$ et $\times$ la somme de polynômes, le produit par un scalaire et le produit de polynômes, par $+$, $.$ et $\times$ le produit de matrices, et par +, ., et o la somme d'endomorphismes, le produit par un scalaire et le produit de composition des endomorphismes, nous avons :

$(P + Q)(M) = P(M) + Q(M)$ $\qquad$ $(P + Q)(u) = P(u) + Q(u)$

$(\lambda.P)(M) = \lambda.P(M)$ $\qquad$ $(\lambda.P)(u) = \lambda.P(u)$

$(P \times Q)(M) = P(M) \times Q(M)$ $\qquad$ $(P \times Q)(u) = P(u) \circ Q(u)$

En effet, si $P = \sum\limits_i a_i X^i$ et $Q = \sum\limits_i b_i X^i$, alors $P + Q = \sum\limits_i (a_i + b_i) X^i$, $\lambda.P = \sum\limits_i \lambda a_i X^i$ et

$P \times Q = \sum\limits_m \sum\limits_{i=0}^{m} a_i b_{m-i} X^m$. De sorte que :

$$(P + Q)(M) = \sum_i (a_i + b_i) M^i = \sum_i a_i M^i + \sum_i b_i M^i = P(M) + Q(M)$$

$$(\lambda.P)(M) = \sum_i \lambda a_i M^i = \lambda.\sum_i a_i M^i = \lambda.P(M)$$

$(P \times Q)(M) = \sum\limits_m \sum\limits_{i=0}^{m} a_i b_{m-i} M^m = \sum\limits_i a_i M^i \times \sum\limits_j b_j M^j$ en développant le membre de droite et en regroupant les termes M^{i+j} tels que $i + j = m$. Les formules sont identiques pour les endomorphismes, sauf que le produit des endomorphismes est noté o. Dans ce dernier cas, on a donc :

$(P \times Q)(u) = P(u) \circ Q(u)$

ou, plus brièvement :

$(PQ)(u) = P(u) \circ Q(u)$

ou encore, pour tout x de E :

$(PQ)(u)(x) = (P(u) \circ Q(u))(x) = P(u)(Q(u)(x))$

En particulier, on a $P(u) \circ Q(u) = Q(u) \circ P(u)$ puisque PQ = QP.

Comme, pour tout polynôme P, $P(u)$ commute avec u, on en déduit que, pour tout polynôme P, $\mathrm{Ker}(P(u))$ et $\mathrm{Im}(P(u))$ sont stables par u.

Malheureusement, si on prend n'importe quel polynôme P, il y a toutes les chances que l'on obtienne des sous-espaces stables triviaux, à savoir $\mathrm{Im}(P(u)) = E$ et $\mathrm{Ker}(P(u)) = \{0\}$ ou l'inverse. Le cas $\mathrm{Im}(P(u)) = E$ et $\mathrm{Ker}(P(u)) = \{0\}$ s'obtient lorsque $P(u)$ est un endomorphisme inversible. Le cas $\mathrm{Im}(P(u)) = \{0\}$ et $\mathrm{Ker}(P(u)) = E$ s'obtient lorsque $P(u) = 0$. On dit alors que P est un polynôme annulateur de u. Ce second cas peut cependant parfois donner des situations intéressantes.

4- Polynômes annulateurs d'un endomorphisme

Début de partie réservée aux PSI/PSI/MP/MP**

a) Définition

Soit u un endomorphisme de E de dimension finie n. L(E) est lui-même de dimension finie n^2. Il en résulte que $(\mathrm{Id}, u, u^2, ..., u^{n^2})$, constitué de $n^2 + 1$ éléments d'un espace de dimension n^2 forme un système lié. Il existe donc des a_k non tous nuls tels que $\sum\limits_k a_k u^k = 0$. Autrement dit, il existe au moins un polynôme P non nul tel que $P(u) = 0$. Nous admettrons qu'il existe d'ailleurs un tel polynôme de degré n seulement, à savoir $P(X) = \det(u - X\,\mathrm{Id})$ ou $\det(X\,\mathrm{Id} - u)$ (ce dernier ayant l'intérêt d'être unitaire), ce qui constitue le :

THEOREME DE CAYLEY-HAMILTON
Soit u un endomorphisme. Le polynôme $\chi_u(X) = \det(X\,\mathrm{Id} - u)$ *s'appelle polynôme caractéristique de u. Il s'agit d'un polynôme annulateur de u.*

Ce théorème est admis. Deux démonstrations en sont données dans l'annexe III.

EXEMPLE 1 : Soit $M = \begin{pmatrix} -1 & 3 \\ -2 & 4 \end{pmatrix}$. Son polynôme caractéristique est : $X^2 - 3X + 2$, et on vérifie que l'on a effectivement $M^2 - 3M + 2I = 0$.

EXEMPLE 2 :
Soit $M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Son polynôme caractéristique vaut $(X - 2)^2(X - 1)$. Mais on peut remarquer que $(M - 2I)(M - I) = 0$, donc $(X - 2)(X - 1)$ annule aussi M. Il peut donc exister d'autres polynômes annulateurs que le polynôme caractéristique.

b) Utilisation d'un polynôme annulateur pour rechercher des sous-espaces stables
Soit P un polynôme annulateur de l'endomorphisme u. Supposons que ce polynôme soit tel qu'il se factorise en deux facteurs $P = P_1P_2$ avec $\deg P_1 \geq 1$ et $\deg P_2 \geq 1$. avec $P_1(u) \neq 0$ et $P_2(u) \neq 0$. Donc $\mathrm{Ker}(P_1(u)) \neq E$ et $\mathrm{Im}(P_1(u)) \neq \{0\}$, et de même pour P_2.
En outre, $P_1(u)$ ne peut être inversible, sinon on aurait $0 = P(u) = P_1(u) \circ P_2(u) \Rightarrow P_2(u) = 0$ en composant à gauche par l'inverse de $P_1(u)$. Et de même, $P_2(u)$ n'est pas inversible. Donc $\mathrm{Ker}(P_1(u)) \neq \{0\}$ et $\mathrm{Im}(P_1(u)) \neq E$, et de même pour P_2.

$\mathrm{Ker}(P_1(u))$ et $\mathrm{Im}(P_1(u))$ sont alors des sous-espaces vectoriels stables non triviaux de u.

Dans l'exemple 1, $X^2 - 3X + 2 = (X - 1)(X - 2)$ donc on pourra chercher par exemple $\mathrm{Ker}(u - I)$ ou $\mathrm{Ker}(u - 2I)$. Ce type de recherche intervient de manière plus systématique dans la diagonalisation des endomorphismes.

c) Utilisation d'un polynôme annulateur pour calculer une puissance de matrice
Donnons deux exemples :

EXEMPLE 1 :
Reprenons l'exemple $M = \begin{pmatrix} -1 & 3 \\ -2 & 4 \end{pmatrix}$ dont un polynôme annulateur est $X^2 - 3X + 2$ et considérons la division euclidienne de X^n par ce polynôme :

$$X^n = (X^2 - 3X + 2)Q + R \qquad \text{où } \deg(R) < 2$$

dont il existe a_n et b_n tels que :

$$X^n = (X^2 - 3X + 2)Q + a_nX + b_n$$

Pour X = 1 et X = 2, racines du polynôme annulateur, on obtient respectivement :

$$1 = a_n + b_n$$
$$2^n = 2a_n + b_n$$

d'où :

$$a_n = 2^n - 1$$

$b_n = 2 - 2^n$

Ainsi :

$X^n = (X^2 - 3X + 2)Q + (2^n - 1)X + 2 - 2^n$

Pour X = M, et sachant que $M^2 - 3M + 2I = 0$, on obtient :

$M^n = (2^n - 1)M + (2 - 2^n)I$

EXEMPLE 2 :

Considérons $M = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix}$ dont le polynôme caractéristique est $X^2 - 2X + 1 = (X - 1)^2$ qui donc également un polynôme annulateur de M. Considérons la division euclidienne de X^n par ce polynôme. Comme ci-dessus, il existe a_n et b_n tels que :

$X^n = (X - 1)^2Q + a_nX + b_n$

Pour X = 1 racine du polynôme annulateur, on obtient :

$1 = a_n + b_n$

Mais ici, 1 est racine double. On obtient une autre relation en prenant la dérivée des polynômes :

$nX^{n-1} = (X - 1)S + a_n$ où $S = 2Q + (X - 1)Q'$

Pour X = 1, on obtient :

$n = a_n$

d'où :

$a_n = n$

$b_n = 1 - n$

Ainsi :

$X^n = (X - 1)^2Q + nX + 1 - n$

Pour X = M, et sachant que $(M - I)^2 = 0$, on obtient :

$M^n = nM + (1 - n)I$

<u>d) Utilisation d'un polynôme annulateur pour trouver l'inverse d'une matrice</u>

Si on dispose d'un polynôme annulateur d'une matrice M $a_0 + a_1X + ... + a_pX^p$ avec $a_0 \neq 0$, alors on peut conclure que M est inversible et trouver facilement son inverse. En effet :

$a_0I + a_1M + ... + a_pM^p = 0$

donc $I = \dfrac{-a_1M - ... - a_pM^p}{a_0} = M \times \dfrac{(-a_1I - a_2M - ... - a_pM^{p-1})}{a_0}$

donc l'inverse de M est $-\dfrac{(a_1I + a_2M + ... + a_pM^{p-1})}{a_0}$.

EXEMPLE 1 :

Pour $M = \begin{pmatrix} -1 & 3 \\ -2 & 4 \end{pmatrix}$ vérifiant $M^2 - 3M + 2I = 0$, on a $M \times \dfrac{(M - 3I)}{-2} = I$ donc $M^{-1} = \dfrac{3I - M}{2}$. Il est intéressant de noter que cette expression correspond à l'expression $M^n = (2^n - 1)M + (2 - 2^n)I$ trouvée plus haut pour $n = -1$, alors qu'a priori, cette expression n'a été prouvée que pour $n \geq 0$.

EXEMPLE 2 :

Pour $M = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix}$ vérifiant $M^2 - 2M + I = 0$, on a $M \times (2I - M) = I$ donc $M^{-1} = 2I - M$. Ici aussi, on constate qu'on obtient cette expression en posant $n = -1$ dans la relation $M^n = nM + (1 - n)I$ valide a priori pour $n \geq 0$.

Fin de la partie réservée aux PSI/PSI/MP/MP*. Retour à la partie commune PSI/MP/PC*

5- Formes linéaires et hyperplans

Une forme linéaire est un élément de $L(E,\mathbb{K})$. On la note le plus souvent avec une lettre grecque.

PROPOSITION

Soit H *un sous-espace vectoriel d'un espace vectoriel* E, *de dimension finie ou non. Il y a équivalence entre :*

(i) H *admet un supplémentaire de dimension* 1

(ii) H *est le noyau d'un forme linéaire non nulle* φ

H *s'appelle un hyperplan. Les formes linéaires* φ *dont* H *est le noyau sont définies à une constante multiplicative près.*

Démonstration :

(ii) $\Rightarrow$ (i)

Soit u est un vecteur n'appartenant pas à H (et donc tel que $\varphi(u) \neq 0$). Notons D la droite engendrée par u et montrons que $E = D \oplus H$.

Pour tout x de E : $x = \dfrac{\varphi(x)}{\varphi(u)} u + (x - \dfrac{\varphi(x)}{\varphi(u)} u)$ avec $\dfrac{\varphi(x)}{\varphi(u)} u \in D$ et $x - \dfrac{\varphi(x)}{\varphi(u)} u \in H$. Donc $x \in D + H$.

Par ailleurs, la somme est directe, car si $x \in D \cap H$, alors $\exists\ \lambda$, $x = \lambda u$ et $\varphi(x) = 0$ donc $\varphi(\lambda u) = 0 = \lambda\varphi(u)$. Comme $\varphi(u) \neq 0$, on a $\lambda = 0$ et donc $x = 0$.

(i) $\Rightarrow$ (ii)

Réciproquement, un sous-espace vectoriel H admettant comme supplémentaire une droite D de vecteur directeur n est le noyau d'un forme linéaire. En effet, si $x = x_H + \lambda n$ est la décomposition de x, il suffit de définir $\varphi(x) = \lambda$. $\varphi(x) = 0$ est une équation de H.

On généralise ainsi l'équation $ax + by + cz = 0$ d'un plan en dimension 3.

Si $\varphi(x) = 0$ est une équation de H, toute autre équation de H est proportionnelle à celle-ci. En effet, soit ψ est une autre forme linéaire, nulle sur H, et soit u un vecteur n'appartenant pas à H et D la droite vectorielle engendrée par u. La décomposition donnée plus haut

$$x = \underbrace{\frac{\varphi(x)}{\varphi(u)} u}_{\text{élément de D}} + \underbrace{(x - \frac{\varphi(x)}{\varphi(u)} u)}_{\text{élément de H}}$$

permet de conclure que ::

$$\psi(x) = \varphi(x) \frac{\psi(u)}{\varphi(u)} = \text{Cte} \times \varphi(x)$$

Plusieurs remarques :

En dimension finie, le theorème du rang permet de dire que $\dim \text{Ker}(\varphi) = \dim E - 1$, montrant de nouveau l'équivalence entre (i) et (ii).

Si u est choisi de façon que $\varphi(u) = 1$, alors la décomposition se simplifie en :

$$x = \varphi(x)\, u + (x - \varphi(x)\, u)$$

La décomposition ci-dessus généralise le cas de E, espace euclidien : si n est unitaire et orthogonal à l'hyperplan H, on a :

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

et $\quad x = \langle x,n\rangle n + x - \langle x,n\rangle n$

De même que la décomposition $x = \langle x,n\rangle n + x - \langle x,n\rangle n$ permet de trouver facilement l'expression de la projection orthogonale sur une droite de vecteur directeur unitaire n (à savoir $\langle x,n\rangle\, n$), de même la décomposition $x = \varphi(x)u + (x - \varphi(x)u)$ permet de trouver rapidement l'expression de la projection parallèlement à l'hyperplan d'équation $\varphi(x) = 0$, sur la droite de vecteur directeur u tel que $\varphi(u) = 1$, (à savoir $\varphi(x)u$)

EXEMPLES :

❑ Expression de la projection sur la droite de vecteur directeur $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ parallèlement au plan d'équation $2x + y - z = 0$. Cette expression vaut $\dfrac{2x + y - z}{5}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$.

❑ Soit E l'espace vectoriel des fonctions continues sur $[-1,1]$ et H le sous-espace vectoriel des fonctions s'annulant en 0. H est un hyperplan. En effet, soit $\varphi : E \to \mathbb{R}$ définie pour f dans E par $\varphi(f) = f(0)$. φ est une forme linéaire, et $H = \mathrm{Ker}(\varphi)$.
Soit u une fonction de E n'appartenant pas à H, par exemple la fonction constante égale à 1. Alors u endendre une droite D supplémentaire de H. Toute fonction f se décompose sous la forme :

$$f = \frac{\varphi(f)}{\varphi(u)}\, u + \left(f - \frac{\varphi(f)}{\varphi(u)}\, u\right) = f(0) + (f - f(0))$$

III : Diagonalisation

1- Exemples de problèmes

Considérons les problèmes suivants :

i) Calculer M^n, où M est une matrice carrée, par exemple :

$$M = \begin{pmatrix} 8 & 0 & 9 \\ -3 & -1 & -3 \\ -6 & 0 & -7 \end{pmatrix}$$

ii) Trouver les suites vérifiant les relations suivantes :

x_0, y_0 et z_0 sont donnés

$$\text{pour tout } k : \begin{cases} x_{k+1} = 8x_k + 9z_k \\ y_{k+1} = -3x_k - y_k - 3z_k \\ z_{k+1} = -6x_k - 7z_k \end{cases}$$

iii) Trouver les fonctions vérifiant le système différentiel suivant :

$x(0)$, $y(0)$ et $z(0)$ sont donnés

$$\text{pour tout } t \text{ réel} : \begin{cases} x'(t) = 8x(t) + 9z(t) \\ y'(t) = -3x(t) - y(t) - 3z(t) \\ z'(t) = -6x(t) - 7z(t) \end{cases}$$

Les problèmes i) et ii) sont équivalents. En effet, dans ii), on peut poser :

$$X_k = \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} \text{élément de } \mathbb{R}^3.$$

On a alors la relation de récurrence suivante entre v_{k+1} et v_k :

$$X_{k+1} = MX_k, \text{ M étant la matrice du i),}$$

d'où il est facile de tirer, par récurrence, que : $X_k = M^k X_0$.

La difficulté de ces problèmes tient au fait que la matrice utilisée n'est pas diagonale. Comparons en effet les trois problèmes précédents aux trois problèmes suivants :

ibis) Calculer D^k, où D est une matrice diagonale, par exemple :

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

iibis) Trouver les suites vérifiant les relations suivantes :

x_0, y_0 et z_0 sont donnés ;

$$\text{pour tout } k : \begin{cases} x_{k+1} = -x_k \\ y_{k+1} = -y_k \\ z_{k+1} = 2z_k \end{cases}$$

iiibis) Trouver les fonctions vérifiant le système différentiel suivant :

$x(0)$, $y(0)$ et $z(0)$ sont donnés ;

$$\text{pour tout } t \text{ réel} : \begin{cases} x'(t) = -x(t) \\ y'(t) = -y(t) \\ z'(t) = 2z(t) \end{cases}$$

La résolution est immédiate ; on a respectivement :

$$D^k = \begin{pmatrix} (-1)^k & 0 & 0 \\ 0 & (-1)^k & 0 \\ 0 & 0 & 2^k \end{pmatrix} \qquad \begin{cases} x_k = (-1)^k x_0 \\ y_k = (-1)^k y_0 \\ z_k = 2^k z_0 \end{cases} \qquad \begin{cases} x(t) = e^{-t}\, x(0) \\ y(t) = e^{-t}\, y(0) \\ z(t) = e^{2t}\, z(0) \end{cases}$$

Le principe de la diagonalisation consiste à passer, si possible d'une matrice quelconque (et d'un problème général) à une matrice diagonale (et à un problème plus simple).

2- Valeurs propres, vecteurs propres, sous-espaces propres

DEFINITION :

Un endomorphisme u d'un espace vectoriel E *est dit diagonalisable s'il existe une base de* E *dans laquelle la matrice de u est diagonale.*

Une matrice carrée M est dite diagonalisable si elle est semblable à une matrice diagonale D.

M est semblable à D s'il existe P inversible telle que $M = PDP^{-1}$. M et D représentent le même endomorphisme dans des bases différentes, la matrice de passage de l'une à l'autre étant P. En pratique, M est donnée, et on cherche D, si c'est possible. Si $(e_1, e_2, ..., e_n)$ est la base (si elle existe) dans laquelle la matrice de u est D, diagonale, de la forme

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & ... \\ 0 & \lambda_2 & 0 & ... \\ ... & & ... & \\ ... & & 0 & \lambda_n \end{pmatrix}$$

on a $u(e_i) = \lambda_i e_i$. On est donc amené à s'intéresser aux vecteurs transformés en un multiple d'eux-même par u.

DEFINITION

Soit u un endomorphisme d'un espace vectoriel E. *On appelle vecteur propre v de u tout vecteur non nul de* E *pour lequel il existe un scalaire λ tel que :*

$$u(v) = \lambda v$$

λ s'appelle valeur propre relative au vecteur propre v. λ est une valeur propre de u.
λ étant une valeur propre, on appelle sous-espace propre associé à λ le sous-espace

$$E_\lambda = \{ v \in E \mid u(v) = \lambda v \}$$

Si v est vecteur propre, λ est unique. En effet :

$$u(v) = \lambda v \text{ et } u(v) = \lambda' v \Rightarrow \lambda v = \lambda' v \Rightarrow \lambda = \lambda' \text{ car } v \text{ est non nul.}$$

L'ensemble des valeurs propres d'un endomorphisme u s'appelle le spectre de u, noté $\mathrm{Sp}(u)$.
E_λ est bien un sous-espace vectoriel. Il suffit de constater que $E_\lambda = \mathrm{Ker}(u - \lambda \mathrm{Id})$

On remarque que l'ensemble des vecteurs propres associés à λ n'est autre que $E_\lambda - \{0\}$. λ est donc une valeur propre associée à un vecteur propre non nul si et seulement si $\mathrm{Ker}(u - \lambda\mathrm{Id}) \neq \{0\}$, ce qui est vérifié si et seulement si $u - \lambda\mathrm{Id}$ est non inversible, et donc si et seulement si $\det(u - \lambda\mathrm{Id}) = 0$. Il s'agit d'un polynôme en λ dont on peut montrer que le degré vaut $n = \dim E$ (voir en fin de chapitre le paragraphe trigonalisation). Ainsi, les valeurs propres se trouvent en cherchant les racines du polynôme $\det(u - X\mathrm{Id})$, polynôme qu'on a déjà rencontré plus haut.

PROPOSITION
$\det(X\mathrm{Id} - u) = \chi_u(X)$ *s'appelle polynôme caractéristique de l'endomorphisme u. Ses racines sont les valeurs propres de u. L'ordre de multiplicité de cette valeur en tant que racine de χ_u s'appelle ordre de multiplicité d'une valeur propre.*

EXEMPLES :
Cherchons les valeurs propres de quelques applications classiques :

❑ Les homothéties $\lambda\mathrm{Id}$: La seule valeur propre est λ. L'espace tout entier est sous-espace propre.

❑ Les projecteurs : Les valeurs propres sont 0 et 1, le sous-espace propre associé à la valeur propre 0 étant le noyau du projecteur (la direction parallélement à laquelle on projette). Le sous-espace propre associé à la valeur propre 1 est l'image du projecteur (le sous-espace sur lequel on projette).

❑ Les symétries : Les valeurs propres sont –1 et 1, le sous-espace propre associé à la valeur propre –1 étant la direction parallèlement à laquelle on effectue la symétrie. Le sous-espace propre associé à la valeur propre 1 est le sous-espace par rapport auquel s'effectue la symétrie.

3- Méthode pratique

Pour diagonaliser une matrice ou un endomorphisme, on procède comme suit :

a) Recherche des valeurs propres

On calcule $\chi_u(X) = \det(X\,\mathrm{Id} - u)$ ou, plus commodément, $\det(u - X\,\mathrm{Id})$ puis on cherche les racines de ce polynôme. Ce sont les valeurs propres. Pour une matrice M de $\mathcal{M}_n(\mathbb{K})$, on calcule $\det(M - XI_n)$.

EXEMPLE : $M = \begin{pmatrix} 8 & 0 & 9 \\ -3 & -1 & -3 \\ -6 & 0 & -7 \end{pmatrix}$

$$\det(M - XI) = \begin{vmatrix} 8-X & 0 & 9 \\ -3 & -1-X & -3 \\ -6 & 0 & -7-X \end{vmatrix} = -(X-2)(X+1)^2$$

Les valeurs propres sont –1 valeur propre double, et 2 valeur propre simple.

b) Recherche des sous–espaces propres

Pour chaque valeur propre λ trouvée, on résout le système $(u - \lambda \mathrm{Id})v = 0$. L'ensemble des vecteurs v solutions est le sous–espace propre E_λ associé à λ.

EXEMPLE : $M = \begin{pmatrix} 8 & 0 & 9 \\ -3 & -1 & -3 \\ -6 & 0 & -7 \end{pmatrix}$

- Sous–espace propre associé à la valeur propre –1. On résout le système :

$$(M + I)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \Leftrightarrow \begin{cases} 9x + 9z = 0 \\ -3x - 3z = 0 \\ -6x - 6z = 0 \end{cases} \Leftrightarrow x + z = 0$$

Il s'agit du plan vectoriel engendré par $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

- Sous–espace propre associé à la valeur propre 2. On résout le système :

$$(M-2I)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \Leftrightarrow \begin{cases} 6x + 9z = 0 \\ -3x - 3y - 3z = 0 \\ -6x - 9z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 3t \\ y = -t \\ z = -2t \end{cases}$$

Il s'agit de la droite vectorielle engendrée par $\begin{pmatrix} 3 \\ -1 \\ -2 \end{pmatrix}$

c) Recherche d'un base de vecteurs propres

Dans chaque E_λ, on choisit une base. Nous verrons dans le paragraphe suivant que la réunion des vecteurs ainsi obtenus forme un système libre (nous verrons que les sous-espaces propres sont en somme directe). Dans le cas précédent, les trois vecteurs trouvés forment notre base de vecteurs propres. Donc M est diagonalisable, semblable à la matrice diagonale $D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

La matrice de passage est $P = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ -1 & 0 & -2 \end{pmatrix}$. On a $M = PDP^{-1}$ avec $P^{-1} = \begin{pmatrix} -2 & 0 & -3 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$.

On peut maintenant résoudre les questions du III-1).

i) Calculer M^k avec $M = \begin{pmatrix} 8 & 0 & 9 \\ -3 & -1 & -3 \\ -6 & 0 & -7 \end{pmatrix}$

Comme $M = PDP^{-1}$, avec $D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ -1 & 0 & -2 \end{pmatrix}$ et $P^{-1} = \begin{pmatrix} -2 & 0 & -3 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$, on a :

$M^k = PD^kP^{-1}$. D'où :

$$M^k = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ -1 & 0 & -2 \end{pmatrix} \begin{pmatrix} (-1)^k & 0 & 0 \\ 0 & (-1)^k & 0 \\ 0 & 0 & 2^k \end{pmatrix} \begin{pmatrix} -2 & 0 & -3 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2(-1)^{k+1} + 3.2^k & 0 & 3(-1)^{k+1} + 3.2^k \\ (-1)^k - 2^k & (-1)^k & (-1)^k - 2^k \\ 2(-1)^k - 2^{k+1} & 0 & 3(-1)^k - 2^{k+1} \end{pmatrix}$$

On vérifie que, pour $k = 0$, on trouve bien $M^k = I$, et pour $n = 1$, on trouve $M^k = M$.

ii) Les solutions du problème III–1–ii) sont :

$$x_k = (2(-1)^{k+1} + 3.2^k)\, x_0 + (3(-1)^{k+1} + 3.2^k)\, z_0$$
$$y_k = ((-1)^k - 2^k)\, x_0 + (-1)^k\, y_0 + ((-1)^k - 2^k)\, z_0$$
$$z_k = (2(-1)^k - 2^{k+1})\, x_0 + (3(-1)^k - 2^{n+1})\, z_0$$

iii) Résolvons le problème du III–1–iii). Soit $X(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$.

$X'(t) = MX(t)$

$\Leftrightarrow \quad X'(t) = PDP^{-1}X(t)$

$\Leftrightarrow \quad P^{-1}X'(t) = DP^{-1}X(t)$

$\Leftrightarrow \quad Y'(t) = DY(t)$ où l'on a posé $Y(t) = P^{-1}Y(t)$

On a donc $Y(0) = P^{-1}X(0) = \begin{pmatrix} -2 & 0 & -3 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix} = \begin{pmatrix} -2x(0) - 3z(0) \\ x(0) + y(0) + z(0) \\ x(0) + z(0) \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix}$

En outre, si on pose $Y(t) = \begin{pmatrix} u(t) \\ v(t) \\ w(t) \end{pmatrix}$, l'équation $Y' = DY$ conduit au système $\begin{cases} u' = -u \\ v' = -v \\ w' = 2w \end{cases}$ qui conduit aux solutions $Y(t) = \begin{pmatrix} e^{-t}u_0 \\ e^{-t}v_0 \\ e^{2t}w_0 \end{pmatrix}$. D'où $X(t) = PY(t) = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ -1 & 0 & -2 \end{pmatrix} \begin{pmatrix} e^{-t}u_0 \\ e^{-t}v_0 \\ e^{2t}w_0 \end{pmatrix} = \begin{pmatrix} e^{-t}u_0 + 3e^{2t}w_0 \\ e^{-t}v_0 - e^{2t}w_0 \\ -e^{-t}u_0 - 2e^{2t}w_0 \end{pmatrix}$

D'où :

$$x(t) = (-2e^{-t} + 3e^{2t})\, x(0) + (-3e^{-t} + 3e^{2t})\, z(0)$$
$$y(t) = (e^{-t} - e^{2t})\, x(0) + e^{-t}\, y(0) + (e^{-t} - e^{2t})\, z(0)$$
$$z(t) = (2e^{-t} - 2e^{2t})\, x(0) + (3e^{-t} - 2e^{2t})\, z(0)$$

4- Propriétés des sous-espaces propres et des valeurs propres

PROPOSITION :

Soit λ *une valeur propre de* u, *de vecteur propre* x. *Alors, pour tout polynôme* P, $P(u)(x) = P(\lambda)x$

Démonstration :

Le lecteur vérifiera que, par récurrence sur k, on a pour tout k, $u^k(x) = \lambda^k x$. Il en résulte que, si $P = \sum_{k\geq 0} a_k X^k$, alors :

$$P(u)(x) = \sum_{k\geq 0} a_k\, u^k(x) = \sum_{k\geq 0} a_k\, \lambda^k\, x = P(\lambda)x$$

Tout polynôme annulateur de u permet d'avoir une idée des valeurs propres de u :

PROPOSITION :

Soit u un endomorphisme (ou M une matrice carrée) et P *un polynôme annulmateur de u* (*ou de* M).
Alors les valeurs propres λ *de u (ou de M) vérifient nécessairement racines de* P.

Démonstration :
Soit λ une valeur propre de u, de vecteur propre x (non nul, donc). On a alors :

$$P(u)(x) = P(\lambda)x$$

Or $P(u) = 0$, donc $P(\lambda)x = 0$, et comme $x \neq 0$, $P(\lambda) = 0$

Le spectre de u est donc inclus dans l'ensemble des racines de P.

EXEMPLE : si $u^2 - 3u + 2\mathrm{Id} = 0$, alors $\mathrm{Sp}(u) \subset \{1,2\}$. Il est possible cependant qu'une seule de ces deux valeurs soit valeur propre de u (par exemple, si $u = \mathrm{Id}$).

PROPOSITION :
Soient $\lambda_1, \lambda_2, \dots, \lambda_m$ *m valeurs propres distinctes. Alors la somme des sous–espaces vectoriels* $E_{\lambda_1} + E_{\lambda_2} + \dots + E_{\lambda_m}$ *est directe.*

Démonstration 1 :
Elle se fait par récurrence sur m.
Montrons cette proposition pour $m = 2$. Il suffit de montrer que $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$. Soit v élément de $E_{\lambda_1} \cap E_{\lambda_2}$. On a alors :

$$u(v) = \lambda_1 v \text{ et } u(v) = \lambda_2 v \Rightarrow \lambda_1 v = \lambda_2 v$$
$$\Rightarrow (\lambda_1 - \lambda_2)v = 0$$
$$\Rightarrow v = 0 \text{ car } \lambda_1 \text{ et } \lambda_2 \text{ sont distincts.}$$

Supposons la propriété vraie pour $m-1$, montrons la pour m.
Il suffit de montrer que 0 se décompose de manière unique en somme d'éléments de E_{λ_i}, autrement dit, que l'on a :

$$\forall i,\ v_i \in E_{\lambda_i} \text{ et } 0 = v_1 + \dots + v_m \Rightarrow \text{pour tout } i,\ v_i = 0$$

En effet, on applique u. On obtient :

$$0 = u(v_1) + \dots + u(v_m)$$
$$\Rightarrow \quad 0 = \lambda_1 v_1 + \dots + \lambda_m v_m \quad (*)$$

On a, par ailleurs, en multipliant par λ_m :

$$0 = v_1 + \dots + v_m$$
$$\Rightarrow \quad 0 = \lambda_m v_1 + \dots + \lambda_m v_m \quad (**)$$

En retranchant (*) et (**), on obtient :

$$0 = (\lambda_1 - \lambda_m)v_1 + \dots + (\lambda_{m-1} - \lambda_m)v_{m-1}$$

$\Rightarrow$ $\forall i \in \{1,\dots,m-1\}$ $(\lambda_1 - \lambda_i)v_i = 0$ car les $m-1$ sous–espaces E_{λ_i} sont en somme directe (hypothèse de récurrence)
$\Rightarrow$ $\forall i \in \{1,\dots,m-1\}$ $v_i = 0$ car les valeurs propres sont distinctes.

En reportant dans (*), on en tire aussi que $v_m = 0$

Démonstration 2 :
Considérons une relation

$$0 = v_1 + \dots + v_m$$

avec v_i élément de E_{λ_i}. Si on applique $m-1$ fois l'endomorphisme u, on obtient successivement :

$$0 = \lambda_1 v_1 + \ldots + \lambda_m v_m$$
$$0 = \lambda_1^2 v_1 + \ldots + \lambda_m^2 v_m$$
$$0 = \lambda_1^3 v_1 + \ldots + \lambda_m^3 v_m$$
$$\ldots$$
$$0 = \lambda_1^{m-1} v_1 + \ldots + \lambda_m^{m-1} v_m$$

Cherchons les valeurs possibles des v_i en raisonnant composante par composante sur une base quelconque. Les coefficients du système sont $\begin{pmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_m \\ \ldots & & \ldots \\ \lambda_1^{m-1} & \ldots & \lambda_m^{m-1} \end{pmatrix}$ qui est une matrice de Vandermonde, de rang m, car les coefficients λ_i sont distincts. Les seules solutions au système sont donc les solutions nulles, et tous les v_i sont nuls.

Démonstration 3 :
Considérons une relation

$$0 = v_1 + \ldots + v_m$$

avec v_i élément de E_{λ_i}. Soit P le polynôme $(X - \lambda_2)(X - \lambda_3)\ldots(X - \lambda_m)$ et appliquons l'endomorphisme $P(u)$ aux deux membres de l'égalité. Compte tenu du fait que v_i est vecteur propre associé à la valeur propre λ_i, on a :

$$P(u)(v_1) = P(\lambda_1)v_1 = (\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)\ldots(\lambda_1 - \lambda_m)v_1$$
$$P(u)(v_2) = P(\lambda_2)v_2 = 0$$
$$\ldots$$
$$P(u)(v_m) = P(\lambda_m)v_m = 0$$

Donc, en sommant tous les termes :

$$0 = P(u)(v_1 + \ldots + v_m)$$
$$= (\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)\ldots(\lambda_1 - \lambda_m)v_1$$

Comme les λ_i sont distincts entre eux, on a $(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)\ldots(\lambda_1 - \lambda_m) \neq 0$ et donc $v_1 = 0$.
On procèdera de même pour montrer que $v_i = 0$ en appliquant cette fois l'endomorphisme $P(u)$ avec :

$$P(X) = (X - \lambda_2)(X - \lambda_3) \ldots (X - \lambda_{i-1})(X - \lambda_{i+1}) \ldots (X - \lambda_m)$$

Le théorème n'affirme pas que la somme des sous-espaces propres donne E tout entier. C'est d'ailleurs un obstacle à la diagonalisation. Les conditions de diagonalisation sont étudiées au III-6.

COROLLAIRE
Toute famille de m vecteurs propres associés à m valeurs propres deux à deux distinctes forme une famille libre.

EXEMPLE :
Pour tout n, la famille $(1, \cos(x), \cos(2x), \ldots, \cos(nx))$ forme une famille libre. En effet, chaque fonction est vecteur propre de l'opérateur D^2 (dérivée seconde) avec des valeurs propres distinctes.

PROPOSITION
Tout sous-espace propre E_λ *d'un endomorphisme u est stable par u. Si v commute avec u, alors* E_λ *est stable également par v.*

Démonstration :
E_λ est stable par u :

$$x \in E_\lambda \Rightarrow u(x) = \lambda x \text{ et } \lambda x \in E_\lambda$$
$$\Rightarrow u(x) \in E_\lambda$$

E_λ est stable par v si $v \circ u = u \circ v$:

$$x \in E_\lambda \Rightarrow u(x) = \lambda x$$
$$\Rightarrow v(u(x)) = \lambda v(x) = u(v(x))$$
$$\Rightarrow v(x) \in E_\lambda$$

5- Propriétés du polynôme caractéristique

On a appelé polynôme caractéristique d'un endomorphisme u le polynôme $\chi_u(X) = \det(X\,\mathrm{Id} - u)$. Le polynôme caractéristique d'une matrice M élément de $\mathcal{M}_n(\mathbb{K})$ est $\chi_M(X) = \det(XI_n - M)$. Les racines de ces polynômes sont les valeurs propres de u ou de M. Deux matrices semblables ont évidemment mêmes valeurs propres, puisque le déterminant de deux matrices semblables est le même. Si M est semblable à N, on a :

$$\det(M - XI_n) = \det(P(N - XI_n)P^{-1}) = \det(P)\det(N - XI_n)\det(P^{-1}) = \det(N - XI_n)$$

Le déterminant d'une matrice étant égal à celui de sa transposée, on déduit également la proposition suivante :

PROPOSITION :

La transposée d'une matrice possède les mêmes valeurs propres que la matrice elle–même.

Il y a un rapport entre l'ordre de multiplicité des valeurs propres en tant que racine du polynôme caractéristique et la dimension du sous-espace propre correspondant :

PROPOSITION

L'ordre de multiplicité de la valeur propre λ est supérieur ou égal à la dimension q du sous–espace propre.

En effet, si on choisit une base de E en prenant une base du sous–espace propre, complétée en une base de E, la matrice de u possèdera q λ sur la diagonale. On aura, par blocs :

$$M = \begin{pmatrix} \lambda I_q & B \\ O & C \end{pmatrix}$$

Le polynôme caractéristique est alors égal à $(\lambda - X)^q \det(C - XI)$. Il se peut que λ soit également racine de $\det(C - XI)$, de sorte que son ordre de multiplicité est au moins q.

PROPOSITION

Si F *est un sous-espace vectoriel stable par u, alors le polynôme caractéristique de la restriction de u à* F *divise le polynôme caractéristique de u.*

En effet, si on prend une base de F que l'on complète en une base de E, la matrice de u dans cette base est triangulaire par blocs, de la forme $M = \begin{pmatrix} A & B \\ O & C \end{pmatrix}$. Le polynôme caractéristique de u est $\det(XI_n - M)$, celui de $u_{|F}$ est $\det(XI_q - A)$, où $q = \dim(F)$, et ce dernier divise le premier puisque :

$$\det(XI_n - M) = \det(XI_q - A)\det(XI_{n-q} - C)$$

Outre le fait d'être le polynôme dont les racines sont les valeurs propres de u, le polynôme caractérisque jouit d'une autre propriété. C'est un polynôme annulateur de u (Cayley-Hamilton, énoncé plus haut).

6- Conditions de diagonalisation

Nous étudions dans ce chapitre des conditions nécessaires ou suffisantes pour qu'un endomorphisme soit diagonalisable.

PROPOSITION
Soit u un endomorphisme de E. *Il y a équivalence entre :*

(i) *u est diagonalisable*

(ii) E *est la somme des sous-espaces propres*

(iii) *Le polynôme caractéristique est scindé, et l'ordre de multiplicité de chaque valeur propre est égale à la dimension du sous-espace propre.*

(iv) *Il existe un polynôme annulateur de u, scindé à racines simples.*

(v) *Le polynôme* $\prod_{\lambda \in \mathrm{Sp}(u)} (X - \lambda)$ *est un polynôme annulateur de u.*

Un polynôme est scindé s'il se factorise en facteurs du premier degré (éventuellement avec multiplicité des racines).

Démonstration :
Supposons u diagonalisable. Alors il existe une base de vecteurs propres. Si l'on regroupe côte à côte les vecteurs propres ayant même valeur propre, et si la valeur propre λ_i apparaît k_i fois, la matrice de u dans cette base est diagonale par blocs de la forme $\begin{pmatrix} \lambda_1 I_{k_1} & 0 & \dots & 0 \\ 0 & \lambda_2 I_{k_2} & \dots & 0 \\ \dots & & \dots & \\ 0 & 0 & \dots & \lambda_p I_{k_p} \end{pmatrix}$ où I_{k_i} est une matrice identité $k_i \times k_i$. On peut alors vérifier par le calcul les points suivants :

❑ La recherche du sous-espace propre $\mathrm{Ker}(u - \lambda_i \mathrm{Id})$ correspond précisément au sous-espace vectoriel engendré par les vecteurs de base de valeur propre λ_i. Cet espace est donc de dimension k_i, et comme $k_1 + k_2 + \dots + k_p$ est égal à l'ordre de la matrice et donc à la dimension de E, on a :

$$\dim(E) = \dim(E_{\lambda_1}) + \dim(E_{\lambda_2}) + \dots + \dim(E_{\lambda_p})$$

et donc $E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_p}$

❑ Le polynôme caractéristique vaut $\det(X\mathrm{Id} - u) = (X - \lambda_1)^{k_1}(X - \lambda_2)^{k_2}\dots(X - \lambda_p)^{k_p}$, donc le polynôme caractéristique est scindé et l'ordre de multiplicité k_i de la valeur propre λ_i est égal à la dimension du sous-espace propre.

❑ On vérifiera enfin que $\prod_{i=1}^{p} (X - \lambda_i)$ est un polynôme annulateur de l'endomorphisme. En effet, chaque matrice $M - \lambda_i I_n$ possède des 0 à la place des λ_i. Ce polynôme annulateur est scindé à racines simples.

On a ainsi montré que (i) $\Rightarrow$ (ii) et (iii) et (iv) et (v)

Réciproquement :

(ii) $\Rightarrow$ (i) Si E est somme des sous–espaces propres, il suffit de prendre dans chaque sous-espace propre une base pour obtenir ainsi une base de E.

(iii) $\Rightarrow$ (ii) Si n est la dimension de E et si le polynôme caractéristique s'écrit $(X - \lambda_1)^{k_1}...(X - \lambda_p)^{k_p}$, on a $n = k_1 + ... + k_p$ car le degré de $\det(X\mathrm{Id} - u)$ est égal au nombre de X sur la diagonale du déterminant et est donc égal à la dimension de l'espace E. Si on suppose en outre que $\dim E_{\lambda_i} = k_i$ alors $\dim(E) = \dim(E_{\lambda_1}) + \dim(E_{\lambda_2}) + ... + \dim(E_{\lambda_p})$ donc $E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus ... \oplus E_{\lambda_p}$ et on obtient (ii). Il y a donc équivalence entre (i), (ii) et (iii)

(v) $\Rightarrow$ (iv) est trivial puisque $\prod\limits_{\lambda \in \mathrm{Sp}(u)} (X - \lambda)$ est scindé à racines simples.

(iv) $\Rightarrow$ (v).
Soit $P = (X - \lambda_1)...(X - \lambda_p)$ un polynôme annulateur de u, à racines distinctes. On a donc :

$$0 = P(u) = (u - \lambda_1\mathrm{Id}) \circ (u - \lambda_2\mathrm{Id}) \circ ... \circ (u - \lambda_p\mathrm{Id})$$

Si l'un des λ_i n'est pas une valeur propre, alors $u - \lambda_i\mathrm{Id}$ est inversible, et en composant par son inverse, on obtient un autre polynôme annulateur sans le facteur $X - \lambda_i$. En procédant ainsi pour tout λ_i qui n'est pas une valeur propre, on se ramène à un polynôme annulateur dont toutes les racines sont des valeurs propres. Par ailleurs, toute valeur propre λ soit nécessairement figurer parmi les racines de P. En effet, si x est vecteur propre (non nul) de u pour la valeur propre λ, on a :

$$0 = P(u)(x) = P(\lambda)x$$

donc $\quad P(\lambda) = 0$
Il y a donc équivalence entre (iv) et (v)

v) $\Rightarrow$ ii). C'est l'implication la plus délicate. Nous donnons trois démonstrations possibles :

<u>Démonstration 1</u> :
Montrons d'abord le lemme suivant : soit A un polynôme de la forme $(X - \lambda)B(X)$ avec $B(\lambda) \neq 0$, alors $\mathrm{Ker}(A(u)) = \mathrm{Ker}(u - \lambda\mathrm{Id}) \oplus \mathrm{Ker}(B(u))$.
La somme est en effet directe : si x appartient à $\mathrm{Ker}(u - \lambda\mathrm{Id}) \cap \mathrm{Ker}(B(u))$ alors $u(x) = \lambda x$ et $B(u)(x) = B(\lambda)x$. Par ailleurs x appartient à $\mathrm{Ker}(B(u))$ donc $B(u)(x) = 0$ donc $B(\lambda)x = 0$, et puisque $B(\lambda) \neq 0$, $x = 0$.
La somme est égale à $\mathrm{Ker}(A(u))$. En effet, la division euclidienne de $B(X)$ par $X - \lambda$ donne $B(X) = (X - \lambda)Q(X) + B(\lambda)$, donc, $B(u) = (u - \lambda\mathrm{Id}) \circ Q(u) + B(\lambda)\mathrm{Id}$. Pour tout vecteur x de $\mathrm{Ker}(A(u))$, on a donc :

$$B(u)(x) = ((u - \lambda\mathrm{Id}) \circ Q(u))(x) + B(\lambda)x$$

$$\Rightarrow \qquad x = \frac{1}{B(\lambda)}\,(B(u)(x) - ((u - \lambda\mathrm{Id}) \circ Q(u))(x))$$

Or on a $A(u)(x) = 0 = ((u - \lambda\mathrm{Id}) \circ B(u))(x))$. Donc $B(u)(x)$ appartient à $\mathrm{Ker}(u - \lambda\mathrm{Id})$. De même $((u - \lambda\mathrm{Id}) \circ Q(u))(x)$ appartient à $\mathrm{Ker}(B(u))$ car $(B(u) \circ (u - \lambda\mathrm{Id}) \circ Q(u))(x) = (Q(u) \circ A(u))(x) = 0$. Donc $\mathrm{Ker}(A(u)) = \mathrm{Ker}(u - \lambda\mathrm{Id}) \oplus \mathrm{Ker}\, B(u)$.

Considérons maintenant un polynôme $P = (X - \lambda_1)...(X - \lambda_p)$, les λ_i étant distincts, et montrons par récurrence sur p que $\mathrm{Ker}(P(u)) = \bigoplus\limits_{i=i}^{p} \mathrm{Ker}(u - \lambda_i\mathrm{Id})$. Pour $p = 1$, $P(u) = u - \lambda_1\mathrm{Id}$ et l'égalité $\mathrm{Ker}(P(u)) = \mathrm{Ker}(u - \lambda_1\mathrm{Id})$ est triviale. Si la relation est vraie au rang $p - 1$, prenons $A = P$, $\lambda = \lambda_p$, $B = (X - \lambda_1)...(X - \lambda_{p-1})$ (avec donc $B(\lambda) \neq 0$) et appliquons le lemme précédent. On a :

$$\mathrm{Ker}(\mathrm{P}(u)) = \mathrm{Ker}(\mathrm{B}(u) \oplus \mathrm{Ker}(u - \lambda_p\mathrm{Id})$$

Il suffit d'appliquer l'hypothèse de récurrence sur B pour conclure.

Si P est de plus un polynôme annulateur de u, alors $\mathrm{Ker}(\mathrm{P}(u)) = \mathrm{E}$. Cela montre que, si $\prod_{\lambda \in \mathrm{Sp}(u)} (\mathrm{X} - \lambda)$ est un polynôme annulateur de u alors $\mathrm{E} = \bigoplus_{\lambda \in \mathrm{Sp}(u)} \mathrm{Ker}(u - \lambda\mathrm{Id})$, somme directe de sous-espaces propres.

Démonstration 2 :

On remarque que, u et v étant deux endomorphismes, $\dim \mathrm{Ker}(v \circ u) \leq \dim \mathrm{Ker}(u) + \dim \mathrm{Ker}(v)$. En effet, notons $\mathrm{F} = \mathrm{Ker}(v \circ u)$. Le théorème du rang donne :

$$\dim(\mathrm{F}) = \dim(\mathrm{Im}(u|_\mathrm{F})) + \dim(\mathrm{Ker}(u|_\mathrm{F}))$$

où $u|_\mathrm{F}$ désigne la restriction de u à F. Or $\mathrm{Im}(u|_\mathrm{F})$ est inclus dans $\mathrm{Ker}(v)$, et $\mathrm{Ker}(u|_\mathrm{F})$ est inclus dans $\mathrm{Ker}(u)$, d'où le résultat. Par récurrence, on a :

$$\begin{aligned}\dim(\mathrm{Ker}(f_1 \circ f_2 \circ \dots \circ f_p)) &\leq \dim(\mathrm{Ker}(f_1 \circ f_2 \circ \dots \circ f_{p-1})) + \dim(\mathrm{Ker}(f_p)) \\ &\leq \dim(\mathrm{Ker}(f_1)) + \dots + \dim(\mathrm{Ker}(f_p))\end{aligned}$$

Appliquons ce résultat à $(u - \lambda_1\mathrm{Id}) \circ \dots \circ (u - \lambda_p\mathrm{Id}) = 0$, où les λ_i sont les valeurs propres de u. On a alors :

$$\dim(\mathrm{Ker}((u - \lambda_1\mathrm{Id}) \circ \dots \circ (u - \lambda_p\mathrm{Id}))) \leq \dim(\mathrm{Ker}(u - \lambda_1\mathrm{Id})) + \dots + \dim(\mathrm{Ker}(u - \lambda_p\mathrm{Id}))$$

mais comme $(u - \lambda_1\mathrm{Id}) \circ \dots \circ (u - \lambda_p\mathrm{Id}) = 0$ par hypothèse, $\mathrm{Ker}((u - \lambda_1\mathrm{Id}) \circ \dots \circ (u - \lambda_p\mathrm{Id})) = \mathrm{E}$, et par ailleurs, les $\mathrm{Ker}(u - \lambda_i\mathrm{Id})$ sont les sous-espaces propres E_{λ_i}. On obtient donc :

$$\dim(\mathrm{E}) \leq \dim(\mathrm{E}_{\lambda_1}) + \dots + \dim(\mathrm{E}_{\lambda_p})$$

mais on sait par ailleurs que ces sous–espaces propres sont en somme directe, et donc leur somme est de dimension inférieure ou égale à $\dim(\mathrm{E})$. Les deux membres sont donc égaux, et u est diagonalisable.

Démonstration 3 :

Montrons que tout x se décompose en une somme $x = x_1 + \dots + x_p$, $x_i \in \mathrm{Ker}(u - \lambda_i\mathrm{Id})$, les λ_i étant les valeurs propres distinctes de u.

$$x = x_1 + \dots + x_p$$

Cherchons déjà quelle forme doivent nécessairement prendre les x_i, en supposant que la décomposition existe. Si on applique $(u - \lambda_1\mathrm{Id}) \circ \dots \circ (u - \lambda_{i-1}\mathrm{Id}) \circ (u - \lambda_{i+1}\mathrm{Id}) \circ \dots \circ (u - \lambda_n\mathrm{Id})$ aux deux membres, on obtient (revoir la démonstration 3 qui montre que les sous-espaces propres sont en somme directe) :

$$((u - \lambda_1\mathrm{Id}) \circ \dots \circ (u - \lambda_{i-1}\mathrm{Id}) \circ (u - \lambda_{i+1}\mathrm{Id}) \circ \dots \circ (u - \lambda_n\mathrm{Id}))(x) = (\lambda_i - \lambda_1)\dots(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})\dots(\lambda_i - \lambda_n)\, x_i$$

donc nécessairement :

$$x_i = \frac{(u - \lambda_1\mathrm{Id}) \circ \dots \circ (u - \lambda_{i-1}\mathrm{Id}) \circ (u - \lambda_{i+1}\mathrm{Id}) \circ \dots \circ (u - \lambda_p\mathrm{Id})}{(\lambda_i - \lambda_1)\dots(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})\dots(\lambda_i - \lambda_p)}(x)$$

Réciproquement, posons donc $\mathrm{L}_i(\mathrm{X}) = \dfrac{(\mathrm{X} - \lambda_1)\dots(\mathrm{X} - \lambda_{i-1})(\mathrm{X} - \lambda_{i+1})\dots(\mathrm{X} - \lambda_p)}{(\lambda_i - \lambda_1)\dots(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})\dots(\lambda_i - \lambda_p)}$, et prenons $x_i = \mathrm{L}_i(u)(x)$. On a bien $x_i \in \mathrm{Ker}(u - \lambda_i\mathrm{Id})$ car :

$$(u - \lambda_i\mathrm{Id})(x_i) = ((u - \lambda_i\mathrm{Id}) \circ \frac{(u - \lambda_1\mathrm{Id}) \circ \dots \circ (u - \lambda_{i-1}\mathrm{Id}) \circ (u - \lambda_{i+1}\mathrm{Id}) \circ \dots \circ (u - \lambda_p\mathrm{Id})}{(\lambda_i - \lambda_1)\dots(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})\dots(\lambda_i - \lambda_p)})(x_i)$$

et au numérateur, on reconnaît l'endomorphisme $(u - \lambda_1\mathrm{Id}) \circ (u - \lambda_2\mathrm{Id}) \circ \ldots \circ (u - \lambda_p\mathrm{Id})$ qui, par hypothèse, est nul.

Il reste à vérifier que, pour tout x, la somme des x_i vaut x, ou que $\sum_{i=1}^{p} \mathrm{L}_i(u)(x) = x$, ou que $\sum_{i=1}^{p} \mathrm{L}_i(u) = \mathrm{Id}$. Il suffit pour cela de montrer que $\sum_{i=1}^{p} \mathrm{L}_i(\mathrm{X}) = 1$. On remarque que $\mathrm{L}_i(\lambda_i) = 1$ et que $\mathrm{L}_i(\lambda_j) = 0$ si $j \neq i$. Le polynôme $\sum_{i=1}^{p} \mathrm{L}_i(\mathrm{X}) - 1$ admet donc les p racines distinctes $\lambda_1, \ldots, \lambda_p$ et est de degré au plus $p - 1$, donc il est nul. Donc $\sum_{i=1}^{p} \mathrm{L}_i(\mathrm{X}) = 1$. CQFD

La propriété (iv) est très puissante. Soit u un endomorphisme tel que $u^3 - 7u + 6\mathrm{Id} = 0$. u annule le polynôme $\mathrm{X}^3 - 7\mathrm{X} + 6$ qui se factorise sous la forme $(\mathrm{X} - 2)(\mathrm{X} - 1)(\mathrm{X} + 3)$. On sait donc que u est diagonalisable, l'ensemble des valeurs propres étant une partie de $\{-3, 1, 2\}$.

EXEMPLE : $\mathrm{M} = \begin{pmatrix} 8 & 0 & 9 \\ -3 & -1 & -3 \\ -6 & 0 & -7 \end{pmatrix}$. On pourra vérifier que M annule le polynôme $(\mathrm{X} - 2)(\mathrm{X} + 1)$. Donc M est diagonalisable.

On insistera sur le fait que la propriété (iv) exige que le polynôme soit ***scindé à racines simples***. Si le polynôme est seulement ***scindé***, l'endomorphisme n'est pas nécessairement diagonalisable. On peut cependant montrer qu'il est trigonalisable, i.e. il existe une base dans laquelle la matrice de u est triangulaire.

La négation de la propriété (iii) indique pour quelles raisons un endomorphisme n'est pas diagonalisable :

❑ Sur $\mathbb{R}$, il se peut que u ait des valeurs propres complexes et non réelles, autrement dit, le polynôme caractéristique n'est pas scindé. Par exemple, si $u^2 + \mathrm{Id} = 0$, alors les valeurs propres vérifient nécessairement $\lambda^2 + 1 = 0$, ce qui prouve qu'il n'y a pas de valeurs propres réelles. Si on se place sur $\mathbb{C}$, alors il y a deux valeurs propres potentielles : i et $-i$. u est d'ailleurs diagonalisable sur $\mathbb{C}$ puisqu'il y possède un polynôme annulateur scindé à racines simples. Ainsi, la possibilité de diagonalisation dépend du corps de base sur lequel on se place. Il est clair que le spectre de u dans $\mathbb{R}$, ensemble des solutions dans $\mathbb{R}$ du polynôme $\det(\mathrm{X}\,\mathrm{Id} - u) = 0$, est inclus dans le spectre de u dans $\mathbb{C}$, ensemble des solutions dans $\mathbb{C}$ du même polynôme.

EXEMPLE 1 : Diagonaliser $\mathrm{M} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, avec $\theta \neq k\pi$.

Le polynôme caractéristique vaut $\mathrm{X}^2 - 2\cos\theta\mathrm{X} + 1$. Il ne possède pas de racine sur $\mathbb{R}$. Il n'est donc pas diagonalisable sur $\mathbb{R}$. Si on se place sur $\mathbb{C}$, les racines sont $e^{\pm i\theta}$. Le sous-espace propre associé à la valeur propre $e^{i\theta}$ est la droite de vecteur directeur $\begin{pmatrix} 1 \\ -i \end{pmatrix}$. Le sous-espace propre associé à la valeur propre $e^{-i\theta}$ est la droite de vecteur directeur $\begin{pmatrix} 1 \\ i \end{pmatrix}$. M est diagonalisable sur $\mathbb{C}$, semblable à $\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$ et la matrice de passage est $\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$.

EXEMPLE 2 : Diagonaliser la matrice M = $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

Le polynôme caractéristique vaut det(M – XI) = $-X^3 + 3X^2 - 3X + 2 = -(X-2)(X^2 - X + 1)$

Si le corps de base est $\mathbb{R}$, alors seule 2 est valeur propre, avec seulement une droite de vecteurs propres. M n'est donc pas diagonalisable.

Si on se place sur $\mathbb{C}$, les valeurs propres sont 2, $\frac{1}{2} - \frac{i\sqrt{3}}{2}$ ou $\frac{1}{2} + \frac{i\sqrt{3}}{2}$ donnant chacune une droite comme sous-espace propre. Disposant de trois vecteurs propres, l'endomorphisme est diagonalisable sur $\mathbb{C}$.

❑ Supposons maintenant que le polynôme caractéristique soit scindé, et admette donc n valeurs propres comptées avec leur ordre de multiplicité, où n = dim(E). Cet ordre est supérieur ou égal à la dimension du sous-espace propre correspondant. Il est possible qu'il lui soit strictement supérieur, auquel cas u n'est pas diagonalisable. Ainsi, si λ est valeur propre double, on s'attend à trouver un plan de vecteurs propres. Si on ne trouve qu'une droite, il manquera au moins un vecteur pour former une base de vecteurs propres.

EXEMPLE : Diagonaliser la matrice M = $\begin{pmatrix} 7 & -11 & -2 \\ -3 & -1 & -6 \\ -1 & 1 & 6 \end{pmatrix}$

Le polynôme caractéristique vaut : det(M – XI_3) = $-X^3 + 12X^2 - 256 = -(X+4)(X-8)^2$

- Sous–espace propre associé à la valeur propre 8. On résout le système :

$$(M - 8I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \Leftrightarrow \begin{cases} -x - 11y - 2z = 0 \\ -3x - 9y - 6z = 0 \\ -x + y - 2z = 0 \end{cases} \Leftrightarrow \begin{cases} x + 11y + 2z = 0 \\ y = 0 \\ x + 2z = 0 \end{cases} \Leftrightarrow \begin{cases} y = 0 \\ x + 2z = 0 \end{cases}$$

Il s'agit de la droite vectorielle engendrée par $\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$. La valeur propre étant double, il manquera un vecteur.

- Sous–espace propre associé à la valeur propre –4. On résout le système :

$$(M + 4I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \Leftrightarrow \begin{cases} 11x - 11y - 2z = 0 \\ -3x + 3y - 6z = 0 \\ -x + y + 10z = 0 \end{cases} \Leftrightarrow \begin{cases} x = y \\ z = 0 \end{cases}$$

Il s'agit de la droite vectorielle engendrée par $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

La somme des sous–espaces propres étant de dimension 2, la matrice n'est pas diagonalisable.

Voici, pour terminer ce paragraphe des conditions ***<u>suffisantes</u>*** (et non plus nécessaires et suffisantes) pour qu'un endomorphisme soit diagonalisable.

PROPOSITION :
Soit u un endomorphisme d'un espace vectoriel E de dimension n. Si u admet n valeurs propres distinctes, alors u est diagonalisable.

<u>Démonstration</u> :
Il existe n sous–espaces propres, tous de dimension supérieure ou égale à 1. La dimension de la somme est donc supérieure ou égale à n. Comme elle ne peut excéder celle de E, la dimension de la somme est exactement égale à n, chaque sous–espace étant de dimension 1. La somme est donc bien égale à E et u est diagonalisable.

La formulation de la proposition précédente est évidemment équivalente à la suivante :
PROPOSITION :
Si le polynôme caractéristique de u est scindé et ne possède que des racines simples, alors u est diagonalisable.

On a enfin :
PROPOSITION
Soit u diagonalisable et F *un sous-espace vectoriel de* E *stable par u. Alors la restriction de u à* F *est diagonalisable.*

<u>Démonstration</u> :
u annule un polynôme scindé dont toutes les racines sont simples. Il en est a fortiori de même de $u|_F$ qui annule le même polynôme.

7- Condition de trigonalisation

PROPOSITION :
Un endomorphisme est trigonalisable si et seulement si son polynôme caractéristique est scindé.

<u>Démonstration</u> :
Si u est trigonalisable, il existe une base dans laquelle la matrice de u est triangulaire. Si sa diagonale porte les coefficients λ_1, ..., λ_n, alors $\chi_u(X) = (X - \lambda_1)(X - \lambda_2)...(X - \lambda_n)$ qui est scindé (éventuellement à racines multiples si plusieurs λ_i coïncident).

Montrons la réciproque par récurrence sur la dimension de l'espace E. Si dim(E) = 1, il n'y a rien à montrer. Supposons que la proposition est montrée pour les espaces vectoriels de dimension $n - 1$. Soit E de dimension n et u un endomorphisme dont le polynôme caractéristique est scindé. Il possède donc au moins une valeur propre λ_1. Soit e_1 un vecteur propre associé. Complétons e_1 en une base $(e_1, \varepsilon_2, ..., \varepsilon_n)$. Soit $F = \text{Vect}(\varepsilon_2, ..., \varepsilon_n)$ et p le projecteur sur F parallèlement à la droite $\text{Vect}(e_1)$. Soit enfin $v = p \circ u$. La matrice de u dans la base $(e_1, \varepsilon_2, ..., \varepsilon_n)$ est de la forme :

$$M = \begin{pmatrix} \lambda_1 & B \\ 0 & C \end{pmatrix}$$

où B est une ligne de $n - 1$ coefficients, C n'est autre que la matrice de v dans la base $(\varepsilon_2, ..., \varepsilon_n)$, à gauche de laquelle se trouve une colonne de 0. On a donc :

$$\chi_u(X) = (X - \lambda_1)\chi_v(X)$$

et comme χ_u est scindé, χ_v l'est également. D'après l'hypothèse de récurrence, v est trigonalisable et il existe une base $(e_2, ..., e_n)$ dans laquelle la matrice T de v est triangulaire. Pour tout $i \geq 2$, on décompose ensuite $u(e_i)$ sur $\mathrm{Vect}(e_1) \oplus \mathrm{F}$:

$$\exists\, \alpha_i,\ u(e_i) = \alpha_i e_1 + p \circ u(e_i) = \alpha_i e_1 + v(e_i)$$

La matrice de u dans la base $(e_1, ..., e_n)$ est alors $\begin{pmatrix} \lambda_1 & \alpha \\ 0 & \mathrm{T} \end{pmatrix}$ où α est la ligne $(\alpha_2 \ ... \ \alpha_n)$, et cette matrice est bien triangulaire.

En annexe II, on montre un résultat plus général : s'il existe un polynôme annulateur scindé de u, alors u est trigonalisable.

CONSEQUENCES :

❑ Sur $\mathbb{C}$, tout polynôme est scindé, donc tout endomorphisme sur un espace vectoriel complexe est trigonalisable.

❑ Toute matrice M de $\mathcal{M}_n(\mathbb{C})$ est trigonalisable dans $\mathbb{C}$. Comme le polynôme caractéristique ou la trace sont invariant par changement de base, en raisonnant sur la matrice trigonalisée T, on voit que :

$$\mathrm{Tr(M)} = \mathrm{Tr(T)} = \sum_{i=1}^{n} \lambda_i \qquad \text{où les } \lambda_i \text{ sont les valeurs propres complexes M.}$$

$$\det(\mathrm{M}) = \det(\mathrm{T}) = \prod_{i=1}^{n} \lambda_i$$

$$\chi_{\mathrm{M}}(\mathrm{X}) = \chi_{\mathrm{M}}(\mathrm{T}) = \prod_{i=1}^{n} (\mathrm{X} - \lambda_i) = \mathrm{X}^n - \mathrm{Tr(M)}\mathrm{X}^{n-1} + ... + (-1)^n \det(\mathrm{M}).$$

$$\deg(\chi_{\mathrm{M}}) = n$$

En effet, les coefficients de la diagonale de la matrice trigonalisée T sont précisément les valeurs propres. Dans le cas où M est réelle, les formules précédentes sont valides à condition de calculer ses valeurs propres dans $\mathbb{C}$. Plus généralement :

$$\mathrm{Tr(M^2)} = \mathrm{Tr(T^2)} = \sum_{i=1}^{n} \lambda_i^2 \qquad \text{car la matrice } \mathrm{T}^2 \text{ a pour coefficients diagonaux les } \lambda_i^2$$

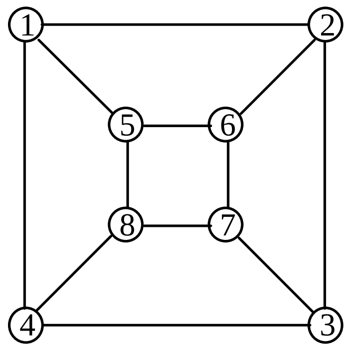
$$\mathrm{Tr(M^3)} = \sum_{i=1}^{n} \lambda_i^3 \qquad \text{etc...}$$

Annexe I : une application du théorème du rang en S.I. ou en physique

a) Cycles indépendants, nombre cyclomatique

En Sciences Industrielles, un mécanisme est modélisé par des solides indéformables S_1, S_2, ..., S_n ayant entre eux des liaisons, traduisant les mouvements possibles entre les pièces. Notons L_{ij} une liaison entre le solide S_i et le solide S_j. Il n'existe bien évidemment pas des liaisons entre tous les solides du mécanisme, de sorte que L_{ij} n'est pas nécessairement défini pour tout i et j. On a par ailleurs $\mathrm{L}_{ij} = \mathrm{L}_{ji}$ lorsqu'une telle liaison existe.

Le mécanisme est symbolisé par un graphe de structure, ayant des sommets et des arêtes. Il y a autant de sommets que de solides (soit s ce nombre), et une arête relie le sommet S_i et S_j si et seulement si une liaison L_{ij} existe. Autrement dit, les arêtes représentent les liaisons entre solides. Il y a donc autant d'arêtes que de liaisons (soit a ce nombre). Voici un exemple fictif de graphe de structure, où les sommets (ou solides) sont numérotés de 1 à 8 :



Nous supposerons toujours un graphe de structure connexe, c'est-à-dire qu'on peut se rendre d'un sommet quelconque à un autre sommet en suivant un chemin du graphe (sinon, cela voudrait dire que l'on dispose de deux mécanismes disjoints !!).

De tels graphes de structures se trouvent également en électrocinétique, comme modélisation de circuits électrique, où les arêtes sont plutôt appelées branches du circuit et les sommets sont appelés noeuds.

Un cycle est un chemin fermé, i.e. commençant et finissant au même sommet. Nous supposerons en outre qu'un tel chemin ne passe pas deux fois par la même arête. Par ailleurs, nous ne définissons pas de sens de parcours privilégié d'un circuit : il peut se parcourir dans un sens ou dans l'autre.

En S.I., un tel cycle est parfois appelé boucle de solides, et en électrocinétique, un cycle est appelé maille du circuit. Voici des exemples de cycles définis dans le graphe de structure précédent : 1-4-8-5-1 ou 1-4-8-7-3-2-1. On peut définir la notion de cycles indépendants, et de nombre cyclomatique. Mathématiquement, cette notion d'indépendance est rigoureusement identique à la notion de système de vecteurs indépendants dans un espace vectoriel. Précisons cela.

Le corps de base utilisé sera inhabituel. Il s'agit du corps constitué des deux éléments 0 et 1, avec comme règle de la somme 1 + 1 = 0. Ce corps est noté $\mathbb{Z}/2\mathbb{Z}$. On peut le voir aussi comme le corps des deux éléments {pair, impair} avec les règles d'opérations usuelles sur les nombres pairs et impairs. La caractéristique essentielle de ce corps est que :

$$\forall x \in \mathbb{Z}/2\mathbb{Z}, x + x = 0$$

Considérons maintenant F espace vectoriel de dimension s (le nombre de sommets) sur $\mathbb{Z}/2\mathbb{Z}$, dont une base sera notée $(S_1, ..., S_s)$. Autrement dit, les sommets désignent les vecteurs d'une base de F. Un vecteur quelconque de F s'écrit $S_{i_1} + S_{i_2} + ... + S_{i_k}$ où les indices $i_1, i_2, ..., i_k$ sont distincts (car s'il y a deux indices identiques, on a $S_i + S_i = (1 + 1)\ S_i = 0\ S_i = 0$).

Considérons également E espace vectoriel de dimension a (le nombre d'arêtes) sur $\mathbb{Z}/2\mathbb{Z}$, dont une base sont les vecteurs L_{ij}.Un vecteur quelconque de E s'écrit $\sum L_{ij}$ où les indices sont également distincts.

Définissons enfin une application linéaire δ de E dans F, appelée *bord* en donnant les images des vecteurs de base de E. Si L_{ij} est une liaison entre S_i et S_j alors $\delta(L_{ij}) = S_i + S_j$. Par exemple, dans le graphe de structure dessiné ci-dessus :

$$\delta(L_{56}) = S_5 + S_6$$

$$\delta(L_{56} + L_{67} + L_{73}) = \delta(L_{56}) + \delta(L_{67}) + \delta(L_{73}) = S_5 + S_6 + S_6 + S_7 + S_7 + S_3$$

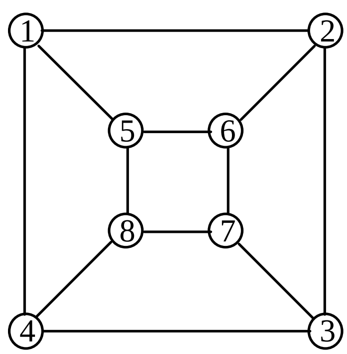
$$= S_5 + S_3 \quad \text{puisque } S_6 + S_6 = 0 = S_7 + S_7$$

On comprend pourquoi δ s'appelle *bord*. On a calculé le bord (i.e. les extrémités) du chemin 5-6-7-3.

Quelle est l'image de δ ? Un sommet isolé ne peut être dans cette image, car $\delta(L_{ij})$ fait intervenir deux sommets et il en résulte que $\delta(\sum L_{ij})$ est combinaison linéaire d'un nombre pair de sommets (même après simplification). L'application δ ne peut être surjective. Montrons que son rang est s–1 (un de moins que le nombre de sommets). Il suffit de montrer que, par exemple, le système libre $(S_1 + S_2, S_1 + S_3, ..., S_1 + S_s)$ est constitué de vecteurs de Im(δ). Or nous avons supposé le graphe connexe, ce qui signifie que, pour tout i, il existe un chemin allant de S_1 à S_i. Cela ne signifie rien d'autre que le bord de ce chemin est $S_1 + S_i$ qui est donc bien dans Im(δ).

Quel est le noyau de δ ? Les éléments du noyau ne sont autres que les cycles. En effet, un chemin est dans le noyau si et seulement si son bord est nul, et donc si et seulement si le chemin commence et finit au même sommet : il s'agit bien des cycles. C'est le cas par exemple de $L_{14} + L_{48} + L_{85} + L_{51}$, dont l'image par δ est nul. Les cycles dits indépendants sont précisément ceux qui sont linéairement indépendants au sens mathématique du terme dans l'espace vectoriel E. Le nombre de cycles indépendants est la dimension de Ker(δ). Or le théorème du rang nous donne la dimension de Ker(δ). C'est $\mu = a - \text{rg}(\delta) = a - (s - 1) = a - s + 1$. Ce nombre est dit nombre cyclomatique du graphe. C'est le nombre maximal de cycles indépendants.
Dans l'exemple précédent, $\mu = 12 - 8 + 1 = 5$. Il y a cinq cycles indépendants, par exemple :

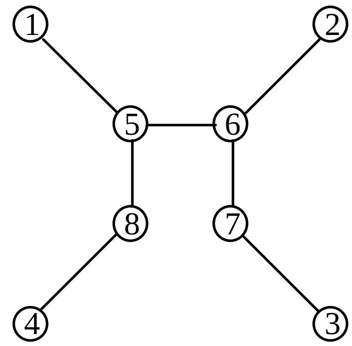
$$C_1 = L_{15} + L_{58} + L_{84} + L_{41}$$
$$C_2 = L_{12} + L_{26} + L_{65} + L_{51}$$
$$C_3 = L_{26} + L_{67} + L_{73} + L_{32}$$
$$C_4 = L_{37} + L_{78} + L_{84} + L_{43}$$
$$C_5 = L_{14} + L_{43} + L_{32} + L_{21}$$



Tout autre cycle est combinaison linéaire de ces cycles, par exemple :

$$L_{56} + L_{67} + L_{78} + L_{85} = C_1 + C_2 + C_3 + C_4 + C_5$$
$$L_{15} + L_{56} + L_{67} + L_{73} + L_{32} + L_{21} = C_2 + C_3$$

On peut montrer que le nombre cyclomatique s'interprête aussi comme le nombre minimal d'arêtes à supprimer pour qu'il n'y ait plus de cycles. Dans l'exemple précédent, où $\mu = 5$, il suffit d'enlever 5 arêtes bien choisies pour ne plus avoir de cycles, par exemple les arêtes L_{12}, L_{23}, L_{34}, L_{45} et L_{87}.



Le graphe final obtenu s'appelle un arbre.

<u>b) Utilisation du nombre cyclomatique</u>

En S.I., le nombre cyclomatique intervient dans l'analyse géométrique ou cinématique du mécanisme. Chaque liaison fait intervenir un ou plusieurs paramètres géométriques (respectivement cinématique), par exemple l'angle formé entre deux solides dans le cas d'une liaison pivot, ou la distance entre deux solides dans le cas d'une liaison glissière (respectivement la vitesse angulaire ou la vitesse linéaire). Dans le cas d'un cycle (ou boucle de solides), il existe nécessairement une relation entre les différentes paramètres de chaque liaison pour pouvoir refermer la boucle. Il convient donc d'écrire toutes les relations entre paramètres afin d'éliminer les paramètres superflus. Pour cela, il est inutile d'établir une telle relation pour chaque cycle du graphe : il suffit de le faire pour les cycles indépendants. Le nombre cyclomatique indique combien il y a de tels cycles indépendants et donc combien il y a de relations à établir entre paramètres.

En électrocinétique, on dispose d'un circuit constitué de a branches (ou arêtes) et de s noeuds (ou sommets), chaque branche étant constituée de générateurs et de résistances. Le problème est d'établir les valeurs de l'intensité du courant dans chaque branche, ou la différence de potentiel entre deux noeuds.

❑ La méthode brute consiste à dire que l'on dispose de $2a$ inconnues élémentaires, à savoir les a intensités dans chacune des branches et les a différences de potentiels aux bornes de chacune des branches. Pour résoudre le système, il convient donc d'établir $2a$ équations indépendantes qui sont :

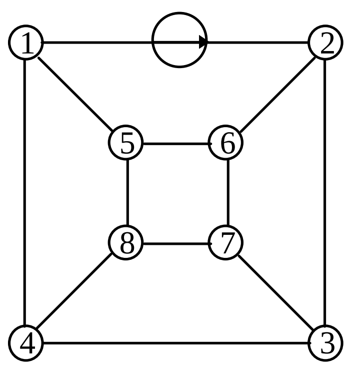
- Les a relations entre tensions et intensité dans chaque branche résultant de la loi d'Ohm (du type U = RI – E, avec U tension aux bornes de la branche, I intensité du courant dans la branche, E force électromotrice dans la branche)
- Les s–1 relations résultant de la loi des noeuds $\sum_j I_{ij} = 0$ pour chaque noeud i du circuit sauf un, où I_{ij} désigne l'intensité du courant allant de i vers j. On notera qu'il n'y a que s–1 équations car les s équations $\sum_j I_{ij} = 0$ écrites pour chaque sommet i sont liées entre elles. La somme de ces s équations est en effet nulle : $\sum_i \sum_j I_{ij} = \sum_{i<j} I_{ij} + I_{ji} = 0$ car $I_{ij} + I_{ji} = 0$. Il suffit donc d'écrire s–1 lois des noeuds pour que la dernière soit automatiquement vérifiée.
- Il reste encore $a - s + 1$ équations à écrire. On reconnaît là le nombre cyclomatique. Les dernières relations consistent en la loi des mailles appliquée sur chaque cycle indépendant : $\sum U = 0$, où l'on

fait la somme des tensions de chaque branche constituant la maille. La somme est nulle puisqu'on revient au point de départ. Toute autre loi des mailles appliquée à un autre cycle se déduira par combinaison linéaire de la loi des mailles appliquée aux cycles indépendants. Il est donc inutile de l'écrire.

❑ La méthode des tensions indépendantes consiste à partir des $s - 1$ tensions inconnues s'appliquant à chacune des branches de l'arbre qu'on obtiendrait à partir du circuit initial si on enlevait $a - s + 1$ branches bien choisies. Ces $s - 1$ tensions sont appelées les tensions indépendantes du circuit. Les tensions s'appliquant sur les $a - s + 1$ autres branches s'en déduisent en appliquant la loi des mailles $\sum U = 0$ à chacune des $a - s + 1$ maille indépendante. Pour chaque branche on en déduit l'intensité du courant qui la parcourt, exprimée en fonction des tensions indépendantes. Il suffit d'écrire ensuite les $s - 1$ relations données par la loi des noeuds pour avoir à résoudre un système de $s - 1$ équations indépendantes (les lois des noeuds) à $s - 1$ inconnues (les tensions indépendantes).

❑ La méthode des intensités indépendantes consiste à attribuer à chacune des $a - s + 1$ mailles indépendantes une intensité fictive. Ces $a - s + 1$ intensités sont appelées intensités indépendantes. Le courant réel parcourant une branche est la somme algébrique des intensités indépendantes des mailles auxquelles appartient la branche. On en déduit l'expression de la tension appliquée aux bornes de chaque branche en fonction des intensités indépentantes. La loi des noeuds est automatiquement vérifiée. Il suffit d'écrire ensuite la loi des mailles $\sum U = 0$ pour chaque maille indépendante pour avoir à résoudre un système de $a - s + 1$ équations (la loi des mailles) à $a - s + 1$ inconnues (les intensités indépendantes).

EXEMPLE :
Considérons un circuit électrique en forme d'armature cubique, chaque arête du cube étant une branche de résistance R. On place dans l'une des branches un générateur de tension E. On demande l'intensité du courant dans chacune des branches.



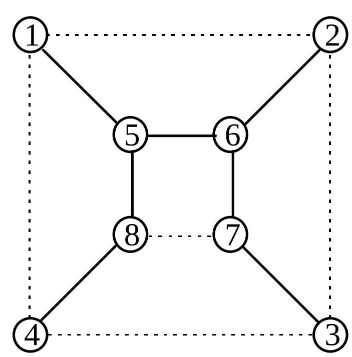
La première méthode de résolution donne le système suivant de 24 équations à 24 inconnues (on note I_{ij} l'intensité de courant du sommet *i* vers le sommet *j*, et U_{ij} la différence de potentiel entre le sommet *i* et le sommet *j*) :

$$\begin{cases} U_{12} = RI_{12} - E \\ U_{15} = RI_{15} \\ U_{14} = RI_{14} \\ U_{23} = RI_{23} \\ U_{26} = RI_{26} \\ U_{34} = RI_{34} \\ U_{37} = RI_{37} \\ U_{48} = RI_{48} \\ U_{56} = RI_{56} \\ U_{58} = RI_{58} \\ U_{67} = RI_{67} \\ U_{78} = RI_{78} \\ U_{12} + U_{26} - U_{56} - U_{15} = 0 \\ U_{23} + U_{37} - U_{67} - U_{26} = 0 \\ U_{34} + U_{48} - U_{78} - U_{37} = 0 \\ U_{15} + U_{58} - U_{48} - U_{14} = 0 \\ U_{12} + U_{23} + U_{34} - U_{14} = 0 \\ I_{12} + I_{15} + I_{14} = 0 \\ -I_{12} + I_{23} + I_{26} = 0 \\ -I_{23} + I_{34} + I_{37} = 0 \\ -I_{14} - I_{34} + I_{48} = 0 \\ -I_{15} + I_{56} + I_{58} = 0 \\ -I_{26} - I_{56} + I_{67} = 0 \\ -I_{37} - I_{67} + I_{78} = 0 \end{cases}$$

Les 12 premières équations expriment la loi d'Ohm, les 5 suivantes la loi des mailles, les 7 dernières la loi des noeuds. Les courageux trouveront que :

$$I_{15} = I_{14} = -I_{23} = -I_{26} = -\frac{5E}{24R},\ I_{67} = I_{37} = -I_{48} = -I_{58} = \frac{E}{24R},\ I_{34} = -I_{56} = \frac{E}{6R},\ I_{12} = \frac{5E}{12R},\ I_{78} = \frac{E}{12R}$$

La deuxième méthode consiste à prendre comme inconnues les tensions indépendantes U_{15}, U_{56}, U_{26}, U_{67}, U_{37}, U_{58} et U_{48}, correspondant à l'arbre :



Les sept équations à écrire résultent des sept lois des noeuds, exprimées en fonction des tensions indépendantes :

$I_{12} + I_{15} + I_{14} = 0 \quad \Leftrightarrow \quad E + U_{15} + U_{56} - U_{26} + 2U_{15} + U_{58} - U_{48} = 0$

$-I_{12} + I_{23} + I_{26} = 0 \quad \Leftrightarrow \quad -E - U_{15} - U_{56} + 3U_{26} + U_{67} - U_{37} = 0$

$-I_{23} + I_{34} + I_{37} = 0 \quad \Leftrightarrow \quad -U_{26} - 2U_{67} + 3U_{37} - U_{56} + U_{58} - U_{48} = 0$

$-I_{14} - I_{34} + I_{48} = 0 \quad \Leftrightarrow \quad -U_{15} - 2U_{58} + 3U_{48} - U_{37} + U_{67} + U_{56} = 0$

$-I_{15} + I_{56} + I_{58} = 0 \quad \Leftrightarrow \quad -U_{15} + U_{56} + U_{58} = 0$

$-I_{26} - I_{56} + I_{67} = 0 \quad \Leftrightarrow \quad -U_{26} - U_{56} + U_{67} = 0$

$-I_{37} - I_{67} + I_{78} = 0 \quad \Leftrightarrow \quad -U_{37} - 2U_{67} - U_{56} + U_{58} = 0$

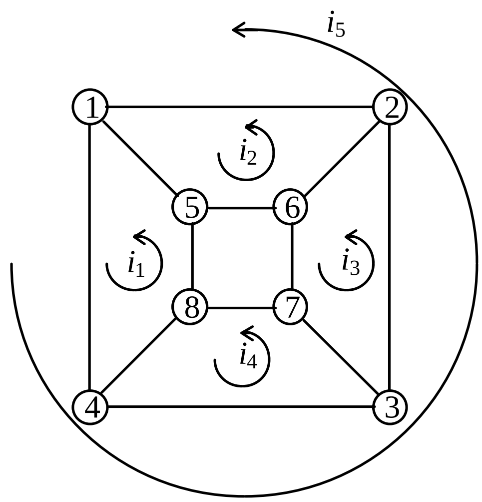
Le système conduit à la solution :

$$U_{15} = -U_{26} = -\frac{5E}{24},\ U_{58} = -U_{37} = U_{48} = -U_{67} = -\frac{E}{24},\ U_{56} = -\frac{E}{6}$$

à partir de laquelle on peut déduire les autres tensions, puis les intensités des courants.

La troisième méthode consiste à attribuer un courant de maille inconnu à chacune des cinq mailles indépendantes que nous avons notées C_1, C_2, C_3, C_4 et C_5 dans le a), à savoir :

$C_1 = L_{15} + L_{58} + L_{84} + L_{41}$
$C_2 = L_{12} + L_{26} + L_{65} + L_{51}$
$C_3 = L_{26} + L_{67} + L_{73} + L_{32}$
$C_4 = L_{37} + L_{78} + L_{84} + L_{43}$
$C_5 = L_{14} + L_{43} + L_{32} + L_{21}$



Notons ces cinq courants de maille i_1, i_2, i_3, i_4 et i_5, comptés positivement dans le sens trigonométrique. Le courant réel I_{12} par exemple est égal à $-i_2 - i_5$. Le courant I_{37} est égal à $i_4 - i_3$, etc... Les cinq équations à écrire sont les lois des mailles le long précisément des cinq cycles indépendants choisis, exprimés en fonction des courants de maille :

$U_{12} + U_{26} - U_{56} - U_{15} = 0 \quad \Leftrightarrow \quad -4i_2 - i_5 - \frac{E}{R} + i_3 + i_1 = 0$

$U_{23} + U_{37} - U_{67} - U_{26} = 0 \quad \Leftrightarrow \quad -4i_3 - i_5 + i_4 + i_2 = 0$

$U_{34} + U_{48} - U_{78} - U_{37} = 0 \quad \Leftrightarrow \quad -4i_4 - i_5 + i_1 + i_3 = 0$

$U_{15} + U_{58} - U_{48} - U_{14} = 0 \quad \Leftrightarrow \quad -4i_1 + i_2 + i_4 - i_5 = 0$

$U_{12} + U_{23} + U_{34} - U_{14} = 0 \quad \Leftrightarrow \quad -i_2 - 4i_5 - \frac{E}{R} - i_3 - i_4 - i_1 = 0$

Le système conduit à la solution :

$$i_1 = \frac{E}{24R},\ i_2 = -\frac{E}{6R},\ i_3 = \frac{E}{24R},\ i_4 = \frac{E}{12R},\ i_5 = -\frac{E}{4R}$$

On en déduit alors toutes les autres intensités, par exemple :

$I_{12} = -i_2 - i_5 = \frac{5E}{12R}$

$I_{37} = i_4 - i_3 = \frac{E}{24R}$

Annexe II : Trigonalisation des matrices

*Début de partie réservée aux MP/MP**

a) Polynômes premiers entre eux et identité de Bezout

Soient P et Q deux polynômes et considérons $\mathcal{I} = \{AP + BQ \mid A \in \mathbb{K}[X], B \in \mathbb{K}[X]\}$. Il n'est pas difficile de vérifier qu'il s'agit d'un idéal. $\mathcal{I}$ s'appelle idéal engendré par P et Q. Mais étant un idéal, il est en fait engendré par un seul polynôme D, élément de $\mathcal{I}$ de degré minimal. Ainsi :

$$\mathcal{I} = \{AP + BQ \mid A \in \mathbb{K}[X], B \in \mathbb{K}[X]\} = \{QD \mid Q \in \mathbb{K}[X]\}$$

Cherchons le rapport entre P, Q et D ?

- D étant élément de $\mathcal{I}$ est lui-même de la forme D = AP + BQ.
- Inversement, P et Q étant eux-mêmes des éléments de $\mathcal{I}$ sont des multiples de D. D est un polynôme diviseur commun de P et Q.
- Montrons qu'il s'agit du plus grand diviseur commun. Soit U un autre polynôme diviseur commun de P et Q. Cela signifie que P et Q vont se factoriser par U, $P = UP_0$ et $Q = UQ_0$ et qu'il en est de même de $D = AP + BQ = (AP_0 + BQ_0)U$. Ainsi, tout diviseur commun de P et Q divise D. D est le PGCD de P et Q.

EXEMPLE : $P = (X - 2)^2(X - 1)$ et $Q = (X + 3)(X - 2)(X - 1)$. Le PGCD de P et Q est $D = (X - 2)(X - 1)$. Il existe A et B tel que :

$$D = AP + BQ$$
$$\Leftrightarrow \quad (X - 2)(X - 1) = (X - 2)^2(X - 1)A + (X + 3)(X - 2)(X - 1)B$$
$$\Leftrightarrow \quad 1 = (X - 2)A + (X + 3)B$$

Il suffit de prendre $A = -\dfrac{1}{5}$ et $B = \dfrac{1}{5}$

Deux polynômes P et Q sont premiers entre eux si leur PGCD D vaut 1. On note $P \wedge Q = 1$. Il existe alors A et B tels que AP + BQ = 1 (identité de Bézout). Inversement, deux polynômes P et Q vérifiant cette relation sont premiers entre eux, puisque leur PGCD D devra diviser 1 = AP + BQ donc être égal à 1.

Voici un exemple d'utilisation de l'identité de Bezout, le théorème de Gauss :

$$\text{Si } R \text{ divise } PQ \text{ et si } R \wedge P = 1 \text{ alors } R \mid Q$$

En effet, il existe A et B tels que $1 = AR + BP \Rightarrow Q = RAQ + BPQ$ mais PQ se factorise par R, par exemple PQ = RS. Donc Q = R(AQ + BS) et R divise Q.

b) Théorème de décomposition des noyaux

Revenons à nos sous-espaces stables. Nous avons vu que, pour tout polynôme P, $\mathrm{Ker}(P(u))$ est stable par u. Si P se factorise sous la forme $P = P_1P_2$ avec P_1 et P_2 premiers entre eux, nous allons montrer que $\mathrm{Ker}(P(u)) = \mathrm{Ker}(P_1(u)) \oplus \mathrm{Ker}(P_2(u))$. On utilise pour cela l'identité de Bézout. Il existe A et B tels que $1 = AP_1 + BP_2$. Il est aisé de vérifier que $\mathrm{Ker}(P_1(u))$ et $\mathrm{Ker}(P_2(u))$ sont inclus dans $\mathrm{Ker}(P(u))$. De plus :

- La somme est directe :

Si x est élément de $\mathrm{Ker}(P_1(u)) \cap \mathrm{Ker}(P_2(u))$. L'identité de Bézout, appliquée à u donne :

$$\mathrm{Id} = (AP_1)(u) + (BP_2)(u) = A(u) \circ P_1(u) + B(u) \circ P_2(u)$$

Puis appliquée à x, on obtient :

$$x = A(u)(P_1(u)(x)) + B(u)(P_2(u)(x)) = 0 \text{ car } P_1(u)(x) = P_2(u)(x) = 0$$

- La somme vaut $\mathrm{Ker}(P(u))$.

Soit x élément de Ker(P(u)). L'égalité Id = A(u) o $P_1(u)$ + B(u) o $P_2(u)$ donne cette fois :

$$x = A(u)(P_1(u)(x)) + B(u)(P_2(u)(x))$$

avec $A(u)(P_1(u)(x))$ élément de $\mathrm{Ker}(P_2(u))$ et $B(u)(P_2(u)(x))$ élément de $\mathrm{Ker}(P_1(u))$. En effet :

$$P_2(u)(A(u)(P_1(u)(x))) = A(u)((P_1P_2)(u)(x)) = A(u)(P(u)(x)) = 0 \text{ car } P(u)(x) = 0$$

et de même pour $B(u)(P_2(u)(x))$

Ce théorème est particulièrement intéressant lorsqu'on l'applique à un polynôme annulateur de u puisque Ker(P(u)) donne alors E tout entier.

EXEMPLE 1 : M = $\begin{pmatrix} 5 & -1 \\ 3 & 1 \end{pmatrix}$ associé à u dans la base canonique de $\mathbb{K}^2$.

Soit P = $X^2 - 6X + 8 = (X - 2)(X - 4)$.

Ker(P(u)) = E

Ker(u – 2Id) est la droite engendrée par $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$

Ker(u – 4Id) est la droite engendrée par $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Les deux sous-espaces sont en somme directe. Dans la base $\begin{pmatrix} 1 \\ 3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, la matrice de u est $\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$.

EXEMPLE 2 : M = $\begin{pmatrix} -7 & 3 & -2 \\ -13 & 6 & -3 \\ 14 & -5 & 5 \end{pmatrix}$ et P = $X^3 - 4X^2 + 5X - 2 = (X - 2)(X - 1)^2$.

Ker(P(u)) = E

Ker(u – 2Id) est l'ensemble des $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ vérifiant :

$$\begin{cases} -9x + 3y - 2z = 0 \\ -13x + 4y - 3z = 0 \\ 14x - 5y + 3z = 0 \end{cases} \Leftrightarrow \begin{cases} z = -3x \\ x = y \end{cases}$$

Il s'agit de la droite engendrée par $\begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix}$. On a $M\begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix}$

Ker((u – Id)^2) est le plan d'équation $-3x + y - z = 0$, de base $\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. On a :

$$M\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$M\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$$

La matrice de u dans la base $\begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ est $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 0 \end{pmatrix}$.

Le théorème de décomposition s'étend de proche en proche au produit P = $P_1P_2...P_k$ de polynômes premiers entre eux deux à deux. En effet :

$$\mathrm{Ker}(P(u)) = \mathrm{Ker}(P_1(u)) \oplus \mathrm{Ker}(P_2...P_k(u))$$

$= \mathrm{Ker}(P_1(u)) \oplus \mathrm{Ker}(P_2(u)) \oplus \mathrm{Ker}(P_3 \dots P_k(u))$
$= \dots$
$= \mathrm{Ker}(P_1(u)) \oplus \mathrm{Ker}(P_2(u)) \oplus \dots \oplus \mathrm{Ker}(P_k(u))$

Le raisonnement repose sur le fait que $P_1 \wedge P_2 \dots P_k = 1$. Cela résulte du fait que les identités de Bezout donnent respectivement :

$P_1 \wedge P_2 = 1 \Rightarrow \exists\, A_2$ et B_2, $B_2P_2 = 1 - A_2P_1$
$P_1 \wedge P_3 = 1 \Rightarrow \exists\, A_3$ et B_3, $B_3P_3 = 1 - A_3P_1$
...
$P_1 \wedge P_k = 1 \Rightarrow \exists\, A_k$ et B_k, $B_kP_k = 1 - A_kP_1$
$\Rightarrow$ $B_2 \dots B_kP_2 \dots P_k = (1 - A_2P_1) \dots (1 - A_kP_1)$ de la forme $1 - AP_1$ en développant
$\Rightarrow$ $P_1 \wedge P_2 \dots P_k = 1$

c) Décomposition de E en sous-espaces stables

Une méthode pour décomposer E en sous-espaces stables est la suivante :

i) Déterminer un polynôme P annulateur de l'endomorphisme u, par exemple en calculant $P(X) = \det(X\,\mathrm{Id} - u)$ (théorème de Cayley-Hamilton, admis)

ii) Factoriser ce polynôme en produit de facteurs irréductibles distincts $R_1, \dots, R_m$ de sorte que $P = R_1^{p_1} \times R_2^{p_2} \times \dots R_m^{p_m}$

iii) Poser $P_1 = R_1^{p_1}$, $P_2 = R_2^{p_2}$, ... $P_m = R_m^{p_m}$. Les P_i sont premiers entre eux deux à deux, et $E = \mathrm{Ker}(P(u)) = \mathrm{Ker}(P_1(u)P_2(u) \dots P_m(u))$ de sorte que $E = E_1 \oplus \dots \oplus E_m$ avec $E_i = \mathrm{Ker}(P_i(u))$.

Cette méthode échoue si $m = 1$, autrement dit si P est puissance d'un facteur irréductible. C'est par exemple le cas des endomorphismes nilpotents, pour lesquels il existe une puissance p telle que $u^p = 0$.

d) Diagonalisation

Nous retrouvons un premier résultat sur la diagonalisation. S'il existe un polynôme annulateur de u ***scindé à racines simples***, $P = (X - \lambda_1) \dots (X - \lambda_k)$, alors u est diagonalisable. En effet, la méthode précédente conduit à :

$$E = \mathrm{Ker}(u - \lambda_1 I) \oplus \mathrm{Ker}(u - \lambda_2 I) \oplus \dots \oplus \mathrm{Ker}(u - \lambda_m I)$$

et si on prend une base de chacun de ces sous-espaces vectoriels, alors la matrice de u sera de la forme $\begin{pmatrix} \Lambda_1 & O & \dots & O \\ O & \Lambda_2 & \dots & O \\ \dots & & \dots & \\ O & \dots & O & \Lambda_m \end{pmatrix}$ où les Λ_k sont des blocs ayant des λ_k sur la diagonale.

e) Cas des endomorphismes nilpotents

Un endomorphisme nilpotent v vérifie $v^p = 0$ pour une certaine puissance de p. Un polynôme annulateur est donc X^p. La méthode du c) ne permet donc pas de décomposer E en somme directe de sous-espaces vectoriels stables. Cependant, il existe une suite croissante de sous-espaces vectoriels stables. Soit p la plus petite puissance de v qui s'annule. On vérifiera facilement que :

$$\{0\} \subset \mathrm{Ker}(v) \subset \mathrm{Ker}(v^2) \subset \dots \subset \mathrm{Ker}(v^p) = E$$

En fait, les inclusions précédentes sont strictes. En effet, si jamais on a :

$\mathrm{Ker}(v^k) = \mathrm{Ker}(v^{k+1})$ avec $k < p$

alors, $x \in \mathrm{Ker}(v^{k+2}) \Rightarrow v^{k+2}(x) = 0 \Rightarrow v^{k+1}(v(x)) = 0 \Rightarrow v(x) \in \mathrm{Ker}(v^{k+1}) \Rightarrow v(x) \in \mathrm{Ker}(v^k) \Rightarrow v^{k+1}(x) = 0 \Rightarrow x \in \mathrm{Ker}(v^{k+1})$ donc $\mathrm{Ker}(v^{k+2}) = \mathrm{Ker}(v^{k+1})$ et par récurrence $= \mathrm{Ker}(v^p) = E$, ce qui contredit la minimalité de p.

On prend alors une base de Ker(v), complétée en une base de Ker(v^2), ... complétée en une base de Ker(v^p). Du fait que, si x appartient à Ker(v^k), son image par v appartient à Ker(v^{k-1}), on obtient donc une matrice par blocs ayant la forme suivante :

$$\begin{pmatrix} O & A & \dots & & \\ O & O & B & \dots & \\ \dots & & \dots & & \\ O & \dots & O & C & \\ O & O & \dots & O & \end{pmatrix}$$

Inversement, on vérifie facilement que toute matrice de cette forme est nilpotente.

f) Trigonalisation

Supposons qu'un polynôme annulateur de u (par exemple son polynôme caractéristique) soit scindé, mais pas à racines simples, c'est-à-dire de la forme $P(X) = \prod_{i=1}^{p} (X - \lambda_i)^{k_i}$, où les λ_i sont distincts.

Posons $P_i = (X - \lambda_i)^{k_i}$ et $E_i = \mathrm{Ker}(P_i(u))$. Les P_i étant premiers entre eux deux et le produit des P_i s'annulant sur E, on en déduit que E est la somme directe des E_i, comme vu dans le b).

Si on se restreint à l'un de ces sous–espaces E_i, $v = u - \lambda_i\mathrm{Id}$ vérifie $v^{k_i} = 0$, donc v est nilpotent. Or nous avons vu dans le e) qu'il existe une base de E_i dans laquelle la matrice de v est triangulaire supérieure, avec des 0 sur la diagonale. Il en résulte que la matrice de u restreinte à E_i, est triangulaire supérieure avec des λ_i sur la diagonale. Quand on regroupe les bases des divers E_i pour obtenir une base de E, on obtient une base dans laquelle la matrice de u est triangulaire supérieure.

Ainsi, tout endomorphisme ayant un polynôme annulateur ***scindé*** est trigonalisable.

C'est en particulier le cas de tout endomorphisme lorsque le corps de base est $\mathbb{C}$, puisque tout polynôme sur $\mathbb{C}$ est scindé (Théorème de d'Alembert)

g) Théorème de Cayley-Hamilton

On déduit du théorème de trigonalisation (éventuellement en se plaçant sur $\mathbb{C}$) une démonstration du théorème de Cayley-Hamilton. Toute matrice A est semblable à une matrice triangulaire complexe T de la forme :

$$T = \begin{pmatrix} T_1 & O & \dots & O \\ O & T_2 & \dots & O \\ \dots & & \dots & \\ O & \dots & O & T_m \end{pmatrix}$$

avec $T_i = \begin{pmatrix} \lambda_i & * & \dots & * \\ 0 & \lambda_i & \dots & * \\ \dots & & \dots & \\ 0 & \dots & 0 & \lambda_i \end{pmatrix}$, matrice d'ordre k_i.

Le polynôme caractéristique de T, et donc de A, est $(X - \lambda_1)^{k_1}(X - \lambda_2)^{k_2}\dots(X - \lambda_m)^{k_m}$. Il s'agit donc de montrer que :

$$(A - \lambda_1 I)^{k_1}(A - \lambda_2 I)^{k_2}\dots(A - \lambda_m I)^{k_m} = 0$$

$$\Leftrightarrow \quad P(T - \lambda_1 I)^{k_1}(T - \lambda_2 I)^{k_2}\dots(T - \lambda_m I)^{k_m}P^{-1} = 0$$

$$\Leftrightarrow \quad (T - \lambda_1 I)^{k_1}(T - \lambda_2 I)^{k_2}\dots(T - \lambda_m I)^{k_m} = 0.$$

Or $T_i - \lambda_i I = \begin{pmatrix} 0 & * & \dots & * \\ 0 & 0 & \dots & * \\ \dots & & \dots & \\ 0 & \dots & 0 & 0 \end{pmatrix}$ est une matrice nilpotente, chaque puissance décalant le triangle supérieur d'un rang. Il en résulte que $(T_i - \lambda_i I)^{k_i} = O$. On en déduit que :

$$(T - \lambda_1 I)^{k_1} \text{ est de la forme } \begin{pmatrix} O & O & \dots & O \\ O & (T_2 - \lambda_1 I)^{k_1} & \dots & O \\ \dots & & \dots & \\ O & \dots & O & (T_m - \lambda_1 I)^{k_1} \end{pmatrix}$$

$$(T - \lambda_2 I)^{k_2} \text{ est de la forme } \begin{pmatrix} (T_1 - \lambda_2 I)^{k_2} & O & \dots & O \\ O & O & \dots & O \\ \dots & & \dots & \\ O & \dots & O & (T_m - \lambda_2 I)^{k_2} \end{pmatrix}$$

...

$$(T - \lambda_m I)^{k_m} \text{ est de la forme } \begin{pmatrix} (T_1 - \lambda_m I)^{k_m} & O & \dots & O \\ O & (T_2 - \lambda_m I)^{k_m} & \dots & O \\ \dots & & \dots & \\ O & \dots & O & O \end{pmatrix}$$

Quand on multiplie ces matrices, on obtient la matrice nulle.
*Fin de partie réservée aux MP/MP**

Annexe III : Le théorème de Cayley-Hamilton

Le théorème de Cayley-Hamilton énonce que :

Si $\chi_u(X) = \det(X\,\mathrm{Id} - u)$, *alors* $\chi_u(u) = 0$.

Ce théorème est facile à vérifier en dimension 2. En effet, si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est la matrice de u, alors :

$$\chi_u(X) = \begin{vmatrix} X - a & -b \\ -c & X - d \end{vmatrix} = X^2 - (a + d)X + ad - bc$$

Il est alors facile de vérifier que $A^2 - (a + d)A + (ad - bc)I = 0$.

Le cas général est plus délicat. Il convient d'abord de prendre conscience que, si on définit le polynôme $\chi(X) = \det(XI_n - A)$, il est incorrect de dire que :

$$\chi(A) = \det(AI_n - A) = \det(A - A) = \det(0) = 0$$

ne serait-ce que parce que le résultat final 0 est le scalaire 0, alors que le membre de gauche $\chi(A)$ est une matrice. Mais il y a une erreur plus profonde. Dans $\det(XI_n - A)$, le produit de X par I_n est un produit externe de la matrice I_n par l'élément X de $\mathbb{K}[X]$. Il représente donc la matrice $\begin{pmatrix} X & 0 & \dots & 0 \\ 0 & X & \dots & 0 \\ \dots & & \dots & \\ 0 & \dots & 0 & X \end{pmatrix}$. Remplacer X par A ne revient donc absolument pas à calculer AI_n qui est un produit interne de deux matrices de $\mathcal{M}_n(\mathbb{K})$. Le remplacement de X par A dans XI_n pourrait à la rigueur s'interpréter comme définissant une matrice par blocs $\begin{pmatrix} A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \dots & & \dots & \\ 0 & \dots & 0 & A \end{pmatrix}$, mais cette dernière matrice est un élément de $\mathcal{M}_{n^2}(\mathbb{K})$ et non de $\mathcal{M}_n(\mathbb{K})$. Comment alors la retrancher de la matrice $A = (a_{ij})$ qui, elle, est dans $\mathcal{M}_n(\mathbb{K})$? On peut enfin remarquer que A et ${}^tA$ ont le même polynôme caractéristique (car le déterminant est le même pour une matrice et sa transposée) et donc qu'on a aussi bien $\chi(A) = 0$ que $\chi({}^tA) = 0$. Or si on calcule $\chi({}^tA)$ en remplaçant de façon erronée X par ${}^tA$ dans $\det(XI_n - A)$, on obtient non pas $\chi({}^tA) = 0$, mais $\det({}^tA - A)$ qui est non nul en général.

Dans tous les cas, ces tentatives ne traduisent pas ce que veut dire le théorème de Cayley-Hamilton. Celui-ci énonce que l'on calcule d'abord le déterminant $\det(XI_n - A)$, puis, ayant obtenu une expression polynomiale, que l'on remplace X par A. Le calcul matriciel que l'on effectue alors conduit à la matrice nulle.

Nous donnons dans cette annexe une démonstrations du théorème de Cayley-Hamilton basée sur la comatrice, notion du programme de MPSI. Soit A la matrice de u. Considérons la matrice $B(X) = XI_n - A$, à coefficients polynomiaux ou même, si on le souhaite à coefficients dans le corps $\mathbb{K}(X)$ des fractions rationnelles. Considérons également la transposée de la comatrice de $XI_n - A$, que nous appellerons $C(X)$, de sorte que $B(X)C(X) = \det(B(X))I$, où $\det(B(X))$ n'est autre que $\chi_u(X)$. Rappelons que les coefficients de $C(X)$ étant formés à partir de sous–déterminants $n-1 \times n-1$ de $B(X)$, ces coefficients sont donc des polynômes en X. Ainsi :

$$B(X)C(X) = \chi_u(X)I$$

Regroupons dans $C(X)$ les termes constants, puis les termes en X, etc..., ce qui revient à écrire $C(X)$ sous forme de combinaison linéaire de matrices C_i :

$$C(X) = C_0 + XC_1 + ... + X^{n-1}C_{n-1}$$

On a alors, en développant le produit $B(X)C(X)$:

$$\begin{aligned}\chi_u(X)I &= (XI_n - A)(C_0 + XC_1 + ... + X^{n-1}C_{n-1}) \\ &= -AC_0 + X(C_0 - AC_1) + ... + X^{n-1}(C_{n-2} - AC_{n-1}) + X^nC_{n-1}\end{aligned}$$

Justifions maintenant le résultat suivant :
Si $T_0 + XT_1 + ... + X^nT_n = U_0 + XU_1 + ... + X^nU_n$, où les T_i et les U_i sont des matrices carrés, alors pour tout i, $T_i = U_i$. En effet, l'égalité précédente implique l'égalité formelle des polynômes constituant les termes (i,j) de chacun des deux membres. Les coefficients de ces polynômes sont donc égaux, ce qui signifie que $T_0 = U_0$, ... $T_n = U_n$.
Supposons que $\chi_u(X) = a_0 + a_1X + ... + a_nX^n$. On a donc, d'après l'identification ci-dessus entre les coefficients de $\chi_u(X)I_n$ et ceux de $B(X)C(X)$

$$\begin{aligned}a_0I &= -AC_0 \\ a_1I &= C_0 - AC_1 \\ &\ldots \\ a_{n-1}I &= C_{n-2} - AC_{n-1} \\ a_nI &= C_{n-1}\end{aligned}$$

Si on multiplie a_1I par A, a_2I par A^2, ..., a_nI par A^n et si on ajoute membre à membre, on obtient :

$$a_0I + a_1A + ... + a_nA^n = 0$$

ce qui est le résultat cherché.

◆