ELSEVIER

Journal of Pure and Applied Algebra 156 (2001) 127–145

JOURNAL OF PURE AND APPLIED ALGEBRA

www.elsevier.com/locate/jpaa

Efficient algorithms for computing the characteristic polynomial in a domain

Jounaïdi Abdeljaoued[a,*], Gennadi I. Malaschonok[b]

[a] *Ecole Supérieure des Sciences et Techniques (ESSTT), Dept. Math. Info, University of Tunis II, 5 Avenue Taha Hussein, Tunis 1008, Tunisia*

[b] *Tambov State University, 392622 Tambov, Russia*

Received 1 October 1998; received in revised form 28 June 1999

Communicated by M.-F. Roy

Abstract

Two new sequential methods are given for computing the characteristic polynomial of an endomorphism of a free finite rank-n module over a domain, that require $O(n^3)$ ring operations with exact divisions. © 2001 Elsevier Science B.V. All rights reserved.

MSC: 15A21; 15–04; 65F05

1. Introduction

Let $\mathscr{R}$ be a domain that is, a commutative ring with identity and without zero divisors. We assume that $\mathscr{R}$ is equipped with an algorithm allowing *exact division*. This means that if two elements a and b of $\mathscr{R}$ are given (a being different from zero) such that $b = ac$ with $c \in \mathscr{R}$, then this algorithm can exhibit the exact quotient c. Let $\mathscr{R}^{n\times m}$ denote the set of $n \times m$ matrices with entries in $\mathscr{R}$ and $\mathscr{M}$ be an n-rank free module over $\mathscr{R}$.

Let f be an endomorphism of $\mathscr{M}$ and p_f its characteristic polynomial. If $\boldsymbol{I}_n$ denotes the $n \times n$ identity matrix and if $\boldsymbol{A} = (a_{ij}) \in \mathscr{R}^{n\times n}$ is the matrix of f in terms of some given basis of $\mathscr{M}$, then $p_f(\boldsymbol{x})=\det(\boldsymbol{A}-\boldsymbol{x}\boldsymbol{I}_n)$. One of the important problems of computational commutative algebra is to find effective methods for the computation of characteristic polynomials. Until now, the best theoretical algorithms for solving this

* Corresponding author.

E-mail addresses: jounaidi@esstt.rnu.tn (J. Abdeljaoued), malaschonok@math-iu.tambov.su (G.I. Malaschonok).

PII: S0022-4049(99)00158-9

problem in a *domain* are:

- the Hessenberg sequential method [10,6,1] with $O(n^3)$ arithmetic operations in the fraction field $\mathscr{K}$ of $\mathscr{R}$.
- the Csanky parallel algorithm [7] improved by Preparata and Sarwate [15] and based on the method of the french astronomer Le Verrier [11] with a family of arithmetic circuits of size $O(n^{\alpha+1/2})$ and depth $O(\log^2 n)$ where $O(n^{\alpha})$ is the sequential complexity of the $n \times n$ matrix multiplication in $O(\log n)$ parallel time. [1]

In the case of an *arbitrary commutative ring*, the best parallel algorithms are the Chistov one [5] and the Improved Berkowitz Algorithm [1] with size $O(n^{\alpha+1}\log n)$ and depth $O(\log^2 n)$. The sequential version of the last algorithm has the best practical behaviour for the computation of the characteristic polynomial in an arbitrary commutative ring [2].

The aim of this paper is to describe two new efficient sequential methods with $O(n^3)$ ring operations (additions, subtractions, multiplications and exact divisions). The first one is the Quasi-triangular method and the second one is the tri-diagonal method.

As in the case of Hessenberg's method, they proceed by reducing the given matrix $\boldsymbol{A}$ to a particular upper quasi-triangular (Hessenberg) form, similar to $\boldsymbol{A}$ and therefore having the same characteristic polynomial as $\boldsymbol{A}$. For the Quasi-triangular method, it is possible to have an arbitrary initial location of zero elements in the given matrix $\boldsymbol{A}$.

The main difference with Hessenberg's method lies in the fact that our algorithms allow simplifications at each step (by exact division) while the usual Hessenberg method (on inputs from a ring with exact division) got to operate with fractions that cannot be simplified to a ring element even if one allows exact division. In this way, Hessenberg's method carries over operations with fractions until a final simplification (because the characteristic polynomial has coefficients in the given ring) is achieved. In our algorithms, all operations are then done with elements of the ring $\mathscr{R}$ instead of the much more costly field operations of the Hessenberg method.

Next section presents some basic facts about the quasi-triangular and the tri-diagonal matrices. Section 3 gives a brief exposition of the Dodgson method (also known as the Jordan–Bareiss method) for computing determinants. Sections 4 and 5 give a detailed exposition of our two new algorithms and our main results are stated and proved there. Section 6 is devoted to the study of the sequential complexity of the algorithms. At the end of the article, it is included Appendix A where a simple example is given that shows the transformations which are performed during the reduction step in each algorithm.

2. Quasi-triangular and tri-diagonal matrices

Definition 1. A matrix $\boldsymbol{H} = (h_{ij}) \in \mathscr{R}^{n\times n}$ is said to be upper quasi-triangular (resp. lower quasi-triangular) if $h_{ij} = 0$ for $j - i \geq 2$ (resp. $i - j \geq 2$) that is, the matrix $\boldsymbol{H}$

[1] The current estimation of α is due to Winograd and Coppersmith (1987), with $\alpha < 2.376$.

(resp. the transpose $\boldsymbol{H}^{\mathrm{T}}$) has the following form:

$$\begin{pmatrix} h_{11} & h_{12} & \dots & h_{1,n-1} & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2,n-1} & h_{2n} \\ 0 & h_{32} & \dots & h_{3,n-1} & h_{3n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & h_{n,n-1} & h_{nn} \end{pmatrix}.$$

Definition 2. A tri-diagonal matrix is a matrix which is both upper and lower quasi-triangular.

For abbreviation, we will write QT instead of "quasi-triangular" and TD instead of "tri-diagonal". By expanding the determinant of an upper (resp. lower) QT matrix along its last row (resp. column), it is easy to check the following fact:

Proposition 1. *Let $\boldsymbol{H}_k$ $(1 \le k \le n)$ denote the leading principal submatrix of order k extracted from the quasi-triangular matrix $\boldsymbol{H} = (h_{ij}) \in \mathscr{R}^{n\times n}$ and $\boldsymbol{H}_0 = 1$.*

The determinant $\det(\boldsymbol{H}_n)$ of H is given by the recursion

$$\det(\boldsymbol{H}_n) = h_{nn} \det(\boldsymbol{H}_{n-1}) + \sum_{i=1}^{n-1} h_{i,n} \left(\prod_{j=i+1}^{n} (-h_{j,j-1}) \right) \det(\boldsymbol{H}_{i-1}). \tag{2.1}$$

If H is tri-diagonal, then we have

$$\det(\boldsymbol{H}_n) = h_{nn} \det(\boldsymbol{H}_{n-1}) - h_{n,n-1} h_{n-1,n} \det(\boldsymbol{H}_{n-2}). \tag{2.2}$$

The characteristic polynomial $\boldsymbol{P}_k(\boldsymbol{x})$ of the $k \times k$ QT (resp. TD) matrix $\boldsymbol{H}_k$ is the determinant of the characteristic matrix $\boldsymbol{H}_k - \boldsymbol{x}\boldsymbol{I}_k$ which is also QT (resp. TD).

It follows that the recursive formula (2.1) (resp. (2.2)) of Proposition 1 above gives the following recurrent relationships between the characteristic polynomials $\boldsymbol{P}_k(\boldsymbol{x})$ $(2 \le k \le n)$ where $\boldsymbol{P}_0(\boldsymbol{x}) = 1$ and $\boldsymbol{P}_1(x) = h_{11} - x$:

$$\boldsymbol{P}_k(\boldsymbol{x}) = (h_{kk} - \boldsymbol{x})\boldsymbol{P}_{k-1}(\boldsymbol{x}) + \sum_{i=1}^{k-1} h_{i,k} \left(\prod_{j=i+1}^{k} (-h_{j,j-1}) \right) \boldsymbol{P}_{i-1}(\boldsymbol{x}) \quad \text{if } \boldsymbol{H} \text{ is QT}, \tag{2.3}$$

$$\boldsymbol{P}_k(\boldsymbol{x}) = (h_{kk} - \boldsymbol{x})\boldsymbol{P}_{k-1}(\boldsymbol{x}) - h_{k,k-1} h_{k-1,k} \boldsymbol{P}_{k-2}(\boldsymbol{x}) \quad \text{if } \boldsymbol{H} \text{ is TD}. \tag{2.4}$$

Using these formulae, we also get the total cost of the computations in terms of arithmetic operations. More precisely, let $S(k)$ be the number of additions/subtractions (resp. multiplications) required for computing the sequence of characteristic polynomials

$\boldsymbol{P}_1(x), \boldsymbol{P}_2(x), \ldots, \boldsymbol{P}_k(x)$ for a QT matrix; then (2.3) shows that

$$S(k) = S(k-1) + (k-1) + \sum_{i=1}^{k-1} (k-i)$$

$$\left(\text{resp. } S(k) = S(k-1) + 2(k-1) + \sum_{i=1}^{k-2} i \right).$$

In both cases, we have $S(1) = 0$ and $S(k) = S(k-1) + \frac{1}{2}(k-1)(k+2)$ for $2 \le k \le n$.

We thus get $S(n) = \frac{1}{6}n(n-1)(n+4) \sim \frac{1}{6}n^3$ multiplications in $\mathscr{R}$ and the same number of additions/subtractions for computing the characteristic polynomial $\boldsymbol{P}_n(\boldsymbol{x})$ of an $n \times n$ QT matrix. For a TD matrix, formula (2.4) gives the recursion $S(k) = S(k-1) + 2(k-1)$ where $S(k)$ is the number of additions/subtractions (resp. multiplications) required for computing the characteristic polynomials $\boldsymbol{P}_{k-1}(\boldsymbol{x})$ and $\boldsymbol{P}_k(\boldsymbol{x})$ $(1 \le k \le n)$. Having $S(1) = 0$ in both cases, we conclude that $s(n) = n(n-1)$. We have thus proved the following result.

Proposition 2. *The computation of the characteristic polynomial of a quasi-triangular* (*resp. tri-diagonal*) *matrix with entries in* $\mathscr{R}$ *requires asymptotically* $(\frac{1}{3})n^3$ (resp. $2n^2$) *arithmetic operations of the ring* $\mathscr{R}$ *consisting in* $(\frac{1}{6})n(n-1)(n+4)$ (*resp.* $n(n-1)$) *additions/subtractions and the same number of multiplications.*

3. The Dodgson elimination process

Let $\boldsymbol{A} = (a_{ij}) \in \mathscr{R}^{n \times m}$ be an $n \times m$ matrix over the domain $\mathscr{R}$.

Let $a_{i,j}^{(1)} = a_{ij}$ for $1 \le i \le n$ and $1 \le j \le m$.

For $2 \le k \le \min(n, m)$ and $1 \le j \le m$, let $a_{i,j}^{(k)} = a_{i,j}^{(k-1)}$ if $1 \le i \le k-1$ and let

$$a_{i,j}^{(k)} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,k-1} & a_{1,j} \\ \vdots & \ddots & \vdots & \vdots \\ a_{k-1,1} & \cdots & a_{k-1,k-1} & a_{k-1,j} \\ a_{i,1} & \cdots & a_{i,k-1} & a_{i,j} \end{vmatrix} \quad \text{if } k \le i \le n. \tag{3.1}$$

In the case of $k \le i \le n$ and $1 \le j \le m$, $a_{i,j}^{(k)}$ is the minor of order k obtained by bordering the leading principal minor of order $(k-1)$ of the matrix $\boldsymbol{A}$ with the matching elements of the ith row and the jth column of $\boldsymbol{A}$.

Let r be the rank of $\boldsymbol{A}$. We assume that the r first leading principal minors of $\boldsymbol{A}$ are non-zero. The Dodgson elimination process can be expressed as a sequence of $r-1$ stages which construct the sequence of matrices

$$\boldsymbol{A}_{\mathrm{u}}^{(k)} = (a_{i,j}^{(k)}) \in \mathscr{R}^{n \times m}$$

by using the Sylvester determinant identity [3,4,9]

$$a_{i,j}^{(k+1)}a_{k-1,k-1}^{(k-1)} = \begin{vmatrix} a_{k,k}^{(k)} & a_{k,j}^{(k)} \\ a_{i,k}^{(k)} & a_{i,j}^{(k)} \end{vmatrix} \tag{3.2}$$

(with the convention that $a_{0,0}^{(0)} = 1$). The resulting matrix

$$\boldsymbol{A}_{\mathrm{u}} = \boldsymbol{A}_{\mathrm{u}}^{(r)}$$

is upper triangular because $a_{i,j}^{(k)} = 0$ whenever $1 \le j \le k-1$ and $j < i \le n$ (which is due to Definition (3.1)). See [8,13,14,9,4,16,12] for more details.

An equivalent way of stating the Dodgson process is to apply the recurrence relation

$$\boldsymbol{A}_{\mathrm{u}}^{(k+1)} = D_{k-1}^{-1}\tilde{L}_k\boldsymbol{A}_{\mathrm{u}}^{(k)}$$

(true for all k between 1 and $r-1$) in which the following notations are used:

(1) $\boldsymbol{I}_k$ denotes the identity matrix of order k (for $1 \le k \le n$).

(2) $\delta_k = a_{k,k}^{(k)}$ denotes the pivot element of the kth step of the Dodgson elimination process (for $1 \le k \le r-1$) and $\delta_0 = a_{0,0}^{(0)} = 1$.

(3)

$$v_k = \begin{pmatrix} a_{k+1,k}^{(k)} \\ \vdots \\ a_{n,k}^{(k)} \end{pmatrix}$$

is the submatrix of the matrix $\boldsymbol{A}_{\mathrm{u}}^{(k)}$ obtained by taking the $n-k$ last elements of the kth column (here $1 \le k \le n-1$).

(4) $\tilde{L}_1 = \tilde{\ell}_1$ and

$$\tilde{L}_k = \begin{pmatrix} \boldsymbol{I}_{k-1} & 0 \\ 0 & \tilde{\ell}_k \end{pmatrix}$$

for $2 \le k \le r-1$ with

$$\tilde{\ell}_k = \begin{pmatrix} 1 & 0 \\ -v_k & \delta_k\boldsymbol{I}_{n-k} \end{pmatrix}$$

for $1 \le k \le r-1$ ($\tilde{\ell}_k$ being a square matrix of order $n-k+1$).

(5)

$$D_k = \begin{pmatrix} \boldsymbol{I}_{k+1} & 0 \\ 0 & \delta_k\boldsymbol{I}_{n-k-1} \end{pmatrix}$$

for $0 \le k \le r-1$ ($\delta_0 = 1$ gives $D_0 = \boldsymbol{I}_n$).

The Dodgson process can be extended to $n-1$ stages instead of the $r-1$ effective ones by putting $\delta_k = 1$ and $\tilde{L}_k = D_k = \boldsymbol{I}_n$ for $r \le k \le n-1$.

We shall call this elimination process the row-Dodgson process.

Similarly, applied to the columns of the matrix $\boldsymbol{A}$ the Dodgson process yields a lower trianglular matrix $\boldsymbol{A}_{\mathrm{l}} = \boldsymbol{A}_{\mathrm{l}}^{(r)}$ due to the recurrence relation $\boldsymbol{A}_{\mathrm{l}}^{(k+1)} = \boldsymbol{A}_{\mathrm{l}}^{(k)}\tilde{U}_kC_{k-1}^{-1}$, where $\tilde{U}_k$ and C_k are defined analogously to $\tilde{L}_k$ and D_k. However, we should note that if matrix $\boldsymbol{A}$ is not square ($n \neq m$), the matrix $\boldsymbol{I}_{m-k}$ should replace $\boldsymbol{I}_{n-k}$ in the definition of $\tilde{U}_k$ and C_k.

We shall call this elimination process the column-Dodgson process.

We now have the relations $\boldsymbol{A}_{\mathrm{u}} = \tilde{L}\boldsymbol{A}$ and $\boldsymbol{A}_{\mathrm{l}} = \boldsymbol{A}\tilde{U}$ where $\tilde{L}$ (resp. $\tilde{U}$) is the product of the matrices associated to the successive steps of the Dodgson process on the rows (resp. columns) of $\boldsymbol{A}$. In the case where we need to search for a non-zero pivot, the matrix corresponding to this step will be multiplied by a permutation matrix.

Now, using the above notations, we associate the $n-1$ operators

$$\tilde{\boldsymbol{L}}_k : \mathscr{R}^{n\times m} \to \mathscr{R}^{n\times n} \quad (1 \le k \le n-1)$$

to the row-Dodgson process over $n \times m$ matrices with entries in the domain $\mathscr{R}$.

$\tilde{\boldsymbol{L}}_k$ is defined by $\tilde{\boldsymbol{L}}_k(\boldsymbol{A}) = \tilde{L}_k$, where $\tilde{L}_k$ is the matix defined from matrix $\boldsymbol{A}$ as in item (4) above.

4. The quasi-triangular algorithm (QTA)

The main idea of the QTA is to apply the Dodgson elimination process for reducing a given square matrix $\boldsymbol{A} = (a_{ij}) \in \mathscr{R}^{n\times n}$ to a similar QT matix. It is assumed that the integer n is not smaller than 3.

4.1. QTA without pivoting

Let us consider the following partition of the matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \boldsymbol{h} \\ \boldsymbol{v} & \boldsymbol{b} \end{pmatrix}$$

where $\boldsymbol{h} \in \mathscr{R}^{1\times(n-1)}$, $\boldsymbol{v} \in \mathscr{R}^{(n-1)\times 1}$, $\boldsymbol{b} \in \mathscr{R}^{(n-1)\times(n-1)}$.

Let $\boldsymbol{G}$ be some square matrix of order $n-1$ with entries in $\mathscr{R}$. We apply the row-Dodgson triangularization process to the rectangular matrix $[\boldsymbol{v}, \boldsymbol{bG}]$ which belongs to $\mathscr{R}^{(n-1)\times n}$. The columns of $\boldsymbol{G}$ are to be conveniently chosen one by one and step by step during this process. Using the notations above, we consider the matrices $\tilde{L}_k = \tilde{\boldsymbol{L}}_k([\boldsymbol{v}, \boldsymbol{bG}])$ $(1 \le k \le n-2)$ as defined in the previous section.

The first step of the Dodgson process amounts to leftmultiplying $\boldsymbol{A}$ by $D_0^{-1}\tilde{L}_1$ where $D_0 = \boldsymbol{I}_{n-1}$ and

$$\tilde{L}_1 = \begin{pmatrix} 1 & 0 \\ -v_1 & \delta_1 \boldsymbol{I}_{n-2} \end{pmatrix}.$$

Here $\delta_1 = a_{21}$ and

$$v_1 = \begin{pmatrix} a_{31} \\ \vdots \\ a_{n1} \end{pmatrix}.$$

It is clear that the matrix $\tilde{L}_1$ is determined by the first column $\boldsymbol{v}$ of $[\boldsymbol{v}, \boldsymbol{bG}]$ and does not depend on the columns of $\boldsymbol{G}$. We can thus choose the first column of $\boldsymbol{G}$ so that the first column of $\boldsymbol{bG}$ is equal to the first column of $\boldsymbol{b}L_1$ where

$$L_1 = \begin{pmatrix} \delta_1 & 0 \\ v_1 & \boldsymbol{I}_{n-2} \end{pmatrix}.$$

The second step of the considered Dodgson process amounts to leftmultiplying by $D_1^{-1}\tilde{L}_2$ where

$$D_1 = \begin{pmatrix} \boldsymbol{I}_2 & 0 \\ 0 & \delta_1 \boldsymbol{I}_{n-3} \end{pmatrix} \quad \text{and} \quad \tilde{L}_2 = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{\ell}_2 \end{pmatrix}$$

with

$$\tilde{\ell}_2 = \begin{pmatrix} 1 & 0 \\ -v_2 & \delta_2 \boldsymbol{I}_{n-3} \end{pmatrix}.$$

The matrix $\tilde{L}_2$ only depends on the first two columns of $[\boldsymbol{v}, \boldsymbol{bG}]$ and does not depend on the last $n-2$ columns of $\boldsymbol{bG}$. One can choose the second column of $\boldsymbol{G}$ so that the first two columns of $\boldsymbol{bG}$ are the same as those of $\boldsymbol{b}L_1L_2$ where

$$L_2 = \begin{pmatrix} 1 & 0 \\ 0 & \ell_2 \end{pmatrix}$$

with

$$\ell_2 = \begin{pmatrix} \delta_2 & 0 \\ v_2 & \boldsymbol{I}_{n-3} \end{pmatrix}.$$

We continue in the same way in order to obtain the matrices which correspond to the kth step of the Dodgson process $(3 \le k \le n-2)$ that is, the matrices

$$D_{k-1} = \begin{pmatrix} \boldsymbol{I}_{k+1} & 0 \\ 0 & \delta_{k-1} \boldsymbol{I}_{n-k-2} \end{pmatrix}, \qquad \tilde{L}_k = \begin{pmatrix} \boldsymbol{I}_{k-1} & 0 \\ 0 & \tilde{\ell}_k \end{pmatrix}$$

where

$$\tilde{\ell}_k = \begin{pmatrix} 1 & 0 \\ -v_k & \delta_k \boldsymbol{I}_{n-1-k} \end{pmatrix} \quad \text{and} \quad L_k = \begin{pmatrix} \boldsymbol{I}_{k-1} & 0 \\ 0 & \ell_k \end{pmatrix}$$

with

$$\ell_k = \begin{pmatrix} \delta_k & 0 \\ v_k & \boldsymbol{I}_{n-1-k} \end{pmatrix}.$$

These matrices only depend on the first k columns of $[\boldsymbol{v}, \boldsymbol{bG}]$ that is, they only depend on $\boldsymbol{v}$ and the first $k-1$ columns of $\boldsymbol{b}L_1L_2 \cdots L_{k-1}$.

So, we can choose the kth column of $\boldsymbol{G}$ in such a way that the first k columns of $\boldsymbol{bG}$ are the same as the first k columns of $\boldsymbol{b}L_1L_2 \cdots L_k$. At the end of this process and provided that all the pivots δ_k are different from zero (Dodgson without pivoting), we obtain $\boldsymbol{G} = L_1L_2 \cdots L_{n-2}$. Accordingly, the matrix $\tilde{L}[\boldsymbol{v}, \boldsymbol{b}L] = [\tilde{L}\boldsymbol{v}, \tilde{L}\boldsymbol{b}L]$ where $\tilde{L} = D_{n-3}^{-1}\tilde{L}_{n-2} \cdots D_1^{-1}\tilde{L}_2 D_0^{-1}\tilde{L}_1$ and $L = L_1L_2 \cdots L_{n-2}$ has an upper triangular form.

Now, let

$$\tilde{\mathscr{L}}_k = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{L}_k \end{pmatrix}, \qquad \mathscr{L}_k = \begin{pmatrix} 1 & 0 \\ 0 & L_k \end{pmatrix}, \qquad \mathscr{D}_k = \begin{pmatrix} 1 & 0 \\ 0 & D_k \end{pmatrix},$$

$$\tilde{\mathscr{L}} = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{L} \end{pmatrix} \quad \text{and} \quad \mathscr{L} = \begin{pmatrix} 1 & 0 \\ 0 & L \end{pmatrix}.$$

Obviously, $\tilde{\mathscr{L}} = \mathscr{D}_{n-3}^{-1}\tilde{\mathscr{L}}_{n-2}\cdots\mathscr{D}_1^{-1}\tilde{\mathscr{L}}_2\mathscr{D}_0^{-1}\tilde{\mathscr{L}}_1$, $\mathscr{L} = \mathscr{L}_1\mathscr{L}_2\cdots\mathscr{L}_{n-2}$ and the matrix

$$\tilde{\mathscr{L}}\boldsymbol{A}\mathscr{L} = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{L} \end{pmatrix}\begin{pmatrix} a_{11} & \boldsymbol{h} \\ \boldsymbol{v} & \boldsymbol{b} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & L \end{pmatrix} = \begin{pmatrix} a_{11} & \boldsymbol{h}L \\ \tilde{L}\boldsymbol{v} & \tilde{L}\boldsymbol{b}L \end{pmatrix}$$

is upper quasi-triangular because $[\tilde{L}\boldsymbol{v}, \tilde{L}\boldsymbol{b}L]$ is upper triangular. On the other hand, the matrix $\mathscr{T} = \tilde{\mathscr{L}}\mathscr{L}$ is diagonal and non-singular.

Indeed, $\mathscr{T} = \mathrm{diag}(1, \delta_1, \delta_1\delta_2, \ldots, \delta_{n-3}\delta_{n-2}, \delta_{n-2})$. This comes from the fact that $\tilde{\ell}_k\ell_k = \delta_k\boldsymbol{I}_{n-k}$ and that $\tilde{L}_kL_k$ commutes with the matrices L_{k+1} and $\tilde{L}_{k+1}$ for all k between 1 and $n-3$. The matrix $\mathscr{T}^{-1}\tilde{\mathscr{L}}\boldsymbol{A}\mathscr{L}$ is similar to $\boldsymbol{A}$ since $\mathscr{T}^{-1}\tilde{\mathscr{L}}\mathscr{L} = \boldsymbol{I}_n$.

Provided that all the pivot elements are different from 0, we have actually proved the following result.

Theorem 1. *Let $\boldsymbol{A} = (a_{ij}) \in \mathscr{R}^{n\times n}$ and $\boldsymbol{A}^{[k]} = (a_{ij}^{[k]}) \in \mathscr{R}^{n\times n}$ the sequence of matrices recursively defined by $\boldsymbol{A}^{[1]} = \boldsymbol{A}$ and $\boldsymbol{A}^{[k+1]} = \mathscr{D}_{k-1}^{-1}\tilde{\mathscr{L}}_k\boldsymbol{A}^{[k]}\mathscr{L}_k$ (with $a_{k+1,k}^{[k]} \neq 0$) for $1 \leq k \leq n-2$.*

Then the matrix $\boldsymbol{A}^{[n-1]} = \tilde{\mathscr{L}}\boldsymbol{A}\mathscr{L}$ is quasi-triangular and $\mathscr{T}^{-1}\boldsymbol{A}^{[n-1]}$ is similar to $\boldsymbol{A}$.

It follows that the characteristic polynomial of $\boldsymbol{A}$ is

$$p_A(x) = \det(\boldsymbol{A} - x\boldsymbol{I}_n) = \det(\mathscr{T}^{-1}\boldsymbol{A}^{[n-1]} - x\boldsymbol{I}_n) = (\det\mathscr{T})^{-1}\det(\boldsymbol{A}^{[n-1]} - x\mathscr{T}).$$

Thus $p_A(x)$ is obtained by applying the recursive formula (2.1) given in Section 2 to the quasi-triangular matrix $\boldsymbol{A}^{[n-1]} - x\mathscr{T}$ and then performing the exact division of $\det(\boldsymbol{A}^{[n-1]} - x\mathscr{T})$ by $\det\mathscr{T}$.

The sequential process described above, starting with the given matrix $\boldsymbol{A}^{[1]} = \boldsymbol{A}$ and computing the matrices $\boldsymbol{A}^{[2]}, \ldots, \boldsymbol{A}^{[n-1]}$ as well as the characteristic polynomial of $\boldsymbol{A}$, is called the *Quasi-Triangular Algorithm without pivoting*.

This algorithm may be summarized in the following steps:

Reduce to the quasi-triangular form:
- initialize by setting $\mathscr{D}_0 = \Delta_0 = \boldsymbol{I}_n$,
- compute the sequence of matrices $\boldsymbol{A}^{[k+1]} = \mathscr{D}_{k-1}^{-1}\tilde{\mathscr{L}}_k\boldsymbol{A}^{[k]}\mathscr{L}_k$ and that of diagonal ones $\Delta_k = \mathscr{D}_{k-1}^{-1}\tilde{\mathscr{L}}_k\Delta_{k-1}\mathscr{L}_k = \mathrm{diag}(1, \delta_1, \delta_1\delta_2, \ldots, \delta_{k-2}\delta_{k-1}, \delta_k\boldsymbol{I}_{n-k})$ (for k from 1 to $n-2$);
- end with the upper quasi-triangular matrix $\boldsymbol{A}_\mathrm{u} = \boldsymbol{A}^{[n-1]}$ and the diagonal matrix $\mathscr{T} = \Delta_{n-1}$.

Compute the characteristic polynomial of $\boldsymbol{A}$:

- compute the determinant of the quasi-triangular matrix $\boldsymbol{A}_{\mathrm{u}} - \boldsymbol{x}\mathscr{T}$ by using the Hessenberg recursion (2.1);
- divide $\det(\boldsymbol{A}_{\mathrm{u}} - \boldsymbol{x}\mathscr{T})$ by $\det(\mathscr{T})$ for getting the characteristic polynomial of the input matrix $\boldsymbol{A}$.

All divisions in this algorithm are exact because the reduction step amounts to performing successively the Dodgson elimination process with exact divisions to the rows of matrix $\boldsymbol{A}\mathscr{L}_1 \cdots \mathscr{L}_{n-1}$ and the last division by $\det(\mathscr{T})$ is also exact because the quotient is nothing but the characteristic polynomial of $\boldsymbol{A} \in \mathscr{R}^{n\times n}$.

4.2. QTA with pivoting

Theorem 1 is still true if we drop the assumption $a^{[k]}_{k+1,k} \neq 0$ (note that $a^{[k]}_{k+1,k} = \delta_k$). However, if the pivot element δ_k is equal to 0 for some k $(1 \le k \le n-2)$ and if $v_k = 0$, then we must put $\boldsymbol{A}^{[k+1]} = \boldsymbol{A}^{[k]}$. This means that the kth step of the Dodgson process must be skipped. Now, if $\delta_k = 0$ with $v_k \neq 0$ and $a^{[k]}_{ik} \neq 0$ for some $i \ge k+2$, then we set $\delta_k = a^{[k]}_{ik}$ and $\boldsymbol{A}^{[k+1]} = \mathscr{D}^{-1}_{k-1}\tilde{\mathscr{L}}_k P_{i,k+1}\boldsymbol{A}^{[k]}P_{i,k+1}\mathscr{L}_k$, where $P_{i,j}$ denotes the permutation matrix obtained from the identity matrix $\boldsymbol{I}_n$ by swapping the ith and jth rows (or columns) of $\boldsymbol{I}_n$.

5. Tri-diagonal algorithm (TDA)

5.1. TDA without pivoting

The main step of the algorithm we are going to describe here consists in reducing the given matrix $\boldsymbol{A} = (a_{ij}) \in \mathscr{R}^{n\times n}$ to a tri-diagonal matrix which is similar to $\boldsymbol{A}$.

The idea is to alternate row and column transformations by using the Dodgson elimination method so that Sylvester's identities (3.2) give exact divisions at each step of the process. Starting with the matrix $\boldsymbol{A} = \boldsymbol{A}^{[1]}$ and $\mathscr{D}_0 = \mathscr{T}_0 = \mathscr{S}_0 = \boldsymbol{I}_n$, the kth stage of the tri-diagonal algorithm (TDA) consists in computing the matrix $\boldsymbol{B}^{[k]} = \mathscr{D}^{-1}_{k-1}\tilde{\mathscr{L}}_k\mathscr{T}^{-1}_{k-1}\boldsymbol{A}^{[k]}\mathscr{L}_k$ by using the kth step of the Dodgson row process and then performing the kth step of the Dodgson column process on the matrix $\boldsymbol{B}^{[k]}$ that is, computing the matrix $\boldsymbol{A}^{[k+1]} = \mathscr{U}_k\boldsymbol{B}^{[k]}\mathscr{S}^{-1}_{k-1}\tilde{\mathscr{U}}_k\mathscr{C}^{-1}_{k-1}$.

In consequence, for $1 \le k \le n-2$: $\boldsymbol{A}^{[k+1]} = (\mathscr{U}_k\mathscr{D}^{-1}_{k-1}\tilde{\mathscr{L}}_k\mathscr{T}^{-1}_{k-1})\,\boldsymbol{A}^{[k]}(\mathscr{L}_k\mathscr{S}^{-1}_{k-1}\tilde{\mathscr{U}}_k\mathscr{C}^{-1}_{k-1})$.

Let $\delta_k = a^{[k]}_{k+1,k}$ (resp. $\eta_k = b^{[k]}_{k,k+1}$) be the pivot elements of the Dodgson row process (resp. column process). Then the matrices $\mathscr{U}_k$, $\tilde{\mathscr{U}}_k$ are defined by

$$\mathscr{U}_k = \begin{pmatrix} \boldsymbol{I}_k & 0 \\ 0 & u_k \end{pmatrix}, \qquad \tilde{\mathscr{U}}_k = \begin{pmatrix} \boldsymbol{I}_k & 0 \\ 0 & \tilde{u}_k \end{pmatrix}$$

where

$$u_k = \begin{pmatrix} \eta_k & h_k \\ 0 & \boldsymbol{I}_{n-k-1} \end{pmatrix} \quad \text{and} \quad \tilde{u}_k = \begin{pmatrix} 1 & -h_k \\ 0 & \eta_k\boldsymbol{I}_{n-k-1} \end{pmatrix}$$

with

$$h_k = [b^{[k]}_{k,k+2}, \dots, b^{[k]}_{k,n}] \in \mathscr{R}^{1\times(n-k-1)}.$$

For $1 \le k \le n-2$, the matrix $\mathscr{T}_k$ (resp. $\mathscr{S}_k$) is obtained from $\boldsymbol{I}_n$ by replacing the $(k+1)$th (resp. the $(k+2)$th) diagonal element by δ_k. Note that leftmultiplying (resp. rightmultiplying) by $\mathscr{T}_{k-1}^{-1}$ (resp. $\mathscr{S}_{k-1}^{-1}$) amounts to dividing the kth row (resp. the $(k+1)$th column) of the multiplied matrix by δ_{k-1} $(0 \le k \le n-2)$ with the convention that $\delta_0 = \eta_0 = 1$.

Putting $\mathscr{Y}_k = \mathscr{U}_k \mathscr{D}_{k-1}^{-1} \tilde{\mathscr{L}}_k \mathscr{T}_{k-1}^{-1}$, $\mathscr{Z}_k = \mathscr{L}_k \mathscr{S}_{k-1}^{-1} \tilde{\mathscr{U}}_k \mathscr{C}_{k-1}^{-1}$, we get $\boldsymbol{A}^{[k+1]} = \mathscr{Y}_k \boldsymbol{A}^{[k]} \mathscr{Z}_k$ and

$$\boldsymbol{A}^{[n-1]} = (\mathscr{Y}_{n-2}\mathscr{Y}_{n-3}\cdots\mathscr{Y}_1)\ \boldsymbol{A}\ (\mathscr{Z}_1\mathscr{Z}_2\cdots\mathscr{Z}_{n-2}). \tag{5.1}$$

The fact that the obtained matrix $\boldsymbol{A}^{[n-1]}$ is tri-diagonal and that all divisions occurring in this process are exact comes from the Sylvester identities used during the Dodgson elimination process. Now, the matrix $\boldsymbol{A}_{\mathrm{t}} = \mathscr{T}_{n-2}^{-1}\boldsymbol{A}^{[n-1]}\mathscr{S}_{n-2}^{-1}$ is also tri-diagonal and we have $\boldsymbol{A}_{\mathrm{t}} = \mathscr{Y}\boldsymbol{A}\mathscr{Z}$ where $\mathscr{Y} = \mathscr{T}_{n-2}^{-1}(\mathscr{Y}_{n-2}\mathscr{Y}_{n-3}\cdots\mathscr{Y}_1)$ and $\mathscr{Z} = (\mathscr{Z}_1\mathscr{Z}_2\cdots\mathscr{Z}_{n-2})\mathscr{S}_{n-2}^{-1}$.

Let us prove that the matrix $\mathscr{Y}\mathscr{Z}$ is diagonal and equal to the matrix

$$\mathscr{T} = \mathrm{diag}(1, \eta_1, \eta_1\eta_2, \dots, \eta_{n-3}\eta_{n-2}, \eta_{n-2}). \tag{5.2}$$

More precisely, for $1 \le k \le n-2$, let $\Delta_k = \mathscr{Y}_k\mathscr{Y}_{k-1}\cdots\mathscr{Y}_1\mathscr{Z}_1\mathscr{Z}_2\cdots\mathscr{Z}_k$ and let $\Lambda_k \in \mathscr{R}^{k\times k}$ and $\Gamma_k \in \mathscr{R}^{(n-k)\times(n-k)}$ denote the diagonal matrices

$$\Lambda_k = \mathrm{diag}(1, \eta_1, \eta_1\eta_2, \dots, \eta_{k-2}\eta_{k-1}) \quad \text{and} \quad \Gamma_k = \begin{pmatrix} \delta_k\eta_k\eta_{k-1} & 0 \\ 0 & \delta_k\eta_k\boldsymbol{I}_{n-k-1} \end{pmatrix}.$$

We will establish by induction on k that

$$\Delta_k = \begin{pmatrix} \Lambda_k & 0 \\ 0 & \Gamma_k \end{pmatrix}. \tag{5.3}$$

Obviously, this is true for $k=1$ since

$$\Delta_1 = \begin{pmatrix} 1 & 0 \\ 0 & \delta_1\eta_1\boldsymbol{I}_{n-1} \end{pmatrix}.$$

Assuming that (5.3) holds for $k-1 \ge 1$, we will prove it for k.

We have

$$\Delta_k = \mathscr{Y}_k\Delta_{k-1}\mathscr{Z}_k = \mathscr{U}_k\mathscr{D}_{k-1}^{-1}\tilde{\mathscr{L}}_k\mathscr{T}_{k-1}^{-1}\begin{pmatrix} \Lambda_{k-1} & 0 \\ 0 & \Gamma_{k-1} \end{pmatrix}\mathscr{L}_k\mathscr{S}_{k-1}^{-1}\tilde{\mathscr{U}}_k\mathscr{C}_{k-1}^{-1}.$$

The following successive computations can be easily checked:

$$\mathscr{T}_{k-1}^{-1}\Delta_{k-1} = \begin{pmatrix} \Lambda_k & 0 \\ 0 & \delta_{k-1}\eta_{k-1}\boldsymbol{I}_{n-k} \end{pmatrix},$$

$$\tilde{\mathscr{L}}_k(\mathscr{T}_{k-1}^{-1}\Delta_{k-1})\mathscr{L}_k = \begin{pmatrix} \Lambda_k & 0 \\ 0 & \delta_{k-1}\eta_{k-1}\delta_k\boldsymbol{I}_{n-k} \end{pmatrix},$$

$$\mathscr{D}_{k-1}^{-1}(\tilde{\mathscr{L}}_k\mathscr{T}_{k-1}^{-1}\Delta_{k-1}\mathscr{L}_k)\mathscr{S}_{k-1}^{-1} = \begin{pmatrix} \Lambda_k & 0 \\ 0 & \eta_{k-1}\delta_k \boldsymbol{I}_{n-k} \end{pmatrix},$$

$$\mathscr{U}_k(\mathscr{D}_{k-1}^{-1}\tilde{\mathscr{L}}_k\mathscr{T}_{k-1}^{-1}\Delta_{k-1}\mathscr{L}_k\mathscr{S}_{k-1}^{-1})\tilde{\mathscr{U}}_k = \begin{pmatrix} \Lambda_k & 0 \\ 0 & \eta_{k-1}\eta_k\delta_k \boldsymbol{I}_{n-k} \end{pmatrix},$$

and finally

$$(\mathscr{U}_k\mathscr{D}_{k-1}^{-1}\tilde{\mathscr{L}}_k\mathscr{T}_{k-1}^{-1}\Delta_{k-1}\mathscr{L}_k\mathscr{S}_{k-1}^{-1}\tilde{\mathscr{U}}_k)\mathscr{C}_{k-1}^{-1} = \begin{pmatrix} \Lambda_k & 0 \\ 0 & \Lambda_k \end{pmatrix} = \Delta_k.$$

Now,

$$\Gamma_{n-2} = \begin{pmatrix} \delta_{n-2}\eta_{n-2}\eta_{n-3} & 0 \\ 0 & \delta_{n-2}\eta_{n-2} \end{pmatrix}$$

and multiplying the $n \times n$ diagonal matrix Δ_{n-2} by $\mathscr{T}_{n-2}^{-1}$ on the left and by $\mathscr{S}_{n-2}^{-1}$ on the right amounts to dividing the two last rows (or columns) of the given matrix by δ_{n-2}.

From this and (5.3), it follows that

$$\mathscr{Y}\mathscr{Z} = \mathscr{T}_{n-2}^{-1}\begin{pmatrix} \Lambda_{n-2} & 0 \\ 0 & \Gamma_{n-2} \end{pmatrix}\mathscr{S}_{n-2}^{-1} = \operatorname{diag}(1, \eta_1, \eta_1\eta_2, \ldots, \eta_{n-3}\eta_{n-2}, \eta_{n-2}) = \mathscr{T}.$$

Using the notations above, we have thus proved:

Theorem 2. *Let* $\boldsymbol{A} = (a_{ij}) \in \mathscr{R}^{n\times n}$ *and* $\boldsymbol{A}^{[k]} = (a_{ij}^{[k]}) \in \mathscr{R}^{n\times n}$ *the sequence of matrices recursively defined by* $\boldsymbol{A}^{[1]} = \boldsymbol{A}$ *and* $\boldsymbol{A}^{[k+1]} = \mathscr{Y}_k\boldsymbol{A}^{[k]}\mathscr{Z}_k$ *for* $1 \le k \le n-2$. *Let* $\boldsymbol{A}_{\rm t} = \mathscr{T}_{n-2}^{-1}\,\boldsymbol{A}^{[n-1]}\mathscr{S}_{n-2}^{-1}$.

Then the matrix $\mathscr{T}^{-1}\boldsymbol{A}_{\rm t}$ *is tri-diagonal and similar to* $\boldsymbol{A}$, *provided that all the pivot elements are different from zero.*

Accordingly, the tri-diagonal algorithm (without pivoting) for computing the characteristic polynomial of a given matrix $\boldsymbol{A} \in \mathscr{R}^{n\times n}$ consists in the two following steps:

Reduce to the tri-diagonal form: Starting with the matrix $\boldsymbol{A}^{[1]} = \boldsymbol{A}$,

- compute the sequence $\boldsymbol{A}^{[k+1]} = \mathscr{Y}_k\boldsymbol{A}^{[k]}\mathscr{Z}_k$ for k from 1 to $n-2$ and the diagonal matrix $\mathscr{T}$;
- divide the $(n-1)$th row and the nth column of $\boldsymbol{A}^{[n-1]}$ by δ_{n-2} for getting $\boldsymbol{A}_{\rm t}$ (these are exact divisions).

Compute the characteristic polynomial of $\boldsymbol{A}$:

- compute the determinant of the tri-diagonal matrix $\boldsymbol{A}_{\rm t} - \boldsymbol{x}\mathscr{T}$ by using the Hessenberg recursion (2.2);
- divide $\det(\boldsymbol{A}_{\rm t} - \boldsymbol{x}\mathscr{T})$ by $\det(\mathscr{T})$ (exact divisions) for getting the characteristic polynomial of $\boldsymbol{A}$.

5.2. TDA with partial pivoting

If for some k $(1 \le k \le n-2)$ both $a_{k+1,k}^{[k]}$ and v_k are zero, then we set $\boldsymbol{B}^{[k]} = \boldsymbol{A}^{[k]}$. This amounts to skipping the kth step of the row-Dodgson process.

If $b^{[k]}_{k,k+1}=0$ and $h_k=0$, then we put $\boldsymbol{A}^{[k+1]}=\boldsymbol{B}^{[k]}$ that is, we skip the column-Dodgson kth step of our algorithm.

If $a^{[k]}_{k+1,k}=0$ and $v_k \neq 0$ with $a^{[k]}_{ik} \neq 0$ for some $i \geq k+2$, then we take $\delta_k = a^{[k]}_{ik}$ and we swap the rows i and $k+1$ and the columns i and $k+1$ of the matrix $\boldsymbol{A}^{[k]}$ before continuing. This means that we replace $\boldsymbol{A}^{[k]}$ by $\boldsymbol{P}_{i,k+1}\boldsymbol{A}^{[k]}\boldsymbol{P}_{i,k+1}$ where $\boldsymbol{P}_{i,k+1}$ is the permutation matrix associated to these row and column permutations.

In these cases, Theorem 2 still holds and yields the algorithm that we call "*TDA with partial pivoting*". There is only one case where we cannot assure the existence of a general criteria for moving the non-zero pivot to the appropriate location without losing the zeros of the previous step. It is the case when $b^{[k]}_{k,k+1}=0$ and $h_k \neq 0$ for some k $(1 \leq k \leq n-2)$.

6. Complexity of the algorithms

The reduction to a similar quasi-triangular or tri-diagonal matrix is the most expensive step for each of the two algorithm QTA and TDA.

The kth step $(1 \leq k \leq n)$ of this reduction to a QT (resp. TD) matrix consists in computing the matrix $A^{[k+1]} = \mathscr{D}^{-1}_{k-1}\tilde{\mathscr{L}}_k\boldsymbol{A}^{[k]}\mathscr{L}_k$ (resp. $\boldsymbol{A}^{[k+1]} = \mathscr{Y}_k\boldsymbol{A}^{[k]}\mathscr{Z}_k$) and the diagonal matrix Δ_k. The matrix $\boldsymbol{A}^{[k+1]}$ is obtained from the matrix $\boldsymbol{A}^{[k]}$ by using $n(n-k-1)+(n-k)(n-k-1)$ (resp. $2(n-k-1)(2n-k)$) additions/subtractions, $n(n-k)+2(n-k)(n-k-1)$ (resp. $2n(n-k)+4(n-k)(n-k-1)$) multiplications and $(n-k)(n-k-1)$ (resp. $2(n-k)(n-k-1)$) exact divisions.

It follows that the whole reduction step of the QTA (resp. TDA) requires $(\frac{5}{6})n(n-1)(n-2)$ (resp. $(\frac{5}{3})n(n-1)(n-2)$) additions/subtractions, $(\frac{1}{6})n(7n^2-9n-4)$ (resp. $(\frac{1}{3})n(7n^2-15n+5)$) multiplications and $(\frac{1}{3})n(n-1)(n-2)$ (resp. $(\frac{2}{3})n(n+1)(n-2)$) exact divisions.

Asymptotically, the reduction step takes $(\frac{7}{3})n^3$ ring operations for QTA (with $(\frac{1}{3})n^3$ exact divisions) whereas it takes $(\frac{14}{3})n^3$ (with $(\frac{2}{3})n^3$ exact divisions) for the TDA algorithm.

On the other hand, as stated at the end of Section 2 (Proposition 2) with a slight modification due to the fact that for QTA (resp. TDA) we first compute the determinant of $\boldsymbol{A}_{\mathrm{u}} - \boldsymbol{x}\mathscr{T}$ (resp. $\boldsymbol{A}_{\mathrm{t}} - \boldsymbol{x}\mathscr{T}$) and then divide the result by $\det(\mathscr{T})$, the computation of the characteristic polynomial asymptotically takes $(\frac{1}{3})n^3$ ring operations for QTA and only $2n^2$ for TDA with n exact divisions in both cases. From this we conclude that the sequential time for each of QTA and TDA is $\mathrm{O}(n^3)$. Comparing this result with the known bounds of the Hessenberg method for computing the characteristic polynomial (details on these bounds can be found in [1, pp. 48–52, 6, pp. 55–56], we obtain the following sequential times (see Table 1) where the field operations ($+/-$ and $*$) have been translated into ring operations while field divisions are considered without taking into account gcd computations.

Remark. Processing the successive steps of QTA and TDA leads to a progressive growth of the intermediate coefficients (for integer inputs). A rough idea of this growth

Table 1
Asymptotic complexity

Computation steps	Arithmetic operations	QTA	TDA	Hessenberg †
Reduction step	Additions/subtractions	$\sim (\frac{5}{6})n^3$	$\sim (\frac{5}{3})n^3$	$\sim (\frac{5}{6})n^3$
	Multiplications	$\sim (\frac{7}{6})n^3$	$\sim (\frac{7}{3})n^3$	$\sim (\frac{25}{6})n^3$
	Divisions	$\sim (\frac{1}{3})n^3$	$\sim (\frac{2}{3})n^3$	$\sim (\frac{5}{3})n^3$
	Total 1	$\sim (\frac{7}{3})n^3$	$\sim (\frac{14}{3})n^3$	$\sim (\frac{20}{3})n^3$
Characteristic polynomial	Additions/subtractions	$\sim (\frac{1}{6})n^3$	$\sim n^2$	$\sim (\frac{1}{6})n^3$
	Multiplications	$\sim (\frac{1}{6})n^3$	$\sim n^2$	$\sim (\frac{5}{6})n^3$
	Divisions	$\sim n$	$\sim n$	$\sim (\frac{1}{3})n^3$
	Total 2	$\sim (\frac{1}{3})n^3$	$\sim 2n^2$	$\sim (\frac{4}{3})n^3$
Total cost	(Total 1 + Total 2)	$\sim (\frac{8}{3})n^3$	$\sim (\frac{14}{3})n^3$	$\sim 8n^3$

can be given by the bit-size τ of $\det(\mathscr{T})$, the determinant of the diagonal matrix $\mathscr{T}$ which summarizes the successive transformations (multiplications and exact divisions). Since $\det(\mathscr{T}) = \delta_1^2\delta_2^2\dots\delta_{n-2}^2$ (resp. $\det(\mathscr{T}) = \eta_1^2\eta_2^2\dots\eta_{n-2}^2$) for QTA (resp. TDA), τ is bounded by $2(\beta(\delta_1)+\beta(\delta_2)+\cdots+\beta(\delta_{n-2}))$ (resp. $2(\beta(\eta_1)+\beta(\eta_2)+\cdots+\beta(\eta_{n-2}))$) where $\beta(m)$ denotes the bit-size of the integer m.

Let β_k $(1 \le k \le n-2)$ be the maximum bit-size of the matrix $\boldsymbol{A}^{[k]}$ entries.

Observing that $\beta_k \le 2\beta_{k-1} + \beta(n)$ for QTA and $\beta_k \le 3\beta_{k-1} + 2\beta(n)$ for TDA, one gets the rough upper bounds: $\tau \le (\beta_1 + \beta(n))2^{n-1}$ for QTA and $\tau \le 2(\beta_1+\beta(n))3^{n-1}$ for TDA. Here β_1 is the maximum bit-size of the coefficients in the input matrix $\boldsymbol{A} \in \mathscr{R}^{n\times n}$.

7. Conclusion

It is worthwhile to point out that the main advantage of our two algorithms (QTA and TDA) is that all computations are made in the underlying ring instead of the field of fractions where arithmetic operations are more costly.

This allows us to overcome the major drawbacks of the Hessenberg method which uses divisions without automatic simplification and suffers from the necessity of dealing with the hard problem of gcd computations for avoiding the enormous growth of the intermediate coefficients.

In case of integer coefficients and in spite of the relatively important growth of intermediate results, our algorithms are much well adapted to modular calculations. This may be done by combining the well-known Hadamard inequalities for bounding the coefficients of the characteristic polynomial with the Chinese Remainder Theorem.

Acknowledgements

The authors are greatly indebted to the Mega 98 Committee who made this work possible especially, Marie-Françoise Coste-Roy and Tomas Recio for their advice and their active interest in the publication of this paper. The authors are also grateful to James Davenport and Laureano Gonzalez Vega for their kind help. Special thanks to the referee for several helpful comments and suggestions that improved the clarity of this presentation.

Appendix A An example for QTA and TDA

Below, we show the successive transformations of the reduction step in each of the two algorithms QTA and TDA by applying them to the same matrix

$$A = \begin{bmatrix} -3 & 5 & -4 & 2 & 1 \\ 2 & -1 & 3 & 0 & 2 \\ 5 & 3 & 1 & -3 & 0 \\ 1 & 2 & 4 & -1 & -5 \\ 2 & 1 & -3 & 0 & 2 \end{bmatrix}.$$

A.1. Application of QTA

Starting with $\delta_0 = 1$ and $A^{[1]} = A$, $\delta_1 = 2$, we obtain

$$\mathscr{L}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathscr{D}_0^{-1}\tilde{\mathscr{L}}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & -5 & 2 & 0 & 0 \\ 0 & -1 & 0 & 2 & 0 \\ 0 & -2 & 0 & 0 & 2 \end{bmatrix},$$

$$\mathscr{D}_0^{-1}\tilde{\mathscr{L}}_1\mathscr{L}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix},$$

$$A^{[2]} = \begin{bmatrix} -3 & -6 & 242 & 2 & 1 \\ 2 & 17 & -311 & 0 & 2 \\ 0 & -69 & 1363 & -6 & -10 \\ 0 & 0 & -15\,138 & 96 & 459 \\ 0 & 0 & 6872 & -156 & -260 \end{bmatrix}, \qquad \delta_2 = -69,$$

$$\mathscr{L}_1\mathscr{L}_2=\begin{bmatrix}1&0&0&0&0\\0&2&0&0&0\\0&5&-69&0&0\\0&1&9&1&0\\0&2&-52&0&1\end{bmatrix},\qquad \mathscr{D}_1^{-1}\tilde{\mathscr{L}}_2\mathscr{D}_0^{-1}\tilde{\mathscr{L}}_1=\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&-5&2&0&0\\0&57&-9&-69&0\\0&-61&52&0&-69\end{bmatrix},$$

$$\mathscr{D}_1^{-1}\tilde{\mathscr{L}}_2\mathscr{D}_0^{-1}\tilde{\mathscr{L}}_1\mathscr{L}_1\mathscr{L}_2=\begin{bmatrix}1&0&0&0&0\\0&2&0&0&0\\0&0&-138&0&0\\0&0&0&-69&0\\0&0&0&0&-69\end{bmatrix},$$

$$A^{[3]}=\begin{bmatrix}-3&-6&242&-23\,404&1\\2&17&-311&13\,744&2\\0&-69&1363&22\,108&-10\\0&0&-15\,138&1\,701\,000&459\\0&0&0&295\,517\,616&-11\,328\end{bmatrix},\qquad \delta_3=-15\,138,$$

$$\mathscr{L}=\begin{bmatrix}1&0&0&0&0\\0&2&0&0&0\\0&5&-69&0&0\\0&1&9&-15\,138&0\\0&2&-52&6872&1\end{bmatrix},\qquad \tilde{\mathscr{L}}=\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&-5&2&0&0\\0&57&-9&-69&0\\0&-7706&10\,512&-6872&-15\,138\end{bmatrix},$$

$$\mathscr{T}=\begin{bmatrix}1&0&0&0&0\\0&2&0&0&0\\0&0&-138&0&0\\0&0&0&1\,044\,522&0\\0&0&0&0&-15\,138\end{bmatrix}=\tilde{\mathscr{L}}\mathscr{L},$$

$$A^{[4]}=A_{\mathrm{u}}=\begin{bmatrix}-3&-6&242&-23\,404&1\\2&17&-311&13\,744&2\\0&-69&-1363&22\,108&-10\\0&0&-15\,138&1\,701\,000&459\\0&0&0&295\,517\,616&-11\,328\end{bmatrix}.$$

The characteristic polynomial of A is the same as the characteristic polynomial of $\mathscr{T}^{-1}A_{\mathrm{u}}$ that is, $p_A(x)=-x^5-2x^4-x^3+3x^2-179x+972$.

A.2. Application of TDA

$$\delta_0=1,\qquad \eta_0=1,$$

$$A^{[1]} = \mathrm{A} = \begin{bmatrix} -3 & 5 & -4 & 2 & 1 \\ 2 & -1 & 3 & 0 & 2 \\ 5 & 3 & 1 & -3 & 0 \\ 1 & 2 & 4 & -1 & -5 \\ 2 & 1 & -3 & 0 & 2 \end{bmatrix},$$

$$\mathscr{Y}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 10 & -8 & 4 & 2 \\ 0 & -5 & 2 & 0 & 0 \\ 0 & -1 & 0 & 2 & 0 \\ 0 & -2 & 0 & 0 & 2 \end{bmatrix}, \qquad \mathscr{Z}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 8 & -4 & -2 \\ 0 & 5 & 14 & -10 & -5 \\ 0 & 1 & 4 & -8 & -1 \\ 0 & 2 & 8 & -4 & -8 \end{bmatrix},$$

$\delta_1 = 2, \qquad \eta_1 = -6,$

$$\Delta_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -12 & 0 & 0 & 0 \\ 0 & 0 & -12 & 0 & 0 \\ 0 & 0 & 0 & -12 & 0 \\ 0 & 0 & 0 & 0 & -12 \end{bmatrix},$$

$$A^{[2]} = \begin{bmatrix} -3 & -6 & 0 & 0 & 0 \\ -12 & 140 & 368 & -400 & -164 \\ 0 & -69 & -198 & 174 & 129 \\ 0 & 9 & 6 & -6 & 63 \\ 0 & -52 & -136 & 104 & 52 \end{bmatrix},$$

$$\mathscr{Y}_2\mathscr{Y}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 5 & -4 & 2 & 1 \\ 0 & 19\,182 & -12\,696 & 13\,800 & 5658 \\ 0 & 57 & -9 & -69 & 0 \\ 0 & -61 & 52 & 0 & -69 \end{bmatrix},$$

$$\mathscr{Z}_1\mathscr{Z}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & -242 & 4656 & 1602 \\ 0 & 5 & -398 & 4740 & 1176 \\ 0 & 1 & -148 & -1888 & 1170 \\ 0 & 2 & -86 & -544 & 5646 \end{bmatrix},$$

$\delta_2 = -69, \qquad \eta_2 = -5116$

$$\Delta_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -6 & 0 & 0 & 0 \\ 0 & 0 & -21\,18\,024 & 0 & 0 \\ 0 & 0 & 0 & 353\,004 & 0 \\ 0 & 0 & 0 & 0 & 353\,004 \end{bmatrix},$$

$$A^{[3]} = \begin{bmatrix} -3 & -6 & 0 & 0 & 0 \\ -6 & 70 & -5116 & 0 & 0 \\ 0 & 353\,004 & -29\,186\,448 & 246\,634\,704 & 196\,727\,832 \\ 0 & 0 & 45\,414 & -2\,004\,936 & -2\,968 \\ 0 & 0 & -20\,616 & 1\,485\,296 & 1\,611\,912 \end{bmatrix},$$

$$\mathscr{Y}_3\mathscr{Y}_2\mathscr{Y}_1 =$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 5 & -4 & 2 & 1 \\ 0 & -278 & 184 & -200 & -82 \\ 0 & 19\,628\,657\,424 & 76\,406\,693\,472 & -162\,328\,528\,224 & -129\,481\,126\,992 \\ 0 & 23\,118 & -31\,536 & 20\,616 & 45\,414 \end{bmatrix},$$

$$\mathscr{Z}_1\mathscr{Z}_2\mathscr{Z}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & -242 & -2\,585\,808 & 971\,132\,976 \\ 0 & 5 & -398 & -2\,768\,376 & 1\,197\,839\,664 \\ 0 & 1 & -148 & 1\,592\,208 & -1\,230\,537\,744 \\ 0 & 2 & -86 & -1\,328\,880 & 2\,396\,769\,264 \end{bmatrix},$$

$$\delta_3 = 45\,414, \qquad \eta_3 = 1\,500\,719\,904,$$

$$\Delta_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -6 & 0 & 0 & 0 \\ 0 & 0 & 30\,696 & 0 & 0 \\ 0 & 0 & 0 & -348\,674\,297\,072\,829\,696 & 0 \\ 0 & 0 & 0 & 0 & 68\,153\,693\,720\,256 \end{bmatrix},$$

$$A^{[4]} = \begin{bmatrix} -3 & -6 & 0 & 0 & 0 \\ -6 & 70 & -5\,116 & 0 & 0 \\ 0 & -5\,116 & 422\,992 & 1\,500\,719\,904 & 0 \\ 0 & 0 & 68\,153\,693\,720\,256 & 86\,878\,836\,123\,554\,304 & 1\,113\,140\,858\,205\,990\,016\,512 \\ 0 & 0 & 0 & 197\,200\,190\,016 & -58\,898\,476\,017\,792 \end{bmatrix},$$

$$\mathscr{Y} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 5 & -4 & 2 & 1 \\ 0 & -278 & 184 & -200 & -82 \\ 0 & 432\,216 & 1\,682\,448 & -3\,574\,416 & -2\,851\,128 \\ 0 & 23\,118 & -31\,536 & 20\,616 & 45\,414 \end{bmatrix},$$

$$\mathscr{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & -242 & -2\,585\,808 & 21\,384 \\ 0 & 5 & -398 & -2\,768\,376 & 26\,376 \\ 0 & 1 & -148 & 1\,592\,208 & -27\,096 \\ 0 & 2 & -86 & -1\,328\,880 & 52\,776 \end{bmatrix},$$

$$\tau = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -6 & 0 & 0 & 0 \\ 0 & 0 & 30\,696 & 0 & 0 \\ 0 & 0 & 0 & -7\,677\,683\,028\,864 & 0 \\ 0 & 0 & 0 & 0 & 1\,500\,719\,904 \end{bmatrix} = \mathscr{Y}\mathscr{Z},$$

$$A_{\mathrm{t}} = \begin{bmatrix} -3 & -6 & 0 & 0 & 0 \\ -6 & 70 & -5116 & 0 & 0 \\ 0 & -5116 & 422\,992 & 1\,500\,719\,904 & 0 \\ 0 & 0 & 1\,500\,719\,904 & 1\,913\,040\,827\,136 & 539\,722\,611\,072 \\ 0 & 0 & 0 & 197\,200\,190\,016 & -1\,296\,923\,328 \end{bmatrix} = \mathscr{Y}A\mathscr{Z}.$$

This characteristic polynomial of A is the same as the characteristic polynomial of $\tau^{-1}A_t$ that is,

$$p_A(x) = -x^5 - 2x^4 - x^3 + 3x^2 - 179x + 972.$$

References

[1] J. Abdeljaoued, Algorithmes rapides pour le calcul du polynôme caractéristique, Thèse de l'Université de Franche-Comté, 1997.

[2] J. Abdeljaoued, Berkowitz Algorithm, Maple and computing the characteristic polynomial in an arbitrary commutative ring, Comput. Algebra MapleTech 4 (3) (1997).

[3] A.G. Akritas, E.K. Akritas, G.I. Malaschonok, Various proofs of Sylvester's (determinant) identity, Math. Comput. Simulation 42 (1996) 585–593.

[4] E.H. Bareiss, Sylvester's identity and multistep integer-preserving Gaussian elimination, Math. Comput. 22 (1968) 565–578.

[5] A.L. Chistov, Fast parallel calculation of the rank of matrices over a field of arbitrary characteristic, Proceedings of the FCT'85, Springer Lecture Notes in Computer Science, vol. 199, 1985, pp. 147–150.

[6] H. Cohen, A Course in Computational Algebraic Number Theory, Graduate Texts in Maths, vol. 138, Springer, Berlin, 1993.

[7] L. Csanky, Fast parallel inversion algorithms, SIAM J. Comput. 5 (4) (1976) 618–623.

[8] C.L. Dodgson, Condensation of determinants, being a new and brief method for computing their arithmetic values, Proc. Roy. Soc. Lond. A 15 (1866) 150–155.

[9] E. Durand, Solutions Numériques des Équations Algébriques, Tome II, Masson, Paris, 1961.

[10] D.K. Faddeev, V.N. Faddeeva, Computational Methods of Linear Algebra, Freeman, San Francisco, 1963.

[11] U.J.J. Le Verrier, Sur les variations séculaires des éléments elliptiques des sept planètes principales: Mercure, Venus, La Terre, Mars, Jupiter, Saturne et Uranus, J. Math. Pures Appl. 4 (1840) 220–254.

[12] G.I. Malaschonok, Solution of a system of linear equations in an intergal domain, USSR J. Comput. Math. Math. Phys. 23 (1983) 1497–1500.

[13] G.I. Malaschonok, Algorithms for the solution of systems of linear equations in commutative rings, in: T. Mora, C. Traverso (Eds.), Effective Methods in Algebraic Geometry, Progress in Mathematics, vol. 94, Birkhauser, Basel, 1991, pp. 289–298.

[14] G.I. Malaschonok, Recursive method for the solution of systems of linear equations, in: A. Sydow (Ed.), Computational Mathematics, Proceedings of the 15th IMACS World Congress, vol. I, Wissenschaft & Technik Verlag, Berlin, 1997, pp. 475–480.
[15] F.P. Preparata, D.V. Sarwate, An improved parallel processor bound in fast matrix inversion, Inform. Process. Lett. 7 (3) (1978) 148–150.
[16] T. Sasaki, H. Murao, Efficient Gaussian elimination method for symbolic determinants and linear systems, ACM Trans. Math. Software 8 (4) (1968) 277–289.