Computing the Frobenius Normal Form of a Sparse Matrix

Gilles Villard

CNRS-LMC, BP53 F38041 Grenoble cedex 9, France

Abstract. We probabilistically determine the Frobenius form and thus the characteristic polynomial of a matrix $A \in \mathbb{F}^{n\times n}$ by $O(\mu n \log(n))$ multiplications of A by vectors and $O\left(\mu n^2 \log^2(n) \log\log(n)\right)$ arithmetic operations in the field $\mathbb{F}$. The parameter μ is the number of distinct invariant factors of A, it is less than $3\sqrt{n}/2$ in the worst case. The method requires $O(n)$ storage space in addition to that needed for the matrix A.

1 Introduction

The known complexity estimates of the computation of the characteristic polynomial and *a fortiori*, of the Frobenius normal form of special – sparse or black box – square matrices A over a field $\mathbb{F}$, seem to not be satisfactory. We refer to Kaltofen [8, Open Problem 3] and to Pan et al. [16, 15] for discussions on this subject and survey of current solutions. We denote by $\mathcal{M}(n)$ the number of operations in $\mathbb{F}$ required for $n \times n$ matrix multiplications. The characteristic polynomial of a general matrix A can be computed at cost of $O(n^3)$ or $O(\mathcal{M}(n)\log n)$ operations by the method of Keller-Gehrig [10]. The Frobenius normal form can be computed in $O(n^3)$ as achieved by Storjohann [19, 20] and Storjohann and the author [21] while the randomized Las Vegas algorithms of Giesbrecht [7] and of Eberly [3, §4.3] give the best known asymptotic complexity $O(\mathcal{M}(n)\log n)$. The problem we address is to reduce these estimates when A is sparse or more generally, given by a fast (faster than $O(n^2)$) matrix-vector multiplication.

If we rely on the bound $\mathcal{M}(n) \le 2n^3 - n^2$, the worst case estimate for sparse matrices was $O(n^3)$, not better than in the general case. Our paper decreases the bound to $O(\mu n \log(n))$ products of A by vectors and $O\left(\mu n^2 \log^2(n) \log\log(n)\right)$ arithmetic operations in $\mathbb{F}$ where μ is the number of distinct invariant factors of A. Since μ is less than $3\sqrt{n}/2$, we gain a factor almost $\sqrt{n}$ if multiplying A by a vector costs $n \log^{O(1)}$ operations. The algorithm is Monte Carlo randomized, it succeeds with high probability if the field contains a large enough number of distinct elements compared to the dimension of the matrix (see Theorem 3).

Faster solutions exist for particular classes of sparse matrices. A first particular class is given by matrices A defined with their $s(n)$-separator families, the characteristic polynomial can be computed in $O(n^2) + n\mathcal{M}(s(n))$ operations as shown by Reif [17] or Pan et al. [16, 15]. Up to a logarithmic factor with the standard matrix multiplication, we reach this estimate even when $s(n) = O(\sqrt{n})$.

In this framework of separable matrices it is not known how to compute the Frobenius normal form.

Another particular class of matrices is formed by those having few invariant factors. Let us call ν the number of non-trivial (non equal to one) invariant factors of A. If $\nu = 1$, *i.e.* if the characteristic polynomial is equal to the minimum polynomial, then Wiedemann [22] has shown that the characteristic polynomial is computed at cost of $O(n)$ multiplications of A by vectors and $O(n^2)$ additional operations in $\mathbb{F}$ (which could be generalized to any value of ν using a block method). The Eberly's algorithm [4] does not require the knowledge of ν, has cost sensitive to ν and applies to the computation of the Frobenius form. Eberly obtains the normal form by $O(n)$ multiplications of A and A^T by vectors and $O(\nu n^2)$ operations over $\mathbb{F}$ with an additional requirement of $O(n^2)$ storage space. His method also provides a corresponding transition matrix. Our result gives a better cost for large values of ν and thus in the worst case where $\nu = O(n)$, but we do not provide a transition matrix. By reducing the problem to computing minimum polynomials we require only $O(n)$ storage space for elements of $\mathbb{F}$ in addition to the storage of A.

The paper is organized as follows. The Frobenius form of A is computed as the list of the μ distinct invariant factors together with the multiplicities m_i, $1 \leq i \leq \mu$, at which they appear in the characteristic polynomial. We first study in Section 2 the effect of a perturbation $A + B$, with B of rank k, on the Frobenius form of A. Section 3 extends the result to random Toeplitz perturbations B on which we rely for the final complexity. We then prove in Section 4 that computing any of the invariant factors reduces to computing the minimum polynomials of A and of $A + B$ for B of rank given by the index of the target invariant factor. Finally, the Frobenius form itself is computed at Section 5. It is found by a binary search of the μ invariant factor degree changes using random Toeplitz perturbations. We refer to Gantmacher for the classical definitions of the *invariant factors* of matrices over $\mathbb{F}^{n\times n}$ or $\mathbb{F}[x]^{n\times n}$ in relation with the *Frobenius normal form* [6, Chap. 7 §5] or with the *Smith normal form* [6, Chap. 6 §2].

2 Rank-k Perturbations

For a given matrix $A \in \mathbb{F}^{n\times n}$, many relationships are known between the invariant factors of A and those of $A + B$, in particular we have:

Lemma 1. *If $s_1, \ldots, s_n$ are the invariant factors of A and if B has rank k then the invariant factors $\sigma_1, \ldots, \sigma_n$ of $A + B$ satisfies:*

$$s_i | \sigma_{i+k}, \; i = 1, \ldots, n - k. \tag{1}$$

Proof. Let U and V be unimodular matrices in $\mathbb{F}[x]^{n\times n}$ such that $U(x - A)V$ is in Smith normal form $S = \mathrm{diag}(s_1, s_2, \ldots, s_n)$ and let R be unimodular in $\mathbb{F}[x]^{n\times n}$ such that the last $n - k$ rows of $R(UBV)$ are zero. We define the matrix

T, equivalent to $x - A - B$, by:

$$T = RS - R(UBV) = \begin{bmatrix} \times & \ldots & \ldots & \times \\ \vdots & & & \vdots \\ \times & \ldots & \ldots & \times \\ \alpha_{1,1}s_1 & \alpha_{1,2}s_2 & \ldots & \alpha_{1,n}s_n \\ \vdots & & & \vdots \\ \alpha_{n-k,1}s_1 & \ldots & \ldots & \alpha_{n-k}s_n \end{bmatrix} \in \mathbb{F}[x]^{n\times n}$$

where the "$\times$" stand for any elements in $\mathbb{F}[x]$ and the $\alpha_{i,j}$'s are in $\mathbb{F}[x]$. Let $\overline{T}$ be the matrix formed by the last $n-k$ rows of T and let its first $n-k$ linearly independent columns be those indexed by $j_1, \ldots, j_{n-k}$. Considering the left Hermite normal form of $\overline{T}$, we know that for a matrix $\overline{W} \in \mathbb{F}[x]^{(n-k)\times n}$, $\overline{T}$ is left equivalent to $\operatorname{diag}(s_{j_1}, \ldots, s_{j_{n-k}})\overline{W}$. Thus for $W \in \mathbb{F}[x]^{n\times n}$, T is equivalent to $\operatorname{diag}(1, \ldots, 1, s_{j_1}, \ldots, s_{j_{n-k}})W$. Since the invariant factors of two non-singular matrices divide the invariants factors of the product – see Newman [14, Theorem II.14], the invariants factors $\sigma_1, \ldots, \sigma_n$ of T and thus of $x - A - B$ are respectively divisible by $1, \ldots, 1, s_{j_1}, \ldots, s_{j_{n-k}}$. Thus s_{j_i} divides σ_{i+k} and since $j_i \geq i$ implies that s_i divides s_{j_i}, s_i also divides σ_{i+k} for $1 \leq i \leq n-k$. $\square$

In addition to this knowledge on the invariant factors, informations on the characteristic polynomial may be obtained following Lidskii's theory [12] and the construction of Moro et al.[13, §2]. Let $\tilde{A} \in \mathbb{F}^{n\times n}$ be in Jordan normal form $\operatorname{diag}(J, J')$ where

$$J = \operatorname{diag}(J_1^{r_1},\ J_2^{r_1},\ \ldots, J_{q_1}^{r_1},\ \ldots\ldots, J_1^{r_l},\ J_2^{r_l},\ \ldots, J_{q_l}^{r_l}) \in \mathbb{F}^{r\times r}$$

gives all the Jordan blocks

$$J_i^{r_j} = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{F}^{r_j\times r_j}$$

arranged in decreasing dimensions $r_1 > r_2 > \ldots > r_l$, associated to a given eigenvalue λ. We denote by u_i^j and v_i^j, $1 \leq j \leq l$ and $1 \leq i \leq q_j$, the canonical vectors which are right (column) and left (row) eigenvectors in $\mathbb{F}^n$ associated to the block $J_i^{r_j}$. For an integer $1 \leq s \leq l$ and $\tilde{B} \in \mathbb{F}^{n\times n}$, we define Φ_s by:

$$\Phi_s = \begin{bmatrix} v_1^1 \\ \hline \vdots \\ \hline v_{q_1}^1 \\ \hline \vdots \\ \hline v_1^s \\ \hline \vdots \\ \hline v_{q_s}^s \end{bmatrix} \cdot \tilde{B} \cdot \left[\, u_1^1 \middle| \ldots \middle| u_{q_1}^1 \middle| \ldots \middle| u_1^s \middle| \ldots \middle| u_{q_s}^s \,\right] \in \mathbb{F}^{d_s\times d_s},\ d_s = q_1 + \ldots + q_s.$$

Let us note that Φ_s is a sub-matrix of Φ_{s+1} and of $\tilde{B}$. Now, for an integer k, $1 \leq k \leq n$, we define Ψ_k^* and Ψ_k by:

$$\begin{cases} \Psi_k^* = [\ ],\ \Psi_k = \Phi_1 \quad \text{if } 1 \leq k \leq q_1, \\ \Psi_k^* = \Phi_{s-1} \text{ and } \Psi_k = \Phi_s \quad \text{if } q_1 + \ldots + q_{s-1} < k \leq q_1 + \ldots + q_s, \\ \Psi_k^* = \Psi_k = \Phi_l \quad \text{if } q_1 + \ldots + q_l < k \leq n. \end{cases}$$

As previously, Ψ_k^* is a sub-matrix of Ψ_k, they are both sub-matrices of $\tilde{B}$. Lidskii's conditions on particular minors of Ψ_k and thus of $\tilde{B}$ give relations between the characteristic polynomial of $\tilde{A}$ and the one of $\tilde{A} + \epsilon\tilde{B}$:

Lemma 2. *Let $\tilde{B}$ be of rank k. Let Σ_k be the sum of all determinants of $k \times k$ principal sub-matrices of Ψ_k that contain Ψ_k^*, taking the sum of the $k \times k$ principal minors of Ψ_k if $1 \leq k \leq q_1$ and $\Sigma_k = \Sigma_{q_1+\ldots+q_l}$ for $k > q_1 + \ldots + q_l$. Then for $k \leq q_1 + \ldots + q_l$ we have*

$$\chi_\lambda(x, \epsilon) = \det\Big((x + \lambda) - (\tilde{A} + \epsilon\tilde{B})\Big) = (-1)^k c' \Sigma_k \epsilon^k x^{\rho(k)} + \beta(x, \epsilon) x^{\rho(k)+1} \quad (2)$$

where c' is nonzero in the algebraic closure $\overline{\mathbb{F}}$ of $\mathbb{F}$, $\rho(k)$ is the sum of the dimensions of the $q_1 + \ldots + q_l - k$ smallest (last) Jordan blocks of J and β is in $\overline{\mathbb{F}}[x, \epsilon]$. For $k > q_1 + \ldots + q_l$, taking $\rho(k) = 0$, we have

$$\chi_\lambda(x, \epsilon) = \det\Big((x + \lambda) - (\tilde{A} + \epsilon\tilde{B})\Big) = \Big((-1)^{\tilde{k}} c' \Sigma_{\tilde{k}} + \alpha(\epsilon)\epsilon\Big) \epsilon^{\tilde{k}} + x\beta(x, \epsilon) \quad (3)$$

where $\tilde{k} = q_1 + \ldots + q_l$, c' and β are as above and $\alpha \in \overline{\mathbb{F}}[\epsilon]$ has degree $k - \tilde{k} - 1$.

Proof. Since B has rank k there is no power of ϵ greater than k in the determinant, this gives the bound on the degree of α. We begin with the case $k \leq q_1 + \ldots + q_l$. By Lemma 1, with $\tilde{B}$ of rank k, at least $q_1 + \ldots + q_l - k$ invariant factors of $\tilde{A}$ are factors of the characteristic polynomial of $\tilde{A} + \epsilon\tilde{B}$ thus $x^{\rho(k)}$ divides $\chi_\lambda(x, \epsilon)$. We adapt the arguments of Moro et al. [13, Theorem 3.1] to show that the terms in $x^{\rho(k)}\epsilon^i$, $i < k$, are zero and to compute the coefficient of ϵ^k.

For a given i we begin to compute the lowest possible power $x^{\rho(i)}$ involved with ϵ^i. The terms in $\chi_\lambda(x, \epsilon)$ are products of n factors that are elements of $(x + \lambda) - (\tilde{A} + \epsilon\tilde{B})$ in different rows and columns. To produce ϵ^i, $n - i$ factors free of ϵ must be taken in $x + \lambda - J$ or $x + \lambda - J'$. We look at the possible contributions for the term in $x^{\rho(i)}\epsilon^i$. The factors may be "x" or "-1" from the part corresponding to $-J$, or "x", "$\lambda - \lambda'$" or "-1" from the part corresponding to $-J'$ where λ' stands for any eigenvalue of J'. Define $c_{\lambda'} = \lambda - \lambda' \neq 0$. The "$x$" contributing to $x^{\rho(i)}$ must come from the $-J$ part because otherwise the corresponding terms with "$c_{\lambda'}$" would contradict the fact that $\rho(i)$ is minimum. Let j be the number of "ϵ" contributing to ϵ^i and taken in the $-J$ part. If β factors "-1" are taken in the $-J$ part from γ different Jordan blocks then also $\beta + \gamma$ factors "x" are excluded (from the same rows and columns) to choose the remaining factors in the $-J$ part. Since $r - j - \beta$ factors "x" from the $-J$ part will be in the term we have $(r-j-\beta)+(\beta+\gamma)=r-j+\gamma \leq r$ and $\gamma \leq j$. To produce the lowest possible power $\rho(i)$ of x, the term must be formed by taking all the "-1" from j blocks among the largest Jordan blocks of J and j must be as large as possible thus equal to i. The term must be completed by $r-i-\beta$ factors "x" from the $-J$ part and by all the factors "$c_{\lambda'}$" from the $-J'$ part to not introduce additional "ϵ".

We may conclude as done by Moro et al. [13]. The lowest power $\rho(i)$ is the sum of the dimensions of the $q_1+\ldots+q_l-i$ smallest blocks of J. Thus for $i<k$ there is no term in $x^{\rho(k)}\epsilon^i$. If we delete the rows and the columns corresponding to the factor "-1" and "x" from the $-J$ part for the term in $x^{\rho(k)}\epsilon^k$, the remaining elements are in $-\epsilon\Psi_k$. The different possible ways to choose i largest Jordan blocks in J give the different $k\times k$ principal minors to sum and give $(-1)^k\Sigma_k$. The constant c' is the product $\prod_{\lambda'}(\lambda-\lambda')$ taken over all the eigenvalues (with their multiplicities) of J'.

For $k>q_1+\ldots+q_l$, similar arguments show that the lowest power of ϵ in the constant term of χ_λ is $\tilde{k}=q_1+\ldots+q_l$ and give the corresponding coefficient. $\square$

We now apply both Lemma 1 and Lemma 2 for a general $A+\epsilon_0 B$:

Theorem 1. *Let P be such that $P^{-1}AP$ is in Jordan normal form. Let $\lambda_1,\ldots,\lambda_\delta$ be the distinct eigenvalues of A. Let B be of rank k. We denote by $\Gamma^B_{k,\lambda_j}(\epsilon)$ the factor of $\epsilon^k x^{\rho(k)}$ in (2) or of $\epsilon^{\tilde{k}}x^0$ in (3) for $\lambda=\lambda_j$, $1\leq j\leq\delta$, and for $\tilde{B}=P^{-1}BP$. If the invariant factors of A are $s_1,\ldots,s_n$ and if*

$$\Delta_{A,B}(\epsilon_0)=\epsilon_0\prod_{j=1}^{\delta}\Gamma^B_{k,\lambda_j}(\epsilon_0)\neq 0 \qquad (4)$$

for ϵ_0 in $\mathbb{F}^*$, then the invariant factors of $A+\epsilon_0B$ are $t_1,\ldots,t_2,s_1t_{k+1},\ldots,s_{n-k}t_n$ where the t_i's are polynomials in $\mathbb{F}[x]$ relatively prime to the characteristic polynomial of A thus also to $s_1,\ldots,s_{n-k}$.

Proof. The existence of the t_i's is given by Lemma 1. We have to prove their relative primeness to the characteristic polynomial of A. This may be deduced locally at each eigenvalue λ_j, $1\leq j\leq\delta$. Indeed, if (4) holds then by Lemma 2, the valuation of $\chi_{\lambda_j}(x,\epsilon)$ in x is exactly $\rho(k)$. Since $\rho(k)$ is the valuation of $\prod_{i=1}^{n-k}s_i(x+\lambda_j)$, λ_j cannot be a zero of t_i, $1\leq i\leq n$, otherwise the valuation would be strictly greater. $\square$

3 Rank-k Toeplitz Perturbations

For any given matrix A, we first prove that condition (4) is generically satisfied when B is a product of two Toeplitz matrices:

Lemma 3. *Let $\zeta_1,\ldots,\zeta_{n+k-1}$ and $\xi_1,\ldots,\xi_{n+k-1}$ be $2(n+k-1)$ distinct indeterminates over $\mathbb{F}$. Let $\mathcal{B}=\mathcal{U}\mathcal{V}$ be the product of the two Toeplitz matrices*

$$\mathcal{U}=\begin{bmatrix} \zeta_n & \zeta_{n+1} & \cdots & \zeta_{n+k-1} \\ \zeta_{n-1} & \zeta_n & & \zeta_{n+k-2} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \zeta_2 & \ddots & & \zeta_{k+1} \\ \zeta_1 & \zeta_2 & \cdots & \zeta_k \end{bmatrix}, \quad \mathcal{V}=\begin{bmatrix} \xi_k & \xi_{k+1} & \cdots\cdots\cdots & \xi_{n+k-1} \\ \xi_{k-1} & \xi_k & \ddots\ \ddots\ \ddots & \xi_{n+k-2} \\ \vdots & & \ddots\ \ddots & \vdots \\ \xi_1 & \xi_2 & \cdots\cdots\cdots & \xi_n \end{bmatrix} \tag{5}$$

of $\mathbb{F}[\zeta_1,\ldots,\zeta_{n+k}]^{n\times k}$ and of $\mathbb{F}[\xi_1,\ldots,\xi_{n+k}]^{k\times n}$. Then for any given $A\in\mathbb{F}^{n\times n}$, $\Delta_{A,\mathcal{B}}$ is a nonzero polynomial in $\overline{\mathbb{F}}[\zeta_1,\ldots,\xi_{n+k-1},\epsilon]$. Its degree in ϵ is at most $n(k-1)+1$ and its coefficient $\Delta^{(1)}_{A,\mathcal{B}}$ of degree 1 in ϵ is nonzero in $\overline{\mathbb{F}}[\zeta_1,\ldots,\xi_{n+k-1}]$ of total degree at most $2n$.

Proof. We denote by Σ^{B}_{k,λ_j} the quantity Σ_k of Lemma 2 for $\lambda=\lambda_j$, $1\le j\le\delta$, and $\tilde{B}=P^{-1}BP$. We prove that $\Delta^{(1)}_{A,\mathcal{B}}$ and thus $\Delta_{A,\mathcal{B}}$ is nonzero. It is equivalent to show that any Σ^{B}_{k,λ_j} is nonzero since – up to a nonzero constant – their product gives $\Delta^{(1)}_{A,\mathcal{B}}$. By definition, $\Sigma^{\mathcal{B}}_{k,\lambda_j}$ is a sum of $k\times k$ (or $\tilde{k}\times\tilde{k}$) minors built on different rows and columns of $P^{-1}\mathcal{B}P$. For a matrix A and sets I,J of indexes, we denote by $A_{I,J}$ the determinant of the sub-matrix of A built on the rows whose indexes are in I and columns whose of indexes in J. By the Binet-Cauchy formula – see Gantmacher [6, p9], the $k\times k$ minor of $P^{-1}\mathcal{B}P$ built on sets I and J is:

$$\mathcal{D}=\sum_{\substack{L=\{l_1,\ldots,l_k\}\\ 1\le l_1<\ldots<l_k\le n}}\ \sum_{\substack{M=\{m_1,\ldots,m_k\}\\ 1\le m_1<\ldots<m_k\le n}} P^{-1}_{I,L}\mathcal{B}_{L,M}P_{M,J}$$

Now, if the terms of the sum $\Sigma^{\mathcal{B}}_{k,\lambda_j}$ are built on $(I_1,J_1),\ldots(I_\iota,J_\iota)$ then

$$\Sigma^{\mathcal{B}}_{k,\lambda_j}=\sum_{i=1}^{\iota}\mathcal{D}_i=\sum_{i=1}^{\iota}\sum_{L}\sum_{M}P^{-1}_{I_i,L}\mathcal{B}_{L,M}P_{M,J_i}=\sum_{L}\sum_{M}\left(\sum_{i=1}^{\iota}P^{-1}_{I_i,L}P_{M,J_i}\right)\mathcal{B}_{L,M}.$$

Using ideas similar to those Kaltofen and Pan [9, Theorem 2], to prove that the latter sum is non zero we observe that there exist sets L_0 and M_0 with $\sum_i P^{-1}_{I_i,L_0}P_{M_0,J_i}\ne 0$ and that the $\mathcal{B}_{L,M}$'s are linearly independent over $\mathbb{F}$. Indeed, if $\sum_i P^{-1}_{I_i,L}P_{M,J_i}=0$ for all L and M then $\sum_i(P^{(-1)}TP)_{I_i,J_i}=0$ for any matrix T. But taking $T=PEP^{-1}$ with E a matrix whose only $k\times k$ nonzero minor is $E_{I_{i_0},J_{i_0}}$ for some i_0, $1\le i_0\le\iota$, $\sum_i(P^{(-1)}TP)_{I_i,J_i}=E_{I_{i_0},J_{i_0}}\ne 0$. Therefore, L_0 and M_0 exist as announced. For the independence of the $\mathcal{B}_{L,M}$'s, let us first notice that:

$$\mathcal{B}_{L,M}=(\mathcal{U}\mathcal{V})_{L,M}=\mathcal{U}_{L,\{1,\ldots,k\}}\mathcal{V}_{\{1,\ldots,k\},M}.$$

We may view these minors of $\mathcal{U}$ and $\mathcal{V}$ as polynomials in $\mathbb{F}[\zeta_1, \dots, \zeta_{n+k-1}]$ and $\mathbb{F}[\xi_1, \dots, \xi_{n+k-1}]$ with lexicographically ordered terms using the variable orders

$$\zeta_{n-k+1} > \dots > \zeta_2 > \zeta_1, \quad \xi_{n-k+1} > \dots > \xi_2 > \xi_1.$$

For given $L = \{l_1, \dots, l_k\}$ and $M = \{m_1, \dots, m_k\}$, the diagonal terms

$$\zeta_{n-l_1+1} \cdot \zeta_{n-l_2+2} \cdot \dots \cdot \zeta_{n-l_k+k} \quad \text{and} \quad \xi_{k+m_1-1} \cdot \xi_{k+m_2-2} \cdot \dots \cdot \xi_{m_k}$$

are the lexicographically smallest terms in the minor expansions of $\mathcal{U}_{L,\{1,\dots,k\}}$ and of $\mathcal{V}_{\{1,\dots,k\},M}$ and uniquely correspond to L and M thus the polynomials $\mathcal{B}_{L,M}$ are linearly independent over $\mathbb{F}$. This establishes that $\Delta^{(1)}_{A,\mathcal{B}}$ and $\Delta_{A,\mathcal{B}}$ are nonzero polynomials. Moreover, each $\Sigma^{\mathcal{B}}_{k,\lambda_j}$ has degree at most two times the number of Jordan blocks with eigenvalue λ_j thus their product has total degree at most $2n$. By the definition of the $\Gamma^{\mathcal{B}}_{k,\lambda_j}$ in Theorem 1 and by the degree bound for α in (3), $\Gamma^{\mathcal{B}}_{k,\lambda_j}$ has degree in ϵ at most k minus the number d_{λ_j} of Jordan blocks with eigenvalue λ_j thus the product $\Delta_{A,\mathcal{B}}$ has degree in ϵ at most $1 + \sum_j (k - d_{\lambda_j}) \leq nk - n + 1$. □

Lemma 3 indicates that the condition of application of Theorem 1 is generically satisfied for a special class of matrix B, for random such perturbations it follows that:

Theorem 2. *Let A be a matrix in $\mathbb{F}^{n\times n}$ with invariant factors $s_1, \dots, s_n \in \mathbb{F}[x]$ and let S be a finite subset of $\mathbb{F}$. Let $U \in \mathbb{F}^{n\times k}$ and $V \in \mathbb{F}^{k\times n}$ be Toeplitz matrices – as in (5) – with entries chosen uniformly and independently from S. The minimum polynomial of $A+B = A+UV$ is $s_{n-k}t$, where $t \in \mathbb{F}[x]$ has degree less than $\sum_{n-k+1}^{n} \deg s_i$ and is relatively prime to the minimum polynomial of A, with probability at least $1 - (nk+n+1)/|\mathsf{S}|$.*

Proof. By Lemma 3, $\Delta^{(1)}_{A,\mathcal{B}}$ has degree bounded by $2n$. Therefore, using the Schwartz-Zippel Lemma [18, Lemma 1], if the indeterminates $\zeta_1, \dots, \zeta_{n-k+1}$, $\xi_1, \dots, \xi_{n-k+1}$ are replaced by uniformly and independently chosen elements in S to give $B = UV \in \mathbb{F}^{n\times n}$ then $\Delta^{(1)}_{A,B}$ is nonzero with probability at least $1 - 2n/|\mathsf{S}|$. The same argument then gives that for a uniformly and independently chosen ϵ_0 in S, $\Delta_{A,B}(\epsilon_0)$ is nonzero with probability at least $1 - (2n + nk - n + 1)/|\mathsf{S}| \geq 1 - (nk+n+1)/|\mathsf{S}|$. Theorem 1 implies that the minimum polynomial of $A + \epsilon_0 B$ is $s_{n-k}t$, where t is relatively prime to the characteristic polynomial of A. In case of successful random choices, ϵ_0 must be nonzero thus the probability bound is valid for $A + B$. Theorem 1 also implies that the remaining invariant factors are $t_1, t_2, \dots, s_1 t_{k+1}, \dots, s_{n-k-1} t_{n-1}$, their product has degree at least $\sum_1^{n-k-1} \deg s_i$ which, taking the degree of s_{n-k} into account, proves the degree bound for t. □

4 Computing One of the Invariant Factors

For convenience we denote by f_i the $(n-i+1)$-th invariant factor of A. If ν invariant factors are non-trivial (not equal to one) then the polynomials f_i are the characteristic polynomials of the ν blocks of the Frobenius form of A, $f_{i+1}|f_i$ and if $\deg f_i = d_i$ then $\sum_{i=1}^{\nu} d_i = n$. Theorem 2 reduces the computation of f_i to the computation of the minimum polynomial of A and of $A+B$ for B of rank $i-1$:

Function InvFact

> *Input*: an index $1 \leq i \leq n$.
> *Output*: the invariant factor f_i.
>
> Compute the minimum polynomial f_1 of A and if $i = 1$ return f_1
> Choose random Toeplitz matrices U and V of rank $k = i - 1$ as in (5)
> Compute g the minimum polynomial of $A + UV$
> Return $\gcd(f_1, g)$. □

Since the function InvFact consists in computing two minimum polynomials, many results are available for its cost analysis, especially:

Lemma 4. *Let* S *be a subset of* $\mathbb{F}$. *The degree* d *and the coefficients of the minimum polynomial of* $A \in \mathbb{F}^{n\times n}$ *may be probabilistically computed by* $O(d)$ *multiplications of* A *by vectors and* $O(dn)$ *arithmetic operations in* $\mathbb{F}$. *The algorithm returns correct answers with probability at least* $1 - 2d/|\mathsf{S}|$ *and requires* $O(n)$ *space in addition to that required for the matrix coefficient.*

Proof. The version of Lanczos's method given by Lambert [11, Algorithm 3.5.1], for two input vectors u_{curr} and v_{curr}, provides a factor of the minimum polynomial within the announced cost and space bounds. As proved by Kaltofen and Pan [9, Lemma 2], if u_{curr} and v_{curr} are randomly chosen with entries in S then the computed factor coincides with the minimum polynomial with probability more than $1 - 2d/|\mathsf{S}|$. □

This lemma together with the function InvFact gives the following.

Lemma 5. *Let* $A \in \mathbb{F}^{n\times n}$ *and* $\mathsf{S} \subset \mathbb{F}$. *We may probabilistically compute the* $(n-i+1)$*-th invariant factor* f_i *of* A *by computing* $O(d_1+d_2+\ldots+d_i)$ *multiplications of* A *by vectors and* $O\left((d_1 + d_2 + \ldots + d_i)n \log(n) \log\log(n)\right)$ *arithmetic operations in* $\mathbb{F}$. *The algorithm returns* f_i *with probability at least* $1-(n^2+5n+1)/|\mathsf{S}|$.

Proof. By Theorem 2, with probability at least $1-(n^2+n+1)/|\mathsf{S}|$, the minimum polynomial g of $A + UV$ is $f_i t$ with $d = \deg g \leq d_1 + \ldots + d_i$. In addition, t is relatively prime to f_1 and $f_i|f_1$ thus $\gcd(f_1, g) = \gcd(f_1, f_i t) = f_i$. The cost of the matrix-vector multiplications in the computation of g dominates the overall cost. By Lemma 4, $O(d)$ multiplications of $A+UV$ by vectors are needed. For a vector u, $(A+UV)u = Au + (U(Vu))$, the cost of one multiplication is thus the cost of one multiplication by A plus two times the cost of multiplying a Toeplitz matrix by a vector. This latter operation reduces to polynomial multiplication – see Bini and Pan [1, p133] – and costs $O(n \log(n) \log\log(n))$ using the algorithm of Cantor and Kaltofen [2]. This proves the first assertion. By Theorem 2 and Lemma 4, the probability that the algorithm fails can be bounded by $((n^2+n+1) + 2 \times 2d)/|\mathsf{S}| \leq (n^2 + 5n + 1)/|\mathsf{S}|$. □

5 Computing the Frobenius Normal Form

The Frobenius normal form of A is computed as the set of the invariant factors of A. Each invariant factor may be computed by the function $\mathsf{InvFact}$. To reduce the overall cost of the computation we remark that in case of repeated invariant factors it is sufficient to explicitly compute only one of them and their number. For A having exactly μ distinct invariant factors ϕ_i, our algorithm $\mathsf{Frobenius}$ will precisely compute few copies of each ϕ_i (a logarithmic number) together with the powers m_i at which they appear in the characteristic polynomial, so that:

$$\chi_A = \phi_1^{m_1}\phi_2^{m_2}\ldots\phi_\mu^{m_\mu}, \quad \phi_{i+1}|\phi_i,\ 1 \le i < \mu. \tag{6}$$

We denote by $\phi_i^{(1)}, \ldots, \phi_i^{(m_i)}$ the m_i copies of each ϕ_i, $1 \le i \le \mu$. We develop a binary search using rank-k perturbations to detect that – for every ϕ_i – a rank-$k_i^{(1)}$ perturbation provides $\phi_i^{(1)}$ and that a rank-$k_i^{(m_i)}$ provides $\phi_i^{(m_i)} = \phi_i^{(1)}$. From where everything will be known since then the multiplicities m_i must be $k_i^{(m_i)} - k_i^{(1)} + 1$. To know that f_j is $\phi_i^{(1)}$ or $\phi_i^{(m_i)}$ for some i is equivalent to know that $f_{j-1} \ne f_j$ or $f_j \ne f_{j+1}$ ($f_0 = f_{n+1} = 0$). This leads to the function $\mathsf{SearchThresholds}$ which recursively computes, for two different invariant factors f_l and f_m, $l < m$, a decomposition of $f_l f_{l+1} \ldots f_m$

$$\begin{cases} f_l f_{l+1} \ldots f_m = \psi_1^{(1)} \ldots \psi_1^{(l_1)} \psi_2^{(1)} \ldots \psi_2^{(l_2)} \ldots \psi_\kappa^{(1)} \ldots \psi_\kappa^{(l_\kappa)}, \\ \psi_1^{(1)} = f_l, \quad \psi_\kappa^{(l_\kappa)} = f_m, \\ \psi_i^{(1)} = \psi_i^{(2)} = \ldots = \psi_i^{(l_i)} = \psi_i, \quad 1 \le i \le \kappa, \\ \psi_{i+1}|\psi_i, \quad 1 \le i < \kappa, \end{cases} \tag{7}$$

as a partial decomposition (6).

Function $\mathsf{SearchThresholds}$

Input: $[(l, f_l), (m, f_m)]$, $l < m$.
Output: $[l_1, \psi_1, \ldots, l_\kappa, \psi_\kappa]$ ** *The decomposition (7)* **

If $l = m - 1$
  if $f_l = f_m$ then Return $[2, f_l]$
(a) if $f_l \ne f_m$ then Return $[1, f_l, 1, f_m]$
$k := \lceil (l+m)/2 \rceil$
(b) $f_k := \mathsf{InvFact}(k)$
(c) If $f_l \ne f_k$ then $[k_1, \psi_1, \ldots, k_r, \psi_r] := \mathsf{SearchThresholds}((l, f_l), (k, f_k))$
(d) else $r := 1$, $[k_1, \psi_1] := [k - l + 1, f_l]$
(c') If $f_k \ne f_m$ then $[m_1, \tilde{\psi}_1, \ldots, m_s, \tilde{\psi}_s] := \mathsf{SearchThresholds}((k, f_k), (m, f_m))$
(d') else $s := 1$, $[m_1, \tilde{\psi}_1] := [m - k + 1, f_m]$
** *Here*, $\psi_r = \tilde{\psi}_1$ **
Return $[k_1, \psi_1, \ldots, k_r + m_1 - 1, \psi_r, m_2, \tilde{\psi}_2, \ldots, m_s, \tilde{\psi}_s]$. □

The function basically finds all the threshold indexes j such that $f_j \ne f_{j+1}$, $l \le j < m$. They are found at step (a) – with $j = l$, $j + 1 = m$ – which is the deepest level of the recursion. The multiplicities l_i of the invariants factors at accumulated at each level when it is discovered – at steps (d) and (d') – that for two indexes – (l, k) or (k, m) – the corresponding invariant factors are identical. Let us bound the number of recursive calls to the function and thus the number of invariant factor computations by $\mathsf{InvFact}$ at step (b). Since k is an index bound of the created search intervals, the function may go through (b) with $f_l \neq f_k$ and $f_k \neq f_m$ at most $\kappa - 2$ times. In this latter situation, a new invariant factor f_k is found. Then, at its deeper levels, the binary search will refine at most two intervals with index bounds corresponding to invariant factors identical to f_k. The function will thus go $O(\sum_{i=1}^{\kappa} \log l_i)$ times through (b) and either through (c) or (c'), this may be bounded by $O(\kappa \log(m - l))$.

Since decompositions (6) and (7) coincide for $l = 1$ and $m = n$, the Frobenius normal form can be computed by:

Algorithm Frobenius

Input: $A \in \mathbb{F}^{n \times n}$.
Output: the Frobenius normal form of A (decomposition (6)).

$f_1 := \mathsf{InvFact}(1)$
$f_n := \mathsf{InvFact}(n)$
$\mathsf{SearchThresholds}((1, f_1), (n, f_n))$. $\square$

Before giving the final theorem we may run the algorithm on an example.

Example 1. Let us compute the Frobenius normal form of

$$A = \begin{bmatrix} 2 & 0 & 3 & 0 & 3 & 0 \\ 0 & 2 & -1 & 0 & -1 & 0 \\ 0 & 0 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \in \mathbb{Q}^{6 \times 6}.$$

The first call to $\mathsf{InvFact}$ gives the minimum polynomial $f_1 = x^2 - 3x + 2$ of A. For Toeplitz matrices

$$U_5 = \begin{bmatrix} 1 & 0 & 3 & 1 & -1 \\ 2 & 1 & 0 & 3 & 1 \\ 4 & 2 & 1 & 0 & 3 \\ -2 & 4 & 2 & 1 & 0 \\ 3 & -2 & 4 & 2 & 1 \\ 5 & 3 & -2 & 4 & 2 \end{bmatrix}, V_5 = \begin{bmatrix} 3 & 1 & -4 & 2 & 1 & 2 \\ 0 & 3 & 1 & -4 & 2 & 1 \\ -1 & 0 & 3 & 1 & -4 & 2 \\ 3 & -1 & 0 & 3 & 1 & -4 \\ 1 & 3 & -1 & 0 & 3 & 1 \end{bmatrix}$$

of rank $n-1=5$, the second call to InvFact computes the minimum polynomial g of $A+U_5V_5$:

$$g(x) = x^6 + 28\,x^5 + 612\,x^4 - 16993\,x^3 - 499329\,x^2 + 1938360\,x - 434012.$$

Since g is relatively prime to f_1, $f_n = f_6 = \gcd(f_1, g) = 1$. The call to the recursive function SearchThresholds then leads to the execution:

· SearchThresholds$((1, x^2-3x+2), (6, 1))$
  $k=4$, $f_4 = x-2$
  ·· SearchThresholds$((1, x^2-3x+2), (4, x-2))$
    (*) $k=3$, $f_3 = x-2$
    ··· SearchThresholds$((1, x^2-3x+2), (3, x-2))$
      $k=2$, $f_2 = x-2$
      ···· SearchThresholds$((1, x^2-3x+2), (2, x-2))$
        $f_1 \neq f_2$
        (a) Return $[1, x^2-3x+2, 1, x-2]$
      ···· (d') $f_2 = f_3$, found $[2, x-2]$
      Return $[1, x^2-3x+2, 2, x-2]$
    ··· (d') $f_3 = f_4$, found $[2, x-2]$
    Return $[1, x^2-3x+2, 3, x-2]$
  ·· SearchThresholds$((4, x-2), (6, 1))$
    $k=5$, $f_5 = x-2$
    ··· (d) $f_4 = f_5$, found $[2, x-2]$
    ··· SearchThresholds$((5, x-2), (6, 1))$
      $f_5 \neq f_6$
      (a) Return $[1, x-2, 1, 1]$
    Return $[2, x-2, 1, 1]$
  Return $[1, x^2-3x+2, 4, x-2, 1, 1]$.

The numbers of dots indicate the levels of the recursion and the labels (a), (d) or (d') – corresponding to the definition of SearchThresholds – are given for the deepest levels. The (*) indicates one of the steps (b) and thus one of the internal calls to InvFact. Here, for $k=3$, InvFact(3) generates two Toeplitz matrices of rank $k-1=2$, for instance:

$$U_2 = \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ 0 & 1 \\ 3 & 0 \\ -1 & 3 \\ -1 & -1 \end{bmatrix}, V_2 = \begin{bmatrix} 1 & 6 & -1 & 2 & 1 & 3 \\ 4 & 1 & 6 & -1 & 2 & 1 \end{bmatrix}.$$

The minimum polynomial of $A+U_2V_2$ is

$$g = x^4 - 42\,x^3 + 349\,x^2 + 80\,x - 1236,$$

this implies that $f_3 = \gcd(f_1, g) = x - 2$.

The resulting list $[1, x^2 - 3x + 2, 4, x - 2, 1, 1]$ provides the invariant factors of A:

$$f_1 = x^2 - 3x + 2, f_2 = f_3 = f_4 = f_5 = x - 2, f_6 = 1$$

according to identities (7). $\square$

Theorem 3. *Let $A \in \mathbb{F}^{n\times n}$ and $\mathsf{S} \subset \mathbb{F}$. If A has μ distinct invariant factors then we may probabilistically compute the Frobenius normal form of A by computing $O(\mu n \log(n))$ multiplications of A by vectors and $O\left(\mu n^2 \log^2(n) \log\log(n)\right)$ arithmetic operations in $\mathbb{F}$. Since μ is always less than $3\sqrt{n}/2$, this gives $O(n^{3/2}\log(n))$ multiplications of A by vectors plus $O\left(n^{5/2}\log^2(n)\log\log(n)\right)$ operations in the worst case. The algorithm returns the correct answer with probability at least $1 - O(n^{5/2}\log n)/|\mathsf{S}|$ and requires $O(n)$ storage space in addition to that necessary to store the input matrix.*

Proof. Taking into account the invariant factors equal to one, the number μ of distinct invariant factors must satisfy:

$$(\mu - 1) + (\mu - 2) + \ldots + 1 \leq \deg(\phi_1) + \deg(\phi_2) + \ldots + \deg(\phi_{\mu-1}) \leq n,$$

thus μ is less than $3\sqrt{n}/2$. The execution of Frobenius will generate $O(\sqrt{n}\log(n))$ calls to InvFact. Lemma 5 then imply the cost and storage assertion. In the same way, the failure probability is in $O(n^2 \times \sqrt{n}\log n)$. $\square$

Our strategy, based on rank-k perturbations combined to binary searches, may be applied in other situations. A paper of Eberly, Giesbrecht and the author [5] demonstrates the technique over the integers.

Acknowledgements. Grateful thanks to Erich Kaltofen for his questions.

References

1. D. Bini and V. Pan. *Polynomial and matrix computations.* Birkhäuser, 1994.
2. D.G. Cantor and E. Kaltofen. On fast multiplication of polynomials over arbitrary algebras. *Acta Informatica*, 28(7):693–701, 1991.
3. W. Eberly. Asymptotically Efficient Algorithms for the Frobenius Form. Technical report, Department of Computer Science, University of Calgary, Canada, TR2000/649/01, may 2000.
4. W. Eberly. Black Box Frobenius Decompositions over Small Fields. In *Proc. International Symposium on Symbolic and Algebraic Computation, St Andrews, Scotland*, ACM Press, august 2000.
5. W. Eberly, M. Giesbrecht, and G. Villard. Computing the determinant and Smith form of an integer matrix. In *Proc. 41st Annual IEEE Symposium on Foundations of Computer Science, Redondo Beach, CA*, 2000.
6. F.R. Gantmacher. *Théorie des matrices.* Dunod, Paris, France, 1966.
7. M. Giesbrecht. Nearly optimal algorithms for canonical matrix forms. *SIAM Journal on Computing*, 24(5):948–969, 1995.

8. E. Kaltofen. Challenges of symbolic computation my favorite open problems, 1998. Manuscript submitted for publication, North Carolina State Univ., Dept. Mathematics.
9. E. Kaltofen and V. Pan. Processor efficient parallel solution of linear systems over an abstract field. In *Proc. 3rd Annual ACM Symposium on Parallel Algorithms and Architecture*. ACM-Press, 1991.
10. W. Keller-Gehrig. Fast algorithms for the characteristic polynomial. *Theoretical Computer Science*, 36:309–317, 1985.
11. R. Lambert. *Computational aspects of discrete logarithms.* PhD thesis, University of Waterloo, Ontario, Canada, 1996.
12. V.D. Lidskii. Perturbation theory of non-conjugate operators. *U.S.S.R. Comput. Maths. Math. Phys.*, 1:73–85, 1965.
13. J. Moro, J. Burke, and M. Overton. On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *SIAM J. Matrix Anal. Appl.*, 18:793–817, 1997.
14. M. Newman. *Integral Matrices.* Academic Press, 1972.
15. V.Y. Pan and Z.Q. Chen. The complexity of the matrix eigenproblem. In *The 21rst Annual ACM Symposium on Theory of Computing, Atlanta, Georgia*, pages 507–516. ACM Press, May 1999.
16. V.Y. Pan, Z.Q. Chen, and A. Zheng. The complexity of the algebraic eigenproblem. MSRI Preprint 1998-071, Mathematical Sciences Research Institute, California, USA, dec. 1998.
17. J.H. Reif. Efficient parallel computation of the characteristic polynomial of a sparse, separable matrix. In *The 36th Annual IEEE Conf. Found. Comp. Sc., Milwaukee, WI*, pages 123–132, October 1995.
18. J.T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27:701–717, 1980.
19. A. Storjohann. An $O(n^3)$ algorithm for Frobenius normal form. In *International Symposium on Symbolic and Algebraic Computation, Rostock, Germany*, pages 101–104. ACM Press, August 1998.
20. A. Storjohann. PhD Thesis, Institut für Wissenschaftliches Rechnen, ETH-Zentrum Zürich, Switzerland, to appear, 2000.
21. A. Storjohann and G. Villard. Algorithms for similarity transforms. In *The Seventh Rhine Workshop on Computer Algebra, Bregenz, Austria*, March 2000.
22. D. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Transf. Inform. Theory*, IT-32:54–62, 1986.