Polynômes d'endomorphismes et applications

Jean-François Burnol, 10 septembre 2010

Dans toute cette fiche il sera question d'endomorphismes sur des K-espaces vectoriels de **dimensions finies**.

Postface : (15 septembre) bon je me suis visiblement laissé entraîner par mon élan. Je n'en suis pas trop mécontent car cela m'a permis de voir que je connaissais des choses sans même savoir comment, ni d'où, cela venait, et heureusement qu'il y a Internet qui m'a permis de finaliser en y trouvant les noms appropriés (par exemple : « matrices compagnons », « décomposition de Frobenius ») pour toutes ces choses. Au final il y a donc ici certaines des notions qui sont probablement un peu trop poussées pour être directement utiles à l'Oral, mais elles pourraient bien l'être pour l'Écrit. En effet, cette fiche est devenue un

Mini-Traité sur la réduction des endomorphismes

où tout ce que vous n'avez jamais voulu savoir sur la décomposition de Frobenius occupe une douzaine de pages alors que le plus raisonnable et convenable Lemme des Noyaux n'est qu'à peine évoqué à la toute fin.

1 Rapides rappels basiques, qu'il faut connaître de toute façon

Le théorème de Cayley-Hamilton peut s'aborder de multiples manières. Je voudrais rappeler d'abord brièvement les aspects les plus basiques de la décomposition des endomorphismes, qui ont en particulier comme conséquence une preuve du théorème de Cayley-Hamilton. Pour ce dernier (que l'on peut voir comme un énoncé sur les matrices carrées) on peut d'emblée supposer que le corps K est algébriquement clos. Cette hypothèse permet lorsque l'on a un espace vectoriel V de dimension finie muni d'un endomorphisme f de travailler avec les valeurs propres (attention : $V \neq \{0\}$). Donc soit λ une valeur propre de f, $\dim \mathrm{Ker}(f-\lambda) > 0$ (l'endomorphisme identité Id_V est sous-entendu dans la notation : $f - \lambda = f - \lambda \mathrm{Id}_V$). La suite $K_n = \mathrm{Ker}(f-\lambda)^n$ (avec $K_0 = \{0\}$) est croissante pour l'inclusion. Il existe un plus petit indice N (et $N \geq 1$) avec $K_N = K_{N+1}$. Comme les inclusions précédentes sont strictes on a $\dim K_N \geq N$, une remarque qui sera utile plus tard. À partir de $K_N = K_{N+1}$ il est facile de voir que $K_N = K_{N+1} = K_{N+2} = \ldots$ (je le laisse en exercice). Soit $J_n = \mathrm{Im}(f-\lambda)^n$.

La suite des images $J_n = \mathrm{Im}(f-\lambda)^n$ est décroissante pour l'inclusion et se stabilise aussi à partir de $n = N$, en utilisant le théorème du rang pour le calcul de la dimension de J_n. Si $x \in K_N \cap J_N$ alors $x = (f-\lambda)^N(y)$ et $(f-\lambda)^N(x) = 0$ donc $(f-\lambda)^{2N}(y) = 0$ donc $(f-\lambda)^N(y) = 0$ donc $x = 0$. Donc K_N et J_N sont en somme directe et comme leurs dimensions sont complémentaires (théorème du rang) on obtient la décomposition en somme directe $V = K_N \oplus J_N$. Sur J_N, $f-\lambda$ est injectif (puisque le noyau est K_1 et $K_1 \cap J_N = \{0\}$), donc λ n'est pas valeur propre de f sur J_N. On en déduit par récurrence sur $\dim V$ qu'il existe une décomposition (unique d'ailleurs) $V = \oplus V_\lambda$ indexée par les valeurs propres λ et $f-\lambda$ est nilpotent sur V_λ. **Les espaces V_λ sont les « espaces caractéristiques » de l'endomorphisme** f. On peut choisir une base de V_λ de sorte que f y soit représenté par une une matrice triangulaire supérieure avec des λ sur la diagonale ($f-\lambda$ est nilpotente sur V_λ, et la trigonalisabilité d'un endomorphisme nilpotent N se montre facilement par récurrence sur la dimension de l'espace, puisque tout passage à un quotient redonne un endomorphisme nilpotent, ou encore en utilisant une forme linéaire non nulle L avec $N^*(L) = 0$, puisque N^* est aussi nilpotent, et alors l'hyperplan $L(x) = 0$ est stable par N). En combinant les bases des différents V_λ, on obtient une matrice triangulaire représentant f, et on en déduit que le polynôme caractéristique $\det(X-f)$ vaut $\prod_\lambda (X-\lambda)^{\dim V_\lambda}$. Comme $\dim V_\lambda \geq N_\lambda$ (« remarque utile » faite précédemment), l'endomorphisme $\prod_\lambda (f-\lambda)^{\dim V_\lambda}$ vaut zéro sur chaque V_λ donc sur V tout entier : d'où le théorème de Cayley-Hamilton.

Tout ce que je viens de dire, il faut absolument le connaître. Dans la suite on s'intéresse plus généralement à tous les polynômes en l'endomorphisme f. Paradoxalement peut-être, dans un premier temps au moins on n'y utilisera pas ce qui précède, encore que la boucle sera bouclée vers la fin de notre épopée. Et on ne veut pas être restreint à un corps algébriquement clos. Ceci élargit le champ de vision, et le langage le plus adapté serait celui des modules sur l'anneau $K[X]$, mais j'ai par masochisme voulu l'éviter (ça me paraissait impossible mais en fin de compte la Didactique habitue à tous les sacrifices, par usure).

2 Polynômes minimaux (au pluriel)

Soit V un K-espace vectoriel et $f \in \mathrm{End}(V)$ un endomorphisme. Il existe un (unique) morphisme de K-algèbres de $K[X]$ sur $\mathrm{End}(V)$ qui envoie X sur f. On note $P(f)$ l'image du polynôme P, car en effet il s'agit de substituer f à X dans P.

Le noyau est un idéal de $K[X]$, idéal non nul, car $K[X]$ est un K espace vectoriel de dimension infinie tandis que $\mathrm{End}(V)$ est de dimension finie. Donc il est engendré par un unique polynôme unitaire M_f : on l'appelle polynôme (annulateur) minimal de f. C'est donc le polynôme unitaire de plus bas degré

vérifiant $M_f(f) = 0$. Par exemple si $f = 0$ alors $M_f = X$, sauf si $V = \{0\}$ auquel cas $M_f = 1$. Le cas $M_f = 1$ ne peut survenir que si V est l'espace nul. Si $\dim V > 0$ alors $\deg M_f > 0$.

On peut (ici $\dim V = N > 0$) représenter f par une matrice carrée A de taille N, et M_f est le polynôme unitaire de plus bas degré avec $M_f(A) = 0$, donc on peut aussi parler du polynôme minimal $M_A = M_f$ de la matrice A. Comme pour tout polynôme P on a $P(SAS^{-1}) = SP(A)S^{-1}$, le polynôme minimal $M_{SAS^{-1}} = M_A$ est un invariant de similitude de la matrice A. Notons aussi l'énoncé important suivant :

Proposition 1. **Soit A une matrice carrée de taille N à coefficients dans un corps K. Le polynôme minimal de A reste inchangé par le passage à un sur-corps L de K.**

Preuve : nous représentons les matrices carrées de taille N comme des lignes avec N^2 coefficients. Pour chaque n, appliquons l'algorithme du pivot de Gauss aux $n+1$ lignes I_N, A, ..., A^n, algorithme qui procède par combinaisons linéaires et permutations de lignes. Le nombre de lignes non nulles à la fin est égal à la dimension d_n de l'espace vectoriel engendré par I_N, A, ..., A^n. Le degré du polynôme minimal est le plus petit n avec $d_n < 1 + n$. Comme l'algorithme du pivot ne dépend que des coefficients et non du corps ambiant L, le degré du polynôme minimal reste inchangé. Donc le polynôme minimal reste inchangé (puisque celui sur L divise dans $L[X]$ celui sur K). On peut aussi faire la preuve en utilisant une K-forme linéaire $\lambda : L \to K$ avec $\lambda(1) = 1$, mais la preuve d'existence de λ nécessite le Lemme de Zorn si L est de dimension infinie sur K. Cependant si $K = \mathbf{R}$ et $L = \mathbf{C}$ cette approche marche très bien avec $\lambda(z) = \mathrm{Re}(z)$: si $P \in L[X]$ annule A alors $\lambda(P) \in K[X]$ aussi, et $\lambda(P)$ est unitaire si P l'est. On retrouve que le degré du polynôme minimal ne peut pas décroître.

On peut aussi associer à tout vecteur x l'idéal dans $K[X]$ des polynômes P avec $P(f)(x) = 0$, d'où un polynôme minimal M_x (aussi appelé polynôme annulateur et bien sûr il faudrait plutôt écrire $M_{f,x}$). On a $\deg M_x > 0$ sauf si $x = 0$. Certainement M_x divise M_f puisque $M_f(f)(x) = 0$.

Nous essayons maintenant de comprendre la relation entre M_f et les M_x. Soit $(e_1, \ldots, e_N)$ une base de V. Notons M_1, ..., M_N les polynômes minimaux associés aux vecteurs de la base et soit P leur PPCM (que l'on prend unitaire lui aussi). On a $P(f)(e_j) = 0$ pour tout j donc $P(f) = 0$ donc P est un multiple (au sens de la division des polynômes) du polynôme minimal M_f. Mais par ailleurs chaque M_j divise M_f donc leur PPCM divise M_f. Au final : $M_f = P = \mathrm{ppcm}(M_1, \ldots, M_N)$.

En fait, et cela sera notre premier énoncé très significatif, on peut toujours trouver un vecteur x avec $M_x = M_f$. Nous verrons cela plus tard, pour le moment faisons une pause pour montrer le théorème de Cayley-Hamilton d'une façon très simple.

3 Une démonstration super géniale du théorème de Cayley Hamilton

Soit A une matrice carrée et $P(X) = \det(X - A)$ son polynôme caractéristique. On a :

$$P(A) = \det(A - A) = \det 0 = 0 \tag{1}$$

Ceux qui sont satisfaits par cette preuve peuvent arrêter la lecture ici, car il est peu probable que la suite leur soit utile ! (soit parce qu'ils sont d'excellents mathématiciens et savent comment convertir la chose ci-dessus en une vraie preuve, soit au contraire parce qu'ils sont de très modestes mathématiciens pour lesquels c'était déjà une vraie preuve).

4 Matrices compagnons et Cayley-Hamilton

Il y a une situation qui sera fondamentale pour nous : on dit que V est cyclique (pour f) s'il existe un vecteur x tel que V soit engendré par x, $f(x)$, $f^2(x)$, ... Calculons dans ce cas le polynôme minimal de f. D'abord si $x = 0$ alors $V = \{0\}$ et $M_f = 1$. Sinon il existe un plus grand entier $n \geq 1$ tel que x, $f(x)$, ..., $f^{n-1}(x)$ soient linéairement indépendants. Donc $f^n(x)$ est une combinaison linéaire des précédents :

$$f^n(x) = a_{n-1} f^{n-1}(x) + \cdots + a_0\, x \tag{2}$$

Par récurrence pour $m \geq n$, $f^m(x)$ est aussi combinaison linéaire de ces n vecteurs. Donc ces vecteurs forment une base de V. Notons

$$P(X) = X^n - a_{n-1}X^{n-1} - \cdots - a_0 \tag{3}$$

On a $P(f)(x) = 0$, donc $P(f)(f(x)) = f(P(f)(x)) = 0$, etc...Donc P est tel que $P(f) = 0$ et est ainsi un multiple (au sens de la division des polynômes) de M_f. Par ailleurs pour tout polynôme non nul Q de degré strictement inférieur à n, on a $Q(f)(x) \neq 0$ puisque x, $f(x)$, ..., $f^{n-1}(x)$ sont linéairement indépendants. Donc M_f est de degré au moins n et finalement $M_f = P$.

Écrivons la matrice de f dans la base $(x, f(x), \ldots, f^{n-1}(x))$. Elle est

$$C(P) = \begin{pmatrix} 0 & & & & & a_0 \\ 1 & 0 & & & & a_1 \\ & 1 & 0 & & & a_2 \\ & & \ddots & \ddots & & \\ & & & 1 & 0 & a_{n-2} \\ & & & & 1 & a_{n-1} \end{pmatrix} \tag{4}$$

Une telle matrice s'appelle "matrice compagnon" (du polynôme unitaire P).

Proposition 2. Le polynôme caractéristique de la matrice compagnon est égal à P.

Preuve : par exemple par récurrence en développant le déterminant $\det(XI_n - C(P))$ par rapport à la première colonne. Nous ré-aborderons ce point plus tard pour en donner une version plus sophistiquée.

Nous pouvons maintenant donner une preuve rapide du théorème de Cayley-Hamilton :

Théorème 1. Soit V un espace vectoriel non nul et f un endomorphisme de polynôme caractéristique P. Alors $P(f) = 0$.

Preuve : soit x un vecteur non nul, et V_x le plus petit sous-espace stable par f et contenant x, c'est-à-dire l'espace cyclique engendré par x et ses images sous f. Notons g la restriction de f à V_x. On sait (en complétant une base de V_x par le théorème de la base incomplète et en faisant un calcul de déterminant par bloc) que le polynôme caractéristique P_x de g divise celui de f. Or P_x par ce qui précède n'est autre que le polynôme minimal de x, en particulier $P_x(f)(x) = 0$. Donc $P(f)(x) = 0$. Donc $P(f) = 0$. Attention, en dimension zéro le théorème est vrai mais de manière un peu spéciale : on doit convenir $P = 1$ (par exemple parce qu'un produit sur un ensemble vide est toujours pris égal à 1, de même qu'une somme sur un ensemble vide est toujours prise égale à zéro) donc $P(f) = \mathrm{Id}$ et on a $\mathrm{Id}_V = 0_V$ comme identité d'endomorphismes sur l'espace vectoriel nul V.

Il est utile d'étudier de plus près les espaces cycliques.

Proposition 3. Soit P_1 et P_2 deux polynômes unitaires. S'il existe un isomorphisme $\phi : K[X]/(P_1) \simeq K[X]/(P_2)$ de K-espaces vectoriels vérifiant $\forall Q\ \phi(XQ) = X\phi(Q)$, alors $P_1 = P_2$.

Preuve : notons $V_1 = K[X]/(P_1)$ et soit $f_1 \in \mathrm{End}(V_1)$ la multiplication par X et notons $V_2 = K[X]/(P_2)$ et soit $f_2 \in \mathrm{End}(V_2)$ la multiplication par X. Le polynôme minimal M_1 de f_1 est P_1 et le polynôme minimal M_2 de f_2 est P_2. Or, pour tout polynôme T, on $\phi \circ T(f_1) = T(f_2) \circ \phi$. Donc $M_1(f_2) = 0$ et M_2 divise M_1. En échangeant les rôles, M_1 divise M_2 et finalement $M_1 = M_2$ c'est-à-dire $P_1 = P_2$. On peut aussi dire qu'à cause de l'isomorphisme ϕ, la multiplication par X a même polynôme caractéristique dans les deux espaces V_1 et V_2, or dans le premier sa matrice dans la base canonique $(1, X, \ldots, X^{n-1})$ est la matrice compagnon $C(P_1)$ et dans le second c'est $C(P_2)$. Donc $P_1 = P_2$.

L'énoncé qui suit sera utilisé par la suite :

Proposition 4. Soit $V = K[X]/(P)$ et f l'endomorphisme de multiplication par X. Pour tout polynôme unitaire T, le noyau $\mathrm{Ker}\, T(f)$ est le sous espace cyclique engendré par $\frac{P}{\mathrm{pgcd}(T,P)}$ et le polynôme minimal associé est $\mathrm{pgcd}(T, P)$. En particulier :

$$\dim \mathrm{Ker}\, T(f) = \deg \mathrm{pgcd}(T, P) \tag{5}$$

Preuve : soit T un polynôme unitaire, déterminons $\mathrm{Ker}\, T(f)$. Tout d'abord si $x \in V$ est (la classe d') un polynôme Q alors tout bêtement $T(f)(x)$ est (la classe) du polynôme TQ. Donc pour que $T(f)(x) = 0$, il faut et il suffit que P divise TQ. Notons $D = \mathrm{pgcd}(T, P)$, $P = DA$, $T = DB$. Pour que DA divise QDB il est nécessaire et suffisant que A divise QB mais A est premier à B donc cela équivaut à ce que A divise Q. Donc le noyau de $T(f)$ est composé des vecteurs x de la forme $Q = AR$ (modulo P). Il s'agit donc du sous-espace cyclique engendré par A $(= \frac{P}{\mathrm{pgcd}(T,P)})$. Si l'on multiplie A par un polynôme non nul de degré strictement inférieur à celui de $D = \mathrm{pgcd}(T, P)$, on obtient un polynôme non nul de degré strictement inférieur à celui de P donc un élément non nul de $K[X]/(P)$. Donc le polynôme minimal de A est $\mathrm{pgcd}(T, P)$. Fin de la preuve.

5 Existence d'un vecteur de polynôme minimal maximal

Nous prouvons maintenant le résultat crucial suivant :

Théorème 2. **Étant donné dans V des vecteurs $x_1, \ldots, x_k$, de polynômes minimaux associés $M_1, \ldots, M_k$, il existe un vecteur x dans V de polynôme minimal associé $M_x = \mathrm{ppcm}(M_1, \ldots, M_k)$.**

Preuve : il suffit de faire le cas $k = 2$ car le cas général en résulte par une récurrence (et le fait que $\mathrm{ppcm}(M_1, M_2, M_3) = \mathrm{ppcm}(\mathrm{ppcm}(M_1, M_2), M_3)$ etc...). Supposons donc $k = 2$ et aussi, dans un premier temps que M_1 et M_2 sont premiers entre eux. Alors je dis que $x = x_1 + x_2$ convient. En effet si $0 = P(f)(x) = P(f)(x_1) + P(f)(x_2)$ alors $0 = (PM_2)(f)(x_1) + 0$ donc M_1 divise PM_2 donc M_1 divise P. De même M_2 divise P donc M_1M_2 divise P. Donc M_1M_2 divise M_x. Mais bien sûr $(M_1M_2)(f)(x) = 0$ donc M_x divise M_1M_2. Donc $M_x = M_1M_2$.

Dans le cas général on fait quelque chose d'un peu rusé. Il n'y a qu'un nombre fini de polynômes unitaires irréductibles distincts divisant M_1 ou M_2. Notons-les $\pi_1, \ldots, \pi_k$. Ainsi :

$$M_1 = \pi_1^{a_1} \ldots \pi_k^{a_k} \qquad M_2 = \pi_1^{b_1} \ldots \pi_k^{b_k} \tag{6}$$

Soit $J \subset \{1, \ldots, k\}$ l'ensemble des j avec $a_j \geq b_j$ et posons (rappel : par convention un produit vide vaut 1) :

$$N_1 = \prod_{j \in J} \pi_j^{a_j} \qquad N_2 = \prod_{j \in J^c} \pi_j^{b_j} \tag{7}$$

Par construction $\mathrm{ppcm}(M_1, M_2) = N_1N_2$, et N_1 et N_2 sont premiers entre eux. Enfin écrivons $M_1 = N_1T_1$, $M_2 = N_2T_2$, et $x_1' = T_1(f)(x_1)$ et $x_2' = T_2(f)(x_2)$.

Je prétends que le polynôme minimal de x'_1 est N_1 et celui de x'_2 est N_2. En effet si $P(f)T_1(f)(x_1) = 0$ alors PT_1 est divisible par $M_1 = N_1T_1$ donc N_1 divise P. Et réciproquement $N_1(f)(x'_1) = N_1(f)T_1(f)(x_1) = M_1(f)(x_1) = 0$. De même pour x'_2. Donc par ce que l'on a fait avant, le polynôme minimal de $x'_1 + x'_2$ est $N_1N_2 = \text{ppcm}(M_1, M_2)$, comme voulu. Ce qui termine cette preuve.

6 Décomposition

Si l'on reprend donc notre discussion d'un endomorphisme f sur un K-espace vectoriel V non nul on sait à ce stade qu'il existe des vecteurs x tels que $M_x = M_f$, puisque nous avions vu précédemment $M_f = \text{ppcm}(M_{e_1}, \ldots, M_{e_n})$ pour toute base $(e_1, \ldots, e_n)$ de V. La deuxième étape cruciale est :

Théorème 3. *Soit x tel que $M_x = M_f$. Alors le sous-espace cyclique V_x de V engendré par x et ses images sous f admet un complémentaire W stable par f.*

La preuve en est assez rusée. Tout d'abord, si V est l'espace nul le théorème est vrai bien que pas très exaltant. Si V n'est pas l'espace nul on a $\deg M_f > 0$, donc $\deg M_x > 0$ et x n'est pas nul. Soit $n = \dim V_x \geq 1$, V_x possède comme base x, $f(x)$, ..., $f^{n-1}(x)$. Choisissons une forme linéaire L sur V avec les contraintes :

$$L(x) = L(f(x)) = \cdots = L(f^{n-2}(x)) = 0 \quad \text{et} \quad L(f^{n-1}(x)) = 1 \tag{8}$$

Bien sûr si $n = 1$ il faut juste comprendre $L(x) = 1$. Maintenant on pose $L_1 = L$ et $L_2 = f^*(L_1)$, ..., $L_n = (f^*)^{n-1}(L_1)$. C'est-à-dire :

$$L_k(y) = L(f^k(y)) \tag{9}$$

Finalement j'écris l_1, ..., l_n pour les restrictions de L_1, ..., L_n à V_x. Dans la base de V_x^* duale de la base $(x, \ldots, f^{n-1}(x))$ de V_x, les formes linéaires l_1, ..., l_n s'écrivent de manière triangulaire inversible (plus précisément la matrice des coefficients de l_n, ..., l_1 est triangulaire inférieure avec des 1 sur la diagonale). Elles sont donc linéairement indépendantes.

Donc les n formes linéaires L_1, ..., L_n sur V sont linéairement indépendantes. Leur noyau commun W est donc de codimension n, de plus son intersection avec V_x est aussi de codimension n dans V_x donc cette intersection est nulle, donc en fait $V = V_x \oplus W$. Il reste à prouver que W est stable sous f. Mais si $y \in W$ et $z = f(y)$ alors tout d'abord :

$$L_1(z) = L_2(y) = 0, L_2(z) = L_3(y) = 0, \ldots, L_{n-1}(z) = L_n(y) = 0 \tag{10}$$

et finalement $L_n(z) = L_1(f^n(y))$ or le polynôme minimal M_f est de degré n donc $f^n(y)$ est combinaison linéaire de y, $f(y)$, ..., $f^{n-1}(y)$. Comme L_1 s'annule sur

eux, on a bien $L_n(z) = 0$. Donc $z \in W$, ce qu'il fallait démontrer. Le Théorème est établi.

Si V n'est pas l'espace nul nous avons vu que le x n'est pas un vecteur nul, donc W, supplémentaire de V_x est un sous-espace propre de V (éventuellement nul). Si W n'est pas l'espace nul, nous pouvons itérer la construction, et ainsi de suite jusqu'à obtenir un espace nul. Notons M_1 le polynôme minimal de f sur V, M_2 celui de sa restriction au sous-espace propre W, etc.... Ainsi M_2 divise M_1, etc...Ainsi, par récurrence sur la dimension, on obtient le théorème de structure (première partie : existence) :

Théorème 4. Soit f un endomorphisme d'un K-espace vectoriel $V \neq \{0\}$. Il existe une décomposition de V en somme directe de sous-espaces cycliques non nuls

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_p \tag{11}$$

de polynômes minimaux associés M_1, ..., M_p, avec M_{j+1} diviseur de M_j pour $1 \leq j < p$. Le polynôme minimal de f est M_1 et le polynôme caractéristique $\det(X - f)$ est le produit $M_1 M_2 \cdots M_p$.

La formule pour le polynôme caractéristique résulte du calcul des polynômes caractéristiques pour les espaces cycliques. On note comme conséquence immédiate le Théorème de Cayley-Hamilton, puisque notre formule prouve que le polynôme minimal divise le polynôme caractéristique.

La décomposition du théorème n'est pas du tout unique, par contre, et c'est là la deuxième partie du théorème de structure, le nombre p et les polynômes M_1, ..., M_p, eux sont déterminés de manière unique :

7 Unicité

Théorème 5. Soit $p \geq 1$ un entier et $M_p | M_{p-1} | \cdots | M_1$ des polynômes unitaires de degrés strictement positifs. Et soit $q \geq 1$ un entier et $N_q | N_{q-1} | \cdots | N_1$ des polynômes unitaires de degrés strictement positifs. S'il existe un isomorphisme

$$\phi : K[X]/(M_1) \oplus \cdots \oplus K[X]/(M_p) \simeq K[X]/(N_1) \oplus \cdots \oplus K[X]/(N_q) \tag{12}$$

commutant à la multiplication par X alors $q = p$, $N_p = M_p$, ..., $N_1 = M_1$.

Soit $V = K[X]/(M_1) \oplus \cdots \oplus K[X]/(M_p)$ et notons f la multiplication par X dans V. Et soit $W = K[X]/(N_1) \oplus \cdots \oplus K[X]/(N_q)$ et g la multiplication par X dans W. Si T est un polynôme quelconque on a la formule :

$$\phi \circ T(f) = T(g) \circ \phi \tag{13}$$

qui généralise $\phi \circ f = g \circ \phi$ faisant partie des hypothèses de l'énoncé.

Ainsi $\phi(\operatorname{Ker} T(f)) = \operatorname{Ker} T(g)$ et $\dim \operatorname{Ker} T(f) = \dim \operatorname{Ker} T(g)$. Or, pour T unitaire, on a, d'après la Proposition 4,

$$\dim \operatorname{Ker} T(f) = \sum_{1 \leq j \leq p} \deg \operatorname{pgcd}(T, M_j) \leq p \deg T \tag{14}$$

l'égalité n'étant atteinte que si T divise tous les M_j, c'est-à-dire si T divise M_p. Je rappelle maintenant que dans notre énoncé on a imposé $\deg M_p > 0$. Je peux donc considérer la fonction ψ qui, aux polynômes unitaires non constants T associe

$$\psi(T) = \frac{\dim \operatorname{Ker} T(f)}{\deg T} = \frac{\dim \operatorname{Ker} T(g)}{\deg T} \tag{15}$$

Cette fonction est majorée par p et atteint son maximum p en $T = M_p$ et ses diviseurs unitaires non constants. Et elle est majorée par q et atteint son maximum q en $T = N_q$ et ses diviseurs. Donc $p = q$, et $M_p = N_q$ car ce polynôme est caractérisé comme étant celui de plus haut degré réalisant le maximum de la fonction ψ.

Notons donc T ce polynôme unitaire non constant $M_p = N_q$. L'isomorphisme $\phi : V \simeq W$ passe au quotient en un isomorphisme $\phi : V/\operatorname{Ker} T(f) \simeq W/\operatorname{Ker} T(g)$. De plus le noyau de la restriction de $M_p(f)$ au module cyclique $K[X]/(M_j)$ est engendré par le polynôme M_j/M_p, et on a l'isomorphisme commutant à la multiplication par X :

$$(K[X]/(M_j))/(M_j/M_p) \simeq K[X]/(M_j/M_p) \tag{16}$$

donc

$$V/\operatorname{Ker} T(f) \simeq \oplus_{1 \leq j \leq p'} K[X]/(M_j/M_p) \tag{17}$$

où l'on a arrêté la somme au plus grand indice j avec $\deg M_j > \deg M_p$. Il est possible qu'il n'y ait aucun tel indice auquel cas $M_1 = M_2 = \cdots = M_p$. Mais alors $V/\operatorname{Ker} T(f) = 0$ donc $W/\operatorname{Ker} T(g) = 0$ donc on a aussi $N_1 = N_2 = \cdots = N_q$ et on a déjà vu $p = q$ et $M_p = N_q$. Donc soit la preuve s'arrête là, soit on s'est ramené à un isomorphisme entre deux sommes directes non nulles d'espaces cycliques non nuls, avec un nombre inférieur de termes. Par hypothèse de récurrence, $M_j/M_p = N_j/N_q$ pour $j \leq p' = q'$ et le Théorème est démontré.

Si l'on revient à la situation d'un espace vectoriel V non nul et d'un endomorphisme f, et que l'on représente V comme somme directe de sous-espaces cycliques non nuls :

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_p \tag{18}$$

de polynômes minimaux associés M_1, ..., M_p vérifiant $M_p | M_{p-1} | \cdots | M_1$, tout choix de vecteurs $x_1 \in V_1$, ..., $x_p \in V_p$ les engendrant comme espaces cycliques donne un isomorphisme :

$$V \simeq K[X]/(M_1) \oplus \cdots \oplus K[X]/(M_p) \tag{19}$$

qui fait correspondre à f la multiplication par X. Une autre décomposition du même type donne un autre isomorphisme :

$$V \simeq K[X]/(N_1) \oplus \cdots \oplus K[X]/(N_q) \tag{20}$$

et par le Théorème que nous venons de prouver on a donc en fait $p = q$ et $M_j = N_j$, pour $1 \leq j \leq p$.

Attention, dans tout cela on a fait un usage essentiel de la condition que les M_j forment une chaîne décroissante pour la divisibilité des polynômes : l'unicité serait complètement fausse sans cette condition. Et par ailleurs, ce sont uniquement les polynômes M_j qui sont uniques, et pas la décomposition elle-même de V en sous-espaces cycliques.

8 Invariants de similitude

Soit f un endomorphisme de V, et soit θ un isomorphisme du K-espace vectoriel V. Soit $g = \theta \circ f \circ \theta^{-1}$ conjugué à f. Si $V = V_1 \oplus V_2 \oplus \cdots \oplus V_p$ est une décomposition en espaces cycliques pour f alors $V = \theta(V_1) \oplus \cdots \oplus \theta(V_p)$ est une décomposition en espaces cycliques pour g. En effet pour chaque V_j, on a $\theta : V_j \simeq \theta(V_j)$ avec $g(\theta(x)) = \theta(f(x))$, donc $\theta(V_j)$ est aussi cyclique pour g avec comme générateur $\theta(x_j)$ si x_j est générateur pour f de V_j, et les polynômes minimaux sont les mêmes. Donc le nombre p et les polynômes $M_1, \ldots, M_p$ sont les mêmes pour f et g : ce sont des invariants de similitude.

Réciproquement si f et g donnent lieu à des décompositions de V aux mêmes polynômes minimaux associés, alors

$$\phi_1, \phi_2 : V \simeq K[X]/(M_1) \oplus \cdots \oplus K[X]/(M_p) \tag{21}$$

avec $\phi_1 \circ f = X \circ \phi_1$, $\phi_2 \circ g = X \circ \phi_2$, donc avec $\theta = \phi_2^{-1} \circ \phi_1$ on a :

$$\theta \circ f \circ \theta^{-1} = \phi_2^{-1} \circ X \circ \phi_2 = g \tag{22}$$

L'endomorphisme f est donc déterminé à similitude près par la connaissance des polynômes minimaux $M_1, \ldots, M_p$.

Si l'on passe après choix d'une base de V à des matrices $N \times N$ à coefficients dans K, on obtient le théorème suivant :

Théorème 6. Soit K un corps. Toute matrice carrée A à coefficients dans K est semblable à une matrice composée de blocs diagonaux qui sont des matrices compagnons $C(M_1), \ldots, C(M_p)$ avec M_{j+1} diviseur de M_j :

$$\exists S \qquad SAS^{-1} = \begin{pmatrix} C(M_1) & 0 & \ldots & \ldots & 0 \\ 0 & C(M_2) & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & 0 & C(M_p) \end{pmatrix} \tag{23}$$

De plus deux telles matrices « réduites », c'est-à-dire sous cette forme diagonale par blocs, avec M_{j+1} diviseur de M_j, ne peuvent être semblables que si elles sont identiques. Le polynôme unitaire M_1 est le polynôme minimal de la matrice A, et le produit $M_1 \dots M_p$ est le polynôme caractéristique $\det(X - A)$.

Terminologie : les polynômes M_j sont les **facteurs invariants** de la matrice A (ou de l'endomorphisme f). La décomposition de l'espace vectoriel V en somme directe d'espace cycliques (avec condition de divisibilité sur les polynômes minimaux) ou la réduction de la matrice A s'appelle la **Décomposition de Frobenius.** Je donnerai plus loin des « formules » pour les facteurs invariants, mais surtout un algorithme de calcul, plus efficace.

9 Quelques conséquences du Théorème de Décomposition de Frobenius

Proposition 5. Si deux matrices carrées A et B à coefficients dans un corps K sont semblables sur une extension L du corps K alors elles sont semblables sur K.

Preuve : on peut supposer A et B réduites sur K au sens du Théorème précédent. Si l'on remplace K par un sur-corps L, elles restent réduites en tant que matrices à coefficients dans L. Donc si elles sont semblables sur L elles sont identiques. Donc les matrices d'origine étaient semblables sur K.

Exercice : prouvez-le directement lorsque $L = \mathbf{C}$ et $K = \mathbf{R}$.

Proposition 6. Soit K un corps. Toute matrice carrée à coefficients dans K est semblable (sur K) à sa transposée.

Preuve : il suffit de le montrer pour une matrice compagnon C(P), $n = \deg P \geq 1$. Soit $V = K[X]/(P)$, et la base $e_1 = 1$, $e_2 = X$, ..., $e_n = X^{n-1}$. La matrice C(P) est la matrice dans cette base de la multiplication f par X dans V. Considérons f^* agissant sur le dual V^*. Soit T un polynôme. Il est clair que $T(f^*)$ est l'adjoint de T(f). Donc $T(f^*)$ est nul si et seulement si T(f) est nul, donc le polynôme minimal de f^* est celui de f, à savoir P. Il existe par notre théorème général une forme linéaire L avec $M_L = P$ (en fait il suffit de prendre par exemple le L tel que $L(e_n) = 1$ et tous les autres nuls) et le sous-espace cyclique engendré par L est déjà de dimension égale à celle de V^*, donc en fait V^* est cyclique pour f^*, de polynôme minimal P. Donc pour une certaine base de V^* la matrice de f^* est aussi C(P). Par contre, on sait d'une manière générale que si l'on prend comme base de V^* la base duale de $(e_1, \dots, e_n)$ alors la matrice représentant f^* est la transposée de la matrice C(P) qui représentait f. Donc C(P) est semblable à sa transposée.

10 Magie noire avec les matrices compagnons

Soit $P = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ un polynôme unitaire de degré au moins 1, et soit $C(P)$ sa matrice compagnon :

$$C(P) = \begin{pmatrix} 0 & 0 & 0 & \cdots & \cdots & -a_0 \\ 1 & 0 & 0 & \cdots & \cdots & -a_1 \\ 0 & 1 & 0 & \cdots & \cdots & -a_2 \\ \cdots & \cdots & \ddots & \ddots & \cdots & \cdots \\ 0 & \cdots & \cdots & 1 & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} \tag{24}$$

On sait que $\det(X - C(P)) = P$, et on va préciser cet énoncé d'une manière assez remarquable.

Théorème 7. Il existe deux matrices $n \times n$ U_P et V_P, à coefficients dans l'anneau $K[X]$, de déterminants 1, et telle que l'identité suivante soit vérifiée :

$$U_P \cdot (X - C(P)) \cdot V_P = \begin{pmatrix} P & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} \tag{25}$$

La matrice de droite est diagonale avec $P, 1, \ldots, 1$ sur la diagonale. En particulier cela redonne $\det(X - C(P)) = P$.

Preuve : tout d'abord si l'on a une matrice (carrée ou rectangulaire) avec des lignes $L_1, \ldots, L_n$, alors remplacer L_i par $L_i + tL_j$ ($j \neq i$) se réalise en multipliant à gauche par la matrice carrée $n \times n$ avec des 1 sur la diagonale et comme unique autre entrée non nulle un t à l'intersection de la i^e ligne et de la j^e colonne. De même si les colonnes sont $C_1, \ldots, C_m$, remplacer C_i par $C_i + tC_j$ se réalise en multipliant à droite par la matrice carrée $n \times n$ avec des 1 sur la diagonale et comme unique autre entrée non nulle un t à l'intersection de la i^e colonne et de la j^e ligne. Regardons la matrice $X - C(P)$:

$$X - C(P) = \begin{pmatrix} X & 0 & \cdots & \cdots & a_0 \\ -1 & X & 0 & \cdots & a_1 \\ 0 & -1 & X & \cdots & a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & -1 & X & a_{n-2} \\ 0 & 0 & 0 & -1 & X + a_{n-1} \end{pmatrix} \tag{26}$$

On fait $L_{n-1} \to L_{n-1} + XL_n$, puis $L_{n-2} \to L_{n-2} + XL_{n-1}$, puis etc...et enfin $L_1 \to L_1 + XL_2$ (à ce stade L_2 est une ligne commençant par -1 puis il y a

$n-2$ zéros et enfin un certain polynôme à la fin). Ceci correspond à multiplier $X - C(P)$ par la gauche par une certaine matrice (triangulaire supérieure avec des 1 sur la diagonale) U_P et donne comme résultat :

$$U_P \cdot (X - C(P)) = \begin{pmatrix} 0 & 0 & \cdots & \cdots & P \\ -1 & 0 & 0 & \cdots & *_1 \\ & \ddots & \ddots & \cdots & \cdots \\ & & -1 & 0 & *_{n-2} \\ & & & -1 & X + a_{n-1} \end{pmatrix} \tag{27}$$

(28)

$$\Longrightarrow U_P \cdot (X - C(P)) \cdot \begin{pmatrix} 1 & 0 & \cdots & \cdots & *_1 \\ & 1 & 0 & \cdots & *_2 \\ & & \ddots & \ddots & \vdots \\ & & & 1 & X + a_{n-1} \\ & & & & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & \cdots & P \\ -1 & 0 & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \vdots \\ & & -1 & 0 & 0 \\ & & & -1 & 0 \end{pmatrix}$$

Il ne reste plus qu'à multiplier à droite par la matrice (de déterminant 1) :

$$\begin{pmatrix} 0 & -1 & & & \\ & 0 & -1 & & \\ & & \ddots & \ddots & \\ & & & 0 & -1 \\ 1 & & & & 0 \end{pmatrix} \tag{29}$$

Théorème 8. Soit A une matrice carrée sur un corps K, de taille $N \times N$. Soit M_1, ..., M_p ses facteurs invariants (de Frobenius) et posons $M_{p+1} = \cdots = M_N = 1$ (on a nécessairement $p \leq N$, car p est le nombre de blocs compagnons dans la réduction de A à la forme de Frobenius). Il existe des matrices U et V à coefficients dans l'anneau $K[X]$, de déterminants des scalaires non nuls, telles que

$$U \cdot (X - A) \cdot V = \mathrm{diag}(M_1, M_2, \ldots, M_N) \tag{30}$$

Preuve : il existe S inversible à coefficents dans K avec $S^{-1}AS$ égale à une suite diagonale de blocs compagnons. Ensuite on applique le théorème précédent à chaque bloc : en assemblant tout par blocs le long de la diagonale on a presque ce que l'on cherche, sauf qu'il ne reste plus qu'à conjuguer le tout par une matrice de permutation afin de mettre les 1 à la fin, et d'aligner les M_1, ..., M_p à la suite les uns des autres au début de la diagonale.

11 Une formule théorique pour les facteurs invariants

Théorème 9. Soit A une matrice carrée de taille $N \times N$ à coefficients dans un corps K. Soit M_1, ..., M_p ses facteurs invariants (de Frobenius) et posons $M_{p+1} =$

$\cdots = M_N = 1$. **Alors, pour $1 \leq j \leq N$, le produit**

$$M_j M_{j+1} \dots M_N \tag{31}$$

est le PGCD des mineurs de taille $N - j + 1$ de la matrice $X - A$ à coefficients dans l'anneau $K[X]$.

On peut donc obtenir le polynôme minimal M_1, en calculant les N^2 mineurs de taille $N - 1$ de la matrice $X - A$, puis leur PGCD, disons D et alors

$$M_1 = \frac{\det(X - A)}{D} \tag{32}$$

Sûrement pas l'algorithme le moins gourmand en calculs !

Preuve du Théorème : soit $\mathcal{A}$ un anneau commutatif quelconque. Si A est une matrice $n \times m$, notons $J_k(A)$ l'idéal de $\mathcal{A}$ engendré par les $\binom{n}{k}\binom{m}{k}$ mineurs de taille $k \times k$. Je dis que $J_k(BA) \subset J_k(A)$ pour toute matrice B (de taille $p \times n$). En effet chaque ligne de BA est une combinaison linéaire des lignes de A, donc par multilinéarité du déterminant, tout mineur de taille k de BA est une combinaison linéaire des mineurs de taille k de A. En particulier si B est une matrice carrée inversible (c'est-à-dire de déterminant une unité de $\mathcal{A}$), on a $J_k(BA) = J_k(A)$ car $A = B^{-1}BA$. De même pour la multiplication à droite, qui s'interprète comme des combinaisons de colonnes. Nous appliquons ceci avec $\mathcal{A} = K[X]$, et à l'identité du Théorème 8 :

$$U \cdot (X - A) \cdot V = \mathrm{diag}(M_1, M_2, \dots, M_N) \tag{33}$$

où U et V sont, comme il y est dit, des matrices inversibles à coefficients dans $\mathcal{A}$. Maintenant, lorsque l'on prend un mineur d'une matrice diagonale, déterminant de l'intersection de k lignes par k colonnes, si l'on retient la ligne i, il faut aussi retenir la colonne i, car sinon on aurait une ligne nulle. Donc les seuls mineurs non nuls sont centrés sur la diagonale. Finalement, compte tenu des relations de divisibilité $M_N | \cdots | M_1$, le « plus petit » (au sens de la divisibilité) mineur de taille k de $\mathrm{diag}(M_1, M_2, \dots, M_N)$ est celui en bas à droite, qui vaut $M_{N-k+1} \cdots M_N$. Fin de la preuve.

12 Un algorithme pour le calcul des facteurs invariants

Supposons que l'on ait une identité

$$U' \cdot (X - A) \cdot V' = \mathrm{diag}(M'_1, M'_2, \dots, M'_N) \tag{34}$$

avec des matrices U' et V' à coefficients dans $K[X]$, de déterminants scalaires et non nuls, et des polynômes M'_j vérifiant $M'_N | \cdots | M'_1$. Alors $M_j = M'_j$. En

effet, la preuve du théorème précédent s'applique à l'identique et montre que les produits $M'_j M'_{j+1} \dots M'_N$, donc les M'_j ne dépendent que de la matrice A.

Il suffit donc de réduire $X - A$ à une telle forme diagonale, par des combinaisons réversibles de lignes, de colonnes, et des permutations de lignes et de colonnes. Prenons comme point de départ n'importe quelle matrice P à coefficients dans l'anneau euclidien $K[X]$. Regardons les entrées de la première ligne et de la première colonne. Si elles ne sont pas toutes nulles, il existe une entrée non nulle de degré minimal. Par des permutations soit de lignes soit de colonnes on la met en position $(1,1)$. Puis on applique l'algorithme de division euclidienne à chacune des autres entrées de la première ligne et de la première colonne pour les remplacer par leurs restes. S'il subsiste un reste non nul, on recommence, et ainsi de suite. En un nombre fini d'étapes, on les a toutes réduites à zéro, sauf peut-être l'entrée en position $(1,1)$. On recommence ensuite à partir de $(2,2)$. Etc...En un nombre fini d'étapes on a réduit la matrice à une autre qui n'a d'entrées non nulles que sur la diagonale principale. Par des permutations on peut mettre les zéros à la fin. Il reste à se débrouiller pour obtenir les relations de divisibilité : il faut que l'entrée en $(1,1)$ soit multiple de toutes les autres non nulles, etc.... C'est là où intervient la **formule magique** suivante, valable dans tout anneau commutatif intègre (les coefficients sont dans son corps des fractions) :

$$(35) \qquad 0 \neq d = an + bm \implies \begin{pmatrix} \frac{m}{d} & \frac{n}{d} \\ -a & b \end{pmatrix} \begin{pmatrix} n & 0 \\ 0 & m \end{pmatrix} \begin{pmatrix} b\frac{m}{d} & -1 \\ a\frac{n}{d} & 1 \end{pmatrix} = \begin{pmatrix} \frac{mn}{d} & 0 \\ 0 & d \end{pmatrix}$$

Supposons en particulier que $n = N$ et $m = M$ soient des polynômes non nuls, et $D = AN + BM$ leur PGCD donné par une identité de Bezout. Alors on a réduit $\begin{pmatrix} N & 0 \\ 0 & M \end{pmatrix}$ à $\begin{pmatrix} \mathrm{ppcm}(N,M) & 0 \\ 0 & \mathrm{pgcd}(N,M) \end{pmatrix}$ en la multipliant à gauche et à droite par des matrices de déterminants 1. En appliquant ceci aux entrées $(1,1)$ et $(2,2)$ puis $(1,1)$ et $(3,3)$ etc..., on met le PPCM (des termes non nuls) en $(1,1)$, puis on recommence à partir de $(2,2)$, etc... La réduction de la matrice P de $K[X]$ est achevée.

En appliquant cet algorithme à la matrice $X - A$ on obtient au final les invariants de Frobenius M_1, ..., M_p de la matrice A.

13 Décomposition de Jordan d'un endomorphisme nilpotent

Dans le cas d'un endomorphisme nilpotent N, le polynôme minimal et tous les facteurs invariants (qui en sont des diviseurs) sont des monômes X^k. La matrice compagnon $C(X^k)$ est conjuguée à sa transposée (il suffit d'ailleurs de

passer à la base $(X^{k-1}, \dots, 1)$ de $K[X]/(X^k)$) :

$$C(X^k)^t = J_k := \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \tag{36}$$

Donc la décomposition de Frobenius dans ce cas donne une décomposition en blocs de Jordan, dont le nombre et les tailles sont des invariants de similitude de l'endomorphisme nilpotent N. Le polynôme minimal est X^K avec K la taille du plus grand bloc, et on a $M_1 = M_2 = \cdots = M_q$ avec q le nombre de fois qu'il y a un bloc de taille K, puis M_{q+1} est X^L avec L la taille maximale de ceux qui restent etc...

14 Décomposition en espaces caractéristiques d'un endomorphisme scindé

Théorème 10. Le polynôme minimal M_f et le polynôme caractéristique $P_f = \det(X - f)$ ont les mêmes zéros dans K. L'un est scindé sur K si et seulement si l'autre l'est aussi.

Preuve : par Cayley-Hamilton $M_f | P_f$ donc tout zéro de M_f dans K est un zéro de P_f. Réciproquement si $P_f(\lambda) = 0$ alors λ est valeur propre donc il existe x non nul avec $f(x) = \lambda x$. Le polynôme minimal de x est $X - \lambda$. Donc M_f est un multiple de $X - \lambda$, autrement dit $M_f(\lambda) = 0$. Ceci montre que M_f et P_f ont les mêmes zéros dans K.

Si P_f est scindé, M_f qui en est un diviseur le sera aussi. Réciproquement, si M_f est scindé, les autres facteurs invariants M_2, ..., M_p qui en sont des diviseurs le sont aussi donc le produit $P_f = M_1 \cdots M_p$ est scindé.

On peut montrer (M_f scindé $\implies$ $\det(X - f)$ scindé) sans avoir à invoquer la théorie des facteurs invariants. Par exemple, par récurrence sur la dimension de la manière suivante. Soit x un vecteur non nul et V_x l'espace cyclique qu'il engendre. Notons W le quotient V/V_x et g l'endomorphisme quotient. On a

$$\det(X - f) = \det{}_{V_x}(X - f) \det(X - g) \tag{37}$$

Le polynôme minimal de g divise celui de f donc est scindé, et par hypothèse de récurrence on a donc $\det(X - g)$ scindé. Par ailleurs on sait que $\det_{V_x}(X - f)$ est le polynôme minimal de x et donc lui aussi divise M_f et est scindé. Donc $\det(X - f)$ est scindé. Le Théorème est démontré.

On dit qu'un endomorphisme f est scindé sur le corps K si son polynôme minimal M_f (ou son polynôme caractéristique) est scindé :

$$M_f = \prod (X - \lambda_j)^{m_j}, \qquad \lambda_j \in K, m_j > 0 \tag{38}$$

$$P_f = \prod (X - \lambda_j)^{n_j}, \qquad \lambda_j \in K, n_j \geq m_j \tag{39}$$

Reprenons maintenant le raisonnement fait dans notre toute première section. À la valeur propre λ on peut associer l'espace caractéristique V_λ, par exemple par la formule

$$V_\lambda = \mathrm{Ker}(f - \lambda)^{\dim(V)} \tag{40}$$

En effet, la chaîne $K_n = \mathrm{Ker}(f - \lambda)^n$ vérifie $\dim K_n \geq n$ tant qu'elle est strictement croissante, donc il y a un indice $n \leq \dim(V)$ avec $K_n = K_{n+1}$ et il est facile de voir alors $K_n = K_{n+1} = K_{n+2} = \ldots$. L'espace V_λ possède, comme nous l'avons vu, un supplémentaire canonique stable par f qui est $W = \mathrm{Im}(f-\lambda)^{\dim(V)}$. Sur ce W, λ n'est plus valeur propre. Le polynôme minimal (ou caractéristique) de f sur W divise celui sur V donc est scindé, et on peut raisonner par récurrence, car il y aura une nouvelle valeur propre pour continuer. Ainsi :

Théorème 11. À tout endomorphisme scindé f sur V est associé une décomposition de V en la somme directe des espaces caractéristiques $V_\lambda = \mathrm{Ker}(f - \lambda)^{\dim V}$ associés aux valeurs propres de f.

Notez qu'à la différence de la décomposition de Frobenius de V en somme directe d'espaces cycliques, il y a ici une unicité au niveau même de la décomposition vectorielle.

Même si f n'est pas scindé, on peut tout de même faire le raisonnement en utilisant les valeurs propres disponibles. En effet soit λ une première valeur propre, on a, en notant m_λ et n_λ les multiplicités de λ dans M_f et P_f, et par notre raisonnement antérieur :

$$V = V_\lambda \oplus W \qquad W = \mathrm{Im}(f - \lambda)^{m_\lambda} = \mathrm{Im}(f - \lambda)^{n_\lambda} \tag{41}$$

Le polynôme caractéristique $P_f = \det(X - f)$ vaut $(X - \lambda)^{n_\lambda} \cdot \det_W(X - f)$. Or $\mathrm{Ker}(f - \lambda) \cap W = \{0\}$ donc λ n'est pas valeur propre de f sur W, donc λ n'est pas racine de $\det_W(X - f)$. Par contre toute autre racine de P_f est racine de $\det_W(X - f)$. On peut itérer et au final on a une décomposition en sous-espaces stables par f :

$$V = \oplus_\lambda V_\lambda \oplus W \qquad W = \mathrm{Im} \prod (f - \lambda)^{m_\lambda} \tag{42}$$

La formule $W = \mathrm{Im} \prod (f - \lambda)^{m_\lambda}$ demande une petite justification que je laisse en exercice.

Il est possible de caractériser intrinsèquement l'espace W :

Théorème 12. Tout sous-espace $Z \subset V$ stable par f et sur lequel f n'a aucune valeur propre est inclus dans W, qui est donc le plus grand avec cette propriété.

Preuve : comme Z est stable par f il est stable par $T(f) = \prod (f-\lambda)^{m_\lambda}$. Comme chaque $f - \lambda$ est injectif sur Z, le produit $T(f)$ est injectif, donc bijectif sur Z. Ainsi $Z = T(f)(Z) \subset W$.

15 Décomposition de Dunford d'un endomorphisme scindé

C'est une version un peu plus abstraite de la décomposition

$$V = \oplus V_\lambda \tag{43}$$

en sous-espaces caractéristiques. Notons d l'endomorphisme qui vaut λId sur V_λ. Évidemment d est diagonalisable, $f = d + n$ et n est nilpotent. De plus d et f commutent (évident puisqu'il suffit de le vérifier sur chaque V_λ ceux-ci étant stables par d et par f). Donc d et n commutent. On a le complément suivant, dit Théorème de Dunford :

Théorème 13. Soit f un endomorphisme scindé. Il existe une unique décomposition $f = d + n$ avec d diagonalisable, n nilpotent et $dn = nd$. De plus d et n sont des polynômes en f.

Preuve : on a montré l'existence (on verra après pour la dernière assertion). Supposons $f = d' + n'$ avec $d'n' = n'd'$. Donc $d'f = fd'$. Soit μ une valeur propre de d'. Ainsi f laisse stable l'espace propre $W_\mu = \mathrm{Ker}(d' - \mu)$. Mais $f - d'$ restreint à cet espace est à la fois n' et $f - \mu$. Donc $f - \mu$ est nilpotent sur W_μ donc, d'abord μ est une valeur propre λ de f et ensuite $W_\mu \subset V_\lambda$. Comme d' est diagonalisable, V est la somme directe des W_μ, et comme on sait déjà $V = \oplus V_\lambda$ la seule possibilité est que $W_\mu = V_\mu$ et que toutes les valeurs propres de f sont aussi valeurs propres de d'. Ainsi $d' = d$.

Pour montrer que d (et par conséquent n) est un polynôme en f, il suffit d'utiliser le théorème des restes chinois. Posons $N = \dim V$ et soit $\lambda_1, \ldots, \lambda_p$ les valeurs propres de f.

$$\exists P \in K[X], \forall j, \qquad P \equiv \lambda_j \mod (X - \lambda_j)^N \tag{44}$$

Sur V_λ on a $(f - \lambda)^N = 0$ donc $P(f) = \lambda \mathrm{Id}_{V_\lambda} = d$. Donc $P(f) = d$ sur V tout entier. Fin de la preuve.

Comme corollaire, notons que tout endomorphisme g qui commute avec f commutera avec d et n puisque ceux-ci sont des polynômes en f.

16 Endomorphismes diagonalisables

On s'inspire d'une partie de notre dernière démonstration pour prouver :

Théorème 14. La restriction d'un endomorphisme diagonalisable f à un sous-espace stable $W \subset V$ est diagonalisable.

Preuve : soit $\lambda_1, \ldots, \lambda_p$ les valeurs propres distinctes de f sur V et $V_i = \mathrm{Ker}(f - \lambda_i)$ les espaces propres. Considérons les polynômes :

$$P_i = \frac{\prod_{j \neq i}(X - \lambda_j)}{\prod_{j \neq i}(\lambda_i - \lambda_j)} \tag{45}$$

En fait $P_i(f)$ est la projection sur V_i parallèlement à $\oplus_{j \neq i} V_j$ puisque $P_i(f)$ s'annule sur les V_j, $j \neq i$ et est l'identité sur V_i. Comme W est stable par f il est stable par $P_i(f)$. Il en résulte l'identité

$$W \cap V_i = P_i(f)(W) \tag{46}$$

Comme tout vecteur s'écrit

$$x = P_1(f)(x) + \cdots + P_p(f)(x) \tag{47}$$

on obtient la décomposition :

$$W = (W \cap V_1) \oplus (W \cap V_2) \oplus \cdots \oplus (W \cap V_p) \tag{48}$$

Donc la restriction de f à W est diagonalisable. De plus en prenant pour chaque j un complémentaire dans V_j de $W \cap V_j$, on obtient aussi :

Théorème 15. Tout sous-espace $W \subset V$ qui est stable sous f, avec f un endomorphisme diagonalisable de V, possède un complémentaire stable sous f.

Il est possible de caractériser les endomorphismes diagonalisables par leur polynôme minimal :

Théorème 16. Un endomorphisme f est diagonalisable si et seulement si son polynôme minimal M_f est scindé à racines simples sur le corps K.

Preuve : si f est diagonalisable il est clair que $M_f = \prod_j (X - \lambda_j)$, le produit portant sur les valeurs propres de f. Réciproquement, si M_f est scindé on peut utiliser la théorie expliquée précédemment de la décomposition $V = \oplus V_\lambda$, et comme $f - \mu$ est inversible sur V_λ pour tout $\mu \neq \lambda$, de $M_f(f) = 0$ résulte $f - \lambda = 0$ sur V_λ. Donc f est diagonalisable.

17 Endomorphismes semi-simples

On dit que f est semi-simple si tout sous-espace $W \subset V$ stable par f admet un complémentaire stable par f. De même une matrice carrée A est dite semi-simple si l'endomorphisme associé de K^N est semi-simple.

Théorème 17. Un endomorphisme f est semi-simple si et seulement si son polynôme minimal M_f ne possède pas de facteur carré.

Preuve : supposons $M_f = \pi^2 N$ avec π un polynôme unitaire irréductible (non constant ; mais je pense que c'est la convention usuelle sur les irréductibles). Notons $W = \mathrm{Ker}\,\pi(f)$. Si f était semi-simple, il existerait un complémentaire stable Z. Sur Z, $\pi(f)$ est injectif donc de $M_f(f) = 0$ résulte $N(f) = 0$. Donc $\pi(f)N(f)$ est nul sur W et sur Z. Donc πN annule f, contradiction.

Supposons maintenant au contraire que M_f est le produit $\pi_1 \dots \pi_q$ de polynômes unitaires irréductibles distincts. Ils sont premiers deux à deux, et on peut utiliser le théorème des restes chinois. Il existe des polynômes P_1, ..., P_q tels que :

$$P_i \equiv 1 \mod \pi_i \qquad \text{et } \forall j \neq i \; P_i \equiv 0 \mod \pi_j \tag{49}$$

Sur $\mathrm{Ker}\,\pi_j(f)$ on a $P_i(f) = 0$ pour $i \neq j$, par contre $P_i(f) = \mathrm{Id}$ sur $\mathrm{Ker}\,\pi_i(f)$. Supposons qu'on ait une relation

$$0 = x_1 + \cdots + x_q \qquad \text{avec } \forall i \; x_i \in \mathrm{Ker}\,\pi_i(f) \tag{50}$$

En appliquant $P_i(f)$ on obtient $x_i = 0$. Donc les $\mathrm{Ker}\,\pi_i(f)$ sont en somme directe. Soit $P = P_1 + \cdots + P_q$. On a $P \equiv 1 \mod \pi_i$ pour tout i donc $P - 1$ est un multiple de M_f, donc $P(f) = \mathrm{Id}$ et

$$\forall x \in V \quad x = x_1 + \cdots + x_q \qquad \text{avec } \forall i \; x_i = P_i(f)(x) \tag{51}$$

Il est clair que $\pi_i P_i$ est nul modulo π_j pour tous les $j = 1, \dots, q$ donc M_f divise $\pi_i P_i$, donc $\pi_i(f)P_i(f)(x) = 0$ pour tout x et par conséquent $P_i(f)(x) \in \mathrm{Ker}\,\pi_i(f)$. Nous avons donc prouvé la décomposition en somme directe

$$V = \oplus_i \mathrm{Ker}\,\pi_i(f) \tag{52}$$

et le fait que $P_i(f)$ est la projection sur $\mathrm{Ker}\,\pi_i(f)$ parallèlement aux autres. Soit enfin $W \subset V$ stable par f. Alors W est stable par $P_i(f)$ donc $W \cap \mathrm{Ker}\,\pi_i(f) = P_i(f)(W)$, ce qui montre que les $P_i(f)(W)$ sont en somme directe. Comme $\mathrm{Id} = \sum_i P_i(f)$, on a

$$W = \oplus_i (W \cap \mathrm{Ker}\,\pi_i(f)) \tag{53}$$

et pour le problème de trouver un complémentaire Z de W stable par f, on s'est donc ramené au cas où le polynôme minimal de f est un irréductible π. On

sait dans ce cas que l'anneau $L = K[X]/(\pi)$ est un corps. Définissons une loi de multiplication externe par L sur V par les formules

$$\lambda \cdot x = Q(f)(x) \qquad \text{avec } \lambda \equiv Q \mod \pi \tag{54}$$

Ceci est bien défini puisque $\pi(f)(x) = 0$. Il est clair que $1 \cdot x = x$, que $(\lambda - \mu) \cdot x = \lambda \cdot x - \mu \cdot x$, $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$, $\lambda \cdot (x - y) = \lambda \cdot x - \lambda \cdot y$, etc... Donc cela fait de V un L-espace vectoriel (et la multiplication externe par L est compatible à celle par K). Tout sous K-espace vectoriel $W \subset V$ stable par f est en fait un sous L-espace vectoriel. On prend Z un L-espace vectoriel complémentaire de W dans V. Ce Z est donc stable par f et la décomposition $V = W \oplus Z$ vaut aussi au sens des K-espaces vectoriels. Ceci termine la preuve de l'implication (si M_f n'a pas de facteur carré alors f est semi-simple) et donc la preuve du Théorème.

Théorème 18. Soit K un corps **de caractéristique nulle** et A une matrice carrée de taille $N \geq 1$ à coefficients dans K. Les assertions suivantes sont équivalentes :

1. A est semi-simple,
2. il existe un corps $L \supset K$ sur lequel A est diagonalisable,
3. A est diagonalisable sur la clôture algébrique $\overline{K}$ de K,
4. le polynôme minimal de A n'a que des racines simples dans $\overline{K}$.

Preuve : soit M le polynôme minimal de A. On a prouvé que A est semi-simple si et seulement si $M = \prod_i \pi_i$ avec des polynômes irréductibles distincts. Pour $i \neq j$ on a une identité de Bezout $A\pi_i + B\pi_j = 1$ qui prouve que π_i et π_j ne peuvent avoir de racine commune dans aucun corps $L \supset K$. Et par ailleurs, comme K est de caractéristique nulle, on a $\deg \pi_i' = \deg \pi_i - 1$, $\pi_i' \neq 0$. Donc $\operatorname{pgcd}(\pi_i, \pi_i') = 1$ et π_i ne peut avoir, par le même argument, de racine commune avec π_i' dans aucun corps $L \supset K$, autrement dit il ne peut y avoir dans L que des racines simples. Donc, si A est semi-simple, son polynôme minimal ne peut avoir que des racines simples dans tout corps $L \supset K$, en particulier dans la clôture algébrique $\overline{K}$. On a donc par nos théorèmes précédents (1) $\Longrightarrow$ (4) $\Longrightarrow$ (3) $\Longrightarrow$ (2). Si A est diagonalisable sur L son polynôme minimal M (dont on sait qu'il ne dépend pas du corps) est scindé à racines simples sur L. On remplace L par le sous-corps L_0 extension finie de K engendrée par ces racines, et on sait qu'il existe un K plongement $L_0 \to \overline{K}$, donc M est scindé sur $\overline{K}$ à racines simples. Il ne peut donc pas y avoir de facteur carré dans sa décomposition en polynômes irréductibles de $K[X]$, et par conséquent A est semi-simple, ce qui établit (2) $\Longrightarrow$ (1).

Note : en caractéristique positive notre propriété de « semi-simplicité » n'est pas stable par extension de corps. Soit $K = \mathbf{Z}/2\mathbf{Z}(t)$ et $A = \left(\begin{smallmatrix} 0 & t \\ 1 & 0 \end{smallmatrix}\right)$. Elle est semi-simple sur K mais pas sur $\overline{K}$ (et n'est diagonalisable sur aucun sur-corps L de K).

Exercice. Montrer que la décomposition de Dunford sur $\mathbf{C}$ de tout $A \in \mathrm{Mat}_n(\mathbf{R})$ est définie sur $\mathbf{R}$: $A = D + N$ avec $D, N \in \mathrm{Mat}_n(\mathbf{R})$, $DN = ND$, N nilpotente, et D semi-simple.

Fin par épuisement des combattants. Le fameux Lemme des Noyaux manque mais on a fait à plusieurs reprises des preuves (à chaque fois qu'on a invoqué le théorème des restes chinois) de cas particuliers, et la méthode employée marche pour l'énoncé plus général.