RAPPORT TECHNIQUE

Some algorithms for matrix polynomials

G. Villard

RT157 Part 1 **Mars 96**

Résumé

Poursuivant une série d'articles sur le calcul de formes normales de matrices, nous nous intéressons ici à la complexité parallèle du calcul de transformations associées. Pour une matrice B de $\mathcal{M}_{n,n}(K)$, où K est un corps commutatif, nous démontrons que le problème de calculer une matrice de transformation P telle que $F = P^{-1}BP$ soit sous forme normale de Frobenius, est dans $\mathcal{NC}_K^2$. Nous établissons ensuite un fait analogue pour le calcul de matrices unimodulaires $U(x)$ et $V(x)$ telles que, étant donnée $A(x)$ dans $\mathcal{M}_{n,m}(K[x])$, $S(x) = U(x)A(x)V(x)$ soit sous forme normale de Smith. Pour des corps K concrets tels que le corps de rationnels ou les corps de Galois, ces deux problèmes sont dans $\mathcal{NC}^2$.
Par synthèse avec certains précédents résultats, nous avons donc démontré que les problèmes de calculer une réduction à la forme de Jordan sur une extension algébrique, et des réductions aux formes de Frobenius et de Smith sur le corps initial, sont tous dans $\mathcal{NC}_K^2$. En corollaire nous donnons aussi le premier algorithme séquentiel polynomial pour le calcul de la forme normale de Smith sur $K[x]$, valable pour un corps arbitraire.

Mots clefs : algorithmes parallèles, $\mathcal{NC}_K^2$, formes normales, transformations unimodulaires, transformations semblables.

FAST PARALLEL ALGORITHMS FOR MATRIX REDUCTION TO NORMAL FORMS

Gilles VILLARD *

Abstract. Continuing a series of articles on normal forms of matrices, we investigate fast parallel algorithms to compute the corresponding transformations. Given a matrix B in $\mathcal{M}_{n,n}(K)$, where K is an arbitrary commutative field, we establish that computing a similarity transformation P such that $F = P^{-1}BP$ is under Frobenius normal form can be done in $\mathcal{NC}_K^2$. Using a reduction to this first problem, a similar fact is then proved for the Smith normal form $S(x)$ of a polynomial matrix $A(x)$ in $\mathcal{M}_{n,m}(K[x])$; to compute unimodular matrices $U(x)$ and $V(x)$ such that $S(x) = U(x)A(x)V(x)$ can be done in $\mathcal{NC}_K^2$. We get that over concrete fields such as the rationals, these problems are in $\mathcal{NC}^2$.

Using our previous results we have thus established that the problems of computing transformations over a field extension for the Jordan normal form, and transformations over the input field for the Frobenius and the Smith normal form are all in $\mathcal{NC}_K^2$. As a corollary we establish a polynomial-time sequential algorithm to compute transformations for the Smith form over $K[x]$.

Keywords: Parallel algorithms, $\mathcal{NC}_K^2$, normal forms, unimodular matrices, similarity matrices.

1 Introduction

The classical problem of computing canonical form of matrices is widely addressed in the literature and has many applications in various areas. In this paper we deal with the Frobenius and the Jordan normal forms of matrices over a commutative field K, and with the Smith normal form over a ring $K[x]$ of univariate polynomials.

For the theoretical aspects the reader may refer to [18,6]. The forms are well understood if considered as giving informations about a module decomposition [12]. From a practical point of view, fast sequential and parallel algorithms are known to compute the normal forms themselves. The problems of computing the Frobenius and the Smith form are in class $\mathcal{P}$(sequential polynomial-time problems) and in classes $\mathcal{NC}_K$and $\mathcal{NC}$. We refer to [4] for the definitions of the boolean parallel complexity class $\mathcal{NC}$ and to [8] for the arithmetic parallel complexity class $\mathcal{NC}_K$.

Concerning the computation of transformations for the forms, the problems were not so well overcomed. Efficient probabilistic solutions have been given but no deterministic algorithm was known to compute a transformation for the Frobenius form in $\mathcal{NC}_K$. The same question was also open for the Smith form: how to compute unimodular transformations fast and deterministically in parallel? How to obtain transformations over $K[x]$ even for small fields? Our new results will strongly rely on previous results in [23,25] that will be often refered throughout the paper.

We will first focus on the problem of computing a transformation for the Frobenius form. We propose a reduction of this problem to the one of computing a transformation for the Jordan normal form and we use known solutions for this latter problem. This will establish a fast parallel deterministic algorithm.

Next we will compute transformations over $K[x]$ for the Smith normal form. We will see that the problem is somehow equivalent to the previous one. Indeed, we compute unimodular transformations from similarity transformations for an associated Frobenius

*LMC-IMAG, 46 Av. F. Viallet, F-38031 Grenoble Cedex, Gilles.Villard@imag.fr.

form and conversely. This will give a fast parallel and a polynomial-time sequential solution.

1.1 The Smith and Frobenius normal forms

The normal forms we deal with are from the following two theorems of Smith and Frobenius. In the following the identity matrix of dimension n will be denoted by I_n or by I if the dimension can be deduced from the context.

Definition 1.1 *A square matrix of polynomials in $K[x]$ is said to be unimodular if its determinant is a nonzero element in K. It is easily seen that a matrix is unimodular if and only if it has an inverse over $K[x]$.*

Theorem 1.1 *If $A(x)$ is a $n \times m$ matrix of polynomials of rank r, there exist unimodular matrices $U(x)$ of dimension n and $V(x)$ of dimension m such that the only nonzero entries of $S(x) = U(x)A(x)V(x)$ are certain diagonal ones, the $s_{i,i}(x)$ for $1 \leq i \leq r$, these latter are monic and $s_{i,i}(x)$ is a factor of $s_{i+1,i+1}(x)$ for $1 \leq i < r$. The diagonal form satisfying these conditions is unique.*

We call this unique form $S(x)$ the Smith normal form of $A(x)$. Further, σ polynomials among the $s_{i,i}(x)$, let for $r - \sigma + 1 \leq i \leq r$, are nonunit: we will call them the nontrivial invariant factors of $A(x)$ and denote them by $s_1(x), \ldots, s_\sigma(x)$. The square diagonal submatrix of $S(x)$ consisting of the σ nontrivial invariant factors will be denoted by $S_\sigma(x)$.

Theorem 1.2 *If B is a $n \times n$ matrix with entries in K, B is similar to a matrix F, $F = P^{-1}BP$, which is quasi-companion, the polynomials associated to the companion blocks are the nontrivial invariant factors of $xI - B$.*

This unique form F is called the Frobenius normal form of B.

1.2 Previous algorithms and limitations

We overview the known methods and algorithms before introducing our new approach.

The first polynomial-time algorithms to compute the *Frobenius form* has been independently proposed in [17] and in [22]. They are polynomial-time in the dimension of the matrix, and also in the coefficient lengths for concrete fields such as the field of the field of the rationals Q or GF_q, the finite field with q elements. These algorithms consist in elimination processes and are consequently highly sequential. They can be adjusted to compute an associated transformation matrix over K even for small fields [19]. The n steps of the elimination process can be avoided by randomization to give Monte-Carlo or Las Vegas algorithms. The key idea used in [21,10] is that a random construction of a transformation matrix leads with high probability to the form. This approach gives fast parallel probabilistic algorithms [10].

Algorithms to compute the *Smith normal form* has been first given for matrices of integers. Computing the diagonalization by repeated triangularizations of the matrix and of its transpose leads to polynomial-time algorithms [5,16]. The same approach has been proposed for polynomial matrices [15]; this bounds the degrees of the polynomials involved during the calculations, but seems to be inadequate to bound the coefficients of those polynomials in particular over the rational polynomials. The first polynomial-time algorithm to compute the Smith form over $Q[x]$ appeared in [13], it is based on the Chinese remainder algorithm. Then it has been established in [26] that the form and associated unimodular transformations can be computed over any ring $K[x]$, with coefficient lengths remaining polynomially bounded over $Q[x]$. A drawback of this latter method is that a field extension is needed if K has less than $2d\min\{n,m\}+1$ elements, where d is the degree of the polynomials in input, and consequently the unimodular transformations are not obtained over $K[x]$ but involve elements of the extension. As for the computation of the Frobenius form, randomization can remove the sequential iterations (successive triangularizations). This has been used in [13,14,24]. The key idea is somehow equivalent to the one used for the Frobenius form: after a random transformation of the input matrix only one triangularization is sufficient with high probability. This gives a probabilistic parallel solution for the problem and speed the sequential methods themselves.

All the above algorithms are sequential elimination processes that can be randomized to give fast sequential and parallel solutions. A different approach has been only recently used for algorithmic purposes. It is based on the following theorem about characteristic subspaces of a matrix [6].

Theorem 1.3 *Let B a $n \times n$ matrix with entries in K and let $\chi(x)$ be its characteristic polynomial. If $\chi_1(x)$ and $\chi_2(x)$ are relatively prime such that $\chi(x) = \chi_1(x)\chi_2(x)$ then*

$$K^n = \ker \chi_1(B) \oplus \ker \chi_2(B).$$

If $e_1, e_2, \ldots, e_{n_1}$ and $f_1, f_2, \ldots, f_{n_2}$ are bases for $\ker \chi_1(B)$ and $\ker \chi_2(B)$, let P be the $n \times n$ matrix which columns are the e_i's and the f_i's in this order, then

$$P^{-1}BP = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}$$

is quasi-diagonal, the blocks B_1 and B_2 are of dimension n_1 and n_2.

This fact shows us that provided a factorization of the characteristic polynomial is known, B can be brought into a corresponding special form. This approach is used in [1] for sequential algorithms over finite fields using an irreducible decomposition of the characteristic polynomial. But it is not clear that these algorithms can lead to fast parallel ones and it may be hard to factor fast in parallel.

The approach can also be used to develop fast algorithms over any fields as it has been simultaneously shown in [11] in sequential and in [23,25] for parallel aspects. Introducing the linear factors $x - \lambda_i$, $1 \leq i \leq l$, of $\chi(x)$ and using an arithmetic on algebraic numbers, the Jordan, the Frobenius and the Smith normal forms can be computed fast in parallel in a deterministic way. But it was still an open question, even with this approach, to compute transformation matrices over K for the Frobenius and the Smith form.

1.3 A new approach

Previous algorithms to compute the Frobenius or the Smith form fall into three main categories:

- *highly sequential algorithms* (iterative processes): the Frobenius form and transformations can be computed over K; there is no polynomial-time algorithm (in the dimension of the matrix and in the degree of its entries over an abstract field, and in the coefficient lengths over the rationals) working over any field K and obtaining the Smith form and transformations over $K[x]$;

- *probabilistic algorithms*: transformations for the Frobenius form are computed over

K; transformations for the Smith form are computed with elements in a field extension if K is too small;

- *fast parallel deterministic algorithms*: there is no algorithm that computes transformation matrices over K either for the Frobenius or the Smith normal form.

To answer the raised questions, we propose new reductions between problems. The paper is organized as follows: we begin at §2 with basic reminders about the Hermite normal form (echelon form of the input matrix) and the Jordan normal form which will be needed as intermediate forms. In addition, for computing the Smith form, two applications of the Hermite form will allow us to focus on square nonsingular matrices for the rest of the paper. Then at §3 we present an algorithm for computing transformations for the Frobenius form. We reduce the problem to the computation of transformation matrices for the Jordan form. This establishes that the problem $\mathcal{REDUCTION\ TO\ FROBENIUS\ FORM\ OVER}$ K (the normal form and a similarity transformation) is in $\mathcal{NC}_K^2$. Then at §4 and §5, considering K^n as a $K[x]$-module, we show how the problems of computing transformations for the Frobenius form over K and transformations for the Smith form over $K[x]$ can be viewed as equivalent. We solve the former from the latter and *vice versa*. From these results we derive an algorithm at §6 and show that the problem $\mathcal{REDUCTION\ TO\ SMITH\ FORM\ OVER}$ $K[x]$ (the normal form and unimodular transformations) is in $\mathcal{NC}_K^2$. This will also establish, as a corollary, a sequential polynomial-time algorithm which is independent of K.

2 Hermite and Jordan normal form: technical results

Two intermediate forms will play an important role: the Hermite normal form to compute the Smith one and the Jordan normal form to compute the Frobenius one.

2.1 Hermite normal form computation

Definition 2.1 *Two matrices $A(x)$ and $A'(x)$ of polynomials are said to be right (resp. left) equivalent if there exist a unimodular matrix $V(x)$ (resp. $U(x)$) such that $A(x) = A'(x)V(x)$ (resp. $A(x) = U(x)A'(x)$). If $A(x) = U(x)A'(x)V(x)$ then $A(x)$ and $A'(x)$ are said to be equivalent.*

The Hermite normal form has been originally developed for square matrices [18], it is triangular and unique if the matrix is nonsingular. We are going to follow a treatment found in [14] giving the canonical form for right or left equivalence of arbitrary matrices over a principal ideal domain. For arbitrary rectangle matrices, only an echelon form can be obtained, by extension we will also call it Hermite normal form.

Definition 2.2 *A matrix $H(x)$ of rank r is in (column) Hermite normal form if:*

- r nonzero columns precede zero columns,

- the tailing nonzero element in each column is monic and of row index strictly lower than the row indexes of the tailing elements of the following columns,

- in each row which contains the tailing nonzero element of some column, the entries following that entry are of lower degree.

With such a definition, each matrix in $\mathcal{M}_{n,m}(K[x])$ is right equivalent to a unique matrix under Hermite normal form; associated unimodular transformations are not unique. If $H(x)$ is square and nonsingular it is upper-triangular. Fast parallel algorithms to compute the form are given in [13,27].

Theorem 2.1 *The Hermite normal form $H(x)$ of a polynomial matrix $A(x)$ and a unimodular matrix $R(x)$ such that $A(x) = H(x)R(x)$ can be computed in $\mathcal{NC}^2_K$. Over the rationals or finite fields GF_q this can be done in $\mathcal{NC}^2$.*

Transposing everything in the above, we can define a left Hermite normal form with similar properties. With the following corollary we will restrict ourselves to the case of nonsingular square matrices.

Corollary 2.1 *Each matrix $A(x)$ in $\mathcal{M}_{n,m}(K[x])$ is equivalent to a matrix*

$$\begin{bmatrix} I & 0 & 0 \\ 0 & H_\tau(x) & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{1}$$

where $H_\tau(x)$ is a square nonsingular matrix of dimension τ under Hermite normal form and which diagonal entries are at least of degrees 1. One can compute form (1) and associated left and right unimodular multipliers in $\mathcal{NC}^2_K$ ($\mathcal{NC}^2$ over Q and GF_q).

Proof. Let r be the rank of $A(x)$. Applying two times theorem 2.1 we compute the left Hermite normal form of $A(x)$ and the Hermite normal form of this latter, we get:

$$A(x) = U_1(x) \begin{bmatrix} H_r(x) & 0 \\ 0 & 0 \end{bmatrix} V_1(x)$$

where $H_r(x)$ is square of rank r under Hermite normal form. Further, $H_1(x)$ has non zero diagonal entries with say $r-\tau$ unit ones and in each row which contains a unit element on the diagonal, all the other entries are zero. By elementary row operations every entries above a unit diagonal one can be zeroed. Then by row and column permutations we can construct an identity submatrix with the $r - \tau$ unit diagonal entries, the other entries give $H_\tau(x)$ and the form (1) is obtained. □

2.2 Jordan normal form computation

We now consider the parallel construction of the Jordan normal form of a matrix B in $\mathcal{M}_{n,n}(K)$.

Definition 2.3 *Each matrix B in $\mathcal{M}_{n,n}(K)$ is similar to a unique (up to permutation) quasi-diagonal matrix J which blocks are banded matrices $\mathcal{J}_k(\lambda_i)$ in $\mathcal{M}_{k,k}(K)$ of the form:*

$$\begin{bmatrix} \lambda_i & 1 & 0 & \ldots \\ 0 & \lambda_i & \ddots & 0 \\ \vdots & & \ddots & 1 \\ 0 & \ldots & 0 & \lambda_i \end{bmatrix}$$

where λ_i, $1 \leq i \leq l$, is an eigenvalue of B. This form is called the Jordan form of B.

From [23] we know that the form can be computed fast in parallel. More precisely, we can compute a symbolic Jordan form in $\mathcal{M}_{n,n}(K[\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l])$: this form gives the structure of J with indeterminates $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l$ that take the place of the distinct eigenvalues. Each indeterminate $\tilde{\lambda}_i$ is associated with a polynomial $\Lambda_i(x)$ in $K[x]$, with the understanding that Λ_i is a representation of the corresponding eigenvalue λ_i, *i.e.* $\Lambda_i(\lambda_i) = 0$; $\Lambda_i(x)$ is a *generalized eigenvalue.* The Λ_i's are divisors of the characteristic polynomial of B. Clearly, this symbolic form is not unique, different choices are possible for the Λ_i's. Following [14,23] we only need to distinguish between eigenvalues having Jordan blocks with different structures. We are going to consider symbolic Jordan form corresponding to Λ_i's such that: if there is a dimension k such that λ_i and λ_j do not have the same number of Jordan blocks of dimension k then Λ_i and Λ_j are relatively prime, otherwise the representations are the same.

Theorem 2.2 *The problem of computing a symbolic Jordan form in $\mathcal{M}_{n,n}(K[\tilde{\lambda}_1,\ldots,\tilde{\lambda}_l])$ of a matrix B in $\mathcal{M}_{n,n}(K)$ is in $\mathcal{NC}_K^2$. Over the rationals or finite fields GF_q the problem is in $\mathcal{NC}^2$. The indeterminates are associated to generalized eigenvalues, $\Lambda_i(x)$, that are equal if and only if the corresponding eigenvalues have the same Jordan structure and are relatively prime otherwise.*

Proof. Up to the time complexity, this is theorem 7 of [23]. As shown there, the property on the $\Lambda_i(x)$'s is satified by computing a gcd-free basis. Using the algorithm in [9] this is done in $O(\log^2 n)$ arithmetic or boolean steps using polynomially many processors. □

3 A transformation for the Frobenius form

From the previous theorem concerning the Jordan form, we are going to develop an algorithm to compute a transformation P for the Frobenius form. At this point, we may emphasized that it seems hard to compute P over K fast in parallel directly, *i.e.* without using the Jordan form, for an arbitrary field. Even though this is probably easy if K is such that polynomial factorization is in $\mathcal{NC}_K$. We will need the following result of [25] to compute the form itself.

Theorem 3.1 *The problem of computing the Frobenius normal form F in $\mathcal{M}_{n,n}(K)$ of a matrix B in $\mathcal{M}_{n,n}(K)$ is in $\mathcal{NC}_K^2$. Over the rationals or finite fields GF_q the problem is in $\mathcal{NC}^2$.*

Let us give an idea of our method. Using at first a classical construction [28], we obtain a transformation matrix involving the eigenvalues (indeterminates representing them), then a transformation over the input field is computed. For the former matrix, we begin by computing a transformation L for the Jordan normal form of B (lemma 3.1 and lemma 3.2), $J = L^{-1}BL$, then we consider a transformation M from the Jordan form to the Frobenius form (lemma 3.3), $F = M^{-1}JM$, and finally we will take $P = LM$. We will have to ensure that P can be obtained over K. This will rely on the particular structure of L and on that of M.

The matrix L is computed *conformable to J i.e.* the structure of L and the way the $\tilde{\lambda}_i$ appear matches the structure of J [11,22]. Each Jordan block of dimension k of J is associated to k columns in L, the corresponding vectors form a so called *Jordan chain of length k* [6].

Definition 3.1 *Let L in $\mathcal{M}_{n,n}(K[\tilde{\lambda}_1,\ldots,\tilde{\lambda}_l])$ be a transformation from B to its symbolic Jordan normal form, L is said to be conformable to J if:*

- to each blocks $\mathcal{J}_k(\tilde{\lambda}_i)$ composed of the columns indexed, for a fixed j_0, from j_0+1 to j_0+k in J correspond the columns indexed from j_0+1 to j_0+k in L; the corresponding k column vectors form a Jordan chain of length k;

- these columns of L depends only on $\tilde{\lambda}_i$: their entries are elements of $K[\tilde{\lambda}_i]$;

- if two indeterminates $\tilde{\lambda}_{i_1}$ and $\tilde{\lambda}_{i_2}$ are associated to the same generalized eigenvalue

$\Lambda_i(x)$*, the entries in the columns of* L *corresponding to the blocks* $\mathcal{J}_k(\tilde\lambda_{i_1})$ *and the entries in the columns corresponding to* $\mathcal{J}_k(\tilde\lambda_{i_2})$ *are polynomials with the same coefficients.*

Example 3.1 Let the Jordan normal form of B be composed of two blocks of dimension two and of two blocks of dimension one: $J = \operatorname{diag}(\mathcal{J}_2(\lambda_2), \mathcal{J}_2(\lambda_2), \lambda_3, \lambda_4)$. Then $L(\tilde\lambda_1, \tilde\lambda_2, \tilde\lambda_3, \tilde\lambda_4)$ given by

$$\begin{bmatrix} -1+4\,\tilde\lambda_1 & 23/2+9\,\tilde\lambda_1 & -1+4\,\tilde\lambda_2 & 23/2+9\,\tilde\lambda_2 & -28\,\tilde\lambda_3-8 & -28\,\tilde\lambda_4-8 \\ \tilde\lambda_1-1 & -25/2-9\tilde\lambda_1/2 & -5\,\tilde\lambda_2-1 & -25/2-9\,\tilde\lambda_2/2 & -4-12\,\tilde\lambda_3 & -4-12\,\tilde\lambda_4 \\ \tilde\lambda_1 & 8+11\,\tilde\lambda_1/2 & 3\,\tilde\lambda_2 & 8+11\,\tilde\lambda_2/2 & -4-12\,\tilde\lambda_3 & -4-12\,\tilde\lambda_4 \\ \tilde\lambda_1 & -57-30\,\tilde\lambda_1 & -18\,\tilde\lambda_2 & -57-30\,\tilde\lambda_2 & -8\,\tilde\lambda_3 & -8\,\tilde\lambda_4 \\ 0 & 3 & 0 & 3 & 4\,\tilde\lambda_3 & 4\,\tilde\lambda_4 \\ \tilde\lambda_1 & -29/2-9\,\tilde\lambda_1 & 1-4\,\tilde\lambda_2 & -29/2-9\,\tilde\lambda_2 & -4\,\tilde\lambda_3 & -4\,\tilde\lambda_4 \end{bmatrix}$$

is conformable to J. Two blocks of columns of L which entries are polynomials in $\tilde\lambda_1$ and in $\tilde\lambda_2$ respectively, give the Jordan chains of length 2 associated to the Jordan blocks of dimension 2 in J. The two indeterminates are associated to a unique generalized eigenvalue $\Lambda_{1,2}(x)$. The two other indeterminates $\tilde\lambda_3$ and $\tilde\lambda_4$ each give one column in L, they are associated to a unique generalized eigenvalue $\Lambda_{3,4}(x)$. □

As a corollary of theorem 2.2, using nullspace computations over algebraic numbers [23], an associated transformation matrix represented by a matrix $L(\tilde\lambda_1, \ldots, \tilde\lambda_l)$ in $\mathcal{M}_{n,n}(K[\tilde\lambda_1, \ldots, \tilde\lambda_l])$ can also be computed. The following classical results are derived from [6]. Let $B^*(x)$ denote the reduced adjoint matrix of $x - B$:

$$(x-B)^{-1} = B^*(x)/\psi(x)$$

where $\psi(x)$ is the minimum polynomial of B. To construct the Jordan chains giving the columns of L, one can take, up to constants $1/k!$, certain linear combinations of the column vectors of the successive derivatives of $B^*(x)$. To ensure the algorithm works over any field, as done in [7] for the squarefree decomposition, we define:

$$(x^n)^{[k]} = C_n^k x^{n-k}, \quad n, k \geq 0.$$

By linearity, this gives a mapping $K[x] \to K[x]$ such that for any polynomial $a(x)$,

$$(a(x))^{[k]} = \frac{1}{k!}\frac{d^k a}{dx^k}, \quad k! \neq 0. \tag{2}$$

Lemma 3.1 *For each* B *in* $\mathcal{M}_{n,n}(K)$ *a transformation matrix* $L(\tilde\lambda_1, \ldots, \tilde\lambda_l)$ *for the symbolic jordan form, conformable to* J*, can be computed in* $\mathcal{NC}^2_K$*. Over the rationals or finite fields the problem is in* $\mathcal{NC}^2$*.*

Proof. We refer to [6] for this construction. Let $[j_1, j_2, \ldots, j_k]$ be k consecutive columns of a Jordan block $\mathcal{J}_k(\lambda_i)$ of dimension k of J. The corresponding columns $[l_1, l_2, \ldots l_k]$ of L constitute a Jordan chain of length k associated to λ_i:

$$(B - \lambda_i I) l_1 = 0, \quad (B - \lambda_i I) l_i = l_{i-1}, \ 2 \leq i \leq k. \tag{3}$$

We first focus on the computation of the Jordan chains of length k for any given k and λ_i, thus working over the algebraic extension $K(\lambda_i)$. Then we will see that symbolically the transformation matrix can be computed as announced.

If the successive transforms of the adjoint $B^*(x)$, using (2) entry-wise, are denoted by $B^{[k]}$:

$$B^{[k]}(x) = (B^*(x))^{[k]}, \quad k \geq 0 \tag{4}$$

we know [6] that

$$(B - \lambda_i I)B^{[0]}(\lambda_i) = 0, \quad (B - \lambda_i I)B^{[i]}(\lambda_i) = B^{[i-1]}(\lambda_i), \; 2 \leq i \leq k. \tag{5}$$

We denote by κ the length of the longest Jordan chain associated to λ_i.

We first assume $k < \kappa$, $k = \kappa$ will be a particular case of this general situation. Let r_{k+1} be the rank of $B^{[\kappa-k-1]}(\lambda_i)$ and let $N_k(\lambda_i)$ be an invertible matrix such that the first r_{k+1} columns of $B^{[\kappa-k-1]}(\lambda_i)N_k(\lambda_i)$ are linearly independent and the others are null. From (5) the last $n - r_{k+1}$ columns of $B^{[\kappa-k]}(\lambda_i)N_k(\lambda_i)$ are eigenvectors. They belong to chains of lengths greater than k [6]. Now, we isolate the one belonging to chains of length k exactly. If r_k denotes the rank of $B^{[\kappa-k]}$, they are $r_k - 2r_{k+1}$ such eigenvectors. They are found by computing a maximal linearly independent set of vectors in the span of the last $n - r_{k+1}$ column ones of $B^{[\kappa-k]}(\lambda_i)N_k(\lambda_i)$ and linearly independent with respect to the first r_k columns of $B^{[\kappa-k-1]}(\lambda_i)N_k(\lambda_i)$. We assume that this is done by computing a transformation matrix $\tilde{N}_k(\lambda_i) = N_k(\lambda_i)M_k(\lambda_i)$ such that the columns $r_{k+1}+1, \ldots, r_k - r_{k+1}$ of $B^{[\kappa-k]}(\lambda_i)\tilde{N}_k(\lambda_i)$ satisfy this property. Once these $n_k = r_k - 2r_{k+1}$ eigenvectors belonging to the chains of length k are obtained, the chains themselves are easily derived. Indeed, let l_1 be one of these vectors, being the column c_1 of $B^{[\kappa-k]}(\lambda_i)\tilde{N}_k(\lambda_i)$, then the remaining vectors $[l_2, \ldots, l_k]$ of the associated chain (given by (3)) are the c_1-th columns of $B^{[\kappa-j]}(\lambda_i)\tilde{N}_k(\lambda_i)$, $k-1 \leq j \leq 1$. The set of the chains constructed this way for all lengths k, $1 \leq k \leq n$ and all eigenvalues λ_i, $1 \leq i \leq l$, give the columns of a transformation matrix from B to its Jordan form [6]. For $k = \kappa$, the computation is a particular case of the above construction. The eigenvectors are directly found to form a maximal linearly independent set of columns of $B^{[0]}(x) = (B^*(x))^{[0]} = B^*(x)$. We compute them using $\tilde{N}_\kappa(x)$ such that the first r_κ columns of $B^{[0]}(x)\tilde{N}_\kappa(x)$ are linearly independent. The rest of the chains is deduced as previously.

We now perform the computation symbolically. The Jordan chains are going to be given by polynomials with equal coefficients for all the roots of a given generalized eigenvalue $\Lambda_i(x)$. Chains can be constructed for all k and i simultaneously, we restrict ourselves to a given k and a given $\Lambda_i(x)$. By definition of the $\Lambda_i(x)$'s, all their roots are eigenvalues with the same Jordan structure, the ranks r_{k+1} and r_k are also independent of the choice of the eigenvalue represented by $\Lambda_i(x)$. We may thus apply directly a parallel arithmetic on algebraic numbers as introduced in [23]. The eigenvalues are represented as polynomials in $K[x]/(\Lambda_i(x))$. The matrices $B^{[k]}(x)$ given by (4) are computed in parallel modulo $\Lambda_i(x)$. Then we apply proposition 4 in [23] (for maximal linearly independent set of columns and nullspace computation over algebraic numbers) to compute $N_k(x)$ and $M_k(x)$, with the understanding that $\tilde{N}_k(x)$ is a suitable matrix for all the roots of $\Lambda_i(x)$. The $r_k - 2r_{k+1}$ target eigenvectors are read off $B^{[k]}(x)\tilde{N}_k(x)$. For the rest of the chains, we simply pick up the corresponding columns in $B^{[\kappa-j]}(x)\tilde{N}_k(x)$, $k-1 \leq j \leq 1$. Subsituting the symbol x by the symbols of the symbolic eigenvalues $\tilde{\lambda}_j$ whcich representation is Λ_i we get the associated blocks of columns in $L(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l)$. For all the roots of a given generalized eigenvalue, the entries of the vectors of the associated chains are polynomials with the same coefficients since this true by construction [23] for the matrices $N_k(x)$ and $M_k(x)$. The matrix $L(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l)$ has been computed conformable to the Jordan form. Finally, the announced complexity is valid since it holds for matrix product, for a maximal linearly independent set of columns and for the nullspace [23]. □

We give a complement to this lemma that will be useful to show that the target transformation $P = LM$ will be actually computed over K (theorem 3.2). This will be proven using Newton's identities [12] over any fields. The only complication requiring some extra care will concern computations over fields of characteristic p, $p > 0$.

Lemma 3.2 *Let K be a field of characteristic p, $p > 0$. Let B and L be as in lemma 3.1 and assume $\Lambda_i(x)$ to be a generalized eigenvalue representing eigenvalues associated to Jordan blocks $\mathcal{J}_{kp^\alpha}(\lambda_j)$ ($k \geq 1$ and $\alpha \geq 0$) which dimensions are only multiples of p^α. In particular we assume that $\Lambda_i(x)$ is a polynomial in x^{p^α}* i.e. *can be written as $\bar{\Lambda}_i(x^{p^\alpha})$. Then, the entries in $L(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l)$ of the ends of chain (last vectors of the Jordan chains) associated to $\Lambda_i(x)$, are polynomials in x^{p^α}.*

Proof. This is a direct consequence of the previous proof. Let us compute the last vectors of the chains of length kp^α; again, the longest chains with $kp^\alpha = k_m p^\alpha = \kappa$ will be particular cases. The ends of chain are columns of

$$E_k = B^{[\kappa-1]}(x) N_{kp^\alpha}(x) M_{kp^\alpha}(x).$$

It is sufficient to prove the claim of the lemma respectively for the three matrices of the above right-hand term.

First consider $B^{[\kappa-1]}(x)$. We view it as a matrix polynomial (of degree $d-1$ if the minimum polynomial $\psi(x)$ is of degree d):

$$\begin{array}{rcl} B^{[\kappa-1]}(x) & = & (B^*(x))^{[\kappa-1]} = (x^{d-1})^{[\kappa-1]} + B_{d-2}(x^{d-2})^{[\kappa-1]} + \ldots + B_{\kappa-1} \\ & = & C_{d-1}^{\kappa-1} x^{d-\kappa} + C_{d-2}^{\kappa-1} B_{d-2} x^{d-\kappa-1} + \ldots + B_{\kappa-1}. \end{array} \qquad (6)$$

But since K is of characteristic p, $C_j^{\kappa-1} = C_j^{k_m p^\alpha - 1}$ is equal to zero if $j \neq k' p^\alpha - 1$ for some integer k'. If $j = k' p^\alpha - 1$ then $j - (\kappa - 1) = (k' - k_m) p^\alpha$ and only such powers may appear in (6).

Now, concerning $M_{kp^\alpha}(x)$. This requires to very briefly describe the procedures in [3,23] to compute a maximal linearly independent set of columns. Over an abstract field, for a set of columns $B = (B_1, \ldots, B_n)$, a maximal set is constructed by taking the columns j such that rank $(B_1, \ldots, B_{j-1}) <$ rank $(B_1, \ldots, B_j)$. Over algebraic numbers, this implies some extra work to ensure that the corresponding transformation matrix $M_{kp^\alpha}(x)$ is independent of the root of Λ_i (and further that L is conformable to J). Indeed, even if the rank of B is the same for all the roots, the choice of the columns may depend on them. We refer to [23] for a satifying solution that consists in weighting the columns by suitable factors of the generalized eigenvalue. These latter factors can be chosen polynomials in x^{p^α} as Λ_i is. This yields a matrix $M_{kp^\alpha}(x)$ that satifies the claimed property.

Finally, for $N_{kp^\alpha}(x)$. The nullspace of a matrix can be obtained [3] by computing a maximal linearly independent set of columns, then a maximal linearly independent set of rows and by inverting the corresponding submatrix. As indicated in lemma 3.1, to compute a chain of length kp^α, this operation is done on $B^{[\kappa - kp^\alpha - 1]}(x)$ ($k < k_m$). For the same reason than for matrix (6) only p^α-th powers of x appear in $B^{[\kappa - kp^\alpha - 1]}(x)$. The property is preserved by matrix inversion and is true for $N_{kp^\alpha}(x)$. For the longest chains, we just take $\tilde{N}_\kappa(x) = M_\kappa(x)$ and the property also holds. □

From [28,22] we now compute a transformation between a companion matrix and its Jordan normal form. This will be applied to sub-blocks of the companion blocks of the target Frobenius form.

Lemma 3.3 *Let C_χ be a $n \times n$ companion matrix with characteristic polynomial $\chi(x)$. If $\chi(x)$ is a q-th power of a squarefree polynomial of degree l (C_χ has l distinct eigenvalues with the same multiplicity) then the Jordan form of C_χ is*

$$J = diag\left(\mathcal{J}_q(\lambda_1), \mathcal{J}_q(\lambda_2), \ldots, \ldots, \mathcal{J}_q(\lambda_l)\right).$$

A transformation matrix M in $\mathcal{M}_{n,n}(K[\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l])$ from J to C_χ is the $n \times n$ matrix

$$(M_q(\tilde{\lambda}_i))_{i=1}^{l}$$

where

$$M_q(x) = \begin{bmatrix} 0 & \ldots & 0 & 1 & \ldots & C_{q-1}^{n-2}x^{n-q-1} & C_{q-1}^{n-1}x^{n-q} \\ \vdots & & & & & & \\ 0 & 1 & 2x & 3x^2 & \ldots & C_1^{n-2}x^{n-3} & C_1^{n-1}x^{n-2} \\ 1 & x & x^2 & x^3 & \ldots & x^{n-2} & x^{n-1} \end{bmatrix} \in M_{q,n}(K[x]). \quad (7)$$

We now prove the main fact of this section by first reducing the general case to the situation of lemma 3.3, where the Frobenius form is given by a unique invariant factor, power of a squarefree polynomial.

Theorem 3.2 *Let K be a commutative field. The problem of computing the Frobenius normal form F of a matrix B in $\mathcal{M}_{n,n}(K)$ and a similarity transformation P in $\mathcal{M}_{n,n}(K)$ such that $F = P^{-1}BP$ is in $\mathcal{NC}_K^2$. Over the rationals or finite fields GF_q the problem is in $\mathcal{NC}^2$.*

Proof. The set of the invariant factors of $xI - B$ *i.e.* the Frobenius form of B is pre-computed using theorem 3.1. If F has σ companion blocks, let the Jordan blocks of J be numbered by increasing dimensions, considering that some blocks are of dimension 0 to have exactly σ blocks for each eigenvalue λ_i, $1 \leq i \leq l$. We denote these blocks by $\mathcal{J}^{(j)}(\lambda_i)$, $1 \leq i \leq l$ and $1 \leq j \leq \sigma$. For any fixed j, $1 \leq j \leq \sigma$, the companion block C_{s_j} of F associated to the j-th nontrivial invariant factor $s_j(x)$ is constructed from the blocks $\mathcal{J}^{(j)}(\lambda_i)$, $1 \leq i \leq l$. Thus each block C_{s_j} is computed independently of the others from J; in addition, by lemma 3.1, the transformation matrix L can be splitted into blocks of columns corresponding to each $\mathcal{J}^{(j)}(\lambda_i)$ and C_{s_j}. Consequently, we can restrict ourselves to the computation of one of the blocks of F, let this block be the j_0-th one and be denoted by C_s with characteristic polynomial $s(x) = s_{j_0}(x)$ the j_0-th nontrivial invariant factor. By definition, each distinct eigenvalue of B is associated with a unique Jordan block in C_s.

To use the particular structure of conformable matrices we split $s(x)$ with respect to the generalized eigenvalues $\Lambda_i(x)$'s computed by theorem 2.2. This leads to compute a quasi-companion matrix $\bar{F}$ more refined than the Frobenius form. Each block of $\bar{F}$ will involve eigenvalues belonging to the same generalized eigenvalues and thus leading to the same dimension of Jordan block in C_s as required to apply lemma 3.3. From a transformation for $\bar{F}$, a transformation for F will be easily computed. If there is d distinct $\Lambda_i(x)$ we split $s(x)$ in d factors:

$$s(x) = \chi_1(x)\chi_2(x)\ldots\chi_d(x)$$

where

$$\chi_i(x) = \gcd\left(s(x), \Lambda_i(x))^n\right), \; 1 \le i \le d. \tag{8}$$

All the roots of $\chi_i(x)$, $1 \le i \le d$, are eigenvalues of B with the same Jordan structure. In addition, from identity (8), if it is nontrivial then $\chi_i(x)$ has the same roots than $\Lambda_i(x)$ all with multiplicity the dimension, say q, of the corresponding unique Jordan block:

$$\chi_i(x) = \Pi_j \left(x - \lambda_j^{(i)}\right)^q$$

the product being taken on the roots $\lambda_j^{(i)}$ of $\chi_i(x)$. If we compute a transformation from B to the quasi-diagonal matrix $\bar{F}$ composed of the companion blocks associated to the $\chi_i(\lambda)$

$$\bar{F} = \mathrm{diag}\left(C_{\chi_1}, \ldots, C_{\chi_d}\right)$$

a transformation matrix for F is readily obtained: as cyclic vector it suffices to take the sum of the cyclic vectors giving the blocks C_{χ_i}. Since L can be splitted into blocks of columns corresponding to each C_{χ_i} (lemma 3.1), we can finally restrict ourselves to the case where F is a companion block associated to a $\chi_i(x)$. From now on, we may thus assume that F is a companion matrix of dimension n and that its characteristic polynomial is a polynomial $\chi(x)$ which roots have the same multiplicity q. Let $l' = n/q$ denote the number of distincts eigenvalues of F. The current generalized eigenvalue is $\Lambda_i(x)$.

We apply lemma 3.1 and lemma 3.3 to compute transformation matrices L and M in $\mathcal{M}_{n,n}(K[\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{l'}])$. A symbolic transformation $\tilde{P}$ from B to F is

$$\tilde{P} = LM = \left[\begin{array}{c|c|c} L_q(\tilde{\lambda}_1) & \ldots & L_q(\tilde{\lambda}_{l'}) \end{array}\right] \left[\begin{array}{c} M_q(\tilde{\lambda}_1) \\ \hline \ldots \\ \hline M_q(\tilde{\lambda}_{l'}) \end{array}\right] \tag{9}$$

where the $n \times q$ matrix $L_q(x)$ is formed by a Jordan chain of length q and where the $q \times n$ matrix $M_q(x)$ is given by identity (7). This construction gives a matrix $\tilde{P}$ in $\mathcal{M}_{n,n}(K[\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{l'}])$, it remains to deduce a matrix P in $\mathcal{M}_{n,n}(K)$. Actually, substituting the symbols in $\tilde{P}$ by the eigenvalues leads to such a matrix P in $\mathcal{M}_{n,n}(K)$. Indeed, the entries of $\tilde{P}$ are polynomials of $K[\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{l'}]$; simplifying them using $\chi(x)$ ($\chi(\tilde{\lambda}_i) = 0$) and using (9) we get homogeneous polynomials of the form

$$\tilde{P}_{ij} = p_{ij}^{(0)} + p_{ij}^{(1)}(\tilde{\lambda}_1 + \ldots + \tilde{\lambda}_{l'}) + p_{ij}^{(2)}(\tilde{\lambda}_1^2 + \ldots + \tilde{\lambda}_{l'}^2) + \ldots + p_{ij}^{(n-1)}(\tilde{\lambda}_1^{n-1} + \ldots + \tilde{\lambda}_{l'}^{n-1}) \tag{10}$$

where the $p_{ij}^{(k)}$'s are constants in K. The indeterminates $\tilde{\lambda}_i$ stands for the eigenvalues λ_i, consequently in the algebraic closure of K, the matrix P with entries

$$P_{ij} = p_{ij}^{(0)} + p_{ij}^{(1)}(\lambda_1 + \ldots + \lambda_{l'}) + p_{ij}^{(2)}(\lambda_1^2 + \ldots + \lambda_{l'}^2) + \ldots + p_{ij}^{(n-1)}(\lambda_1^{n-1} + \ldots + \lambda_{l'}^{n-1}) \tag{11}$$

is a transformation matrix for the Frobenius form. We denote by Σ_k the power sums of the distinct eigenvalues: $\Sigma_k = \lambda_1^k + \ldots + \lambda_{l'}^k, \;\; 1 \le k \; \le n-1$.

Different cases arise depending on q. Firstly, if $q \neq 0$ in K – either K is of characteristic $p = 0$ or $\gcd(q,p) = 1$ – then the Σ_k's are elements of K. They can be computed as $\Sigma_k = \bar{\Sigma}_k / q$ where the $\bar{\Sigma}_k$'s are the power sums of the zeros of $\chi(x)$ (themselves obtained by Newton's identities [12]). And, as claimed, P is a matrix in $\mathcal{M}_{n,n}(K)$:

$$P_{ij} = p_{ij}^{(0)} + p_{ij}^{(1)}\Sigma_1 + \ldots + p_{ij}^{(n-1)}\Sigma_{n-1} \; \in K. \tag{12}$$

Secondly, if the division by q is not allowed. We are going to use lemma 3.2. Indeed, we can restrict ourselves to $\tilde{P}_1 = \{\tilde{P}_{i1}\}$, the first column vector of $\tilde{P}$: we know the j-th column vector is computed from the first as $\tilde{P}_j = B^{j-1}\tilde{P}_1$ since we build a transformation for a companion block. Now, from the particular form of M (lemma 3.3), $\tilde{P}_1$ is the sum of the ends of Jordan chains of length q. It is thus sufficient to prove that the entries of this sum are elements of K. We show that only selected power sums appear and that we can bypass the problem of the division by q.

Let $q = \rho p^\beta$ and β be maximal *i.e.* $1 \le \rho < p$. If the assumption of lemma 3.2 is true with $\alpha = \beta$, only multiples of p^β appear as exponents in relations (10) and (11) for the $\tilde{P}_{i1}$'s. The involved power sums are the Σ_{kp^β}'s. In a way analogous than in the regular case ($q \neq 0$), we first compute power sums $\bar{\Sigma}_{kp^\beta}$ with multiplicity ρ from

$$\chi_{p^\beta}(x) = \chi(x^{1/p^\beta}), \tag{13}$$

and the Σ_{kp^β}'s are derived by division by ρ. If the assumption of lemma 3.2 is not true with $\alpha = \beta$, let $\bar{\beta}$ be the greatest integer sucht that the assumption is true. For some other invariant factor $\bar{s}(x)$ involving the same eigenvalues than $\chi(x)$ – but with a lower multiplicity $\bar{q}$ – we are led to the previous situation. We have $\bar{q} = \bar{\rho}p^{\bar{\beta}}$ with $\bar{\beta}$ maximal. Thus, with $\bar{s}(x)$ and

$$\bar{\chi}(x) = \gcd\left(\bar{s}(x), \Lambda_i(x))^n\right),$$

the involved power sums can be computed as done above using relation (13), and used as in (12) to show that the sums of the ends of chain and thus P have entries in K.

To conclude the proof we have to verify that computations can be done fast in parallel. We first get the generalized eigenvalues $\Lambda_i(x)$'s, the invariant factors of $B - xI$ from theorems 2.2 and 3.1. Then using (8) we simultaneously compute the $\chi_i(x)$ and solve the problem for each corresponding companion blocks. The transformation $\tilde{P}$ is obtained using lemmata 3.1 and 3.3. It remains to compute the power sums. In the regular case, the sums with multiplicities are computed by Newton's identities as shown in [2]. For the general case, one can first search for the greatest exponent $\bar{\beta}$ to apply lemma 3.2. This clearly be done in $O(\log n)$ polynomial divisions between the $\chi_i(x)$'s having the same roots but with different multiplicities. For concrete fields such as the rationals or finite fields, the problem is in $\mathcal{NC}^2$since the algorithm is a fixed number of solutions of problems in $\mathcal{NC}^2$. □

Example 3.2 We take the data of example 3.1 and follow the proof above. There are two generalized eigenvalues: $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ each give one block with $J_1^{(2)}(x) = (x^2+x+1)^2$; $\tilde{\lambda}_3$ and $\tilde{\lambda}_4$ each give one block with $J_1^{(1)}(x) = x^2 + 3x + 1$. The Frobenius form has one block C_{s_1} with $s_1(x) = (x^2 + x + 1)^2(x^2 + 3x + 1)$. Thus at first, we compute a transformation for the matrix $\bar{F}$ that has two companion blocks corresponding to the polynomials: $\chi_1(x) = J_1^{(2)}(x)$, $\chi_2(x) = J_1^{(1)}(x)$. A transformation matrix L for the Jordan form has been given previously. For M, using lemma 3.3 we take:

$$\begin{bmatrix} 0 & 1 & 2\tilde{\lambda}_1 & 3\tilde{\lambda}_1^2 & 0 & 0 \\ 1 & \tilde{\lambda}_1 & \tilde{\lambda}_1^2 & \tilde{\lambda}_1^3 & 0 & 0 \\ 0 & 1 & 2\tilde{\lambda}_2 & 3\tilde{\lambda}_2^2 & 0 & 0 \\ 1 & \tilde{\lambda}_2 & \tilde{\lambda}_2^2 & \tilde{\lambda}_2^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \tilde{\lambda}_3 \\ 0 & 0 & 0 & 0 & 1 & \tilde{\lambda}_4 \end{bmatrix}$$

Thus a transformation for $\bar{F}$ (we omit the third and the fourth columns which entries are polynomials in Σ_1, Σ_2, Σ_3 and Σ_4):

$$LM = \begin{bmatrix} 23+9\,\Sigma_1 & -2+31\,\Sigma_1/2+9\,\Sigma_2 & \ldots & \ldots & -28\,\Sigma_1'-16 & -28\,\Sigma_2'-8\,\Sigma_1' \\ -25-9\,\Sigma_1/2 & -35\,\Sigma_1/2-2-9\,\Sigma_2/2 & \ldots & \ldots & -8-12\,\Sigma_1' & -4\,\Sigma_1'-12\,\Sigma_2' \\ 16+11\,\Sigma_1/2 & 11\,\Sigma_1+11\,\Sigma_2/2 & \ldots & \ldots & -8-12\,\Sigma_1' & -4\,\Sigma_1'-12\,\Sigma_2' \\ -114-30\,\Sigma_1 & -75\,\Sigma_1-30\,\Sigma_2 & \ldots & \ldots & -8\,\Sigma_1' & -8\,\Sigma_2' \\ 6 & 3\,\Sigma_1 & \ldots & \ldots & 4\,\Sigma_1' & 4\,\Sigma_2' \\ -29-9\,\Sigma_1 & 2-37\,\Sigma_1/2-9\,\Sigma_2 & \ldots & \ldots & -4\,\Sigma_1' & -4\,\Sigma_2' \end{bmatrix}$$

where $\Sigma_1 = \lambda_1+\lambda_2 = -1$, $\Sigma_2 = \lambda_1^2+\lambda_2^2 = -1$, $\Sigma_3 = \lambda_1^3+\lambda_2^3 = 2$, $\Sigma_4 = \lambda_1^4+\lambda_2^4 = -1$, and $\Sigma_1' = \lambda_3+\lambda_4 = -3$, $\Sigma_2' = \lambda_3^2+\lambda_4^2 = 7$. Consequently,

$$LM = \begin{bmatrix} 14 & -\frac{53}{2} & 1/2 & 41 & 68 & -172 \\ -\frac{41}{2} & 20 & \frac{31}{2} & -\frac{95}{2} & 28 & -72 \\ 21/2 & -\frac{33}{2} & -3 & \frac{57}{2} & 28 & -72 \\ -84 & 105 & 33 & -192 & 24 & -56 \\ 6 & -3 & -3 & 6 & -12 & 28 \\ -20 & \frac{59}{2} & 5/2 & -47 & 12 & -28 \end{bmatrix}.$$

We have obtained a rational transformation for $\bar{F}$. Finally, a transformation P for F is constructed from LM by taking as cyclic vector the sum of the first and the fifth column vectors of LM. □

4 $K[x]$-modules for the Smith and the Frobenius forms

This section is intended to point out the correspondence between transformations for the Frobenius normal form of a constant matrix and transformations for the Smith normal form of a polynomial matrix. This will lead, at §5, to a reduction of the latter problem to the former and, at §6, to an algorithm based on this reduction and on theorem 3.2.

Our approach is an extension of the one in [20] for computing normal forms of matrices. In this latter paper, following the classical approach [12], the author computes transformations for the Frobenius normal form of a constant matrix, from transformations for the Smith normal form of an associated polynomial matrix. The algorithm is based on a similar correspondence to the one between the Frobenius form of a matrix B and the Smith form of $xI - B$ (theorem 1.2). This section is intended to recall these basic facts. Then we will show that the converse approach also is valid, even if it is less usual, to compute transformations for the Smith form from transformations for the Frobenius one.

4.1 K^n as $K[x]$-module

The ongoing presentation is derived from standard results in ([12], §3.10), the proofs are omitted. Let B be a matrix in $\mathcal{M}_{n,n}(K)$. We make K^n with basis (u_i), $1 \le i \le n$, a $K[x]$-module by defining the action of x on a vector $v \in K^n$ as:

$$xv = Bv.$$

If we call $N_0 = (xI - B)K[x]^n$ the submodule of $K[x]^n$ generated by the columns of $xI - B$, N_0 is the kernel of p_0:

$$\begin{cases} p_0 \ : K[x]^n \to K^n \\ p_0({}^t[g_1, g_2, \ldots, g_n]) = g_1 u_1 + g_2 u_2 + \ldots + g_n u_n \end{cases}$$

and K^n is isomorphic to $K[x]^n/N_0$. The following diagram commutes:

$$\begin{array}{ccc} K[x]^n/N_0 & \stackrel{p_0}{\to} & K^n \\ \downarrow x & & \downarrow B \\ K[x]^n/N_0 & \stackrel{p_0}{\to} & K^n \end{array}$$

Now we consider a matrix $H(x)$ equivalent to $xI-B$ with $H(x) = U_0(x)(xI-B)V_0(x)$ where $U_0(x)$ and $V_0(x)$ are invertible in $\mathcal{M}_{n,n}(K[x])$. We may assume (corollary 2.1) that $H(x)$ has the quasi-diagonal form:

$$\begin{bmatrix} I_{n-\tau} & \\ & H_\tau(x) \end{bmatrix} \tag{14}$$

where $I_{n-\tau}$ is the identity matrix of dimension $n - \tau$ and $H_\tau(x)$ is under Hermite normal form of rank τ, with diagonal entries of degrees $d_1, \ldots, d_\tau$ at least 1. As previously, if N is the submodule of $K[x]^\tau$ generated by the columns of $H_\tau(x)$, since $H(x)$ and $xI - B$ are equivalent, N is isomorphic to N_0 and we can make K^n isomorphic to $K[x]^\tau/N$. To construct a corresponding commuting diagram we choose as natural K-basis of $K[x]^\tau/N$:

$$\begin{array}{r} \overline{(1, 0, \ldots, 0)}, \overline{(x, 0, \ldots, 0)}, \ldots, \overline{(x^{d_1-1}, 0, \ldots, 0)}, \ldots \\ \ldots, \overline{(0, 0, \ldots, 1)}, \overline{(0, 0, \ldots, x)}, \ldots, \overline{(0, 0, \ldots, x^{d_\tau-1})}, \end{array} \tag{15}$$

and we denote by p the coordinate function with respect to this basis. We have:

$$\begin{array}{ccc} K[x]^\tau/N & \stackrel{p}{\to} & K^n \\ \downarrow x & & \downarrow C \\ K[x]^\tau/N & \stackrel{p}{\to} & K^n \end{array} \tag{16}$$

where, by construction, C is under *polycyclic* or *shift-Hessenberg* form. More precisely, C is quasi upper triangular with τ diagonal blocks. If the entries of $H_\tau(x)$ are denoted by:

$$\begin{array}{rcl} h_{i,j}(x) & = & a_{i,0}^{(j)} + a_{i,1}^{(j)} x + \ldots + a_{i,d_i-1}^{(j)} x^{d_i-1}, \quad 1 \le i < j \le \tau \\ h_{i,i}(x) & = & a_{i,0}^{(i)} + a_{i,1}^{(i)} x + \ldots + a_{i,d_i-1}^{(i)} x^{d_i-1} + x^{d_i}, \quad 1 \le i \le \tau, \end{array}$$

the diagonal blocks $C_{i,i}$ of C are the companion blocks associated to the $h_{i,i}(x)$; the upper diagonal blocks $C_{i,j}$ have non zero entries only in their last column which is equal to $[-a_{i,0}^{(j)}, -a_{i,1}^{(j)}, \ldots, -a_{i,d_i-1}^{(j)}]$.

We can do the same thing for the submatrix $S_\sigma(x)$ of the Smith normal form of $xI - B$ formed by the σ nontrivial invariant factors,

$$S(x) = \begin{bmatrix} I & 0 \\ 0 & S_\sigma(x) \end{bmatrix}$$

the associated polycyclic form is the Frobenius normal form F of B. If we call N' the submodule of $K[x]^\sigma$ generated by the columns of $S_\sigma(x)$, the diagram (16) becomes:

$$\begin{array}{ccc} K[x]^\sigma/N' & \overset{p'}{\rightarrow} & K^n \\ \downarrow x & & \downarrow F \\ K[x]^\sigma/N' & \overset{p'}{\rightarrow} & K^n \end{array} \tag{17}$$

where p' denotes the coordinate function according to the natural K-basis of $K[x]^\sigma/N'$:

$$\begin{array}{c} \overline{(1,0,\ldots,0)}, \overline{(x,0,\ldots,0)}, \ldots, \overline{(x^{\delta_1-1},0,\ldots,0)}, \ldots \\ \ldots, \overline{(0,0,\ldots,1)}, \overline{(0,0,\ldots,x)}, \ldots, \overline{(0,0,\ldots,x^{\delta_\sigma-1})}, \end{array} \tag{18}$$

the δ_i being the degrees of the nontrivial invariant factors $s_i(x)$, $1 \leq i \leq \sigma$. We summarize in the following lemma.

Lemma 4.1 *Let $H_\tau(x)$ be a matrix in $\mathcal{M}_{\tau,\tau}(K[x])$ of full rank which determinant is of degree n. We assume $H_\tau(x)$ to be under Hermite normal form with diagonal entries of degrees at least 1. Let $S(x)$ be its Smith normal form, $S(x) = U(x)H_\tau(x)V(x)$, and $S_\sigma(x)$ be the submatrix of $S(x)$ given by the σ nontrivial invariant factors. The columns of $H_\tau(x)$ and of $S_\sigma(x)$ generates the submodules N and N' of $K[x]^\tau$ and $K[x]^\sigma$. To $H_\tau(x)$ and $S_\sigma(x)$ we respectively associate matrices C and F in $\mathcal{M}_{n,n}(K)$: C has polycyclic form and F has Frobenius form. Then the following $K[x]$-module isomorphism diagram commutes*

$$\begin{array}{ccc} K[x]^\tau/N & \overset{\phi_u}{\rightarrow} & K[x]^\sigma/N' \\ \downarrow p & & \downarrow p' \\ K^n & \overset{P^{-1}}{\rightarrow} & K^n \end{array} \tag{19}$$

where ϕ_u is the restriction of the isomorphism associated to $U(x)$ and P, invertible in $\mathcal{M}_{n,n}(K)$, satisfies $F = P^{-1}CP$. The isomorphisms p and p' are the coordinate functions with respect to the bases

$$\begin{array}{c} \overline{(1,0,\ldots,0)}, \overline{(x,0,\ldots,0)}, \ldots, \overline{(x^{d_1-1},0,\ldots,0)}, \ldots, \overline{(0,0,\ldots,x^{d_\tau-1})}, \\ \overline{(1,0,\ldots,0)}, \overline{(x,0,\ldots,0)}, \ldots, \overline{(x^{\delta_1-1},0,\ldots,0)}, \ldots, \overline{(0,0,\ldots,x^{\delta_\sigma-1})} \end{array}$$

of $K[x]^\tau/N$ and $K[x]^\sigma/N'$.

Proof. Combining diagrams (16), diagram (17) and $S(x) = U(x)H_\tau(x)V(x)$ leads to the diagram (19) with p and p'. Taking $P = p \text{ o } \phi_u^{-1} \text{o } p'^{-1}$, we have:

$$P(xu) = xP(u),\ u \in K^n, \text{ i.e. } PFu = CPu,$$

and,

$$F = P^{-1}CP$$

as announced. □

We notice that only $U(x)$ plays a role in the above constructions, $H_\tau(x)$ and $S(x)$ can be considered up to a unimodular matrix $V(x)$. We also remark that from P only ϕ_u, the restriction of the isomorphism associated to $U(x)$ from $K[x]^\tau/N$ to $K[x]^\sigma/N'$, can be directly computed. We will need a "completion phase" to recover a satisfying matrix $U(x)$ from ϕ_u (see §5). Now we assume the assumptions of lemma 4.1 are satisfied. We first look at the informations given by diagram (19) about the relations between the matrices P^{-1} and $U(x)$ and we compute P^{-1} from $U(x)$.

4.2 From $U(x)$ to a transformation for the Frobenius normal form

To compute the columns of P^{-1} from $U(x)$, let us take K^n with the canonical basis rewritten as

$$(c_1^{(0)},\ldots,c_1^{(d_1-1)},\ldots\ldots,c_\tau^{(0)},\ldots,c_\tau^{(d_\tau-1)})$$

and let

$$(e_1^{(0)},\ldots,e_1^{(d_1-1)},\ldots\ldots,e_\tau^{(0)},\ldots,e_\tau^{(d_\tau-1)})$$

denotes the basis (15). For i, $1\le i\le\tau$, and j, $0\le j\le d_i-1$, we have:

$$\begin{aligned} P^{-1}c_i^{(j)} &= (p' \text{ o } \phi_u \text{ o } p^{-1})c_i^{(j)} = (p' \text{ o } \phi_u)e_i^{(j)} = (p' \text{ o } \phi_u)(x^j e_i^{(0)}) \\ &= x^j(p' \text{ o } \phi_u)e_i^{(0)} = x^j p' U_i = x^j \bar{U}_i = F^j \bar{U}_i \end{aligned}$$

where U_i is the vector of the last σ entries of the i-th column of $U(x)$ taken in N' *i.e.* modulo the columns of $S_\sigma(x)$, and $\bar{U}_i$ is the image of U_i under p'. Since only $\sigma\le\tau$ invariant factors are non constant, only the last σ rows of $U(x)$ are involved to compute P^{-1} (the others are zeroed by the modulo). From the above identities and applying p' column-wise, we express P^{-1} by:

$$P^{-1} = p'\left(\left[U_1,\ldots,x^{d_1-1}U_1,\ldots\ldots,U_\tau,\ldots,x^{d_\tau-1}U_\tau\right] \bmod \begin{bmatrix} s_1(x) \\ \ldots \\ s_\sigma(x)\end{bmatrix}\right). \qquad (20)$$

The modulo is taken component-wise. Equivalently, let the entries of the last σ columns of $U^{-1}(x)$ be denoted by $u_i^{(j)}(x)$, $1\le j\le\sigma$ and $1\le i\le\tau$, then [12] σ cyclic vectors P_j to build P are given by:

$$P_j = u_1^{(j)}(x)c_1^{(0)} + \ldots + u_\tau^{(j)}(x)c_\tau^{(0)}, \quad 1\le j\le\sigma. \qquad (21)$$

These relations can be easily used to compute a transformation P for the Frobenius form, once a transformation $U(x)$ for the Smith form is known. This has been used in [20].

Example 4.1 Let $S_2(x)$ $(\tau=\sigma=2)$ the Smith normal form of $H_2(x)$ be given by:

$$\begin{aligned} S_2(x) &= U(x)H_2(x)V(x) \\ &= \begin{bmatrix} 1 & 0 \\ 2x-x^2 & 1 \end{bmatrix}\begin{bmatrix} x^2-1 & x+1 \\ 0 & x^2-x-2\end{bmatrix}\begin{bmatrix} -1 & -1 \\ x & x-1\end{bmatrix} \\ &= \begin{bmatrix} x+1 & 0 \\ 0 & (x-2)\left(x^2-1\right)\end{bmatrix}. \end{aligned}$$

To $H_2(x)$ we associate the matrix C under polycyclic form:

$$C = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Looking at the degrees of the diagonal entries of $C(x)$, for K-basis of $K[x]^\tau/N$ we take $(1,0)$, $(x,0)$, $(0,1)$, $(0,x)$. A transformation P^{-1} is constructed from the columns U_1, xU_1, U_2 et xU_2 reduced modulo the columns of $S(x)$. Identity (20) is:

$$P^{-1} = p'\left(\left[\begin{array}{cc|cc} 1 & x & 0 & 0 \\ \hline 2x-x^2 & 2x^2-x^3 & 1 & x \end{array}\right] \bmod \begin{bmatrix} x+1 \\ x^3-2x^2-x+2 \end{bmatrix}\right)$$

thus,

$$U(x) = \left[\begin{array}{c|c} 1 & 0 \\ \hline 2x - x^2 & 1 \end{array}\right] \quad \leftrightarrow \quad P^{-1} = \left[\begin{array}{cc|cc} 1 & -1 & 0 & 0 \\ \hline 0 & 2 & 1 & 0 \\ 2 & -1 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{array}\right].$$

The Frobenius form F of C satisfies:

$$F = P^{-1} \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} P = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

The matrix P^{-1} is completely determined by the choice of $U(x)$. □

5 Transformations for the Smith normal form

We still assume we are in the scope of application of lemma 4.1. Now, the problem is to compute a transformation $U(x)$ for the Smith form from a transformation P for the Frobenius form. As underlined previously, only ϕ_u, the restriction of the isomorphism associated to $U(x)$ from $K[x]^\tau/N$ to $K[x]^\sigma/N'$ is known. This will readily gives us matrices $\bar{U}_{\sigma,\tau}(x)$ and $M_{\sigma,\tau}(x)$ in $\mathcal{M}_{\sigma,\tau}(K[x])$ such that

$$\bar{U}_{\sigma,\tau}(x) H_\tau(x) = S_\sigma(x) M_{\sigma,\tau}(x).$$

Then, $\bar{U}_{\sigma,\tau}(x)$ will be completed and modified to give a unimodular matrix $U(x)$.

5.1 Partial construction

We use the natural representation for the elements of $K[x]^\tau/N$ (resp. $K[x]^\sigma/N'$) by elements of $K[x]^\tau$ (resp. $K[x]^\sigma$) taken modulo $H_\tau(x)$ (resp. modulo $S_\sigma(x)$). Let $(e_1, \ldots, e_\tau)$ and $(f_1, \ldots, f_\sigma)$ denote the constant vectors of the bases (15) and (18), they are minimal sets of generators of $K[x]^\tau/N$ and $K[x]^\sigma/N'$. Using diagram (19) and $\phi_u = p'^{-1} \text{ o } P^{-1} \text{ o } p$ we compute the matrix

$$\bar{U}_{\sigma,\tau}(x) = [\phi_u(e_1), \ldots, \phi_u(e_\tau)] \in \mathcal{M}_{\sigma,\tau}(K[x]). \tag{22}$$

With the above representation of the elements of $K[x]^\tau/N$, $\bar{U}_{\sigma,\tau}(x)$ is well defined. Since ϕ_u is an ismomorphism and since the σ vectors $(f_1, \ldots, f_\sigma)$ generates $K[x]^\sigma/N'$, the rank of $\bar{U}_{\sigma,\tau}(x)$ is σ. By construction, any column of $H_\tau(x)$ is sent, under the homomorphism associated to $\bar{U}_{\sigma,\tau}(x)$: $K[x]^\tau \to K[x]^\sigma$, into the module generated by the columns of $S_\sigma(x)$. Thus there exists a matrix $M_{\sigma,\tau}(x)$ in $\mathcal{M}_{\sigma,\tau}(K[x])$ such that

$$\bar{U}_{\sigma,\tau}(x) H_\tau(x) = S_\sigma(x) M_{\sigma,\tau}(x). \tag{23}$$

If $\sigma = \tau$ and $\bar{U}_{\sigma,\tau}(x)$ is unimodular, we have $S(x) = S_\sigma(x)$, obviously we can take $U(x) = \bar{U}_{\sigma,\tau}(x)$. And $V(x) = M_{\tau,\tau}^{-1}(x)$ is a satisfying right multiplier.

Unfortunately, when $\sigma = \tau$, $\bar{U}_{\sigma,\tau}(x)$ may be non unimodular and further, the generic case is $\sigma < \tau$. Anyway, we show below that $\bar{U}_{\sigma,\tau}(x)$ satisfies to conditions allowing to modify and complete it to get a correct transformation matrix. In some sens (the dimensions of the matrices do not necessarily correspond), $\bar{U}_{\sigma,\tau}(x)$ and $S_\sigma(x)$ are left coprime.

5.2 Completion to a unimodular matrix

In the following, $\bar{U}_{\sigma,\tau}(x)$ has been computing by (22). The next proposition will be proven using two additional lemmata.

Proposition 5.1 *Let $\bar{U}_{\sigma,\tau}(x)$ be defined by identity (22), it satisfies (23):*

$$\bar{U}_{\sigma,\tau}(x)H_\tau(x) = S_\sigma(x)M_{\sigma,\tau}(x).$$

Then $\bar{U}_{\sigma,\tau}(x)$ can be completed and modified to give unimodular matrices $U(x)$ and $V(x)$ such that

$$U(x)H_\tau(x)V(x) = S(x)$$

is the Smith normal form of $H(x)$.

The first lemma is a trivial construction.

Lemma 5.1 *Let $\bar{U}_{\sigma,\tau}(x)$ be a matrix of rank σ in $\mathcal{M}_{\sigma,\tau}(K[x])$, with $0 \le \sigma \le \tau$. Let $\left[\bar{H}_{\sigma,\sigma}(x), 0\right]$ be its Hermite normal form. Then there exists a bordering matrix $\bar{B}_{\tau-\sigma,\sigma}(x)$ in $\mathcal{M}_{\tau-\sigma,\tau}(K[x])$ such that the Hermite normal form of*

$$\bar{U}(x) = \begin{bmatrix} \bar{B}_{\tau-\sigma,\sigma}(x) \\ \bar{U}_{\sigma,\tau}(x) \end{bmatrix}$$

is

$$\begin{bmatrix} I_{\tau-\sigma} & 0 \\ 0 & \bar{H}_{\sigma,\sigma}(x) \end{bmatrix}.$$

In particular, if $\bar{H}_{\sigma,\sigma}(x) = I_\sigma$, $\bar{U}_{\sigma,\tau}(x)$ has been completed to a unimodular matrix $U(x) = \bar{U}(x)$.

Proof. Let $R(x)$ be a unimodular transformation in $\mathcal{M}_{\tau,\tau}(K[x])$ such that

$$\bar{U}_{\sigma,\tau}(x) = \left[0, \bar{H}_{\sigma,\sigma}(x)\right] R(x).$$

We can take

$$\begin{bmatrix} \bar{B}_{\tau-\sigma,\sigma}(x) \\ \bar{U}_{\sigma,\tau}(x) \end{bmatrix} = \begin{bmatrix} I_{\tau-\sigma} & 0 \\ 0 & \bar{H}_{\sigma,\sigma}(x) \end{bmatrix} R$$

by construction, $\bar{B}_{\tau-\sigma,\sigma}$ satisfies the lemma. □

The second lemma may be viewed a skew Bezout's identity.

Lemma 5.2 *We consider matrices in $\mathcal{M}_{\tau,\tau}(K[x])$. Let $\bar{U}(x)$ and $S(x)$ be two left coprime non singular matrices, $S(x)$ is assumed to be under Smith normal form, then there exist a diagonal matrix $T(x)$ and a matrix $W(x)$ such that*

$$U(x) = T(x)\bar{U}(x) + S(x)W(x) \tag{24}$$

is unimodular.

Proof. The matrices $\bar{U}(x)$ and $S(x)$ are left coprime, there exist [18] two matrices $X(x)$ and $Y(x)$ such that:

$$\bar{U}(x)X(x) + S(x)Y(x) = I_\tau. \tag{25}$$

We are going to prove the result in two steps. First we show that the above Bezout's identity (25) on matrix polynomials can be translated on well chosen scalar polynomials.

Next, the Bezout's coefficients obtained from these latter identities on polynomials will give target matrices $T(x)$ and $W(x)$. We compute the Hermite normal form $H_u(x)$ of $\bar{U}(x)$:

$$\bar{U}(x) = H_u(x)R(x).$$

Let $h_1(x), \dots, h_\tau(x)$ be the diagonal entries of $H_u(x)$, the $s_{1,1}(x), \dots, s_{\tau,\tau}(x)$ are the invariant factors, (25) may be rewritten as:

$$\begin{bmatrix} h_1 & \times & \times & \times \\ 0 & h_2 & \times & \times \\ \vdots & & \ddots & \times \\ 0 & \dots & 0 & h_\tau \end{bmatrix} RX = \begin{bmatrix} 1 - s_{1,1}y_{1,1} & -s_{1,1}y_{1,2} & \dots & -s_{1,1}y_{1,\tau} \\ -s_{2,2}y_{2,1} & 1 - s_{2,2}y_{2,2} & \dots & -s_{2,2}y_{2,\tau} \\ & & \ddots & \\ -s_{\tau,\tau}y_{\tau,1} & -s_{\tau,\tau}y_{\tau,2} & \dots & 1 - s_{\tau,\tau}y_{\tau,\tau} \end{bmatrix}.$$

For all i, $1 \leq i \leq \tau$, let us consider the minor computed with the last $\tau - i + 1$ rows and columns of the right-hand side matrix:

$$\begin{vmatrix} h_i & \times & \times & \times \\ 0 & h_{i+1} & \times & \times \\ \vdots & & \ddots & \times \\ 0 & \dots & 0 & h_\tau \end{vmatrix} \tilde{r}_i = \begin{vmatrix} 1 - s_{i,i}y_{i,i} & -s_{i,i}y_{i,i+1} & \dots & -s_{i,i}y_{i,\tau} \\ -s_{i+1,i+1}y_{i+1,i} & 1 - s_{i+1,i+1}y_{i+1,i+1} & \dots & -s_{i+1,i+1}y_{i+1,\tau} \\ & & \ddots & \\ -s_{\tau,\tau}y_{\tau,i} & -s_{\tau,\tau}y_{\tau,i+1} & \dots & 1 - s_{\tau,\tau}y_{\tau,\tau} \end{vmatrix}$$

where $\tilde{r}_i$ denotes the minor computed with the last $\tau - i + 1$ rows and columns of RX. Since $s_{i,i}(x)$ is a factor of $s_{j,j}(x)$ for $i \leq j \leq \tau$, there exist polynomials $\tilde{y}_i(x)$ such that:

$$h_i(x)\left(h_{i+1}(x) \dots h_\tau(x)\tilde{r}_i(x)\right) = 1 - s_{i,i}(x)\tilde{y}_i(x), \ 1 \leq i \leq \tau,$$

this is simply a Bezout's identity showing that $h_i(x)$ and $s_{i,i}(x)$ are relatively prime. Taking $t_i(x) = h_{i+1}(x) \dots h_\tau(x)\tilde{r}_i(x)$, as announced we get the traduction of (25) on scalar polynomials:

$$h_i(x)t_i(x) + s_{i,i}(x)\tilde{y}_i(x) = 1, \ 1 \leq i \leq \tau. \tag{26}$$

We now construct the matrices $T(x)$ and $W(x)$, clearly this can be done with any Bezout's coefficients satisfying (26); $T(x)$ is the diagonal matrix defined by:

$$T(x) = \operatorname{diag}\left(t_1(x), \dots, t_\tau(x)\right)$$

and $W(x)$ is:

$$W(x) = \tilde{Y}(x)R(x)$$

where $\tilde{Y}(x) = \operatorname{diag}\left(\tilde{y}_1(x), \dots, \tilde{y}_\tau(x)\right)$. It remains to check that $U(x) = T(x)\bar{U}(x) + S(x)W(x)$, is unimodular. We have:

$$U(x) = \left(T(x)H_u(x) + S(x)\tilde{Y}(x)\right) R(x)$$

the matrix $T(x)H_u(x) + S(x)\tilde{Y}(x)$ is triangular with diagonal identity, since $R(x)$ is unimodular so is $U(x)$. □

We are now ready to prove the proposition.

Proof of proposition 5.1. A matrix $U(x)$ is constructed by applying lemma 5.1 and lemma 5.2 with $\bar{U}_{\sigma,\tau}(x)$. We essentially need to show that it satisfies a left coprimeness criterion with $S(x)$.

As for the construction of $\bar{U}_{\sigma,\tau}(x)$, using diagram (19) and $\phi_u^{-1} = p^{-1} \text{o } P \text{ o } p'$ we compute the matrix

$$\bar{U}_{\tau,\sigma}^{(-1)}(x) = \left[\phi_u^{-1}(f_1), \ldots, \phi_u^{-1}(f_\sigma)\right] \in \mathcal{M}_{\tau,\sigma}(K[x]).$$

Since ϕ_u is an isomorphism, we know that:

$$\pi\left(\bar{U}_{\sigma,\tau}(x)\bar{U}_{\tau,\sigma}^{(-1)}(x)\right) = I_\sigma$$

where $\pi : K[x]^\sigma \to K[x]^\sigma$ is applied colum-wise and associates to a vector the representation of the corresponding class in $K[x]^\sigma/N'$, *i.e.* π reduces modulo $S_\sigma(x)$. Consequently, there exists a matrix $Y_\sigma(x)$ in $\mathcal{M}_{\sigma,\sigma}(K[x])$ such that:

$$\bar{U}_{\sigma,\tau}(x)\bar{U}_{\tau,\sigma}^{(-1)}(x) + S_\sigma(x)Y_\sigma(x) = I_\sigma.$$

Following lemma 5.1 we consider:

$$\left[0, \bar{H}_{\sigma,\sigma}(x)\right] R(x)\bar{U}_{\tau,\sigma}^{(-1)}(x) + S_\sigma(x)Y_\sigma(x) = I_\sigma$$

where $\bar{H}_{\sigma,\sigma}(x)$ and $S_\sigma(x)$ are left coprime. This coprimeness remains true for the matrices $\operatorname{diag}\left(I_{\tau-\sigma}, \bar{H}_{\sigma,\sigma}(x)\right)$ and $S(x) = \operatorname{diag}\left(I_{\tau-\sigma}, S_\sigma(x)\right)$. In the same way, $\bar{U}$ constructed by lemma 5.1 and $S(x)$ are left coprime. To finish we can apply lemma 5.2 to $\bar{U}$, we get

$$U(x) = T(x)\bar{U}(x) + S(x)W(x).$$

It remains to prove that $U(x)$ is a transformation matrix for the Smith form. From

$$\bar{U}_{\sigma,\tau}(x)H_\tau(x) = S_\sigma(x)M_{\sigma,\tau}(x)$$

it can be easily deduced that there exist $M(x)$ such that

$$\bar{U}H_\tau(x) = S(x)M(x).$$

Now for $U(x)$,

$$\begin{aligned} U(x)H_\tau(x) &= T(x)\bar{U}(x)H_\tau(x) + S(x)W(x)H_\tau(x) \\ &= T(x)S(x)M(x) + S(x)W(x)H_\tau(x) \end{aligned}$$

and using that $T(x)$ and $S(x)$ commute since they are diagonal we obtain:

$$\begin{aligned} U(x)H_\tau(x) &= S(x)\left(T(x)M(x) + W(x)H_\tau(x)\right) \\ &= S(x)M'(x). \end{aligned}$$

It is obvious that $M'(x)$ must be unimodular so we can take $V(x) = (M'(x))^{-1}$. $\square$

Example 5.1 Let

$$H_2(x) = \begin{bmatrix} x^2-1 & 1 \\ 0 & x^2-x-2 \end{bmatrix}, \quad S_1(x) = \left[\left(x^2-x-2\right)\left(x^2-1\right)\right].$$

A transformation matrix P for the Frobenius form of the constant matrix C associated to $H_2(x)$ is given by:

$$F = P^{-1}CP = \left[\begin{array}{cc|cc} -6 & 2 & -3 & 0 \\ -1 & -5 & 1 & -3 \\ 4 & -4 & 0 & 1 \\ -1 & 3 & 0 & 0 \end{array}\right] \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{3}{32} & \frac{9}{32} & \frac{27}{32} & \frac{49}{32} \\ \frac{0}{32} & \frac{3}{32} & \frac{9}{32} & \frac{27}{32} \\ \frac{-1}{2} & \frac{-1}{2} & \frac{-3}{2} & \frac{-5}{2} \\ \frac{-1}{4} & \frac{-3}{4} & \frac{-5}{4} & \frac{-11}{4} \end{bmatrix}.$$

Here, $\tau = 2$ and $\sigma = 1$. Following (22) we get:

$$\bar{U}_{1,2}(x) = \left[\begin{array}{cc} -6 - x + 4\,x^2 - x^3 & x - 3 \end{array}\right].$$

Notice that $\bar{U}_{1,2}(x)$ cannot be directly completed to a unimodular matrix since the greatest divisor of its entries is $x - 3 \neq 1$. As explained in lemma 5.1 to complete $\bar{U}_{1,2}(x)$ and in lemma 5.2 to modify it, we compute a unimodular matrix $R(x)$ for the Hermite normal form of $\bar{U}_{1,2}(x)$:

$$\bar{U}_{1,2}(x) = \left[\begin{array}{cc} 0 & \bar{H}_{1,1} \end{array}\right] R(x) = \left[\begin{array}{cc} 0 & x - 3 \end{array}\right] \left[\begin{array}{cc} x - x^2 & 1 \\ x - x^2 + 2 & 1 \end{array}\right]$$

$x - 3$ is relatively prime to $s_{2,2}(x)$, if $t_2(x)$ and $\tilde{y}_2(x)$ are corresponding Bezout's coefficients, this leads to

$$U(x) = (\left[\begin{array}{cc} 1 & 0 \\ 0 & t_2(x) \end{array}\right] \left[\begin{array}{cc} 1 & 0 \\ 0 & x - 3 \end{array}\right] + \left[\begin{array}{cc} 1 & 0 \\ 0 & s_{2,2}(x) \end{array}\right] \left[\begin{array}{cc} 1 & 0 \\ 0 & \tilde{y}_2(x) \end{array}\right])R(x)$$

so in fact we can take $U(x) = R(x)$. The Hermite normal form of $U(x)H_2(x)$ is

$$S(x) = \left[\begin{array}{cc} 1 & 0 \\ 0 & (x^2 - x - 2)\,(x^2 - 1) \end{array}\right]$$

that is the Smith normal form of $H_2(x)$. □

6 Reduction to Smith normal form: the complexity

From §4 we know that a reduction (F, P^{-1}) to the Frobenius form over K is readily computed from a reduction $(S(x), U(x))$ to the Smith form over $K[x]$. Indeed, the transformation P^{-1} is computed by taking the columns of $U(x)$ modulo $S(x)$ (see relation (20). Conversely, from §5, in addition to P^{-1}, we have used Bezout's identities (reductions to Hermite normal form) to obtain relation (24) for the "lifting" of $U(x)$.

However, the two problems have the same parallel time complexity. In particular, next theorem concerning the Smith form of polynomial matrices is equivalent to theorem 3.2 concerning the Frobenius form of scalar matrices by virtue of above remarks.

Theorem 6.1 *Let K be a commutative field. The problem of computing the Smith normal form $S(x)$ of a matrix $A(x)$ of degree d in $\mathcal{M}_{n,m}(K[x])$ and unimodular transformations $U(x)$ in $\mathcal{M}_{n,n}(K[x])$ and $V(x)$ in $\mathcal{M}_{m,m}(K[x])$ such that $S(x) = U(x)A(x)V(x)$ is in $\mathcal{NC}^2_K$. Over the rationals or finite fields GF_q the problem is in $\mathcal{NC}^2$.*

Proof. We simply associate to $A(x)$ a constant matrix B which Frobenius form F gives $S(x)$ and from a transformation P for F (theorem 3.2) we compute a transformation $U(x)$ (proposition 5.1).

We first determine the dimension ν of an appropriate matrix B. Let us remind that if N_A is the module generated by the columns of $A(x)$, this dimension is the dimension of $K[x]^n/N_A$ as K-vector space. Applying corollary 2.1 we bring $A(x)$ into form (1):

$$\left[\begin{array}{ccc} I & 0 & 0 \\ 0 & H_\tau(x) & 0 \\ 0 & 0 & 0 \end{array}\right] = U_1(x)A(x)V_1(x).$$

It is now sufficient to deal with the regular matrix $H_\tau(x)$. If $S_\tau(x)$ is its Smith form with $S_\tau(x) = U_2(x)H_\tau(x)V_2(x)$, we have

$$S(x) = \begin{bmatrix} I & 0 & 0 \\ 0 & S_\tau(x) & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & U_2(x) & 0 \\ 0 & 0 & I \end{bmatrix} U_1(x)A(x)V_1(x) \begin{bmatrix} I & 0 & 0 \\ 0 & V_2(x) & 0 \\ 0 & 0 & I \end{bmatrix}.$$

We are led to form (14), we choose for B the polycyclic form associated to $H_\tau(x)$ by lemma 4.1. In particular, the dimension ν is the degree of the determinant of $H_\tau(x)$ which is also the determinant of $S_\tau(x)$. If $A(x)$ is of rank r, ν is the degree of the greatest common divisor of the $r \times r$ nonzero minors of $A(x)$.

Once B is known, applying theorem 3.2, we compute a transformation P and its inverse. Also F and $S(x)$ are known. Then a matrix $\bar{U}_{\sigma,\tau}$, residue modulo $S(x)$ of $U_2(x)$, is constructed by identity (22) and finally unimodular matrices $U_2(x)$ and $V_2(x)$ – and further, $U(x)$ and $V(x)$ – are given by proposition 5.1.

Computations can clearly be done in $\mathcal{NC}^2_K$. The problem reduces to the Frobenius normal form computation of the matrix which dimension ν is lower than $d\min(n,m)$. The only cost that has not been given yet, is the cost of the lifting of the residue matrix to get $U(x)$ at proposition 5.1. This consists in computing the Hermite normal form $H_u(x)$ of $\bar{U}_{\sigma,\tau}$ (in $\mathcal{NC}^2_K$from [13,27]) and to compute Bezout's indentities between the diagonal entries of $H_u(x)$ and the diagonal entries of $S(x)$ (using the algorithm in [3]). By reduction to a fixed number of problems in $\mathcal{NC}^2$the problem is in $\mathcal{NC}^2$for concrete fields. □

Example 6.1 *We calculate a reduction to Smith form of*

$$A(x) = \begin{bmatrix} 1 & 0 & -2 & x \\ 3 & x^2-1 & -5 & 4x - x^3 \\ 3x^2 - 2x - 6 & 0 & 3x - 5x^2 + 10 & 3x^3 - 2x^2 - 6x \\ x+3 & x^2-1 & -2x-5 & x^2 - x^3 + 4x \end{bmatrix}.$$

By two Hermite normal form reductions (on the left and on the right) we get $H(x) = U_1(x)A(x)V_1(x)$ under form (1):

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & x^2-1 & 1 & 0 \\ 0 & 0 & x^2-x-2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1-x^2 & -x & 0 & x \\ 0 & 1 & 0 & 0 \\ -x & 0 & 1 & 0 \\ -x & -1 & 0 & 1 \end{bmatrix} A(x) \begin{bmatrix} 2x-5 & 0 & 2 & -x \\ -2x & 1 & 0 & x \\ -3 & 0 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix}.$$

Here the K-vector space for B is of dimension $\nu = 4$, the 2×2 polynomial regular submatrix $H_2(x)$ of $H(x)$ is the matrix we have considered at example 5.1.

As trivial corollary we obtain a polynomial-time sequential algorithm. This result improves the previous ones on the subject since no polynomial-time algorithm, reducing to Smith normal form and computing multipliers over the input field K, independently of K, was known even in sequential.

Corollary 6.1 *The problem of computing the Smith normal form of a polynomial matrix $A(x)$ and associated unimodular transformations $U(x)$ and $V(x)$ is in $\mathcal{P}$ over $Q[x]$ or $GF_q[x]$. For any commutative field K, the problem can be solved in a polynomial number of operations in K.*

Conclusion

As pointed out by von zur Gathen [8], *complexity studies aim at understanding how the solutions of problems depends on the ground field; from this point of view the most satisfactory methods do not depend on the field at all.* We has shown that such methods and fast parallel algorithms can be found for reducing matrices to canonical forms. The next step could be to look for efficient solutions *i.e.* using a number of processors close to the sequential time. Even if we have shown the problems to be in $\mathcal{NC}_K$, our solutions for rational forms still need to deal, at least symbolically, with field extensions. To avoid this may be a key point toward efficiency.

References

1. D. Augot and P. Camion. The minimal polynomials, characteristic subspaces, normal bases and the Frobenius form. Technical Report 2006, INRIA France, August 1993.
2. D. Bini and V. Pan. *Polynomial and matrix computations.* Birkhäuser, 1994.
3. A. Borodin, J. von zur Gathen, and J. Hopcroft. Fast parallel matrix and gcd computations. *Information and Control*, 52:241–256, 1982.
4. S.A. Cook. A taxonomy of problems with fast parallel algorithms. *Inf. Control*, 64:2–22, 1985.
5. M.A. Frumkin. Polynomial time algorithms in the theory of linear diophantine equations. In *Fundamentals of Computation Theory*, pages 386–392. LNCS 56, Springer, New-York, 1977.
6. F.R. Gantmacher. *Théorie des matrices.* Dunod, Paris, France, 1966.
7. J. von zur Gathen. Parallel algorithms for algebraic problems. *SIAM J. Comp.*, 13(4):802–824, 1984.
8. J. von zur Gathen. Parallel arithmetic computations: a survey. In *Proc. 12th Int. Symp. Math. Found. Comput. Sci.*, pages 93–112. LNCS 233, Springer Verlag, 1986.
9. T. Gautier. *Calcul formel et parallélisme : conception du système* `Givaro` *et applications au calcul dans les extensions algébriques.* PhD thesis, INPG, Grenoble, France, 1996.
10. M. Giesbrecht. Nearly optimal algorithms for canonical matrix forms. *SIAM Journal on Computing*, 1994. To appear.
11. T. Gómez-Díaz. *Quelques applications de l'évaluation dynamique* . PhD thesis, Université de Limoges, France, 1994.
12. N. Jacobson. *Basic Algebra I.* W.H. Freeman and Company, 1974.
13. E. Kaltofen, M.S. Krishnamoorthy, and B.D. Saunders. Fast parallel computation of Hermite and Smith forms of polynomials matrices. *SIAM J. Alg. Disc. Meth.*, 8 4, pp 683-690, 1987.
14. E. Kaltofen, M.S. Krishnamoorthy, and B.D. Saunders. Parallel algorithms for matrix normal forms. *Linear Algebra and its Applications*, 136:189–208, 1990.
15. R. Kannan. Solving systems of linear equations over polynomials. *Theoretical Computer Science*, 39:69–88, 1985.
16. R. Kannan and A. Bachem. Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix. *SIAM J. Comput.*, 8 4, pp 499-507, 1979.
17. H. Lüneburg. *On rational form of endomorphims : a primer to constructive algebra.* Wissenschaftsverlag, Mannheim, 1987.
18. C.C. MacDuffee. *The theory of matrices.* Chelsea, New-York, 1956.

19. K. Martín and J.M. Olazábal. An algorithm to compute the change basis for the rational form of K-endomorphisms. *Extracta Mathematicae*, 6:89–91, 1992.

20. T. Mulders. Computation of normal forms for matrices. In Algoritmen In De Algebra, A Seminar on Algebraic Algorithms, 1993. University of Nijmegen, Netherlands.

21. P. Ozello. A probalistic algorithm to compute the Frobenius form of a matrix. Technical Report RR 653M, IMAG, Grenoble, France, March 1987.

22. P. Ozello. *Calcul exact des formes de Jordan et de Frobenius d'une matrice*. PhD thesis, Université Scientifique et Médicale de Grenoble, France, 1987.

23. J.L. Roch and G. Villard. Parallel computations with algebraic numbers, a case study: Jordan normal form of matrices. In *Parallel Architectures and Languages Europe 94, Athens Greece*, LNCS 817, July 1994.

24. A. Storjohann. Computation of Hermite and Smith normal forms of matrices. Master's thesis, 1994. University of Waterloo, Canada.

25. G. Villard. Fast parallel computation of the Smith normal form of polynomial matrices. In *International Symposium on Symbolic and Algebraic Computation, Oxford, UK*, pages 312 – 317. ACM Press, July 1994.

26. G. Villard. Generalized subresultants for computing the Smith normal form of polynomial matrices. *Journal of Symbolic Computation*, 20:269–286, 1995.

27. G. Villard. Computing Popov and Hermite forms of polynomial matrices. In *International Symposium on Symbolic and Algebraic Computation, Zurich, Suisse*. ACM Press, July 1996.

28. J.H Wilkinson. *The algebraic eigenvalue problem*. Oxforf University Press, 1965.

APACHE: List of Research Reports

All these reports are available via ftp: `ftp.imag.fr:/pub/APACHE/RAPPORT`

1: Présentation du projet APACHE (en français)
B. Plateau et al. - October 1993 (revised version November 94)

2: Matrix Transpose for Block Allocations on Processor Networks
C. Calvin, and D. Trystram - November 1993

3: Optimal Scheduling of UECT trees on two identical processors
F. Guinand and D. Trystram - November 1993

4: Some Ergodic Results on Stochastic Iterative DED
J-M. Vincent - December 1993

5: Optimal parallel execution of complete binary trees and grids into most popular interconnection networks
E. Bampis, J-C. Konig and D. Trystram - December 1993

6: The numerical solution of stochastic automata networks
W. Stewart, K. Atif and B. Plateau - November 1993

7: Parallel algebraic computations: about the Jordan normal form
J-L. Roch and G. Villard - February 1994

8: Parallel algebraic computations: about the Smith normal form
G. Villard - February 1994

9: Gossiping in Torus with Circuit-Switched Routing
C. Calvin, S. Perennes and D. Trystram - April 1994

10: Parallel Matrix-vector Product on Rings with a minimum of Communications
L. Colombet, P. Michallon and D. Trystram - June 1994

11: ATHAPASCAN-0a sur PVM 3 : définition et mode d'emploi (en français)
M. Christaller - June 1994

12: Introduction au modèle de programmation par processus communicants : deux exemples PVM et MPI (en français)
C. Calvin and L. Colombet - July 1994

13: Analysis of gossiping algorithms in torus with restricted bufferization capabilities
J-C. Konig, P.S. Rao and D. Trystram - November 1994

14: PAC++ v2.0 User and developer manual
T. Gautier, J.L. Roch and G. Villard - November 1994

15: Performance Evaluation of Parallel Systems: the Scope Extensible Interactive Environment
Yves Arrouye, December 1994

16: Fast parallel algorithms for matrix reduction to normal forms
G. Villard - June 1995 (Updated Feb. 96)

17: Worst Case Analysis of Lawler's Algorithm for Scheduling Trees with Communication Delays
F. Guinand, C. Rapine and D. Trystram - July 1995

18: Optimized execution replay mechanism for RPC-based parallel programming models
A. Fagot, J. Chassin de Kergommeaux - July 1995

19: Fast matrix polynomial computations
G. Villard - February 96