Approche polynomiale des systemes lineaires stationnaires causaux.Application a la representation cyclique

Laurent Baratchart

IRIA

CENTRE DE ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.:954 90 20

Rapports de Recherche

N° 143

APPROCHE POLYNOMIALE DES SYSTÈMES LINÉAIRES STATIONNAIRES CAUSAUX. APPLICATION À LA REPRÉSENTATION CYCLIQUE

Laurent BARATCHART

Juillet 1982

APPROCHE POLYNOMIALE DES SYSTEMES LINEAIRES STATIONNAIRES CAUSAUX. APPLICATION A LA REPRESENTATION CYCLIQUE.

Laurent BARATCHART

(Décembre 1981).

Abstract

This report consists of 7 paragraphs and 2 appendices. Only paragraph 6, together with appendix 1, and some parts of paragraph 2, may claim for some originality in content. However, to achieve coherence, they have been included in a more substantial development, which deals, in a formal style, with realization theory of linear, time invariant, causal systems over fields. Only the discrete-time case is described here. Paragraph 0 summarizes some classical results which provide the necessary mathematical basis, while paragraph 1 introduces the notion of a system and of an associated state-space. Paragraph 2 examines the realization problem, and the set of its solutions. Paragraph 3 is very short and introduces duality. Paragraph 4 proves some results about divisibility between matrices, which are of use in paragraph 5, where we develop the so-called "Führmann realization" and the canonical forms of a given system. Paragraph 6 deals with causal factorization, especially in the cyclic case. Appendix 1 generalizes some results from paragraph 6. Appendix 2 is a brief comment concerning the Führmann realization. Except for paragraph 0 which gives no proofs, the paper is self contained.

Résumé

Ce rapport comprend 7 paragraphes et 2 appendices. Seuls le paragraphe 6, l'appendice 1,quelques bribes du paragraphe 2, peuvent prétendre à une certaine originalité de contenu. Pour en faciliter l'abord cependant, on les a inclus dans un ensemble plus étoffé qui expose dans un style formel un certain nombre de résultats de la théorie de la réalisation des systèmes linéaires stationnaires causaux sur un corps. Seul le cas discret est illustré ici. Après un paragraphe 0 de rappels d'algèbre classiques, le paragraphe 1 introduit la notion de système et d'état. Le paragraphe 2 examine le problème de la réalisation et la multiplicité de ses solutions. Le paragraphe 3 est très court et introduit la notion de dualité. Le paragraphe 4 expose quelques résultats sur la divisibilité des matrices qui sont utilisés au paragraphe 5 pour l'examen de la représentation de Führmann et des formes canoniques d'un système. Le paragraphe 6 traite de la factorisation causale, et se spécialise au cas cyclique. L'appendice 1 lui fait suite en évoquant le cas général. L'appendice 2 est un bref commentaire sur la représentation de Führmann. Aucune connaissance n'est à priori requise pour lire l'ensemble.

PARAGRAPHE 0

Ce paragraphe introduit les préliminaires mathématiques auxquels on fera référence dans la suite. Essentiellement, il énonce le théorème de structure des modules de type fini sur les anneaux principaux et quelques corollaires. Les théorèmes 1 et 2 sont fréquemment invoqués, notamment aux paragraphes 2,5,6. Aucune démonstration n'est donnée, on pourra se reporter à [BI], [JA], [LG] pour celles-ci.

0. Préliminaires.

Définition 1 :

Un ensemble A est un anneau s'il est muni de deux lois, addition et multiplication, (on note $x+y$ la somme et xy le produit de 2 éléments), telles que :

i) $(A,+)$ est un groupe commutatif.

ii) la multiplication est associative et a un élément neutre.

iii) $\forall (x,y,z) \in A \quad (x+y)z = xz+yz$ et $z(x+y) = zx+zy$.

A est dit commutatif si $xy = yx \quad \forall (x,y) \in A^2$.

A est dit intègre si $xy = 0 \Rightarrow \begin{matrix} x = 0 \\ \text{ou} \\ y = 0. \end{matrix}$

Définition 2 :

Un sous-anneau I d'un anneau A commutatif est un idéal si $\forall x \in A$, $y \in I$, $xy \in I$. Par exemple ($\forall a \in A$), l'ensemble (a) des multiples de a est un idéal, appelé engendré par a.

Définition 3 :

Un anneau commutatif intègre est dit principal si et seulement si tout idéal est engendré par un élément.

Exemple :

L'anneau des entiers $\mathbb{Z}$, ou celui des polynômes en une indéterminée sur un corps K, $K[x]$, sont des anneaux principaux.

Définition 4 :

Un groupe additif commutatif M est dit module sur l'anneau A, s'il est muni d'une opération : $A \times M \to M$ telle que $(a+b)x = ax+bx$, $a(x+y) = ax+ay$, $a(bx) = ab(x)$, et $1.x = x \quad \forall_b^a \in M$.

Remarque :

Un module sur un corps est un espace vectoriel. Mais la structure d'espace vectoriel est très particulière parmi les structures de module. Par exemple tout espace vectoriel admet une base. Ceci n'est pas le cas en général dans un module et amène la définition suivante.

Définition 5 :

Un module est dit libre s'il admet une base (un système libre et générateur).

Proposition 1 :

Si M est un module libre le cardinal de toutes les bases est le même, et c'est la dimension de M (éventuellement infinie), ou encore son rang.

Exemple :

Si $p \in \mathbb{Z} - \{0\}$, $\frac{\mathbb{Z}}{p\mathbb{Z}}$ n'est pas libre sur $\mathbb{Z}$; $\frac{K[z]}{(p)}$, où (p) est l'idéal engendré par $p \in K[z]$, non plus (sauf si p=0). Par contre $K^r[z] = \underbrace{K[z] \times K[z] \times \ldots \times K[z]}_{r \text{ fois}}$ est libre, une base étant par exemple la base canonique.

Un autre phénomène qui peut intervenir dans un module et non dans un espace vectoriel est le suivant : il peut arriver que pour $a \in A$ et $x \in M$, on ait $ax = 0$. Sans pour autant avoir a=0 ou x=0. Ceci amène une autre notion :

Définition 6 :

Un élément $x \neq 0$ d'un module M sur un anneau A est dit de torsion si il existe $a \neq 0 \in A$ tel que ax=0. Un module est dit de torsion si tous ses éléments sont de torsion.

Exemple :

Si $p \in \mathbb{Z} - \{0\}$, $\frac{\mathbb{Z}}{p\mathbb{Z}}$ est de torsion, en tant que module sur Z : p est un annulateur. De même $\frac{K[z]}{(p)}$. Un module libre est sans torsion.

Introduisons à présent l'analogue des applications linéaires dans les espaces vectoriels :

Définition 7 :

Soient M et M' deux modules sur le même anneau A. Une application $f : M \to M'$ est dite homomorphisme de modules entre M et M' si :

$$\forall_{y}^{x} \in M,\ a \in A : \begin{array}{l} f(x+y) = f(x)+f(y) \\ f(ax) = af(x). \end{array}$$

Les deux théorèmes qui viennent explicitent la structure des modules sur un anneau principal, pourvu qu'ils soient de "type fini", c'est-à-dire qu'ils soient engendrés par un nombre fini d'éléments (sans qu'il y ait nécessairement indépendance, ce qui en fait une notion distincte des modules libres).

Introduisons la notion de module quotient : si M est un module sur un anneau A, M' un sous module, on considère la relation d'équivalence dans M : $x \sim y$ si et seulement $x-y \in M'$.

On notera $\bar{x}$ la classe de x par cette relation. Alors si on définit $\bar{x}+\bar{y} = \overline{x+y}$ et $a\bar{x} = \overline{ax}$ pour $x,y \in M$, $a \in A$, on voit que l'ensemble des classes d'équivalence est muni d'une structure de module sur A. (Le seul point à vérifier en réalité est que ces relations sont bien définies, c'est-à-dire ne dépendent pas du représentant choisi dans la classe. Ceci résulte de ce que M' est un sous module de M). L'ensemble des classes d'équivalences muni de sa structure de module est appelé module quotient de M par M' et noté $\frac{M}{M'}$.

L'application $\rho : M \to \frac{M}{M'}$, définie par $\rho(x) = \bar{x}$ est un homomorphisme de modules qui est surjectif. On l'appelle surjection canonique de M sur $\frac{M}{M'}$.

Théorème 1 :

Soit L un module libre sur un anneau principal A, et M un sous module, il est libre de rang inférieur à celui de L; s'il est de rang fini n, il existe alors une base B de L, n éléments e_i de B, et n éléments non nuls α_i de A ($1 \leq i \leq n$) tels que :

a) Les $\alpha_i e_i$ forment une base de M.

b) α_i divise α_{i+1} pour $1 \leq i \leq n-1$.

De plus le module M' ayant pour base (e_i) et les idéaux principaux $A\alpha_i$ sont déterminés de façon unique par ces conditions ; $\frac{M'}{M}$ est le sous module de Torsion de $\frac{L}{M}$ et est isomorphe à la somme directe des A modules $\frac{A}{(\alpha_i)}$; enfin $\frac{L}{M}$ est somme directe de $\frac{M'}{M}$ et d'un module libre isomorphe à $\frac{L}{M'}$.

Corollaire 1 :

Sur un anneau A principal, tout module de torsion de type fini n est isomorphe à une somme directe $\sum_{\oplus i=1}^{n} \frac{A}{(\alpha_i)}$ avec $\alpha_1 | \alpha_2 | \ldots | \alpha_{i+1}$, $\alpha_i \neq 0$ ("|" signifie "divise")

Remarque :

Si certains α_i valent 1 (sont inversibles dans A), les termes correspondant de la somme sont nuls, i.e. ne contribuent pas.

Corollaire 2 :

Tout module de type fini sur un anneau principal est somme directe d'un module libre et d'un module de torsion.

Définition 8 :

Une matrice de $A^{p\times p}$ inversible dans $A^{p\times p}$ est dite unimodulaire. Cela revient à dire que son déterminant est inversible dans A.

Définition 9 :

On appelle opération élémentaire sur les lignes d'une matrice $M \in A^{p\times m}$ l'une des opérations suivantes :

i) Echanger deux lignes.

ii) Ajouter à une ligne une autre multipliée par un élément de A.

iii) Multiplier une ligne par un élément inversible de A.

Faire une telle opération revient à multiplier à gauche par une matrice unimodulaire, comme on s'en convaincra aisément (cf.[JA]). Une telle matrice est dite "élémentaire".

On définit de même des opérations élémentaires sur les colonnes qui correspondent à des multiplications à droite par des matrices unimodulaires. On peut montrer que toute matrice unimodulaire est produit de matrices élémentaires (conséquence aisée du théorème 2 à venir).

Théorème 2 :

Soit A un anneau principal, M une matrice de type p×n sur A.

i) Il existe U unimodulaire dans $A^{p\times p}$, telle que :

$$UM = \begin{pmatrix} \alpha_{11} & & X \\ & \alpha_{22} & \\ 0 & & \ddots \end{pmatrix}$$

avec U produit de transformations élémentaires sur les lignes (cf. Définition 7).

Si A est par exemple l'anneau K[z] des polynômes sur un corps K, on peut choisir $d° a_{ij} < d° a_{jj}$.

ii) Il existe $V^{m\times m}$ unimodulaire dans $A^{m\times m}$, telle que :

$$MV = \begin{pmatrix} a'_{11} & 0 \\ X & \ddots \end{pmatrix}$$

V produit de transformations élémentaires sur les colonnes.

Si A = K[z], on peut choisir $d° a_{ij} < d° a_{ii}$.

iii) Il existe $U,V \in A^{p\times p}$ et $A^{m\times m}$, unimodulaires, telles que :

$$UMV = \begin{pmatrix} a_{11} & & 0 \\ & a_{22} & \\ 0 & & \ddots \end{pmatrix}$$

avec $a_{ii} | a_{i+1,i+1}$, et U et V produits de matrices élémentaires. Les a_{ii} sont alors les invariants du module $M.A^n$ dans A^p.

Corollaire 1 :

Si A est principal, L un A. module libre de base finie (u_j) $1 \leq j \leq k$, M un sous module de L, (x_ℓ) un système de générateurs de M et (α_i) $(1 \leq i \leq n)$ les facteurs invariants de M par rapport à L, alors pour $1 \leq m \leq n$, $\delta m = \alpha_1 \ldots \alpha_m$ est un p.g.c.d. des mineurs d'ordre m de la matrice dont les colonnes sont formées avec les coordonnées des (x_ℓ) par rapport à la base (u_j).

Remarque 1 :

Ceci relie les a_{ii} du théorème, partie iii), avec les éléments initiaux de M.

Remarque 2 :

Les invariants de M dans L, dépendent de L : si L=M, ils sont par exemple tous inversibles.

Corollaire 2 :

Si D est une matrice régulière de $A^{k \times k}$ (i.e. dét D $\neq 0$), alors $\frac{A^k}{DA^k}$ est un module de torsion et réciproquement.

Corollaire 3 :

Si $A = K[z]$, $D \in A^{k \times k}$ est régulière, la dimension de $\frac{A^k}{DA^k}$ comme espace vectoriel sur K est égale à degré (dét D).

Remarque :

Ceci se voit aisément en écrivant ($\approx$ signifie isomorphe).

$\frac{A^k}{DA^k} \approx \sum\limits_{\oplus} \frac{K[z]}{(a_{ii})}$ où les a_{ii} sont les invariants de D. La dimension en tant qu'espace vectoriel sur K de $\frac{K[z]}{(a_{ii})}$ est d° a_{ii} (penser à la division euclidienne), donc dim $(\frac{A^k}{DA^k}) = \Sigma$ d° (a_{ii}) = d° (dét D).

Donnons pour terminer deux résultats sur les modules quotients qu'on utilisera aux §.2 et 5.

Soit $M_1 \supset M_2 \supset M_3$ une "tour" de modules :

Considérons l'application $f : \frac{M_1}{M_3} \to \frac{M_1}{M_2}$ définie par $f(\bar{x}_3) = \bar{x}_2$ où "$\bar{x}_i$" indique la classe de $x \in M_1$ dans $\frac{M_1}{M_i}$. Son noyau est $\frac{M_2}{M_3}$. Elle est surjective. Ainsi on a :

Proposition 2 :

(Isomorphisme du double-quotient).

On a un isomorphisme $\bar{f} : \dfrac{\frac{M_1}{M_3}}{\frac{M_2}{M_3}} \to \frac{M_1}{M_2}$

Donnons nous à présent deux sous-modules M_1 et M_2 d'un même module M. M_1+M_2 est un module (ensemble des sommes x_1+x_2, avec $x_1 \in M_1$ et $x_2 \in M_2$).

Considérons l'application $g : M_1 \to \frac{M_1+M_2}{M_2}$ définie par $g(x) = \bar{x}_2$, classe de x dans le module quotient de droite. Son noyau est $M_1 \cap M_2$. Elle est surjective. On a alors :

Proposition 3 :

On a un isomorphisme :

$$\bar{g} : \frac{M_1}{M_1 \cap M_2} \to \frac{M_1+M_2}{M_2} \quad .$$

PARAGRAPHE 1

Ce paragraphe introduit quelques formalismes, qui serviront ensuite, essentiellement celui des séries formelles (commutatives en une variable) et la notion de système linéaire stationnaire vue à travers ces formalismes. On y développe rapidement le concept d'état et de représentation interne au moyen de manipulations algébriques élémentaires (essentiellement le quotientage), dans un ordre d'idées analogue au chapitre 10 de [KFA]. On en déduit les relations entre rationnalité et dimension d'état par une application du théorème 1 §.0. On prouve enfin la relation $\mathcal{H} = H(zI-F)^{-1}G$. Un automaticien n'y trouvera aucun résultat inconnu.

1. La représentation de Nérode.

A) Pour fournir un support physique intuitif aux considérations élémentaires qui suivent, on a adopté le cadre des systèmes linéaires stationnaires discrets, que nous décrivons brièvement. Les ingrédients sont les suivants :

i) Le corps des séries formelles :

Si K est un corps, qui dans toute la suite sera désigné par cette lettre comme le "corps de base", on en construit un autre qui le contient, dont les éléments sont des séries $\sum_{i>N} a_i z^{-i}$, dont les coefficients a_i sont dans K, et dont tous les coefficients en puissances positives sont nuls à partir d'un certain rang comme l'indique le N<i où N est un entier relatif non spécifié, et qui varie avec chaque série : il ne s'agit pas d'une borne uniforme pour les degrés. En d'autres termes, une série S s'écrit :

$$S = a_{-N} z^{N} + a_{-(N-1)} z^{N-1} + \ldots a_0 + a_1 z^{-1} + \ldots$$

On a annoncé que l'ensemble, des séries de ce type, qu'on notera K[[z]] est un corps. Il faut donc décrire les opérations dont on le munit :

- L'addition se fait naturellement, composante à composante.

- La multiplication nécessite plus d'explications : on l'effectue comme une multiplication normale en groupant les puissances de z égales, mais encore faut il s'assurer que les coefficients ainsi obtenus sont des sommes finies dans K, et non des séries. Le coefficient d'ordre j du produit de $S_1 = \sum_{i>N_1} a_i z^{-i}$ et $S_2 = \sum_{i>N_2} b_i z^{-i}$ est $\sum_{i+k=-j} a_i b_k$. Il faut donc s'assurer que seul un nombre fini de couples (i,k) donnent une contribution non nulle à cette somme. Mais $i=-j-k$ entraîne que les seuls termes importants sont obtenus pour $i \leq -j-N_2$ car si $k<N_2$ les termes b_k sont nuls. Comme i lui même peut être supposé $\geq N_1$, seul un nombre fini de i interviennent, donc un nombre fini de k à cause de la relation $i+k=-j$. On voit donc que le caractère "fini à gauche" des séries considérées est essentiel pour définir la multiplication.

Les propriétés assurant que K[[z]] est un corps sont toutes triviales sauf peut être l'existence de l'inverse d'une série $S = \sum_{i>N} a_i z^{-i}$ avec $a_N \neq 0$. Pour s'en convaincre, écrivons l'inverse éventuel $S^{-1} = \sum_{i>N'} b_i z^{-i}$. La relation $SS^{-1} = 1$ impose alors :

$$N' = -N$$

$$b_{N'} = \frac{1}{a_N}$$

$$b_{N'+1}a_N + b_{N'}a_{N+1} = 0 \Leftrightarrow b_{N'+1} = \frac{1}{a_N}\left(-b_{N'}a_{N+1}\right)$$

$$b_{N'+2}a_N + b_{N'+1}a_{N+1} + b_{N'}a_{N+2} = 0 \Leftrightarrow b_{N'+2} = \frac{1}{a_N}\left(-b_{N'+1}a_{N+1} - b_{N'}a_{N+2}\right)$$

On voit qu'on a un système triangulaire qu'on résout de proche en proche.

L'interprétation d'une série est la suivante : la suite a_N, a_{N+1}, ... est la suite des valeurs prises par un signal à des instants discrets (par exemple toutes les secondes !), la présence de z servant à indexer le temps : l'instant i correspond à la puissance z^{-i}. On peut se demander pourquoi cette inversion de signe, i.e. pourquoi on ne note pas $\Sigma a_i z^i$. C'est évidemment là affaire de convention. Cependant, la nôtre, qui peut paraître la plus compliquée, se trouve devenir plus commode par la suite, dans notre optique qui est de faire jouer un rôle central à la structure de module sur K[z]. En outre les relations entre causalité et valuation (voir ci-après) se présentent aussi plus naturellement.

Enfin remarquons qu'aucune idée de sommation pour certaines valeurs de la variable z n'est sous-jacente dans l'écriture : les séries sont "formelles".

ii) La valuation de K[[z]] :

Si $S = a_{-N}z^{N} + a_{-N+1}z^{N-1} + \dots$ avec $a_{-N} \neq 0$, appelons N le degré d_S de la série S. C'est le prolongement naturel du degré habituel sur les polynômes. On a évidemment $d_{S1+S2} \leq \sup\{d_{S1},d_{S2}\}$, et $d_{S1.S2} = d_{S1} + d_{S2}$. Posons alors $v(s) = \exp(d_S)$, on a :

$$\begin{cases} v(S1+S2) \leq \sup(v(S1),v(S2)) \quad (*) \\ \\ v(S1.S2) = v(S1).v(S2) \end{cases}$$

v est une valeur absolue sur K[[z]], et même une valuation (cela signifie qu'elle vérifie une inégalité triangulaire "gordini" comme le montre (*)). Cette valuation permet de mesurer la "taille" d'une série, mais c'est une taille d'une nature un peu particulière : une série est d'autant plus petite qu'elle "commence plus tard". Une série S est dite propre si $v(s) < 1$, ce qui signifie que S "commence" à l'instant 1 au plus tôt.

Remarquons qu'une fraction rationelle s'identifie naturellement à une série : Si q est un polynôme on peut écrire avec la division en puissances décroissantes $\frac{1}{q} = \sum_{i \geq d°q} a_i z^{-i}$ (il s'agit du calcul de l'inverse effectué plus tôt). Pour une fraction $\frac{p}{q}$, son degré est manifestement d° P - d° Q. Il est facile de vérifier que l'identification précédente est une identification algébrique : K(z), corps des fractions rationelles s'identifie naturellement à un sous corps de K[[z]]. On peut même vérifier que K[[z]] est le complété de K(z) pour la valuation v.

iii) L'espace vectoriel K^p[[z]] :

K^p[[z]] dont les éléments peuvent s'écrire comme des vecteurs à p composantes (qui sont elles-mêmes des séries) est un espace vectoriel de dimension p sur K[[z]] pour les opérations naturelles. Remarquons que K^p[[z]] s'identifie naturellement avec l'ensemble des séries formelles dont les coefficients sont des vecteurs de K^p. On utilisera indifféremment ces deux notions. Ainsi les éléments de K^p[[z]] s'interpèteront aussi comme des signaux indexés sur le temps, dont les valeurs sont des vecteurs à p composantes.

iv) La notion de système dynamique linéaire stationnaire discret et causal :

Pour nous le mot système dynamique signifie qu'il va s'agir d'une application ψ d'un ensemble de signaux d'entrée $K^p[[z]]$ dans un ensemble de signaux, de sortie $K^m[[z]]$. Ensuite le mot linéaire indique que c'est une application K-linéaire de

$K^p[[z]] \to K^m[[z]]$: $\psi(S1+S2) = \psi(S1) + \psi(S2)$ et $\psi(\lambda S1) = \lambda\psi(S1)$ pour $\lambda \in K$.

Le mot stationnaire signifie que la correspondance $S1 \to \psi(S1)$ ne dépend pas de l'instant où elle a lieu, donc que le moment où on "injecte" le signal S1 importe peu : cela se traduit par $\psi(z.S1) = z\psi(S1)$ qui entraîne aussi $\psi(S1) = z\psi(\frac{S1}{z}) \Rightarrow \psi(\frac{S1}{z}) = \frac{1}{z}\,\psi(S1)$. Arrivés à ce point, on constate que ψ est K(z) linéaire, c'est-à-dire que c'est une application linéaire de $K^p[[z]]$ dans $K^m[[z]]$, chacun étant considéré comme espace vectoriel sur le corps K(z) des fractions rationelles. (Distinguer K[z] = polynômes, et K(z) = fractions rationelles). On aimerait en déduire que ψ est K[[z]] linéaire. C'est hélas faux en général. Sans vouloir discuter ici ce problème en détails, indiquons qu'il est lié au problème de la continuité de ψ en tant qu'opérateur de $K^p[[z]] \to K^m[[z]]$ (vus comme e.v. sur K(z)), et au fait que K[[z]] n'est pas algébrique sur K(z). Cf.[MU]. C'est ici qu'intervient le fait que ψ est causal. Cela signifie que la sortie ne peut pas "commencer avant" l'entrée qui la produit. Ceci impose que ψ est K[[z]] linéaire, chose qu'on peut prouver ainsi : Soit S une série de K[[z]], U une entrée de $K^p[[z]]$. On veut prouver $\psi(SU) = S\psi(U)$. A cause de la linéarité par rapport à la multiplication par z, on peut toujours supposer que d° U = 0. Prouvons l'égalité du coefficient d'ordre k de ces deux séries de $K^m[[z]]$: posons $S = P.z^{-k} + S'_{k+1}$ où P est un polynôme et S'_{k+1} une série de degré $\leq -(k+1)$. On a $\psi(SU) = \psi(P.z^{-k}U+S'_{k+1}U)$ donc $\psi(SU) = \psi(Pz^{-k}U) + \psi(S'_{k+1}U)$. Le degré du deuxième terme, puisque ψ est causale, est majoré par $-(k+1)$, donc les termes de degrés $\geq -k$ des deux séries $\psi(SU)$ et $\psi(Pz^{-k}U)$ sont égaux, et en particulier le terme d'ordre k (il est flanqué d'une puissance -k). Un raisonnement en tout point analogue montre que les termes de degré $\geq -k$ des séries $Pz^{-k}\psi(U)$ et $S\psi(U)$ sont les mêmes. Il ne reste plus qu'à utiliser la linéarité sur les fractions rationelles pour conclure car $\psi(Pz^{-k}U) = Pz^{-k}\psi(U)$.

Remarque :

On se rend compte qu'une preuve analogue marche si au lieu de supposer ψ causale, on suppose seulement qu'il existe $N \in \mathbb{Z}$ tel que $d° \psi(S) \leq N + d° S$, autrement dit si on met une "borne uniforme" à la non causalité. En terme de valuation, cela s'écrit $v(\psi S) \leq (e^N)v(S)$. Il faut relier ce qui a été dit précédemment avec le fait que ceci est exactement la condition de continuité d'un opérateur linéaire de $K^p[[z]] \to K^m[[z]]$, si on munit ces espaces de la norme du sup des valuations des composantes (c'est ce qu'il faut comprendre lorsqu'on écrit $v(U)$ avec $U \in K^p[[z]]$). D'ailleurs comme $K[[z]]$ est complet, il n'y a qu'une topologie d'espace vectoriel normé sur $K[[z]]$ pour un espace de dimension finie, donc le choix de cette norme est indifférent cf.[LG].

En résumé, ψ peut donc se représenter, dans les bases canoniques

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \leftarrow i^{\text{ème}} \text{ ligne} \\ \\ \\ \\ \\ \end{matrix}$$

de $K^p[[z]]$ et $K^m[[z]]$ par une matrice qu'on appèlera son transfert, de dimension $m \times p$, dont les éléments sont des séries de $K[[z]]$, qu'on notera $H = (h_{ij})$ $1 \leq i \leq m$, $1 \leq j \leq p$. L'image d'une entrée

$$U = \begin{bmatrix} u_1 \\ \vdots \\ u_p \end{bmatrix}$$

où les u_i sont des séries de $K[[z]]$ sera

$$Y = H.U. = \begin{bmatrix} \Sigma h_{1j} u_j \\ \vdots \\ \Sigma h_{mj} u_j \end{bmatrix}$$

produit matriciel habituel.

B) Construction de l'état.

Plaçons-nous à l'instant 0, par exemple, et supposons qu'une entrée $U = \sum_{i>N} u_i z^{-i}$ soit en train d'opérer sur le système H. Cela signifie que H a déjà "absorbé" les termes $U_{-N}, U_{-N+1}, \ldots, U_0$. La causalité de H entraîne que la sortie à l'instant 1 ne dépend que de ces termes : elle serait la même si on tronquait U à ce moment. Pour parler vaguement, on dit (cf.[FA]) que l'état est une "quantité" résumant tout le passé du système, permettant de connaître la sortie à venir si on connaît les entrées à venir. Autrement dit, l'information nécessaire pour "prendre les commandes", à l'instant considéré, ici l'instant 0. On peut toujours prendre comme état l'entrée $P = U_{-N}z^{+N} + \ldots + U_0$, car elle résume le passé puisqu'elle le décrit exhaustivement. En fait il est d'autres entrées qui produiraient les mêmes sorties pour des instants futurs, et il suffit d'en connaître une. Si Q est une telle entrée (polynomiale car le reste ne nous intéresse pas d'après la causalité) on voit par linéarité que (P-Q) est une entrée polynomiale générant une sortie, elle aussi, polynomiale. On vient de dire que l'on peut prendre pour état indifféremment P ou Q. Si on appelle V l'ensemble des entrées v polynomiales provoquant des sorties polynomiales, on peut donc prendre pour état l'un quelconque des $U+v$, $v \in V$. Ainsi ce qu'il importe de connaître c'est la classe $\bar{U}$ de U pour la relation d'équivalence : $P \sim Q \Leftrightarrow P-Q \in V$.

V est un module sur K[z] (trivial car si on multiplie $v \in V$ par un polynôme, la sortie est à fortiori polynomiale), sous module de $K^p[z]$. (qui n'est lui-même qu'une partie de $K^p[[z]]$, celle des séries qui n'ont pas de puissances négatives en z). Appelons $X = \frac{K^p[z]}{V}$ le module quotient. Notre état à l'instant 0 va être $\bar{U}$, classe de U dans X.

Pour connaître la sortie à l'instant 1, il suffit de calculer la sortie engendrée par un représentant quelconque de $\bar{U}$, et de prendre son coefficient d'ordre-1 : d'après leur définition le résultat ne dépendra pas du représentant choisi. Si S est une série, $S = \Sigma a_i z^{-i}$, on note $(.)_k$ l'application $S \rightarrow a_k = (S)_k$. L'application qui à $\bar{u} \in X$ fait correspondre la sortie à l'instant 1 sera notée $H : X \rightarrow K^m$.

Connaissant à présent l'état à l'instant 0 et l'entrée à l'instant 1, on peut calculer l'état à l'instant 1 qui lui-même permettra de connaître la sortie à l'instant 2, par la même application H d'après la stationnarité.

On a x_1, état à l'instant 1 qui vaut : (*) $x_1 = zp+u_1$. En effet si on se place à l'instant 1, cela revient à supposer, d'après la stationnarité que l'entrée U a "commencé" un instant plus tôt, et donc c'est la classe de $U_{-N}z^{N+1} + U_{-N+1}z^N + \ldots + U_0 z + U_1$ qu'il faut à présent prendre pour état. Appelons alors F l'application $X \xrightarrow{F} X$ $F(x) = zx$ qui signifie par définition des opérations sur le module quotient X que pour un représentant u de $\bar{u}$, on calcule $\overline{zu}$ (qui ne dépend pas du représentant). Appelons $G : K^p \to X$ définie par $v \to \bar{v}$, classe de v dans X. Alors l'équation (*) se lit : $x_1 = Fx_0 + Gu_1$ car par définition $\bar{P} = x_0$. Si on adjoint $y_1 = Hx_0$, ces équations deviennent familières à tout automaticien. La stationnarité implique alors les relations plus générales :

$$\begin{cases} x_{i+1} = Fx_i + Gu_{i+1} \\ y_{i+1} = Hx_i . \end{cases}$$

Le triplet (H,F,G) est appelé réalisation de H.

C) Dimension de l'espace d'état et rationnalité.

Jetons un coup d'oeil au théorème 1 du §.0. $K^p[z]$ est un module libre (une base possible sont les e_i de la base canonique) sur $K[z]$ qui est principal. V est un sous-module de type fini (puisque $K^p[z]$ l'est) de $K^r[z]$. On est donc exactement dans le cas d'application du théorème pour décrire la structure de X : il est isomorphe à une somme directe $\sum_{\oplus} \frac{K[z]}{(p_i)}$ où $p_1 | p_2 | \ldots$ ("|" signifie "divise") et où certains p_i peuvent être nuls. (Par convention, $0|0$). On sait alors, que X est de dimension finie sur K si et seulement si il est de torsion, c'est-à-dire si il existe un polynôme q non nul tel que $qx = 0 \quad \forall x \in X$. Alors on peut énoncer le :

Théorème 1 :

X est de dimension finie sur K si et seulement si H est rationnelle, c'est-à-dire si tous ses éléments sont des fractions rationnelles.

i) Si $\mathcal{H}$ est rationnelle :

Soit q un p.p.c.m. des dénominateurs des éléments h_{ij} de $\mathcal{H}$. Il est clair que $\forall x \in K^p[z]$, $qx \in V$. Ainsi $\overline{qx} = q\bar{x} = 0$. Donc X est de torsion, et donc de dimension finie sur K.

ii) Si X est de torsion :

Soit q un annulateur. Le produit $\mathcal{H}.qe_i$ est donc polynomial par définition de q. Ce produit est $q.\begin{pmatrix} h_{1i} \\ h_{2i} \\ \vdots \\ h_{mi} \end{pmatrix}$. Donc qh_{ki} est polynomial pour $1 \leq k \leq m$.

Ceci signifie que h_{ki} est rationnel, et comme c'est vrai pour tout i, on a terminé.

Remarque 1 :

Le théorème 1 du §.0. donne des renseignements sur l'espace d'état X même lorsque celui-ci n'est pas de dimension finie. On ne s'étendra cependant pas davantage sur le cas général et on supposera désormais que X est de K-dimension finie, c'est-à-dire que $\mathcal{H}$ est rationnelle.

Remarque 2 :

Dans le cas $\mathcal{H}$ rationnelle, les applications H,F,G qui sont évidemment K linéaires, peuvent être représentées par des matrices scalaires, si on a fait choix de bases, puisqu'on est de dimension finie. C'est là le gain de l'introduction de l'état, car $K^p[z]$ n'est pas de K-dimension finie (d'où l'intérêt, en l'occurence, de quotienter).

Théorème 2 :

Une base quelconque de X (espace d'état), et les bases canoniques de K^p et K^m étant choisies, appelons H,F,G les matrices associées aux applications précédemment définies, on a l'égalité entre matrices rationnelles :

$$\mathcal{H} = H(zI-F)^{-1}G.$$

Preuve :

Pour prouver cette égalité, il suffit de montrer que ces deux matrices ont même effet sur tout vecteur de K^p, puisque cet espace contient la base canonique. Posons, pour $u_0 \in K^p$ $\mathcal{H}u_0 = \sum_{i\geq 1} A_i z^{-i}$ (la condition $i\geq 1$ est celle de causalité). A_i est la sortie (dans K^m) générée par u_0, à l'instant i. Par ailleurs on a $(zI-F)^{-1} = z^{-1}(I-\frac{F}{z})^{-1}$. Donc :

$$(zI-F)^{-1} = z^{-1}\sum_{t=0}^{\infty} F^t z^{-t} = \sum_{t=0}^{\infty} F^t z^{-(t+1)}$$

Notons que cette égalité a lieu formellement entre séries à coefficients dans $K^{n\times n}$ si n est la dimension de l'espace d'état, cet ensemble étant un anneau (pour les opérations décrites au début de ce paragraphe), et l'élément $zI-F$ s'y trouvant être inversible. Mais $(zI-F)^{-1}Gu_0$ est un vecteur de $K^n(z)$ (fractions rationnelles) dont le développement en série est $\sum_{t=0}^{\infty} F^t Gu_0 z^{-(t+1)}$; ce fait est évident, u_0 étant scalaire. Notons que cette formule aurait un sens même si u_0 ne l'était pas, mais la somme devrait être interprêtée comme une série convergente au sens de la valuation.
Des considérations analogues montrent que le développement en série de $H(zI-F)^{-1}Gu_0$ dans $K^m[[z]]$ est $\sum_{t=0}^{\infty} HF^t Gu_0 z^{-(t+1)}$. A présent il résulte immédiatement des définitions de H,F,G que l'état à l'instant t est $F^t Gu_0$, (et 0 si $t<0$) et la sortie à l'instant t+1 est $HF^t Gu_0$. Ainsi $\sum_{t=0}^{\infty} HF^t Gu_0 z^{-(t+1)} = \sum_{t=1}^{\infty} HF^{t-1} Gu_0 z^{-t}$ est bien la sortie engendrée par u_0.

Remarque 1 :

Le résultat de ce paragraphe est que si $\mathcal{H}$ est rationnelle, il existe des matrices H,F,G, avec F carrée, et scalaires, telles que $\mathcal{H} = H(zI-F)^{-1}G$. Enoncée ainsi, cette affirmation ne paraît pas évidente.

Remarque 2 :

Il est facile de prouver : $\mathcal{H}$ rationnelle $\Leftrightarrow \dim_K X < \infty$, sans utiliser le théorème de structure des modules sur les anneaux principaux. C'est seulement un peu plus long.

PARAGRAPHE 2

Dans ce paragraphe, on étudie de plus près le rôle de la représentation définie au §.1. parmi les autres représentations. On mentionne son caractère universel dans certaines catégories, et partant, son caractère essentiellement unique. On caractérise également les matrices F des triplets (H,F,G).

On a fait libre usage des théorèmes du §.0, et le caractère fastidieux de certaines preuves a provoqué une rédaction moins détaillée que dans le reste de ce rapport.

2. <u>Le problème de la réalisation</u>.

A) Il est posé par l'étude des triplets de matrices de dimensions ($m\times n$, $n\times n$, $n\times p$), à coefficients scalaires (H,F,G) telles que $H(zI-F)^{-1}G = \mathcal{H}$, matrice rationnelle $m\times p$. On a vu au paragraphe 1 que la famille de ces triplets est non vide. En fait on verra même par la suite qu'on en a trouvé un élément particulier.

Si on possède un tel triplet, et si on considère une entrée $U = \sum_{i\geq N} u_i z^{-i}$, on peut introduire les suites x_k de K^n et y_k de K^m définie par :

$$x_k = 0 \quad \text{si} \quad k<N$$

$$\begin{cases} x_{k+1} = Fx_k + Gu_{k+1} \\ y_{k+1} = Hx_k \end{cases}$$

Les mêmes considérations qu'au §.1. montrent alors que $Y = \Sigma y_i z^{-i}$ est la sortie associée à l'entrée U, pour le système $\mathcal{H}$. Un triplet du type (H,F,G) ci-dessus sera dit réalisation de $\mathcal{H}$.

B) Si Y est un espace vectoriel de dimension finie sur K, F un endomorphisme de Y, on munira le couple (Y,F) de sa structure classique de module sur $K[z]$ en définissant l'opération externe ainsi :

$$\forall P(z) \in K[z],\ \forall y \in Y,\ P(z).y = P(F)y \quad \text{où} \quad P(F) \text{ est}$$

le polynôme P de l'endomorphisme F. Rappelons que dans ce cas, les polynômes invariants (cf.§.o) du $K[z]$ - module ainsi constitué sont les polynômes invariants de F (ceux dont on lit les coefficients dans la forme compagnon). Réciproquement, si on dispose d'un $K[z]$ - module de torsion de type fini, on définit avec la multiplication par z : $x \to z.x$ un endomorphisme d'espace vectoriel de dimension finie sur K, dont les polynômes invariants sont ceux du $K[z]$ - module initial. Il y a donc équivalence entre les deux notions :

- $K[z]$ - module de torsion de type fini.

- Espace vectoriel de dimension finie sur K, muni d'un endomorphisme F.

On les emploiera indifféremment.

C) Soit une réalisation (H,F,G) de $\mathcal{H}$. Définissons deux applications ϕ et ψ de la manière suivante :

Soit Y l'espace d'état (celui sur lequel opère F)

$$\phi : K^P[z] \to Y : \Sigma u^i z^i \overset{\phi}{\to} \Sigma F^i G u_i$$

ϕ fait correspondre à une entrée polynomiale l'état dans lequel elle amène le système à l'instant 0. ϕ est manifestement un homomorphisme de $K[z]$ modules (cf.B)).

$$\psi : Y \to \frac{K^m[[z]]}{K^m[z]} \qquad y \overset{\psi}{\to} \overline{\left(\sum_{i=1}^{\infty} HF^{i-1} y z^{-i} \right)}$$

ψ fait correspondre à l'état y la sortie qu'il génère si on n'exerce plus d'entrée sur le système, ou plus exactement la classe de cette sortie dans le $K[z]$ - module quotient $\frac{K^m[[z]]}{K^m[z]}$ (cela signifie qu'on ne s'intéresse pas à la partie polynomiale des signaux de sortie, ou encore qu'on ne s'intéresse qu'à l'avenir). ψ est aussi un homomorphisme de $K[z]$ modules (la présence du "—" est indispensable pour qu'il en soit ainsi) .

Enfin, introduisons l'application Λ :

$$\Lambda : K^P[z] \to \frac{K^m[[z]]}{K^m[z]} \qquad u \overset{\Lambda}{\to} \overline{\mathcal{H}u}$$

Λ fait correspondre à une entrée u (polynomiale) la classe de la sortie qu'elle engendre dans $\frac{K^m[[z]]}{K^m[z]}$.

Proposition 2 :

Soit (H,F,G) un triplet, ϕ et ψ les morphismes qui en sont déduits par le Théorème 1. Alors :

$$[(H,F,G) \text{ minimal}] \Leftrightarrow [\phi \text{ surjective, } \psi \text{ injective}]$$

Preuve :

Si ϕ n'est pas surjective, son image $\mathcal{Im}\ \phi$ est un sous-module propre de Y, donc $\dim_K \mathcal{Im}\ \phi < \dim_K Y$. En considérant la restriction de ψ à $\mathcal{Im}\ \phi$, on a un espace d'état de dimension inférieure et Y n'est pas minimal.

Si ψ n'est pas injective, soit Ker ψ son noyau. Alors soit ρ la surjection canonique : $Y \overset{\rho}{\to} \frac{Y}{\text{Ker } \psi}$. En composant G par ρ, et en définissant ψ de façon naturelle sur $\frac{Y}{\text{Ker } \psi}$, on voit qu'on peut prendre ce module quotient comme espace d'état, et il est de dimension inférieure à Y. De nouveau, Y n'est pas minimal.

Réciproquement, si ϕ est surjective et ψ injective, la dimension $\dim_K Y$ est exactement celle de $\frac{K^P[z]}{V}$ où V (cf.§.1) est le sous-module des entrées générant des sorties polynomiales, autrement dit le noyau de Λ. Ceci définit bien la dimension de Y de façon unique, car si Y' est tel que $\dim_K Y' < \dim_K Y$, on déduit de Y', comme dans la première partie un espace d'état de dimension inférieure ou égale tel que ϕ et ψ associées soient surjectives et injectives. On aboutirait alors à la contradiction:

$$\dim_K \frac{K^P[z]}{V} < \dim_K \frac{K^P[z]}{V} .$$

En particulier, on voit que la représentation de Nérode exhibée au paragraphe 1 est minimale. Bien que le résultat suivant soit une conséquence d'un théorème à venir, on peut l'établir à part. C'est celui qui figure dans le chapitre 10 de [KFB] sous le nom de "Fundamental theorem of linear realization theory".

Proposition 1 :

Le diagramme suivant est commutatif :

$$\begin{tikzcd} & Y \arrow[dr, "\psi"] & \\ K^P[z] \arrow[ur, "\phi"] \arrow[rr, "\Lambda"] & & \dfrac{K^m[[z]]}{K^m[z]} \end{tikzcd} \qquad (*)$$

Preuve :

Résulte immédiatement des définitions de (H,F,G).

Il y a une réciproque.

Théorème 1 :

La donnée d'une réalisation (H,F,G) de $\mathcal{H}$ est équivalente à celle de ϕ et ψ, homomorphisme de K[z] - modules, rendant commutatif le diagramme (*), où (ϕ,ψ) et (H,F,G) sont reliés comme indiqué ci-dessus.

Preuve :

On vient de voir que (H,F,G) donne ϕ et ψ comme souhaité. Réciproquement, si ϕ et ψ remplissent les conditions, posons $G : K^P \to Y \quad u_0 \to \phi(u_0)$ et $H : Y \to K^m \quad y \to (\psi y)_1$ (sélection du coefficient d'ordre−1). Enfin, prenons pour F la multiplication par z dans le K[z] module Y, il est trivial par linéarité que ϕ et ψ sont reliés à (H,F,G) comme on le veut. En outre, dire que le diagramme commute, c'est clairement dire que (H,F,G) est une réalisation de $\mathcal{H}$, les matrices $\mathcal{H}$ et $H(zI-F)^{-1}G$ se correspondant visiblement sur les entrées scalaires. On voit que le problème de la réalisation est celui de la factorisation de Λ. Ceci amène naturellement aux considérations de minimalité.

D)

Définition 1 :

Un triplet (H,F,G) sera dit minimal si c'est une réalisation de $\mathcal{H}$ telle que la dimension $\dim_K Y$ comme espace vectoriel sur K (qui est finie par définition) est minimale parmi celles de toutes les réalisations.

<u>Proposition 3</u> :

Tous les triplets minimaux d'un transfert H rationnel sont équivalents en ce sens que si (H_1,F_1,G_1) et (H_2,F_2,G_2) sont de tels triplets, il existe, avec des notations évidentes, un unique isomorphisme Θ de K[z]-modules entre Y_1 et Y_2 tel que le diagramme suivant soit commutatif :

$$\begin{tikzcd}
& Y_1 \arrow[dd, "\Theta"] \arrow[dr, "\psi_1"] & \\
K^P[x] \arrow[ur, "\phi_1"] \arrow[dr, "\phi_1"'] & & \dfrac{K^m[[z]]}{K^m[z]} \\
& Y_2 \arrow[ur, "\psi_2"'] &
\end{tikzcd}$$

<u>Preuve</u> :

Soit W_1 le noyau de ϕ_1 : l'injectivité de ψ_1 montre immédiatement que W_1 consiste en les entrées générant des sorties polynomiales, donc $W_1 = V$.

ϕ_1, surjective, introduit donc un isomorphisme $\tilde{\phi}_1 : \frac{K^P[x]}{V} \to Y_1$. De même pour ϕ_2, on définit $\tilde{\phi}_2$. Alors $\tilde{\phi}_2 . \tilde{\phi}_1^{-1}$ est manifestement le morphisme Θ cherché, son unicité étant évidente, vu l'injectivité des ψ_i.

<u>Remarque 1</u> :

Il est à noter qu'une contrainte de dimension sur K entraîne une similitude dans la structure de module. La proposition dit que les matrices F des triplets minimaux ont toutes la même forme compagnon.

<u>Remarque 2</u> :

Les conditions ϕ surjective et ψ injective sont appelées "commandabilité" et "observabilité". Elles se traduisent visiblement, si $\dim_K Y = n$, par :

$$\text{rg}[G, FG, F^2G, \dots, F^{n-1}G] = n$$

et $$\text{rg} \begin{bmatrix} H \\ HF \\ HF^2 \\ \vdots \\ HF^{n-1} \end{bmatrix} = n$$

conditions bien connues des automaticiens.

En effet $[G,FG,\ldots,F^{n-1}G]$ sont des colonnes de générateurs sur K de $\mathcal{I}m\ \phi$, tandis que $\begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix}$ $[y] = 0$ où $y \in Y$ est exactement la condition "$\psi(y) = 0$" (avec Cayley-Hamilton).

E) Donnons-nous un couple (Y,F) où Y est un e.v. sur K de dimension finie et F un endomorphisme (ou ce qui revient au même donnons-nous un $K[z]$ - module de torsion de type fini). Alors :

<u>Définition 2</u> :

Le couple (Y,F) est dit espace d'état pour $\mathcal{H}$ s'il existe H et G tels que (H,F,G) soit une réalisation de $\mathcal{H}$, ou, ce qui revient au même s'il existe (ϕ,ψ) rendant commutatif le diagramme désormais habituel. Remarquons que (ϕ,ψ) (ou H,G) ne sont pas unique pour un couple (Y,F) donné. Par exemple si ρ est un isomorphisme de : $Y \to Y$ alors $\rho . \phi$ et $\psi . \rho^{-1}$ conviennent. On dira que deux tels couples sont équivalents (il s'agit effectivement d'une relation d'équivalence). On peut alors reposer la question sous la forme : (Y,F) donné, n'existe-t-il qu'un couple (ϕ,ψ) à une équivalence près ? Sans autres hypothèses sur ϕ et ψ, la réponse est non. Par exemple si $Y = \mathcal{I}m\ \phi \oplus M$ où M est un module arbitraire, il est clair que la définition de ψ sur M l'est aussi. Cependant on a un élément de réponse avec la proposition suivante :

<u>Proposition 4</u> :

Si la dimension $\dim_K Y$ est minimale, tous les couples ϕ,ψ sont équivalents.

<u>Preuve</u> :

Résulte immédiatement de la preuve de la proposition 3.

Le résultat suivant fournit une caractérisation des espaces d'états d'un transfert $\mathcal{H}$.

Théorème 2 :

Soit $X = \sum_{\oplus\, i=1}^{n} \frac{K[z]}{(a_i)}$ la décomposition cyclique de l'espace d'état minimal (elle est la même pour tous d'après la proposition 3). Soit $Y = \sum_{\oplus\, i=1}^{r} \frac{K[z]}{(b_i)}$ la décomposition cyclique d'un module de torsion de type fini sur $K[z]$. Y (en tant que module) est espace d'état pour H si et seulement si $r \geq n$ et il existe une sous-suite $b_{i_1}, \dots, b_{i_r}$ des b_i telle que $a_j | b_{i_j}$. Si c'est le cas, on peut toujours prendre comme sous-suite les n derniers des b_i

Preuve :

Montrons rapidement que les deux conditions sur les invariants données par l'énoncé sont équivalentes. La seconde entraîne évidemment la première. Réciproquement supposons la première.

On peut ordonner les b_{i_j} par divisibilité croissante. En effet, supposons par exemple $i_1 > i_2$. Alors les relations de divisibilité sur les invariants donnent :

$$\begin{matrix} a_1 | b_{i_1} \\ a_2 | b_{i_2} \end{matrix} \quad \underline{\text{et}} \quad \left.\begin{matrix} a_1 | a_2 \\ b_{i_2} | b_{i_1} \end{matrix}\right\} \text{ donc } \left\{\begin{matrix} a_1 | b_{i_2} \\ a_2 | b_{i_1} \end{matrix}\right.$$

on a donc rétabli l'ordre des indices i_1 et i_2. Si on suppose alors les i_j ordonnées, si $a_j | b_{i_j}$, il divise aussi $b_{i_{(j+1)}}$. Donc les b_{i_j} peuvent être pris consécutifs et décalés "vers la droite".

Passons à la démonstration du théorème proprement dit :

i) Supposons la condition remplie : on peut écrire $Y = E \oplus \sum_{\oplus\, i=1}^{n} \frac{K[z]}{(a_i d_i)}$ où E est un module qui regroupe le "reste de la somme". Choisissons un des termes $\frac{K[z]}{(a_i d_i)}$, et considérons son sous module M_i engendré par d_i. Il est isomorphe à $\frac{K[z]}{(a_i)}$. Donc Y contient un sous-module M isomorphe à $\sum_{\oplus i=1}^{n} \frac{K[z]}{(a_i)}$ c'est $\{d_1\} \oplus \{d_2\} \oplus \ldots \oplus \{d_n\}$. On peut donc construire $\phi : K^P[z] \to M \underset{inj}{\to} Y$ où la dernière flèche représente l'injection canonique. On peut ensuite trouver $\psi : M \to \frac{K^m[[z]]}{K^m[z]}$ de telle sorte que M soit espace d'état, d'ailleurs canonique. Le problème est de prolonger ψ à Y. Il suffit manifestement de le faire pour $\frac{K[z]}{(a_i d_i)}$. Soit S telle que $\overline{S} = \psi(d_i)$ (la classe est prise dans $\frac{K^m[[z]]}{K^m[z]}$). Posons $S' = \frac{S}{d_i}$. Posons $\tilde{\psi}(1) = \overline{S}'$. On voit qu'on prolonge ψ, C.Q.F.D. (on peut toujours mettre par exemple $\psi(E) = 0$).

<u>Remarque</u> :

On utilise le fait que K[[z]] est un corps.

ii) Supposons (Y) espace d'état pour H. Alors nous allons prouver des lemmes :

<u>Lemme 1</u> :

Si sur un anneau principal A on a une tour de modules libres de type fini : $M_1 \subset M_2 \subset M$, si les invariants de M_1 dans M_2 (resp. M_2 dans M) sont $\{\lambda_1,\ldots\lambda_r\}$ (resp. $\{\mu_1,\ldots,\mu_r\}$) certains pouvant être nuls (et alors par convention $0|0$), et si $\{q_1,\ldots,q_r\}$ sont les invariants de M_1 dans M, on a : $\lambda_i | q_i$ (resp. $\mu_i | q_i$).

<u>Preuve</u> :

En représentant des générateurs d'un sous-module comme colonnes d'une matrice de $A^{r\times r}$, on en est réduit à prouver que les invariants $\{q_1,\ldots,q_r\}$ d'un produit $N_1.N_2$, sont divisible respectivement par ceux de N_1 par exemple, soient $\{\lambda_1,\ldots,\lambda_r\}$. Par application du Théorème 2, §.0, on peut supposer que N_1 est diagonale, avec ses invariants sur la diagonale. $N_1.N_2$ est donc déduite de N_2 en multipliant sa $i^{\text{ème}}$ ligne par λ_i. Soit Δ_{k+1} un mineur d'ordre k+1 de N_1N_2. En le développant par rapport à sa dernière ligne on a une expression : $\Delta_{k+1} = \Sigma d_i \Delta_k^i$ où les d_i sont des éléments d'une ligne d'ordre $p \geq k+1$ et les Δ_k^i des mineurs d'ordre k. Si d_k est le p.g.c.d. des mineurs d'odre k de N_1N_2, on voit que $\lambda_{k+1}d_k | \Delta_{k+1}$ donc $\lambda_{k+1}d_k | d_{k+1}$. Ainsi

$\lambda_{k+1} \Big| \dfrac{d_{k+1}}{d_k}$, et d'après le corollaire 1 du Théorème 2, §.0, le lemme est prouvé.

Remarque :

On ne peut pas déduire de ce lemme que $\lambda_i\mu_i | q_i$, et ce dernier résultat est faux comme le montre l'exemple :

$$N_1 = \begin{pmatrix} 4 & 0 \\ 0 & 12 \end{pmatrix} \qquad N_2 = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix} \qquad N_1N_2 = \begin{pmatrix} 24 & 0 \\ 0 & 24 \end{pmatrix}$$

cependant on a le :

Corollaire :

Avec les notations du lemme, si λ_i et μ_i sont premiers entre eux $\lambda_i\mu_i | q_i$. Si en outre N_1 et N_2 sont régulières $\lambda_i\mu_i = q_i$, $\forall i \in \{1,\ldots,r\}$.

Preuve :

Evident par le théorème de Gauss et la multiplicativité du déterminant.

Lemme 2 :

Soit A un anneau principal unitaire, Y un module de torsion de type fini, r, sur A. Soit $a_1,a_2,\ldots,a_r$ une suite d'éléments de A telle que $a_i | a_{i+1}$. Supposons qu'une des deux conditions suivantes soit remplie :

i) Il existe un sous-module $N \subset Y$ tel que si $b_1,b_2,\ldots,b_r$ sont les invariants de N sur A (certains éventuellement égaux à 1), on a $a_i | b_i$.

ii) Il existe un sous-module $N \subset Y$ tel que si $b_1,\ldots,b_r$ sont les invariants (même convention qu'au-dessus) du module quotient $\frac{Y}{N}$, alors on a $a_i | b_i$.

Alors si $d_1,d_2,\ldots,d_r$ sont les invariants de Y sur A, on a $a_i | d_i$.

Preuve :

Supposons ii) remplie :

Représentons Y comme module quotient d'un module libre (cf.§.0) $F = \{e_1\} \oplus \{e_2\} \oplus \ldots \oplus \{e_r\}$, par un sous-module M.

N est alors représentable comme un quotient $\frac{F'}{M}$ avec $F \supset F' \supset M$. On a donc une tour de modules libres $F \supset F' \supset M$. D'après l'isomorphisme du double quotient, $\frac{Y}{N} = \frac{F/M}{F'/M} \approx \frac{F}{F'} \approx \underset{\oplus}{\Sigma} \frac{A}{(b_i)}$, d'après le lemme 1, les invariants de M dans F sont divisibles par les b_i, donc par les a_i d'après l'hypothèse, et ceux sont précisément les invariants de Y sur A.

Supposons i) remplie. Reprenant les mêmes notations que pour ce qui précède, on a une tour $F \supset F' \supset M$, avec $N = \frac{F'}{M} = \Sigma \frac{A}{(b_i)}$. De nouveau l'application du lemme 1 permet de conclure.

Retournons à notre preuve. On peut raisonner par récurrence sur la dimension de Y :

Si cette dimension est minimale, on a fini. Sinon supposons ψ non injective. Soit N son noyau. $\frac{Y}{N}$ est donc espace d'état de dimension inférieure, et il a un jeu d'invariants du type $(\lambda_i a_i)$ d'après l'hypothèse de récurrence (dans cette notation il faut comprendre qu'on a éventuellement "rajouté" des a_i égaux à 1 qui ne contribuent donc pas à la décomposition de X). Appliquons le lemme 2 pour conclure.

Si ϕ est non surjective, l'image N de ϕ est un sous module de Y, qui peut servir d'espace d'état, et dont les invariants, par récurrence sont du type $(a_i \lambda_i)$.

L'application du lemme 2, de nouveau, permet de conclure.

Corollaire :

Etant donné un espace d'état (Y,F) dont le nombre d'invariants $\neq 1$ est $\leq P$, pour H, il existe H et G telles que (H,F,G) soit une réalisation avec ϕ surjective. (Rappelons que H est $m \times p$).

Preuve :

Soit $X = \frac{K^P[z]}{V} = \sum_{\oplus} \frac{K[z]}{(a_i)}$ l'espace canonique. D'après le théorème, on a $Y = \sum_{\oplus i=1}^{p} \frac{K[z]}{(a_i\lambda_i)}$. Appelons h l'application naturelle : $Y \to X$ induite par l'inclusion naturelle $(a_i\lambda_i) \subset (a_i)$. Il s'agit de trouver g rendant commutatif le diagramme :

$$\begin{tikzcd}
K^P[z] \arrow[r] \arrow[dr, "g"'] & X = \frac{K^P[z]}{V} \approx \sum \frac{K[z]}{(a_i)} \\
& Y \arrow[u, "h"']
\end{tikzcd}$$

Or il existe (Théorème 1.§.0) une base de $K^P[z]$, $e_1, \ldots, e_p$, telle que $V = \sum_{\oplus} \{a_i e_i\}$. Posons $g(e_i) = \bar{1}$ dans le $i^{\text{ème}}$ terme de la somme décomposant Y.

Il est clair que ceci définit un homomorphisme g rendant le diagramme commutatif. De plus g est manifestement surjective.

Remarque 1 :

Il peut arriver que (g,h) soit unique à une équivalence près, par exemple si λ_i est premier avec a_p $\forall i$ (par application du théorème du reste chinois). Nous ne savons pas si c'est le cas général.

Remarque 2 :

Si on supprime la contrainte sur le nombre de blocs invariants, on peut encore trouver h et g, mais celle-ci n'est plus nécessairement surjective.

Théorème 3 :

Dans la catégorie des triplets (H,F,G) où ϕ est surjective, et dont les morphismes sont les homomorphismes Θ entre espaces d'état faisant commuter le diagramme :

$$\begin{tikzcd}
 & Y_1 \arrow[dd, "\Theta"] \arrow[dr, "\psi_1"] & \\
K^P[z] \arrow[ur, "\phi_1"] \arrow[dr, "\phi_2"'] & & \dfrac{K^P[[z]]}{K^P[z]} \\
 & Y_2 \arrow[ur, "\psi_2"'] &
\end{tikzcd}$$

la représentation canonique est un objet universel attractif.

Preuve :

Si M est le noyau de ψ, il suffit de considérer la surjection naturelle : $Y \to \frac{Y}{M}$. Manifestement cet homomorphisme est unique car ϕ est surjective, et à une entrée correspond une sortie bien déterminée.

Corollaire :

La représentation canonique est unique à un isomorphisme près.

Remarque 1 :

On est déçu d'être obligé de supposer ϕ surjective. Cependant, si ça n'est pas le cas, Θ, qui existe toujours, n'est plus nécessairement unique.

Remarque 2 :

Tout ce qui précède peut se faire avec ψ injective au lieu de ϕ surjective et en renversant le sens de Θ dans les homomorphismes ci-dessus. C'est une conséquence, par exemple,du § suivant.

PARAGRAPHE 3

Ce paragraphe dit quelques mots du phénomène de dualité, qui est en fait la dualité algébrique élémentaire sur le corps des séries.

3. Dualité.

A) Pour la commodité, abandonnons provisoirement la notation K[[z]] au profit de "S", en rappelant que c'est le corps des séries formelles n'ayant qu'un nombre fini de termes de puissance positive. On note M' la transposée d'une matrice M.

H système dynamique linéaire stationnaire est un endomorphisme : $S^p \to S^m$.

S^p et S^m, espaces vectoriels de dimension finie sur le corps S, peuvent s'identifier, par le biais des bases duales, à leurs duaux respectifs $(S^p)^*$ et $(S^m)^*$.

A H, on associe son transposé ${}^tH$ qui applique $(S^m)^* \to (S^p)^*$. La matrice de ${}^tH$ dans les bases duales de celles prises pour H est la transposée de celle de H : H'.

B) Les résultats suivants sont à peu près évidents :

i) [(H,F,G) réalise H] $\Leftrightarrow$ [(G',F',H') réalise ${}^tH$]

ii) [(Y,F) espace d'état pour H] $\Leftrightarrow$ [(Y, ${}^tF$) espace d'état de ${}^tH$]

iii) [ϕ_H surjective] $\Leftrightarrow$ [$\psi_{{}^tH}$ injective]

iv) [ψ_H injective] $\Leftrightarrow$ [$\phi_{{}^tH}$ surjective]

v) [(H,F,G) minimal pour H] $\Leftrightarrow$ [(${}^tG$, ${}^tF$, ${}^tH$) minimal pour ${}^tH$].

On ne développe pas plus avant ces considérations qui introduisent la notion de système dual largement répandue en automatique et montrent que les diagrammes commutatifs du paragraphe précédent pourraient s'écrire en travaillant sur ψ au lieu de ϕ, et en inversant le sens des flèches.

PARAGRAPHE 4

Dans ce paragraphe on énonce et on prouve des théorèmes de divisibilité élémentaires pour les matrices, qui généralisent ceux connus pour les scalaires, et serviront dans la suite. On s'appuie essentiellement sur le fait que les sous-modules des modules libres sur un anneau principal sont libres aussi, et de rang inférieur (cf. Th.1.§.0).

4. Divisibilité des matrices.

Dans tout ce qui suit A est un anneau principal. Soient M et N deux matrices de $A^{n\times p}$ et $A^{n\times m}$ respectivement. Considérons l'ensemble V des combinaisons MB+NC avec $B \in A^p$ et $C \in A^m$. C'est un sous-module de A^n qui est lui-même un module libre de rang n sur A. Donc V est libre de rang $k \leq n$ (Théorème 1.§.0). Soit $e_1, \ldots, e_k$ une base de V. Soit $D = (e_1, \ldots, e_k)$ la matrice de $A^{n\times k}$ formée des colonnes e_i. Alors on peut écrire :

$$V = D.A^k.$$

De plus avec la contrainte $k = \text{rg } V$, on peut affirmer que D est défini à la multiplication à droite par une matrice inversible dans $A^{k\times k}$ près. En effet si $DA^k = D'A^k \Rightarrow D = D'U'$ et $D' = DU$ avec U et U' dans $A^{k\times k}$. Donc $D = DUU'$. Soit $D(I-UU') = 0$ comme les colonnes de D sont indépendantes sur A (c'est une base de V) on voit que $I-UU' = 0 \Rightarrow u'$ inversible.

Noter qu'on peut écrire $M = DR$ où $R \in A^{k\times p}$ car les colonnes de M sont évidemment dans V. De même $N = D.S$, $S \in A^{k\times m}$. Toute matrice D vérifiant de telles relations est dite diviseur commun à gauche de M et N.

Définition :

D, comme défini dans ce qui précède est appelé plus grand commun diviseur à gauche de M et N. Il est défini à la multiplication à droite par une unité près.

Proposition 1 :

Si E est un diviseur commun à gauche de M et N, alors E divise D à gauche.

Preuve :

Le module $E = EA^r$ (E est supposée $n\times r$) contient V puisqu'il contient les colonnes de M et celles de N. Donc il contient celles de D. Ainsi on a $D = E\Lambda$ pour $\Lambda \in A^{r\times k}$.

Remarque :

Il est un cas dans lequel on peut connaître k à priori, c'est celui où $M \in A^{n\times n}$ (i.e. n=p) est régulière. Dans ce cas le rang de MA^n est n, donc le rang de V qui est supérieur, ne peut qu'être n aussi par conséquent k=n également. Il se trouve que ce cas est le seul que l'on aura à envisager dans ce rapport. On trouvera un algorithme pour déterminer un p.g.c.d. de deux matrices dans [D.V.].

Définition :

M et N seront dites premières entre elles à gauche si $V = A^n$. Ceci est équivalent à dire que D est inversible (donc carrée).

En effet si D est inversible, il est immédiat que $V = A^n$. Réciproquement si $V = A^n$, V contient les colonnes de la matrice identité, et D est donc inversible car $V = DA^k$ et k=n, naturellement.

Proposition 2 :

Si D est un p.g.c.d. à gauche de M et de N, M = DR, N = DS. Alors R et S sont premières entre elles à gauche.

Preuve :

On a $D(RB+SC) = DA^k$ avec $B \in A^p$ et $C \in A^m$. Supposons que $x \in A^k$ soit tel que $x \notin W = \{RB+SC,\ B \in A^p,\ C \in A^m\}$. Comme cependant $Dx \in DW$, $\exists\ w \in W$ tel que $Dx = Dw \Rightarrow D(x-w) = 0$ comme le rang de D est k, $x-w = 0 \Rightarrow x = w$, contradiction.

Théorème 1 :

(Bezout) pour que M et N soient premières entre elles à gauche, il faut et il suffit que pour $B \in A^{p\times n}$ et $C \in A^{m\times n}$ on aie :

$$MB + NC = I \in A^{n\times n}.$$

Preuve :

Si $V = A^n$, cette relation est triviale. Réciproquement si MB+NC = I, multipliant à droite par $x \in A^n$ quelconque, on voit que $V = A^n$.

Remarque :

Tout ce qui est dit à gauche pourrait l'être à droite en transposant tout de manière évidente. On parlera de diviseurs à droite et de matrices premières entre elles à droite.

PARAGRAPHE 5

Dans ce paragraphe, on décrit de façon assez détaillée la représentation de Führmann en tant que lien entre les représentations externes et internes d'un même système. Au paragraphe E), on en déduit un abord assez aisé des problèmes de représentation canonique évoqués dans [KFA] par exemple. Ces problèmes sont décrits plus exhaustivement sous un angle calculatoire dans [HE]. Pour plus de détails et des applications voir [FU] et [AN]. On a essayé d'introduire le point de vue modulaire quand on pensait pouvoir le faire sans complication excessive. Les preuves qui en découlent, gagnent pense-t-on, en simplicité.

5. La représentation de Führmann.

A) La représentation de Führmann est une construction qui utilise une forme factorisée de H du type $ND^{-1}N_1$ où N,D,N_1 sont des matrices polynomiales de dimensions $(m\times r)$, $(r\times r)$, $(r\times p)$ respectivement ; à partir de cette forme elle exhibe un triplet (H,F,G) en termes d'opérations arithmétiques élémentaires sur les matrices N,D,N_1. Autrement dit, elle permet de relier les éléments de H (par le biais d'une factorisation) aux matrices (H,F,G) d'une réalisation. Si dans le cas scalaire cette lecture est si triviale, c'est entre autres parce qu'une fraction rationelle se présente sous une forme toute factorisée (!). Remarquons enfin que toute matrice rationelle se factorise sous une telle forme où on peut même choisir N ou N_1 = identité. En effet, il suffit de choisir un p.g.c.d,p, pour les dénominateurs des éléments de H, et alors on a $H = (pI)^{-1}M = M(pI)^{-1}$ où M est polynomiale de dimension $m\times p$.

Il est une autre remarque dont on fera usage : si N et D ont un facteur commun à droite, i.e. :

$$\left.\begin{aligned} N &= N'R \\ D &= D'R \end{aligned}\right\} , \qquad \text{on a} \quad ND^{-1}N_1 = N'RR^{-1}D'^{-1}N_1 = N'D'^{-1}N_1$$

La même remarque vaut pour les facteurs à gauche de D et N_1. Ainsi on peut toujours supposer que (N,D), et (D,N_1) sont premières entre elles à droite et à gauche respectivement. Ceci est l'analogue du cas scalaire où il semble naturel de simplifier les facteurs commun du numérateur et du dénominateur.

B) Décrivons la représentation annoncée.

Tout d'abord soit $S = \sum_{i>k}^{\infty} a_i z^{-i}$ une série formelle. On appelle Π l'opérateur qui "ne garde que la partie propre", i.e. $\Pi(S) = \sum_{i=1}^{\infty} a_i z^{-i}$.

On ne précise pas la nature des a_i, car elle n'intervient pas : l'opération s'effectuera aussi bien si les a_i sont des scalaires, des vecteurs , des matrices, etc...

Prouvons un lemme qui fera mieux comprendre le rôle de Π :

Lemme 1 :

Soit D une matrice carrée polynomiale régulière de dimension $(r\times r)$. Alors $M = \{Dy,\ y \in K^r[z]\}$ est un sous module de $K^r[z]$. Si on pose $\overline{M} = \frac{K^r[z]}{M}$ (le module quotient), il existe dans chacune des classes d'équivalence qui composent $\overline{M}$ un unique représentant x tel que $D^{-1}x$ soit une série propre. Il s'obtient à partir d'un élément quelconque y de la classe en lui appliquant l'opérateur $\Pi_D : K^r[z] \to K^r[z] : \Pi_D(y) = D\ \Pi\ D^{-1}y$.

Preuve :

On peut écrire $D^{-1}y = P+S$ où P est la partie polynomiale, S la partie propre. Alors par définition, $\Pi_D(y) = DS$. Donc $\Pi_D(y)-y = D(S-D^{-1}y) = -DP$. Comme P est polynomial, ceci montre que $\Pi_D(y)$ et y appartiennent à la même classe modulo M. Par ailleurs $D^{-1}\Pi_D(y) = \Pi D^{-1}y$ est propre par définition de Π. Enfin si x et x' sont dans une même classe et vérifient tous deux $\left.\begin{matrix} D^{-1}x \\ D^{-1}x' \end{matrix}\right\}$ propre, on a $D^{-1}(x-x')$ propre aussi. Or $x-x' = Dz$ où $z \in K^r[z]$ puisqu'ils sont dans la même classe. Donc z propre $\Rightarrow z=0$. Ceci prouve le lemme.

Remarque :

Ce lemme dans le cas scalaire, dit que l'on peut dans $\frac{K[z]}{(P)}$ sélectionner un représentant unique de degré inférieur à celui de P, car c'est le sens de la condition $\frac{x}{p}$ propre ! On le sait bien, c'est le reste de la division euclidienne.

Dès lors il est possible de fournir un support concret au module quotient $\overline{M}$: on identifie chaque classe à son unique représentant x tel que $D^{-1}x$ soit propre et on obtient un ensemble de vecteurs de $K^r[z]$ où l'addition est l'addition de polynômes composantes à composantes puisque si $\left.\begin{matrix} D^{-1}x \\ D^{-1}x' \end{matrix}\right\}$ sont propres, $D^{-1}(x+x')$ l'est aussi.

La multiplication par un polynôme p ne s'effectue pas ainsi car $D^{-1}px = pD^{-1}x$ peut cesser d'être propre même si $D^{-1}x$ l'est (par exemple si $p = \text{dét}(D)$). Par définition de la multiplication dans $\overline{M}$, $p\bar{x} = \overline{px}$. Donc pour que notre identification en soit une, il faut définir $p.x = \Pi_D(px)$. (. est le symbole adopté pour la multiplication dans $\overline{M}$).

L'espace d'état dans la représentation de Führmann sera précisément $\overline{M}$, identifié par la construction précédente à l'ensemble des éléments x de $K^r[z]$ tels que $D^{-1}x$ soit propre, muni des opérations définies ci-dessus :

i) addition normale.
ii) $p.x = \Pi_D(px)$.

Puisque $\overline{M}$ est notre espace d'état avec sa structure de module, cela signifie que l'action de F s'identifie dans $\overline{M}$ avec la multiplication par z :

$$F : \overline{M} \to \overline{M} \qquad x \to z.x$$

décrivons les applications G et H :

$$G : K^p \to \overline{M} \quad : \quad G(u) = \Pi_D N_1 u$$

$H : \overline{M} \to K^m \quad : \quad H(x) = (ND^{-1}x)_1$ (rappelons que cette notation signifie qu'on sélectionne le coefficient du terme en "z^{-1}" dans la série $ND^{-1}x$ dont les coefficients sont dans K^m).

Théorème 1 :

Le triplet (H,F,G) décrit ci-dessus et associé à $\mathcal{H} = ND^{-1}N_1$ est une réalisation de $\mathcal{H}$.

Preuve :

Reprenant l'argument utilisé au théorème 1 du paragraphe 1, il suffit de montrer que $H(zI-F)^{-1}G$ et $\mathcal{H}$ donnent même image de tout vecteur scalaire $u \in K^p$. On a $H(zI-F)^{-1}Gu = \sum_{t=1}^{\infty} HF^{t-1}Guz^{-t}$.

$(HF^{t-1}Gu) = (HF^{t-1}\Pi_D N_1 u) = (H\Pi_D(z^{t-1}\Pi_D N_1 u)) = (H\Pi_D z^{t-1}N_1 u) = (ND^{-1}\Pi_D z^{t-1}N_1 u)_1 =$

$(N\Pi D^{-1}N_1 uz^{t-1})_1$. Par hypothèse on sait que $\mathcal{H}$ est propre, donc si on décompose $D^{-1}N_1 = P+S$ où P est polynomiale et S propre, on a :

$ND^{-1}N_1 = NP + NS \Rightarrow ND^{-1}N_1 = \Pi(ND^{-1}N_1) = \Pi NS.$

$= \Pi N \Pi D^{-1}N_1.$

Donc $(N\Pi D^{-1}N_1 uz^{t-1})_1 = (\Pi N\Pi D^{-1}N_1 uz^{t-1})_1 = (ND^{-1}N_1 uz^{t-1})_1 = (\mathcal{H}u.z^{t-1})_1 = (\mathcal{H}u)_t$

Le résultat en découle donc.

C) On peut se demander dans quels cas la représentation de Führmann est minimale. Un bref examen du cas scalaire permet de mieux comprendre la situation.

Dans le cas $\mathcal{H} = \frac{P}{Q}$ où P,Q sont des polynômes premiers entre eux, tous les automaticiens savent que la matrice F d'un triplet minimal est cyclique, de polynôme minimal Q. Ceci concorde avec les résultats du paragraphe 1 car si P et Q sont premiers entre eux, l'ensemble $\mathcal{V}$ des polynômes u tels que $\mathcal{H}$.u soit polynomial est celui des multiples de Q, donc l'espace d'état d'une représentation canonique est isomorphe en tant que module à $\frac{k[z]}{(Q)}$. On voit donc qu'il est cyclique, de polynôme minimal Q.

Si P et Q ne sont pas premiers entre eux, il n'en va plus ainsi, et la représentation de Führmann associée qui a un espace d'état isomorphe à $\frac{k[z]}{(Q)}$, donc de dimension d°- Q sur K n'est pas minimale puisque l'on peut simplifier un facteur, donc faire baisser le degré du dénominateur. Dans ces conditions il est possible de conjecturer le

<u>Théorème 2</u> :

Le triplet de Führmann (H,F,G) est commandable si et seulement si D et N_1 sont premières entre elles à gauche, et observable si et seulement si N et D sont premières entre elles à droite.

<u>Preuve</u> :

i) <u>Commandabilité</u> :

Soit P une entrée polynomiale. Elle amène le système en l'état $\phi(P) = \Pi_D N_1 P$ d'après la définition de H et F (rappelons que $\phi(\sum_{i=0}^{\ell} a_i z^i) = (\sum_{i=0}^{\ell} F^i G a_i))$ autrement dit les classes atteintes dans $\overline{M}$ sont celles des éléments multiples de $N_1 = \{N_1 P,\ P \in K^p[z]\}$.

Prouver que ϕ est surjective, c'est exactement prouver que toute classe d'équivalence dans $\overline{M}$ a un représentant du type N_1P avec $P \in K^p[z]$. Cela signifie donc que tout élément de $k^r[z]$ peut s'écrire $N_1P + DQ$ avec $Q \in K^r[z]$, c'est donc dire exactement que le module engendré à droite par (N_1 et D) sur ($k^p[z] \times k^r[z]$) est $k^r[z]$ en entier. C'est dire exactement que N_1 et D sont premières entre elles à gauche.

ii) Observabilité :

Il suffit d'invoquer la dualité et d'appliquer ce qui précède. Donnons néanmoins une preuve directe.

Soit $\bar{x} \in \overline{M}$. Evaluons $\psi(\bar{x})$ (rappelons que par définition $\psi(\bar{x}) = \sum_{i=1}^{\infty} (HF^{i-1}x)z^{-i}$). Donc la définition de (H,F,G) impose :

$$\psi(\bar{x}) = \sum_{i=1}^{\infty}(ND^{-1}\Pi_D z^{i-1}x)_1 z^{-i}.$$

Supposons $\psi(\bar{x}) = 0$, cela entraîne $(ND^{-1}\Pi_D z^{i-1}x)_1 = 0 \; \forall i$. Observons alors que:

$$(ND^{-1}\Pi_D z^{i-1}x)_1 = (ND^{-1}z^{i-1}x)_1 = 0.$$

En effet, $z^{i-1}x$ et $\Pi_D z^{i-1}x$ diffèrent d'un multiple de D, donc $ND^{-1}(z^{i-1}x-\Pi_D z^{i-1}x)$ est polynomial. Ainsi $\psi(x) = 0$ est équivalent à $(ND^{-1}z^{i-1}x)_1 = 0 \; \forall i \geq 1$, donc à $ND^{-1}x$ polynomial. (En donnant à i toutes les valeurs, on voit que toute la partie propre de $ND^{-1}x$ doit être nulle). Ainsi dire que ψ est injective, c'est dire l'équivalence dans $k^r[z]$:($ND^{-1}x$ polynomial) $\Leftrightarrow$ ($D^{-1}x$ polynomial), (en effet $\bar{x} = 0$ dans $\overline{M}$ est équivalent à $D^{-1}x$ polynomial pour $x \in k^r[z]$). Il nous reste à prouver que cette relation ne peut tenir que si N et D sont premières entre elles à droite.

i) Supposons qu'il n'en soit pas ainsi : $\left.\begin{matrix} N = N'R \\ D = D'R \end{matrix}\right\}$ avec dét R non constant (sinon R est unimodulaire) soit $y \in k^r[z]$, alors si on pose $x = D'y$, on a $x \in K^r[z]$ et $ND^{-1}x = N'RR^{-1}D'^{-1}D'y = N'y \in K^r[z]$. Mais $D'k^r[z]$ est un sous module de $K^r[z]$ qui contient $Dk^r[z]$ (puisque $D = D'R$) mais en est différent : en effet les dimensions des modules quotients $\frac{K^r[z]}{D'K^n[z]}$ et $\frac{K^r[z]}{DK^r[z]}$ en tant qu'espaces vectoriels sur K sont respectivement d° dét D' et d° dét D qui diffèrent car d° dét R > 0. En prenant $x \in D'K^n[z]$, $x \notin DK^n[z]$, on a un contrexemple à l'équivalence annoncée.

ii) Si N et D sont premières entre elles à droite, alors il existe des matrices polynomiales A et B de dimensions convenables telles que $AD + BN = I$. Il s'ensuit $A + BND^{-1} = D^{-1}$, et si x est tel que $ND^{-1}x$ est polynomial, alors $D^{-1}x$ l'est visiblement aussi. Ceci achève la preuve.

Remarque 1 :

L'introduction du point de vue le moins matriciel et le plus modulaire possible simplifie considérablement les preuves par rapport à celles se trouvant dans [AN].

Remarque 2 :

Si on a comme objectif de prouver la condition de minimalité, on peut procéder ainsi :

i) Si les conditions de "premières entre elles" à gauche et à droite ne sont pas satisfaites, on peut simplifier des facteurs et avoir une factorisation $N'D'^{-1}N_1'$ avec deg dét $D' <$ d° dét D. Donc on aurait une représentation de Führmann dont la dimension d'état serait plus petite

(rappelons que dim $\dfrac{K^r[z]}{DK^r[z]}$ = d° dét D , où la dimension doit s'entendre comme espace vectoriel sur K) donc la représentation initiale ne pouvait pas être minimale.

ii) Si ces conditions à droite et à gauche sont satisfaites, alors la relation de Bezout : $AD + BN = I \Leftrightarrow A + BND^{-1} = D^{-1} \Rightarrow AN_1 + BND^{-1}N_1 = D^{-1}N_1$, et donc $(ND^{-1}N_1)x = \mathcal{H}x$ ne peut être polynomiale que si $D^{-1}N_1x$ l'est. Ainsi il résulte des considérations du paragraphe 1 que la dimension de la réalisation canonique de $\mathcal{H} = (ND^{-1}N_1)$ et de $\mathcal{H} = D^{-1}N_1$ est la même. C'est donc aussi celle de $\mathcal{H}_1' = N_1D'^{-1}$. Comme N_1 et D sont premières entre elles à gauche, N_1' et D' le sont à droite. Ainsi cette dimension minimale est donc en fin de compte celle de $\mathcal{H}_2 = D'^{-1}$ ou encore de $\mathcal{H}_2' = D^{-1}$. Son espace d'état canonique est bien sûr $\dfrac{K^r[z]}{DK^r[z]}$, c'est-à-dire exactement celui de la représentation de Führmann associée à (N,D,N_1).

<u>Corollaire 1</u> :

Si on dispose d'une factorisation $H = ND^{-1}N_1$ où :

i) N et D sont premières entre elles à droite.

ii) N_1 et D_1 sont premières entre elles à gauche.

Alors :

- La dimension d'une réalisation canonique de H est d° dét D.

- Les polynômes invariants de toutes les matrices F d'un triplet canonique (H,F,G) sont les polynômes invariants de la matrice D.

<u>Preuve</u> :

La première assertion découle du théorème 2 et du corollaire 2 du §0, puisque l'on peut choisir comme espace d'état minimal celui de la représentation de Führmann : $\dfrac{K^r[z]}{DK^r[z]}$.

La seconde découle du théorème 2 du §0, toujours en examinant la représentation de Führmann, et en utilisant le fait que toutes les matrices F des triplets minimaux ont même polynômes invariants, ce qui n'est qu'une autre façon de dire que tous les espaces d'états minimaux sont isomorphes en tant que modules (cf.§2).

D) Cas particulier. On remarque que dans la factorisation d'une matrice H selon la remarque figurant dans A) sous la forme $(dI)^{-1}M$, on n'a à considérer que le cas N=I à priori, (ou bien $N_1=I$ si on fait passer $(dI)^{-1}$ à droite de M), mais il peut être intéressant, si on dispose de la forme $D^{-1}N_1$, (resp ND^{-1}) de factoriser D sous la forme $U^{-1}.\mathrm{diag}\{a_i\}V^{-1}$ où U,V sont unimodulaires et les a_i sont les polynômes invariants de D (cf.§0). On peut alors écrire $H = (V)\ \mathrm{diag}\ \{\frac{1}{a_i}\}\ (U\ N_1)$ et comme V et U sont polynomiales on a $H = N^0\ \mathrm{diag}\ \{\frac{1}{a_i}\}\ N_1^0$. Il importe de noter que si D était première à gauche avec N_1, $\mathrm{diag}\{a_i\}$ et N_1 le sont aussi car par Bezout $DA + N_1B = I$, ou encore :

$$UDV(V^{-1}A) + UN_1B = U \Rightarrow (UDV)(V^{-1}AU^{-1}) + (UN_1)(BU^{-1}) = I$$

ce qui, joint au fait que $UDV = \mathrm{diag}\{a_i\}$ prouve l'affirmation. De manière analogue, N^0 et $\mathrm{diag}\{a_i\}$ sont premières entre elles à droite. (Encore plus trivial car N' est carrée unimodulaire, ses diviseurs réguliers sont unimodulaires). Rebaptisons N^0 et N_1^0 en N et N_1.

Ainsi, sous la nouvelle forme où nous l'avons mise, la factorisation $H = N \ \mathrm{diag}\ \{\frac{1}{a_i}\}\ N_1$ est associée à une représentation de Führmann minimale

et on voit que l'espace d'état $\overline{M} = \dfrac{K^r[z]}{\mathrm{diag}\{a_i\}K^r[z]}$ a la structure de module :

$\sum\limits_{\oplus} \frac{K[z]}{(a_i)}$ et donc que, conformément au corollaire 1 : $\dim \overline{M} = \sum\limits_i d° a_i$

(dimension comme e.v. sur K) dans une base convenable, F s'écrit :

$$\begin{pmatrix} B_1 & & 0 \\ & B_2 & \\ 0 & & \ddots \end{pmatrix} \quad \text{avec} \quad B_i = \begin{pmatrix} 0 & 0 & \cdots & a_i^0 \\ 1 & 0 & \cdots & \vdots \\ & 1 & \ddots & \vdots \\ 0 & & \ddots & a_i^{n_i-1} \\ & & 1 & a_i \end{pmatrix} \quad \text{(forme compagnon)}$$

où on pose $n_i = d° a_i$ et $a_i = z^{n_i} + \sum\limits_{k=n_i-1}^{0} a_i^k z_i^k$

Une telle base est par exemple $\mathcal{B} = \{e_{ij}\}_{\substack{1 \leq i \leq r \\ 0 \leq j < n_i}}$ avec :

$$e_{ij} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ z^j \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i^{\text{ème}} \text{ ligne}$$

dans la suite, e_{ij} désignera toujours un tel vecteur.

On peut, pour se faire une idée, décrire comment obtenir H et G, si on munit $\overline{M}$ de la base ci-dessus et K^p et K^m de leurs bases canoniques. Il suffit de se reporter à la définition de la réalisation de Führmann donnée dans ce paragraphe pour constater que :

$\underline{G}$: s'obtient colonne par colonne en réduisant le $i^{\text{ème}}$ élément de la colonne j de $\overline{N}_1$ considérée modulo a_i

(c'est l'opération $\Pi_D\, e_{io}$)

et en écrivant à partir de ce vecteur $\begin{pmatrix} b_0^0 + b_0^1 z + \dots + b_0^{n_1-1} z^{n_1-1} \\ \hline b_r^0 + b_r^1 z + \dots + b_r^{n_r-1} z^{n_r-1} \end{pmatrix}$

la $j^{\text{ème}}$ colonne de G : $\begin{pmatrix} b_0^0 \\ b_0^1 \\ b_0^{n_1-1} \\ \vdots \\ b_{n_r}^0 \\ b_{n_r}^1 \\ \vdots \\ b_n^{n_r-1} \end{pmatrix}$

$\underline{H}$: s'obtient colonne par colonne en considérant la i^e colonne de N, par exemple, en l'appelant h_i, et en considérant le vecteur $\frac{h_i}{a_i}$, dont on écrit le développement en série :

$$\frac{h_i}{a_i} = \Sigma\, h_i^k z^{-k}.$$

Les vecteurs $\{h_i^1, h_i^2, \dots, h_i^{n_i}\}$ fournissent alors n_i colonnes pour H.

Remarque 1 :

Les bases choisies trivialisent la forme de F, mais non celle de H ni celle de G. Il ne s'agit pas de la forme canonique présentée dans [KFA] On examinera ce point plus en détail au sous-paragraphe E).

Remarque 2 :

Il est tout-à-fait possible que certains a_i soient constants. Dans ce cas il n'y a pas de e_{ij} correspondants. Par exemple dans le cas où p=1, ou m=1 (i.e. $\mathcal{H}$ est "scalaire" d'un côté). On peut écrire $\mathcal{H} = \frac{N}{d}$ ou $\frac{N_1}{d}$ ou d est un polynôme premier avec l'ensemble des éléments de N ou N_1. Notre D est alors scalaire (c'est "d") et il n'y a qu'un invariant évidemment : lui-même. Cependant, on peut vouloir, même dans ces cas, écrire une factorisation avec D matriciel, par exemple $\mathcal{H} = (dI)^{-1}N_1$ dans le cas p=1 avec N_1 vecteur colonne. Nos manipulations pour rendre dI et N_1 premières entre elles vont alors nous donner une matrice D, qui est telle que $\frac{K^r[z]}{DK[z]}$ est une représentation minimale de l'espace d'état. Par conséquent D n'a qu'un invariant et c'est le d précédent.

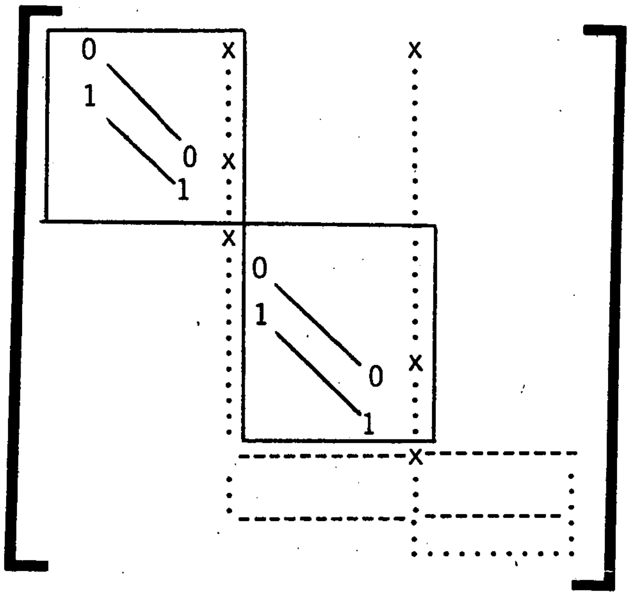
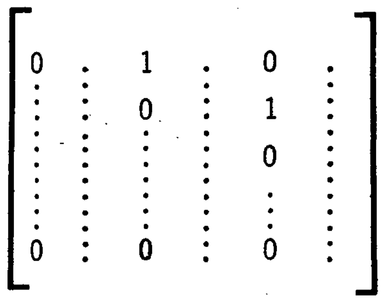
Remarque 3 :

Le cas où il n'y a qu'un invariant est précisément celui où l'espace d'état est cyclique. En particulier c'est ce qui se produit si tous les pôles sont simples comme le montrent immédiatement les conditions de divisibilité sur les invariants $a_i | a_{i+1}$, (a_i divise a_{i+1}). Auquel cas bien sûr , F est diagonalisable.

E) Dans ce sous-paragraphe, on va expliciter un peu les liens entre les formes canoniques des systèmes multivariables évoquées par exemple dans [KFA] et les propriétés des facteurs d'une expression $\mathcal{H} = ND^{-1}$ ou $D^{-1}N_1$. La raison pour laquelle on travaille avec N_1 ou N = Id sur cette question doit être mise en rapport avec le fait que dans ces formes, on se préoccupe de trivialiser les formes de (G,F) ou de (F,H), pris par paires.

En prenant par exemple N = Id, on voit, en considérant la définition du triplet de Führmann, qu'on concentre dans D l'information simultanément relative à F et H. On peut donc conjecturer que ceci sera bien adapté à la recherche d'une forme canonique "observable" pour $\mathcal{H}$.

Prenons donc $H = D^{-1}N_1$, où les hypothèses désormais habituelles de "premières entre elles à gauche" concernant D et N_1 sont supposées satisfaites. On cherche à obtenir une forme canonique observable, c'est-à-dire exhiber une base de l'espace d'état, et une de l'espace de sortie dans lesquelles les matrices F et H s'écrivent :

F =  H = 

où le nombre de "1" dans H est égal au nombre de blocs dans F, et où l'emplacement de ces "1" en colonne coïncide avec celui des colonnes marquées "x" dans F.

On va procéder en deux temps : tout d'abord mettre en évidence une base $\mathcal{B}$ de l'espace d'état qui mette F sous cette forme. Ensuite on examinera la matrice H obtenue et on montrera qu'on peut, moyennant des combinaisons de lignes (changement de base dans l'espace de sortie) la mettre sous la forme voulue.

1) Choix de $\mathcal{B}$.

Il s'agit de trouver k vecteurs x_i dans l'espace d'état $\overline{M}$ tels que les vecteurs $\{z^\ell x_i\}$ $0 \leq \ell < d_i$ soient une base de $\overline{M}$. Une condition nécessaire est bien sûr Σd_i = d° dét D. Posons Δ = dét D.

Afin d'exhiber de tels vecteurs, il est possible de choisir D sous une forme particulière. Choisir, en l'occurence, signifie remplacer D par une matrice du type U.D. où U est unimodulaire : en procédant ainsi, on conserve N = Id. Introduisons à cet effet une nouvelle notion.

Définition 1 :

Une matrice polynomiale D régulière étant donnée, appelons ℓ_i les lignes de D_i appelons degré de ℓ_i le sup,d_i,des degrés intervenant dans ℓ_i. La matrice est dite "ligne-réduite" si $\Sigma d_i = d°$ dét D. Une définition similaire en colonnes donne lieu à une forme "colonne réduite", équivalente à dire que ${}^tD$ est ligne réduite .

Remarque 1 :

On a toujours d° dét D $\leq \Sigma d_i$: il suffit de penser à la façon dont on calcule le déterminant.

Proposition 1 :

Si on appelle D_ℓ la matrice scalaire des coefficients des éléments de plus haut degré dans chaque ligne, dire que D est ligne réduite, c'est exactement dire que D_ℓ est régulière.

Preuve :

Si D_ℓ est régulière, comme le déterminant de D, noté Δ, contient le monôme $z^{\Sigma d_i}$ dét $D\ell$, et que les autres monômes sont de degré inférieur par définition des d_i, on a $d° \Delta \geq \Sigma d_i$, d'où l'égalité.

Si $d° \Delta = \Sigma d_i$, il est impossible, pour la même raison, d'avoir dét $D\ell = 0$.

Proposition 2 :

Soit M un sous module de $K^r[z]$. D'après le (th.1, §0), il est libre, de rang $\leq r$. Il existe une base de M telle que si on appelle N la matrice dont les colonnes sont formées avec les vecteurs de cette base, la matrice des coefficients de ses plus hauts degrés en colonne est de rang plein. On dira de façon abrégée : matrice des plus hauts degrés en colonne.

Preuve :

Par récurrence sur le rang k de M : pour k=1, c'est trivial. Supposons la proposition vraie jusqu'à k-1. Soit $e_1,\dots,e_k$ une base de M. On peut toujours supposer les degrés en colonnes ordonnées : $d° e_i \leq d° e_{i+1}$.

Soit $M' = \{e_1,\ldots,e_{k-1}\}$. On peut supposer d'après l'hypothèse de récurrence que la matrice de plus hauts degrés de $[e_1,\ldots,e_{k-1}]$ est de rang plein. Si celle de $e_1,\ldots,e_{k-1},e_k$ ne l'est pas, on peut en remplaçant e_k par $e_k - \sum_{i=1}^{k-1} p_i e_i$ où les p_i sont des polynômes (le d° de e_k est maximal), faire décroître le degré de e_k. Comme le degré est entier, on peut en répétant l'opération arriver à une base comme dans la proposition.

Corollaire :

Pour toute matrice D régulière, il existe une matrice unimodulaire U telle que DU soit colonne-réduite. De même, il existe V telle que VD soit ligne réduite.

Preuve :

Il suffit de prouver la première affirmation, car on passe à l'autre en transposant.

Si on considère le module M engendré par les colonnes de D, on applique la prop.2 pour trouver une matrice D' colonne réduite dont les colonnes engendrent aussi M. On a donc DU = D' où U est unimodulaire, d'où le résultat.

On a donc montré que l'on peut supposer, dans une factorisation $H = D^{-1}N_1$, que D est ligne réduite. C'est là la forme particulière dont nous parlions plus haut.

Passons à présent à la construction de B. Nous aurons besoin d'un lemme :

Lemme :

Si D ($r\times r$) est ligne réduite et si $\tilde{D}$ est la matrice transposée des cofacteurs de D(i.e. $\tilde{D} = \Delta . D^{-1}$) $\tilde{D}$ est colonne-réduite. Si on pose n = d° Δ et si on note $d_1,d_2,\ldots,d_r$ le degré des lignes de D (toujours supposée ligne-réduite), le degré des colonnes de $\tilde{D}$ est $(n-d_1)$, $(n-d_2)$, ..., $(n-d_r)$.

Preuve :

Rappelons tout d'abord que :

$$D.\tilde{D} = \Delta I \quad (1)$$

où I est la matrice identité de dimension $r \times r$.

Appelons D_ℓ la matrice (scalaire) des plus hauts degrés en ligne de D et $\tilde{D}_\ell$ celle des plus hauts degrés en colonne de $\tilde{D}$; appelons respectivement d_i et $\tilde{d}_j$ ces plus hauts degrés en ligne et en colonne. Enfin, posons :

$$M = D_\ell.\tilde{D}_\ell = (M_{ij}).$$

Si $M_{ij} \neq 0$, le terme de plus haut degré de l'élément (i,j) du produit $D\tilde{D}$ est exactement $d_i+\tilde{d}_j$.

La première colonne $\tilde{C}_1$ de $\tilde{D}_\ell$ est non nulle par définition ($\tilde{D}$ est régulière). Donc, comme D_ℓ est régulière, on a :

$D_\ell\tilde{C}_1 \neq 0$. Or $D_\ell\tilde{C}_1 = \begin{pmatrix} M_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ d'après (1). Donc $M_{11} \neq 0$. De même $M_{ii} \neq 0 \quad \forall i$,

et $M = \text{diag}\{M_{ii}\}$.

Alors, on a toujours d'après (1) :

$$d_i + \tilde{d}_i = n \Rightarrow \tilde{d}_i = n - d_i$$

qui constitue une des affirmations de l'énoncé. La deuxième vient en faisant la somme :

$$\Sigma\tilde{d}_i = n.r - \Sigma d_i$$

Comme $\Sigma d_i = n$ par hypothèse, on a :

$$\Sigma\tilde{d}_i = n(r-1)$$

et comme on a : dét $\tilde{D} = \Delta^{r-1}$, le résultat s'en suit i.e. : $\tilde{D}$ est colonne-réduite.

Ces définitions et résultats étant établis, il nous est possible de choisir B. Rappelons que par définition :

$$e_{ij} = \begin{pmatrix} 0 \\ \vdots \\ z^i \\ 0 \\ 0 \end{pmatrix} \leftarrow i^{\text{ème}} \text{ ligne}$$

Appelons x_i le $i^{\text{ème}}$ vecteur colonne de $\tilde{D}$. Considérons le jeu de vecteurs $e_{i\ell}$ avec $0 \le \ell < d_i$, où les d_i sont les degrés des lignes de D, comme dans les lemmes précédents. Comptons les : il y en a $\Sigma d_i = n = d° \Delta$, donc autant que la dimension de l'espace d'état $\overline{M}$. Ils sont tous représentants de leur classe dans $\overline{M}$. En effet on a :

$$D^{-1} e_{i\ell} = \frac{z^\ell}{\Delta} x_i \text{ qui est propre d'après le lemme car } \ell < d_i.$$

Ainsi ils sont indépendants si et seulement si on a :

$$\Sigma \lambda_i^\ell e_{i\ell} = 0 \Rightarrow \lambda_i^\ell = 0 \ \forall_i^\ell.$$

Mais à cause de la forme de e_i, ceci revient à dire que si $P_i(z) e_i = 0$, nécessairement alors $P_i = 0$. C'est une évidence. Ainsi si on pose $B = \{e_{i\ell}\}$ on a une base de l'espace d'état. Il est trivial que dans cette base, la matrice F a la forme annoncée. Ceci n'est pas exceptionnel : la base exhibée en D) le réalise aussi. Mais la base B a un comportement particulier relativement à H : en effet on voit que par construction, $D^{-1} e_{i\ell}$ est propre pour $\ell \le d_i - 1$ donc $D^{-1} e_{i\ell}$ est de degré au plus "-2" pour $\ell < d_i - 1$. Autrement dit $(D^{-1} e_{i\ell})_1 = 0$ pour $\ell < d_i - 1$, ce qui montre que la matrice H dans les bases B et canonique sur K^m a la forme :

$$\left[\begin{array}{cccc|cccc} & & & x & & & & x \\ & & & x & & & & x \\ 0 & 0 & \dots & \vdots & 0 & 0 & \dots & \vdots \;\dots \\ & & & \vdots & & & & \vdots \\ & & & x & & & & x \end{array}\right] \qquad (2)$$

$$\underbrace{\qquad\qquad}_{d_1} \quad \underbrace{\qquad\qquad}_{d_2}$$

2) Forme de H.

Pour pouvoir mettre H sous la forme annoncée, il suffit donc de prouver que le rang des lignes vaut k qui est par définition le "nombre de blocs" dans F

Remarquer que ce nombre k est en général distinct du nombre r de colonnes, car certains des d_i peuvent être nuls. En effet si ce rang est k on peut quitte à permuter des lignes supposer que les k premières sont indépendantes. La multiplication à gauche par une matrice convenable fournira alors une forme :

$$\left[\begin{array}{cccccc} & 1 & & 0 & & 0 \\ & 0 & & 1 & & \\ & 0 & & 0 & & \\ 0\cdots & \vdots & 0\cdots & x & \cdots & 1 \\ & x & & x & & x \\ & \vdots & & \vdots & & \vdots \\ & x & & x & & x \end{array}\right]\begin{array}{l}\left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} k. \\ \\ \\ \\ \end{array}$$

Alors par combinaisons de lignes adéquates, on parviendra à la forme voulue.

Si les colonnes de $\begin{matrix}\text{"x"}\\ x\\ \vdots\end{matrix}$ de H dans (2) ne sont pas indépendantes sur K, $\exists$ des λ_i non tous nuls tels que $\Sigma\lambda_i\begin{bmatrix}x\\ x\\ \vdots\\ \vdots\end{bmatrix}_i = 0$. Un instant de réflexion montre que ceci signifie exactement : $\exists(\lambda_i)$ non tous nuls, tels que : $\Sigma\lambda_i z^{d_i-1}\frac{x_i}{\Delta}$ soit de degré au plus "-2" (i.e. le terme en z^{-1} disparaît par le jeu de la combinaison linéaire). Donc on a $\Sigma\lambda_i z^{d_i}\frac{x_i}{\Delta}$ propre, ou encore $\Sigma\lambda_i z^{d_i}D^{-1}e_{io}$ propre. Donc $D^{-1}(\Sigma\lambda_i z^{d_i}e_{io})$qui est propre. Ainsi le vecteur $x = \Sigma\lambda_i z^{d_i}e_{io}$ est son propre représentant dans $\overline{M}$. Il est donc combinaison scalaire des $e_{i\ell}$ pour $\ell<d_i$, ce qui est manifestement absurde (car jamais les termes de degré d_i ne pourront être "compensés"). Ceci achève de montrer que c'est bien la forme canonique décrite dans [KFA]que nous obtenons, ou plus exactement sa forme duale. Pour obtenir la forme "commandable", il suffit évidemment de tout transposer. On en déduit le :

Théorème 3 :

Si $H = D^{-1}N_1$ où D et N_1 sont premières entre elles à gauche, et D sous la forme ligne réduite, les degrés d_i des lignes de D sont les indices d'observabilité du système H. En particulier, ils ne dépendent pas de la matrice D vérifiant ces conditions (cf.[KFA]). Si $H = ND^{-1}$ où N et D sont premières entre elles à gauche et D sous la forme colonne réduite, les degrés K_i des colonnes de D sont les indices de commandabilité du système H.

Remarque 1 :

Ceci illustre l'importance accordée dans la littérature [H H], [AN], [HE] au caractère réduit (en ligne ou en colonne) des matrices. On avait déjà rencontré des formes colonnes réduites dans le théorème 2, §.0, où on écrit :

$$D = U \begin{pmatrix} d_{11} & & & \\ & d_{22} & & \\ 0 & & \ddots & \\ & & & d_{mn} \end{pmatrix} \text{avec } d° \, d_{ij} < d° \, d_{jj}.$$

La matrice triangulaire vérifiant ces conditions est évidemment colonne réduite, mais cette réduction est obtenue par multiplication à gauche par U, et ne nous est donc d'aucun secours pour cette question...

Remarque 2 :

L'introduction de la représentation de Führmann et de $\tilde{D}$ introduit un raccourci pour la preuve de ce résultat (on trouve dans [HE] une preuve basée sur des manipulations matricielles uniquement).

Remarque 3 :

On a utilisé le résultat selon lequel les K_i (ou les d_i) ne dépendent que de H comme conséquence d'un résultat de théorie des systèmes. On peut le prouver élémentairement comme suit : supposons $ND^{-1} = N_0 D_0^{-1}$ où les conditions habituelles entre $\left.\begin{matrix} N, D \\ N_0, D_0 \end{matrix}\right\}$ sont vérifiées. On a $ND^{-1}D_0 = N_0$.

Donc pour tout vecteur polynomial x, $ND^{-1}D_0x$ est polynomial. Compte tenu de N et D premières entre elles à droite, ceci impose $D^{-1}D_0x$ polynomial (il suffit d'écrire AN + BD = I). Donc $D^{-1}D_0$ est polynomial . Par symétrie son inverse aussi. Donc elle est unimodulaire. Ainsi $D_0 = DU$ avec U unimodulaire. Les plus grands d_i sont caractérisés d'après ce qui précède comme la longueur maximale d'une chaine $x, zx, \ldots, z^{d_i-1}x$ telle que $z^{\ell}\, x'\, D^{-1}$ soit propre pour $\ell \leq d_i - 1$, leur nombre p comme le nombre maximal de chaines indépendantes qu'on peut ainsi former. Si on remplace x' par $x'U^{-1}$, la définition ne change pas puisque U^{-1} est unimodulaire. Ainsi pour D et D_0 ces quantités sont égales, et ainsi de suite pour les chaines de longueur inférieure.

Remarque 4 :

On peut vérifier que si $H = ND^{-1}$, la représentation de Führmann est exactement celle de Nérode présentée au §.1.

PARAGRAPHE 6

Paramétrisation des systèmes cycliques.

A) <u>Un théorème de factorisation.</u>

Si on se réfère aux théorèmes du §.0, on voit que les matrices polynomiales peuvent être factorisées sous des formes du type UDV, UT, T_0V où D est diagonale et T, T_0 triangulaires, supérieure et inférieure respectivement, tandis que U et V sont unimodulaires, c'est-à-dire inversibles dans leur anneau, ce qui revient à dire que leur déterminant est constant. De plus des conditions de divisibilité entre les éléments, ou des inégalités sur leurs degrés respectifs, peuvent être imposées. On sait aussi que par une factorisation UDV on met en évidence les invariants de la matrice, mais ce n'est pas le cas pour les formes triangulaires décrites ci-dessus.
On propose ici un théorème de factorisation intermédiaire entre ceux-ci, qui permet d'écrire une matrice régulière (i.e. de déterminant non nul) sous la forme U.D.R. où R est régulière scalaire, U unimodulaire, et D triangulaire, mais sous une forme mettant en évidence les invariants.
L'intérêt éventuel d'une telle factorisation sera évoqué dans la suite.
Pour l'instant énonçons le théorème :

<u>Théorème 1</u> :

Soit K un corps infini. M une matrice régulière de type $n \times n$ dont les éléments sont des éléments de $K[z]$. Il existe $U \in K^{n \times n}[z]$ unimodulaire, et $R \in K^{n \times n}$ régulière, telles que $D = UMH = (d_{ij})$ vérifie :

i) $d_{ij} = 0$ si $i > j$ (D est triangulaire supérieure).

ii) d_{ii} divise tous les éléments $d_{k\ell}$ pour $k \geq i$, ceci $\forall i \in [1, \ell]$.
On peut représenter ces conditions sur le schéma ci-dessous en disant que dans la sous-matrice extraite en pointillé, le terme "en haut à gauche" divise tous les autres termes.

$$D = \left(\begin{array}{cc|ccc} d_{11} & & & x & \\ & d_{22} & & & \\ \hline & & d_{ii} & d_{i,i+1} & \\ & & & \ddots & \\ 0 & & 0 & & \ddots \end{array}\right)$$

Donnons une conséquence de ce théorème avant d'en entreprendre la démonstration.

Corollaire :

•Avec les notations ci-dessus, les éléments diagonaux de la matrice D sont les invariants de N.

Preuve :

Appelons Δ_j le p.g.c.d. de tous les mineurs d'ordre j de D (cf.Corollaire 1 du Théorème 2 §.0). La condition ii) du théorème entraîne que $d_{11}.d_{22}\ldots d_{jj}|\Delta_j$ (rappelons que "|" signifie "divise"). Mais comme un des mineurs vaut précisément $d_{11}\ldots d_{jj}$, c'est lui le p.g.c.d. Alors il résulte du corollaire 1 du théorème 2 §.0 que les d_{ii} sont les invariants de D. Comme D est équivalente à M, ce sont aussi ceux de M.

B) Preuve du théorème.

Nous allons d'abord prouver des lemmes, pour lesquels nous nous appuierons sur le théorème suivant que nous rappelons :

Théorème 2 :

Si K est un corps infini, un espace vectoriel sur K ne peut être union finie de sous-espaces vectoriels propres (i.e. distincts de l'espace entier).

Preuve :

Par récurrence sur le nombre n de sous-espaces intervenant dans l'union.

n=1 Il n'y a rien à montrer.

n quelconque : Supposons $V = \bigcup_{i=1}^{n} V_i$ où V est notre espace, les V_i nos sous-espaces. $\forall i \in \{1\ldots n\}$, il existe $x_i \in V_i$, tel que $x_i \notin V_j$, pour $j \neq i$. Sinon V serait union de n-1 des V_i seulement, et par hypothèse de récurrence ils ne seraient pas tous propres. Soient i et j deux indices distincts, x_i et x_j deux éléments comme ci-dessus $\forall\lambda \in K$, $\exists$ $k(\lambda)$ tel que $x_i+\lambda x_j \in V_{k(\lambda)}$.

Comme K est infini, il existe λ_1 et λ_2 distincts tels que $k(\lambda_1) = k(\lambda_2)$. Mais alors $(x_i+\lambda_1 x_j) - (x_i+\lambda_2 x_j) = (\lambda_1-\lambda_2)x_j \in V_{k(\lambda_1)}$, donc $x_j \in Vk(\lambda_1) \Rightarrow k(\lambda_1) = k(\lambda_2) = j$. Mais alors $x_i \in V_j$, une contradiction qui termine la preuve.

Remarque :

Si K est fini, le théorème est faux, puisque, par exemple, un corps fini est union finie de droites. C'est dans ce théorème que se trouve l'origine de la restriction "K infini" dans l'énoncé du théorème 1. A partir de maintenant, K est un corps infini.

Lemme 1 :

Si $p_1,\dots,p_n$ sont premiers entre eux dans K[z], $p_n \neq 0$, il existe $\lambda_1,\lambda_2,\dots,\lambda_{n-1} \in K$ tels que $\lambda_1 p_1 + \dots + \lambda_{n-1}p_{n-1}$ soit premier avec p_n.

Preuve :

Soient $u_1,\dots,u_k$ les facteurs premiers de p_n dans K[z]. Considérons K[z] comme espace vectoriel sur K. Soit V son sous-espace engendré par $(p_1,p_2,\dots,p_{n-1})$. Si la conclusion du lemme est fausse, pour tout élément q de V, il existe i tel que $u_i|q$. Appelons W_i le sous-espace de K[z] formé des polynômes multiples de u_i. Alors posons $V_i = W_i \cap V$. On a donc $V = \bigcup_{i=1}^{k} V_i$. Par application du théorème. $V \subset V_{i0}$ pour un certain i0. Ainsi $u_{i0}|p_i \ \forall i \in \{1,\dots,n-1\}$ et comme $u_{i0}|p_n$ par définition, les (p_i) ne pourraient pas être premiers entre eux. Ceci prouve le lemme.

Lemme 2 :

Dans la conclusion du lemme 1, on peut choisir un des λ_i arbitrairement, pourvu qu'il soit non nul.

Preuve :

Par homothétie cela revient à montrer qu'on peut choisir $\lambda_i=1$, sinon cela impose que $\sum_{j\neq i,n} \lambda_j p_j + p_i$ n'est jamais premier avec p_n donc divisible par un des u_ℓ pour tout jeu de $\lambda_{j,j\neq i,n}$ fixons donc un tel jeu et examinons les quantités : $Q(\mu) = \mu(\sum_{j\neq i,n} \lambda_j p_j) + p_i$, $\mu \in K$. Comme K est infini, il existe μ et μ' distincts tels que $Q(\mu)$ et $Q(\mu')$ soient divisibles par le même u_ℓ. Mais alors par différence, $\sum_{j\neq i} \lambda_j P_j$ l'est aussi. Ainsi $\sum_{j\neq i,n} \lambda_j P_j$ n'est jamais premier avec P_n. En appliquant le lemme 1, on voit que les P_j, $j\neq i,n$ ne sont pas premiers avec p_n. Par ailleurs il existe une combinaison $\sum_{j=1}^{n-1} \lambda_j P_j$ qui est première avec P_n. Donc dans cette combinaison $\lambda_i\neq 0$ (car les autres P_j ont tous un diviseur commun avec P_n). Par homothétie, on peut imposer $\lambda_i = \lambda$, ce qui montre le lemme.

<u>Lemme 3</u> :

Soit $D = \begin{pmatrix} d_1 & d_2 \\ 0 & d_3 \end{pmatrix}$, une matrice 2×2 triangulaire supérieure à coefficients dans $K[z]$, avec $d_3\neq 0$. Alors il existe $U = \begin{pmatrix} \alpha & \beta \\ \delta & \gamma \end{pmatrix} \in K^{2\times 2}[z]$ vérifiant dét $U = 1$ (donc U est unimodulaire) et $R = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$ scalaire vérifiant dét $R = 1$ (donc H est régulière) telles que si on pose $\mu = \text{p.g.c.d.}\ (d_1,d_2,d_3)$ on aie : $UDR = \begin{pmatrix} \mu & \ell \\ 0 & \frac{d_1 d_3}{\mu} \end{pmatrix}$, où μ divise tous les éléments de la matrice.

<u>Preuve</u> :

Calculons UDR :

$$UDR = \begin{pmatrix} \alpha & \beta \\ \delta & \gamma \end{pmatrix}\begin{pmatrix} d_1 & d_2 \\ 0 & d_3 \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} \alpha d_1 & \alpha d_2+\beta d_3 \\ \delta d_1 & \delta d_2+\gamma d_3 \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

$$= \begin{pmatrix} x\alpha d_1+z(\alpha d_2+\beta d_3) & y\alpha d_1+t(\alpha d_2+\beta d_3) \\ x\delta d_1+z(\delta d_2+\gamma d_3) & y\delta d_1+t(\delta d_2+\gamma d_3) \end{pmatrix}$$

La première équation que nous regardons est celle qui assure le caractère triangulaire du produit UDR :

$$x\delta d_1 + z(\delta d_2+\gamma d_3) = 0. \quad (1)$$

Posons $d_1 = d_1'\mu$, $d_2 = d_2'\mu$, $d_3 = d_3'\mu$. Pour satisfaire à (1), il est légitime de choisir :

$$\left.\begin{array}{l} \delta = -zd_3' \\ \gamma = (xd_1'+zd_2') \end{array}\right\} \text{ car (1) qui s'écrit } \mu[\delta(xd_1'+zd_2')+\gamma zd_3'] = 0 \text{ est alors}$$

vérifiée.

Si à présent δ et γ sont premiers entre eux on peut choisir α et β tels que $\alpha\gamma-\beta\delta = 1$, ce qui assure dét U = 1. D'après le lemme 1 et le lemme 2, on peut choisir x' tel que $(x'd_1'+d_2')$ et d_3' soient premiers entre eux. Alors en multipliant par $z\neq 0$ mais à part cela quelconque on trouve α et β (c'est Bezout). On a donc réservé le choix de $z\neq 0$. Le terme (1,1) du produit UDR s'écrit alors :

$$x\alpha d_1 + z(\alpha d_2+\beta d_3) = \mu[\alpha(xd_1'+zd_2')+\beta zd_3'] = \mu[\alpha\gamma-\beta\delta] = \mu.$$

Il reste à assurer $xt-zy = 1$.

On peut par exemple prendre $t=0$ et $y= -\frac{1}{z}$ puisque $z\neq 0$.

L'élément (1,2) de UDR s'écrit alors $y\alpha d_1$.

La valeur du terme (2,2) est alors nécessairement $\frac{d_1d_3}{\mu}$ pour des raisons de déterminants. Il est par ailleurs clair que μ divise tous les éléments de UDH, ce qui était de toute façon évident car on ne fait que combiner linéairement les éléments de D, dont μ était le p.g.c.d.

Remarque :

Le degré de l'élément (1,1) n'a pas augmenté. Il a même décru strictement si d_1 n'était pas lui-même le p.g.c.d. de (d_1,d_2,d_3), autrement dit si d_1 ne divisait pas d_2 et d_3 il est nécessaire ici pour interpréter le cas $d_1=0$ de poser $d^\circ\ 0 = +\infty$).

A partir de maintenant, pour la commodité des notations, on appelera transformation de type T une transformation du type qui nous intéresse, à savoir multiplication à gauche par une unimodulaire U, et à droite par une scalaire régulière R.

Lemme 4 :

Si $D = \begin{pmatrix} d_{11} & & & X \\ & d_{22} & & \\ & 0 & \ddots & \\ & & & d_{nn} \end{pmatrix}$ est dans $K^{n\times n}[z]$ et si pour un certain i,

d_{ii} ne divise pas tous les éléments de la $i^{\text{ème}}$ ligne, on peut trouver une transformation de type T telle que si $D' = UDR = (d'_{ij})$, on ait $d^\circ\ (d'_{ii}) < d^\circ\ (d_{ii})$, ($d^\circ$ (P) signifie "degré de P") et telle que $d'_{kk} = d_{kk}$ pour $k<i$. En outre, pour $k<i$, $d'_{k\ell}$ est une combinaison linéaire des d_{kj}.

Preuve :

Supposons $d_{ii} \not| \ d_{ik}$ ("$\not|$" signifie "ne divise pas"). Tout d'abord ramenons-nous au cas $k = i+1$; en échangeant deux colonnes (transformation du type R) on obtient une matrice du type :

$$\begin{pmatrix} d_{11} & & & & & \\ & \ddots & & & & \\ & & d_{ii} & d_{ik} & & \\ & 0 & & \hdashline & & \\ & & & X & & \\ & & & X & \ddots & \\ & & & \vdots & 0 & d_{nn} \\ & & & X & & \end{pmatrix}$$

faisant alors des transformations élémentaires sur les lignes de la sous-matrice encadrée (Théorème 2.§.0), on lui restitue sa forme triangulaire (opération du type U). A présent considérons la matrice $D_2 = \begin{pmatrix} d_{ii} & d_{i,i+1} \\ 0 & d_{i+1,i+1} \end{pmatrix}$.

Comme $d_{ii} \nmid d_{i,i+1}$, et que D régulière $\Rightarrow d_{i+1,i+1} \neq 0$, on peut d'après le lemme 3 trouver U_2 et R_2 d'ordre 2 telles que $D_2' = U_2 D_2 R_2 = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ avec $d^\circ a < d^\circ d_{ii}$. Considérons alors les matrices :

$$U = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & \boxed{U_2} & \\ & & & \ddots \\ 0 & & & & 1 \end{pmatrix} \qquad R = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & \boxed{R_2} & \\ & & & & 1 \\ 0 & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

Elles définissent bien une transformation du type T. En outre UDR s'écrit :

$$\begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \boxed{U_2} & \\ & & & 1 \\ 0 & & & & 1 \end{pmatrix} \begin{pmatrix} d_{11} & & & X \\ & \ddots & & \\ & & d_{ii} & d_{i,i+1} \\ 0 & & & d_{i+1,i+1} & \\ & & & & \ddots \\ & & & & & d_{nn} \end{pmatrix} \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & [R_2] & \\ & & & 1 \\ 0 & & & & 1 \end{pmatrix}$$

$$= \begin{pmatrix} d_{11} & & & & X \\ & \ddots & & & \\ & & d_{i-1,i-1} & & \\ & & & [D_2'] & \\ & & & & d_{i+2,i+2} \\ 0 & & & & & \ddots \\ & & & & & & d_{nn} \end{pmatrix}$$

Il est clair que les d_{kk} pour k<i n'ont pas changé dans l'opération, et que $d'_{k\ell}$ est une combinaison linéaire des d_{kj} pour k<i. Ce qui montre le lemme .

Lemme 5 :

Si $D = \begin{pmatrix} d_{11} & & X \\ & \ddots & \\ 0 & & d_{nn} \end{pmatrix}$ et si pour un certain i, d_{ii} ne divise pas, tous les d_{jj} pour $j \geq i$, on peut trouver une transformation du type T telle que si $D' = UDR = (d'_{ij})$ on aie $d° \, d'_{ii} < d° \, d_{ii}$, et $d'_{kk} = d_{kk}$ pour k<i.

Preuve :

Montrons qu'on peut se ramener au cas où $d_{ii} \nmid d_{i+1,i+1}$. En effet, soit k le plus petit indice strictement supérieur à i tel que $d_{ii} \nmid d_{kk}$. Alors $d_{k-1,k-1} \nmid d_{kk}$ car $d_{ii} | d_{k-1,k-1}$ par définition. Alors la même manipulation qu'au lemme 4 sur la matrice $\begin{pmatrix} d_{k-1,k-1} & X \\ 0 & d_{kk} \end{pmatrix}$ nous donne D' avec $d'_{k-1,k-1} | d_{kk}$. Observons par ailleurs que $d'_{ii} = d_{ii}$ n'a pas changé. Donc $d'_{ii} \nmid d'_{k-1,k-1}$ car sinon $d'_{ii} = d_{ii} | d'_{k-1,k-1} | d_{kk} \Rightarrow$ contradiction. Ainsi on s'est "rapproché de i" (k a décru de 1). En un nombre fini d'opérations on se ramène au cas k = i+1. De plus on a laissé invariants les d_{kk} pour k<i. A présent il ne reste plus qu'à effectuer la même opération qu'à la preuve du lemme 4 pour voir le degré de d_{ii} diminuer. Les d'_{kk}, pour k<i, restant inchangés, ceci prouve le lemme.

Lemme 6 :

Etant donné $D = \begin{pmatrix} d_{11} & & X \\ & \ddots & \\ 0 & & d_{nn} \end{pmatrix}$, il existe une transformation du type T telle que si on pose $D' = UDR = \begin{pmatrix} d'_{11} & & (d'_{ij}) \\ & \ddots & \\ 0 & & d'_{nn} \end{pmatrix}$ on a $d'_{11} | d'_{ij}$, $\forall ij$.

Preuve :

Par application répétée du lemme 4, on peut supposer $d'_{11}|d'_{1\ell}\ \forall\ell$; appliquant de nouveau ce lemme à la $2^{\text{ème}}$ ligne, pour obtenir une matrice D", on peut supposer $d''_{22}|d''_{2\ell}$, et $d''_{11}|d''_{1\ell}$ car $d''_{11} = d'_{11}$ et $d''_{1\ell}$ est remplacé par une combinaison linéaire de d'_{1k}.

De proche en proche, on peut assurer que $d'_{ii}|d'_{ij}$ avec $d^\circ\, d'_{11} \leq d^\circ\, d_{11}$. Si alors $d'_{11} \nmid d'_{ii}$, le degré de d'_{11} décroît strictement. Ainsi en un nombre fini d'opérations, les conditions du lemme sont remplies.

Preuve du théorème 1 :

D'après le (Théorème 2.§.0), on peut transformer M en une matrice triangulaire supérieure D. Procédons alors par récurrence sur la dimension de D

si n=1 : rien à montrer.

Appelons, dans le cas général, $\overline{D}$ la matrice obtenue en rayant la première ligne et la première colonne de D. Appliquons le lemme 6 : on peut supposer que $d_{11}|d_{ij}$, $\forall ij$. Soient $\overline{U}$ et $\overline{R}$ les matrices de la transformation de type T associée à $\overline{D}$ par l'hypothèse de récurrence. Alors posons

$$U = \begin{pmatrix} 1 & 0 & \cdots \\ 0 & & \\ \vdots & \overline{U} & \end{pmatrix} \quad \text{et} \quad H = \begin{pmatrix} 1 & 0 & \cdots \\ 0 & & \\ \vdots & \overline{H} & \end{pmatrix}$$

$$UDH = \begin{pmatrix} 1 & 0 & \cdots \\ 0 & & \\ \vdots & \overline{U} & \end{pmatrix} \begin{pmatrix} d_{11} & & \\ 0 & & \\ \vdots & & \overline{D} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots \\ 0 & & \\ \vdots & \overline{H} & \end{pmatrix}$$

$$= \begin{pmatrix} d_{11} & & \\ 0 & & \\ \vdots & \overline{U}\,\overline{D} & \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots \\ 0 & & \\ \vdots & \overline{H} & \end{pmatrix} = \begin{pmatrix} d_{11} & X' & X' & \cdots \\ 0 & & & \\ \vdots & & \overline{U}\,\overline{D}\,\overline{H} & \end{pmatrix}$$

Observons que les éléments marqués "X'" sont des combinaisons linéaires des d_{1j}, donc sont divisibles par d_{11}. De même comme d_{11} divisait tous les éléments de $\overline{D}$, il divise tous les éléments de $\overline{U}\,\overline{D}\,\overline{R}$ (ceci achève la preuve du théorème).

C) Application à la paramétrisation des systèmes cycliques causaux.

L'intérêt d'une factorisation du type précédent peut être le suivant : pour paramétrer un système de module d'état donné, il est possible de factoriser le dénominateur D de $H = D^{-1}N$ en $D = U \operatorname{diag}\{a_i\} V$ comme dans le §.5. Cependant la présence de V rend complexe l'analyse de la causalité de $H = V^{-1} \operatorname{diag}\{\frac{1}{a_i}\} N'$ (avec $N' = U^{-1}N$). Si V est scalaire régulière cette causalité se réduit à celle du produit des deux matrices $\operatorname{diag}\{\frac{1}{a_i}\} N'$.

Malheureusement on ne parvient pas en général à une forme diagonale, en imposant V scalaire. On peut néanmoins utiliser celle du théorème, qui donne accès aux invariants. Mais l'analyse de la propreté de $D^{-1}N$ avec $D = \begin{pmatrix} d_{11} & & X \\ & \ddots & \\ 0 & & d_{nn} \end{pmatrix}$ est lourde, quoique possible. Il est cependant un cas dans lequel l'expression se simplifie, c'est celui où il n'y a qu'un invariant non scalaire, autrement dit celui où $\frac{K^n[z]}{DK^n[z]}$ est cyclique. En effet :

Proposition 1 :

Si $\frac{K^n[z]}{DK^n[z]}$ est cyclique, il existe une transformation de type T telle que :

$$UDR = \begin{pmatrix} 1 & 0 & & a_1 \\ & \ddots & & \vdots \\ 0 & & & a_n \end{pmatrix} \qquad \text{avec } d^\circ\, a_i < d^\circ\, a_n$$

Preuve :

Au vu du théorème 1, tout ce qu'il y a à prouver est qu'on peut obtenir par des transformations du type U d° $d_{ij} < d_{jj}$ pour $i<j$. Mais ceci est clair grâce à la division euclidienne (c'est d'ailleurs la base de la démonstration du Théorème 2.§.0 cf[JA]).

Noter qu'alors $(UDR)^{-1}$ s'écrit aisément, c'est :

$$\begin{pmatrix} 1 & & -\frac{a_i}{a_n} \\ & \ddots & \vdots \\ 0 & & \frac{1}{a_n} \end{pmatrix}$$

Autrement dit un système cyclique a une fonction de transfert m×p H factorisable sous la forme :

$$H = R \begin{pmatrix} 1 & 0 & -\frac{a_i}{a_m} \\ 0 & & \vdots \\ & & \frac{1}{a_m} \end{pmatrix} N, \quad \text{où } N \text{ est polynomiale.}$$

Si on appèle $E(\frac{P}{Q})$ la partie entière de la fraction rationelle $\frac{P}{Q}$ (i.e. le quotient, éventuellement nul, de la division euclidienne), alors on a :

Proposition 2 :

$$\left[\begin{pmatrix} 1 & \mathbf{0} & -\frac{a_i}{a_m} \\ & \ddots & \\ 0 & & \frac{1}{a_m} \end{pmatrix} N \text{ propre} \right] \Leftrightarrow \left[N = \begin{pmatrix} E(\frac{a_i b_j}{a_m}) \\ \vdots \\ b_j \end{pmatrix} \right]$$

avec $d° b_j < d° a_m$

Preuve :

Il suffit d'écrire que la colonne j du produit de gauche est propre : appelant n_{ij} les éléments de N, on a :

$$\text{et} \quad \left\{ \begin{array}{l} d^\circ n_{mj} < d^\circ a_m \\ \\ n_{ij} - \dfrac{a_i b_j}{a_m} \text{ propre, donc } n_{ij} = E\left(\dfrac{a_i b_j}{a_m}\right). \end{array} \right.$$

Il peut être intéressant de savoir si $d^\circ a_m$ est le degré de la représentation minimale de H dans une telle factorisation, c'est-à-dire si

$$D = \begin{pmatrix} 1 & & a_i \\ & \ddots \; 0 & \\ 0 & & a_m \end{pmatrix}$$ et N sont premières entre elles à gauche.

On a à ce propos le résultat suivant :

Proposition 3 :

Les matrices :

$$D = \begin{pmatrix} 1 & & & a_1 \\ & \ddots & 0 & \vdots \\ & 0 & \ddots & a_i \\ & & & \vdots \\ & & & a_m \end{pmatrix} \text{ et } N = \begin{pmatrix} E\left(\dfrac{a_i b_j}{a_m}\right) \\ \vdots \\ b_j \end{pmatrix}$$

sont premières entre elles à gauche si et seulement si les b_j et a_m sont premiers entre eux.

Preuve :

Si μ est un diviseur commun non trivial des a_m et des b_j, alors :

$$M = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & \mu \end{pmatrix}$$ est un diviseur commun non trivial à gauche de D et N.

Si les b_j et a_m sont premiers entre eux il existe, d'après le lemme 1 des scalaires $\lambda_1,\ldots,\lambda_p$ tels que $\Sigma\lambda_j b_j$ soit premier avec a_m. Il existe donc une combinaison V des colonnes de N dont la dernière composante est première avec a_m. Un coup d'oeil à la représentation de Führmann permet de voir que ce vecteur V est son propre représentant dans $\overline{M} = \frac{K^m[z]}{DK^m[z]}$. Par ailleurs si p est un polynôme de K[z], $pV \in DK^m[z] \Rightarrow a_m | p$ car tout élément de $DK^m[z]$ a une dernière composante divisible par a_m et celle de V est première avec a_m. Donc V est un générateur de $\overline{M}$ car il est de période a_m. Ceci montre que $\mathrm{Im}\ \phi = \overline{M}$ (cf.§.2) donc que la représentation de Führmann est commandable, et donc que D et N sont premières entre elles, à gauche.

Remarque :

Il est possible de donner une démonstration ne faisant aucune référence à la réalisation. En effet si on pose $\Lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix}$, dire que V est générateur de $\overline{M}$ sur K[z] , c'est dire que tout vecteur de $K^m[z]$ s'écrit $Dx+N(p\Lambda)$ où p est un polynôme de K[x]. Ainsi le module $\{Dx+Ny\}$, $x \in K^m[z]$, $y \in K^p[z]$, est $K^m[z]$ en entier. D'après le paragraphe 4, D et N sont premières entre elles à gauche.

On peut résumer les considérations qui précèdent dans le :

Théorème 3 :

Une fonction de transfert H représente un système cyclique causal de dimension minimale n si et seulement si on peut écrire :

$$H = R \begin{pmatrix} 1 & & -\frac{a_i}{a_m} \\ & \ddots & \\ & & \frac{1}{a_m} \end{pmatrix} \begin{pmatrix} E(\frac{a_i b_j}{a_m}) \\ \vdots \\ b_j \end{pmatrix}$$

avec : R scalaire régulière

$d° a_i < n \quad i \neq m \quad d° b_j < n$

$d° a_m = n$

p.g.c.d. $(a_m, b_j) = 1$.

<u>Preuve</u> :

Le caractère nécessaire résulte de ce qui précède. Le caractère suffisant résulte de la représentation de Führmann.

<u>Corollaire</u> :

Si un système admet pour sa fonction de transfert une factorisation $H = D^{-1}N$ minimale, où D et N ont les formes de la proposition 3, alors :

i) Le système est commandable par l'entrée j si et seulement si b_j et a_m sont premiers entre eux.

ii) Le système est observable par la sortie $i (i<m)$ si et seulement si a_i et a_m sont premiers entre eux.

iii) Le système est toujours observable par sa dernière sortie. (Rappelons qu'il est cyclique).

<u>Preuve</u> :

i) Résulte de l'application de la proposition 3 à la fonction de transfert $[H(0, b_{j0}, 0)]$.

ii) Résulte du théorème 2 du §.5 appliqué à $[{}^t(e_{i0}) \cdot H]$.

iii) Idem.

Avant d'aller plus loin, introduisons pour la commodité quelques notations.

Définition 1 :

Une factorisation du type indiqué dans le théorème 3 sera dite D-factorisation à gauche. Si en outre R = Id, on dira que c'est une D-factorisation simple à gauche. On appelera D-factorisation à droite une factorisation de type transposé, à savoir :

$$\mathcal{H} = N_1 D_1^{-1} R_1$$

avec R_1 scalaire régulière,

$$N_1 = \left(\quad E\left(\frac{d_i c_j}{c_p}\right) \ \dots \ d_i \quad \right) \qquad D_1 = \begin{pmatrix} 1 & & 0 \\ & 0 \ddots & \\ c_1 & \dots & c_p \end{pmatrix}$$

si R_1 = Id, on dira que la D-factorisation à droite est simple.

Nous allons à présent chercher comment la représentation de Führmann permet de déduire une réalisation d'une D-factorisation simple à gauche (et partant d'une D-factorisation quelconque). De façon précise on a le :

Théorème 4 :

Si $D^{-1}N$ est une D-factorisation simple à gauche, avec les notations du théorème 3, on en déduit une réalisation [H,F,G] (en notant pour $p \in K[z]$ $p = \Sigma p^i z^i$ et en supposant $a_p^n = 1$) par les formules :

$$G = \begin{bmatrix} b_1^0 & & b_m^0 \\ \vdots & & \vdots \\ b_1^{n-1} & & b_m^{n-1} \end{bmatrix}$$

$$F = \begin{bmatrix} 0 & \ddots & & -a_p^0 \\ 1 & \ddots & 0 & \vdots \\ 0 & & 1 & -a_p^{n-1} \end{bmatrix} \qquad \text{(forme compagnon)}$$

pour H, enfin, si on pose :

$$M = \begin{bmatrix} a_p^1 & a_p^2 & \cdots\cdots & 1 \\ \vdots & \vdots & \diagup & \\ \vdots & a_p^{n-1} & & \\ a_p^{n-1} & 1 & & 0 \\ 1 & & & \end{bmatrix}$$

on a :

$$H = \begin{bmatrix} -\tilde{H} \\ \\ 00\cdots\cdots 01 \end{bmatrix}, \text{ avec } \tilde{H} = \begin{bmatrix} a_1^o \cdots\cdots a_1^{n-1} \\ \cdots\cdots\cdots\cdots \\ a_{p-1}^o \cdots\cdots a_{p-1}^{n-1} \end{bmatrix} M^{-1} .$$

Preuve :

L'espace d'état d'après les résultats du §.5 est le module quotient $\frac{K^m[z]}{DK^m[z]} = \bar{M}$. Une base en est (e_{mi}) $0 \leq i \leq n-1$. En effet, ces vecteurs sont au nombre de $n = d°$ dêt D, ils sont manifestement indépendants, et ils sont leurs propres représentants dans $\bar{M}$. Il est clair que dans cette base, F (qui est la multiplication par z dans $\bar{M}$) a la forme indiquée. Remarquons que si $V = \begin{bmatrix} Q_1 \\ \vdots \\ Q_m \end{bmatrix} \in K^m[z]$, son représentant dans $\bar{M}$ est

$\begin{bmatrix} 0 \\ \vdots \\ 0 \\ R(Q_m) \end{bmatrix}$ où $R(Q_m)$ est le reste de Q_m dans la division par a_m. Ceci est clair car le vecteur proposé est son propre représentant (son produit par D^{-1} à gauche est propre) et en outre il est équivalent à V car leur différence a une dernière composante multiple de a_m, et donc appartient à $DK^m[z]$;

(en effet la condition de solvabilité de Dx=y, y donné dans $K^m[z]$ est que la dernière composante de y soit divisible par a_m, comme on s'en convaincra aisément). Ainsi on a $\Pi_D N e_{i0} = \begin{pmatrix} 0 \\ \vdots \\ b_i \end{pmatrix}$, ce qui montre que G a la forme indiquée.

Pour déterminer H, il faut déterminer $(D^{-1}e_{mi})_1$, c'est-à-dire $\begin{pmatrix} -\frac{a_i}{a_m} & z^i \\ \frac{z^i}{a_m} & \end{pmatrix}_1$,

pour $i \in \{0,\dots,n-1\}$. Ceci permet déjà de voir que H a la dernière ligne annoncée. Il reste à déterminer les n premiers termes du développement en série de $\frac{a_i}{a_m}$. Posons : $\frac{a_i}{a_m} = \sum_{k=1}^{\infty} \left(\frac{a_i}{a_m}\right)_k z^{-k}$.

La relation $a_m\left(\frac{a_i}{a_m}\right) = a_i$ s'écrit :

$$\sum_r \left(\frac{a_i}{a_m}\right)_r a_m^{k+r} = a_i^k \quad k=0,\dots,n-1.$$

Matriciellement, cette relation s'écrit :

$$\begin{bmatrix} a_1^0 & \dots & a_1^{n-1} \\ \vdots & & \vdots \\ a_{p-1}^0 & \dots & a_{p-1}^{n-1} \end{bmatrix} = \begin{bmatrix} \left(\frac{a_1}{a_m}\right)_1 & \dots & \left(\frac{a_1}{a_m}\right)_m \\ \dots & \dots & \dots \\ \left(\frac{a_{p-1}}{a_m}\right)_1 & \dots & \left(\frac{a_{p-1}}{a_m}\right)_n \end{bmatrix} \begin{bmatrix} a_p^1 & a_p^2 & & 1 \\ \vdots & \vdots & \iddots & \\ a_p^{n-1} & 1 & & \\ 1 & & 0 & \end{bmatrix}$$

Ceci prouve le théorème.

D) Symétrisation - Rôle de R.

La symétrisation est un outil qui va nous permettre de mieux analyser le rôle de R. Cette opération s'énonce ainsi :

si $D^{-1}N$ est une D-factorisation simple à gauche, inéductible, trouver une

D-factorisation simple à droite $N_1D_1^{-1}$,inéductible aussi,telle que :

$$D^{-1}N = N_1D_1^{-1}.$$

On ne peut pas toujours, comme on va le voir, résoudre une telle équation. Cependant :

Proposition 5 :

Avec les notations de la définition (1), alors si b_p et a_m sont premiers entre eux, le problème de symétrisation a une solution unique.

Preuve :

$$D^{-1}N = \begin{pmatrix} 1 & & -\frac{a_i}{a_m} \\ & \ddots & \\ 0 & & \frac{1}{a_m} \end{pmatrix} \begin{pmatrix} E(\frac{a_i b_j}{a_m}) \\ \\ b_j \end{pmatrix} = \begin{pmatrix} -PP(\frac{a_i b_j}{a_m}) \\ \\ \frac{b_j}{a_m} \end{pmatrix}$$

où "PP" signifie "partie propre", pour une fraction rationelle, à savoir ce qui reste lorsqu'on a retiré la partie entière. De manière analogue :

$$N_1D_1^{-1} = \begin{pmatrix} -PP(\frac{d_i c_j}{c_p}) & \frac{d_i}{c_p} \end{pmatrix}$$

Si on cherche à ce que la factorisation $N_1D_1^{-1}$ soit minimale aussi, on est amené à poser :

$$\begin{cases} c_p = a_m \\ b_p = d_m. \end{cases}$$

On veut par ailleurs avoir :

$$\forall j \neq p \qquad \frac{b_j}{a_m} = -PP(\frac{d_m c_j}{c_p}) \Leftrightarrow \frac{b_j}{a_m} = -PP(\frac{b_p c_j}{a_m}).$$

Ceci revient à dire que $a_m | b_p c_j + b_j$. Le problème est donc : peut-on trouver c_j de degré < n tel que $b_j = \lambda a_m - b_p c_j$ pour $\lambda \in K[z]$? La réponse, affirmative, est donnée dans le lemme suivant (cf[RU]) :

Lemme 7 :

Si a_m et b_p sont deux polynômes premiers entre eux où $d° a_m \geq d° b_p$, $d° a_m = n$. Si on appelle V l'espace vectoriel des polynômes $\{P = \lambda a_m + \mu b_p, (\lambda,\mu) \in K^2[z], d° \lambda < n, d° \mu < n\}$, alors $V = K_{2n-1}[z]$, l'ensemble des polynômes de degré strictement inférieur à 2n dans K[z].

Preuve :

Les polynômes $\{z^i a_m\}$, $0 \leq i \leq n-1$ et $\{z^j b_p\}$, $0 \leq j \leq n-1$ sont générateurs de V, et sont au nombre de 2n. Comme $V \subset K_{2n-1}[z]$ qui est de dimension 2n, il suffit de montrer qu'ils sont libres sur K. Ceci revient à dire que si (λ,μ) sont dans $K^{n-1}[z]$, $\lambda a_m + \mu b_p = 0 \Rightarrow \lambda = \mu = 0$. Mais $\lambda a_m = -\mu b_p \Rightarrow a_m | \mu$ par le théorème de Gauss, puisque a_m et b_p sont premiers entre eux. Donc $\mu=0$, car $d° \mu < n$. Mais alors $\lambda=0$. Ceci prouve le lemme. On peut donc trouver c_j, qui est unique.

A présent observons que nous devons avoir :

$$\frac{d_i}{c_p} = -PP\left(\frac{a_i b_p}{a_m}\right) \qquad i \neq m$$

soit : $$\frac{d_i}{a_m} = -PP\left(\frac{a_i b_p}{a_m}\right) \qquad \text{ce qui détermine } d_i.$$

On a donc réussi à déterminer les d_i et le c_j tels que les dernières lignes et colonnes de $D^{-1}N$ et $N_1 D_1^{-1}$ coïncident. Considérons les autres éléments :

modulo (a_m), on a : $d_i c_j \equiv -a_i b_p c_j$ et $a_i b_j \equiv -a_i b_p c_j$ aussi.

Il s'ensuit que $PP\left(\frac{a_i b_j - d_i c_j}{a_m}\right) = 0$. Mais ceci signifie simplement l'égalité des deux factorisations : $D^{-1}N = N_1 D_1^{-1}$. L'unicité est claire d'après le calcul.

Ce résultat étant acquis observons qu'une conséquence du corollaire du théorème 3 est la suivante : dans toute D-factorisation à gauche d'une fonction de transfert cyclique, $\mathcal{H} = RD^{-1}N$, $R^{-1}H$ est observable par sa dernière sortie, c'est-à-dire que les coefficients de la dernière ligne de R^{-1} combinent les lignes de $\mathcal{H}$, de façon que cette combinaison des sorties permette à elle seule d'observer le système. Le théorème suivant montre que cette condition est aussi suffisante pour que R soit associée à une D-factorisation à gauche.

Théorème 5 :

Une condition nécessaire et suffisante pour qu'une matrice R scalaire et régulière participe à une D-factorisation à gauche (resp. à droite) de la fonction de transfert cyclique $\mathcal{H}$ est que la dernière ligne (resp. colonne) de R^{-1}, $(\lambda_1,\ldots,\lambda_m)$, (resp. $\mu_1,\ldots,\mu_p$), soit telle que si (H,F,G) est une réalisation minimale de $\mathcal{H}$, et $(h_1,\ldots,h_m)$ (resp. $g_1,\ldots,g_p$) sont les lignes (resp. colonnes) de H (resp. de G), la paire $(\Sigma\lambda_i h_i, F)$ (resp. $(F,\Sigma\mu_i g_i)$ soit observable. (resp. commandable). Cette notion est indépendante de la réalisation choisie.

L'ensemble Ω_1 (resp. Ω_2) des (λ_i) (resp. (μ_i)) qui conviennent, et qui est non vide d'après la proposition 1, est un ouvert de Zariski dans K^m.

Preuve :

Il suffit de prouver les affirmations concernant l'observabilité, les autres s'en déduisant par dualité. Que les propriétés soient indépendantes de la réalisation est clair. La condition nécessaire découle du corollaire du théorème 3. Prouvons le caractère suffisant : D-factorisons à gauche ${}^t\mathcal{H}\,{}^tR^{-1} = R_0 D^{-1} N$. Comme $R^{-1}\mathcal{H}$ est observable par sa dernière sortie par hypothèse, ${}^t\mathcal{H}\,{}^tR^{-1}$ est commandable par la dernière entrée par dualité. D'après le corollaire du théorème 3 et la proposition 5, on peut écrire $D^{-1}N = N_1 D_1^{-1}$, donc ${}^t\mathcal{H}\,{}^tR^{-1} = (R_0 N_1) D_1^{-1} = (N')D_1^{-1}$. Ainsi on a $\mathcal{H} = R({}^tD_1)^{-1}\,{}^tN'$, qui est une D-factorisation à gauche ce qui prouve la condition suffisante.

Enfin, il est clair que si ${}^t\mathcal{H} = RD^{-1}N$, l'ensemble Ω_1 est composé des (λ_i) tels que l'on aie, avec les notations du théorème 3 $\sum_{i=1}^{m}(\lambda_i b_i)$ premier avec a_m. Ainsi, en considérant le résultant de a_m et $\Sigma\lambda_i b_i$, on voit que les λ_i qui conviennent sont ceux qui n'annulent pas ce résultant (qui est un polynôme non trivial en les λ_i). Par définition Ω_1 est un ouvert de Zariski. Ceci achève la preuve.

Note :

Pour ne pas rompre le fil, et parce qu'on attache moins d'importance à cette partie de l'énoncé, on n'a pas défini la notion de résultant de deux polynômes, ni celle d'ouvert de Zariski. On renvoie pour cela à [L.G] et [Z.S si besoin est.

Le théorème 5 permet d'interpréter, du point de vue de l'automatique, le rôle de R. Il a par contrecoup une conséquence purement algébrique qui renforce légèrement la proposition 1 :

Proposition 6 :

Dans la décomposition de la proposition 1, on peut choisir la matrice R si et seulement si elle est régulière et si sa dernière ligne est un vecteur engendrant du module $\frac{K^n[z]}{{}^tDK^n[z]}$. La régularité est triviale. Prouvons le caractère nécessaire de la deuxième condition : on a ${}^tM = {}^tR\,{}^tD\,{}^tU$.

Considérons le vecteur $\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix}$. Supposons que l'on aie $P\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} = {}^tD(y)$ avec $P \in K[z]$ et $y \in K^m[z]$. Dans ce cas $y = {}^tD^{-1}P\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} = \begin{bmatrix}0\\ 0\\ \vdots\\ \frac{P}{a_m}\end{bmatrix}$. Donc $a_m | P$, et $\begin{bmatrix}0\\ 0\\ \vdots\\ 1\end{bmatrix}$ est engendrant dans $\frac{K^m[z]}{{}^tDK^m[z]}$.

Comme $\frac{K^m[z]}{{}^tDK^m[z]} = \frac{K^m[z]}{{}^tD\,{}^tUK^m[z]}$ (puisque ${}^tU$ est unimodulaire) $\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix}$ est engendrant dans ce second module. Supposons alors $P.{}^tR\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix} = My$. Alors comme ${}^{t}R$ est régulière, on a $P.\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = {}^{t}D\,{}^{t}Uy$; donc $a_m | P$. Ainsi la condition nécessaire est établie en transposant

Pour la suffisance, observons que d'après le théorème 5, on peut factoriser $M^{-1} = R^{-1}D^{-1}.N$

Montrons que N est unimodulaire : rendant les déterminants, on a :

$$\frac{1}{\text{dét } M} = \frac{\text{dét } R^{-1} \,.\, \text{dét } N}{\text{dét } D} .$$

Comme les deux factorisations sont réduites, on a dét M = dét D (c'est le produit des invariants de l'espace d'état), donc dét N est une constante. En inversant, on a le résultat.

Remarque :

Il peut paraître curieux d'utiliser la possibilité de factoriser des matrices rationnelles dans le but de prouver un résultat sur la factorisation de matrices polynomiales, comme on le fait dans la preuve de cette proposition. Nous ne connaissons pas de preuve directe, i.e. travaillant directement sur la construction des matrices R et U.

On peut appliquer le théorème 5 à des classes de matrices diverses selon les cas, par exemple choisir R orthogonale, ou du type $\begin{pmatrix} 1 & & \\ & \ddots & 0 \\ 0 & & 1 \\ x & \dots & x \end{pmatrix}$, etc...

Pour déterminer ce paragraphe on peut énoncer une proposition de caractère plus symétrique que les précédentes :

Proposition 7 :

Avec les notations du théorème 5, si R_1 (resp.R_2) est scalaire régulière et si R_1^{-1} (resp.R_2^{-1}) a sa dernière ligne (resp. colonne) dans Ω_1 (resp. Ω_2), on peut écrire $H = R_1D^{-1}.NR_2 = R_1N_1D_1^{-1}R_2$ qui sont des

D-factorisations à gauche et à droite respectivement, D, N, N_1, D_1 étant alors uniques.

Preuve :

Comme $R_1^{-1} \mathcal{H} R_2^{-1}$ est simultanément observable par sa dernière entrée et commandable par sa dernière sortie, l'existence résulte du théorème 5 et de la proposition 5 conjuguée avec le corollaire du théorème 3. L'unicité résulte de l'expression de $D^{-1}N$ telle qu'elle est calculée dans la preuve de la proposition 5 (puisque a_m est monique), car alors les b_j sont tous égaux, puis les a_i par un raisonnement aisé :

$$PP\left(\frac{a_i b_p}{a_m}\right) = PP\left(\frac{a_i' b_p}{a_m}\right) \Rightarrow a_m | b_p (a_i - a_i')$$

et comme a_m et b_p sont premiers entre eux, $a_m | a_i - a_i' \Rightarrow a_i = a_i'$ pour des raisons de degré. Ensuite N_1 et D_1 sont déterminés uniquement par N et D selon la proposition 5.

Remarques :

Pour conclure ce paragraphe, ajoutons que :

- D'une part le théorème 1 est faux si M n'est pas régulière. Pour s'en convaincre, on peut examiner dans le cas 2x2 les calculs page 62, avec $d_3 = 0$, d_1 et d_2 premiers entre eux, $d° \, d_2 > d° \, d_1$.

- D'autre part, nous ne savons pas pour l'instant décrire un représentant caractéristique dans chaque orbite, ainsi qu'on le fait pour les transformations unimodulaires à droite et à gauche, en exhibant la forme invariante.

REFERENCES

[AN] A.C. ANTOULAS. "A Polynomial Matrix Approach to F mod G- Invariant Subspaces". Diss. Eth. n° 6442. Swiss Federal Institute of Technology, Zurich.

[BI] N. BOURBAKI. Livre II Algèbre Chapitre 7. Hermann.

[DV] DESOER VIDYASAGAR. "Feedback systems. Input output properties". Academic Press. "Electrical Science Series".

[FU] P.A. FUHRMANN. "Algebraic system theory : an Analyst's Point of view". Journal of the Franklin Institute 301 : 521-540.

[HE] M. HEYMANN. "Structure and realization problems in the theory of dynamical systems". CISM Courses and lectures n° 204. Springer-Verlag.

[HH] M. HAUTUS and M. HEYMANN. "Linear Feedback, an Algebraic Approach". SIAM J. Cont. and Opt. 16 : 83 - 105. 1978.

[JA] N. JACOBSON. "Lectures in abstract algebra". T.II. "Linear Algebra". Springer-Verlag.

[KFA] KALMAN, FALB, ARBIB. "Topics in Mathematical System Theory". Mc Graw-Hill.

[LG] S. LANG. "Algebra". Addison-Wesley.

[MU] C. MUTAFFIAN. "Le Défi Algébrique". T.II. Vuibert.

[RU] G. RUCKEBUSCH. "Sur l'approximation rationnelle des Filtres". Centre de Mathématiques Appliquées de l'Ecole Polytechnique. Rapport n° 35.

[ZS] O. ZARISKI and P. SAMUEL. "Commutative Algebra". Vol.I,II. Graduate texts in mathematics. Springer Verlag 2^e^ Ed. 1979.

APPENDICE 1

A propos du cas non cyclique.

Le théorème de factorisation nous a fourni quelques résultats dans le cas cyclique, parce qu'alors la matrice obtenue était d'inverse "simple" et que c'est son inverse que l'on utilisait pour exprimer la fonction de transfert.

Il n'en va plus ainsi dans le cas général, où la matrice est à priori triangulaire pleine, ce qui rend son inversion bien pénible. Il est naturel de chercher à raffiner la forme de la factorisation, c'est-à-dire de chercher à exhiber des représentants plus "simples" des orbites sous l'action du groupe $\mathcal{U}\times\mathcal{R}$ ($\mathcal{U}$: unimodulaires, $\mathcal{R}$: régulière scalaire).

Nous connaissons pour l'instant fort peu de choses sur ces orbites, et aucune forme particulière n'a été exhibée dans le cas général.

Afin d'illustrer, pour terminer, le théorème général, nous allons néanmoins essayer d'analyser la causalité d'une D-factorisation à gauche :

$$\mathcal{H} = D^{-1}N.$$

Posons $D = (d_{ij})$ suivant le théorème, avec :

$$d_{ij} = 0 \quad \text{si} \quad i>j,$$

$$d_{ij} = q_{ij}d_{ii}, \text{ avec } q_{ii} = 1, \text{ pour } \quad i\leq j.$$

Essayons de décrire $D^{-1} = (\ell_{ij})$. On sait qu'elle est triangulaire, donc :

$$\ell_{ij} = 0 \quad \text{si} \quad i>j.$$

Par ailleurs, il est clair que $\ell_{ii} = \frac{1}{d_{ii}}$. Pour exprimer les autres ℓ_{ij}, observons qu'on a :

$\Sigma d_{ik}\ell_{kj} = 0$ si $i \neq j$, soit, en simplifiant à cause du caractère triangulaire, et en tenant compte de l'expression de d_{ik} (qui est une conséquence du théorème) :

$$\ell_{ij} = - \sum_{k=i+1}^{j} q_{ik}\ell_{kj} \qquad \text{pour } i < j.$$

Cette formule donne ℓ_{ij} en fonction des éléments qui sont en dessous de lui sur la colonne. Il faut remarquer qu'une telle relation est obtenue à chaque fois qu'on inverse une matrice triangulaire, mais que, par contre, l'absence de dénominateur dans l'équation provient des conditions de divisibilité du théorème. On peut bien sûr tirer de ces formules une expression explicite de ℓ_{ij} en fonction des éléments de D, mais l'indiçage requiert un certain soin. Notons Σ_{ik} l'ensemble des suites de nombres ordonnés entre i-k et i pour $0<k<i$, qui commencent en i-k et finissent en i (le nombre de points intermédiaires est libre). Si $\sigma \in \Sigma_{ik}$, notons $N(\sigma)$ le nombre de points de σ (y compris les extrémités) diminué de 1. Notons Q_σ le produit $q_{i-k,i1} \cdot q_{i1,i2} \cdots q_{ip,i}$, où $(i-k,i1,\ldots,ip,i)$ sont les points de σ. Alors on voit par récurence que :

$$\ell_{i-k,i} = \frac{1}{d_{ii}} \left[\sum_{\sigma \in \Sigma_{ik}} (-1)^{N(\sigma)} Q_\sigma \right]$$

Posons par définition $\ell_{i-k,i} = \dfrac{1}{d_{ii}} P_{i-k,i}$.

La matrice D^{-1} étant de taille n, appelons E la matrice qui en est déduite en supprimant la première ligne et la première colonne. Soit W l'ensemble des vecteurs polynomiaux N tels que $D^{-1}N$ soit propre, W' l'ensemble des vecteurs polynomiaux N' tels que EN' soit propre. Alors il est manifeste que si $N = \begin{pmatrix} n_1 \\ N' \end{pmatrix} \in W$, $N' \in W'$, et si $N' = \begin{pmatrix} n_2 \\ \vdots \\ n_n \end{pmatrix}$, on a :

$\frac{1}{d_{11}}\left[n_1 + \sum\limits_{k=2}^{n} \frac{P_1k}{M_1k}\, n_k\right]$ propre, avec $M_{1k} = \frac{d_{kk}}{d_{11}}$ (qui est un polynôme car $d_{11} | d_{kk}$). Il est clair que ceci impose :

$$n_1 = - E\left(\sum_{k=2}^{n} \frac{P_1k}{M_1k} n_k\right) + r_1, \qquad d^\circ r_1 < d^\circ d_{11}.$$

Réciproquement, cette condition assure la propreté du produit. Il est clair que la formule obtenue généralise celle du cas cyclique, mais elle n'est pas aussi simple malheureusement (dans le cas cyclique tous les P_{ij}, $j \neq n$, sont nuls). On peut rassembler nos résultats dans le :

Théorème 1 :

Avec les notations précédentes l'espace vectoriel des matrices $N \in K[z]^{n\times p}$ telles que $D^{-1}N$ soit propre, est composé des matrices dont chaque colonne a pour coefficient à ligne k :

$$\sum_{i=k}^{n} \sum_{\sigma\in\Sigma_{i,(i-k)}} (-1)^{N(\sigma)} E\left(\frac{P_{k,i1}}{M_{k,i1}} E\left(\ldots E\left(\frac{P_{ip,i}}{M_{ip,i}} r_i\right)\ldots\right)\right)$$

où $\{k,i1,\ldots,i\}$ sont les éléments de $\sigma \in \Sigma_{i,i-k}$, le coefficient valant par convention 1 pour $i=k$.

Les paramètres étant les r_i, avec $d^\circ r_i < d^\circ d_{ii}$.

Corollaire :

La dimension de cet espace vectoriel est $p \times (\Sigma\, d^\circ d_{ii}) = p(d^\circ \det D)$.

Preuve :

Immédiate d'après la formule ci-dessus.

APPENDICE 2

A propos de la représentation de Führmann.

La preuve, figurant au §.4, que la représentation de Führmann en est bien une, est calculatoire et ne présente pas un lien très net avec les problèmes discutés aux §.1 et 2. Dans cet appendice, nous l'aborderons de façon plus abstraite pour mettre en évidence ce lien.

Supposons donnée une factorisation $H = ND^{-1}N_1$.

En reprenant les notations du §.2, il est clair que si $W = N_1^{-1}(DK^n[z])$ (N_1^{-1} signifie image réciproque). Alors Λ se factorise par $\frac{K^p[z]}{W}$ (ceci provient trivialement de ce que W est un sous-module du module V des entrées polynomiales engendrant des sorties polynomiales aussi). On a donc un diagramme commutatif :

$$\begin{tikzcd} & \frac{K^p[z]}{W} \arrow[dr, "\psi"] & \\ K^p[z] \arrow[ur, "\phi"] \arrow[rr, "\Lambda"] & & \frac{K^m[[z]]}{K^m[z]} \end{tikzcd}$$

avec ϕ surjection canonique et ψ induite par H (i.e. la multiplication par $ND^{-1}N_1$, puis la sélection de la partie propre de la série obtenue). Observons à présent que N_1 induit un homomorphisme injectif :

$$\frac{K^p[z]}{W} \xrightarrow{N_1} \frac{K^n[z]}{DK^n[z]}$$

Si on définit sur $\frac{K^n[z]}{DK^n[z]}$ l'homomorphisme induit par la multiplication par ND^{-1} puis la "prise de la partie propre", que nous notons θ, on a un nouveau diagramme commutatif :

$$\begin{tikzcd}
& \frac{K^p[z]}{W} \arrow[r, "N_1"] & \frac{K^n[z]}{DK^m[z]} \arrow[dr, "\theta"] & \\
K^p[z] \arrow[ur, "\phi"] \arrow[rrr, "\Lambda"'] & & & \frac{K^m[[z]]}{K^m[z]}
\end{tikzcd}$$

En d'autres termes, on peut relever ψ à $\frac{K^n[z]}{DK^n[z]}$, en identifiant $\frac{K^p[z]}{W}$ à son image par N_1.

Il est donc établi que $\frac{K^n[z]}{DK^n[z]}$ peut toujours être module d'état pour H. Il est alors trivial que la représentation construite est celle de Führmann.

On peut noter les conditions pour que celle-ci soit minimale :

- $N_1 \circ \phi$ surjective est équivalent à N_1 surjective puisque ϕ l'est par définition. Cela s'écrit : $N_1 K^p[z] + DK^n[z] = K^n[z]$, c'est-à-dire N_1 et D premières entre elles à gauche (cf.§.4).

- θ injective s'écrit :

($ND^{-1}x$ polynomial) $\Leftrightarrow$ ($D^{-1}x$ polynomial) dans $K^n[z]$. D'après un résultat du §.5, cela signifie exactement, D et N premières entre elles à droite.

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique