ALGÈBRE - LEÇON 142 : PGCD ET PPCM, ALGORITHMES DE CALCUL. APPLICATIONS

SIMON RICHE

1. Commentaires du jury (rapport 2024)

Le candidat doit prendre soin de différencier le cadre théorique des anneaux factoriels ou principaux dans lequel sont définis les PGCD et PPCM et dans lequel s'appliquent les énoncés des théorèmes proposés et le cadre euclidien fournissant les algorithmes. Le champ d'étude de cette leçon ne peut se limiter au cas de $\mathbf{Z}$, mais la leçon peut opportunément s'illustrer d'exemples élémentaires d'anneaux euclidiens, comme $\mathbf{Z}$ et $\mathbf{K}[X]$.

Une part substantielle de la leçon doit être consacrée à la présentation d'algorithmes : algorithme d'Euclide, algorithme binaire, algorithme d'Euclide étendu. Il est possible d'en évaluer le nombre d'étapes dans les pires cas et faire le lien avec les suites de Fibonacci.

Des applications élémentaires sont particulièrement bienvenues : calcul de relations de Bezout, résolutions d'équations diophantiennes linéaires, inversion modulo un entier ou un polynôme, calculs d'inverses dans les corps de ruptures, les corps finis. On peut aussi évoquer le théorème chinois effectif, la résolution d'un système de congruences et faire le lien avec l'interpolation de Lagrange.

Pour aller plus loin, on peut évoquer le rôle de l'algorithme d'Euclide étendu dans de nombreux algorithmes classiques en arithmétique (factorisation d'entiers, de polynômes, etc). Décrire l'approche matricielle de l'algorithme d'Euclide et l'action de $\mathrm{SL}_2(\mathbf{Z})$ sur $\mathbf{Z}^2$ est tout à fait pertinent. On peut aussi établir l'existence d'un supplémentaire d'une droite dans $\mathbf{Z}^2$, ou d'un hyperplan de $\mathbf{Z}^n$, examiner l'éventuelle possibilité de compléter un vecteur de $\mathbf{Z}^n$ en une base. On peut aussi étudier les matrices à coefficients dans un anneau principal ou euclidien, et, de manière plus avancée, la forme normale d'Hermite et son application à la résolution d'un système d'équations diophantiennes linéaires. De même, aborder la forme normale de Smith, et son application au théorème de la base adaptée, permet de faire le lien avec la réduction des endomorphismes *via* le théorème des invariants de similitude. La leçon invite aussi, pour des candidates et candidats familiers de ces notions, à décrire le calcul de PGCD dans $\mathbf{Z}[X]$ et $\mathbf{K}[X, Y]$, avec des applications à l'élimination de variables. On peut rappeler les relations entre PGCD et résultant et montrer comment obtenir le PGCD en échelonnant la matrice de Sylvester. Sur l'approximation diophantienne, on peut enfin envisager le développement d'un rationnel en fraction continue et l'obtention d'une approximation de Padé-Hermite à l'aide de l'algorithme d'Euclide, la recherche d'une relation de récurrence linéaire dans une suite ou le décodage des codes BCH.

Date: Année 2024–2025.

2. Plan

Cette leçon doit commencer par présenter le cadre théorique nécessaire à la considération des PGCD et PPCM (en se limitant au cas des anneaux factoriels, qui est largement suffisant à mon avis), qui se trouve dans de nombreuses références classiques, mais aussi présenter des aspects plus algorithmiques, pour lesquel on pourra consulter [SP].

2.1. **Ce qui doit apparaître.** Définition du PGCD et du PPCM dans un anneau factoriel.

Cas des anneaux principaux : caractérisation en termes d'idéaux et relation de Bezout.

Cas des anneaux euclidiens : calcul algorithmique (algorithme d'Euclide, algorithme d'Euclide étendu).

Application d'une relation de Bezout au calcul des inverses dans les $\mathbb{Z}/n\mathbb{Z}$, dans un corps de rupture.

Résolution d'une équation du type $ax + by = c$ dans $\mathbb{Z}$.

Théorème des restes chinois (effectif).

Application à la résolution de systèmes de congruences.

Contenu d'un polynôme.

Application : si A est factoriel alors $A[X]$ est factoriel.

Analyse de l'algorithme d'Euclide (étendu) [1] :
— présentation matricielle ;
— taille des coefficients de Bezout ;
— nombre maximal d'itérations [2] ;
— complexité.

2.2. **Ce qui peut apparaître.** Résultant de deux polynômes.

Application à l'élimination de variables pour les systèmes d'équations polynomiales. [3]

Matrices à coefficients dans un anneau euclidien.

Forme normale de Smith.

Application aux invariants de similitude.

Algorithme binaire pour le calcul du PGCD d'entiers. [4]

Calcul du PGCD dans un anneau de polynômes à coefficients dans un anneau factoriel en utilisant une pseudo-division euclidienne. [5]

---

1. Voir [SP, §III.4] ou `https://www.math.u-bordeaux.fr/~kbelabas/teach/Agreg/Euclide.pdf`.
2. Ce calcul s'appelle la *théorème de Lamé* ; voir par exemple [Sk, Chap. II, Ex. 2.7] ou [SP, Chap. III, Ex. 4].
3. Voir l'Exercice 2 (tiré de [SP]) ou `https://www.math.u-bordeaux.fr/~kbelabas/teach/Agreg/resultant.pdf`.
4. Voir [SP, Chap. III, Ex. 2, p. 58].
5. Voir [SP, §III.5].

3. Quelques questions bêtes auxquelles il faut absolument savoir répondre rapidement

(1) Si $\mathbb{K}$ est un corps et $\mathbb{L}$ une extension de $\mathbb{K}$, et si $P, Q \in \mathbb{K}[X]$, comparer les PGCD de P et Q dans $\mathbb{K}[X]$ et $\mathbb{L}[X]$. Que peut-on dire concernant le PPCM de P et Q ?

(2) Comment peut-on interpréter l'interpolation de Lagrange comme un système de congruences ?

(3) Pour n, m des entiers positifs, calculer $\text{pgcd}(2^n - 1, 2^m - 1)$.

(4) Soit A un anneau factoriel, de corps des fractions K, et soient $P, Q \in A[X]$. Comparer les PGCD de P et Q dans $A[X]$ et dans $K[X]$. (*Indication* : on pourra considérer la décomposition en produit d'irréductibles pour P et Q.) Montrer au passage que $c(\text{pgcd}(P, Q)) = \text{pgcd}(c(P), c(Q))$ à un inversible près. [6]

4. Exercices

Exercice 1 (Recherche de diviseurs dans $\mathbb{Z}[X]$)**.** Soit $P \in \mathbb{Z}[X]$.

(1) Montrer que si $a \in \mathbb{Z}$ et si $Q \in \mathbb{Z}[X]$ est un diviseur de P (dans $\mathbb{Z}[X]$), alors $Q(a) \mid P(a)$.

(2) Supposons que P est de degré $d \geq 2$, et notons $n = \lfloor d/2 \rfloor + 1$. Fixons n entiers distincts $a_1, \ldots, a_n$. Montrer que le polynôme P est réductible dans $\mathbb{Z}[X]$ si et seulement si il existe des diviseurs $d_1, \ldots, d_n$ de $P(a_1), \ldots, P(a_n)$ respectivement, ni tous égaux à 1 ni tous égaux à -1, tels que le polynôme d'interpolation de $(a_1, \ldots, a_n)$ et $(d_1, \ldots, d_n)$ (c'est-à-dire l'unique polynôme de degré au plus $n - 1$ valant d_i en a_i pour tout i) est à coefficients entiers et divise P.

(3) En déduire un algorithme permettant de tester si P est irréductible dans $\mathbb{Z}[X]$ et de donner un diviseur s'il ne l'est pas.

(4) Appliquer cet algorithme pour factoriser le polynôme $X^4 - X^2 - 2X - 1$ en produit de facteurs irréductibles. (On pourra par exemple choisir les points -1, 0 et 1.)

Référence : [SP, §V.1.3].

Exercice 2 (Résultant de deux polynômes et application à la résolution de systèmes polynomiaux)**.** Soit A un anneau intègre [7], d'anneau des fractions K. On rappelle que si P et Q sont des polynômes à coefficients dans A, de degrés [8] respectifs m et n tels que $n + m > 0$, en notant

$$P(X) = a_m X^m + \cdots + a_1 X + a_0, \quad Q(X) = b_n X^n + \cdots + b_1 X + b_0,$$

6. Pour cette égalité, voir aussi [SP, Proposition III.22].

7. Cette hypothèse est inutile dans les questions ne faisant pas intervenir K.

8. Dans cet exercice, et même si ce n'est pas la convention standard, on considèrera que le polynôme nul a pour degré 0.

la *matrice de Sylvester* de (P,Q) est la matrice carrée de taille $n+m$ définie par

$$\mathrm{Sylv}(P,Q) = \begin{pmatrix} a_m & a_{m-1} & \cdots & \cdots & a_0 & 0 & \cdots & \cdots \\ 0 & a_m & a_{m-1} & \cdots & \cdots & a_0 & 0 & \cdots \\ \vdots & \ddots & \ddots & & & & \ddots & \\ b_n & b_{m-1} & \cdots & b_0 & 0 & \cdots & \cdots & \\ 0 & b_n & b_{n-1} & \cdots & b_0 & 0 & \cdots & \\ \vdots & \ddots & \ddots & & & & \ddots & \end{pmatrix}$$

où le coefficient b_n de la première colonne apparait sur la ligne d'indice $n+1$, et que le *résultant* de (P,Q) est

$$\mathrm{Res}(P,Q) = \det(\mathrm{Sylv}(P,Q)).$$

(1) Montrer que si $a \in A$ et $Q \in A[X]$ est de degré non nul, alors $\mathrm{Res}(a,Q) = a^{\deg(Q)}$.

(2) Montrer que si $P, Q \in A[X]$ sont tels que $\deg(P) + \deg(Q) > 0$, alors $\mathrm{Res}(P,Q) = (-1)^{\deg(P)\deg(Q)}\mathrm{Res}(Q,P)$.

(3) Montrer que les lignes de $\mathrm{Sylv}(P,Q)$ sont les coefficients dans la base naturelle de $K_{n+m-1}[X]$ des polynômes successifs suivants :

$$X^{n-1}P, X^{n-2}P, \cdots, XP, P, X^{m-1}Q, X^{m-2}Q, \cdots, XQ, Q.$$

(4) On suppose que $m \geq n > 0$, et on note R le reste de la division euclidienne de P par Q dans $K[X]$. Montrer qu'avec les notations ci-dessus on a

$$\mathrm{Res}(P,Q) = (-1)^{nm} b_n^{m-\deg(R)} \mathrm{Res}(Q,R).$$

(5) En déduire un algorithme de calcul de $\mathrm{Res}(P,Q)$ basé sur les divisions euclidiennes.

(6) Montrer que P et Q sont premiers entre eux (dans $K[X]$) si et seulement si $\mathrm{Res}(P,Q) \neq 0$.

(7) Soit L un corps algébriquement clos contenant K. Notons $\alpha_1, \cdots, \alpha_m$ les m racines de P dans L (comptées avec multiplicité) et $\beta_1, \cdots, \beta_n$ les n racines de Q dans L. Montrer qu'avec les notations ci-dessus on on a

$$\mathrm{Res}(P,Q) = (a_m)^n (b_n)^m \prod_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} (\alpha_i - \beta_j).$$

(Dans le cas où n ou m est nul, on interprète le produit sur les i,j comme égal à 1.)

(*Indication* : on pourra remarquer que le terme de droite est égal à

$$(-1)^{mn}(b_n)^m \prod_{j=1}^{n} P(\beta_j),$$

puis vérifier que cet élement de A se calcule par le même algorithme que $\mathrm{Res}(P,Q)$.)

(8) Soit B un autre anneau intègre, et soit $\varphi : A \to B$ un morphisme d'anneaux. Supposons que $\deg(P) + \deg(\varphi(Q)) > 0$. Montrer que

$$\varphi(\mathrm{Res}(P,Q)) = (\varphi(a_m))^{n-\deg(\varphi(Q))} \mathrm{Res}(\varphi(P), \varphi(Q)).$$

(9) On suppose dans cette question que $A = \Bbbk[Y_1, \dots, Y_k]$ où $\Bbbk$ est un corps. Les polynômes P et Q peuvent alors s'interpréter comme des polynômes en les $k+1$ variables $Y_1, \dots, Y_k, X$ à coefficients dans $\Bbbk$, et $\mathrm{Res}(P,Q)$ comme un polynôme en $Y_1, \dots, Y_k$. Montrer que si $(\alpha_1, \dots, \alpha_k, \alpha)$ est un point de $\Bbbk^{k+1}$ tel que
$$P(\alpha_1, \dots, \alpha_k, \alpha) = Q(\alpha_1, \dots, \alpha_k, \alpha) = 0,$$
alors
$$\mathrm{Res}(P,Q)(\alpha_1, \dots, \alpha_k) = 0.$$

Référence : [SP, §VII.2–VII.3]. Remarquons que dans la procédure de la dernière question, on est passé d'équations polynomiales en $k+1$ variables à une équation en k variables. On a donc "éliminé une variable". Voir [SP] pour une discussion de cette méthode.

Pour des commentaires, exemples et discussions des applications du résultant, on pourra consulter les vidéos de Philippe Caldero : `https://www.youtube.com/channel/UCZ5bgGfyXy4nnPzV__uJ7Kg`. Pour des preuves différentes de la propriété (6), et d'autres applications, voir [FGN1, Ex. 4.13, "Résultant de deux polynômes"] ou [Go, Chap. 4, §5, Problème 6, p. 219].

Exercice 3 (Suite de Fibonacci)**.** On considère la suite de Fibonacci $(F_n)_{n\geq 0}$ définie en posant $F_0 = 0$, $F_1 = 1$, puis
$$F_{n+2} = F_{n+1} + F_n$$
pour tout $n \geq 0$.

(1) Montrer que pour tous $n \geq 0$ et $k \geq 2$ on a $F_{n+k} = F_k F_{n+1} + F_{k-1} F_n$. (*Indication* : on pourra procéder par récurrence sur k.)

(2) Montrer que si $m \geq n \geq 0$ alors $\mathrm{pgcd}(F_m, F_n) = \mathrm{pgcd}(F_{m-n}, F_n)$.

(3) En déduire que pour tous $m, n \geq 0$ on a $\mathrm{pgcd}(F_m, F_n) = F_{\mathrm{pgcd}(m,n)}$.

Référence : [SP, Chap. III, Ex. 6]. La suite de Fibonacci est importante du point de vue du calcul du PGCD car deux termes successifs de cette suite nécessitent le plus grand nombre d'étapes possible dans l'algorithme d'Euclide, dans un sens qui peut être rendu précis par le Théorème de Lamé.

5. Complément : algèbre linéaire sur les entiers (et les anneaux euclidiens)

Dans cette partie on présente quelques résultats concernant les matrices à coefficients dans des anneaux euclidiens, en se basant notamment sur les notes `https://agreg-maths.univ-rennes1.fr/documentation/docs/alglinent.pdf` de M. Coste et [Co, NQ].

5.1. **Modules.** Ci-dessous on va utiliser quelques notions de base concernant les modules sur un anneau [9] A. Cette notion n'étant pas au programme du concours il est prudent de ne pas aller trop loin dans des considérations délicates sur ces questions, mais il est quand même bien d'avoir quelques notions de base sur cet outil extrèmement utile.

9. Tous les anneaux considérés ci-dessous sont supposés commutatifs et unitaires.

Si A est un anneau et M est un A-module, on rappelle qu'une *base*[10] de M est une famille $(u_1, \dots, u_n)$ d'éléments de M telle que pour chaque $m \in M$ il existe un *unique* n-uplet $(a_1, \dots, a_n)$ d'éléments de A tels que

$$m = a_1 u_1 + \cdots + a_n u_n.$$

Comme dans le cadre des espaces vectoriels sur un corps, on vérifie facilement que la famille $(u_1, \dots, u_n)$ est une base si et seulement si elle vérifie les deux propriétés suivantes :

— elle est génératrice, c'est-à-dire que pour tout $m \in M$ il existe un n-uplet $(a_1, \dots, a_n)$ d'éléments de A tels que $m = a_1 u_1 + \cdots + a_n u_n$;

— elle est libre, c'est-à-dire que si $(a_1, \dots, a_n)$ est un n-uplet d'éléments de A tels que

$$a_1 u_1 + \cdots + a_n u_n = 0,$$

alors $a_1 = \cdots = a_n = 0$.

Une différence fondamentale avec l'algèbre linéaire sur un corps est qu'un module sur un anneau n'admet pas toujours une base. Cependant, s'il en admet une alors toutes les bases ont le même cardinal, comme expliqué dans le lemme suivant.

Lemme 1. Soit A un anneau intègre[11], et soit M un A-module. Supposons que les familles de vecteurs $(a_1, \dots, a_n)$ et $(a'_1, \dots, a'_m)$ sont deux bases de M. Écrivons pour tout $i \in \{1, \dots, n\}$

$$a_i = \sum_{j=1}^{m} \lambda_{i,j} a'_j$$

et pour tout $j \in \{1, \dots, m\}$

$$a'_j = \sum_{k=1}^{n} \mu_{j,k} a_k.$$

Alors $n = m$, et les matrices $(\lambda_{i,j})_{1 \leq i,j \leq n}$ et $(\mu_{i,j})_{1 \leq i,j \leq n}$ sont inversibles dans l'algèbre $\mathrm{M}_n(A)$, et inverses l'une de l'autre.

Démonstration. Tout d'abord, notons que l'existence des coefficients $\lambda_{i,j}$ et $\mu_{j,k}$ est garantie par le fait que nos familles sont génératrices. Pour tout i on a

$$a_i = \sum_{j=1}^{m} \lambda_{i,j} a'_j = \sum_{j=1}^{m} \sum_{k=1}^{n} \lambda_{i,j} \mu_{j,k} a_k = \sum_{k=1}^{n} \left(\sum_{j=1}^{m} \lambda_{i,j} \mu_{j,k} \right) \cdot a_k.$$

Puisque la famille $(a_1, \dots, a_n)$ est libre, ceci implique que pour tous i, k on a

$$\sum_{j=1}^{m} \lambda_{i,j} \mu_{j,k} = \delta_{i,k}.$$

Le même raisonnement avec la famille $(a'_1, \dots, a'_m)$ implique que pour tous i, k on a aussi

$$\sum_{j=1}^{m} \mu_{i,j} \lambda_{j,k} = \delta_{i,k}.$$

10. Ici on se limitera au cas d'une base constituée d'un nombre fini d'éléments. On peut aussi bien sûr travailler avec des bases infinies, mais ce ne sera pas utile pour nous.

11. Cette hypothèse n'est pas nécessaire pour que cet énoncé soit vrai, mais elle permet de simplifier légèrement la preuve

Donc les matrices $(\lambda_{i,j})_{\substack{1\leq i\leq n\\1\leq j\leq m}}$ et $(\mu_{i,j})_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ sont inverses l'une de l'autre comme matrices à coefficients dans A, et donc aussi comme matrices à coefficients dans le corps des fractions K de A. Ceci implique qu'on doit avoir $n = m$, ce qui complète la preuve. □

Si M est un module qui admet une base (un tel module est dit *libre*), le cardinal de n'importe laquelle de ses bases est appelé son *rang*.

5.2. **Vecteurs et pgcd des coefficients.** On suppose à partir de maintenant que A est un anneau euclidien, avec stathme φ. On rappelle que $\mathrm{GL}_n(A)$ désigne l'ensemble des éléments inversibles de l'anneau $\mathrm{M}_n(A)$, c'est-à-dire l'ensemble des matrices dont le déterminant est inversible dans A.

Lemme 2. Soit $a = (a_1, \ldots, a_n) \in A^n$, et soit $d \in A$ un pgcd de $(a_1, \ldots, a_n)$. Alors il existe $M \in \mathrm{GL}_n(A)$ telle que

$$M \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} d \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Démonstration. Remarquons que le pgcd d'un n-uplet d'éléments de A (vu comme orbite de l'action de $A^\times$ sur A par multiplication) ne change pas quand on fait agir sur ce n-uplet une matrice de $\mathrm{GL}_n(A)$. Il suffit donc de démontrer qu'il existe une matrice $N \in \mathrm{GL}_n(A)$ telle que $N \cdot a$ a au plus un coefficient non nul, en première position. (Ce coefficient sera alors nécessairement égal à d à un élément inversible près.)

Si $(a_1, \ldots, a_n) = (0, \ldots, 0)$ il n'y a rien à faire. Sinon, on raisonne par récurrence sur la valeur minimale de φ sur les coefficients non nuls de $(a_1, \ldots, a_n)$. Choisissons i tel que $a_i \neq 0$ et $\varphi(a_i)$ est minimal (parmi les valeurs de φ sur les a_j non nuls). Pour tout $j \neq i$ tel que $a_j \neq 0$, quitte à multiplier $(a_1, \ldots, a_n)$ par une matrice de $\mathrm{GL}_n(A)$, on peut remplacer a_j par le reste d'une division euclidienne par a_i. (En effet, remplacer a_j par $a_j - qa_i$ revient à multiplier notre vecteur par la matrice inversible $\mathrm{I}_n - qE_{ji}$.) Si tous les coefficients d'indice $\neq i$ ainsi obtenus sont nuls, quitte à multiplier par une matrice de permutation on peut faire "remonter" a_i en première position, et on a obtenu un vecteur de la forme voulue. Sinon on a fait strictement diminuer la valeur minimale de φ sur les coefficients non nuls du vecteur, et on conclut par récurrence. □

Remarque 1.

(1) Comme expliqué dans la preuve, le pgcd d'un n-uplet d'éléments de A ne change pas quand on fait agir sur ce n-uplet une matrice de $\mathrm{GL}_n(A)$. En termes savants, le Lemme 2 s'interprète donc en disant que l'application envoyant un vecteur de A^n sur le pgcd de ses coefficients induit une bijection entre l'ensemble des orbites de $\mathrm{GL}_n(A)$ sur A^n et $A/A^\times$. Quand A est un corps on retrouve le fait qu'il n'y a que 2 orbites : celle réduite à $\{0\}$ et celle des vecteurs non nuls. Quand $A = \mathbb{Z}$ les orbites sont en bijection avec les entiers positifs ou nuls. Enfin, quand A est l'algèbre des polynômes sur un corps, ces orbites sont en bijection avec $\mathbb{Z} \cup \{-\infty\}$ (via le degré).

(2) Dans le cas $n = 2$, le procédé décrit dans la preuve précédente est exactement l'algorithme d'Euclide appliqué à a_1 et a_2.

Comme application du Lemme 2 on obtient immédiatement le corollaire classique suivant.

Corollaire 1. Soit $(a_1, \dots, a_n) \in A^n$. Alors il existe une matrice $M \in \mathrm{GL}_n(A)$ dont la première colonne est $(a_1, \cdots, a_n)$ si et seulement si $\mathrm{pgcd}(a_1, \dots, a_n) \in A^\times$.

Démonstration. Si le pgcd est inversible, on peut appliquer le Lemme 2 avec $d = 1$ pour obtenir une matrice $N \in \mathrm{GL}_n(A)$ telle que

$$N \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Alors notre vecteur est la première colonne de la matrice inversible N^{-1}.

Si le pgcd est non inversible, alors le déterminant de toute matrice ayant ce vecteur comme première colonne est divisible par cet élément non inversible (par multilinéarité du déterminant) et ne peut donc pas être inversible. □

Remarque 2. Pour une autre preuve de ce corollaire, plus compliquée (et moins constructive) mais qui a l'avantage de s'appliquer à tous les anneaux principaux, voir [FGN2, Ex. 1.45, "Première colonne d'une matrice inversible de $\mathcal{M}_n(\mathbb{Z})$"].

5.3. **Échelonnement.** Si v est un vecteur non nul de A^n, sa *hauteur*, notée $\mathrm{h}(v)$, est définie comme $n - i$, où i est le plus petit indice d'un coefficient non nul dans v. Si $M \in \mathrm{M}_{n,m}(A)$, et si on note $v_1, \dots, v_m$ ses vecteurs colonnes, on dira que M est *échelonnée suivant les colonnes*[12] s'il existe $k \in \{0, \dots, m\}$ tel que $v_1, \dots, v_k$ sont non nuls avec

$$\mathrm{h}(v_1) > \cdots > h(v_k)$$

et si $v_{k+1} = \cdots = v_m = 0$. (Notons qu'on a alors toujours $k \leq \min(n, m)$.)

Un des intérêts de cette notion est fourni par le lemme suivant.

Lemme 3. Soit $M \in \mathrm{M}_{n,m}(A)$ une matrice échelonnée suivant les colonnes, et soient k et $v_1, \cdots, v_m$ comme ci-dessus.

(1) Les vecteurs $v_1, \dots, v_k$ forment une base du sous-module de A^n engendré par les colonnes de M.

(2) Si $(e_1, \dots, e_n)$ est la base canonique de A^n alors les vecteurs $(e_{k+1}, \dots, e_n)$ forment une base du noyau de M, c'est-à-dire du sous-module $\{v \in A^m \mid M \cdot v = 0\}$.

Démonstration. (1) Les vecteurs $v_1, \dots, v_k$ forment une famille génératrice du sous-module engendré par les colonnes puisque les colonnes suivantes sont nulles. Pour vérifier que cette famille est libre, on remarque que si $\lambda_1, \dots, \lambda_k \in A$ et si i est le plus petit indice tel que $\lambda_i \neq 0$, alors

$$\lambda_1 v_1 + \cdots + \lambda_k v_k$$

est non nul puisque son coefficient d'indice $n - \mathrm{h}(v_i)$ est non nul.

(2) Si $v = (\lambda_1, \dots, \lambda_m)$, alors on a

$$M \cdot v = \lambda_1 v_1 + \cdots + \lambda_k v_k.$$

12. Plutôt que d'échelonner suivant les colonnes on peut aussi bien sûr échelonner suivant les lignes. Cette variante est similaire, et ne sera pas considérée ici.

Puisque la famille $(v_1, \dots, v_k)$ est libre d'après (1), ce vecteur est nul si et seulement si $\lambda_1 = \cdots = \lambda_k = 0$, ce qui implique l'énoncé. □

La proposition suivante affirme que toute matrice à coefficients dans A peut être échelonnée (de façon algorithmique) en la mutlipliant à droite par une matrice inversible.

Proposition 1. Pour tout $M \in \mathrm{M}_{n,m}(A)$, il existe $N \in \mathrm{GL}_m(A)$ telle que MN est échelonnée suivant les colonnes.

Démonstration. On raisonne par récurrence sur n. Si la première ligne de M est nulle, on note M' la matrice de taille $(n-1) \times m$ formée des $n-1$ dernières lignes de M. Par récurrence il existe $Q \in \mathrm{GL}_m(A)$ telle que $M' \cdot Q$ est échelonnée suivant les colonnes, et alors $M \cdot Q$ est également échelonnée suivant les colonnes.

Si la première ligne de M est non nulle, d'après le Lemme 2 il existe $P \in \mathrm{GL}_m(A)$ telle que

$$P \cdot \begin{pmatrix} m_{1,1} \\ m_{1,2} \\ \vdots \\ m_{1,m} \end{pmatrix} = \begin{pmatrix} d \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

pour un $d \in A \smallsetminus \{0\}$. Alors on a

$$(m_{1,1}, \cdots, m_{1,m}) \cdot {}^{\mathrm{t}}P = (d, 0, \cdots, 0),$$

et donc $M \cdot {}^{\mathrm{t}}P$ a pour première ligne $(d, 0, \cdots, 0)$. On note M' la matrice de taille $(n-1)\times(m-1)$ qui forme le coin en bas à droite de la matrice $M \cdot {}^{\mathrm{t}}P$. Par récurrence il existe $Q \in \mathrm{GL}_{m-1}(A)$ telle que $M' \cdot Q$ est échelonnée suivant les colonnes, et en notant Q' la matrice diagonale par blocs de blocs (1) et Q la matrice $M \cdot {}^{\mathrm{t}}P \cdot Q'$ est alors échelonnée suivant les colonnes. □

Pour des exemples concrets de mise en oeuvre de ce procédé, et une description sous forme de (pseudo-)code, on pourra consulter [NQ, §III.3.2].

5.4. **Applications.** La Proposition 1 a tout d'abord la conséquence théorique suivante. [13]

Corollaire 2. Tout sous-A-module de A^n admet une base de cardinal au plus n.

Démonstration. L'anneau A est euclidien, donc principal, et donc *noetherien*, ce qui implique que tout sous-module de A^n est de type fini, c'est-à-dire engendré par un nombre fini de vecteurs. Si maintenant $(v_1, \dots, v_m)$ est une famille génératrice d'un tel sous-module, on considère la matrice M de $\mathrm{M}_{n,m}(A)$ dont les colonnes sont $v_1, \dots, v_m$. Multiplier une matrice à droite par une matrice inversible ne change pas le sous-module engendré par les colonnes, donc une matrice échelonnée obtenue à partir de M par le procédé de la proposition 1 aura le même sous-module engendré par les colonnes que M. D'après le Lemme 3, les colonnes non nulles de cette matrice formeront une base de ce sous-module. Par ailleurs, cette base est de cardinal au plus n d'après la condition d'échelonnement. □

13. Insistons sur le fait que ce corollaire s'applique aux anneaux *euclidiens*, mais pas à tous les anneaux. (Plus généralement, ce résultat est vrai pour un anneau principal ; voir [FGN2, Ex. 1.44, "Bases d'un groupe abélien"] ou [NQ, §III.3.2] pour une preuve—non constructive—dans ce cadre.)

L'échelonnement permet également de résoudre des systèmes linéaires homogènes dans A, en se basant sur l'observation simple suivante.

Lemme 4. Soit $M \in \mathrm{M}_{n,m}(A)$, et soit $N \in \mathrm{GL}_m(A)$ une matrice telle que $M \cdot N$ est échelonnée suivant les colonnes. Notons k le nombre de colonnes non nulles de MN. Alors le noyau de M admet pour base les colonnes de N d'indices $k+1, \dots, m$.

Démonstration. On observe que si $v \in A^m$, alors $M \cdot v = 0$ si et seulement si $(MN) \cdot (N^{-1}v) = 0$. Si on note k le nombre de colonnes non nulles de MN, d'après le Lemme 3(2) cette condition revient à dire que $N^{-1}v$ est une combinaison linéaire de $e_{k+1}, \cdots, e_m$, c'est-à-dire que v est une combinaison linéaire de $Ne_{k+1}, \cdots, Ne_m$. Puisque Ne_j est la j-ième colonne de N, cela implique le résultat voulu. □

Remarque 3. Avec les notations du lemme précédent, les k premières colonnes de N forment une base d'un supplémentaire de $\ker(M)$ dans A^m ; voir [NQ, §3.4] pour les détails.

En raffinant le procédé précédent on peut même résoudre des systèmes linéaires entiers non homogènes ; pour des détails, on renvoie à [NQ, §III.3.3 et §III.3.5].

5.5. **Matrices échelonnées réduites.** L'échelonnement d'une matrice est déjà très utile, mais pour aller plus loin on veut parfois avoir un résultat d'unicité de la matrice échelonnée obtenue. Pour cela il faut fixer quelques données. On commence par choisir un système de représentants $\mathfrak{A} \subset A \smallsetminus \{0\}$ des orbites de l'action de $A^\times$ par multiplication sur $A \smallsetminus \{0\}$, c'est-à-dire un ensemble d'éléments tels que pour tout $a \in A \smallsetminus \{0\}$ il existe un *unique* $a' \in \mathfrak{A}$ et un unique $u \in A^\times$ tels que $a = ua'$. On fixe ensuite, pour tout $a \in \mathfrak{A}$, un système $\mathfrak{R}_a \subset A$ de représentants du quotient $A/(A \cdot a)$, c'est-à-dire un sous-ensemble tel que pour tout $b \in A$ il existe un unique couple $(q, r) \in A \times \mathfrak{R}_a$ tel que $b = aq + r$. Avant d'aller plus loin, expliquons comment on peut fixer de tels choix dans les cas "classiques" :

- si A est un corps, on peut prendre $\mathfrak{A} = \{1\}$ et $\mathfrak{R}_1 = \{0\}$;
- si $A = \mathbb{Z}$, on peut prendre pour $\mathfrak{A}$ l'ensemble des entiers strictement positifs, et pour $\mathfrak{R}_a$ l'ensemble $\{0, \dots, a-1\}$;
- si $A = \Bbbk[X]$ pour un corps $\Bbbk$, on peut prendre pour $\mathfrak{A}$ l'ensemble des polynômes unitaires, et pour $\mathfrak{R}_a$ l'ensemble des polynômes de degré strictement inférieur à $\deg(a)$.

Soit $M \in \mathrm{M}_{n,m}(A)$ une matrice échelonnée suivant les colonnes. Notons k le nombre de colonnes non nulles dans M, notons $v_1, \dots, v_k$ ces colonnes, et posons $f(i) = n - \mathrm{h}(v_i)$ pour $i \in \{1, \dots, k\}$ (de sorte que $f(i)$ est l'indice du premier coefficient non nul de la i-ème colonne). On dira que M est *réduite* si chaque $m_{f(i),i}$ (avec $i \in \{1, \dots, k\}$) appartient à $\mathfrak{A}$ et si pour tout $i \in \{2, \dots, k\}$ et tout $j \in \{1, \dots, i-1\}$ on a $m_{f(i),j} \in \mathfrak{R}_{m_{f(i),i}}$.

Théorème 1. Si $M \in \mathrm{M}_{n,m}(A)$, l'ensemble

$$\{M \cdot N : N \in \mathrm{GL}_m(A)\}$$

contient une unique matrice échelonnée suivant les colonnes et réduite.

Démonstration. Prouvons tout d'abord l'existence d'une telle matrice. En utilisant la Proposition 1 on peut supposer que M est échelonnée suivant les colonnes. On peut ensuite multiplier par une matrice diagonale avec coefficients inversibles pour s'assurer que les $m_{f(i),i}$ sont dans $\mathfrak{A}$, puis considérer pour tout $i \in \{2, \dots, k\}$ (dans

l'ordre croissant) et tout $j \in \{1, \dots, i-1\}$ l'unique écriture $m_{f(i),j} = qm_{f(i),i} + r$ avec $r \in \mathfrak{R}_{m_{f(i),i}}$ et remplacer $m_{f(i),j}$ par r (ce qui revient à multiplier à droite par une matrice de $\mathrm{GL}_m(A)$) pour forcer la condition $m_{f(i),j} \in \mathfrak{R}_{m_{f(i),i}}$.

Pour l'unicité, on doit montrer que si M et M' sont échelonnées suivant les colonnes et réduites, et si $M' = MN$ pour une matrice $N \in \mathrm{GL}_m(A)$, alors $M = M'$. Si une telle matrice existe, alors M et M' ont même rang (vues comme matrices dans $\mathrm{GL}_n(K)$ où K est le corps des fractions de A), et donc le même nombre k de colonnes non nulles. Si on note $v_1, \dots, v_k$ et $v'_1, \dots, v'_k$ ces colonnes dans M et M' respectivement, alors pour $j \in \{1, \dots, k\}$ on a

$$v'_j = \sum_{i=1}^{k} n_{i,j} \cdot v_i,$$

et si $j \geq k+1$ les coefficients $n_{i,j}$ sont nuls pour tout $1 \leq i \leq k$ (voir le Lemme 3(2)).

Puisque la hauteur de $\sum_{i=1}^{k} n_{i,j} \cdot v_i$ est la hauteur de v_l où l est le plus petit indice tel que $n_{l,j} \neq 0$, la condition d'échelonnement impose que $n_{i,j} = 0$ si $i < j$, et que les fonctions "f" sont les mêmes pour M et M'. Puisque la matrice N est inversible et triangulaire inférieure par blocs le bloc $(n_{i,j})_{1 \leq i,j \leq k}$ est inversible, et les éléments $n_{i,i}$ sont donc inversibles pour $i \in \{1, \dots, k\}$. Puisque deux éléments de $\mathfrak{A}$ diffèrent par multiplication par un inversible si et seulement si ils sont égaux, on en déduit que $n_{i,i} = 1$ pour tout $i \in \{1, \dots, k\}$. Enfin, la condition sur les coefficients $m_{f(i),j}$ et $m'_{f(i),j}$ implique que $n_{i,j} = 0$ si $j < i$, et donc finalement que $M = M'$. □

Remarque 4. (1) Encore une fois, le Théorème 1 peut s'interpréter comme une classification des orbites du groupe $\mathrm{GL}_m(A)$ pour son action par multiplication à droite sur $\mathrm{M}_{n,m}(A)$: ces orbites sont en bijection avec les matrices échelonnées suivant les colonnes et réduites. En termes plus concrets, ces orbites sont donc classifiées par les données suivantes :
- un entier $k \in \{1, \dots, \min(n, m)\}$;
- une fonction strictement croissante $f : \{1, \dots, k\} \to \{1, \dots, n\}$;
- des "pivots" $m_{f(i),i}$ appartenenant à $\mathfrak{A}$;
- pour tout $i \in \{2, \dots, k\}$ et tout $j \in \{1, \dots, i-1\}$, un élément $m_{f(i),j} \in \mathfrak{R}_{m_{f(i),i}}$;
- pour tout $i \in \{1, \dots, k\}$ et tout $l \in \{f(i)+1, f(i+1)-1\}$ (où par convention $f(k+1) = n+1$) et $j \in \{1, \dots, i-1\}$ des coefficients $m_{l,j}$ arbitraires dans A.

(2) Dans le cas particulier où A est un corps, le Théorème 1 est très classique, voir par exemple [CG2, Chap. IV, Théorème 2.3.1]. Dans ce cas les "pivots" $m_{f(i),i}$ peuvent être choisis égaux à 1, et les coefficients $m_{f(i),j}$ avec $i \in \{2, \dots, k\}$ et $j < i$ peuvent être choisis nuls.

(3) Dans le cas $A = \mathbb{Z}$, avec les choix "évidents" comme ci-dessus, les matrices échelonnées réduites sont dites en *forme normale d'Hermite*.

5.6. **Applications.** Le principal intérêt de la forme échelonnée réduite est son unicité. Ceci permet par exemple de tester l'égalité de sous-modules de A^n décrits par une famille génératrice, grâce à l'énoncé suivant.

Lemme 5. Soient M et M' deux matrices dans $\mathrm{M}_{n,m}(A)$. Les propriétés suivantes sont équivalentes :

(1) les sous-modules de A^n engendrés par les colonnes de M et M' coincident ;
(2) on a $M \cdot \mathrm{GL}_m(A) = M' \cdot \mathrm{GL}_m(A)$;
(3) les uniques matrices échelonnées suivant les colonnes et réduites dans $M \cdot \mathrm{GL}_m(A)$ et dans $M' \cdot \mathrm{GL}_m(A)$ coincident.

Démonstration. L'équivalence des conditions (2) et (3) est évidente. On a également remarqué au cours de la preuve du Corollaire 2 que la multiplication à droite par une matrice inversible ne change pas le sous-module engendré par les colonnes ; donc la condition (2) implique (1). Enfin, supposons que la condition (1) est vérifiée. Soient N et N' les uniques matrices échelonnées suivant les colonnes et réduites dans $M \cdot \mathrm{GL}_m(A)$ et $M' \cdot \mathrm{GL}_m(A)$ respectivement. Alors les sous-modules engendrés par les colonnes de N et N' coincident, puisqu'ils coincident avec ceux pour M et M' respectivement. Si on note $\overline{N}$ et $\overline{N'}$ les matrices obtenues à partir de N et N' en supprimant les colonnes nulles, alors d'après le Lemme 1 $\overline{N}$ et $\overline{N'}$ ont le même nombre de colonnes, qu'on notera k, et de plus il existe une matrice $P \in \mathrm{GL}_k(A)$ telle que

$$\overline{N'} = \overline{N} \cdot P.$$

Si on note Q la matrice diagonale par blocs avec pour blocs P et I_{m-r}, alors Q est dans $\mathrm{GL}_m(A)$ et on a

$$N' = N \cdot Q,$$

ce qui montre que la condition (2) est vérifiée et achève la preuve. □

Remarque 5. Si on a une méthode pour tester l'égalité de deux sous-modules de A^n, on peut en déduire une méthode pour tester *l'inclusion* de deux sous-modules, en remarquant que si $N_1, N_2 \subset A^n$, alors $N_1 \subset N_2$ si et seulement si $N_1 + N_2 = N_2$.

5.7. **Forme normale de Smith.** La *forme normale de Smith* concerne l'action de $\mathrm{GL}_n(A) \times \mathrm{GL}_m(A)$ sur $\mathrm{M}_{n,m}(A)$ définie par $(P, Q) \cdot M = PMQ^{-1}$. Encore une fois, le théorème suivant donne une description des orbites pour cette action.

Théorème 2. Soit $M \in \mathrm{M}_{n,m}(A)$. Il existe $s \in \{0, \dots, \min(n, m)\}$, des éléments $d_1, \dots, d_s \in A \smallsetminus \{0\}$ tels que

$$d_1 \mid d_2 \mid \cdots \mid d_s$$

et des matrices $P \in \mathrm{GL}_n(A)$, $Q \in \mathrm{GL}_m(A)$ telles que

$$M = P \cdot \begin{pmatrix} d_1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & \vdots & \vdots & & & \vdots \\ \vdots & & \ddots & 0 & \vdots & & & \vdots \\ 0 & \cdots & 0 & d_s & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ \vdots & & & \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} \cdot Q.$$

De plus l'entier s est unique, et la suite $(d_1, \dots, d_s)$ est unique à multiplication par des inversibles près.

La suite $(d_1, \dots, d_s)$ (définie à multiplication par des inversibles près) est appelée la suite des *facteurs invariants* de M. Pour une preuve de ce théorème, et la description d'un algorithme permettant de calculer les facteurs invariants, on pourra consulter par exemple [NQ, §III.4] ou [BMP, §6.4.2].

Remarque 6. (1) Le Théorème 2 est vrai plus généralement pour A un anneau principal. Cependant, dans cette généralité il n'existe pas d'algorithme permettant de calculer les facteurs invariants.

(2) Dans le cas $A = \mathbb{Z}$ on peut imposer que les d_i sont des entiers *positifs* ; la suite des facteurs invariants est alors unique. De même, si A est l'algèbre des polynômes sur un corps, on peut imposer que les d_i sont des polynômes *unitaires*, ce qui permet là encore d'avoir des facteurs invariants uniquement déterminés.

Le Théorème 2 s'applique en particulier pour $A = \Bbbk[X]$ si $\Bbbk$ est un corps. Dans ce cadre, comme expliqué ci-dessus on normalise les facteurs invariants en demandant à ce qu'ils soient unitaires. On peut se poser la question du comportement des facteurs invariants quand on remplace $\Bbbk$ par une extension $\mathbb{K}$. Cette question a une réponse très simple, expliquée dans le lemme suivant.

Lemme 6. Soit $\Bbbk$ un corps, et soit $\mathbb{K}$ une extension de $\Bbbk$. Pour tous $n, m \geq 1$ et toute matrice M de $\mathrm{M}_{n,m}(\Bbbk[X])$, les facteurs invariants de M vue comme matrice de $\mathrm{M}_{n,m}(\Bbbk[X])$ et de $\mathrm{M}_{n,m}(\mathbb{K}[X])$ coincident.

Démonstration. Notons $(p_1, \dots, p_s)$ les facteurs invariants de M vue comme matrice de $\mathrm{M}_{n,m}(\Bbbk[X])$. Alors il existe $P \in \mathrm{GL}_n(\Bbbk[X])$, $Q \in \mathrm{GL}_m(\Bbbk[X])$ telles que

$$M = P \cdot \begin{pmatrix} p_1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & \vdots & \vdots & & & \vdots \\ \vdots & & \ddots & 0 & \vdots & & & \vdots \\ 0 & \cdots & 0 & p_s & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ \vdots & & & \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} \cdot Q.$$

Cette égalité peut être vue comme une égalité dans $\mathrm{M}_{n,m}(\mathbb{K}[X])$. Par unicité dans le Théorème 2, elle montre que $(p_1, \dots, p_s)$ est également la suite des facteurs invariants de M vue comme matrice de $\mathrm{M}_{n,m}(\mathbb{K}[X])$. □

Exercice 4. (1) Montrer que, dans le notations du Théorème 2, d_1 est un pgcd des coefficients de M.

(2) Soient $a, b \in A \setminus \{0\}$, et soient d et m un pgcd et un ppcm de (a, b) respectivement. Montrer que les facteurs invariants de la matrice

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

sont (d, m).

5.8. **Application à la réduction des endomorphismes.** (La présentation de cette partie et de la suivante est fortement inspirée de [De].)

L'application la plus classique du Théorème 2 concerne la réduction des endomorphismes d'un $\Bbbk$-espace vectoriel de dimension finie ou, de façon équivalente, l'étude de la conjugaison dans les matrices carrées à coefficients dans $\Bbbk$. On fixe donc un corps $\Bbbk$ et un entier $n \geq 1$.

Pour commencer on prouve deux lemmes techniques dont on aura besoin plus loin. (On peut voir le premier lemme comme une version de la division euclidienne dans une algèbre de matrices. Notons que cette algèbre n'est pas commutative, donc elle ne peut a fortiori pas être euclidienne.)

Lemme 7. Pour toutes matrices $P \in \mathrm{M}_n(\Bbbk[X])$ et $M \in \mathrm{M}_n(\Bbbk)$, il existe des matrices $Q \in \mathrm{M}_n(\Bbbk[X])$ et $N \in \mathrm{M}_n(\Bbbk)$ telles que

$$P = (X\mathrm{I}_n - M) \cdot Q + N.$$

De même, il existe des matrices $Q' \in \mathrm{M}_n(\Bbbk[X])$ et $N' \in \mathrm{M}_n(\Bbbk)$ telles que

$$P = Q' \cdot (X\mathrm{I}_n - M) + N'.$$

Démonstration. On effectue la preuve dans le premier cas ; le deuxième s'en déduit par transposition. On raisonnera par récurrence sur le degré maximal d d'un coefficient de P. Si $d \in \{0, -\infty\}$ alors P est dans $\mathrm{M}_n(\Bbbk)$, et on peut donc prendre $Q = 0$ et $N = P$. Si $d > 0$ on écrit $P = XR + S$ avec $R \in \mathrm{M}_n(\Bbbk[X])$ dont le degré maximal des coefficients est $d-1$ et S dans $\mathrm{M}_n(\Bbbk)$. Par récurrence il existe $R_1 \in \mathrm{M}_n(\Bbbk[X])$ et $R_2 \in \mathrm{M}_n(\Bbbk)$ telles que

$$R = (X\mathrm{I}_n - M) \cdot R_1 + R_2.$$

Alors on a

$$\begin{aligned} P = XR + S &= X(X\mathrm{I}_n - M) \cdot R_1 + XR_2 + S \\ &= (X\mathrm{I}_n - M) \cdot (XR_1) + (X\mathrm{I}_n - M) \cdot R_2 + MR_2 + S \\ &= (X\mathrm{I}_n - M) \cdot (XR_1 + R_2) + (MR_2 + S); \end{aligned}$$

on peut donc poser $Q = XR_1 + R_2$ et $N = MR_2 + S$. □

Pour une matrice non nulle $P \in \mathrm{M}_n(\Bbbk[X])$, on note $\deg(P)$ le degré maximal d'un coefficient non nul de M.

Lemme 8. Si $P \in \mathrm{M}_n(\Bbbk[X])$ est non nulle et si $M \in \mathrm{M}_n(\Bbbk)$, alors les matrices $P \cdot (X\mathrm{I}_n - M)$ et $(X\mathrm{I}_n - M) \cdot P$ sont non nulles, et elles vérifient

$$\deg(P \cdot (X\mathrm{I}_n - M)) = \deg((X\mathrm{I}_n - M) \cdot P) = \deg(P) + 1.$$

Démonstration. Ce lemme découle des observations simples suivantes :

— si $P \in \mathrm{M}_n(\Bbbk[X])$ est non nulle, alors $(X\mathrm{I}_n) \cdot P$ et $P \cdot (X\mathrm{I}_n)$ son non nulles, et de plus

$$\deg(P \cdot (X\mathrm{I}_n)) = \deg((X\mathrm{I}_n) \cdot P) = \deg(P) + 1;$$

— si $P \in \mathrm{M}_n(\Bbbk[X])$ est non nulle et si $M \in \mathrm{M}_n(\Bbbk)$, alors si $M \cdot P$, resp. $P \cdot M$, est non nulle on a

$$\deg(M \cdot P) \leq \deg(P), \quad \text{resp.} \quad \deg(P \cdot M) \leq \deg(P);$$

— si $P, Q \in \mathrm{M}_n(\Bbbk[X])$ sont non nulles et si $\deg(Q) < \deg(P)$, alors $P + Q$ est non nulle et $\deg(P + Q) = \deg(P)$.

□

Le résultat-clé qui permet de faire le lien entre conjugaison des matrices et forme normale de Smith est le suivant.

Proposition 2. Soient $M_1, M_2 \in \mathrm{M}_n(\Bbbk)$. Alors M_1 et M_2 sont conjuguées dans $\mathrm{M}_n(\Bbbk)$ si et seulement si il existe $P, Q \in \mathrm{M}_n(\Bbbk[X])$ inversibles telles que $X\mathrm{I}_n - M_1 = P \cdot (X\mathrm{I}_n - M_2) \cdot Q$.

Démonstration. Si M_1 et M_2 sont conjuguées il existe $P \in \mathrm{GL}_n(\Bbbk)$ telle que $M_1 = P \cdot M_2 \cdot P^{-1}$. Alors on a

$$X\mathrm{I}_n - M_1 = P \cdot (X\mathrm{I}_n - M_2) \cdot P^{-1},$$

et P est inversible dans $\mathrm{M}_n(\Bbbk[X])$.

Réciproquement, supposons qu'il existe des matrices P, Q inversibles dans $\mathrm{M}_n(\Bbbk[X])$ telles que

$$X\mathrm{I}_n - M_1 = P \cdot (X\mathrm{I}_n - M_2) \cdot Q.$$

On a alors $(X\mathrm{I}_n - M_1) \cdot Q^{-1} = P \cdot (X\mathrm{I}_n - M_2)$. D'après le Lemme 7 on peut écrire

$$P = (X\mathrm{I}_n - M_1) \cdot P_1 + P_2, \qquad Q^{-1} = Q_1 \cdot (X\mathrm{I}_n - M_2) + Q_2$$

avec P_1, Q_1 dans $\mathrm{M}_n(\Bbbk[X])$ et P_2, Q_2 dans $\mathrm{M}_n(\Bbbk)$. On a alors

$$P \cdot (X\mathrm{I}_n - M_2) = (X\mathrm{I}_n - M_1) \cdot P_1 \cdot (X\mathrm{I}_n - M_2) + P_2 \cdot (X\mathrm{I}_n - M_2),$$
$$(X\mathrm{I}_n - M_1) \cdot Q^{-1} = (X\mathrm{I}_n - M_1) \cdot Q_1 \cdot (X\mathrm{I}_n - M_2) + (X\mathrm{I}_n - M_1) \cdot Q_2,$$

et donc

$$(X\mathrm{I}_n - M_1) \cdot (P_1 - Q_1) \cdot (X\mathrm{I}_n - M_2) = -P_2 \cdot (X\mathrm{I}_n - M_2) + (X\mathrm{I}_n - M_1) \cdot Q_2.$$

Puisque le membre de droite a tous ses coefficients de degré au plus 1, en utilisant le Lemme 8 on voit qu'on doit avoir $P_1 = Q_1$, et donc

$$P_2 \cdot (X\mathrm{I}_n - M_2) = (X\mathrm{I}_n - M_1) \cdot Q_2.$$

En identifiant les termes de degrés 1 et 0, on en déduit que

$$P_2 = Q_2 \quad \text{et} \quad P_2 \cdot M_2 = M_1 \cdot Q_2,$$

de sorte que pour conclure il suffit de montrer que Q_2 est inversible.

Pour cela on écrit

$$Q = Q_1' \cdot (X\mathrm{I}_n - M_1) + Q_2',$$

et on observe que

$$\begin{aligned} \mathrm{I}_n = Q^{-1} \cdot Q &= \big(Q_1 \cdot (X\mathrm{I}_n - M_2) + Q_2\big) \cdot Q \\ &= (Q_1 P^{-1}) \cdot (X\mathrm{I}_n - M_1) + Q_2 \cdot \big(Q_1' \cdot (X\mathrm{I}_n - M_1) + Q_2'\big) \\ &= (Q_1 P^{-1} + Q_2 Q_1') \cdot (X\mathrm{I}_n - M_1) + Q_2 Q_2'. \end{aligned}$$

Comme précédemment, pour des raisons de degré on a $Q_1 P^{-1} + Q_2 Q_1' = 0$, et donc $Q_2 Q_2' = \mathrm{I}_n$, ce qui prouve que Q_2 est inversible et achève la preuve. □

Remarque 7. La Proposition 2 montre en particulier que si les matrices $X\mathrm{I}_n - M_1$ et $X\mathrm{I}_n - M_2$ sont équivalentes dans $\mathrm{M}_n(\Bbbk[X])$ (au sens où elles diffèrent par multiplication à gauche et à droite par des matrices inversibles) alors elles sont en fait conjuguées, c'est-à-dire qu'il existe une matrice inversible P telle que $X\mathrm{I}_n - M_1 = P \cdot (X\mathrm{I}_n - M_2) \cdot P^{-1}$. (On peut même alors prendre $P \in \mathrm{GL}_n(\Bbbk)$.) Bien sûr

ceci n'est pas vrai pour des matrices arbitraires dans $\mathrm{M}_n(\Bbbk[X])$, mais exploite la forme très particulière de ces matrices.

La Proposition 2 et le Théorème 2 montrent que deux matrices M_1, M_2 de $\mathrm{M}_n(\Bbbk)$ sont conjuguées si et seulement si les matrices $X\mathrm{I}_n - M_1$ et $X\mathrm{I}_n - M_2$ ont les mêmes facteurs invariants (normalisés en demandant à ce qu'ils soient unitaires). Notons que quand on applique le Théorème 2 à $X\mathrm{I}_n - M$ pour une matrice $M \in \mathrm{M}_n(\Bbbk)$, l'entier "$s$" vaut n ; cela découle du fait que le déterminant de $X\mathrm{I}_n - M$ est non nul, puisqu'il s'agit du polynôme caractéristique de M. La suite $(p_1, \dots, p_r)$ des polynômes *non constants* parmi les facteurs invariants de $X\mathrm{I}_n - M$ est appelée la suite des *invariants de similitude* de M. On a donc obtenu que deux matrices sont conjuguées si et seulement si elles ont les mêmes invariants de similitude. Notons au passage qu'en prenant les déterminants dans l'égalité du Théorème 2 on obtient que

$$p_1 \cdots p_r = \chi_M \tag{1}$$

(où χ_M est le polynôme caractéristique de M). En particulier,

$$\deg(p_1) + \cdots + \deg(p_r) = n.$$

Le comportement des invariants de similitude par extension du corps de coefficients est très simple, et permet de répondre à la question du comportement de la conjugaison des matrices par extension de corps.

Proposition 3. Soit $\mathbb{K}$ une extension de $\Bbbk$ et soit $M \in \mathrm{M}_n(\Bbbk)$. Les invariants de similitude de M vue comme matrice de $\mathrm{M}_n(\Bbbk)$ ou comme matrice de $\mathrm{M}_n(\mathbb{K})$ sont les mêmes. En particulier, deux matrices de $\mathrm{M}_n(\Bbbk)$ sont conjuguées dans $\mathrm{M}_n(\Bbbk)$ si et seulement si elles sont conjuguées dans $\mathrm{M}_n(\mathbb{K})$.

Démonstration. D'après le Lemme 6, les facteurs invariants de $X\mathrm{I}_n - M$ vue comme matrice dans $\mathrm{M}_n(\Bbbk[X])$ ou dans $\mathrm{M}_n(\mathbb{K}[X])$ sont les mêmes. Donc les invariants de similitude de M vue comme matrice de $\mathrm{M}_n(\Bbbk)$ ou de $\mathrm{M}_n(\mathbb{K})$ coincident également.

Pour la deuxième partie de l'énoncé, si deux matrices sont conjuguées dans $\mathrm{M}_n(\Bbbk)$ elles le sont également dans $\mathrm{M}_n(\mathbb{K})$. Réciproquement, si elles sont conjuguées dans $\mathrm{M}_n(\mathbb{K})$ elles ont les mêmes invariants de similitude sur $\mathbb{K}$, et donc sur $\Bbbk$ d'après ce qu'on vient de prouver. Elles sont donc également conjuguées dans $\mathrm{M}_n(\Bbbk)$. □

Remarque 8. La Proposition 3 a une preuve plus simple sous l'hypothèse où $\Bbbk$ est infini, qu'il est bien de connaitre également ; voir par exemple le §6 de `http://www.normalesup.org/~sage/Enseignement/Colles/AlgLin/Determin.pdf`.

5.9. **Matrices compagnons.** Les considérations du §5.8 montrent que chaque classe de conjugaison de matrices de $\mathrm{M}_n(\Bbbk)$ est uniquement caractérisée par les invariants de similitude de n'importe quelle matrice dans cette classe, qui sont une suite $(p_1, \dots, p_r)$ de polynômes unitaires non constants tels que $p_1 \mid \cdots \mid p_r$ et $\sum_{i=1}^r \deg(p_i) = n$. Mais si on veut *classifier* les orbites de $\mathrm{GL}_n(\Bbbk)$ sur $\mathrm{M}_n(\Bbbk)$ par conjugaison, on doit également déterminer quelles suites de polynômes vérifiant ces hypothèses peut apparaitre comme suite des invariants de similitude d'une matrice. La réponse à cette question va être que *toute* telle suite peut apparaitre, et une façon simple de le voir est d'introduire les *matrices compagnons*.

Étant donné un polynôme unitaire non constant

$$P(X) = X^m + a_{m-1}X^{m-1} + \cdots + a_1 X + a_0$$

dans $\Bbbk[X]$, la matrice compagnon associée est la matrice de taille $m \times m$ à coefficients dans $\Bbbk$ définie par

$$C_P := \begin{pmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{m-1}. \end{pmatrix}$$

(Dans le cas $m = 1$, cette matrice doit être interprétée comme étant égale à $(-a_0)$.) L'énoncé suivant peut se vérifier "à la main" en faisant des opérations élémentaires sur les lignes et colonnes de $X\mathrm{I}_m - C_P$, voir [CG1, p. 99–100].

Lemme 9. La matrice C_P admet un unique invariant de similitude, égal à P.

On en déduit l'existence, pour toute suite de polynômes comme ci-dessus, d'une matrice admettant cette suite comme invariants de similitude.

Proposition 4. Pour toute suite $(p_1, \ldots, p_r)$ de polynômes unitaires non constants tels que $p_1 \mid \cdots \mid p_r$ et $\sum_{i=1}^r \deg(p_i) = n$, la matrice (de taille n) diagonale par blocs avec pour blocs successifs $C_{p_1}, \ldots, C_{p_r}$ a pour invariants de similitude $(p_1, \ldots, p_r)$.

Démonstration. Notons M la matrice de l'énoncé. Il suit du Lemme 9 que la matrice $X\mathrm{I}_n - M$ peut s'obtenir par multiplication à gauche et à droite par des matrices inversibles (dans $\mathrm{M}_n(\Bbbk[X])$) à partir de la matrice diagonale avec coefficients successifs

$$\underbrace{1, \ldots, 1}_{\deg(p_1)-1}, p_1, \underbrace{1, \ldots, 1}_{\deg(p_2)-1}, p_2, \ldots, \underbrace{1, \ldots, 1}_{\deg(p_r)-1}, p_r.$$

En conjuguant par une matrice de permutation on peut réordonner ces termes pour obtenir la matrice diagonale avec coefficients

$$\underbrace{1, \ldots, 1}_{n-r}, p_1, \ldots, p_r.$$

Par unicité des facteurs invariants, cela démontre l'énoncé voulu. □

On donc finalement obtenu l'énoncé suivant, qui fournit une classification des classes de conjugaison dans $\mathrm{M}_n(\Bbbk)$.

Théorème 3 (Réduction de Frobenius)**.** L'application envoyant une matrice sur ses invariants de similitude induit une bijection entre l'ensemble des orbites de $\mathrm{GL}_n(\Bbbk)$ pour son action par conjugaison sur $\mathrm{M}_n(\Bbbk)$ et l'ensemble des suites $(p_1, \ldots, p_r)$ de polynômes unitaires de $\Bbbk[X]$ tels que $p_1 \mid \cdots \mid p_r$ et $\sum_{i=1}^r \deg(p_i) = n$. Étant donnée une telle suite de polynômes, un représentant de la classe correspondant à $(p_1, \ldots, p_r)$ est donné par la matrice diagonale par blocs avec pour blocs successifs $C_{p_1}, \ldots, C_{p_r}$.

En général, le calcul des invariants de similitude d'une matrice donnée peut être assez lourd (même s'il existe un algorithme permettant de le faire, sans réfléchir). Le "plus gros" polynôme apparaissant dans cette liste a une interprétation simple, comme suit. (Une autre relation entre les invariants de similitude et les polynômes "usuels" associés à une matrice est donnée par (1).)

Proposition 5. Soit M une matrice de $\mathrm{M}_n(\Bbbk)$, et soient $(p_1, \ldots, p_r)$ ses invariants de similitude. Alors p_r est le polynôme minimal de M.

Démonstration. Puisque le polynôme minimal est invariant par conjugaison, on peut supposer que M est la matrice diagonale par blocs avec pour blocs successifs $C_{p_1}, \dots, C_{p_r}$. Il n'est pas difficile de voir que pour tout polynôme unitaire non constant P, le polynôme minimal de C_P est P. (La preuve est laissée au lecteur ; on peut par exemple utiliser l'Exercice 5 ci-dessous.) Étant donné un polynôme Q, on a alors $Q(M) = 0$ si et seulement si $Q(C_{p_i}) = 0$ pour tout i, ce qui revient à dire que $p_i \mid Q$ pour tout i. La condition de divisibilité dans les invariants de similitude montre que ces conditions sont équivalentes à dire que $p_r \mid Q$, ce qui achève la preuve. □

Pour des exemples explicites de calculs d'invariants de similitude, et quelques applications, on pourra consulter [BMP, §§6.5–6.6].

Exercice 5. Montrer que C_P est la matrice dans une base appropriée de l'endomorphisme du $\Bbbk$-espace vectoriel $\Bbbk[X]/(P)$ donné par la multiplication par X.

Référence : [BMP, Lemme 6.90].

Exercice 6 (Quelques applications classiques des invariants de similitude)**.** Soit $\Bbbk$ un corps.

(1) Montrer que si $M \in \mathrm{M}_n(\Bbbk)$, M et ${}^{\mathrm{t}}M$ ont les mêmes invariants de similitude. En déduire que ces matrices sont conjuguées.

(2) Montrer que pour tout $M \in \mathrm{M}_n(\Bbbk)$, le polynôme minimal et le polynôme caractéristique de M ont les mêmes facteurs irréductibles.

Référence : pour (1), voir [BMP, Application 6.103]. Pour (2), voir [BMP, Application 6.100].

Exercice 7. Dans le cas où $\Bbbk$ est fini, compter le nombre de classes de conjugaison dans $\mathrm{M}_n(\Bbbk)$.

Références

[BMP] V. Beck, J. Malick, G. Peyré, *Objectif agrégation*, 2ème édition, 2005, H&K.

[CG1] P. Caldero, J. Germoni, *Histoires hédonistes de groupes et de géométries, Tome II*, Calvage & Mounet, 2015.

[CG2] P. Caldero et J. Germoni, *Nouvelles histoires hédonistes de groupes et de géométries, Tome I*, Calvage & Mounet, 2017.

[Co] H. Cohen, *A course in computational algebraic number theory*, Springer, 1993.

[De] O. Debarre, *Réduction des endomorphismes*, notes disponibles à l'adresse `https://www.math.ens.fr/~debarre/ReducEndo.pdf`.

[FGN1] S. Francinou, H. Gianella, S. Nicolas, *Oraux X-ENS, mathématiques 1*, Cassini, 2019.

[FGN2] S. Francinou, H. Gianella, S. Nicolas, *Oraux X-ENS, mathématiques 2*, Cassini, 2021.

[Go] X. Gourdon, *Les maths en tête - Algèbre, probabilités, 3ème édition*, Ellipses, 2021.

[NQ] P. Naudin, C. Quitté, *Algorithmique algébrique avec exercices corrigés*, Masson, 1992.

[SP] P. Saux Picart, *Cours de calcul formel. Algorithmes fondamentaux*, Ellipses, 1999.

[Sk] G. Skandalis, *Agrégation interne – Algèbre générale, algèbre linéaire et un peu de géométrie*, Calvage et Mounet, 2017.