Modules projectifs de type fini, applications linéaires croisées et inverses généralisés

Gema M. Diaz–Toca*
Dpto. de Matematicas Aplicada
Universidad de Murcia, Spain
gemadiaz@um.es

Laureano Gonzalez–Vega*
Dpto. de Matemáticas
Univ. of Cantabria, Spain
gvega@matesco.unican.es

Henri Lombardi
Équipe de Mathématiques, UMR CNRS 6623
Univ. de Franche-Comté, France
henri.lombardi@univ-fcomte.fr

Claude Quitté
Laboratoire de Mathématiques, SP2MI,
Université de Poitiers, France
quitte@mathlabo.univ-poitiers.fr

2004 (version plus détaillée)

Résumé

D'une part, nous développons la théorie générale des inverses généralisés de matrices en la mettant en rapport avec la théorie constructive des modules projectifs de type fini. D'autre part nous précisons certains aspects de cette théorie liés au calcul formel et à l'analyse numérique matricielle. Nous démontrons en particulier qu'on peut tester si un $\mathbf{A}$–module de présentation finie est projectif et calculer une matrice de projection correspondante « en temps polynomial ». Plus précisément pour une matrice $A \in \mathbf{A}^{m\times n}$ on peut décider s'il existe un inverse généralisé B pour A (c'est-à-dire une matrice B vérifiant $ABA = A$ et $BAB = B$) et, en cas de réponse positive, calculer un tel inverse généralisé par un algorithme qui utilise $\mathcal{O}(p^6\, q^2)$ opérations arithmétiques (avec $p = \inf(m, n)$, $q = \sup(m, n)$) et un nombre polynomial de tests d'appartenance d'un élément à un idéal engendré par « un petit nombre d'éléments. ».

Introduction

Dans cet article $\mathbf{A}$ désigne un anneau commutatif arbitraire. D'une part, nous développons la théorie générale des inverses généralisés de matrices en la mettant en rapport avec la théorie constructive des modules projectifs de type fini. D'autre part nous précisons certains aspects de cette théorie liés au calcul formel et à l'analyse numérique matricielle.

Nous utiliserons une mesure assez grossière de la complexité des calculs sur machine : cette complexité sera mesurée essentiellement à travers le nombre d'opérations arithmétiques de base dans $\mathbf{A}$.

Nous supposerons en outre souvent qu'il y a sur l'anneau $\mathbf{A}$ un test explicite d'appartenance à un idéal de type fini (l'anneau est « fortement discret » selon la terminologie des mathématiques

*Partially supported by MCyT grant BFM 2002-04402-C02-0.

constructives). Par exemple un corps explicite est fortement discret si et seulement si il possède un test d'égalité à zéro. Nous supposerons aussi que ce test pour « $x \in \langle x_1, \ldots, x_n\rangle$? » (avec la réponse complète en cas d'appartenance) utilise un nombre d'opérations « élémentaires » borné par $\mathcal{O}(n^s)$ (nous ne précisons pas plus la nature exacte de ces opérations). Nous dirons alors que $\mathbf{A}$ est $\mathcal{O}(n^s)$-fortement discret. Notez que le test à zéro utilise donc un nombre d'opérations élémentaires borné par une constante.

Dans la suite une « opération élémentaire » sera ou bien une opération arithmétique de base dans l'anneau, ou bien l'une des opérations élémentaires qui interviennent dans le test d'appartenance à un idéal de type fini.

Par exemple on a facilement.

Lemme 0.1 *Sur un anneau $\mathcal{O}(n^s)$-fortement discret, on a un test pour déterminer si un idéal de type fini $\langle x_1, \ldots, x_n\rangle$ est idempotent et donner, en cas de réponse positive un générateur idempotent de l'idéal. Ce test utilise un nombre d'opérations arithmétiques en $\mathcal{O}(n^4)$ et un nombre d'autres opérations élémentaires en $\mathcal{O}(n^{2s+1})$.*

Preuve Résulte immédiatement du « déterminant trick » qui prouve qu'un idéal de type fini idempotent est engendré par un idempotent. On a besoin du résultat des n tests d'appartenance « $x_i \in \langle x_1, \ldots, x_n\rangle^2$? ». Le $\mathcal{O}(n^4)$ opérations arithmétiques provient du calcul du déterminant qui fournit l'idempotent recherché. □

Un système linéaire sur $\mathbf{A}$, présenté sous forme matricielle $AX = Y$ ($A \in \mathbf{A}^{m\times n}$), est particulièrement « agréable » si on peut calculer une solution (quand il en existe une) en fonction linéaire de Y, autrement dit, quand il existe une matrice $B \in \mathbf{A}^{n\times m}$ telle que $A\,B\,A\,X = A\,X$ pour tout X, i.e. $A\,B\,A = A$. Dans le cas où ceci est possible, nous disons que l'application linéaire définie par A est *localement simple*. Si en outre $B\,A\,B = B$ la matrice B est appelée un *inverse généralisé* de A.

La littérature sur le sujet des inverses généralisés est assez considérable. Nous renvoyons plus particulièrement à [1], [2], [3], [6] ou [15].

Pour ce qui concerne les modules projectifs de type fini qui donnent pour l'essentiel la même théorie sous une forme un peu plus abstraite, nous renvoyons à [11] et pour un traitement élémentaire et constructif à [7].

Nous citons maintenant quelques résultats significatifs obtenus dans le travail présent.

Nous devons d'abord introduire (ou rappeler) quelques définitions.

Soient E et F deux $\mathbf{A}$–modules. Deux applications linéaires $\varphi : E \to F$ et $\varphi^\bullet : F \to E$ sont dites *croisées* si on a :

$$\operatorname{Im}\varphi \oplus \operatorname{Ker}\varphi^\bullet = F\,, \quad \operatorname{Ker}\varphi \oplus \operatorname{Im}\varphi^\bullet = E \tag{1}$$

Nous notons Q_n la matrice diagonale ayant pour coefficient en position (k,k) la puissance t^{k-1} où t est une indéterminée. Si $A \in \mathbf{A}^{m\times n}$ on note A° la matrice ${Q_m}^{-1}\,A\,Q_n$.

L'anneau $\mathbf{A}(t)$ est le localisé $S^{-1}\mathbf{A}[t]$ où S est l'ensemble des polynômes primitifs (i.e., les coefficients engendrent l'idéal $\langle 1\rangle$).

Certains des énoncés qui suivent sont un peu moins précis que dans le texte.

Les deux premiers théorèmes que nous citons doivent sans doute se trouver dans la littérature. Du moins il est raisonnable de penser qu'ils font partie du folklore.

Théorèmes 4.1 et 4.5 *Soient E un $\mathbf{A}$–module projectif de type fini, $\varphi : E \to E$ une application linéaire et $\mathrm{P}_\varphi(Z) = \det(\mathrm{Id}_E + Z\varphi) = 1 + \sum_{\ell\geq 1} d_\ell\, Z^\ell$. Les propriétés suivantes sont équivalentes :*

1. *φ est croisée avec elle-même, et $\operatorname{Im}\varphi$ est un module projectif de rang k.*
2. *φ est de rang $\leq k$ et d_k est inversible.*

3. $\deg \mathrm{P}_\varphi \leq k$, d_k *est inversible et, en définissant* π *par*
$$\pi = d_{k-1}\varphi - d_{k-2}\varphi^2 + \cdots + (-1)^{k-1}\varphi^k,$$
on a les égalités $\pi\,\varphi = d_k\,\varphi$ *et* $\pi^2 = d_k\,\pi$.

Les théorèmes qui suivent sont, à notre connaissance, nouveaux.

Théorèmes 4.3 et 4.4 *Soient E et F deux $\mathbf{A}$–modules projectifs de type fini et deux applications linéaires $\varphi : E \to F$ et $\varphi^\bullet : F \to E$. Posons $\mathrm{P}_{\varphi\varphi^\bullet}(Z) = \det(\mathrm{Id}_F + Z\varphi\varphi^\bullet) = 1 + \sum_{\ell\geq 1} a_\ell\, Z^\ell$. Les propriétés suivantes sont équivalentes :*

1. φ *et* $\varphi^\bullet$ *sont croisées et* $\mathrm{Im}\,\varphi$ *est un module projectif de rang* k.
2. φ *et* $\varphi^\bullet$ *sont de rang* $\leq k$ *et* a_k *est inversible.*
3. $\deg \mathrm{P}_{\varphi\varphi^\bullet} \leq k$, a_k *est inversible et, en définissant* θ *par*
$$\theta = a_{k-1}\,\varphi^\bullet - a_{k-2}\,\varphi^\bullet\varphi\varphi^\bullet + \cdots + (-1)^{k-1}\varphi^\bullet(\varphi\varphi^\bullet)^{k-1},$$
on a les deux égalités $\varphi\,\theta\,\varphi = a_k\,\varphi$ *et* $\varphi^\bullet\,\varphi\,\theta = a_k\,\varphi^\bullet$.

Théorème 4.7 *Soient E et F deux $\mathbf{A}$–modules projectifs de type fini engendrés par n éléments (ou moins), et deux applications linéaires $\varphi : E \to F$ et $\varphi^\bullet : F \to E$. Alors on peut, avec un nombre d'opérations arithmétiques en $\mathcal{O}(n^4)$, et un nombre de tests « $x \in \langle y\rangle$? » en $\mathcal{O}(n^3)$, décider si φ et $\varphi^\bullet$ sont croisées, et en cas de réponse positive calculer des inverses généralisés de φ et $\varphi^\bullet$ en $\mathcal{O}(n^4)$ opérations arithmétiques.*

Théorème 5.5 *Soit une matrice $A \in \mathbf{A}^{m\times n}$. On pose $P(Z,t) = \det(\mathrm{I}_n + ZAA^\circ) = 1 + \sum_{\ell\geq 1} g_\ell(t)Z^\ell$. Les propriétés suivantes sont équivalentes :*

1. A *est localement simple de rang k sur* $\mathbf{A}$.
2. A *est localement simple de rang k sur* $\mathbf{A}(t)$.
3. A *et* A° *sont croisées sur* $\mathbf{A}(t)$, *de rang* k.
4. A *et* A° *sont croisées sur* $\mathbf{A}(t)$, $\deg_Z(P) \leq k$ *et le polynôme* $t^{k(n-k)}g_k(t)$ *est primitif.*
5. $\deg_Z(P) \leq k$, *le polynôme* $t^{k(n-k)}g_k(t)$ *est primitif et si on pose* $B(t) = g_{k-1}(t)A^\circ - g_{k-2}(t)A^\circ AA^\circ + \cdots + (-1)^{k-1}(A^\circ A)^{k-1}A^\circ$, *on a* $A\cdot B(t)\cdot A = g_k(t)\,A$.

Si $\mathbf{A}$ *est un anneau réduit, la dernière condition se simplifie en « $\deg_Z(P) \leq k$ et le polynôme $t^{k(n-k)}g_k(t)$ est primitif ». Lorsque les conditions sont vérifiées $B(t)/g_k(t)$ est un inverse généralisé de A sur l'anneau $\mathbf{A}(t)$.*

Théorème 5.7 *Soit une matrice $A \in \mathbf{A}^{m\times n}$. Les propriétés suivantes sont équivalentes :*

1. A *est localement simple sur* $\mathbf{A}$.
2. A *est localement simple sur* $\mathbf{A}(t)$.
3. A *et* A° *sont croisées sur* $\mathbf{A}(t)$.

Théorème 5.8 *Sur un anneau $\mathbf{A}$ fortement discret, on peut tester si une matrice $A \in \mathbf{A}^{m\times n}$ est localement simple, et en cas de réponse positive, calculer un inverse généralisé de la matrice. Soit $p = \min(m,n)$, $q = \max(m,n)$. Si l'anneau est $\mathcal{O}(n^s)$-fortement discret, ces calculs consomment $\mathcal{O}(p^6\,q^2 + p\,q^4)$ opérations arithmétiques et $\mathcal{O}(p^4\,q + pq^{2s+1})$ autres opérations élémentaires. Avec les mêmes bornes de complexité, on calcule un inverse généralisé de A et des matrices de projection sur le noyau et sur l'image de A.*

Dans nos calculs de complexité, nous avons utilisé les algorithmes de multiplication usuels pour les polynômes et les matrices. On peut donc améliorer les performances en utilisant des algorithmes de multiplication rapide.

Signalons enfin que les preuves de cet article reposent en partie sur des identités de Cramer généralisées (voir sections 1.1 et 2) dont nous avons eu du mal à trouver la trace dans la littérature. Nous remercions à ce sujet d'une part les statisticiens indiens et d'autre part Mustapha Rais pour un exposé à Poitiers dans lequel il interprétait les résultats de [4] au moyen de la théorie des invariants.

1 Identités de Cramer et premier inverse généralisé

1.1 Formules de Cramer usuelles et inusuelles

Une matrice $A \in \mathbf{A}^{m\times n}$ sera dite *de rang* $\leq k$ si tous les mineurs d'ordre $k+1$ sont nuls. Pour une matrice $A \in \mathbf{A}^{m\times n}$ nous noterons $A_{\alpha,\beta}$ la matrice extraite sur les lignes $\alpha = \{\alpha_1, \ldots, \alpha_r\} \subset \{1, \ldots, m\}$ et les colonnes $\beta = \{\beta_1, \ldots, \beta_s\} \subset \{1, \ldots, n\}$.

Si B est une matrice carrée d'ordre n, nous notons $\widetilde{B}$ ou $\operatorname{Adj} B$ la matrice cotransposée (on dit parfois adjointe). La forme élémentaire des identités de Cramer s'écrit alors $B\,\widetilde{B} = \widetilde{B}\,B = \det B\ \mathrm{I}_n$.

Supposons la matrice A de rang $\leq k$. Soit $V \in \mathbf{A}^{m\times 1}$ un vecteur colonne tel que $(A\,|\,V)$ soit aussi de rang $\leq k$. Appelons A_j la j-ème colonne de A. Soit $\mu_{\alpha,\beta} = \det(A_{\alpha,\beta})$ le mineur d'ordre k de la matrice A extrait sur les lignes $\alpha = \{\alpha_1, \ldots, \alpha_k\}$ et les colonnes $\beta = \{\beta_1, \ldots, \beta_k\}$. Pour $j = 1, \ldots, k$ soit $\nu_{\alpha,\beta,j}$ le déterminant de la même matrice extraite, à ceci près que la colonne j a été remplacée par la colonne extraite de V sur les lignes α. Alors on obtient pour chaque couple (α, β) de multi-indices et chaque $j \in \{1, \ldots, k\}$ une identité de Cramer :

$$\mu_{\alpha,\beta}\, V = \sum\nolimits_{j=1}^{k} \nu_{\alpha,\beta,j}\, A_{\beta_j} \tag{2}$$

due au fait que le rang de la matrice $(A_{1..m,\beta}\,|\,V)$ est $\leq k$. Ceci peut se relire comme suit :

$$\begin{aligned}
\mu_{\alpha,\beta}\, V &= \begin{bmatrix} A_{\beta_1} & \ldots & A_{\beta_k} \end{bmatrix} \begin{bmatrix} \nu_{\alpha,\beta,1} \\ \vdots \\ \nu_{\alpha,\beta,k} \end{bmatrix} = \\
&= \begin{bmatrix} A_{\beta_1} & \ldots & A_{\beta_k} \end{bmatrix} \operatorname{Adj}(A_{\alpha,\beta}) \begin{bmatrix} v_{\alpha_1} \\ \vdots \\ v_{\alpha_k} \end{bmatrix} = \\
&= A\,(\mathrm{I}_n)_{1..n,\beta}\operatorname{Adj}(A_{\alpha,\beta})\,(\mathrm{I}_m)_{\alpha,1..m}\,V
\end{aligned}$$

Ceci nous conduit à introduire la notation suivante

Notation 1.1 Nous notons $\mathcal{P}_{k,\ell}$ l'ensemble des parties à k éléments de $\{1, \ldots, \ell\}$. Pour $A \in \mathbf{A}^{m\times n}$ et $\alpha \in \mathcal{P}_{k,m}$, $\beta \in \mathcal{P}_{k,n}$ nous notons

$$\operatorname{Adj}_{\alpha,\beta}(A) := (\mathrm{I}_n)_{1..n,\beta}\operatorname{Adj}(A_{\alpha,\beta})\,(\mathrm{I}_m)_{\alpha,1..m}\,.$$

L'égalité précédente s'écrit alors :

$$\mu_{\alpha,\beta}\, V \;=\; A \operatorname{Adj}_{\alpha,\beta}(A)\, V \tag{3}$$

Comme conséquence on obtient, toujours sous l'hypothèse que A est de rang $\leq k$:

$$\mu_{\alpha,\beta}\, A \;=\; A \operatorname{Adj}_{\alpha,\beta}(A)\, A \tag{4}$$

Voici un exemple de l'égalité $\mu_{\alpha,\beta}\, V \;=\; A\,\mathrm{Adj}_{\alpha,\beta}(A)\,V$ pour voir la matrice $\mathrm{Adj}_{\alpha,\beta}(A)$. Supposons que nous avons le système linéaire :

$$\begin{bmatrix} 5 & -5 & 7 \\ 9 & 0 & 2 \\ 13 & 5 & -3 \end{bmatrix} X = \begin{bmatrix} 26 \\ 6 \\ -14 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix},$$

avec $\mathrm{rg}(A) = \mathrm{rg}(A\,|\,V) = 2$. Prenons $\alpha = \{1,2\}$ et $\beta = \{2,3\}$, alors :

$$\mu_{\alpha,\beta} = \begin{vmatrix} -5 & 7 \\ 0 & 2 \end{vmatrix},\ \sigma_{\alpha,\beta,1} = \begin{vmatrix} 26 & 7 \\ 6 & 2 \end{vmatrix},\ \sigma_{\alpha,\beta,2} = \begin{vmatrix} -5 & 26 \\ 0 & 6 \end{vmatrix},\ \mathrm{Adj}(A_{\alpha,\beta}) = \begin{bmatrix} 2 & -7 \\ 0 & -5 \end{bmatrix},$$

$$(\mathrm{I}_3)_{1..3,\beta} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},\quad (\mathrm{I}_3)_{\alpha,1..3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},\quad \mathrm{Adj}_{\alpha,\beta}(A) = \begin{bmatrix} 0 & 0 & 0 \\ 2 & -7 & 0 \\ 0 & -5 & 0 \end{bmatrix},$$

et

$$\begin{aligned} \mu_{\alpha,\beta}\, V &= \sigma_{\alpha,\beta,1}\, A_2 + \sigma_{\alpha,\beta,2}\, A_3 = \begin{bmatrix} A_2 & A_3 \end{bmatrix} \mathrm{Adj}(A_{\alpha,\beta}) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \\ &= A \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \mathrm{Adj}(A_{\alpha,\beta}) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} V = A\,\mathrm{Adj}_{\alpha,\beta}(A)\,V \end{aligned}$$

Définition 1.2 *Soit $A \in \mathbf{A}^{m\times n}$, les* idéaux déterminantiels de la matrice A *sont les idéaux*

$$\mathcal{D}_k(A) \;:=\; \textit{l'idéal engendré par les mineurs d'ordre } k \textit{ de la matrice } A$$

où k est un entier arbitraire. Pour $k \le 0$ les mineurs sont par convention égaux à 1, pour $k > \min(m,n)$ ils sont par convention égaux à 0. Si A est la matrice d'une application linéaire φ les idéaux $\mathcal{D}_k(A)$ ne dépendent que de φ et sont donc aussi appelés idéaux déterminantiels de l'application linéaire φ.

Les identités de Cramer vues précédemment fournissent des congruences qui ne sont soumises à aucune hypothèse : il suffit par exemple de lire (3) dans l'anneau quotient $\mathbf{A}/\mathcal{D}_{k+1}(A\,|\,V)$ pour obtenir la congruence (5).

Lemme 1.3 *Avec les notations précédentes mais sans aucune hypothèse sur la matrice A ou le vecteur V on a pour $\alpha \in \mathcal{P}_{k,m}$, $\beta \in \mathcal{P}_{k,n}$:*

$$\mu_{\alpha,\beta}\, V \;\equiv\; A\,\mathrm{Adj}_{\alpha,\beta}(A)\,V \qquad \mathrm{mod}\quad \mathcal{D}_{k+1}(A\,|\,V) \tag{5}$$

$$\mu_{\alpha,\beta}\, A \;\equiv\; A\,\mathrm{Adj}_{\alpha,\beta}(A)\,A \qquad \mathrm{mod}\quad \mathcal{D}_{k+1}(A)\,. \tag{6}$$

Une conséquence immédiate de l'identité de Cramer (4) est l'identité suivante moins usuelle.

Proposition 1.4 *Soit $A \in \mathbf{A}^{m\times n}$ de rang $\le k$ avec $\mathcal{D}_k(A) = \langle 1\rangle$. Précisément supposons*

$$\sum\nolimits_{\alpha\in\mathcal{P}_{k,m},\beta\in\mathcal{P}_{k,n}} c_{\alpha,\beta}\,\mu_{\alpha,\beta} = 1 \quad \textit{et posons} \quad B \;=\; \sum\nolimits_{\alpha\in\mathcal{P}_{k,m},\beta\in\mathcal{P}_{k,n}} c_{\alpha,\beta}\,\mathrm{Adj}_{\alpha,\beta}(A)\,.$$

Alors

$$A\,B\,A = A\,. \tag{7}$$

En conséquence $A\,B$ est une projection et $\mathrm{Im}\,A = \mathrm{Im}\,AB$ est facteur direct dans $\mathbf{A}^m$.

L'identité suivante est encore plus miraculeuse (voir [2] théorème 5.5).

Proposition 1.5 (Prasad et Robinson) *Avec les hypothèses et les notations de la proposition précédente, si* $\forall \alpha, \alpha' \in \mathcal{P}_{k,m}$, $\forall \beta, \beta' \in \mathcal{P}_{k,n}$ $c_{\alpha,\beta}\, c_{\alpha',\beta'} = c_{\alpha,\beta'}\, c_{\alpha',\beta}$, *alors*

$$B\,A\,B = B\,. \tag{8}$$

1.2 Applications linéaires simples et lemme de la liberté

Nous ne savons pas s'il existe une terminologie officielle pour la notion suivante.

Définition 1.6 *Une application linéaire* $\varphi : E \to F$ *entre deux* $\mathbf{A}$*–modules libres de dimensions finies est dite* simple (de rang k) *si, pour des bases convenables* $(e_1, \dots, e_n)$ *et* $(f_1, \dots, f_m)$ *de* E *et* F *on a :* $\varphi(e_i) = f_i$ *si* $i \leq k$ *et* $\varphi(e_i) = 0$ *si* $i > k$.

Il revient au même de dire que $\operatorname{Ker}\varphi$ et $\operatorname{Im}\varphi$ sont libres et admettent des supplémentaires libres. Ou encore que la matrice de φ sur des bases arbitraires de E et F s'écrit $A = U\,\mathrm{I}_{k,m,n}\,V$ avec U et V inversibles, et $\mathrm{I}_{k,m,n} \in \mathbf{A}^{m\times n}$ est de la forme $\begin{bmatrix} \mathrm{I}_k & 0 \\ 0 & 0 \end{bmatrix}$.

Si on pose $B = V^{-1}\,\mathrm{I}_{k,n,m}\,U^{-1}$ on a immédiatement

$$A\,B\,A = A \quad \text{et} \quad B\,A\,B = B\,.$$

L'application linéaire $x \mapsto ax$ de $\mathbf{A}$ dans $\mathbf{A}$ est simple si et seulement si a est nul ou inversible.

Le rang d'une application linéaire simple est bien défini dès que l'anneau n'est pas trivial. Avec l'anneau trivial par contre, toutes les applications linéaires sont simples, de tous rangs (cette remarque est nécessaire pour admettre sans réticence le lemme 1.7 ainsi que le point 8 du théorème 3.3).

Le lemme suivant (voir [7]) est immédiat.

Lemme 1.7 (lemme de la liberté) *Soit* $\varphi : E \to F$ *une application linéaire de rang* $\leq k$ *entre deux* $\mathbf{A}$*–modules libres de dimensions finies. Soit* A *une matrice représentant* φ *sur des bases de* E *et* F. *Soit* μ *un mineur d'ordre* k *de* A. *Si* μ *est inversible,* φ *est simple de rang* k. *En particulier* φ *est toujours simple de rang* k *sur l'anneau* $\mathbf{A}[1/\mu]$.

1.3 Systèmes fondamentaux d'idempotents orthogonaux

Un système fondamental d'idempotents orthogonaux (sfio) est une famille finie $(r_i)_{1\leq i\leq n}$ qui vérifie $r_i r_j = 0$ si $i \neq j$ et $\sum_{i=1}^n r_i = 1$. Il revient au même de se donner un tel système dans $\mathbf{A}$ ou de se donner un isomorphisme $\mathbf{A} \to \mathbf{A}_1 \times \cdots \times \mathbf{A}_n$. L'idempotent r_i dans $\mathbf{A}$ correspond alors au « vecteur » $(0, \dots, 0, 1, 0, \dots 0)$ avec 1 en position i dans $\mathbf{A}_1 \times \cdots \times \mathbf{A}_n$. Chaque $\mathbf{A}_i$ est isomorphe à $\mathbf{A}[1/r_i] \simeq \mathbf{A}\,/\langle 1 - r_i\rangle$, ou encore à l'idéal $r_i\mathbf{A}$ qu'on considère comme un anneau unitaire en prenant r_i comme élément neutre pour la multiplication (attention, ce n'est pas un sous anneau de $\mathbf{A}$, parce que le neutre n'est pas le même).

Dans nos énoncés, nous ne supposerons pas que tous les r_i dans un sfio sont non nuls. Cela nous simplifie la vie (et les énoncés) notamment lorsqu'on n'a pas de test d'égalité à 0 dans l'anneau considéré. Il faut simplement se rappeler que l'anneau $\mathbf{A}[1/0]$ est trivial pour comprendre pourquoi les énoncés restent justes.

Une généralisation naturelle de la notion d'application linéaire simple lorsque l'anneau possède des idempotents est la suivante.

Définition 1.8 *Une application linéaire* $\varphi : E \to F$ *entre deux* $\mathbf{A}$*–modules libres de dimensions finies est dite* quasi-simple *si, pour des bases convenables* $(e_1, \dots, e_n)$ *et* $(f_1, \dots, f_m)$ *de* E *et* F

on a : $\varphi(e_i) = r_i f_i$ $(1 \leq i \leq \inf(m,n))$ *où les* r_i *sont des idempotents vérifiant* $r_i r_{i+1} = r_{i+1}$ $(1 \leq i < \inf(m,n))$*, et si et* $\varphi(e_i) = 0$ *pour* $i > \inf(m,n)$.

Posons $r_0 = 1$, $r_{\inf(m,n)+1} = 0$ et $s_i = r_i - r_{i+1}$ $(0 \leq i \leq \inf(m,n))$. Alors les s_i forment un sfio et φ devient simple de rang k lorsqu'on étend les scalaires à l'anneau $\mathbf{A}[1/s_k]$.

Réciproquement il est facile de voir qu'une application linéaire qui devient simple chaque fois qu'on localise en les éléments d'un sfio est quasi-simple.

1.4 Inverses généralisés et applications linéaires croisées

Dans les sections suivantes, nous donnerons plusieurs généralisations du résultat des propositions 1.4 et 1.5, qui nous donnent notre premier « inverse généralisé ». La terminologie concernant les inverses généralisés ne semble pas entièrement fixée. Nous adoptons celle de [6]. Dans [2] l'auteur utilise le terme « reflexive g-inverse » :

Définition 1.9 *Soient* E *et* F *deux* $\mathbf{A}$*–modules, et une application linéaire* $\varphi : E \to F$*. Une application linéaire* $\psi : F \to E$ *est appelée un* inverse généralisé *de* φ *si on a* $\varphi \circ \psi \circ \varphi = \varphi$ *et* $\psi \circ \varphi \circ \psi = \psi$.

Dans ces conditions, on vérifie que $\varphi\,\psi$ et $\psi\,\varphi$ sont des projections, que $\operatorname{Im}\varphi = \operatorname{Im}\varphi\,\psi$, $\operatorname{Im}\psi = \operatorname{Im}\psi\,\varphi$, $\operatorname{Ker}\varphi = \operatorname{Ker}\psi\,\varphi$, $\operatorname{Ker}\psi = \operatorname{Ker}\varphi\,\psi$, et donc $E = \operatorname{Ker}\varphi \oplus \operatorname{Im}\psi$ et $F = \operatorname{Ker}\psi \oplus \operatorname{Im}\varphi$.

Si on a une application linéaire ψ_1 vérifiant $\varphi\,\psi_1\,\varphi = \varphi$ on obtient un inverse généralisé en posant $\psi = \psi_1\,\varphi\,\psi_1$.

Le lemme suivant décrit les inverses généralisés d'une application linéaire simple.

Lemme 1.10 *Soit* E *et* F *des modules libres de dimensions finies et* $\varphi : E \to F$ *une application linéaire simple dont la matrice sur des bases fixées est* $A = U\,\mathrm{I}_{r,m,n}\,V$ *(*U *et* V *sont inversibles, cf. définition 1.6). Alors les inverses généralisés de* φ *sont toutes les applications linéaires ayant (sur les mêmes bases) une matrice* $B \in \mathbf{A}^{n\times m}$ *de la forme suivante (avec* $C \in \mathbf{A}^{r\times(m-r)}$ *et* $D \in \mathbf{A}^{(n-r)\times r}$*) :*

$$B = V^{-1} \begin{bmatrix} \mathrm{I}_r & C \\ D & DC \end{bmatrix} U^{-1}$$

Si ψ est un inverse généralisé de φ, alors φ et ψ sont croisées.

Réciproquement, la connaissance d'une application linéaire croisée avec φ permet de calculer un inverse généralisé de φ. En effet si $\varphi^\bullet$ est croisée avec φ, φ se restreint en un isomorphisme φ_0 de $\operatorname{Im}\varphi^\bullet$ sur $\operatorname{Im}\varphi$ et $\varphi^\bullet$ se restreint en un isomorphisme $\varphi_0^\bullet$ de $\operatorname{Im}\varphi$ sur $\operatorname{Im}\varphi^\bullet$. On a :

$$\begin{array}{rclcrcl} \operatorname{Im}\varphi & = & \operatorname{Im}\varphi\varphi^\bullet\,, & \qquad & \operatorname{Ker}\varphi^\bullet & = & \operatorname{Ker}\varphi\varphi^\bullet\,, \\ \operatorname{Ker}\varphi & = & \operatorname{Ker}\varphi^\bullet\varphi\,, & & \operatorname{Im}\varphi^\bullet & = & \operatorname{Im}\varphi^\bullet\varphi\,. \end{array} \tag{9}$$

Notons $\pi_{\operatorname{Im}\varphi} : F \to F$ la projection sur $\operatorname{Im}\varphi$ parallèlement à $\operatorname{Ker}\varphi^\bullet$. On définit l'application linéaire $\psi : F \to E$ par

$$\forall y \in F \qquad \psi(y) = \varphi_0^{-1}(\pi_{\operatorname{Im}\varphi}(y)). \tag{10}$$

Il est alors clair que ψ convient comme « inverse généralisé de φ via $\varphi^\bullet$ » au sens du théorème suivant.

Théorème et définition 1.11 *Si* $\varphi : E \to F$ *et* $\varphi^\bullet : F \to E$ *sont croisées il existe une unique application linéaire* $\psi : F \to E$ *vérifiant les deux conditions :*

1. $\varphi \circ \psi$ *est la projection sur* $\operatorname{Im}\varphi$ *parallèlement à* $\operatorname{Ker}\varphi^\bullet$ *;*
2. $\psi \circ \varphi$ *est la projection sur* $\operatorname{Im}\varphi^\bullet$ *parallèlement à* $\operatorname{Ker}\varphi$.

Cette application linéaire ψ peut être aussi caractérisée par les 4 égalités suivantes :

$$\varphi\circ\psi\circ\varphi=\varphi,\qquad \psi\circ\varphi\circ\psi=\psi,\qquad \varphi^\bullet\circ\varphi\circ\psi=\varphi^\bullet,\qquad \psi\circ\varphi\circ\varphi^\bullet=\varphi^\bullet. \tag{11}$$

Nous dirons que ψ est l'inverse généralisé de φ via $\varphi^\bullet$ *et nous le noterons* $\psi=\mathrm{Ig}(\varphi,\varphi^\bullet)=\varphi^{\dagger_{\varphi^\bullet}}$.

Preuve Il nous reste à voir que les quatre égalités suffisent.
Puisque $\varphi\,\psi\,\varphi=\varphi$ et $\psi\,\varphi\,\psi=\psi$ on a $E=\mathrm{Ker}\,\varphi\oplus\mathrm{Im}\,\psi$, $F=\mathrm{Ker}\,\psi\oplus\mathrm{Im}\,\varphi$, $\varphi\,\psi$ est la projection sur $\mathrm{Im}\,\varphi$ parallèlement à $\mathrm{Ker}\,\psi$ et $\psi\,\varphi$ est la projection sur $\mathrm{Im}\,\psi$ parallèlement à $\mathrm{Ker}\,\varphi$. Il nous suffit donc de montrer que $\mathrm{Ker}\,\varphi^\bullet=\mathrm{Ker}\,\psi$ et $\mathrm{Im}\,\varphi^\bullet=\mathrm{Im}\,\psi$.
La troisième égalité implique $\mathrm{Ker}\,\psi\subset\mathrm{Ker}\,\varphi^\bullet$. On conclut $\mathrm{Ker}\,\varphi^\bullet=\mathrm{Ker}\,\psi$ en remarquant que $F=\mathrm{Ker}\,\psi\oplus\mathrm{Im}\,\varphi=\mathrm{Ker}\,\varphi^\bullet\oplus\mathrm{Im}\,\varphi$.
De même la dernière égalité implique $\mathrm{Im}\,\varphi^\bullet\subset\mathrm{Im}\,\psi$ et on conclut de la même façon. □

Le théorème précédent correspond à la définition donnée par Moore, dans le cas d'espaces vectoriels hermitiens, avec pour $\varphi^\bullet$ la conjuguée φ^*, ce qui donne des projections orthogonales et l'inverse de Moore-Penrose.

Lemme 1.12 *Si φ et $\varphi^\bullet$ sont croisées, alors $\varphi\varphi^\bullet$ est croisée avec elle-même (même chose pour $\varphi^\bullet\varphi$). En outre si $\theta=\mathrm{Ig}(\varphi\varphi^\bullet,\varphi\varphi^\bullet)$ alors $\varphi\varphi^\bullet\theta=\theta\varphi\varphi^\bullet$, $\theta=\mathrm{Ig}(\varphi^\bullet,\varphi)\,\mathrm{Ig}(\varphi,\varphi^\bullet)$, $\varphi^\bullet\theta=\mathrm{Ig}(\varphi,\varphi^\bullet)$ et $\theta\varphi=\mathrm{Ig}(\varphi^\bullet,\varphi)$.*

On a une caractérisation purement équationnelle de la situation du théorème 1.11, à condition d'introduire les deux inverses généralisés.

Proposition 1.13 *Soient E et F deux $\mathbf{A}$–modules et deux applications linéaires $\varphi:E\to F$ et $\varphi^\bullet:F\to E$.*

1. *Si φ et $\varphi^\bullet$ sont croisées, posons $\psi=\mathrm{Ig}(\varphi,\varphi^\bullet)$ et $\psi^\bullet=\mathrm{Ig}(\varphi^\bullet,\varphi)$. On a :*

$$\begin{array}{llll}\varphi\circ\psi\circ\varphi&=\varphi & \varphi^\bullet\circ\psi^\bullet\circ\varphi^\bullet&=\varphi^\bullet & \varphi\circ\psi&=\psi^\bullet\circ\varphi^\bullet\\ \psi\circ\varphi\circ\psi&=\psi & \psi^\bullet\circ\varphi^\bullet\circ\psi^\bullet&=\psi^\bullet & \psi\circ\varphi&=\varphi^\bullet\circ\psi^\bullet\end{array} \tag{12}$$

2. *Réciproquement si ψ et $\psi^\bullet$ vérifient les égalités (12), alors φ et $\varphi^\bullet$ sont croisées, $\psi=\mathrm{Ig}(\varphi,\varphi^\bullet)$ et $\psi^\bullet=\mathrm{Ig}(\varphi^\bullet,\varphi)$.*

Un cas particulier est le suivant :

Proposition 1.14 *Soit E un $\mathbf{A}$–module et une application linéaire $\varphi:E\to E$.*

1. *Si φ est croisée avec elle-même, posons $\psi=\mathrm{Ig}(\varphi,\varphi)$. On a :*

$$\varphi\circ\psi\circ\varphi=\varphi\qquad \psi\circ\varphi\circ\psi=\psi\qquad \varphi\circ\psi=\psi\circ\varphi \tag{13}$$

2. *Réciproquement si ψ vérife les égalités (13), alors φ est croisée avec elle-même et $\psi=\mathrm{Ig}(\varphi,\varphi)$.*

Dans [2], lorsque sont vérifiées les égalités (13), ψ est appelé un « group inverse » de φ.

1.5 Le cas des modules de type fini

Nous développons maintenant un petit peu d'algèbre linéaire sur les modules de type fini, en donnant quelques résultats bien connus pour les espaces vectoriels de dimension finie qui généralisent de manière parfois inattendue.

Proposition 1.15 ([9] chap. III, exo. 9 p. 80) *Soit E un $\mathbf{A}$–module de type fini et $\varphi:E\to E$ une application linéaire surjective. Alors φ est un isomorphisme.*

Proposition 1.16 *Soit E un $\mathbf{A}$–module de type fini et une application linéaire $\varphi : E \to E$. Les propriétés suivantes sont équivalentes :*

1. $E = \operatorname{Im}\varphi \oplus \operatorname{Ker}\varphi$ *(i.e. φ est croisée avec elle-même).*
2. $E = \operatorname{Im}\varphi + \operatorname{Ker}\varphi$
3. $\operatorname{Im}\varphi = \operatorname{Im}\varphi^2$.

Preuve 1 implique clairement 2 et 3.
2 implique 3 : Tout $x \in E$ s'écrit $x = \varphi(y) + z$ avec $\varphi(z) = 0$ donc tout $\varphi(x) \in \operatorname{Im}\varphi$ s'écrit $\varphi^2(y)$.
3 implique 2 : Si $\varphi(x) = \varphi(\varphi(y))$ alors $\varphi(x - \varphi(y)) = 0$ donc $x = \varphi(y) + z$ avec $\varphi(z) = 0$.
2 et 3 impliquent 1 : Soit $\varphi_0 : \operatorname{Im}\varphi \to \operatorname{Im}\varphi$ obtenue par restriction de φ. Le module $\operatorname{Im}\varphi$ est de type fini puisque E est de type fini. Mais φ_0 est surjective par hypothèse. Donc, par la proposition 1.15 φ_0 est bijective. Ceci implique clairement $\operatorname{Ker}\varphi \cap \operatorname{Im}\varphi = 0$. □

De la même façon :

Proposition 1.17 *Soient E et F deux $\mathbf{A}$–modules de type fini. Des applications linéaires $\varphi : E \to F$ et $\varphi^\bullet : F \to E$ telles que $\operatorname{Im}\varphi^\bullet + \operatorname{Ker}\varphi = E$ et $\operatorname{Im}\varphi + \operatorname{Ker}\varphi^\bullet = F$ sont croisées.*

Preuve Si $\operatorname{Im}\varphi^\bullet + \operatorname{Ker}\varphi = E$ alors $\operatorname{Im}\varphi \simeq E/\operatorname{Ker}\varphi \simeq \operatorname{Im}\varphi^\bullet/(\operatorname{Ker}\varphi \cap \operatorname{Im}\varphi^\bullet)$. De manière symétrique $\operatorname{Im}\varphi^\bullet$ est isomorphe à un quotient de $\operatorname{Im}\varphi$. En composant ces deux isomorphismes on trouve que $\operatorname{Im}\varphi$ est isomorphe à un quotient de lui-même par un sous-module plus grand que $\operatorname{Ker}\varphi \cap \operatorname{Im}\varphi^\bullet$. La proposition 1.15 implique donc que $\operatorname{Ker}\varphi \cap \operatorname{Im}\varphi^\bullet = 0$. Même chose pour $\operatorname{Ker}\varphi^\bullet \cap \operatorname{Im}\varphi$. □

2 Interprétation de l'inverse généralisé avec des identités de Cramer

Lorsqu'on a une matrice carrée A d'ordre n, il y a deux manières très différentes de calculer sa matrice cotransposée $\operatorname{Adj} A$. La première consiste à calculer ses coefficients qui sont, au signe près, des mineurs d'ordre $n-1$ de A. La seconde consiste à utiliser le théorème de Cayley-Hamilton qui nous fournit un polynôme $Q(X)$, facilement déduit du polynôme caractéristique, vérifiant $AQ(A) = \det A\ \mathrm{I}_n$. Alors $\operatorname{Adj} A = Q(A)$. Cette coïncidence peut être vue comme une famille d'identités algébriques remarquables. Dans cette section nous généralisons ce résultat « en rang $k < n$ ». Les choses sont cependant un peu plus délicates et il est plus pratique de travailler avec deux applications linéaires.

Dans cette section E et F sont des modules libres de dimensions finies. On considère deux applications linéaires $\varphi : E \to F$ et $\varphi^\bullet : F \to E$. On ne suppose pas a priori que φ et $\varphi^\bullet$ sont croisées. Soient A et $A^\bullet$ des matrices pour φ et $\varphi^\bullet$ sur des bases fixées de E et F. On note pour simplifier $\mu_{\alpha,\beta} = \det(A_{\alpha,\beta})$ et $\mu^\bullet_{\beta,\alpha} = \det(A^\bullet_{\beta,\alpha})$.
La formule de Binet-Cauchy montre que :

Lemme 2.1 *Si p est la plus petite des dimensions de E et F et si $\det(\mathrm{Id}_E + Z\,\varphi^\bullet\varphi) = 1 + a_1 Z + \cdots + a_p Z^p$ alors, pour tout $k \le p$:*

$$a_k = \sum\nolimits_{\alpha\in\mathcal{P}_{k,m},\beta\in\mathcal{P}_{k,n}} \mu^\bullet_{\beta,\alpha}\,\mu_{\alpha,\beta}\,.$$

Notation 2.2 *On reprend les hypothèses précédentes. On notera* $\mathcal{G}^{(k)}_{\varphi^\bullet}(\varphi) = \mathrm{d}_k(\varphi^\bullet\varphi) = a_k$. *Nous les appelerons des* coefficients de Gram mixtes. *Enfin nous définissons* $\mathrm{Adj}_{\varphi^\bullet,k}(\varphi)$ *et* $\mathrm{Adj}^{(k)}_{\varphi^\bullet}(\varphi)$ *par :*

$$\mathrm{Adj}_{\varphi^\bullet,k}(\varphi) \;=\; \sum\nolimits_{\alpha\in\mathcal{P}_{k,m},\beta\in\mathcal{P}_{k,n}} \mu^\bullet_{\beta,\alpha}\,\mathrm{Adj}_{\alpha,\beta}(\varphi) \tag{14}$$

$$\mathrm{Adj}^{(k)}_{\varphi^\bullet}(\varphi) \;=\; a_{k-1}\,\varphi^\bullet - a_{k-2}\,\varphi^\bullet\varphi\,\varphi^\bullet + \cdots + (-1)^{k-1}(\varphi^\bullet\varphi)^{k-1}\varphi^\bullet\,. \tag{15}$$

A priori l'application linéaire que nous avons notée $\mathrm{Adj}_{\varphi^\bullet,k}(\varphi)$ dépend du choix des bases de E et F. Nous allons voir bientôt qu'il n'en est rien. En effet nous allons montrer :

Théorème 2.3 *On a toujours :*

$$\mathrm{Adj}^{(k)}_{\varphi^\bullet}(\varphi) \;=\; \mathrm{Adj}_{\varphi^\bullet,k}(\varphi) \tag{16}$$

En fait ces applications linéaires sont aussi égales au gradient *de la fonction* $\varphi \mapsto \mathrm{d}_k(\varphi^\bullet\varphi)$ *(notez bien que* $\varphi^\bullet$ *est ici une constante).*

Pour préciser la dernière phrase, nous devons donner la définition du gradient d'une fonction polynomiale $\mathcal{L}(E,F) \to \mathbf{A}$ (c'est-à-dire une fonction qui est donnée par un polynome en les entrées de la matrice A de $\varphi \in \mathcal{L}(E,F)$ une fois choisies des bases de E et F). Il ne s'agit de rien d'autre que la différentielle de la fonction, traduite sous forme d'un élément $\theta \in \mathcal{L}(F,E)$ en utilisant la dualité canonique entre $\mathcal{L}(E,F)$ et $\mathcal{L}(F,E)$ donnée par la forme bilinéaire « trace du produit ».

Définition 2.4 *Soit* $a : \mathcal{L}(E,F) \to \mathbf{A}$ *une fonction polynomiale. On appelle gradient de* a *au point* φ *et on note* $\nabla(a)(\varphi)$ *l'unique application linéaire* $\theta \in \mathcal{L}(F,E)$ *telle que* $a(\varphi+\epsilon) = a(\varphi) + \mathrm{Tr}(\theta\,\epsilon) + \mathcal{O}(\epsilon^{(2)})$, *où* $\mathcal{O}(\epsilon^{(2)})$ *désigne sous forme abrégée une fonction polynomiale de* ϵ *sans terme constant ni terme du premier degré.*

Preuve du théorème 2.3 On utilise un fait bien connu et deux lemmes qui s'en déduisent simplement. Nous donnons les preuves pour faciliter la lecture de l'article.

Fait 2.5 *Pour un endomorphisme* ψ *d'un* $\mathbf{A}$*–module* F *libre de rang* m *on a*

$$\nabla(\det)(\psi) = \mathrm{Adj}(\psi)\,.$$

Preuve du fait 2.5
Voici une première preuve. Raisonnons avec des matrices carrées. On a $\det(\mathrm{I}_m+H) = 1+\mathrm{Tr}\,H+\mathcal{O}(H^{(2)})$ où $\mathcal{O}(H^{(2)})$ est comme dans la définition 2.4. Donc $\nabla(\det)(\mathrm{I}_m) = \mathrm{I}_m$. Si A est inversible on a $\det(A+H) = \det(A)\det(\mathrm{I}_m + A^{-1}H) = \det(A) + \det(A)\mathrm{Tr}\,A^{-1}H + \mathcal{O}(H^{(2)})$. Comme $\det(A)\mathrm{Tr}\,A^{-1}H = \mathrm{Tr}(\det(A)A^{-1}H)$ et $\mathrm{Adj}\,A = \det(A)A^{-1}$ cela donne $\nabla(\det)(A) = \mathrm{Adj}\,A$. On conclut en remarquant qu'on vient de démontrer, sous la condition « A inversible » une identité algébrique dans laquelle on n'a pas précisé le contenu exact du terme $\mathcal{O}(H^{(2)})$. Mais puisqu'il s'agit bien d'une identité algébrique, il suffisait de la démontrer pour A dans un ouvert de $\mathbb{Q}^{n\times n}$. Une autre preuve est la suivante : si A_i (resp. H_i) désigne la i-ème colonne de A (resp. de H), il est clair que la différentielle de det au point A est l'application linéaire

$$H \mapsto \det(H_1, A_2, \ldots, A_n) + \cdots + \det(A_1, \ldots, A_{n-1}, H_n)\,.$$

Par ailleurs l'égalité

$$\det(H_1, A_2, \ldots, A_n) + \cdots + \det(A_1, \ldots, A_{n-1}, H_n) = \mathrm{Tr}(\mathrm{Adj}(A)\,H)$$

résulte clairement des identités de Cramer (cf. par exemple notre section 1.1). □

Un premier corollaire immédiat est le lemme suivant.

Lemme 2.6 *On fixe des bases de E et F. Le gradient de la fonction $\mathcal{L}(E,F) \to \mathbf{A} : \varphi \mapsto \mu_{\alpha,\beta} = \det(A_{\alpha,\beta})$ (où A est la matrice de φ) au point φ est l'endomorphisme « cotransposé en (α,β) » ayant pour matrice $\mathrm{Adj}_{\alpha,\beta}(A)$.*

Preuve du lemme 2.6
Raisonnons avec des matrices. Notons $J_\beta = (\mathrm{I}_n)_{1..n,\beta}$ et $P_\alpha = (\mathrm{I}_m)_{\alpha,1..m}$ les deux matrices telles que $P_\alpha\, A\, J_\beta = A_{\alpha,\beta}$. Puisque l'application $\lambda : A \mapsto A_{\alpha,\beta}$ est linéaire, la différentielle de $A \mapsto \det A_{\alpha,\beta}$ calculée au point A pour l'accroissement H est donnée par

$$\mathrm{Tr}(\mathrm{Adj}(A_{\alpha,\beta})(\lambda(H))) = \mathrm{Tr}(\mathrm{Adj}(A_{\alpha,\beta})\, P_\alpha\, H\, J_\beta) = \mathrm{Tr}(J_\beta\, \mathrm{Adj}(A_{\alpha,\beta})\, P_\alpha\, H) = \mathrm{Tr}(\mathrm{Adj}_{\alpha,\beta}(A)\, H)\,.$$

Autrement dit

$$\nabla(M \mapsto \det M_{\alpha,\beta})(A) = \mathrm{Adj}_{\alpha,\beta}(A)\,.$$

□

Vu le lemme 2.1, un corollaire de ce lemme est que l'application linéaire $\mathrm{Adj}_{\varphi^\bullet,k}(\varphi)$ est le gradient de la fonction $\varphi \mapsto \mathrm{d}_k(\varphi^\bullet\varphi)$. En particulier, malgré les apparences de sa définition, cette application linéaire ne dépend que de φ, $\varphi^\bullet$ et k, et non des bases choisies.
L'autre lemme, bien connu en théorie des invariants (voir par exemple [14, 16, 17]), est :

Lemme 2.7 *Pour un endomorphisme ψ d'un $\mathbf{A}$–module libre F on a*

$$\nabla(\mathrm{d}_k)(\psi) = \mathrm{d}_{k-1}(\psi)\,\mathrm{Id}_F - \mathrm{d}_{k-2}(\psi)\,\psi + \mathrm{d}_{k-3}(\psi)\,\psi^2 + \cdots + (-1)^{k-1}\psi^{k-1}\,.$$

Preuve du lemme 2.7
On pose $\mathbf{B} = \mathcal{L}(F,F)$ et on identifie $\mathbf{B}[X]$ avec $\mathcal{L}(F[X],F[X])$. On considère $\mathbf{B}$ comme une $\mathbf{A}$-algèbre et $\mathbf{B}[X]$ comme une $\mathbf{A}[X]$- algèbre. Nous dérivons la fonction

$$\delta : \mathbf{B} \to \mathbf{A}[X] : \psi \mapsto \det(1_\mathbf{B} + X\psi) = 1_\mathbf{A} + \mathrm{d}_1(\psi)\,X + \cdots + \mathrm{d}_m(\psi)\,X^m\,.$$

Cette fonction est obtenue en composant la fonction affine

$$\mathbf{B} \to \mathbf{B}[X] : \psi \mapsto 1_\mathbf{B} + X\psi$$

et la fonction det : $\mathbf{B}[X] \to \mathbf{A}[X]$. Calculons cette différentielle au point ψ pour un accroissement ϵ. Nous obtenons l'application linéaire

$$\epsilon \mapsto \mathrm{Tr}(\mathrm{Adj}(1_\mathbf{B} + X\psi)\,(X\,\epsilon)) = \mathrm{Tr}((X\,\mathrm{Adj}(1_\mathbf{B} + X\psi))\,\epsilon)\,.$$

L'application linéaire $\eta(X) = \mathrm{Adj}(1_\mathbf{B} + X\psi) = 1_\mathbf{B} + \eta_1\,X + \cdots + \eta_{m-1}\,X^{m-1}$ (avec les $\eta_i \in \mathbf{B}$) est donc égale à

$$1_\mathbf{B} + \nabla(\mathrm{d}_2)(\psi)\,X + \cdots + \nabla(\mathrm{d}_m)(\psi)\,X^{m-1}\,.$$

Elle vérifie

$$\eta(X)\,(1_\mathbf{B} + X\psi) = \det(1_\mathbf{B} + X\psi)\,1_\mathbf{B} = (1_\mathbf{A} + \mathrm{d}_1\,X + \cdots + \mathrm{d}_m\,X^m)\,1_\mathbf{B}\,.$$

On obtiendra donc η comme élément de $\mathbf{B}[X]$ en faisant dans $\mathbf{B}[X]$ la division par puissances croissantes du polynôme $1_\mathbf{B} + \mathrm{d}_1\,X + \cdots + \mathrm{d}_n\,X^m$ par $1_\mathbf{B} + X\psi$ (le fait que la division est exacte, i.e., le reste est nul, fournit l'une des preuves usuelles du théorème de Cayley-Hamilton). Et cela donne le résultat annoncé. □

On déduit enfin du lemme 2.7 que l'application linéaire $\mathrm{Adj}^{(k)}_{\varphi^\bullet}(\varphi)$ est le gradient de la fonction $\varphi \mapsto \mathrm{d}_k(\varphi^\bullet\varphi)$. Nous devons en effet dériver la fonction obtenue en composant la fonction linéaire $\varphi \mapsto \varphi^\bullet\varphi$ et la fonction d_k : le gradient correspondant est bien $(\nabla(\mathrm{d}_k)(\varphi^\bullet\varphi))\,\varphi^\bullet$. □

Le théorème qui suit nous sera particulièrement utile dans la section 5.2.

Théorème 2.8 *On a avec les notations précédentes si φ est de rang $\leq k$:*

$$\varphi \circ \mathrm{Adj}^{(k)}_{\varphi^\bullet}(\varphi) \circ \varphi \quad = \quad a_k\,\varphi. \tag{17}$$

Preuve Une conséquence immédiate de l'identité de Cramer (6) est :

$$\varphi \circ \mathrm{Adj}_{\varphi^\bullet,k}(\varphi) \circ \varphi \quad \equiv \quad a_k\,\varphi \qquad \mathrm{mod} \quad \mathcal{D}_{k+1}(\varphi). \tag{18}$$

On conclut par le théorème 2.3. □

Un cas particulier de la formule (19) qui suit est la formule 2.13 dans [13]. La signification est qu'un inverse généralisé d'une application linéaire φ calculé en utilisant une application linéaire $\varphi^\bullet$ croisée avec φ donne la solution du système linéaire correspondant $AX = V$ (A est la matrice de φ) sous forme d'une moyenne pondérée d'identités de Cramer du type (3) page 4.

Théorème 2.9 *Si φ et $\varphi^\bullet$ sont croisées de rang k, alors*

$$\mathrm{Ig}(\varphi,\varphi^\bullet) = a_k^{-1}\,\mathrm{Adj}_{\varphi^\bullet,k}(\varphi) \tag{19}$$

Preuve Si φ et $\varphi^\bullet$ sont croisées de rang k on obtient en appliquant la formule (23) page 19 l'égalité

$$\mathrm{Ig}(\varphi,\varphi^\bullet) = a_k^{-1}\,\mathrm{Adj}^{(k)}_{\varphi^\bullet}(\varphi).$$

On conclut par le théorème 2.3. □

Notons cependant que pour le calcul de l'inverse généralisé ce n'est pas la formule (14) qui peut servir en pratique, mais plutôt la formule (15).

3 Modules projectifs de type fini

Cette section résume un certain nombre de résultats plus ou moins classiques. On trouve la plupart d'entre eux très bien exposés dans [11]. Pour des preuves entièrement constructives ont peut consulter [7]. Rappelons qu'un module est dit *projectif de type fini* s'il est isomorphe à un facteur direct dans un $\mathbf{A}$–module libre de dimension finie.

3.1 Idéaux de Fitting et applications linéaires localement simples

Si $G \in \mathbf{A}^{m\times n}$, le module $\mathrm{Coker}(G)$ est dit *de présentation finie*. Plus généralement on dit que la matrice G est *une présentation d'un module M* si on a des générateurs $g_1,\ldots,g_m$ de M et si l'application $\mathbf{A}^m \to M$ qui envoie la base canonique sur les g_i identifie M et $\mathrm{Coker}\,G$, c'est-à-dire encore si les colonnes de G engendrent le module des relations entre les g_i.

Définition 3.1 *Si G est une matrice de présentation d'un module M donné par m générateurs liés par n relations, les* idéaux de Fitting du module M *sont les idéaux*

$$\mathcal{F}_k(M) := \mathcal{D}_{m-k}(G)$$

où k est un entier arbitraire. Ces idéaux ne dépendent que de M et non de la présentation choisie pour M.

Définition 3.2 *Des éléments $x_1, \ldots, x_\ell$ de $\mathbf{A}$ sont dit* comaximaux *s'ils engendrent $\mathbf{A}$ comme idéal, c'est-à-dire si une combinaison linéaire des x_i est égale à 1. On dit encore que le vecteur $(x_1, \ldots, x_\ell)$ est* unimodulaire *et que le polynôme $P(Z) = \sum_k x_k Z^{k-1}$ est* primitif.

Un but essentiel de l'article présent est de réaliser avec un petit nombre d'opérations élémentaires les équivalences annoncées dans le théorème suivant. Cela peut être compris comme donnant une solution uniforme et en temps raisonnable pour les systèmes linéaires suffisamment « bien conditionnés », à l'image de ce que fait l'analyse numérique matricielle au moyen des décompositions en valeurs singulières (SVD) et des inverses de Moore-Penrose.

Ce théorème est pour l'essentiel dans [11] (lemme 1 page 8, exercice 7 page 49 et théorème 18 page 122) et dans [2] (voir aussi [7]). La preuve dans [2] est complètement explicite, contrairement à celle dans [11]. De manière un peu surprenante, [2] le fait remonter à ... 1994 !

Théorème 3.3 *Soit une application linéaire $\varphi : \mathbf{A}^n \to \mathbf{A}^m$ et A sa matrice sur les bases canoniques. Les propriétés suivantes sont équivalentes :*

1. $\operatorname{Im}\varphi$ *est facteur direct dans* $\mathbf{A}^m$.
2. $\operatorname{Coker}\varphi$ *est un module projectif de type fini.*
3. $\operatorname{Im}\varphi$ *est facteur direct dans* $\mathbf{A}^m$, $\operatorname{Ker}\varphi$ *est facteur direct dans* $\mathbf{A}^n$ *et si H est un supplémentaire de* $\operatorname{Ker}\varphi$, φ *réalise un isomorphisme de H sur* $\operatorname{Im}\varphi$.
4. *Il existe* $\varphi^\bullet : \mathbf{A}^m \to \mathbf{A}^n$ *telle que* $\mathbf{A}^n = \operatorname{Ker}\varphi \oplus \operatorname{Im}\varphi^\bullet$ *et* $\mathbf{A}^m = \operatorname{Ker}\varphi^\bullet \oplus \operatorname{Im}\varphi$.
5. *Il existe* $\psi : \mathbf{A}^m \to \mathbf{A}^n$ *vérifiant* $\varphi \circ \psi \circ \varphi = \varphi$.
6. *Il existe* $\psi : \mathbf{A}^m \to \mathbf{A}^n$ *vérifiant* $\varphi \circ \psi \circ \varphi = \varphi$ *et* $\psi \circ \varphi \circ \psi = \psi$.
7. *Chaque idéal déterminantiel* $\mathcal{D}_k(\varphi)$ *est idempotent.*
8. *Chaque idéal déterminantiel* $\mathcal{D}_k(\varphi)$ *est engendré par un idempotent* e_k. *Soit alors* $r_k = e_k - e_{k+1}$. *Les* r_k *forment un système fondamental d'idempotents orthogonaux. Pour tout mineur μ d'ordre k de A, sur le localisé* $\mathbf{A}[1/(r_k\,\mu)]$ *l'application linéaire φ devient simple de rang k.*
9. *L'application linéaire φ devient simple après localisation en des éléments x_i comaximaux.*
10. *L'application linéaire φ devient simple après localisation en n'importe quel idéal maximal.*

En particulier un module de présentation finie est projectif si et seulement si ses idéaux de Fitting sont idempotents. Le point 10 est à part, car il n'implique les autres qu'avec l'aide de l'axiome du choix. Les autres équivalences sont constructives.

Les équivalences 1 $\Leftrightarrow$ 2 $\Leftrightarrow$ 3 $\Leftrightarrow$ 4 sont naturelles.

Pour passer de 4 à 5 on constate que les restrictions de φ et $\varphi^\bullet$ aux sous-modules $\operatorname{Im}\varphi$ et $\operatorname{Im}\varphi^\bullet$ sont des isomorphismes. Nous donnerons un calcul « rapide » de ψ à partir de φ et $\varphi^\bullet$ dans la section 4.

Pour passer de 5 à 6 on remarque que si ψ_1 vérifie 5 alors $\psi = \psi_1 \circ \varphi \circ \psi_1$ vérifie 6.

Dans les conditions du 6, $\varphi \circ \psi$ est la projection sur $\operatorname{Im}\varphi$ parallèlement à $\operatorname{Ker}\psi$ et $\psi \circ \varphi$ est la projection sur $\operatorname{Im}\psi$ parallèlement à $\operatorname{Ker}\varphi$.

Le point 7 implique le point 8 de manière immédiate en tenant compte du lemme 1.7. Le point 8 implique trivialement le point 9 et celui-ci implique trivialement le point 10.

Pour montrer que 9 implique 6 on considère l'égalité $ABA = A$ (où A et B sont des matrices pour φ et ψ) comme une équation où l'inconnue est B : elle est facile à résoudre dans le cas où A définit une application linéaire simple, donc elle est résolue localement. Il reste à recoller les solutions en utilisant la combinaison linéaire des x_i égale à 1.

On peut aussi passer assez directement de 8 à 6 grâce à l'identité de Cramer (4). Si on a

$$\sum_{\alpha\in\mathcal{P}_{k,m},\beta\in\mathcal{P}_{k,n}} c_{\alpha,\beta}\,\mu_{\alpha,\beta} = e_k\,,$$

alors $\sum_{\alpha\in\mathcal{P}_{k,m},\beta\in\mathcal{P}_{k,n}} c_{\alpha,\beta}\,\mu_{\alpha,\beta}\,r_k = r_k$ et puisque la matrice est de rang $\leq k$ sur $\mathbf{A}[1/r_k]$ l'identité (4) fonctionne. Il suffit alors de poser :

$$B \;=\; \sum_k \left(\sum\nolimits_{\alpha\in\mathcal{P}_{k,m},\beta\in\mathcal{P}_{k,n}} r_k\,c_{\alpha,\beta}\,\mathrm{Adj}_{\alpha,\beta}(A)\right)\,. \tag{20}$$

Les matrices qui vérifient les propriétés du théorème 3.3 sont celles qui définissent les « meilleurs » systèmes linéaires : ceux pour lesquels on peut exprimer une solution du système linéaire comme une fonction linéaire du second membre. Ce sont aussi les systèmes pour lesquels on a une bonne description de l'image et du noyau, aussi bien du point de vue direct que du point de vue dual. Ces matrices, déjà intensivement étudiées dans [11] sont dites « regular » dans [2] et [12]. Cette terminologie remonte à Von Neuman, qui a étudié les anneaux (non nécessairement commutatifs) dans lesquels tous les éléments possèdent un « inverse généralisé » : il a utilisé pour cela le terme d'élément régulier. L'ennui est que pour une matrice carrée, « régulière » signifie inversible dans la terminologie courante.

Nous proposons la terminologie suivante, que nous étendrons au cas des modules projectifs de type fini.

Définition 3.4 *Une application linéaire φ entre modules libres de dimensions finies qui vérifie les conditions équivalentes du théorème 3.3 sera appelée une application linéaire* localement simple. *Si $\mathcal{D}_k(\varphi) = \langle 1\rangle$ et $\mathcal{D}_{k+1}(\varphi) = \langle 0\rangle$ on dira qu'il s'agit d'une* application linéaire (localement simple) de rang k. *Si on ne précise pas le rang mais qu'il existe, on dit qu'il s'agit d'une* application linéaire de rang constant. *Une* matrice localement simple (resp. de rang constant) *est la matrice d'une application linéaire localement simple (resp. de rang constant).*

Quand il existe, le rang d'une application linéaire localement simple est bien défini dès que l'anneau n'est pas trivial.

Sur un anneau sans autre idempotent que 1 et 0 toute matrice localement simple est de rang constant.

3.2 Rang d'un module projectif de type fini

Définition 3.5 *Soit un module projectif de type fini E engendré par n éléments, et ψ un endomorphisme de E.*

1. *Le déterminant de ψ, si $E\oplus N\simeq\mathbf{A}^n$, est défini par* $\det\psi := \det(\psi\oplus \mathrm{Id}_N)$.
2. *Nous notons*

$$\mathrm{P}_\psi(X) = \det(\mathrm{Id}_E + X\psi) = 1 + \mathrm{d}_1(\psi)X + \cdots + \mathrm{d}_n(\psi)X^n$$

   En particulier $\mathrm{d}_1(\psi) = \mathrm{Tr}(\psi)$ *est appelé la trace de ψ. On pose aussi* $\mathrm{d}_0(\psi) = 1$ *et pour* $p > n$, $\mathrm{d}_p(\psi) = 0$.
3. *Le polynôme caractéristique de ψ sera noté* $\mathrm{C}_\psi(X) = \det(X\mathrm{Id}_E - \psi)$.
4. *On note $\Gamma_\psi(X)$ le polynôme défini par* $\mathrm{C}_{-\psi}(-X) = -X\Gamma_\psi(X) + \det(\psi)$, *et $\Gamma_\psi(\psi)$ s'appelle* l'endomorphisme cotransposé de ψ. *Nous le notons $\widetilde{\psi}$ ou* $\mathrm{Adj}\,\psi$.

5. *Le déterminant de la multiplication par X sur le module E, noté $\mathrm{R}_E(X)$, est appelé le « polynôme multiplicatif du module ».*

Dans la définition précédente il est sous-entendu que le déterminant est « bien défini » : il ne dépend ni de l'entier n ni de la décomposition $E \oplus N \simeq \mathbf{A}^n$.

L'endomorphisme ψ est inversible si et seulement si $\det \psi$ est inversible. Le théorème de Cayley-Hamilton est valable pour les modules projectifs de type fini.

Si un anneau possède des idempotents un module projectif de type fini n'a pas forcément un rang bien défini. C'est le « polynôme multiplicatif » qui remplace le rang.

Théorème 3.6 *Soit E un module projectif de type fini isomorphe à l'image d'une matrice de projection $P \in \mathbf{A}^{m\times m}$ (de sorte que $\mathrm{I}_m - P$ est une matrice de présentation de E).*

1. *Le polynôme $\mathrm{R}_E(X) = \sum_i r_i X^i$ est égal à $\det(\mathrm{I}_m + (X-1)P)$. C'est un polynôme* multiplicatif *: il vérifie $\mathrm{R}_E(XY) = \mathrm{R}_E(X)\mathrm{R}_E(Y)$ et $\mathrm{R}_E(1) = 1$. Cela signifie que les r_i forment un système fondamental d'idempotents orthogonaux (sfio). On a $r_0 = \det(\mathrm{I}_m - P)$ et l'idéal $\langle r_0\rangle$ est l'annulateur de E.*
2. *Le module est dit de rang k si $\mathrm{R}_E(X) = X^k$. Il est dit de rang $\leq k$ si $r_{k+1} = \ldots = r_m = 0$ c'est-à-dire encore si tous les mineurs d'ordre $k+1$ de la matrice P sont nuls. S'il a un rang k le module E est dit de rang constant. Quand il existe, le rang d'un module projectif de type fini est bien défini dès que l'anneau n'est pas trivial.*
3. *Le localisé $\mathbf{A}_{r_k} = \mathbf{A}[1/r_k]$ est isomorphe à $\mathbf{A}/\langle 1 - r_k\rangle$. Le localisé E_{r_k} est isomorphe au sous-module $r_k E$. Il est de rang k en tant que $\mathbf{A}_{r_k}$-module. Le module E est somme directe des « composantes » $r_k E$ ($k > 0$).*
4. *Les idéaux de Fitting $\mathcal{F}_k(E)$ sont liés aux idempotents r_k définis via le polynôme multiplicatif R_E par la relation : $\mathcal{F}_k(E) = \left\langle \sum_{\ell\leq k} r_\ell \right\rangle$.*
5. *Les $\binom{m}{k}$ mineurs diagonaux d'ordre k de P sont les $\mu_{\alpha,\alpha} = \mu_{\alpha,\alpha}(P)$ pour $\alpha \in \mathcal{P}_{k,m}$ et on a $\mathrm{d}_k(P) = \sum_{\alpha\in\mathcal{P}_{k,m}} \mu_{\alpha,\alpha}$. Alors $r_k\,\mathrm{d}_k(P) = r_k$ et, pour chaque $\alpha \in \mathcal{P}_{k,m}$, le module E devient libre de rang k lorsqu'on localise en $r_k\,\mu_{\alpha,\alpha}$.*

Des résultats utiles qui améliorent légèrement la proposition 1.15 sont donnés dans la proposition suivante et son corollaire (pour une preuve constructive voir [7]).

Proposition 3.7 *Soit $\varphi : E \to F$ une application linéaire* surjective *entre modules projectifs de type fini de même polynôme multiplicatif $\mathrm{R}_E = \mathrm{R}_F$, alors φ est un isomorphisme. Ceci s'applique en particulier s'ils ont même rang constant.*

Corollaire 3.8 *Soit F un module projectif de type fini. Si $F_1 \oplus G_1 = F = F_2 \oplus G_2$ avec $F_1 \subset F_2$ on a :*

$$\mathrm{R}_{F_1} = \mathrm{R}_{F_2} \iff \mathrm{R}_{G_1} = \mathrm{R}_{G_2} \iff F_1 = F_2$$

3.3 Quand les modules de rang constant sont libres

Pour un anneau $\mathbf{A}$ il revient au même de dire que toutes les matrices localement simples de rang constant sont simples, ou que tous les modules projectifs de rang constant sont libres.

Signalons quelques cas importants où ceci se produit.

- $\mathbf{A}$ est un anneau local.
- $\mathbf{A}$ est zéro-dimensionnel (i.e., $\forall x \in \mathbf{A}$, $\exists y \in \mathbf{A}$, $\exists n \in \mathbb{N}$ $x^{n+1} = yx^n$).
- Le quotient de $\mathbf{A}$ par son radical de Jacobson $\mathrm{Rad}(\mathbf{A})$ est zéro-dimensionnel (rappelons que si $\mathcal{U}_\mathbf{A}$ désigne le groupe des unités de $\mathbf{A}$, $\mathrm{Rad}(\mathbf{A}) = \{\, x \in \mathbf{A} \mid 1 + x\mathbf{A} \subset \mathcal{U}_\mathbf{A} \,\}$).

– **A** est « fortement U-irréductible », i.e., pour tout polynôme $P = \sum_{i=0}^{n} a_i x^i \in \mathbf{A}[X]$ primitif (i.e., tel que $\langle a_0, \ldots, a_n \rangle = \langle 1 \rangle$), il existe $x \in \mathbf{A}$ tel que $\langle P(x) \rangle = \langle 1 \rangle$.

– $\mathbf{A} = \mathbf{B}[X_1, \ldots, X_n]$ où **B** est un anneau de Bezout, i.e., tout idéal de type fini de **B** est principal.

Les trois premiers cas sont traités constructivement dans [7].

Il semble qu'on ne connaisse pas pour le moment de preuve constructive pour le dernier cas, qui est une extension remarquable du théorème de Quillen-Suslin, due à Lequain et Simis [8]. Une telle preuve fournirait un algorithme pour transformer une matrice A localement simple de rang constant r en une matrice $\mathrm{I}_{r,m,n} = Q_1 A Q_2$ avec Q_1 et Q_2 inversibles.

Le quatrième cas est assez facile. Rappelons comment cela fonctionne. On part d'une matrice localement simple $A = (a_{ij}) \in \mathbf{A}^{m \times n}$ de rang $r \geq 1$. On va la diagonaliser par des changements de base de la source et du but. Imaginons que nous multiplions chaque ligne n°k par X^{k-1} puis chaque colonne n°ℓ par $X^{(\ell-1)m}$, et considérons le polynôme $P(X) = \sum_{k\ell} a_{k\ell} X^{(\ell-1)m+k-1}$. Par hypothèse ce polynôme est primitif. Soit $x \in \mathbf{A}$ tel que $P(x)$ est inversible. Dans la matrice A on ajoute à la première ligne L_1 les lignes $x^{k-1} L_k$ ($k > 1$). Puis dans la matrice obtenue, on ajoute à la première colonne C_1 les colonnes $x^{(\ell-1)m} C_\ell$ ($\ell > 1$). Alors en position $(1,1)$ on trouve $P(x)$ qui peut servir de pivot de Gauss. On termine par induction.

Notez que l'algorithme esquissé ci-dessus utilise un nombre raisonnable d'opérations élémentaires si le caractère fortement U-irréductible de l'anneau est rendu explicite au moyen d'un nombre raisonnable d'opérations élémentaires (nous ne cherchons pas à formaliser la chose).

Pour tout anneau **A** il y a une extension fidèlement plate et fortement U-irréductible de **A** qui est le localisé « de Nagata » $\mathbf{A}(X) = S^{-1}\mathbf{A}[X]$ où $S \subset \mathbf{A}[X]$ est le monoïde formé par les polynômes primitifs. En effet, si $P(T) = \sum_i Q_i(X) T^i$ est tel que $\sum_i B_i(X) Q_i(X)$ soit un polynôme primitif alors pour $k > \sup_i(\deg_X(Q_i))$, $P(X^k)$ est lui-même un polynôme primitif.

Il n'est pas étonnant que cet anneau $\mathbf{A}(X)$ joue un role crucial dans la suite pour nos calculs uniformes (en temps raisonnable).

3.4 Applications linéaires localement simples entre modules projectifs de type fini

Définition 3.9 *Soient E et F deux $\mathbf{A}$–modules projectifs de type fini, et une application linéaire $\varphi : E \to F$.*

1. *Les* idéaux déterminantiels de l'application linéaire φ *sont les idéaux*
$$\mathcal{D}_k(\varphi) \;:=\; \textit{l'idéal engendré par les } \det(\lambda \circ \varphi \circ \theta)$$
où k est un entier arbitraire, $\lambda : F \to \mathbf{A}^k$ et $\theta : \mathbf{A}^k \to E$ sont arbitraires.
2. *L'application linéaire φ est dite de rang $\leq k$ si $\mathcal{D}_{k+1}(\varphi) = 0$.*
3. *L'application linéaire φ est dite localement simple si* $\operatorname{Im}\varphi$ *est facteur direct dans F. Elle est dite localement simple de rang k si en outre le module projectif* $\operatorname{Im}\varphi$ *est de rang k.*

Le calcul de $\mathcal{D}_k(\varphi)$ se fait comme suit. Supposons que $E \oplus E_1 \simeq \mathbf{A}^n$, $F \oplus F_1 \simeq \mathbf{A}^m$, que $\Pi_E : \mathbf{A}^n \to E$ est la projection sur E parallèlement à E_1, $\iota_F : F \to \mathbf{A}^m$ est l'injection naturelle et que $\varphi' = \iota_F \circ \varphi \circ \Pi_E$. Alors $\mathcal{D}_k(\varphi) = \mathcal{D}_k(\varphi')$. Si $P \in \mathbf{A}^{n \times n}$ est la matrice de la projection sur E parallèlement à E_1 (c'est-à-dire de $\pi_E = \iota_E \circ \Pi_E$) et $Q \in \mathbf{A}^{m \times m}$ est la matrice de la projection sur F parallèlement à F_1 alors, la matrice A de φ' vérifie $QAP = A$ (et cette égalité caractérise les matrices du type φ'). On dira que la matrice A *représente* φ (via les isomorphismes $E \oplus E_1 \simeq \mathbf{A}^n$ et $F \oplus F_1 \simeq \mathbf{A}^m$).

De manière générale tout calcul sur les modules projectifs de type fini se ramène à un calcul sur des matrices.

Une application linéaire φ entre modules projectifs de type fini est localement simple de rang k si et seulement si $\mathcal{D}_k(\varphi) = \langle 1\rangle$ et $\mathcal{D}_{k+1}(\varphi) = 0$. Elle est localement simple si et seulement si tous ses idéaux déterminantiels sont engendrés par des idempotents. Plus généralement on peut recopier le théorème 3.3.

Théorème 3.10 *Les propriétés suivantes pour une application linéaire $\varphi : E \to F$ entre modules projectifs de type fini sont équivalentes.*

1. $\operatorname{Im}\varphi$ *est facteur direct dans* F.
2. $\operatorname{Coker}\varphi$ *est un module projectif de type fini.*
3. $\operatorname{Im}\varphi$ *est facteur direct dans* F, $\operatorname{Ker}\varphi$ *est facteur direct dans* E *et si* H *est un supplémentaire de* $\operatorname{Ker}\varphi$, φ *réalise un isomorphisme de* H *sur* $\operatorname{Im}\varphi$.
4. *Il existe* $\varphi^\bullet : F \to E$ *telle que* $E = \operatorname{Ker}\varphi \oplus \operatorname{Im}\varphi^\bullet$ *et* $F = \operatorname{Ker}\varphi^\bullet \oplus \operatorname{Im}\varphi$.
5. *Il existe* $\psi : F \to E$ *vérifiant* $\varphi\circ\psi\circ\varphi = \varphi$.
6. *Il existe* $\psi : F \to E$ *vérifiant* $\varphi\circ\psi\circ\varphi = \varphi$ *et* $\psi\circ\varphi\circ\psi = \psi$.
7. *Chaque idéal déterminantiel* $\mathcal{D}_k(\varphi)$ *est idempotent.*
8. *Chaque idéal déterminantiel* $\mathcal{D}_k(\varphi)$ *est engendré par un idempotent* e_k. *Soit alors* $r_k = e_k - e_{k+1}$. *Les* r_k *forment un système fondamental d'idempotents orthogonaux. Pour tout mineur* μ *d'ordre* k *d'une matrice* A *qui représnte* φ, *sur le localisé* $\mathbf{A}[1/(r_k\,\mu)]$ *l'application linéaire* φ *devient simple de rang* k.
9. *L'application linéaire* φ *devient simple après localisation en des éléments* x_i *comaximaux.*
10. *L'application linéaire* φ *devient simple après localisation en n'importe quel idéal maximal.*

4 Applications linéaires croisées et inverses généralisés pour les modules projectifs de type fini

Dans toute la section 4, E et F sont des modules projectifs de type fini.

4.1 Un critère pour les applications linéaires croisées avec elles-mêmes

Voici une généralisation d'un résultat usuel pour les espaces vectoriels de dimension finie. Il s'agit ici d'une conséquence importante du théorème de Cayley-Hamilton.

Théorème 4.1 *Soit $\varphi : E \to E$ un endomorphisme. Notons $d_j = \mathrm{d}_j(\varphi)$. Les propriétés suivantes sont équivalentes :*

1. φ *est de rang* $\leq k$ *et* d_k *est inversible.*
2. φ *est croisée avec elle-même et de rang* k.

Lorsque ces conditions sont vérifiées, la projection $\pi : E \to E$ sur $\operatorname{Im}\varphi$ parallèlement à $\operatorname{Ker}\varphi$ vérifie :

$$d_k\,\pi = d_{k-1}\,\varphi - d_{k-2}\,\varphi^2 + \cdots + (-1)^{k-1}\varphi^k\,. \tag{21}$$

En outre l'inverse généralisé $\psi = \operatorname{Ig}(\varphi,\varphi)$ vérifie

$$d_k\,\psi = d_{k-1}\,\pi - d_{k-2}\,\varphi + d_{k-3}\,\varphi^2 + \cdots + (-1)^{k-1}\varphi^{k-1}\,. \tag{22}$$

Preuve Le point délicat est : 1 implique 2.
On a $d_k \in \mathcal{D}_k(\varphi)$ donc $\mathcal{D}_k(\varphi) = \langle 1 \rangle$, et $\mathcal{D}_{k+1}(\varphi) = 0$ par hypothèse. Donc φ est localement simple de rang k. Soit K un supplémentaire de $\operatorname{Im}\varphi$ dans E. Sur cette somme directe φ est « triangulaire » avec une « matrice » du type :

$$\begin{bmatrix} \varphi_0 & \varphi' \\ 0_{\operatorname{Im}\varphi,K} & 0_{K,K} \end{bmatrix}$$

où $\varphi_0 : \operatorname{Im}\varphi \to \operatorname{Im}\varphi$ est la restriction de φ. Donc $\det(\mathrm{Id}_E + X\varphi) = \det(\mathrm{Id}_{\operatorname{Im}\varphi} + X\varphi_0)$. On obtient $\det\varphi_0 = d_k$ et donc φ_0 est inversible. Ceci implique tout d'abord $\operatorname{Im}\varphi \cap \operatorname{Ker}\varphi = 0$. Ensuite tout $x \in E$ s'écrit $x_1 + x_2$ où $x_1 = \varphi_0^{-1}(\varphi(x)) \in \operatorname{Im}\varphi$ et $x_2 = x - x_1 \in \operatorname{Ker}\varphi$. Donc $E = \operatorname{Im}\varphi \oplus \operatorname{Ker}\varphi$. On peut donc remplacer K par $\operatorname{Ker}\varphi$ et la « matrice » ci-dessus devient « diagonale » ($\varphi' = 0_{\operatorname{Ker}\varphi,\operatorname{Im}\varphi}$). Le théorème de Cayley-Hamilton appliqué à φ_0 donne

$$d_k \,\mathrm{Id}_{\operatorname{Im}\varphi} = \varphi_0 \left(d_{k-1} - d_{k-2}\,\varphi_0 + d_{k-3}\,\varphi_0^2 + \cdots + (-1)^{k-1}\varphi_0^{k-1} \right)$$

ce qui implique facilement les égalités voulues. □

4.2 Applications linéaires croisées entre modules projectifs de type fini

Dans toute la suite de la section 4 on considère deux applications linéaires $\varphi : E \to F$ et $\varphi^\bullet : F \to E$.

Supposons tout d'abord φ et $\varphi^\bullet$ croisées. Nous reprenons les notations de la section 1.4.

Soit $\varphi_1 : \operatorname{Im}\varphi \to \operatorname{Im}\varphi$ l'automorphisme linéaire défini par $\varphi_1 = \varphi_0\,\varphi_0^\bullet$. C'est la restriction de $\varphi\,\varphi^\bullet$ à $\operatorname{Im}\varphi$. Définissons de même $\varphi_1^\bullet : \operatorname{Im}\varphi^\bullet \to \operatorname{Im}\varphi^\bullet$ par $\varphi_1^\bullet = \varphi_0^\bullet\,\varphi_0$. Si le rang de φ est $\leq k$ on obtient alors :

$$\mathrm{P}_{\varphi_1^\bullet}(Z) \;=\; \mathrm{P}_{\varphi^\bullet\varphi}(Z) \;=\; \mathrm{P}_{\varphi\varphi^\bullet}(Z) \;=\; \mathrm{P}_{\varphi_1}(Z) \;=\; 1 + a_1 Z + \cdots + a_k Z^k$$

où $a_j = \mathrm{d}_j(\varphi^\bullet\varphi)$. Si φ est de rang constant k alors il en va de même pour φ_0, φ_1, $\varphi^\bullet$, $\varphi\,\varphi^\bullet$ etc.... De sorte que a_k est un élément inversible de $\mathbf{A}$.

On a la réciproque suivante importante, qui est un analogue du théorème 4.1.

Théorème 4.2 *Notons $a_j = \mathrm{d}_j(\varphi\varphi^\bullet)$. Les propriétés suivantes sont équivalentes :*

1. *φ et $\varphi^\bullet$ sont croisées de rang k.*
2. *φ et $\varphi^\bullet$ sont de rang $\leq k$ et a_k est inversible.*

Preuve Il faut montrer la réciproque. On a $a_k \in \mathcal{D}_k(\varphi\varphi^\bullet) \subset \mathcal{D}_k(\varphi)\,\mathcal{D}_k(\varphi^\bullet)$. Puisque a_k est inversible $\mathcal{D}_k(\varphi) = \mathcal{D}_k(\varphi^\bullet) = \langle 1 \rangle$ donc φ et $\varphi^\bullet$ sont localement simples de rang k.
Le théorème 4.1 montre en outre que $\varphi\varphi^\bullet : F \to F$ est croisée avec elle même, de rang k. Même chose pour $\varphi^\bullet\varphi : E \to E$.
On a donc la situation suivante : $F_1 = \operatorname{Im}\varphi\,\varphi^\bullet \subset F_2 = \operatorname{Im}\varphi \subset F$ sont 2 modules de rang k en facteur direct dans F. On peut donc appliquer le corollaire 3.8 : $F_1 = F_2$. Ainsi $\operatorname{Im}\varphi\,\varphi^\bullet = \operatorname{Im}\varphi$ et symétriquement $\operatorname{Im}\varphi^\bullet\,\varphi = \operatorname{Im}\varphi^\bullet$.
De la même façon on a $K_1 = \operatorname{Ker}\varphi^\bullet \subset K_2 = \operatorname{Ker}\varphi\,\varphi^\bullet \subset E$. Ils sont tous deux en facteur direct avec un supplémentaire de rang k. On peut donc appliquer le corollaire 3.8 : $K_1 = K_2$.
Finalement on obtient $E = \operatorname{Im}\varphi^\bullet \oplus \operatorname{Ker}\varphi$ et $F = \operatorname{Im}\varphi \oplus \operatorname{Ker}\varphi^\bullet$. □

4.3 Calcul théorique d'un inverse généralisé : le cas du rang constant

En utilisant le théorème de Cayley-Hamilton on démontre comme pour le théorème 4.1 le résultat suivant.

Théorème 4.3 (projections sur l'image et sur le noyau et inverse généralisé en rang constant k)
Si φ et $\varphi^\bullet$ sont croisées de rang k, avec $a_j = \mathrm{d}_j(\varphi\varphi^\bullet)$, on a :

1. *L'inverse généralisé de φ via $\varphi^\bullet$ est donné par*

$$\psi = \mathrm{Ig}(\varphi,\varphi^\bullet) = \varphi^{\dagger_{\varphi^\bullet}} = a_k^{-1}\left(a_{k-1}\varphi^\bullet - a_{k-2}\varphi^\bullet\varphi\varphi^\bullet + \cdots + (-1)^{k-1}(\varphi^\bullet\varphi)^{k-1}\varphi^\bullet\right). \quad (23)$$

2. *La projection sur le sous-espace $I = \mathrm{Im}\,\varphi \subseteq F$ parallèlement à $\mathrm{Ker}\,\varphi^\bullet$ est égale à $\pi_I = \varphi\,\psi$.*
3. *La projection sur le sous-espace $I^\bullet = \mathrm{Im}\,\varphi^\bullet \subseteq E$ parallèlement à $\mathrm{Ker}\,\varphi$ est égale à $\pi_{I^\bullet} = \psi\,\varphi$. Et la projection sur le noyau de φ parallèlement à $\mathrm{Im}\,\varphi^\bullet$ est $\mathrm{Id}_E - \pi_{I^\bullet}$.*

On obtient aussi l'équivalence générale suivante :

Théorème 4.4 *Notons $a_j = \mathrm{d}_j(\varphi\varphi^\bullet)$. Les propriétés suivantes sont équivalentes :*

1. *φ et $\varphi^\bullet$ sont croisées de rang k.*
2. *$a_h = 0$ pour $h > k$, a_k est inversible et, en définissant θ par*

$$\theta = a_{k-1}\,\varphi^\bullet - a_{k-2}\,\varphi^\bullet\varphi\varphi^\bullet + \cdots + (-1)^{k-1}\varphi^\bullet(\varphi\varphi^\bullet)^{k-1},$$

on a les deux égalités $\varphi\,\theta\,\varphi = a_k\,\varphi$ et $\varphi^\bullet\,\varphi\,\theta = a_k\,\varphi^\bullet$.

Preuve Il reste à montrer que 2 implique 1. Posons

$$\psi = a_k^{-1}\theta \quad \text{et} \quad \psi^\bullet = a_k^{-1}\left(a_{k-1}\varphi - a_{k-2}\varphi\varphi^\bullet\varphi + \cdots + (-1)^{k-1}(\varphi\varphi^\bullet)^{k-1}\varphi\right).$$

Un calcul simple montre que les six égalités de la proposition 1.13 sont satisfaites. □

On peut utiliser le test précédent pour savoir si une application linéaire est croisée avec elle-même et de rang k. Une légère variante, sans doute plus efficace du point de vue du calcul est obtenue de façon analogue en s'appuyant sur le théorème 4.1 et la proposition 1.14 :

Théorème 4.5 *On suppose $F = E$. Notons $d_j = \mathrm{d}_j(\varphi)$. Les propriétés suivantes sont équivalentes :*

1. *φ est croisée avec elle-même, de rang k.*
2. *$d_h = 0$ pour $h > k$, d_k est inversible et, en définissant π par*

$$\pi = d_{k-1}\,\varphi - d_{k-2}\,\varphi^2 + \cdots + (-1)^{k-1}\varphi^k,$$

on a les égalités $\pi\,\varphi = d_k\,\varphi$ et $\pi^2 = d_k\,\pi$.

Le théorème 4.4 conduit au résultat de complexité suivant, lorsque les modules E et F sont donnés par des matrices de projection P_E et P_F dont ils sont les images, et φ et $\varphi^\bullet$ sont données par des matrices A et $A^\bullet$ vérifiant $P_F A P_E = A$ et $P_E A^\bullet P_F = A^\bullet$.

Théorème 4.6 *Soient E et F deux $\mathbf{A}$–modules projectifs de type fini engendrés par n éléments (ou moins), et deux applications linéaires $\varphi : E \to F$ et $\varphi^\bullet : F \to E$. Alors on peut, avec $\mathcal{O}(n^4)$ opérations arithmétiques, un test « $1 \in \langle y\rangle$? » et $\mathcal{O}(n^2)$ tests « $x = 0$? », décider si φ et $\varphi^\bullet$ sont croisées et de rang k, et en cas de réponse positive calculer les inverses généralisés $\mathrm{Ig}(\varphi,\varphi^\bullet)$ et $\mathrm{Ig}(\varphi^\bullet,\varphi)$ en utilisant $\mathcal{O}(n^4)$ opérations arithmétiques.*

4.4 Cas où le rang n'est pas constant

Les résultats de la section 4.3 se généralisent en cassant l'anneau en des composantes convenables données par un sfio.

L'idée générale est la suivante : si $\mathrm{R}_{\mathrm{Im}\,\varphi}(X) = \sum_i r_i X^i$ et $\mathrm{P}_{\varphi\varphi^\bullet}(Z) = 1 + a_1 Z + \cdots + a_n Z^n = P(Z)$, chaque polynôme $P_k = r_k P$ doit être de degré k avec pour coefficient de Z^k un élément inversible dans $\mathbf{A}[1/r_k] \simeq r_k\mathbf{A}$ et ceci permet de calculer les r_i lorsqu'on connaît P. Plus précisément

- On peut retrouver les r_i à partir des a_i. Par exemple on doit avoir $\langle r_n\rangle = \langle a_n\rangle = \langle a_n^2\rangle$: on teste si $\langle a_n\rangle = \langle a_n^2\rangle$, en cas de réponse positive avec $a_n = b_n a_n^2$, alors $r_n = a_n b_n$, puis on recommence avec $(1 - r_n)P$ pour trouver r_{n-1} et ainsi de suite.
- On pose $\varphi_k = r_k\,\varphi$ et $\varphi_k^\bullet = r_k\,\varphi^\bullet$. Par le théorème 4.6 on peut tester si φ_k et $\varphi_k^\bullet$ sont croisées et de rang k sur l'anneau $\mathbf{A}[1/r_k]$, et en cas de réponse positive calculer l'inverse généralisé $\mathrm{Ig}(\varphi_k, \varphi_k^\bullet)$.
- On termine en recollant tout ceci : $\mathrm{Ig}(\varphi, \varphi^\bullet) = \sum_{k=1}^n r_k\,\mathrm{Ig}(\varphi_k, \varphi_k^\bullet)$.

Le calcul le plus long dans toute cette affaire est celui du polynôme P qui se fait en $\mathcal{O}(n^4)$ opérations arithmétiques dans $\mathbf{A}$.

La procédure entière est explicite si on dispose d'un test de divisibilité dans $\mathbf{A}$, c.-à-d. un test pour « $x \in \langle y\rangle$? » qui donne un z tel que $x = zy$ en cas de réponse positive. En faisant « $x - y \in \langle 0\rangle$? » on a aussi un test pour l'égalité dans $\mathbf{A}$.

On résume la situation dans le théorème suivant, en supposant que les modules E et F sont donnés par des matrices de projection P_E et P_F dont ils sont les images, et que φ et $\varphi^\bullet$ sont données par des matrices A et $A^\bullet$ vérifiant $P_F A P_E = A$ et $P_E A^\bullet P_F = A^\bullet$.

Théorème 4.7 *Soient E et F deux $\mathbf{A}$–modules projectifs de type fini engendrés par n éléments (ou moins), et deux applications linéaires $\varphi : E \to F$ et $\varphi^\bullet : F \to E$. Alors on peut, avec un nombre d'opérations arithmétiques en $\mathcal{O}(n^4)$, et un nombre de tests « $x \in \langle y\rangle$? » en $\mathcal{O}(n^3)$, décider si φ et $\varphi^\bullet$ sont croisées, et en cas de réponse positive calculer les inverses généralisés $\mathrm{Ig}(\varphi, \varphi^\bullet)$ et $\mathrm{Ig}(\varphi^\bullet, \varphi)$ en $\mathcal{O}(n^4)$ opérations arithmétiques.*

5 Calcul pratique d'un inverse généralisé s'il en existe un.

Dans cette section nous généralisons au cas d'un anneau commutatif $\mathbf{A}$ le travail que nous avons fait dans [4] en vue de la résolution uniforme des systèmes linéaires sur un corps arbitraire, en nous appuyant sur un calcul uniforme du rang d'une matrice dû à Mulmuley [10]. Il s'agit de la possibilité de calculer efficacement un inverse généralisé d'une application linéaire entre $\mathbf{A}$–modules libres lorsqu'il en existe un. En termes plus abstraits : lorsqu'on connaît une matrice de présentation pour un module E, on est capable de tester si ce module est projectif de type fini et en cas de réponse positive, de fournir une matrice de projection dont l'image est isomorphe à E (l'isomorphisme est explicite). Tout ceci avec des calculs assez efficaces, c'est-à-dire ici *en temps polynomial*.

Considérons une application linéaire φ entre deux $\mathbf{A}$–modules libres E et F de dimensions respectives n et m. Notre but est de donner un test pour savoir si φ est localement simple, et, en cas de réponse positive, de calculer en temps polynomial un inverse généralisé de la matrice.

Bien que nous traitions uniquement le cas des modules libres, il ne serait pas difficile de généraliser au cas où E et F sont des modules projectifs de type fini.

Nous nous limitons au point de vue purement matriciel, (c'est le point de vue où des bases ont été fixées dans E et F). Nous introduisons une indéterminée t. Nous considérons une forme quadratique $\Phi_{t,n}$ sur $E' = \mathbf{A}(t)^n$ et une forme quadratique $\Phi_{t,m}$ sur $F' = \mathbf{A}(t)^m$:

$$\begin{array}{rcl} \Phi_{t,n}(\xi_1,\ldots,\xi_n) & = & {\xi_1}^2 + t\,{\xi_2}^2 + \cdots + t^{n-1}\,{\xi_n}^2 \\ \Phi_{t,m}(\zeta_1,\ldots,\zeta_m) & = & {\zeta_1}^2 + t\,{\zeta_2}^2 + \cdots + t^{m-1}\,{\zeta_m}^2 \end{array}$$

Nous notons les « produits scalaires » correspondants par $\langle\cdot,\cdot\rangle^t_{E'}$ et $\langle\cdot,\cdot\rangle^t_{F'}$. Nous notons Q_n et Q_m les matrices (diagonales) de ces formes sur les bases canoniques.

L'application linéaire $\varphi : E \to F$ donne lieu à une application linéaire $E' \to F'$ que nous notons encore φ et qui est définie par la même matrice sur les bases canoniques. Il existe alors une unique application linéaire $\varphi^\circ : F' \to E'$ vérifiant :

$$\forall x \in E' \quad \forall y \in F' \qquad \langle \varphi(x), y\rangle^t_{F'} = \langle x, \varphi^\circ(y)\rangle^t_{E'} \tag{24}$$

La matrice A° de φ° sur les bases canoniques est alors

$$A^\circ \;=\; {Q_n}^{-1}\;{}^{\mathrm{t}}\!A\;Q_m\,, \tag{25}$$

puisqu'on doit avoir pour tous $X \in \mathbf{A}(t)^{n\times 1}$, $Y \in \mathbf{A}(t)^{m\times 1}$: ${}^{\mathrm{t}}(A\,X)\,Q_m\,Y \;=\; {}^{\mathrm{t}}\!X\,Q_n\,(A^\circ\,Y)$.

On vérifie que $(AB)^\circ = B^\circ\,A^\circ$ et $(A^\circ)^\circ = A$.

En pratique si $A = (a_{i,j})$ on obtient $A^\circ = (t^{j-i}\,a_{j,i})$, par exemple :

$$A \;=\; \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{bmatrix}, \quad A^\circ \;=\; \begin{bmatrix} a_{11} & t\,a_{21} & t^2\,a_{31} \\ t^{-1}\,a_{12} & a_{22} & t\,a_{32} \\ t^{-2}\,a_{13} & t^{-1}\,a_{23} & a_{33} \\ t^{-3}\,a_{14} & t^{-2}\,a_{24} & t^{-1}\,a_{34} \\ t^{-4}a_{15} & t^{-3}a_{25} & t^{-2}a_{35} \end{bmatrix}$$

5.1 Idéaux de Gram et idéaux déterminantiels

Définition 5.1 *Soit une matrice $A \in \mathbf{A}^{m\times n}$. On définit les polynômes de Laurent $\mathcal{G}'_k(A)(t) = g_k(t) \in \mathbf{A}[t,1/t]$, et les* coefficients de Gram généralisés de A, *$\mathcal{G}'_{k,\ell}(A) = g_{k,\ell}$ comme suit :*

$$\left\{ \begin{array}{rcl} \mathrm{P}_{AA^\circ}(Z) & = & 1 + g_1(t)\,Z + \cdots + g_m(t)\,Z^m \\ g_k(t) & = & t^{-k(n-k)}\left(\sum_{\ell=0}^{k(m+n-2k)} g_{k,\ell}\,t^\ell\right) \end{array} \right. \tag{26}$$

Autrement dit $g_k(t) = \mathrm{d}_k(AA^\circ)$. Nous définissons aussi $\mathcal{G}'_0(A) = 1$ et $\mathcal{G}'_\ell(A) = 0$ pour $\ell > m$. Les idéaux de Gram de la matrice A *sont les idéaux $\mathcal{C}_k(A)$ définis par :*

$$\mathcal{C}_k(A) \;:=\; \textit{l'idéal engendré par les } g_{h,\ell} \textit{ pour tous les } h \geq k.$$

Proposition 5.2 *On a $\sqrt{\mathcal{C}_k(A)} = \sqrt{\mathcal{D}_k(A)}$. Plus précisément, avec un entier r qui ne dépend que de (m,n,k) on a :*

$$\mathcal{D}_k(A)^r \subset \mathcal{C}_k(A) \subset \mathcal{D}_k(A)^2 \subset \mathcal{D}_k(A).$$

Preuve Les inclusions $\mathcal{C}_k(A) \subset \mathcal{D}_k(A)^2 \subset \mathcal{D}_k(A)$ sont claires. Dans le cas des corps on sait que $\mathcal{C}_k(A) = \mathcal{D}_k(A)$ (cette égalité est essentiellement une reformulation du résultat de Mulmuley, cf. [4, 10]). On peut donc conclure par le Nullstellensatz formel qu'il existe un entier $r > 0$ tel que $\mathcal{D}_k(A)^r \subset \mathcal{C}_k(A)$. Avoir un tel r explicitement demande un peu plus de travail. Nous nous en dispenserons car nous n'en aurons pas besoin pour nos calculs « en temps polynomial ». □

Notez que les idéaux déterminantiels de A° sont égaux à ceux de A.

Corollaire 5.3 *Si A est localement simple, les idéaux de Gram de A sont égaux à ses idéaux déterminantiels et sont engendrés par des idempotents.*

Corollaire 5.4 *Réciproquement :*

1. *Si $\mathcal{D}_{k+1}(A) = 0$ et $\mathcal{C}_k(A) = \langle 1\rangle$, A est localement simple de rang k.*
2. *Si $\mathbf{A}$ est réduit (i.e. 0 est le seul élément nilpotent), si $\mathcal{C}_{k+1}(A) = 0$ et $\mathcal{C}_k(A) = \langle 1\rangle$, A est localement simple de rang k.*
3. *Si $\mathbf{A}$ est réduit et si les idéaux de Gram de A sont engendrés par des idempotents, alors A est localement simple.*
4. *Si $\mathbf{A}$ est réduit et a pour seuls idempotents 0 et 1 (en particulier si $\mathbf{A}$ est intègre) la matrice A est localement simple si et seulement si il existe un entier k vérifiant : $\mathcal{C}_{k+1}(A) = 0$ et $\mathcal{C}_k(A) = \langle 1\rangle$.*

Notez que la condition « $\mathcal{C}_{k+1}(A) = 0$ et $\mathcal{C}_k(A) = \langle 1\rangle$ » revient à dire que $\mathrm{P}_{AA^\circ}(Z)$ est de degré $\leq k$ en Z et que son coefficient $g_k(t)$ est inversible dans $\mathbf{A}(t)$.

5.2 Calcul pratique d'un inverse généralisé

Rappelons que notre but est de donner un test rapide pour savoir si une matrice est localement simple, et, en cas de réponse positive, de calculer un inverse généralisé de la matrice.

Le cas du rang constant

C'est par exemple sûrement le cas si l'anneau n'a pas d'autre idempotent que 0 et 1.

Théorème 5.5 *Soit une matrice $A \in \mathbf{A}^{m\times n}$. On rappelle que $A^\circ = Q_n^{-1}(t)\, A\, Q_m(t)$ et que $\mathrm{P}_{AA^\circ}(Z) = \det(\mathrm{I}_n + ZAA^\circ) = 1 + \sum_{1\leq \ell\leq n} g_\ell(t)\, Z^\ell$. Les propriétés suivantes sont équivalentes :*

1. *A est localement simple de rang k sur $\mathbf{A}$.*
2. *A est localement simple de rang k sur $\mathbf{A}(t)$.*
3. *A est simple de rang k sur $\mathbf{A}(t)$.*
4. *A est de rang $\leq k$ et le polynôme $t^{k(n-k)}g_k(t)$ est primitif.*
5. *A et A° sont croisées sur $\mathbf{A}(t)$, de rang k.*
6. *A et A° sont croisées sur $\mathbf{A}(t)$, $\deg_Z(\mathrm{P}_{AA^\circ}) \leq k$ et $t^{k(n-k)}g_k(t)$ est primitif.*
7. *$\deg_Z(\mathrm{P}_{AA^\circ}) \leq k$, le polynôme $t^{k(n-k)}g_k(t)$ est primitif et on a $A\,\mathrm{Adj}^{(k)}_{A^\circ}(A)\,A = g_k(t)\,A$.*

Si $\mathbf{A}$ est un anneau réduit, la condition 7 se simplifie en « $\deg_Z(\mathrm{P}_{AA^\circ}) \leq k$ et le polynôme $t^{k(n-k)}g_k(t)$ est primitif ». Lorsque les conditions sont vérifiées la matrice $\mathrm{Adj}^{(k)}_{A^\circ}(A)/g_k(t)$ est l'inverse généralisé de A via A° sur l'anneau $\mathbf{A}(t)$.

Preuve Le fait que 2 implique 3 a été expliqué dans la section 3.3.
De la caractérisation de 1 par le fait que $\mathcal{D}_{k+1}(A) = 0$ et $\mathcal{D}_k(A) = \langle 1\rangle$, on déduit facilement l'équivalence de 1 et 2.
Le corollaire 5.3 montre que 1 implique que le polynôme $t^{k(n-k)}g_k(t)$ est primitif. En particulier 1 implique 4. Il montre aussi l'équivalence de 6 et 5.
Le théorème 4.2 montre que 4 implique 5, lequel implique clairement 2.
On a donc l'équivalence des points 1 à 6.
L'équivalence de 2 et 7 résulte du théorème 2.8. En effet ce théorème nous dit que $A\,\mathrm{Adj}^{(k)}_{A^\circ}(A)\,A$

$\equiv\ g_k(t)\,A \bmod \mathcal{D}_{k+1}(A)$. Donc si A est de rang $\leq k$ on a l'égalité. Par ailleurs si on a l'égalité, A est localement simple sur $\mathbf{A}(t)$ (condition 5 dans le théorème 3.3), et le rang est fourni par le corollaire 5.3.
Le cas réduit a déja été vu (proposition 5.2 et corollaire 5.4). □

Décrivons maintenant un algorithme « rapide » pour savoir si une matrice est localement simple de rang constant, et, en cas de réponse positive, pour calculer un inverse généralisé de la matrice. Cet algorithme fonctionne en utilisant la caractérisation 7 dans le théorème précédent.

Comme nous nous intéressons pour le moment uniquement au rang constant, les seuls tests dont nous aurons besoin sont les suivants : le test d'égalité à 0 dans $\mathbf{A}$ et le test « $1 \in \langle x_1, \ldots, x_n\rangle$? ».

On procède comme suit.

1. On calcule A° défini par l'égalité (25).
2. On calcule les polynômes de Gram $g_k(t)$ définis par l'égalité (26). Ceci se fait en calculant le polynôme caractéristique de $A\,A^\circ$ si $m \leq n$ ou celui de $A^\circ\,A$ si $m > n$.
3. On cherche la plus grande valeur de k pour laquelle l'idéal $\mathcal{C}_k(A)$ est non nul. Pour cela on teste les polynômes de Gram généralisés $g_\ell(t)$ pour des valeurs décroissantes de ℓ et on s'arrête au premier non nul. Pour le plus grand k tel que $\mathcal{C}_k(A) \neq 0$ on teste si $\mathcal{C}_k(A) = \langle 1\rangle$. Si la réponse est négative A n'est pas localement simple de rang constant. Si la réponse est positive et si l'anneau est réduit alors $\mathcal{D}_{k+1}(A)$ est nul et la matrice est localement simple de rang k.
4. Dans tous les cas, si la réponse est positive, on calcule la matrice $B = \mathrm{Adj}^{(k)}_{A^\circ}(A)$ à coefficients dans $\mathbf{A}[t, 1/t]$ donnée par l'égalité (15) page 10. Ensuite on teste si :

$$A\,B(t)\,A = g_k(t)\,A \tag{27}$$

(ce test est inutile si l'anneau est réduit). Nous savons déjà (théorème 2.8) que

$$A\,B(t)\,A \ \equiv\ g_k(t)\,A \mod \mathcal{D}_{k+1}(A)$$

En cas de réponse négative, $\mathcal{D}_{k+1}(A) \neq 0$ et A n'est pas localement simple. En cas de réponse positive A est localement simple, au moins sur l'anneau $\mathbf{A}(t)$ car $t^{k(n-k)}\,g_k(t)$ est un polynôme primitif.

5. Il nous reste à calculer un inverse généralisé de A à coefficients dans $\mathbf{A}$. L'égalité (27) peut être lue en chaque degré t^ℓ (avec $-k(n-k) \leq \ell \leq k(m-k)$) comme une égalité dans $\mathbf{A}^{m\times n}$: $A\,B_\ell\,A = g_{k,\ell}\,A$. Comme on connaît une combinaison linéaire $\sum_\ell \alpha_\ell\,g_{k,\ell}$ des coefficients de $g_k(t)$ qui est égale à 1, la combinaison linéaire correspondante des égalités en chaque degré ℓ nous donne la matrice $B' = \sum_\ell \alpha_\ell\,B_\ell$ qui vérifie $A\,B'\,A = A$.

Ici il semble peu probable que l'on ait aussi $B'\,A\,B' = B'$, sauf si les α_ℓ sont obtenus en spécialisant t (ce qui peut se faire si $g_k(\tau)$ est inversible pour une valeur particulière de τ). De toute façon, on peut toujours remplacer B' par $B'\,A\,B'$ pour avoir un vrai inverse généralisé.

Voyons maintenant la complexité de cet algorithme.

Nous n'utilisons ni la multiplication rapide des matrices, ni celle des polynômes, qui, naturellement, amélioreraient de façon substantielle les bornes calculées.

Voici notre calcul des bornes, étape par étape.

1. Coût négligeable.

2. Posons $p = \inf(m,n)$. Le calcul du polynôme caractéristique consomme $\mathcal{O}(p^4)$ opérations arithmétiques dans $\mathbf{A}[t]$ portant sur des polynômes de degré $\leq p\,(n+m)$ et donc $\mathcal{O}(p^6\,(n+m)^2)$ opérations arithmétiques dans $\mathbf{A}$.
3. Cette étape consomme $\mathcal{O}((k\,(n+m-2k))^s)$ opérations élémentaires.
4. Le nombre d'opérations arithmétiques est en $\mathcal{O}(p^5\,(n+m)^2)$, et le test consomme $\mathcal{O}((p^3\,(n+m))$ opérations élémentaires.
5. Nombre d'opérations arithmétiques négligeable par rapport aux étapes 2 ou 4.

Résumons.

Théorème 5.6 *Soit* $\mathbf{A}$ *un anneau avec test d'égalité à* 0 *et test* « $1 \in \langle x_1,\ldots,x_n\rangle$? ». *On peut tester si une matrice* $A \in \mathbf{A}^{m\times n}$ *est localement simple de rang constant, et en cas de réponse positive, calculer un inverse généralisé de la matrice. Soit* $p = \min(m,n)$, $q = \max(m,n)$. *Si le premier test consomme une opération élémentaire et si le deuxième consomme un nombre d'opérations élémentaires en* $\mathcal{O}(n^s)$, *ces calculs consomment* $\mathcal{O}(p^6\,q^2)$ *opérations arithmétiques et* $\mathcal{O}(p^3\,q+q^{2s})$ *autres opérations élémentaires. Avec les mêmes bornes de complexité, on calcule un inverse généralisé de* A *et des matrices de projection sur le noyau et sur l'image de* A.

Dès que $p \geq 2$ ces bornes sont, pour q assez grand, bien meilleures que celles obtenues si on exécute (naïvement) un algorithme qui calcule tous les mineurs de la matrice.

Le cas général

Le cas général se ramène au cas précédent : cf. lemme 0.1 et section 4.4.
On obtient donc.

Théorème 5.7 *Soit une matrice* $A \in \mathbf{A}^{m\times n}$. *Les propriétés suivantes sont équivalentes :*

1. A *est localement simple sur* $\mathbf{A}$.
2. A *est localement simple sur* $\mathbf{A}(t)$.
3. A *est quasi-simple sur* $\mathbf{A}(t)$.
4. A *et* A° *sont croisées sur* $\mathbf{A}(t)$.

Signalons que le lemme 0.1 peut être amélioré en raison du fait suivant : le produit de deux idéaux de type fini localement principaux donnés respectivement par n et m générateurs est un idéal de type fini donné par $n+m-1$ générateurs (voir par exemple [5]). Ceci nous permet d'obtenir la complexité suivante.

Théorème 5.8 *Sur un anneau* $\mathbf{A}$ *fortement discret, on peut tester si une matrice* $A \in \mathbf{A}^{m\times n}$ *est localement simple, et en cas de réponse positive, calculer un inverse généralisé de la matrice. Soit* $p = \min(m,n)$, $q = \max(m,n)$. *Si l'anneau est* $\mathcal{O}(n^s)$*-fortement discret, ces calculs consomment* $\mathcal{O}(p^6\,q^2+p\,q^4)$ *opérations arithmétiques et* $\mathcal{O}(p^4\,q+pq^{2s+1})$ *autres opérations élémentaires. Avec les mêmes bornes de complexité, on calcule un inverse généralisé de* A *et des matrices de projection sur le noyau et sur l'image de* A.

5.3 Les statisticiens indiens

Les numériciens puis les statisticiens ont développé une théorie des « inverses généralisés » d'abord pour le cas des corps $\mathbb{R}$ et $\mathbb{C}$ mais ensuite pour des anneaux commutatifs arbitraires. Cette théorie est essentiellement l'équivalent des théorèmes 3.3 et 3.6, avec des préoccupations particulières de calculs explicites et de formules précises. Ce sont les statisticiens indiens qui ont développé le plus cette théorie.

En fait cette convergence n'est pas fortuite. A la base, il y a le fait que numériquement, $\mathbb{R}$ et $\mathbb{C}$ ne se comportent « pas vraiment » comme des corps, à cause de la difficulté du test à zéro (voire son impossibilité) qui est la source de phénomènes d'instabilité, liés par exemple au calcul de l'inverse d'un nombre trop proche de 0.

Dans [13] la formule (19) est établie pour le cas suivant : la matrice de $\varphi^\bullet$ est de la forme $M\,{}^{\mathrm{t}}A^\star N$ où $x \mapsto x^\star$ est un automorphisme involutif de l'anneau $\mathbf{A}$, supposé intègre. Nous n'avons pas trouvé la formule (19) elle-même dans le cas le plus général, mais cela ne signifie pas qu'elle n'existe pas dans la littérature « indienne ». Il y a deux livres de référence pour ces écrits : [2] et [15].

De manière surprenante, nous avons rarement trouvé chez les statisticiens indiens de formules analogues à l'équation (23) du théorème 4.3 (il y en a une dans [15] dans un cas particulier) mais plutôt des formules du style (14) et (19).

Références

[1] Ben-Israel, A. and Greville, T. *Generalized Inverses : Theory and Applications.* New York : Wiley, 1977. 2

[2] Bhaskara Rao K. *The Theory of Generalized Inverses over a Commutative Ring.* Taylor & Francis. Londres, 2002. 2, 6, 7, 8, 13, 14, 25

[3] D. Bini and V. Y. Pan : *Polynomial and matrix computations.* Progress in Theoretical Computer Science, Birkhäuser, 1994. 2

[4] Diaz-Toca G., Gonzalez-Vega L., Lombardi H. : *Generalizing Cramer's Rule : Solving uniformly linear systems of equations.* SIAM Journal on Matrix Analysis and Applications. **27** n°3 (2005), 621–637. 4, 20, 21

[5] Ducos L., Lombardi H., Quitté C., Salou M. *Théorie algorithmique des anneaux arithmétiques, des anneaux de Prüfer et des anneaux de Dedekind.* Journal of Algebra. **281**, (2004), 604–650. 24

[6] Lancaster P. & Tismenetsky M. *The Theory of Matrices, 2/e* Academic Press (1985) 2, 7

[7] Lombardi H., Quitté C. *Théorie constructive élémentaire des modules projectifs de type fini.* Rapport technique 2002. http ://hlombardi.free.fr/publis/QLPTF1.pdf 2, 6, 12, 13, 15, 16

[8] Lequain, Y., Simis, A. *Projective modules over $R[X_1, ..., X_n]$, R a Prüfer domain.* J. Pure Appl. Algebra **18** (2) (1980), 165–171. 16

[9] Mines R., Richman F., Ruitenburg W. *A Course in Constructive Algebra.* Universitext. Springer-Verlag, (1988). 8

[10] K. Mulmuley : *A fast parallel algorithm to compute the rank of a matrix over an arbitrary field.* Combinatorica, **7**/1, 101–104, 1987. 20, 21

[11] Northcott D. *Finite free resolutions.* Cambridge tracts in mathematics No 71. Cambridge University Press, (1976). 2, 12, 13, 14

[12] Prasad K. *Generalized Inverses of Matrices over Commutative Rings.* Linear Algebra Appl. **211** (1994), 35–52. 14

[13] Prasad K., Bapat R. *The generalized Moore-Penrose inverse.* Linear Algebra Appl. **165** (1992), 59–69. 12, 25

[14] Mustapha Rais, Thèse (1970) : *Distributions homogènes sur des espaces de matrices*, Bulletin de la Société Mathématique de France, supplément au numéro de Juin 1972, mémoire n°30 11

[15] Rao C., Mitra S. *Generalized Inverses of Matrices and its Applications.* John Wiley & Sons (1971). 2, 25

[16] Turnbull H., *On differentiating a matrix*, Proc. Edinb. Math. Soc. **2** vol. 1, part 2, (1928), 111–128. 11

[17] Turnbull H., *The theory of determinants, matrices and invariants*, Dover (1960). 11

Table des matières