IMA Journal of Numerical Analysis (2014) Page 1 of 37
doi:10.1093/imanum/drnxxx

Backward stability of polynomial root-finding using Fiedler companion matrices

FERNANDO DE TERÁN† , FROILÁN M. DOPICO‡ , AND JAVIER PÉREZ§
DEPARTAMENTO DE MATEMÁTICAS, UNIVERSIDAD CARLOS III DE MADRID,
AVDA. UNIVERSIDAD 30, 28911 LEGANÉS, SPAIN

[Received on 26 March 2014]

Computing roots of scalar polynomials as the eigenvalues of Frobenius companion matrices using backward stable eigenvalue algorithms is a classical approach. The introduction of new families of companion matrices allows for the use of other matrices in the root-finding problem. In this paper, we analyze the backward stability of polynomial root-finding algorithms via Fiedler companion matrices. In other words, given a polynomial $p(z)$, the question is to determine whether the whole set of computed eigenvalues of the companion matrix, obtained with a backward stable algorithm for the standard eigenvalue problem, are the set of roots of a nearby polynomial or not. We show that, if the coefficients of $p(z)$ are bounded in absolute value by a moderate number, then algorithms for polynomial root-finding using Fiedler matrices are backward stable, and Fiedler matrices are as good as the Frobenius companion matrices. This allows us to use Fiedler companion matrices with favorable structures in the polynomial root-finding problem. However, when some of the coefficients of the polynomial is large, companion Fiedler matrices may produce larger backward errors than Frobenius companion matrices, although in this case neither Frobenius nor Fiedler matrices lead to backward stable computations. To prove this we obtain explicit expressions for the change, to first order, of the characteristic polynomial coefficients of Fielder matrices under small perturbations. We show that, for all Fiedler matrices except the Frobenius ones, this change involves quadratic terms in the coefficients of the characteristic polynomial of the original matrix, while for the Frobenius matrices it only involves linear terms. We present extensive numerical experiments that support these theoretical results. The effect of balancing these matrices is also investigated.

Keywords: roots of polynomials; eigenvalues; characteristic polynomial; Fiedler companion matrices; backward stability, conditioning

1. Introduction

Let $p(z)$ be a monic polynomial of degree n,

$$p(z) := z^n + \sum_{k=0}^{n-1} a_k z^k, \tag{1.1}$$

with $a_k \in \mathbb{C}$, for $k = 0,\dots,n-1$. The *first* and *second Frobenius companion matrices of* $p(z)$ are defined as

$$C_1 := \begin{bmatrix} -a_{n-1} & -a_{n-2} & \cdots & -a_1 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad C_2 := \begin{bmatrix} -a_{n-1} & 1 & 0 & \cdots & 0 \\ -a_{n-2} & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -a_1 & 0 & 0 & \cdots & 1 \\ -a_0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{1.2}$$

and they have the property that $p(z) = \det(zI - C_1) = \det(zI - C_2)$. Hence, the eigenvalues of both C_1 and C_2 coincide with the roots of $p(z)$. As a consequence, the root-finding problem for scalar monic polynomials (1.1) can be reformulated as an eigenvalue problem. However, these two problems present relevant differences from the numerical

†Corresponding author. Email: fteran@math.uc3m.es
‡Email: dopico@math.uc3m.es
§Email: jpalvaro@math.uc3m.es

point of view. In particular, those regarding conditioning and backward errors. The difference in this setting relies on the fact that, due to perturbations, the companion matrix may become a dense matrix, which has not the structure of a companion matrix any more. In other words, small perturbations of the companion matrix might not correspond to equally small perturbations of the associated polynomial.

To be more precise, a standard way to compute the roots of $p(z)$ is just by computing the eigenvalues of C_1 (or C_2). This is, for instance, the way followed by the MATLAB command `roots`, after balancing the Frobenius matrix. The MATLAB command `roots` then uses the QR-algorithm on the Frobenius matrix to get its eigenvalues. Though this may not be the best way to address the polynomial root-finding problem, from the point of view of efficiency and storage (see, for instance, Moler (1991)), it has been extensively used because of the advantages of the QR algorithm (robustness and backward stability). Nonetheless, to overcome the mentioned drawbacks on the efficiency (measured in number of operations) and storage, several fast variants of the QR method have been proposed, which take advantage of the structure of the companion matrix (see, for instance, Aurentz *et al.* (2013); Bini *et al.* (2004, 2005, 2010); Calvetti *et al.* (2002); Chandrasekaran *et al.* (2008); Gemignani (2007); Van Barel *et al.* (2010); Zhlobich (2012)), but none of them has been proved to be stable. In a different line of research, also variants of C_1, C_2 have been proposed, devoted to improve the accuracy in the case of multiple roots, where the standard companion matrix gives less accurate results than for simple roots (see Brugnano & Trigiante (1995); Niu & Sakurai (2003)). In this paper, we are interested in the backward stability of the root-finding problem solved via an eigenvalue backward stable method, but for a wider class of companion matrices (namely, the Fiedler matrices, see Fiedler (2003)). Our work is motivated by Edelman & Murakami (1995) and Toh & Trefethen (1994), which address related issues for the Frobenius matrices.

Let us first focus on the root-finding problem for $p(z)$ using the first Frobenius companion matrix C_1. Since the QR-algorithm is backward stable, the whole ensemble of computed eigenvalues is the whole ensemble of exact eigenvalues of a matrix C_1+E, where E is a dense matrix such that

$$\|E\| = O(u)\|C_1\|, \tag{1.3}$$

for some matrix norm $\|\cdot\|$, and where u denotes the machine epsilon. However, this does not guarantee that these (computed) eigenvalues are the roots of a nearby polynomial of $p(z)$ or, in other words, that the method is backward stable from the point of view of the polynomials. In this paper, we investigate this issue. In order for the method to be backward stable from the point of view of the polynomials in a normwise sense, the computed eigenvalues should be the exact roots of a polynomial $\widetilde{p}(z)$ such that

$$\frac{\|\widetilde{p}-p\|}{\|p\|} = O(u),$$

for some polynomial norm $\|\cdot\|$. As we will see in Section 3, the backward stability of polynomial root-finding algorithms using companion matrices is closely related to the *conditioning* of the characteristic polynomial under perturbations of these matrices. This conditioning can be measured through the first order term of the Taylor expansion of the coefficients of the characteristic polynomial. In Edelman & Murakami (1995) it has been shown that, if

$$\widetilde{p}(z) = \det(zI - C_1 - E) = z^n + \sum_{k=0}^{n-1} \widetilde{a}_k z^k \tag{1.4}$$

then, to first order in (the entries of) E,

$$\widetilde{a}_k - a_k = \sum_{s=0}^{k} \sum_{j=1}^{n-k-1} a_s E_{j-s+k+1,j} - \sum_{s=k+1}^{n} \sum_{j=n-k}^{n} a_s E_{j-s+k+1,j}. \tag{1.5}$$

If the eigenvalues of C_1 are computed with a backward stable algorithm, it may be proved from (1.5) that, to first order in E, the computed eigenvalues are the exact roots of a polynomial $\widetilde{p}(z)$ as in (1.4) such that

$$\frac{\|\widetilde{p}-p\|}{\|p\|} = O(u)\|p\|, \tag{1.6}$$

with E satisfying (1.3). Note that (1.6) does not imply that computing the roots of $p(z)$ using C_1 (or C_2) is a backward stable method from the point of view of the polynomials, since large values of $\|p\|$ can give large backward errors.

This had been already noticed, for instance, in Lemmonier & Van Dooren (2003), where the authors analyze diagonal scalings of the companion matrix to get small backward errors.

A key advantage in using Frobenius companion matrices in the root-finding problem is that they are easily constructible from the polynomial, without performing any arithmetic operation, by means of a uniform template valid for all polynomials. Any uniform template with these properties is what we mean by a *companion* matrix.

In Fiedler (2003), the author expanded the family of companion matrices associated with the monic polynomial $p(z)$. These matrices were named *Fiedler matrices* in De Terán *et al.* (2010). The family of Fiedler matrices includes C_1 and C_2 but, provided that $n \geqslant 3$, it contains some other different matrices and, in fact, many others when n is large. These matrices provide a new tool that could be used instead of C_1 and C_2 for computing the roots of $p(z)$. Some features of Fiedler matrices have been recently studied. For instance, in De Terán *et al.* (2013) the condition numbers for inversion of different Fiedler matrices have been compared, and it has been proved that, in many cases, some of the new Fiedler matrices have better conditioning than C_1 and C_2. Also, in De Terán *et al.* (2014), Fiedler matrices have been used to get new lower and upper bounds for the modulus of the roots of $p(z)$. We provide the formal definition of Fiedler matrices in Section 2. For the moment, the only relevant information is that, to construct them, we only need to know the polynomial $p(z)$ and to fix a bijection $\sigma : \{0,1,\ldots,n-1\} \to \{1,\ldots,n\}$, and that the Fiedler matrices contain, in different positions, exactly the same entries as C_1 and C_2. We denote the Fiedler matrix associated with the polynomial $p(z)$ and the bijection σ by $M_\sigma(p)$, or M_σ for brevity.

A natural question is whether or not computing the roots of $p(z)$ using a Fiedler matrix M_σ and a backward stable eigenvalue algorithm is backward stable from the point of view of the polynomials, that is, whether or not the computed roots are the exact roots of a polynomial $\widetilde{p}(z)$ such that $\|\widetilde{p} - p\| = O(u)\|p\|$. As it happens with the Frobenius matrices, if we compute the roots of $p(z)$ as the eigenvalues of M_σ using a backward stable algorithm (for instance, the QR algorithm), then the computed roots are the exact eigenvalues of $M_\sigma + E$, where $\|E\| = O(u)\|M_\sigma\|$. However, again, this does not guarantee the backward stability from the point of view of the polynomials. The goal of this paper is to analyze this issue.

In order to accomplish this task we need to know how the coefficients of the characteristic polynomial of M_σ change when the matrix is perturbed as $M_\sigma + E$, with E an arbitrary perturbation having no special structure. This change can be estimated, up to first order in E, through the gradients $\nabla a_k(M_\sigma)$, where $a_k(X) : \mathbb{C}^{n^2} \to \mathbb{C}$ is the kth coefficient of the characteristic polynomial of a matrix $X \in \mathbb{C}^{n\times n}$, considered as a function of its entries. In particular, we find explicitly $\nabla a_k(M_\sigma)$ in terms of the coefficients of $p(z)$. This allows us to get, up to first order, a formula for the variation of the characteristic polynomial of M_σ under small perturbations of M_σ. From this formula, we analyze the backward stability of the polynomial root-finding problem solved by applying backward stable eigenvalue algorithms to Fiedler matrices.

To obtain an expression for $\nabla a_k(M_\sigma)$, we first prove that its coordinates coincide with the entries of the $(k+1)$th coefficient of the matrix polynomial $\mathrm{adj}(zI - M_\sigma)$. Then, we get an explicit formula for $\mathrm{adj}(zI - A)$, for A being a Fiedler matrix M_σ. This is a general theoretical result on Fiedler matrices that may be useful in the future to analyze other features of this family of matrices.

For a precedent on the perturbation analysis of the characteristic polynomial, we refer the reader to Ipsen & Rehman (2008). In that paper, several bounds are derived for the variation of the characteristic polynomial of an arbitrary matrix A under perturbations, in terms of symmetric functions of the singular values of A. The bounds there are very pessimistic for general matrices. However, here we take advantage of the sparsity and the structure of the Fiedler matrices to get more specific bounds depending on the coefficients of $p(z)$.

Throughout this paper, if $A \in \mathbb{C}^{n\times n}$ is a matrix, then $\|A\|_\infty$ denotes the usual matrix ∞-norm (see (Higham, 2002, p. 108)). In particular, for a vector $v = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}^T \in \mathbb{C}^n$, we have $\|v\|_\infty = \max\{|v_1|,\ldots,|v_n|\}$. Similarly, for a polynomial $p(z) = \sum_{k=0}^n a_k z^k$ (not necessarily monic), $\|p\|_\infty$ is the norm on the vector space of scalar polynomials of degree less than or equal to n defined as

$$\|p\|_\infty := \max\{|a_n|, |a_{n-1}|, \ldots, |a_1|, |a_0|\}.$$

Notice that, since we deal in this paper with monic polynomials, $a_n = 1$ and we always have $\|p\|_\infty \geqslant 1$.

The main results of this work are Theorem 3.3 and Corollary 3.2. Theorem 3.3 gives, to first order in E, the coefficients of the characteristic polynomial of $M_\sigma + E$, and Corollary 3.2 tells us that if we compute the roots of a monic polynomial $p(z)$ as the eigenvalues of a Fiedler matrix M_σ other than the Frobenius companion matrices using a backward stable eigenvalue algorithm, then the computed roots are the exact roots of a monic polynomial $\widetilde{p}(z)$ such that

$$\frac{\|\widetilde{p}-p\|_\infty}{\|p\|_\infty} = O(u)\|p\|_\infty^2, \tag{1.7}$$

which implies that computing the roots of $p(z)$ using any of the Fiedler matrices of $p(z)$ is not backward stable if $\|p\|_\infty$ is large. For the Frobenius companion matrices, Corollary 3.2 recovers (1.6). In Section 4 we provide numerical experiments that support this theoretical result.

As a consequence of (1.6) and (1.7) we get the following conclusions:

(C1) From the point of view of the normwise backward errors in the (monic) polynomial $p(z)$, any Fiedler matrix can be used for solving the root-finding problem with the same reliability as Frobenius companion matrices when $\|p\|_\infty = O(1)$. In this case, the root-finding problem solved by applying a backward stable eigenvalue algorithm on any Fiedler companion matrix is a backward stable method.

(C2) However, when $\|p\|_\infty$ is large none of the Fiedler matrices leads to a backward stable algorithm for the root-finding problem and, moreover, any Fiedler matrix other than Frobenius companion matrices may produce much larger backward errors than the ones produced when using Frobenius matrices.

Note, in particular, that since $\|p\|_\infty \geqslant 1$, no Fiedler matrix can improve the behavior of Frobenius matrices in the root-finding problem from the point of view of backward errors. Anyway, the particular structure of some Fiedler matrices can make their use more efficient than the use of classical Frobenius companion matrices. For instance, we could take advantage of the pentadiagonal structure of some Fiedler matrices (which exist for any value of n, see De Terán *et al.* (2010)) to devise structured versions of the *LR* algorithm to get its eigenvalues in $O(n^2)$ flops. However, as for all structured methods for the root-finding problem, stability can not yet be guaranteed.

We have also considered the effect of balancing (see Parlett & Reinsch (1969)) Fiedler companion matrices on the backward errors of the root-finding problem for $p(z)$ using a Fiedler matrix M_σ. The numerical experiments carried out in Section 4 indicate that balancing very often improves the backward errors for general polynomials, including some polynomials for which the backward error without balancing is quite large. However, we prove that, when $|a_{n-1}|$ is much larger than $|a_{n-2}|$, the condition number of $p(z)$ using any balanced Fiedler matrix is large, and so is the backward error. Some experiments on polynomials with $|a_{n-1}|$ much larger than $|a_{n-2}|$ show that, indeed, balancing the Fiedler matrices does not guarantee backward stability for the root-finding polynomial problem.

The paper is organized as follows. In Section 2 we introduce Fiedler matrices and their basic properties. In Section 3 we analyze, to first order, the change of the coefficients of the characteristic polynomial of Fiedler matrices under matrix perturbations, and we connect it with the backward error of the polynomial root-finding problem solved via an eigenvalue algorithm. This section contains the main results of the paper. Section 4 is devoted to numerical experiments that illustrate the theoretical results obtained in Section 3. In Section 5 we provide a geometric interpretation of the change, to first order, of the characteristic polynomial of Fiedler matrices in terms of the orbit space under similarity of these matrices. This is motivated by the one in Edelman & Murakami (1995) for Frobenius companion matrices, and gives a decomposition of $\mathbb{C}^{n\times n}$ as the sum of the tangent space to the similarity orbit of a Fiedler matrix and the Sylvester space of matrices associated to this Fiedler matrix. Section 6 presents a summary of the main contributions of the paper.

2. Fiedler matrices. Definition and basic properties

For a given polynomial $p(z)$ as in (1.1), we define the $n\times n$ matrices

$$M_0 := \begin{bmatrix} I_{n-1} & 0 \\ 0 & -a_0 \end{bmatrix} \quad \text{and} \quad M_k := \begin{bmatrix} I_{n-k-1} & & & \\ & -a_k & 1 & \\ & 1 & 0 & \\ & & & I_{k-1} \end{bmatrix}, \quad k = 1,\dots,n-1, \tag{2.1}$$

which are the basic factors used to build all Fiedler matrices. Here and in the rest of the paper I_j denotes the $j\times j$ identity matrix. In Fiedler (2003) Fiedler matrices are constructed as the product

$$M_{i_1}M_{i_2}\cdots M_{i_n},$$

where $(i_1,i_2,\dots,i_n)$ is any possible permutation of the n-tuple $(0,1,\dots,n-1)$. In order to better express certain key properties of this permutation and the resulting Fiedler matrix, in De Terán *et al.* (2010) the authors index the product of the M_i factors in a slightly different way, as it is described in the following definition.

DEFINITION 2.1 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$, with $n \geqslant 2$, and let M_i, for $i = 0, 1, \ldots, n-1$, be the matrices defined in (2.1). Given any bijection $\sigma : \{0, 1, \ldots, n-1\} \to \{1, \ldots, n\}$, the Fiedler matrix of $p(z)$ associated with σ is the $n \times n$ matrix

$$M_\sigma(p) := M_{\sigma^{-1}(1)} \cdots M_{\sigma^{-1}(n)}. \tag{2.2}$$

We want to notice that $\sigma(i)$ in (2.2) describes the position of the factor M_i in the product $M_{\sigma^{-1}(1)} \cdots M_{\sigma^{-1}(n)}$, i.e., $\sigma(i) = j$ means that M_i is the jth factor in the product. We want to note also that the building factors (2.1) of (2.2) depend also on $p(z)$ (to be precise, they depend on its coefficients). However, in this case we do not write explicitly this dependence for the sake of simplicity. For the same reason, we will also drop the dependence on p in M_σ when there is no risk of confusion (namely, until Section 5).

The family of matrices $\{M_k\}_{k=0}^{n-1}$ satisfies the following commutativity relations

$$M_i M_j = M_j M_i \quad \text{for } |i - j| \neq 1. \tag{2.3}$$

It is proved in Fiedler (2003) that all Fiedler matrices of $p(z)$ are similar, so they have $p(z)$ as characteristic polynomial. Frobenius companion matrices of $p(z)$ are particular cases of Fiedler matrices, namely,

$$C_1 = M_{n-1} M_{n-2} \cdots M_1 M_0 \quad \text{and} \quad C_2 = M_0 M_1 \cdots M_{n-2} M_{n-1}\,.$$

Observe that the matrices M_i are symmetric, and therefore the transpose of any Fiedler matrix is another Fiedler matrix, obtained by reversing the order of the M_i factors in (2.2).

The relations (2.3) imply that some Fiedler matrices associated with different bijections σ are equal. For example, for $n = 3$, the Fiedler matrices $M_0 M_2 M_1$ and $M_2 M_0 M_1$ are equal. These relations suggest that the relative positions of the matrices M_i and M_{i+1} in the product M_σ are of fundamental interest in studying Fiedler matrices. This motivates Definition 2.2, partially introduced in De Terán *et al.* (2010).

DEFINITION 2.2 Let $\sigma : \{0, 1, \ldots, n-1\} \to \{1, \ldots, n\}$ be a bijection.

(a) For $i = 0, \ldots, n-2$, we say that σ has a consecution at i if $\sigma(i) < \sigma(i+1)$ and that σ has an inversion at i if $\sigma(i) > \sigma(i+1)$.

(b) The positional consecution-inversion sequence of σ, denoted by PCIS(σ), is the $(n-1)$-tuple $(v_0, \ldots, v_{n-2})$ such that $v_j = 1$ if σ has a consecution at j and $v_j = 0$ otherwise.

REMARK 2.1 We note that σ has a consecution at i, that is $v_i = 1$, if and only if M_i is to the left of M_{i+1} in the product defining the Fiedler matrix M_σ, while σ has an inversion at i, that is $v_i = 0$, if and only if M_i is to the right of M_{i+1} in M_σ. This simple observation on Definition 2.2 will be used freely.

In order to keep the notation in future sections reasonably simple we introduce the following definitions.

DEFINITION 2.3 Let $\sigma : \{0, 1, \ldots, n-1\} \to \{1, \ldots, n\}$ be a bijection with PCIS$(\sigma) = (v_0, v_1, \ldots, v_{n-2})$, then:

(a) The extended positional consecution-inversion sequence of σ, denoted by EPCIS(σ), is the n-tuple $(v_0, v_1, \ldots, v_{n-1})$, where $v_{n-1} = v_{n-2}$.

(b) For $0 \leqslant i \leqslant j \leqslant n-2$, we set

$$\mathfrak{i}_\sigma(i:j) := \sum_{k=i}^{j} (1 - v_k) \qquad \text{and} \qquad \mathfrak{c}_\sigma(i:j) := \sum_{k=i}^{j} v_k$$

for, respectively, the number of inversions and consecutions of σ from i to j. We also set $\mathfrak{i}_\sigma(i:j) := \mathfrak{c}_\sigma(i:j) := 0$ for $i > j$.

The following immediate identities will be used several times along the paper:

$$\mathfrak{i}_\sigma(i:j) + \mathfrak{c}_\sigma(i:j) = j - i + 1, \quad \text{for } 0 \leqslant i \leqslant j \leqslant n-2, \tag{2.4}$$

$$\mathfrak{i}_\sigma(0:i) + \mathfrak{c}_\sigma(0:j) \leqslant n - 1, \quad \text{for } 0 \leqslant i, j \leqslant n-2. \tag{2.5}$$

We close this section with the following notion, not strictly related to Fiedler matrices, that will be used along the paper.

DEFINITION 2.4 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial of degree n. For $d = 0, 1, \ldots, n$, the degree d Horner shift of $p(z)$ is the polynomial $p_d(z) = z^d + a_{n-1} z^{d-1} + \cdots + a_{n-d+1} z + a_{n-d}$.

Notice that the Horner shifts of $p(z)$ satisfy the following recurrence relation

$$\begin{cases} p_0(z) = 1, \quad \text{and} \\ p_d(z) = z p_{d-1}(z) + a_{n-d}, \quad \text{for } d = 1, 2, \ldots, n. \end{cases} \tag{2.6}$$

3. Backward error, conditioning, and first order perturbation terms of the characteristic polynomial

A natural definition of the normwise *backward error* of the computed roots, $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n$, of the monic polynomial (1.1) via a certain algorithm is

$$\eta_\infty(\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n) := \frac{\|\widetilde{p} - p\|_\infty}{\|p\|_\infty},$$

where $\widetilde{p}(z) = \prod_{i=1}^n (z - \widetilde{\lambda}_i)$. Note that this notion of backward error coincides with the relative distance, in the ∞-norm, between the original polynomial $p(z)$ and the monic polynomial $\widetilde{p}(z)$ whose roots are $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n$. The key in our approach is that the roots are computed as the eigenvalues of a (companion) matrix, A, so that the computed roots are the exact eigenvalues of a certain perturbation of A, say $A + E$. In other words, we have $p(z) = \det(zI - A)$ and, following (1.4) for a general companion matrix A, we also have $\widetilde{p}(z) = \det(zI - (A + E))$. Hence, the difference between $p(z)$ and $\widetilde{p}(z)$ can be measured from the variation of the coefficients of the characteristic polynomial of A under small perturbations of A.

Hence, we consider the kth coefficient of the characteristic polynomial of a matrix $X = [x_{ij}] \in \mathbb{C}^{n \times n}$ as a function of the entries of X, $a_k(X) : \mathbb{C}^{n^2} \to \mathbb{C}$, for $k = 0, 1, \ldots, n-1$. Equivalently:

$$\det(zI - X) = z^n + \sum_{k=0}^{n-1} a_k(X) z^k.$$

The function $a_k(X)$ is a multivariable polynomial function of the entries of the matrix X. Therefore, the first order term in E of its Taylor polynomial centered at A is (see, for instance (Grauert & Fritzsche, 1976, Th. 3.8)) for functions of several complex variables)

$$a_k(A + E) = a_k(A) + \sum_{i,j=1}^n \left. \frac{\partial a_k(X)}{\partial x_{ij}} \right|_{X=A} E_{ij} = a_k(A) + \nabla a_k(A) \cdot \mathrm{vec}(E), \qquad \text{for } k = 0, 1, \ldots, n-1, \tag{3.1}$$

where, for a given $m \times n$ matrix $M = [m_{ij}]$, $\mathrm{vec}(M)$ is the *vectorization* of M, namely, the column vector

$$\mathrm{vec}(M) := [m_{11} \ldots m_{m1}\, m_{12} \ldots m_{m2} \ldots m_{1n} \ldots m_{mn}]^T$$

(see (Horn & Johnson, 1985, Def. 4.2.9), for instance), and

$$\nabla a_k(A) = \left[\left. \frac{\partial a_k(X)}{\partial x_{11}} \right|_{X=A} \cdots \left. \frac{\partial a_k(X)}{\partial x_{n1}} \right|_{X=A} \left. \frac{\partial a_k(X)}{\partial x_{12}} \right|_{X=A} \cdots \left. \frac{\partial a_k(X)}{\partial x_{n2}} \right|_{X=A} \cdots \left. \frac{\partial a_k(X)}{\partial x_{1n}} \right|_{X=A} \cdots \left. \frac{\partial a_k(X)}{\partial x_{nn}} \right|_{X=A} \right].$$

Therefore, to first order in E, we have

$$|a_k(A + E) - a_k(A)| = |\nabla a_k(A) \cdot \mathrm{vec}\ (E)|.$$

For any Fiedler matrix M_σ, we will get an explicit expression of $\nabla a_k(M_\sigma)$ in terms of the entries of M_σ or, equivalently, in terms of the coefficients of its characteristic polynomial $p(z)$. The corresponding expression was given in Edelman & Murakami (1995) for the Frobenius companion matrices, which are particular cases of Fiedler matrices. The general expression we provide here, valid for any Fiedler matrix, requires different techniques to the ones employed in Edelman & Murakami (1995).

The following well-know result (known as Jacobi's formula, see Bhatia & Jain (2009)) provides us a description of the gradient of the determinant. We include a short proof here for completeness.

LEMMA 3.1 Let $A \in \mathbb{C}^{n\times n}$, and consider a small perturbation $A+E$, with $E \in \mathbb{C}^{n\times n}$. Then, the function

$$\begin{array}{rccc} \det: & \mathbb{C}^{n\times n} & \longrightarrow & \mathbb{C} \\ & X & \longmapsto & \det(X), \end{array}$$

is analytic in a neighborhood of A, and

$$\det(A+E) = \det(A) + \operatorname{tr}(\operatorname{adj}(A)\,E) + O(\|E\|^2),$$

where $\|\cdot\|$ is any norm in $\mathbb{C}^{n\times n}$, $\operatorname{adj}(A)$ denotes the adjugate matrix of A (see Bernstein (2009)), and $\operatorname{tr}(B)$ denotes the trace of B.

Proof. The function $\det : \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}$ is clearly analytic in a neighborhood of A, since it is a polynomial function on the entries of $X \in \mathbb{C}^{n\times n}$. Moreover, analogously to (3.1), with the function det instead of a_k, we get

$$\det(A+E) = \det(A) + \nabla\det(A)\cdot \operatorname{vec}(E) + O(\| E \|^2).$$

Now, it is straightforward to check that

$$\left.\frac{\partial \det(X)}{\partial x_{ij}}\right|_{X=A} = (\operatorname{adj}(A))_{ji}$$

(see also (Bernstein, 2009, Fact 10.11.21)). The result now follows from the identity $\operatorname{tr}(AB) = \operatorname{vec}(A^T)^T \cdot \operatorname{vec}(B)$, which is valid for every $A, B \in \mathbb{C}^{n\times n}$. □

As an immediate consequence of Lemma 3.1, applied to $p(z) = \det(zI - A)$, we get Proposition 3.1, which gives a description of the gradient of the coefficients of the characteristic polynomial of A and, as a consequence, an expression for the variation of the characteristic polynomial under small perturbations, up to first order.

PROPOSITION 3.1 Let $A \in \mathbb{C}^{n\times n}$ and $z \in \mathbb{C}$. Let us write the adjoint matrix of $zI - A$ as

$$\operatorname{adj}(zI-A) = \sum_{k=0}^{n-1} z^k P_{k+1}, \tag{3.2}$$

with $P_{k+1} \in \mathbb{C}^{n\times n}$, for $k = 0, 1, \dots, n-1$. Let $a_k(X) : \mathbb{C}^{n^2} \to \mathbb{C}$ be the kth coefficient of the characteristic polynomial of a matrix $X = (x_{ij}) \in \mathbb{C}^{n\times n}$, and let $\nabla a_k(A)$ be the gradient of the function $a_k(X)$ evaluated at A. Then, for $k = 0, 1, \dots, n-1$,

$$\nabla a_k(A) = -\left[\operatorname{vec}\,(P_{k+1}^T)\right]^T.$$

As a consequence, if $A+E$ is a small perturbation of A, with $E \in \mathbb{C}^{n\times n}$, then

$$\det(zI-(A+E)) - \det(zI-A) = -\sum_{k=0}^{n-1} z^k \left[\operatorname{vec}(P_{k+1}^T)\right]^T \cdot \operatorname{vec}(E) + O(\|E\|^2) = -\sum_{k=0}^{n-1} z^k \operatorname{tr}(P_{k+1}E) + O(\|E\|^2),$$

where $\|\cdot\|$ is any norm in $\mathbb{C}^{n\times n}$.

Proof. From Lemma 3.1 and (3.2), we have

$$\begin{aligned} \det(zI-(A+E)) &= \det(zI-A) - \operatorname{tr}(\operatorname{adj}(zI-A)E) + O(\|E\|^2) \\ &= \det(zI-A) - \textstyle\sum_{k=0}^{n-1} z^k \operatorname{tr}(P_{k+1}E) + O(\|E\|^2) \\ &= \det(zI-A) - \textstyle\sum_{k=0}^{n-1} z^k \left[\operatorname{vec}(P_{k+1}^T)\right]^T \cdot \operatorname{vec}(E) + O(\|E\|^2), \end{aligned}$$

and the expression for $\nabla a_k(A)$ follows immediately from this. □

Proposition 3.1 tells us that the variation of the characteristic polynomial of $A \in \mathbb{C}^{n\times n}$ is given, to first order, by the trace of $\operatorname{adj}(zI-A)$. This adjugate matrix is an $n\times n$ matrix whose entries are polynomials of degree at most $n-1$ or, equivalently, a matrix polynomial of size $n\times n$ with degree at most $n-1$. Actually, its degree is exactly $n-1$, because of the identity: $(zI-A)\cdot\operatorname{adj}(zI-A) = \det(zI-A)I_n$. In Section 3.1 we give an explicit expression for the entries of $\operatorname{adj}(zI-A)$, for A being an arbitrary Fiedler matrix M_σ. Then, in Section 3.2, we use this information, following Proposition 3.1, to present an explicit expression for the variation, up to first order, of the coefficients of the characteristic polynomial of M_σ or, in other words, an explicit expression for $\nabla a_k(M_\sigma)$.

3.1 *Adjugate matrix of $zI - M_\sigma$*

The main result of this section is Theorem 3.2, which gives an explicit expression for the adjugate matrix of $zI - M_\sigma$. As we have seen in (3.2), this matrix is not a constant matrix, but a matrix polynomial in the variable z. We use the notation $\mathbb{C}^{n\times n}[z]$ for the set of $n\times n$ matrix polynomials.

An explicit expression for the adjugate in the case of first and second Frobenius companion matrices was already known (see (Gantmacher, 1959, Ch. IV §4) or (Edelman & Murakami, 1995, p. 768)):

$$\operatorname{adj}(zI - C_2) = \begin{bmatrix} p_0(z) \\ p_1(z) \\ \vdots \\ p_{n-1}(z) \end{bmatrix} \begin{bmatrix} z^{n-1} & \cdots & z & 1 \end{bmatrix} - p(z) \begin{bmatrix} 0 & & & & & \\ 1 & 0 & & & & \\ z & 1 & \ddots & & & \\ \vdots & z & \ddots & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \\ z^{n-2} & z^{n-3} & \cdots & z & 1 & 0 \end{bmatrix}, \tag{3.3}$$

and $\operatorname{adj}(zI - C_1) = (\operatorname{adj}(zI - C_2))^T$. Here $p_0(z), \ldots, p_{n-1}(z)$ are the Horner shifts introduced in Definition 2.4. Equation (3.3) has a very particular structure: it is a sum of a rank-1 matrix plus a matrix whose (i, j) entry is of the form $p(z)p_{ij}(z)$, where $p_{ij}(z)$ is a polynomial of degree at most $n-2$. We will prove that this structure is shared also by $\operatorname{adj}(zI - M_\sigma)$, for any Fiedler matrix M_σ. For example, if we consider the Fiedler matrix M_σ of a degree-6 monic polynomial $p(z) = z^6 + \sum_{k=0}^{5} a_k z^k$, with $\operatorname{PCIS}(\sigma) = (1,0,1,0,1)$, we will show that

$$\operatorname{adj}(zI - M_\sigma) = \begin{bmatrix} z^2 \\ z^2 p_1(z) \\ z \\ z p_3(z) \\ 1 \\ p_5(z) \end{bmatrix} \begin{bmatrix} z^3 p_0(z) & z^2 & z^2 p_2(z) & z & z p_4(z) & 1 \end{bmatrix} - p(z) \begin{bmatrix} 0 & 0 & 1 & 0 & z & 0 \\ 1 & 0 & p_1(z) & 0 & z p_1(z) & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ z & 1 & p_2(z) & 0 & p_3(z) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ z^2 & z & z p_2(z) & 1 & p_4(z) & 0 \end{bmatrix}.$$

THEOREM 3.2 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a polynomial and $p_d(z)$, for $d = 0, 1, \ldots, n-1$, the degree d Horner shift of $p(z)$. Let $\sigma : \{0, 1, \ldots, n-1\} \to \{1, \ldots, n\}$ be a bijection with $\operatorname{EPCIS}(\sigma) = (v_0, v_1, \ldots, v_{n-1})$ and let M_σ be the Fiedler matrix of $p(z)$ associated with σ. Let $x_\sigma, y_\sigma \in \mathbb{C}^n[z]$ be the polynomial vectors whose kth entry is

$$x_\sigma(k) = \begin{cases} z^{\mathrm{i}_\sigma(0:n-k-1)} p_{k-1}(z) & \text{if } v_{n-k} = 1, \\ z^{\mathrm{i}_\sigma(0:n-k-1)} & \text{if } v_{n-k} = 0, \end{cases} \quad \text{and} \quad y_\sigma(k) = \begin{cases} z^{\mathrm{c}_\sigma(0:n-k-1)} p_{k-1}(z) & \text{if } v_{n-k} = 0, \\ z^{\mathrm{c}_\sigma(0:n-k-1)} & \text{if } v_{n-k} = 1, \end{cases} \tag{3.4}$$

for $k = 1, 2, \ldots, n$, and let $A_\sigma \in \mathbb{C}^{n\times n}[z]$ be the matrix whose (i, j) entry is

$$A_\sigma(i,j) = \begin{cases} 0 & \text{if } v_{n-i} = v_{n-j} = 0 \text{ and } i \geqslant j, \\ z^{\mathrm{i}_\sigma(n-j+1:n-i-1)} & \text{if } v_{n-i} = v_{n-j} = 0 \text{ and } i < j, \\ z^{\mathrm{c}_\sigma(n-i+1:n-j-1)} & \text{if } v_{n-i} = v_{n-j} = 1 \text{ and } i > j, \\ 0 & \text{if } v_{n-i} = v_{n-j} = 1 \text{ and } i \leqslant j, \\ 0 & \text{if } v_{n-i} = 0 \text{ and } v_{n-j} = 1, \\ z^{\mathrm{c}_\sigma(n-i+1:n-j-1)} p_{j-1}(z) & \text{if } v_{n-i} = 1,\ v_{n-j} = 0 \text{ and } i > j, \\ z^{\mathrm{i}_\sigma(n-j+1:n-i-1)} p_{i-1}(z) & \text{if } v_{n-i} = 1,\ v_{n-j} = 0 \text{ and } i < j, \end{cases} \tag{3.5}$$

for $i, j = 1, 2, \ldots, n$. Then,

$$\operatorname{adj}(zI - M_\sigma) = x_\sigma y_\sigma^T - p(z) A_\sigma.$$

Note that x_σ, y_σ and A_σ depend on the variable z, though we drop it for the ease of notation.

Before proving Theorem 3.2 we state and prove some technical lemmas.

LEMMA 3.2 Let x_σ and y_σ be the vectors defined in (3.4), and A_σ be the matrix defined in (3.5). Then, A_σ is the unique $n\times n$ matrix satisfying the following two properties:

(i) The entries of A_σ are polynomials in z, and

(ii) all entries of $x_\sigma y_\sigma^T - p(z)A_\sigma$ are polynomials of degree less than or equal to $n-1$.

Proof. Throughout this proof we use the following notation:

$$q_k(z) := -a_k z^k - a_{k-1} z^{k-1} - \cdots - a_1 z - a_0 = z^{k+1} p_{n-k-1}(z) - p(z),$$

for $k = 0, 1, \ldots, n-1$. Note that $q_k(z)$ is a polynomial of degree k.

To prove that the entries of A_σ are polynomials, it suffices to see that the exponents of the powers of z appearing in the entries of (3.5) are nonnegative. This is immediate by Definition 2.3. To prove that the (i,j) entry of $x_\sigma y_\sigma^T - p(z)A_\sigma$ is a polynomial of degree less than or equal to $n-1$ we consider each case in (3.5) separately.

(1) $v_{n-i} = v_{n-j} = 0$ and $i \geqslant j$: The (i,j) entry of $x_\sigma y_\sigma^T - p(z)A_\sigma$ is equal to

$$x_\sigma(i) y_\sigma(j) - p(z)A_\sigma(i,j) = z^{\mathfrak{i}_\sigma(0:n-i-1) + \mathfrak{c}_\sigma(0:n-j-1)} p_{j-1}(z)$$

which is a polynomial of degree less than or equal to $n-1$, because, using (2.4),

$$\mathfrak{i}_\sigma(0:n-i-1) + \mathfrak{c}_\sigma(0:n-j-1) + j - 1 = \mathfrak{i}_\sigma(0:n-i-1) - \mathfrak{i}_\sigma(0:n-j-1) + n - 1 \leqslant n-1.$$

(2) $v_{n-i} = v_{n-j} = 0$ and $i < j$: Using (2.4), the (i,j) entry of $x_\sigma y_\sigma^T - p(z)A_\sigma$ is equal to

$$\begin{aligned} x_\sigma(i) y_\sigma(j) - p(z)A_\sigma(i,j) &= z^{\mathfrak{i}_\sigma(0:n-i-1) + \mathfrak{c}_\sigma(0:n-j-1)} p_{j-1}(z) - p(z) z^{\mathfrak{i}_\sigma(n-j+1:n-i-1)} \\ &= z^{\mathfrak{i}_\sigma(n-j+1:n-i-1)} (z^{n-j+1} p_{j-1}(z) - p(z)) \\ &= z^{\mathfrak{i}_\sigma(n-j+1:n-i-1)} q_{n-j}(z), \end{aligned}$$

which is a polynomial of degree less than $n-1$, because

$$\mathfrak{i}_\sigma(n-j+1:n-i-1) + n - j \leqslant n - i - 1 < n-1.$$

(3) $v_{n-i} = v_{n-j} = 1$ and $i > j$: Using (2.4), the (i,j) entry of $x_\sigma y_\sigma^T - p(z)A_\sigma$ is equal to

$$\begin{aligned} x_\sigma(i) y_\sigma(j) - p(z)A_\sigma(i,j) &= z^{\mathfrak{i}_\sigma(0:n-i-1) + \mathfrak{c}_\sigma(0:n-j-1)} p_{i-1}(z) - p(z) z^{\mathfrak{c}_\sigma(n-i+1:n-j-1)} \\ &= z^{\mathfrak{c}_\sigma(n-i+1:n-j-1)} (z^{n-i+1} p_{i-1}(z) - p(z)) \\ &= z^{\mathfrak{c}_\sigma(n-i+1:n-j-1)} q_{n-i}(z), \end{aligned}$$

which is a polynomial of degree less than $n-1$, because

$$\mathfrak{c}_\sigma(n-i+1:n-j-1) + n - i \leqslant n - j - 1 < n-1.$$

(4) $v_{n-i} = v_{n-j} = 1$ and $i \leqslant j$: The (i,j) entry of $x_\sigma y_\sigma^T - p(z)A_\sigma$ is equal to

$$x_\sigma(i) y_\sigma(j) - p(z)A_\sigma(i,j) = z^{\mathfrak{i}_\sigma(0:n-i-1) + \mathfrak{c}_\sigma(0:n-j-1)} p_{i-1}(z),$$

which is a polynomial of degree less than or equal to $n-1$, because, using (2.4),

$$\mathfrak{i}_\sigma(0:n-i-1) + \mathfrak{c}_\sigma(0:n-j-1) + i - 1 = \mathfrak{c}_\sigma(0:n-j-1) - \mathfrak{c}_\sigma(0:n-i-1) + n - 1 \leqslant n-1.$$

(5) $v_{n-i} = 0$ and $v_{n-j} = 1$: The (i,j) entry of $x_\sigma y_\sigma^T - p(z)A_\sigma$ is equal to

$$x_\sigma(i) y_\sigma(j) - p(z)A_\sigma(i,j) = z^{\mathfrak{i}_\sigma(0:n-i-1) + \mathfrak{c}_\sigma(0:n-j-1)}$$

which is a polynomial of degree less than or equal to $n-1$, by (2.5).

(6) $v_{n-i} = 1$, $v_{n-j} = 0$ and $i > j$: Using (2.4), the (i,j) entry of $x_\sigma y_\sigma^T - p(z)A_\sigma$ is equal to

$$\begin{aligned} x_\sigma(i) y_\sigma(j) - p(z)A_\sigma(i,j) &= z^{\mathfrak{i}_\sigma(0:n-i-1) + \mathfrak{c}_\sigma(0:n-j-1)} p_{i-1}(z) p_{j-1}(z) - p(z) z^{\mathfrak{c}_\sigma(n-i+1:n-j-1)} p_{j-1}(z) \\ &= z^{\mathfrak{c}_\sigma(n-i+1:n-j-1)} p_{j-1}(z) (z^{n-i+1} p_{i-1}(z) - p(z)) \\ &= z^{\mathfrak{c}_\sigma(n-i+1:n-j-1)} p_{j-1}(z) q_{n-i}(z), \end{aligned}$$

which is a polynomial of degree less than $n-1$, because

$$\mathfrak{c}_\sigma(n-i+1:n-j-1) + j - 1 + n - i \leqslant i - j - 1 + j - 1 + n - i = n-2.$$

(7) $v_{n-i}=1$, $v_{n-j}=0$ and $i<j$: Using (2.4), the (i,j) entry of $x_\sigma y_\sigma^T - p(z)A_\sigma$ is equal to

$$\begin{aligned} x_\sigma(i)y_\sigma(j)-p(z)A_\sigma(i,j) &= z^{\mathfrak{i}_\sigma(0:n-i-1)+\mathfrak{c}_\sigma(0:n-j-1)}p_{i-1}(z)p_{j-1}(z)-p(z)z^{\mathfrak{i}_\sigma(n-j+1:n-i-1)}p_{i-1}(z) \\ &= z^{\mathfrak{i}_\sigma(n-j+1:n-i-1)}p_{i-1}(z)(z^{n-j+1}p_{j-1}(z)-p(z)) \\ &= z^{\mathfrak{i}_\sigma(n-j+1:n-i-1)}p_{i-1}(z)q_{n-j}(z), \end{aligned}$$

which is a polynomial of degree less than $n-1$, because

$$\mathfrak{i}_\sigma(n-j+1:n-i-1)+i-1+n-j \leqslant j-i-1+i-1+n-j = n-2.$$

Now, suppose that there is another matrix B, whose entries are polynomials in z, and such that the entries of the matrix $x_\sigma y_\sigma^T - p(z)B$ are polynomials in z of degree less than or equal to $n-1$. Let $W_1 = x_\sigma y_\sigma^T - p(z)A_\sigma$ and let $W_2 = x_\sigma y_\sigma^T - p(z)B$, then, $W_1 - W_2 = p(z)(B-A_\sigma)$ is a matrix whose entries are polynomials of degree less than or equal to $n-1$, but if $A_\sigma \neq B$, then $p(z)(B-A_\sigma)$ has, at least, one entry which is a polynomial of degree greater than or equal to n, hence $A_\sigma = B$. □

Lemma 3.3 is key to prove Theorem 3.2. It allows us to relate $\operatorname{adj}(zI-M_\sigma)$ with the adjugate of an $(n-1)\times(n-1)$ matrix obtained by deflating $zI-M_\sigma$ in a certain way. In the following, a matrix polynomial $P(z)\in\mathbb{C}^{n\times n}[z]$ is said to be *unimodular* if $\det P(z)$ is a nonzero constant. In other words, $P(z)$ has a polynomial inverse.

LEMMA 3.3 Let $p(z)=z^n+\sum_{k=0}^{n-1}a_kz^k$, let $\sigma:\{0,1,\ldots,n-1\}\to\{1,\ldots,n\}$ be a bijection with $\mathrm{PCIS}(\sigma)=(v_0,v_1,\ldots,v_{n-2})$, let M_σ be the Fiedler matrix of $p(z)$ associated with σ, and define the unimodular matrix polynomials $Q(z),R(z)\in\mathbb{C}^{n\times n}[z]$ as

$$Q(z) := \begin{bmatrix} 1 & 0 & \\ z & 1 & \\ & & I_{n-2} \end{bmatrix} \quad\text{and}\quad R(z) := \begin{bmatrix} 0 & 1 & \\ -1 & p_1(z) & \\ & & I_{n-2} \end{bmatrix}.$$

Then,

(a) if σ has a consecution at $n-2$,

$$Q(z)(zI_n-M_\sigma)R(z) = \begin{bmatrix} 1 & \\ & zI_{n-1}-\widetilde{M}_\rho \end{bmatrix},$$

(b) if σ has an inversion at $n-2$,

$$R(z)^T(zI_n-M_\sigma)Q(z)^T = \begin{bmatrix} 1 & \\ & zI_{n-1}-\widetilde{M}_\rho \end{bmatrix},$$

where $\rho:\{0,1,\ldots,n-2\}\to\{1,\ldots,n-1\}$ is a bijection such that $\mathrm{PCIS}(\rho)=(v_0,v_1,\ldots,v_{n-3})$, and $\widetilde{M}_\rho = \widetilde{M}_{\rho^{-1}(1)}\widetilde{M}_{\rho^{-1}(2)}\cdots\widetilde{M}_{\rho^{-1}(n-1)}$, with

$$\widetilde{M}_k = \begin{bmatrix} I_{n-k-2} & & & \\ & -a_k & 1 & \\ & 1 & 0 & \\ & & & I_{k-1} \end{bmatrix}, \qquad \text{for } k=1,2,\ldots,n-3,$$

and

$$\widetilde{M}_0 = \begin{bmatrix} I_{n-2} & \\ & -a_0 \end{bmatrix}, \qquad \widetilde{M}_{n-2} = \begin{bmatrix} -p_2(z)+z & 1 & \\ 1 & 0 & \\ & & I_{n-3} \end{bmatrix}.$$

Proof. We only prove part (a) because part (b) is similar. So, let us assume that σ has a consecution at $n-2$. Then, using the commutativity relations (2.3), the factors of M_σ can be rearranged until M_{n-1} is adjacent on the right to M_{n-2}, that is, $M_\sigma = XM_{n-2}M_{n-1}Y$, where X,Y are products of M_i matrices, with $i<n-2$. Now, since $Q(z)$ and $R(z)$ commute with M_i, for $i<n-2$, we have

$$
\begin{aligned}
Q(z)(zI_n - M_\sigma)R(z) &= zQ(z)R(z) - XQ(z)M_{n-2}M_{n-1}R(z)Y \\
&= \begin{bmatrix} 0 & z & 0 & \\ -z & z^2 + zp_1(z) & 0 & \\ 0 & 0 & z & \\ & & & zI_{n-3} \end{bmatrix} - X \begin{bmatrix} -1 & z & 0 & \\ -z & z^2 - a_{n-2} & 1 & \\ 0 & 1 & 0 & \\ & & & I_{n-3} \end{bmatrix} Y \\
&= \begin{bmatrix} 0 & z & 0 & \\ -z & z^2 & 0 & \\ 0 & 0 & z & \\ & & & zI_{n-3} \end{bmatrix} - X \left(\begin{bmatrix} -1 & z & 0 & \\ -z & z^2 - z & 0 & \\ 0 & 0 & 0 & \\ & & & 0_{n-3} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & \\ 0 & -p_2(z) + z & 1 & \\ 0 & 1 & 0 & \\ & & & I_{n-3} \end{bmatrix} \right) Y \\
&= \begin{bmatrix} 1 & & & \\ & z & & \\ & & z & \\ & & & zI_{n-3} \end{bmatrix} - X \begin{bmatrix} 0 & 0 & 0 & \\ 0 & -p_2(z) + z & 1 & \\ 0 & 1 & 0 & \\ & & & I_{n-3} \end{bmatrix} Y \\
&= \begin{bmatrix} 1 & \\ & zI_{n-1} \end{bmatrix} - \begin{bmatrix} 0 & \\ & \widetilde{M}_{\rho^{-1}(1)} \widetilde{M}_{\rho^{-1}(2)} \cdots \widetilde{M}_{\rho^{-1}(n-1)} \end{bmatrix} = \begin{bmatrix} 1 & \\ & zI_{n-1} - \widetilde{M}_\rho \end{bmatrix},
\end{aligned}
$$

where we have used that $p_2(z) = zp_1(z) + a_{n-2}$ and the fact that multiplying any matrix of the form $\operatorname{diag}(A, 0_{n-2})$, with $A \in \mathbb{C}^{2\times 2}$, by M_k, for $k = 0, 1, \ldots, n-3$, keeps that matrix unchanged. Finally, notice that the relative positions of the matrices $\widetilde{M}_0, \widetilde{M}_1, \ldots, \widetilde{M}_{n-2}$ in $\widetilde{M}_\rho$ are the same as the relative positions of the matrices $M_0, M_1, \ldots, M_{n-2}$ in M_σ, therefore $\operatorname{PCIS}(\rho) = (v_0, v_1, \ldots, v_{n-3})$. □

REMARK 3.1 Some important observations about the matrix $\widetilde{M}_\rho$ in Lemma 3.3 are in order:

(a) The matrix $\widetilde{M}_i$, for $i = 0, \ldots, n-3$ is obtained from M_i by removing the first row and column.

(b) The matrix $\widetilde{M}_\rho$ can be seen formally as a Fiedler matrix of the polynomial $r(z) := z^{n-1} + \sum_{k=0}^{n-2} b_k z^k$, where $b_{n-2} = p_2(z) - z$ and $b_k = a_k$ for $k = 0, 1, \ldots, n-3$. Notice that $r(z) = p(z)$ for all $z \in \mathbb{C}$. We also want to emphasize that the formal $(n-2)$th coefficient of $r(z)$ is not an scalar, but a polynomial in z.

(c) The formal Horner shifts of $r(z)$ satisfy: $r_0(z) = p_0(z) = 1$ and $r_k(z) = p_{k+1}(z)$ for $k = 1, 2, \ldots, n-2$.

Now, armed with Lemmas 3.2 and 3.3, we are in the position to prove Theorem 3.2.

Proof. (of **Theorem 3.2**) The proof proceeds by induction in n. For $n = 2$ there are only two Fiedler matrices, namely the first and second Frobenius companion matrices. For these two matrices we have

$$
\operatorname{adj}(zI - C_2) = \operatorname{adj}\left(\begin{bmatrix} a_1 + z & -1 \\ a_0 & z \end{bmatrix} \right) = \begin{bmatrix} z & 1 \\ -a_0 & a_1 + z \end{bmatrix} = \begin{bmatrix} 1 \\ p_1(z) \end{bmatrix} \begin{bmatrix} z & 1 \end{bmatrix} - p(z) \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}
$$

and

$$
\operatorname{adj}(zI - C_1) = \operatorname{adj}\left(\begin{bmatrix} a_1 + z & a_0 \\ -1 & z \end{bmatrix} \right) = \begin{bmatrix} z & -a_0 \\ 1 & a_1 + z \end{bmatrix} = \begin{bmatrix} z \\ 1 \end{bmatrix} \begin{bmatrix} 1 & p_1(z) \end{bmatrix} - p(z) \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},
$$

which are the matrices in the statement of Theorem 3.2 with $\operatorname{PCIS}(\sigma) = (1)$ and $\operatorname{PCIS}(\sigma) = (0)$, respectively. Assume that the result is true for Fiedler matrices of size $(n-1) \times (n-1)$. To prove it for size $n \times n$, we have to distinguish two cases, namely, whether σ has a consecution or an inversion at $n-2$. Suppose that σ has a consecution at $n-2$ (the proof when σ has an inversion at $n-2$ is similar and we omit it). Then, by Lemma 3.3, we have that

$$
zI_n - M_\sigma = Q(z)^{-1} \begin{bmatrix} 1 & \\ & zI_{n-1} - \widetilde{M}_\rho \end{bmatrix} R(z)^{-1},
$$

therefore

$$
\begin{aligned}
\operatorname{adj}(zI_n - M_\sigma) &= \operatorname{adj}\left(R(z)^{-1}\right) \operatorname{adj}\left(\begin{bmatrix} 1 & \\ & zI_{n-1} - \widetilde{M}_\rho \end{bmatrix} \right) \operatorname{adj}\left(Q(z)^{-1}\right) = \\
&= R(z) \begin{bmatrix} p(z) & \\ & \operatorname{adj}(zI_{n-1} - \widetilde{M}_\rho) \end{bmatrix} Q(z),
\end{aligned}
$$

where we have used the identities $\operatorname{adj}(AB) = \operatorname{adj}(B)\operatorname{adj}(A)$, $\det R(z) = \det Q(z) = 1$, and $\det(zI_{n-1} - \widetilde{M}_\rho) = p(z)$. By the induction hypothesis

$$\begin{aligned}\operatorname{adj}(zI_n - M_\sigma) &= R(z)\begin{bmatrix} p(z) & \\ & x_\rho y_\rho^T - p(z)A_\rho \end{bmatrix} Q(z) \\ &= R(z)\begin{bmatrix} 0 \\ x_\rho \end{bmatrix}\begin{bmatrix} 0 & y_\rho^T \end{bmatrix} Q(z) - p(z)R(z)\begin{bmatrix} -1 & \\ & A_\rho \end{bmatrix} Q(z).\end{aligned}$$

Note that in the induction step we may see $\widetilde{M}_\rho$ as a Fiedler matrix associated with $r(z) = z^{n-1} + \sum_{k=0}^{n-2} b_k z^k$, with b_i, for $i = 0, \ldots, n-2$, as in Remark 3.1, part (b). To finish the proof it suffices to prove the following three identities:

$$\text{(i)} \quad x_\sigma = R(z)\begin{bmatrix} 0 \\ x_\rho \end{bmatrix}, \quad \text{(ii)} \quad y_\sigma = Q^T(z)\begin{bmatrix} 0 \\ y_\rho \end{bmatrix}, \quad \text{and} \quad \text{(iii)} \quad A_\sigma = R(z)\begin{bmatrix} -1 & \\ & A_\rho \end{bmatrix} Q(z).$$

(i) From the expressions of PCIS(σ) and PCIS(ρ) we have $\mathfrak{i}_\rho(0:k-1) = \mathfrak{i}_\sigma(0:k-1)$, for $k = 1, 2, \ldots, n-2$. Also we have that the Horner shifts corresponding to $\widetilde{M}_\rho$ are $p_0(z), p_2(z), \ldots, p_{n-1}(z)$. These observations imply that $x_\rho(k) = x_\sigma(k+1)$, for $k = 2, 3, \ldots, n-1$ (note that, for the permutation ρ, n must be replaced by $n-1$ in (3.4)). Therefore

$$R(z)\begin{bmatrix} 0 \\ x_\rho \end{bmatrix} = \begin{bmatrix} 0 & 1 & \\ -1 & p_1(z) & \\ & & I_{n-2} \end{bmatrix}\begin{bmatrix} 0 \\ z^{\mathfrak{i}_\rho(0:n-3)} \\ x_\rho(2:n-1) \end{bmatrix} = \begin{bmatrix} z^{\mathfrak{i}_\rho(0:n-3)} \\ z^{\mathfrak{i}_\rho(0:n-3)} p_1(z) \\ x_\rho(2:n-1) \end{bmatrix} = \begin{bmatrix} z^{\mathfrak{i}_\sigma(0:n-2)} p_0(z) \\ z^{\mathfrak{i}_\sigma(0:n-3)} p_1(z) \\ x_\sigma(3:n) \end{bmatrix} = x_\sigma,$$

where we have used, since $v_{n-2} = 1$, that $\mathfrak{i}_\sigma(0:n-3) = \mathfrak{i}_\sigma(0:n-2)$ and $p_0(z) = 1$.

(ii) From the expressions of PCIS(σ) and PCIS(ρ) we have $\mathfrak{c}_\rho(0:k-1) = \mathfrak{c}_\sigma(0:k-1)$, for $k = 1, 2, \ldots, n-2$. We also have that the Horner shifts corresponding to $\widetilde{M}_\rho$ are $p_0(z), p_2(z), \ldots, p_{n-1}(z)$. These observations imply that $y_\rho(k) = y_\sigma(k+1)$, for $k = 2, 3, \ldots, n-1$. Therefore

$$Q(z)^T\begin{bmatrix} 0 \\ y_\rho \end{bmatrix} = \begin{bmatrix} 1 & z & \\ 0 & 1 & \\ & & I_{n-2} \end{bmatrix}\begin{bmatrix} 0 \\ z^{\mathfrak{c}_\rho(0:n-3)} \\ y_\rho(2:n-1) \end{bmatrix} = \begin{bmatrix} z^{\mathfrak{c}_\rho(0:n-3)+1} \\ z^{\mathfrak{c}_\rho(0:n-3)} \\ y_\rho(2:n-1) \end{bmatrix} = \begin{bmatrix} z^{\mathfrak{c}_\sigma(0:n-2)} \\ z^{\mathfrak{c}_\sigma(0:n-3)} \\ y_\sigma(3:n) \end{bmatrix} = y_\sigma,$$

where we have used, since $v_{n-2} = 1$, that $\mathfrak{c}_\sigma(0:n-2) = \mathfrak{c}_\sigma(0:n-3) + 1$.

(iii) We prove this using Lemma 3.2. From (i) and (ii) we know that

$$\operatorname{adj}(zI - M_\sigma) = x_\sigma y_\sigma^T - p(z)R(z)\begin{bmatrix} -1 & \\ & A_\rho \end{bmatrix} Q(z).$$

But the entries of $R(z)\operatorname{diag}(-1, A_\rho)Q(z)$ are polynomials in z and, moreover, the entries of $\operatorname{adj}(zI - M_\sigma)$ are polynomials of degree less than or equal to $n-1$. Therefore, by the uniqueness proved in Lemma 3.2, we get:

$$R(z)\begin{bmatrix} -1 & \\ & A_\rho \end{bmatrix} Q(z) = A_\sigma.$$

□

3.2 *First-order perturbation of the coefficients of the polynomial* $\det(zI - M_\sigma)$

In this section we show how the coefficients of the characteristic polynomial of any Fiedler companion matrix M_σ change when we perturb M_σ with a dense matrix E. More precisely, we give, to first order in E, the coefficients of the characteristic polynomial of $M_\sigma + E$.

THEOREM 3.3 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial, let $\sigma : \{0, 1, \ldots, n-1\} \to \{1, \ldots, n\}$ be a bijection with EPCIS(σ) $= (v_0, v_1, \ldots, v_{n-1})$, let M_σ be the Fiedler companion matrix of $p(z)$ associated with σ, and let $E \in$

$\mathbb{C}^{n\times n}$ be an arbitrary matrix. If the characteristic polynomial of $M_\sigma + E$ is denoted by $\widetilde{p}(z) = z^n + \sum_{k=0}^{n-1} \widetilde{a}_k z^k$, then, to first order in E,

$$\widetilde{a}_k - a_k = -\sum_{i,j=1}^{n} p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1}) E_{ij}, \qquad k = 0, 1, \ldots, n-1, \tag{3.6}$$

where, for $i, j = 1, 2, \ldots, n$, the function $p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})$ is a multivariable polynomial in the coefficients of $p(z)$. More precisely, $p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})$ is equal to:

(a) if $v_{n-i} = v_{n-j} = 0$:

- $a_{k+\mathfrak{i}_\sigma(n-j:n-i)}$,

  if $\quad j \geqslant i$ and $\quad n-k-i+1 \leqslant \mathfrak{i}_\sigma(n-j:n-i) \leqslant n-k$;
- $-a_{k+1-\mathfrak{i}_\sigma(n-i:n-j-1)}$,

  if $\quad j < i$ and $\quad k+1+i-n \leqslant \mathfrak{i}_\sigma(n-i:n-j-1) \leqslant k+1$;
- 0, otherwise;

(b) if $v_{n-i} = v_{n-j} = 1$:

- $a_{k+\mathfrak{c}_\sigma(n-i:n-j)}$,

  if $\quad j \leqslant i$ and $\quad n-k-j+1 \leqslant \mathfrak{c}_\sigma(n-i:n-j) \leqslant n-k$;
- $-a_{k+1-\mathfrak{c}_\sigma(n-j:n-i-1)}$,

  if $\quad j > i$ and $\quad k+1+j-n \leqslant \mathfrak{c}_\sigma(n-j:n-i-1) \leqslant k+1$;
- 0, otherwise;

(c) if $v_{n-i} = 1$ and $v_{n-j} = 0$:

- 1, if $\quad \mathfrak{i}_\sigma(0:n-j-1) + \mathfrak{c}_\sigma(0:n-i-1) = k$,
- 0, otherwise;

(d) if $v_{n-i} = 0$ and $v_{n-j} = 1$:

- $\displaystyle\sum_{l=\max\{0,k+1+j-\mathfrak{c}_\sigma(n-j:n-i-1)-n\}}^{l=\min\{k+1-\mathfrak{c}_\sigma(n-j:n-i-1),i-1\}} -(a_{n+1-i+l}\, a_{k+1-\mathfrak{c}_\sigma(n-j:n-i-1)-l})$,

  if $\quad j > i$ and $\quad k+2+j-i-n \leqslant \mathfrak{c}_\sigma(n-j:n-i-1) \leqslant k+1$;
- $\displaystyle\sum_{l=\max\{0,k+1+i-\mathfrak{i}_\sigma(n-i:n-j-1)-n\}}^{l=\min\{k+1-\mathfrak{i}_\sigma(n-i:n-j-1),j-1\}} -(a_{n+1-j+l}\, a_{k+1-\mathfrak{i}_\sigma(n-i:n-j-1)-l})$,

  if $\quad j < i$ and $\quad k+2+i-j-n \leqslant \mathfrak{i}_\sigma(n-i:n-j-1) \leqslant k+1$;
- 0, otherwise;

where we set $a_n := 1$.

Proof. From Proposition 3.1, the coefficients of the characteristic polynomial of $M_\sigma + E$ satisfy, to first order in E,

$$\widetilde{a}_k - a_k = -\sum_{i,j=1}^{n} P_{k+1}(j,i) E_{ij},$$

where $P_{k+1}(j,i)$ is the (j,i) entry of P_{k+1} which, according to (3.2) is the kth matrix coefficient of the matrix polynomial $\mathrm{adj}(zI - M_\sigma)$. Therefore $p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})$ is the kth coefficient of the (j,i) entry of $\mathrm{adj}(zI - M_\sigma)$. From Theorem 3.2 and the proof of Lemma 3.2, we know that the (j,i) entry of $\mathrm{adj}(zI - M_\sigma)$, in each of the cases considered in the statement, is:

(a) $z^{\mathfrak{i}_\sigma(0:n-j-1)+\mathfrak{c}_\sigma(0:n-i-1)}p_{i-1}(z)$, if $j \geqslant i$, or $z^{\mathfrak{i}_\sigma(n-i+1:n-j-1)}q_{n-i}(z)$, if $j < i$ (see (1) and (2), respectively, in the proof of Lemma 3.2);

(b) $z^{\mathfrak{i}_\sigma(0:n-j-1)+\mathfrak{c}_\sigma(0:n-i-1)}p_{j-1}(z)$, if $j \leqslant i$, or $z^{\mathfrak{c}_\sigma(n-j+1:n-i-1)}q_{n-j}(z)$, if $j > i$ (see (3) and (4) in the proof of Lemma 3.2);

(c) $z^{\mathfrak{i}_\sigma(0:n-j-1)+\mathfrak{c}_\sigma(0:n-i-1)}$ (see (5) in the proof of Lemma 3.2);

(d) $z^{\mathfrak{c}_\sigma(n-j+1:n-i-1)}p_{i-1}(z)q_{n-j}(z)$, if $j > i$, or $z^{\mathfrak{i}_\sigma(n-i+1:n-j-1)}p_{j-1}(z)q_{n-i}(z)$, if $j < i$ (see (6) and (7) in the proof of Lemma 3.2).

Now, it is just a straightforward computation to check that the formulas given in the statement coincide with the kth coefficient of the previous polynomials. □

REMARK 3.2 According to the notation in (3.1), we have

$$\nabla a_k(M_\sigma) = -\left[p_{11}^{(\sigma,k)} \dots p_{n1}^{(\sigma,k)}\, p_{12}^{(\sigma,k)} \dots p_{n2}^{(\sigma,k)} \dots p_{1n}^{(\sigma,k)} \dots p_{nn}^{(\sigma,k)}\right]^T,$$

where we have dropped the dependence on $a_0, \dots, a_{n-1}$ for brevity.

REMARK 3.3 For $k = n-1$, and σ an arbitrary bijection, a direct verification in Theorem 3.3 gives

$$p_{ij}^{(\sigma,n-1)}(a_0, \dots, a_{n-1}) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Then, for any Fiedler matrix M_σ, it follows from (3.6) that

$$a_{n-1}(M_\sigma + E) - a_{n-1}(M_\sigma) = -\sum_{i=1}^{n} E_{ii}.$$

But, since the $(n-1)$th coefficient of the characteristic polynomial of A is equal to $-\mathrm{tr}(A)$, this is a restatement of the well-know identity:

$$\mathrm{tr}(M_\sigma + E) = \mathrm{tr}(M_\sigma) + \mathrm{tr}(E).$$

We want to emphasize that $p_{ij}^{(\sigma,k)}(a_0, a_1, \dots, a_{n-1})$ are always linear or quadratic polynomials in the coefficients $a_0, \dots, a_{n-1}$. They depend, at a first stage, on whether the bijection σ has a consecution or an inversion at $n-i$ and $n-j$. In particular, $p_{ij}^{(\sigma,k)}(a_0, a_1, \dots, a_{n-1})$ can only be quadratic when there is a consecution at $n-j$ and an inversion at $n-i$. This implies the following corollary.

COROLLARY 3.1 Let M_σ be C_1 or C_2 in the statement of Theorem 3.3, then $p_{ij}^{(\sigma,k)}(a_0, a_1, \dots, a_{n-1})$ in (3.6) is a polynomial of degree at most 1 in $a_0, \dots, a_{n-1}$, for all $k = 0, 1, \dots, n-1$, and all $1 \leqslant i, j \leqslant n$. For the remaining Fiedler matrices M_σ, there is always some k and some i, j such that $p_{ij}^{(\sigma,k)}(a_0, a_1, \dots, a_{n-1})$ is a quadratic polynomial in $a_0, a_1, \dots, a_{n-1}$.

Proof. Let us first recall that the bijection associated with C_1 is $\sigma_1 = (\sigma_1(0), \sigma_1(1), \dots, \sigma_1(n-1)) = (n, n-1, \dots, 1)$, whereas the bijection associated with C_2 is $\sigma_2 = (\sigma_2(0), \sigma_2(1), \dots, \sigma_2(n-1)) = (1, 2, \dots, n)$. Hence, σ_1 has no consecutions, whereas σ_2 has no inversions.

Then, it remains to show that, if $\sigma : \{0, 1, \dots, n-1\} \to \{1, \dots, n\}$ is a bijection having a consecution at $n-j$ and an inversion at $n-i$, for some $2 \leqslant i, j \leqslant n$, then there is some $0 \leqslant k \leqslant n-1$ such that $p_{ij}^{(\sigma,k)}(a_0, a_1, \dots, a_{n-1})$ has degree 2. Note, first, that it must be $i \neq j$. Without loss of generality, let us assume that $j > i$. The proof for the case $j < i$ is analogous. We need to prove that, in the sum defining $p_{ij}^{(\sigma,k)}(a_0, a_1, \dots, a_{n-1})$ in the first bullet of case (d) in Theorem 3.3 there is at least one monomial $a_r a_s$ such that $0 \leqslant r, s \leqslant n-1$. More precisely, we need to prove:

(i) There is some $0 \leqslant k \leqslant n-1$ such that $k + 2 + j - i - n \leqslant \mathfrak{c}_\sigma(n-j : n-i-1) \leqslant k+1$.

(ii) There is some l, with $\max\{0, k+1+j-\mathfrak{c}_\sigma(n-j : n-i-1) - n\} \leqslant l \leqslant \min\{k+1-\mathfrak{c}_\sigma(n-j : n-i-1), i-1\}$, such that $0 \leqslant n+1-i+l \leqslant n-1$ and $0 \leqslant k+1-\mathfrak{c}_\sigma(n-j : n-i-1) - l \leqslant n-1$.

For this, it suffices to take $k = \mathfrak{c}_\sigma(n-j:n-i-1)-1 = \mathfrak{c}_\sigma(n-j+1:n-i-1)$ and $l=0$. Note that (ii) is fulfilled for these values of k and l, because $i \geqslant 2$. □

The expressions given in Theorem 3.3 for the variation of the coefficients of the characteristic polynomial of M_σ are involved in general (that is, for arbitrary Fiedler matrices). We will show them explicitly in Section 3.2.2 for some particularly relevant Fiedler matrices, including the Frobenius companion matrices.

The following result describes one property of the polynomials $p_{ij}^{(\sigma,k)}(a_0,\dots,a_{n-1})$ that will be used later.

LEMMA 3.4 Let $p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})$ be the polynomial defined in (3.6). Then:

(a) For $k=0,1,\dots,n-1$,

$$p_{ii}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1}) = \begin{cases} a_{k+1} & \text{if} \quad i \geqslant n-k, \\ 0 & \text{if} \quad i < n-k, \end{cases}$$

with $a_n = 1$.

(b) If σ has a consecution at $n-2$, then $p_{12}^{(\sigma,0)}(a_0,a_1,\dots,a_{n-1}) = -a_0$, and if σ has an inversion at $n-2$, then $p_{21}^{(\sigma,0)}(a_0,a_1,\dots,a_{n-1}) = -a_0$.

Proof. From Theorem 3.3 we have $p_{ii}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1}) = a_{k+1}$, if $n-1 \geqslant k \geqslant n-i$ (namely, if $i \geqslant n-k$), and $p_{ii}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1}) = 0$ otherwise. This proves part (a).

For part (b), if σ has a consecution at $n-2$, then following the notation of Theorem 3.3, we have $v_{n-2} = v_{n-1} = 1$, and $\mathfrak{c}_\sigma(n-2:n-2) = 1$, so part (b) of Theorem 3.3 gives $p_{12}^{(\sigma,0)}(a_0,a_1,\dots,a_{n-1}) = -a_0$. Similarly, if σ has an inversion at $n-2$, then $v_{n-2} = v_{n-1} = 0$, and part (a) of Theorem 3.3 gives $p_{21}^{(\sigma,0)}(a_0,a_1,\dots,a_{n-1}) = -a_0$. □

To identify those indices k for which $\nabla a_k(M_\sigma)$ contains quadratic terms in $a_0,\dots,a_{n-1}$ may be interesting in practice. Notice that the presence of such quadratic terms implies that the sensitivity of the coefficient $a_k(M_\sigma)$ to perturbations of M_σ is quadratic in $a_0,\dots,a_{n-1}$, instead of linear. This implies in turn that, for large values of $a_0,\dots,a_{n-1}$, we can expect much larger changes after small perturbations in these coefficients than in the ones where $\nabla a_k(M_\sigma)$ contains only linear terms. We have seen in Corollary 3.1 that, for all Fiedler matrices but the Frobenius ones, there is always at least one k such that $\nabla a_k(M_\sigma)$ contains quadratic terms. Moreover, the proof of Corollary 3.1 tells us that if i,j are such that σ has a consecution at $n-j$ and an inversion at $n-i$, and $j>i$ (respectively, $j<i$), then for $k = \mathfrak{c}_\sigma(n-j+1:n-i-1)$ (resp., $k = \mathfrak{i}_\sigma(n-i+1:n-j-1)$) the gradient $\nabla a_k(M_\sigma)$ contains quadratic terms. In particular, Lemma 3.5 states that, for all Fiedler matrices but the Frobenius ones, $\nabla a_0(M_\sigma)$ contains always quadratic polynomials in $a_0,\dots,a_{n-1}$

LEMMA 3.5 Let $p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})$ be the polynomial defined in (3.6), and let $t \in \{0,1,\dots,n-3\}$.

(a) If $\mathrm{PCIS}(\sigma) = (v_0,v_1,\dots,v_t = 1, v_{t+1} = 0, v_{t+2} = 0,\dots,v_{n-2} = 0)$ then

$$p_{2,n-t}^{(\sigma,0)}(a_0,a_1,\dots,a_{n-1}) = -a_{n-1}a_0.$$

(b) If $\mathrm{PCIS}(\sigma) = (v_0,v_1,\dots,v_t = 0, v_{t+1} = 1, v_{t+2} = 1,\dots,v_{n-2} = 1)$ then

$$p_{n-t,2}^{(\sigma,0)}(a_0,a_1,\dots,a_{n-1}) = -a_{n-1}a_0.$$

Proof. We prove only part (a) because part (b) is similar. Since $n-t>2$ and $k - \mathfrak{c}_\sigma(n-j+1:n-i-1) = -\mathfrak{c}_\sigma(t+1:n-3) = 0$, from part (d) of Theorem 3.3,we have

$$p_{2,n-t}^{(\sigma,0)}(a_0,a_1,\dots,a_{n-1}) = \sum_{l=\max\{0,-\mathfrak{c}_\sigma(t+1:n-3)-t\}}^{l=\min\{0,1\}} -a_{n-1+l}\, a_{k-\mathfrak{c}_\sigma(t+1:n-3)-l} = -a_{n-1}a_0.$$

□

The main result, from the theoretical point of view, in this section is a direct consequence of Theorem 3.3.

COROLLARY 3.2 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial, and M_σ be a Fiedler companion matrix of $p(z)$. Assume that the roots of $p(z)$ are computed as the eigenvalues of M_σ with a backward stable algorithm i. e., an algorithm that computes the exact eigenvalues of some matrix $M_\sigma + E$, with $\|E\|_\infty = O(u)\|M_\sigma\|_\infty$. Then the computed roots are the exact roots of a polynomial $\widetilde{p}(z)$ such that:

(a) If $M_\sigma = C_1, C_2$,

$$\frac{\|\widetilde{p}-p\|_\infty}{\|p\|_\infty} = O(u)\|p\|_\infty, \tag{3.7}$$

(b) if $M_\sigma \neq C_1, C_2$,

$$\frac{\|\widetilde{p}-p\|_\infty}{\|p\|_\infty} = O(u)\|p\|_\infty^2, \tag{3.8}$$

where u is the machine precision. In other words, the backward error of the computed roots $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n$ is

$$\eta_\infty(\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n) = \begin{cases} O(u)\|p\|_\infty, & \text{if } M_\sigma = C_1, C_2, \\ O(u)\|p\|_\infty^2, & \text{if } M_\sigma \neq C_1, C_2. \end{cases}$$

Proof. If the eigenvalues of M_σ are computed with a backward stable algorithm, the computed eigenvalues are the exact eigenvalues of a matrix $M_\sigma + E$, for some matrix $E \in \mathbb{C}^{n\times n}$ such that $\|E\|_\infty = O(u)\|M_\sigma\|_\infty$. Thus, the computed eigenvalues are the exact roots of the polynomial $\widetilde{p}(z) = z^n + \sum_{k=0}^{n-1} \widetilde{a}_k z^k = \det(zI - M_\sigma - E)$. From Theorem 3.3, to first order in E,

$$|\widetilde{a}_k - a_k| = \left| \sum_{i,j=1}^{n} p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1}) E_{ij} \right| \leqslant \sum_{i,j=1}^{n} \left| p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1}) \right| \cdot |E_{ij}| \leqslant$$
$$\leqslant \left(\max_{1\leqslant i,j\leqslant n} |E_{ij}| \right) \cdot \left(\sum_{i,j=1}^{n} |p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})| \right).$$

Notice, also from Theorem 3.3, that the absolute value of every polynomial $p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})$ is bounded by $n\|p\|_\infty^2$ and that, by Corollary 3.1, the square in the norm of p is necessary in all Fiedler matrices except the Frobenius companion matrices, where it can be replaced by 1. Therefore,

$$\max_{k=0,1,\ldots,n-1} |\widetilde{a}_k - a_k| = \|\widetilde{p} - p\|_\infty = O(u)\|M_\sigma\|_\infty \|p\|_\infty^2 = O(u)\|p\|_\infty^3,$$

where we have used that $\max_{i,j=1,2,\ldots,n} |E_{ij}| = O(u)\|M_\sigma\|_\infty$ and $\|M_\sigma\|_\infty = O(1)\|p\|_\infty$ (see (De Terán *et al.*, 2014, Th. 3.3)). □

3.2.1 *Recursive formula for the derivatives of the coefficient of the characteristic polynomial.* In Section 3.1 we have given an explicit formula for the entries of $\operatorname{adj}(zI - M_\sigma)$, with M_σ being an arbitrary Fiedler matrix. The aim of this subsection is to provide, in Proposition 3.4, a recursive formula for the coefficients of $\operatorname{adj}(zI - A)$ when viewed as a matrix polynomial in z, where $A \in \mathbb{C}^{n\times n}$ is an arbitrary matrix. This is an interesting theoretical result that gives an alternative description of the coefficients of $\operatorname{adj}(zI - A)$ and, as a consequence of Lemma 3.1, of the gradient of the characteristic polynomial of A. But it may also have a practical interest, as it provides a recursive way to construct these coefficients.

PROPOSITION 3.4 (Gantmacher, 1959, Ch. 4, §4) Let $A \in \mathbb{C}^{n\times n}$ and let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be the characteristic polynomial of A. Let the matrices $A_1, A_2, \ldots, A_n \in \mathbb{C}^{n\times n}$ be defined by the following recurrence relation

$$\begin{cases} A_n = I, \quad \text{and} \\ A_k = A \cdot A_{k+1} + a_k I, \quad \text{for } k = n-1, n-2, \ldots, 1. \end{cases} \tag{3.9}$$

Then,

$$\operatorname{adj}(zI - A) = \sum_{k=0}^{n-1} z^k A_{k+1}.$$

We note that, as a consequence of the recursive relations of the Horner shifts (2.6), the matrices A_k are the Horner shifts of $p(z) = \det(zI - A)$ evaluated at A. More precisely:

$$A_k = p_{n-k}(A) = A^{n-k} + a_{n-1}A^{n-k-1} + \cdots + a_{k+1}A + a_k I.$$

With this in mind, Proposition 3.1 gives the following expression for the gradient of the kth coefficient of the characteristic polynomial of A:

$$\nabla a_k(A) = -\left[\mathrm{vec}(p_{n-k-1}(A^T))\right]^T, \quad \text{for} \quad k = 0,1,\ldots,n-1. \tag{3.10}$$

Proposition 3.4 has been used in Edelman & Murakami (1995) to get an explicit formula for the derivatives of the coefficients of $\det(zI-C)$, with C being a Frobenius companion matrix. For this, the authors take advantage of the explicit expression of the matrices A_k defined in (3.9) with $A = C$, which are very simple in this case (see (Edelman & Murakami, 1995, p. 768)). However, for A being an arbitrary Fiedler matrix, the matrices A_k become much more involved, and it is not easy to get an explicit expression of these matrices just using (3.9). For this reason, we have obtained the expression of the entries of $\mathrm{adj}(zI-A)$ by other means. However, Proposition 3.4 gives us an alternative way to get $\mathrm{adj}(zI-A)$ using the Horner shifts of A.

We want to emphasize that, as a consequence of the previous remarks, the polynomial $p_{ij}^{(\sigma,k)}(a_0,a_1,\ldots,a_{n-1})$ in Theorem 3.3 corresponds to the (j,i) entry of the matrix $p_{n-k-1}(M_\sigma)$. In the following section, we display these matrices for some particular relevant cases, including the Frobenius companion matrices. It is also interesting to note that Corollary 3.1 implies that the first and second Frobenius companion matrices are the only Fiedler matrices M_σ for which all Horner shifts $p_k(M_\sigma)$ have entries which are linear multivariable polynomials in the coefficients of $p(z)$. For all other Fiedler matrices M_σ, there is at least one k such that $p_k(M_\sigma)$ contains some quadratic entries.

3.2.2 *Some particular cases.* We obtain in this section the explicit expression (3.6) for particular Fiedler matrices that are, or may be, of interest in practice. We start with the classical Frobenius companion matrices in Theorem 3.5, where we get analogous formulas to the ones obtained in Edelman & Murakami (1995) for the Frobenius companion matrix considered in that paper.

THEOREM 3.5 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial of degree n, let $C = C_1$ or C_2 be the first or second Frobenius companion matrix of $p(z)$, and let $E \in \mathbb{C}^{n\times n}$. If $\widetilde{p}(z) = z^n + \sum_{k=0}^{n-1} \widetilde{a}_k z^k$ is the characteristic polynomial of $C+E$. Then, to first order in E, for $k = 0,1,\ldots,n-1$:

(i) If $C = C_1$:

$$\widetilde{a}_k - a_k = \sum_{s=0}^{k} \sum_{j=1}^{n-k-1} a_s E_{j-s+k+1,j} - \sum_{s=k+1}^{n} \sum_{j=n-k}^{n} a_s E_{j-s+k+1,j}. \tag{3.11}$$

(ii) If $C = C_2$:

$$\widetilde{a}_k - a_k = \sum_{s=0}^{k} \sum_{i=1}^{n-k-1} a_s E_{i,i-s+k+1} - \sum_{s=k+1}^{n} \sum_{i=n-k}^{n} a_s E_{i,i-s+k+1}. \tag{3.12}$$

Proof. For claim (i), we recall that, if σ is the bijection associated with C_1, then $\mathrm{PCIS}(\sigma) = (0,\ldots,0)$. For this bijection, $\mathfrak{i}_\sigma(n-j:n-i) = j-i+1$ holds for $i \leqslant j$. Then, applying part (a) in Theorem 3.3, we get

$$\widetilde{a}_k - a_k = \sum_{\substack{j<i \\ k+1+i-n\leqslant i-j\leqslant k+1}} a_{j-i+1+k} E_{ij} - \sum_{\substack{j\geqslant i \\ n-k-i+1\leqslant j-i+1\leqslant n-k}} a_{j-i+1+k} E_{ij}.$$

With the change of variables $s = j-i+1+k$ the claim is proved.

For claim (ii), we recall that the bijection σ associated with C_2 satisfies $\mathrm{PCIS}(\sigma) = (1,\ldots,1)$. For this bijection, $\mathfrak{c}_\sigma(n-i:n-j) = i-j+1$, when $j \leqslant i$. Then, applying part (b) in Theorem 3.3, we get

$$\widetilde{a}_k - a_k = \sum_{\substack{j>i \\ k+1+j-n\leqslant j-i\leqslant k+1}} a_{i-j+1+k} E_{ij} - \sum_{\substack{j\leqslant i \\ n-k-j+1\leqslant i-j+1\leqslant n-k}} a_{i-j+1+k} E_{ij}.$$

Again, we use the change of variables $s = i-j+1+k$ to get the result . $\square$

According to (3.10), the matrix $p_{n-k-1}(A^T)$ encodes the information about $\nabla a_k(A)$. In particular, the (i,j) entry of $p_{n-k-1}(A^T)$ is the coefficient of E_{ij} in (3.1). In the case of Frobenius companion matrices, these Horner shifts can be computed without too much effort, since they are equal to:

$$p_{n-k-1}(C_1^T) = p_{n-k-1}(C_2) = \left[\begin{array}{ccc|cccc} 0 & \dots & 0 & 1 & & & 0 \\ -a_k & & & a_{n-1} & 1 & & \\ \vdots & \ddots & & \vdots & a_{n-1} & \ddots & \\ -a_1 & \ddots & -a_k & a_{k+1} & \vdots & \ddots & 1 \\ -a_0 & \ddots & \vdots & & a_{k+1} & \ddots & a_{n-1} \\ & \ddots & -a_1 & & & \ddots & \vdots \\ 0 & & -a_0 & 0 & & & a_{k+1} \end{array}\right], \quad \text{for } k = 0, 1, \dots, n-1, \tag{3.13}$$

where the first block-column contains $n-k-1$ columns, and the second block-column contains $k+1$ columns. The reader may check that, indeed, the (i,j) entry of (3.13) is the coefficient of E_{ij} in (3.11). The same happens with the transpose of (3.13) and formula (3.12).

Excluding the Frobenius companion matrices, the simplest Fiedler matrices are those corresponding to bijections with just one inversion (resp., consecution) at 0, and consecutions (resp., inversions) elsewhere. These particular Fiedler matrices present several numerical advantages that may be of interest in new enhancements of the current codes for the Polynomial Eigenvalue Problem (like MATLAB's `polyeig`). To be precise, one of these matrices is

$$F = M_0(M_{n-1}M_{n-2}\cdots M_1) = \begin{bmatrix} -a_{n-1} & 1 & & & \\ -a_{n-2} & 0 & 1 & & \\ \vdots & & \ddots & \ddots & \\ -a_1 & & & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

and the other one is F^T.

THEOREM 3.6 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial of degree n, let $M_\sigma = F$ be the Fiedler companion matrix of $p(z)$ with $\mathrm{PCIS}(\sigma) = (0,1,1,\dots,1)$ and let $E \in \mathbb{C}^{n\times n}$. If $\widetilde{p}(z) = z^n + \sum_{k=0}^{n-1} \widetilde{a}_k z^k$ is the characteristic polynomial of $F+E$, then, to first order in E,

$$\begin{aligned} \widetilde{a}_k - a_k = \quad & \sum_{j=k+1}^{n-1} a_0 a_{n+k+1-j} E_{nj} + \sum_{s=0}^{k} \sum_{i=1}^{n-k-2} a_s E_{i,i+k+1-s} + \sum_{s=1}^{k} a_s E_{n-k-1,n-s} - E_{n-k-1,n} \\ & - \sum_{s=k+1}^{n} \sum_{i=n-k}^{n-1} a_s E_{i,i+k+1-s} - E_{n-k-1,n} - a_{k+1} E_{nn}. \end{aligned} \tag{3.14}$$

Proof. To compute the polynomials $p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})$ in (3.6), according to Theorem 3.3, we distinguish the following cases:

(a) If $j=n$ and $i=1,2,\dots,n-1$, we have $v_{n-j}=0$ and $v_{n-i}=1$ and $\mathfrak{i}_\sigma(0:n-j-1)+\mathfrak{c}_\sigma(0:n-i-1)=n-i-1$. Therefore, $p_{in}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1}) = 1$ if $i=n-k-1$ and $p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})=0$ otherwise.

(b) If $j=n$ and $i=n$, we have $v_{n-j}=v_{n-i}=0$ and $\mathfrak{i}_\sigma(n-j:n-i-1)=0$. Therefore, $p_{nn}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})=a_{k+1}$ if $n-1 \geqslant k \geqslant 0$ and $p_{nn}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})=0$ otherwise.

(c) If $j=1,2,\dots,n-1$ and $i=n$, we have $v_{n-j}=1$ and $v_{n-i}=0$ and $\mathfrak{i}_\sigma(n-i+1:n-j-1)=0$. Therefore, $p_{nj}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1}) = -a_{n+k-j+1}a_0$ if $j \geqslant k+1$ and $p_{nj}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})=0$ if $j<k+1$.

(d) If $i,j=1,2,\dots,n-1$ and $j \leqslant i$, we have $v_{n-i}=v_{n-j}=1$ and $\mathfrak{c}_\sigma(n-i:n-j)=i-j+1$. Therefore, $p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})=a_{k+1+i-j}$ if $n-k-j+1 \leqslant i-j+1 \leqslant n-k$ and $p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})=0$ otherwise. With the change of variable $s=k+1+i-j$ we get $p_{i,i+k+1-s}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})=a_s$ if $k+1 \leqslant s \leqslant n$ and $n-k \leqslant i \leqslant n-1$, and $p_{i,i+k+1-s}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})=0$ otherwise.

(e) If $i,j=1,2,\ldots,n-1$ and $j>i$, we have $v_{n-i}=v_{n-j}=1$ and $\mathfrak{c}_\sigma(n-j:n-i-1)=j-i$. Therefore, $p_{ij}^{(\sigma,k)}(a_0,a_1,\ldots,a_{n-1})=a_{k+1+i-j}$ if $k+1+j-n\leqslant j-i\leqslant k+1$ and $p_{ij}^{(\sigma,k)}(a_0,a_1,\ldots,a_{n-1})=0$ otherwise. With the change of variable $s=k+1+i-j$ we get $p_{i,i+k+1-s}^{(\sigma,k)}(a_0,a_1,\ldots,a_{n-1})=a_s$ if $0\leqslant s\leqslant k$ and $1\leqslant i\leqslant n-k-1$, and $p_{i,i+k+1-s}^{(\sigma,k)}(a_0,a_1,\ldots,a_{n-1})=0$ otherwise.

□

Theorem 3.6 illustrates how a single change in the PCIS of the Frobenius companion matrix, i. e., just to change the position of the factor M_0 in the product defining C_1 and C_2, implies the appearance of quadratic terms in the coefficients of $p(z)$ in the formula for the gradient of the coefficients of the characteristic polynomial (see the first summand in the right-hand-side of (3.14)). As before, this can also be seen by explicitly displaying the Horner shifts evaluated at F:

$$p_{n-k-1}(F)=\left[\begin{array}{cccc|cccc} 0 & & & & 1 & & & 0 \\ -a_k & & & & a_{n-1} & \ddots & & \vdots \\ \vdots & \ddots & & & \vdots & \ddots & 1 & 0 \\ -a_1 & & -a_k & & a_{k+2} & & a_{n-1} & -a_0 \\ -a_0 & \ddots & \vdots & -a_k & a_{k+1} & \ddots & \vdots & -a_0a_{n-1} \\ & \ddots & -a_1 & \vdots & & \ddots & a_{k+2} & \vdots \\ & & -a_0 & -a_1 & & & a_{k+1} & -a_0a_{k+2} \\ & & & 1 & & & & a_{k+1} \end{array}\right], \quad \text{for } k=0,1,\ldots,n-3,$$

$$p_1(F)=\begin{bmatrix} 0 & & & & & 0 \\ -a_{n-2} & 1 & & & & \\ -a_{n-3} & a_{n-1} & 1 & & & \\ \vdots & & a_{n-1} & \ddots & & \\ \vdots & & & \ddots & 1 & \\ -a_1 & & & & a_{n-1} & -a_0 \\ 1 & & & & 0 & a_{n-1} \end{bmatrix}, \quad \text{and} \quad p_0(F)=I.$$

The number of columns in the first block-column of $p_{n-k-1}(F)$ above is $n-k-1$, and the number of columns in the second block column is $k+1$. The reader may check that the (i,j) entry of $p_{n-k-1}(F)^T$ is the coefficient of E_{ij} in (3.14).

Our last example is the case of a pentadiagonal Fiedler matrix. For any $n\geqslant 3$, there are at least four pentadiagonal matrices corresponding to bijections whose PCIS are $(1,0,1,0,\ldots)$, $(0,1,0,1,\ldots)$, $(1,1,0,1,0,\ldots)$, and $(0,0,1,0,1,0,\ldots)$ (see De Terán *et al.* (2010)). Formulas here, as can be seen in Theorem 3.7, become much more involved.

THEOREM 3.7 Let $p(z)=z^n+\sum_{k=0}^{n-1}a_kz^k$ be a monic polynomial of degree n, let M_σ be the Fiedler companion matrix of $p(z)$ with $\mathrm{PCIS}(\sigma)=(1,0,1,0,\ldots)$ and let $E\in\mathbb{C}^{n\times n}$. If $\widetilde{p}(z)=z^n+\sum_{k=0}^{n-1}\widetilde{a}_kz^k$ is the characteristic polynomial of $M_\sigma+E$, then, to first order in E,

$$\begin{aligned}\widetilde{a}_k-a_k=&-\sum_{s=k+1}^{n}\left(\sum_{r=\lceil\frac{n+s}{2}\rceil-k}^{n/2}a_sE_{2(k+r-s)+1,2r-1}+\sum_{r=\lfloor\frac{n+s}{2}\rfloor-k}^{n/2}a_sE_{2r,2(k+r-s+1)}\right)\\ &+\sum_{s=0}^{k}\left(\sum_{r=1}^{\lceil\frac{n+s}{2}\rceil-k-1}a_sE_{2(k+r-s)+1,2r-1}+\sum_{r=1}^{\lfloor\frac{n+s}{2}\rfloor-k-1}a_sE_{2r,2(k+r-s+1)}\right)-\sum_{s=\max\{1,\frac{n}{2}-k\}}^{\min\{\frac{n}{2},n-k-1\}}E_{2s,2(n-k-s)-1}\\ &+\sum_{s=\max\{0,2k-n+2\}}^{k}\left(\sum_{r=1}^{s-k+\frac{n}{2}}\sum_{m=\max\{0,2k+2r-s-n\}}^{\min\{s,2r-2\}}a_{n-2r+2+m}a_{s-m}E_{2r-1,2(k-s+r)}\right.\\ &\left.+\sum_{r=1}^{s-k+\frac{n}{2}-1}\sum_{m=\max\{0,2k+2r-s-n+1\}}^{\min\{s,2r-1\}}a_{n-2r+1+m}a_{s-m}E_{2(k+r-s)+1,2r}\right)\end{aligned}$$

if n is an even number, or

$$\begin{aligned}\widetilde{a}_k - a_k = &-\sum_{s=k+1}^{n}\left(\sum_{r=\lceil\frac{n+s}{2}\rceil-k}^{\frac{n+1}{2}} a_s E_{2r-1,2(k-s+r)+1} + \sum_{r=\lfloor\frac{n+s}{2}\rfloor-k}^{\frac{n-1}{2}} a_s E_{2(k-s+r+1),2r}\right)\\
&+\sum_{s=0}^{k}\left(\sum_{r=1}^{\lceil\frac{n+s}{2}\rceil-k-1} a_s E_{2r-1,2(k-s+r)+1} + \sum_{r=1}^{\lfloor\frac{n+s}{2}\rfloor-k-1} a_s E_{2(k-s+r+1),2r}\right) - \sum_{s=\max\{1,\frac{n+1}{2}-k\}}^{\min\{\frac{n+1}{2},n-k-1\}} E_{2s-1,2(n-k-s)}\\
&+\sum_{s=\max\{0,2k-n+2\}}^{k}\ \sum_{r=1}^{s-k+\frac{n-1}{2}}\left(\sum_{m=\max\{0,2k-n-s+2r+1\}}^{\min\{s,2r-1\}} a_{n-2r+1+m}a_{s-m}E_{2r,2(k-s+r)+1}\right.\\
&\left.+\sum_{m=\max\{0,2k-n-s+2r\}}^{\min\{s,2r-2\}} a_{n-2r+2+m}a_{s-m}E_{2(k-s+r),2r-1}\right)\end{aligned}$$

if n is an odd number.

Proof. We give a sketch of the proof when the degree of $p(z)$ is even, since the odd case is similar. To compute the polynomials $p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})$ in (3.6) we have to distinguish several cases.

(a) If i and j are odd numbers, then we get

$$p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1}) = \begin{cases} a_{k+1+\frac{j-i}{2}}, & \text{if } j \geqslant i \text{ and } n-\frac{i+j}{2} \leqslant k \leqslant n-1-\frac{j-i}{2},\\ -a_{k+1+\frac{j-i}{2}}, & \text{if } j < i \text{ and } \frac{i-j}{2}-1 \leqslant k \leqslant n-1-\frac{i+j}{2},\\ 0, & \text{otherwise.}\end{cases}$$

(b) If i and j are even numbers, then we get

$$p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1}) = \begin{cases} a_{k+1+\frac{i-j}{2}}, & \text{if } j \leqslant i \text{ and } n-\frac{i+j}{2} \leqslant k \leqslant n-1-\frac{i-j}{2},\\ -a_{k+1+\frac{i-j}{2}}, & \text{if } j > i \text{ and } \frac{j-i}{2}-1 \leqslant k \leqslant n-1-\frac{i+j}{2} \quad \text{and}\\ 0, & \text{otherwise.}\end{cases}$$

(c) If i is an odd number and j is an even number, then we get

$$p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1}) = \begin{cases} 1, & \text{if } k = n-\frac{i+j+1}{2},\\ 0, & \text{otherwise.}\end{cases}$$

(d) If i is an even number and j is an odd number, then we get

$$p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1}) = \begin{cases} \displaystyle\sum_{m=\max\{0,k-\frac{j-i-1}{2}-n+j\}}^{\min\{k-\frac{j-i-1}{2},i-1\}} -a_{n-i+1+m}a_{k-\frac{j-i-1}{2}-m}, & \text{if } j > i \text{ and } \frac{j-i-1}{2} \leqslant k \leqslant n-\frac{j-i+3}{2},\\ \displaystyle\sum_{m=\max\{0,k-\frac{i-j-1}{2}-n+i\}}^{\min\{k-\frac{i-j-1}{2},j-1\}} -a_{n-j+1+m}a_{k-\frac{i-j-1}{2}-m}, & \text{if } j < i \text{ and } \frac{i-j-1}{2} \leqslant k \leqslant n-\frac{i-j+3}{2},\\ 0, & \text{otherwise.}\end{cases}$$

The result follows from these formulas for $p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})$, together with some algebraic manipulations and appropriate changes of variables. □

For the pentadiagonal Fiedler matrix we have considered in Theorem 3.7, the matrices $p_{n-k-1}(M_\sigma)$, for $k = 0,1,\dots,n-1$, do not have a simple structure. For illustrative purposes, we include here a 6×6 example. Let M_σ be the Fiedler companion matrix of the polynomial $p(z) = z^6 + \sum_{k=0}^{5} a_k z^k$ associated with a bijection σ such that

$\text{PICS}(\sigma) = (1,0,1,0,1)$. This matrix is

$$M_\sigma = \begin{bmatrix} -a_5 & 1 & 0 & 0 & 0 & 0 \\ -a_4 & 0 & -a_3 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -a_2 & 0 & -a_1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -a_0 & 0 \end{bmatrix}.$$

Then, it can be seen that

$$p_0(M_\sigma) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad p_1(M_\sigma) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -a_4 & a_5 & -a_3 & 1 & 0 & 0 \\ 1 & 0 & a_5 & 0 & 0 & 0 \\ 0 & 0 & -a_2 & a_5 & -a_1 & 1 \\ 0 & 0 & 1 & 0 & a_5 & 0 \\ 0 & 0 & 0 & 0 & -a_0 & a_5 \end{bmatrix},$$

$$p_2(M_\sigma) = \begin{bmatrix} 0 & 0 & -a_3 & 1 & 0 & 0 \\ -a_3 & 0 & -a_2 - a_3a_5 & a_5 & -a_1 & 1 \\ 0 & 1 & a_4 & 0 & 0 & 0 \\ -a_2 & 0 & -a_1 - a_2a_5 & a_4 & -a_0 - a_1a_5 & a_5 \\ 1 & 0 & a_5 & 0 & a_4 & 0 \\ 0 & 0 & -a_0 & 0 & -a_0a_5 & a_4 \end{bmatrix},$$

$$p_3(M_\sigma) = \begin{bmatrix} 0 & 0 & -a_2 & 0 & -a_1 & 1 \\ -a_2 & 0 & -a_1 - a_2a_5 & 0 & -a_0 - a_1a_5 & a_5 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -a_1 & -a_2 & -a_0 - a_1a_5 - a_2a_4 & a_3 & -a_0a_5 - a_1a_4 & a_4 \\ 0 & 1 & a_4 & 0 & a_3 & 0 \\ -a_0 & 0 & -a_0a_5 & 0 & -a_0a_4 & a_3 \end{bmatrix},$$

$$p_4(M_\sigma) = \begin{bmatrix} 0 & 0 & -a_1 & 0 & -a_0 & 0 \\ -a_1 & 0 & -a_0 - a_1a_5 & 0 & -a_0a_5 & 0 \\ 0 & 0 & 0 & 0 & -a_1 & 1 \\ -a_0 & -a_1 & -a_0a_5 - a_1a_4 & 0 & -a_0a_4 - a_1a_3 & a_3 \\ 0 & 0 & 0 & 1 & a_2 & 0 \\ 0 & -a_0 & -a_0a_4 & 0 & -a_0a_3 & a_2 \end{bmatrix},$$

$$p_5(M_\sigma) = \begin{bmatrix} 0 & 0 & -a_0 & 0 & 0 & 0 \\ -a_0 & 0 & -a_0a_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -a_0 & 0 \\ 0 & -a_0 & -a_0a_4 & 0 & -a_0a_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -a_0 & -a_0a_2 & a_1 \end{bmatrix}.$$

Unlike the previous cases C_2 and F, there does not seem to be a simple pattern for $p_{n-k-1}(M_\sigma)$ for arbitrary n, with $\sigma : \{0,1,\ldots,n-1\} \to \{1,\ldots,n\}$ being the bijection such that PCIS(σ)= $(1,0,1,0,\ldots)$.

3.3 *Balancing and backward errors*

Balancing is a standard preprocessing technique for computing the eigenvalues of a given matrix A, which leads, very often, to more accurate results, especially when the entries of A have very different magnitudes (see Parlett & Reinsch (1969)). Actually, balancing is implemented by default as an initial step in the command `eig` for computing eigenvalues in MATLAB. Balancing consists of performing diagonal similarities DAD^{-1} (i. e., with D diagonal) to A, in order to reduce the norm of A by equilibrating as much as possible the $\infty-$norm of all rows and columns. In addition, very frequently balancing reduces the eigenvalue condition numbers (see (Golub & Van Loan, 1996, §7.2.2)).

Balancing first computes in exact arithmetic a matrix $DM_\sigma D^{-1}$, which has the same characteristic polynomial as M_σ, namely $p(z)$. Then a backward stable algorithm is applied to compute the eigenvalues of $DM_\sigma D^{-1}$, so that we get the exact eigenvalues of $DM_\sigma D^{-1} + \widetilde{E}$, with

$$\|\widetilde{E}\| = O(u)\|DM_\sigma D^{-1}\|, \tag{3.15}$$

for some matrix norm $\|\cdot\|$. Now, we can get a crude formula like (3.6) for the change of the coefficients of the characteristic polynomial of $DM_\sigma D^{-1}$ using the identity:

$$\det(zI - DM_\sigma D^{-1} - \widetilde{E}) = \det(zI - M_\sigma - D^{-1}\widetilde{E}D),$$

and applying Theorem 3.3 with the perturbation $D^{-1}\widetilde{E}D$ instead of E. In particular, following the arguments in the proof of Corollary 3.2, we get

$$|\widetilde{a}_k - a_k| \leqslant n^2 \max_{1\leqslant i,j\leqslant n}\left(\left|p_{ij}^{(\sigma,k)}(a_0,a_1,\dots,a_{n-1})\,\frac{d_j}{d_i}\right|\right)\cdot \max_{1\leqslant i,j\leqslant n}|\widetilde{E}_{ij}|,$$

with $\widetilde{E}$ as in (3.15). In this way, we get a formula which provides an "a posteriori" (that is, once the diagonal parameters d_i are known) measure for the backward error of the polynomial root-finding problem using balanced Fiedler matrices.

Though the numerical experiments carried out in Section 4 indicate that balancing usually produces smaller backward errors, we see in Proposition 3.11 that, for any degree, there are infinitely many polynomials for which the condition numbers of all coefficients of the characteristic polynomial of any matrix $DM_\sigma D^{-1}$ are large. This shows that, though in practice balancing Fiedler matrices may be a good strategy for the root-finding problem, there are polynomials, with any degree, for which the strategy does not lead to small backward errors.

3.4 *Conditioning of the characteristic polynomial*

The developments carried out at the beginning of this section are closely related to the conditioning of the characteristic polynomial of the matrix A. The condition number of the characteristic polynomial provides a measure of its sensitivity to perturbations of the matrix. As we have seen at the beginning of this section, this is in turn related with the gradient of the coefficients of the characteristic polynomial. In this subsection, we introduce the condition number (absolute and relative) for the coefficients of the characteristic polynomial, and we relate it with (the norm of) its gradient. In this way, we will see that the backward stability of the polynomial root-finding problem via eigenvalue methods is determined by the conditioning of the characteristic polynomial.

Let us first assume that the entries of the matrix E in (3.1) satisfy $|E_{ij}| \leqslant \varepsilon\|\mathrm{vec}(A)\|_\infty$. Then, using Holder's inequality $|u^T v| \leqslant \|u^T\|_\infty \|v\|_\infty$ (with $\|\begin{bmatrix} u_1 & \dots & u_n \end{bmatrix}\|_\infty = |u_1| + \cdots + |u_n|$)[1], from (3.1) we get, up to first order, the following inequalities:

$$|a_k(A+E) - a_k(A)| \leqslant \|\nabla a_k(A)\|_\infty \cdot \|\mathrm{vec}(E)\|_\infty \leqslant \varepsilon\|\nabla a_k(A)\|_\infty \cdot \|\mathrm{vec}(A)\|_\infty. \tag{3.16}$$

It is straightforward to show that there exists a particular matrix E with $\|\mathrm{vec}(E)\|_\infty = \varepsilon\|\mathrm{vec}(A)\|_\infty$ such that $|\nabla a_k(A)\cdot \mathrm{vec}(E)| = \|\nabla a_k(A)\|_\infty \|\mathrm{vec}(E)\|_\infty$. For this matrix the bound in (3.16) is attained to first order in ε. With this in mind, Proposition 3.8 immediately follows.

PROPOSITION 3.8 Let $A \in \mathbb{C}^{n\times n}$ and $a_k : \mathbb{C}^{n^2} \to \mathbb{C}$ be the kth coefficient of the characteristic polynomial of $X \in \mathbb{C}^{n\times n}$, considered as a function of X. We define the condition numbers $\kappa(a_k, A)$ and $\kappa_{\mathrm{rel}}(a_k, A)$ as

$$\kappa(a_k,A) := \lim_{\varepsilon\to 0}\sup\left\{\frac{|a_k(A+E) - a_k(A)|}{\varepsilon} : \|\mathrm{vec}(E)\|_\infty \leqslant \varepsilon\|\mathrm{vec}(A)\|_\infty\right\} \tag{3.17}$$

and

$$\kappa_{\mathrm{rel}}(a_k,A) := \lim_{\varepsilon\to 0}\sup\left\{\frac{|a_k(A+E) - a_k(A)|}{\varepsilon|a_k(A)|} : \|\mathrm{vec}(E)\|_\infty \leqslant \varepsilon\|\mathrm{vec}(A)\|_\infty\right\}. \tag{3.18}$$

Then

$$\kappa(a_k,A) = \|\nabla a_k(A)\|_\infty \cdot \|\mathrm{vec}(A)\|_\infty \quad \text{and} \quad \kappa_{\mathrm{rel}}(a_k,A) = \frac{\|\nabla a_k(A)\|_\infty \cdot \|\mathrm{vec}(A)\|_\infty}{|a_k(A)|}.$$

[1] Note that, according to the definition of $\|\cdot\|_\infty$ for $m\times n$ matrices, see (Higham, 2002, p. 108), the expressions for $\|u\|_\infty$ and $\|u^T\|_\infty$ are different.

The definition of condition number introduced in (3.17) and (3.18) may look like non-standard, because of the inclusion of vectorizations. However, the presence of $\mathrm{vec}(E)$ is motivated by (3.1). We have included also $\mathrm{vec}(A)$ in the definition to make it more natural. Moreover, due to the identity

$$\|\mathrm{vec}(M_\sigma)\|_\infty = \|p\|_\infty, \tag{3.19}$$

valid for any Fiedler matrix M_σ, this choice will allow us to get a simpler formula for $\kappa(a_k, M_\sigma)$ (see (3.21) below).

Now, Proposition 3.8, together with (3.10), give us the following formulas for $\kappa(a_k, A)$ and $\kappa_{\mathrm{rel}}(a_k, A)$.

COROLLARY 3.3 Let $A \in \mathbb{C}^{n\times n}$ and let $\kappa(a_k, A)$ and $\kappa_{\mathrm{rel}}(a_k, A)$ be the condition numbers defined in (3.17) and (3.18), respectively. Then, for $k = 0, 1, \ldots, n-1$,

$$\kappa(a_k, A) = \|\mathrm{vec}(p_{n-k-1}(A))\|_1 \cdot \|\mathrm{vec}(A)\|_\infty \quad \text{and} \quad \kappa_{\mathrm{rel}}(a_k, A) = \frac{\|\mathrm{vec}(p_{n-k-1}(A))\|_1 \cdot \|\mathrm{vec}(A)\|_\infty}{|a_k(A)|}, \tag{3.20}$$

where $p_{n-k-1}(z)$ is the degree $n-k-1$ Horner shift of the polynomial $p(z) := \det(zI - A)$.

Note that, according to (3.20), the relative and absolute condition numbers depend on the norms of A and the degree $n-k-1$ Horner shift of the characteristic polynomial of A. This Horner shift depends in turn on the coefficients $a_{k+1}, \ldots, a_{n-1}$ of the characteristic polynomial evaluated at A, namely: $p_{n-k-1}(A) = A^{n-k-1} + a_{n-1}(A)A^{n-k-2} + \cdots + a_{k+1}(A)I$.

In particular, when $A = M_\sigma$ is a Fiedler matrix of the polynomial (1.1), formula (3.20) together with Theorem 3.3 and (3.19), give

$$\kappa(a_k, M_\sigma) = \|p\|_\infty \sum_{i,j=1}^{n} |p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})|, \tag{3.21}$$

where $p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})$ are given in Theorem 3.3, and they are polynomials of degree at most 2 in the coefficients of p, namely $a_0, \ldots, a_{n-1}$.

By considering the maximum condition numbers of all coefficients of the characteristic polynomial we arrive to the following notion.

DEFINITION 3.9 Let $A \in \mathbb{C}^{n\times n}$ and set $p(z) = \det(zI - A)$. Let $\kappa(a_k, A)$ and $\kappa_{\mathrm{rel}}(a_k, A)$ be the condition numbers defined in (3.17) and (3.18), respectively. We define the condition number and the relative condition number of the characteristic polynomial of A with respect to perturbations of A as

$$\kappa(p, A) = \max_{k=0,1,\ldots,n-1} \kappa(a_k, A) \quad \text{and} \quad \kappa_{\mathrm{rel}}(p, A) = \max_{k=0,1,\ldots,n-1} \kappa_{\mathrm{rel}}(a_k, A). \tag{3.22}$$

The following result provides bounds for the absolute and relative condition numbers of the characteristic polynomial when A is a Fiedler matrix.

PROPOSITION 3.10 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial, and M_σ be a Fiedler companion matrix of $p(z)$. Let $\kappa(p, M_\sigma)$ and $\kappa_{\mathrm{rel}}(p, M_\sigma)$ be as in (3.22). Then,

$$\|p\|_\infty^2 \leqslant \kappa(p, M_\sigma) \leqslant n^3 \|p\|_\infty^3 \quad \text{and} \quad \frac{\|p\|_\infty^2}{\max\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}} \leqslant \kappa_{\mathrm{rel}}(p, M_\sigma) \leqslant \frac{n^3 \|p\|_\infty^3}{\min\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}}.$$

Moreover, if $C = C_1, C_2$ denotes both the first and second Frobenius companion matrices, then

$$\|p\|_\infty^2 \leqslant \kappa(p, C) \leqslant n^3 \|p\|_\infty^2 \quad \text{and} \quad \frac{\|p\|_\infty^2}{\max\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}} \leqslant \kappa_{\mathrm{rel}}(p, C) \leqslant \frac{n^3 \|p\|_\infty^2}{\min\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}}.$$

Proof. The bound $\kappa(a_k, M_\sigma) \leqslant n^3 \|p\|_\infty^3$ follows immediately from (3.21) and the bound $|p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})| \leqslant n\|p\|_\infty^2$ (see Corollary 3.1), valid for all $i, j = 1, \ldots, n$.

From (3.21) and Lemma 3.4, it follows that $\kappa(a_k, M_\sigma) \geqslant (k+1)|a_{k+1}| \cdot \|p\|_\infty$, for $k = 0, 1, \ldots, n-1$, and $\kappa(a_0, M_\sigma) \geqslant |a_0| \cdot \|p\|_\infty$. Therefore

$$\kappa(p, M_\sigma) = \max_{k=0,1,\ldots,n-1} \kappa(a_k, M_\sigma) \geqslant \|p\|_\infty^2.$$

Finally, from

$$\frac{\kappa(a_k, M_\sigma)}{\max\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}} \leqslant \frac{\kappa(a_k, M_\sigma)}{|a_k|} \leqslant \frac{\kappa(a_k, M_\sigma)}{\min\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}}$$

we get the bounds for $\kappa_{\text{rel}}(p, M_\sigma)$ in the statement.

For the Frobenius companion matrices, we just note that as a consequence of Corollary 3.1, we have $|p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})| \leqslant n\|p\|_\infty$, where σ is the permutation corresponding to either the first or the second Frobenius companion matrix. □

REMARK 3.4 The factor n^3 appearing in all upper bounds in Proposition 3.10 usually overestimates the condition numbers. It is due to an n^2 factor coming from the maximum possible number of nonzero polynomials $p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})$ in the sum of the right-hand side in (3.21). This number is usually much less than n^2. For instance, it is equal to $(k+1)(2n-2k-1)$ for the first and second Frobenius companion matrices, as can be seen from (3.13). It is also $(k+1)(2n-2k-1)$ for the coefficients a_k with $k = 2, \ldots, n-1$, equal to $3n-4$ for a_1 and equal to n for a_0, for the Fiedler matrix F in Theorem 3.6, as can be seen by looking at the matrices $p_{n-k-1}(F)$ in Section 3.2.2.

3.4.1 *Balancing and condition numbers.* Though similar matrices have the same characteristic polynomial, the sensitivity of its coefficients may be quite different. In other words, the condition numbers $\kappa(a_k, A)$ and $\kappa_{\text{rel}}(a_k, A)$ defined in (3.17) and (3.18) are not invariant under diagonal similarity. Since $q(SAS^{-1}) = Sq(A)S^{-1}$, for any polynomial $q(z)$ and any invertible matrix S, formula (3.20) gives

$$\kappa(a_k, SAS^{-1}) = \|\text{vec}(Sp_{n-k-1}(A)S^{-1})\|_1 \|\text{vec}(SAS^{-1})\|_\infty \tag{3.23}$$

and

$$\kappa_{\text{rel}}(a_k, SAS^{-1}) = \frac{\|\text{vec}(Sp_{n-k-1}(A)S^{-1})\|_1 \|\text{vec}(SAS^{-1})\|_\infty}{|a_k(A)|}.$$

The norms of the vectors in the right hand side of the previous expression can be quite different for different matrices S. The optimal balancing for a given A (or, equivalently, a given polynomial $p(z) = \det(zI - A)$) from the point of view of the sensitivity of the characteristic polynomial (or, equivalently, from the point of view of backward errors of the root-finding problem via eigenvalue methods) would be given by some nonsingular diagonal matrix D such that $\kappa_{\text{rel}}(p, DAD^{-1})$ is minimal among all nonsingular diagonal matrices D (see Parlett & Reinsch (1969) for the eigenvalue problem). In the particular case of Fiedler matrices, the following result provides a lower bound for this minimal conditioning.

PROPOSITION 3.11 . Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial, let $\sigma : \{0, 1, \ldots, n-1\} \to \{1, \ldots, n\}$ be a bijection, let M_σ be the Fiedler companion matrix of $p(z)$ associated with σ and let $D \in \mathbb{C}^{n\times n}$ be a diagonal and nonsingular matrix. Then, for $k = 0, 1, \ldots, n-1$,

$$\kappa(a_k, DM_\sigma D^{-1}) \geqslant (k+1)|a_{n-1}| \cdot |a_{k+1}| \quad \text{and} \quad \kappa_{\text{rel}}(a_k, DM_\sigma D^{-1}) \geqslant \frac{(k+1)|a_{n-1}| \cdot |a_{k+1}|}{|a_k|},$$

where we set $a_n = 1$.

Proof. We prove the result for $\kappa(a_k, DM_\sigma D^{-1})$, since the bound for the relative condition number can be obtained just dividing by $|a_k|$. The result is a consequence of the fact that diagonal similarity does not change the diagonal entries of a matrix. From (3.23),

$$\begin{aligned}\kappa(a_k, DM_\sigma D^{-1}) \geqslant& \|\text{diag}(Dp_{n-k-1}(M_\sigma)D^{-1})\|_1 \cdot \|\text{diag}(DM_\sigma D^{-1})\|_\infty = \\ =& \|\text{diag}(p_{n-k-1}(M_\sigma))\|_1 \cdot \|\text{diag}(M_\sigma)\|_\infty .\end{aligned}$$

Now we prove that $\text{diag}(M_\sigma) = (-a_{n-1}, 0, \ldots, 0)$ and $\text{diag}(p_{n-k-1}(M_\sigma)) = (0, \ldots, 0, a_{k+1}, \ldots, a_{k+1})$, where the coefficient a_{k+1} appears $(k+1)$ times.

For the diagonal of M_σ the proof proceeds by induction in n. The case $n = 2$ is immediate, since the only possible M_σ are $\begin{bmatrix} -a_1 & -a_0 \\ 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} -a_1 & 1 \\ -a_0 & 0 \end{bmatrix}$. We assume that the identity is true for Fiedler matrices associated with polynomials of degree $n-1$. For degree n, we have to distinguish two cases.

(a) If σ has a consecution at $n-2$ then, using MATLAB notation for columns and rows, M_σ may be written as,

$$M_\sigma = \begin{bmatrix} -a_{n-1} & 1 & 0 \\ W(:,1) & 0 & W(:,2:n-1) \end{bmatrix},$$

where $W \in \mathbb{C}^{(n-1)\times(n-1)}$ is a Fiedler companion matrix of the polynomial $z^{n-1} + \sum_{k=0}^{n-2} a_k z^k$ (see (De Terán *et al.*, 2013, p. 949)). Therefore, $\mathrm{diag}(M_\sigma) = (-a_{n-1}, 0, W(2,2), W(3,3), \ldots, W(n-1,n-1)) = (-a_{n-1}, 0, \ldots, 0)$, by induction.

(b) If σ has an inversion at $n-2$ then M_σ may be written as

$$M_\sigma = \begin{bmatrix} -a_{n-1} & W(1,:) \\ 1 & 0 \\ 0 & W(2:n-1,:) \end{bmatrix},$$

where $W \in \mathbb{C}^{(n-1)\times(n-1)}$ is a Fiedler companion matrix of the polynomial $z^{n-1} + \sum_{k=0}^{n-2} a_k z^k$ (see (De Terán *et al.*, 2013, p. 949)). Therefore, $\mathrm{diag}(M_\sigma) = (-a_{n-1}, 0, W(2,2), W(3,3), \ldots, W(n-1,n-1)) = (-a_{n-1}, 0, \ldots, 0)$, by induction.

From Lemma 3.4 and equation (3.10), the (i,i) entry of $p_{n-k-1}(M_\sigma)$ is equal to $p_{ii}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1}) = a_{k+1}$, if $n-1 \geqslant k \geqslant n-i$ (that is, $i \geqslant n-k$), and $p_{ii}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1}) = 0$, otherwise. This concludes the proof. □

3.5 *Backward stability for $\|p\|_\infty$ moderate and coefficientwise backward stability*

Corollary 3.2 indicates that computing the roots of scalar polynomials as the eigenvalues of an arbitrary Fiedler matrix is not backward stable if $\|p\|_\infty$ is large, even if we compute the eigenvalues using a backward stable algorithm. This is revealed by the presence of the factor $\|p\|_\infty$ in (3.7) and $\|p\|_\infty^2$ in (3.8). However, when $\|p\|_\infty$ is moderate, (3.8) guarantees backward stability. This fact is in accordance with results in (Van Dooren & Dewilde, 1983, p. 576), where the authors prove that solving matrix *Polynomial Eigenvalue Problems* by applying the QZ algorithm to the Frobenius companion matrix is backward stable, provided that the original matrix polynomial has been previously scaled so that all coefficients have norm less than or equal to 1. For scalar polynomials (not necessarily monic), this condition can be always achieved by dividing all coefficients of the original polynomial $p(z)$ by some sufficiently large number. However, if we want to restrict ourselves to the set of monic polynomials to use the QR algorithm, this is not a valid strategy any more, since we could get a non-monic polynomial after dividing the coefficients of $p(z)$ (monic). In order to keep the polynomial $p(z)$ as in (1.1) within the set of monic polynomials, we can consider another kind of scaling as, for instance:

$$\widehat{p}(z) := \alpha^n p(z/\alpha) = z^n + \sum_{k=0}^{n-1} a_k \alpha^{n-k} z^k.$$

Now, α can be chosen so that $|a_k \alpha^{n-k}| \leqslant 1$, for all $k = 0, 1, \ldots, n-1$. Note that the roots of $p(z)$ can be easily recovered from those of $\widehat{p}(z)$ just dividing by α. Once all coefficients of $\widehat{p}(z)$ have absolute value less than or equal to 1, we can apply the QR algorithm to any Fiedler companion matrix of $\widehat{p}(z)$ to get its roots, and then recover the roots of $p(z)$. However, this does not guarantee that the method is backward stable. It is not difficult to find examples of quadratic polynomials $p(z)$ such that there is a polynomial $\widehat{q}(z)$ with $\|\widehat{p} - \widehat{q}\| = O(u)\|\widehat{p}\|$, but $\|p - q\|/\|p\|$ is $O(1)$, with $q(z) = (1/\alpha^2)\widehat{q}(\alpha z)$.

We want to emphasize that we are not considering in this paper the backward errors of single roots of p, but the backward error of the set of all roots of p. Backward errors of single roots has been considered in Tisseur (2000) for the more general case of matrix Polynomial Eigenvalue Problems. In particular, the backward error of a single computed root $\widetilde{\lambda}$ considered in Tisseur (2000) is:

$$\eta(\widetilde{\lambda}) = \min\left\{\varepsilon : \ (p + \Delta p)(\widetilde{\lambda}) = 0, \quad |\Delta a_i| \leqslant \varepsilon |a_i|, \ i = 0, 1, \ldots, n\right\},$$

where $p(z) = \sum_{k=0}^{n} a_k z^k$, and $\Delta p(z) = \sum_{k=0}^{n} (\Delta a_k) z^k$ are not necessarily monic. It is shown in (Tisseur, 2000, Theorem 7) that, for quadratic matrix polynomials all whose coefficients have 2-norm equal to 1, computing the eigenvalues of its companion pencil (defined in (Tisseur, 2000, p. 347)) with a backward stable algorithm gives a coefficientwise backward stable method for the Quadratic Eigenvalue Problem. Though, as we have mentioned above, we are considering different notions of backward error, this fact seems to be in accordance with Corollary 3.2 when $\|p\|_\infty = 1$ and with the discussion right below.

We also emphasize that the backward stability of polynomial root-finding when $\|p\|_\infty = 1$ does not guarantee small relative backward errors in each coefficient. In other words, we can not guarantee that

$$\max_{k=0,1,\ldots,n-1} \frac{|\widetilde{a}_k - a_k|}{|a_k|} = O(u) \tag{3.24}$$

even in the case $\|p\|_\infty = 1$. In Section 4 we show some numerical experiments where $\|p\|_\infty = 1$ and (3.24) does not hold. However, when $|a_k|$ is moderate, for all $k = 0,1,\ldots,n-1$, and not too close to zero (loosely speaking, of order $\Theta(1)$), then (3.7)–(3.8) imply that (3.24) holds, also in accordance with Tisseur (2000).

4. Numerical experiments

In this section we provide numerical experiments that support our theoretical results. In particular, our goals are: (i) to show whether or not the bounds in (3.7)–(3.8) correctly predicts the dependence on the norm of $p(z)$ of the largest backward error that may be obtained if the roots of $p(z)$ are computed as the eigenvalues of a Fielder matrix with a backward stable eigenvalue algorithm; (ii) to show that if the roots of a polynomial $p(z)$, with moderate coefficients, are computed as the eigenvalues of a Fiedler matrix, then this process is normwise backward stable, regardless of the Fiedler matrix that is used, which implies that, in this situation, any Fiedler matrix can be used for the root-finding problem with the same reliability as the Frobenius companion matrices; (iii) to investigate, from the point of view of backward errors, the effect of balancing Fiedler matrices; and (iv) following Edelman & Murakami (1995), to show that Theorem 3.3 may be used to predict the backward error when the roots of a monic polynomial are computed as the eigenvalues of a Fiedler matrix. Along this section we denote by $u = 2^{-52}$ the machine epsilon in IEEE double precision arithmetic.

Given a monic polynomial $p(z)$ of degree n, we denote by $\{\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_n\}$ the roots of $p(z)$ computed as eigenvalues of a Fiedler matrix M_σ using a backward stable eigenvalue algorithm. If we denote by $\widetilde{p}(z)$ the monic polynomial of degree n whose roots are $\{\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_n\}$, namely, $\widetilde{p}(z) = \prod_{k=0}^{n}(z - \widetilde{\lambda}_k) = z^n + \sum_{k=0}^{n-1} \widetilde{a}_k z^k$, then we are interested in the following quantities:

- the normwise backward error (NBE): $\|\widetilde{p} - p\|_\infty / \|p\|_\infty$, and
- the coefficientwise backward error (CBE): $\max_{k=0,1,\ldots,n-1} (|\widetilde{a}_k - a_k|/|a_k|)$.

In the numerical experiments, we consider monic polynomials of degree 20 and the following Fiedler companion matrices associated with degree-20 polynomials:

(a) the second Frobenius companion matrix $M_{\sigma_1} = C_2$,

(b) the Fiedler matrix M_{σ_2} with $\mathrm{PCIS}(\sigma_2) = (1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0,1,0,1)$, which is a pentadiagonal matrix,

(c) the Fiedler matrix M_{σ_3} with $\mathrm{PCIS}(\sigma_3) = (0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)$, and

(d) the Fiedler matrix M_{σ_4} with $\mathrm{PCIS}(\sigma_4) = (1,1,1,0,0,1,0,0,1,1,0,0,0,1,1,0,1,1,1)$.

Recall that the matrices M_{σ_2} and M_{σ_3} are the Fiedler matrices considered in Theorems 3.7 and 3.6, respectively.

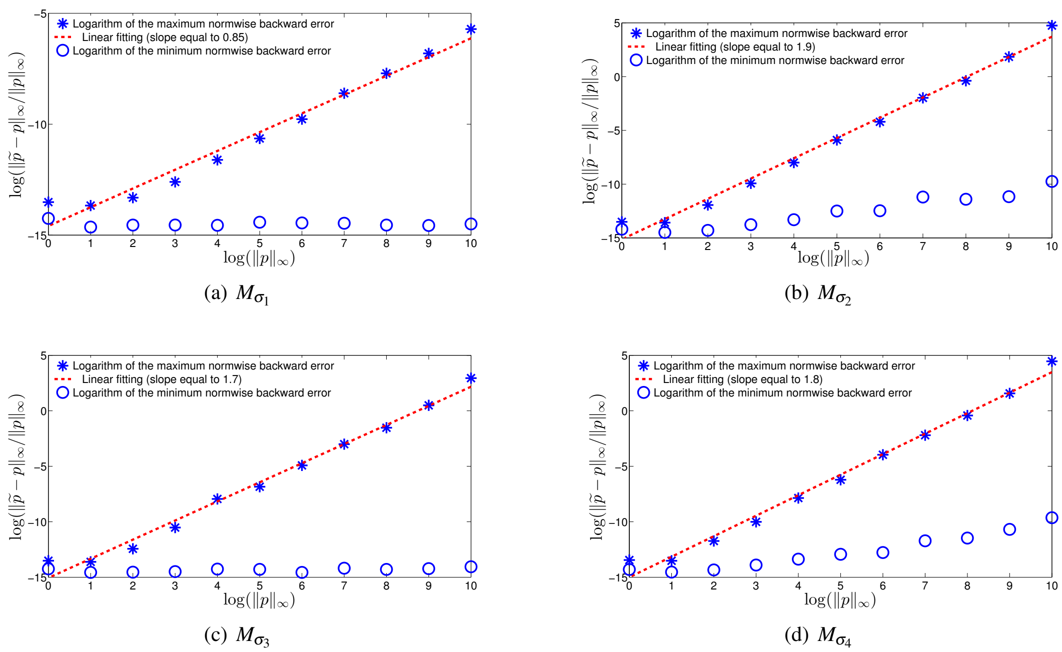
Given a monic polynomial $p(z)$ of degree 20 and a Fiedler matrix M_σ associated with $p(z)$, to compute the polynomial $\widetilde{p}(z)$ we proceed as follows. First, we compute the eigenvalues of M_σ using the function `eig` in MATLAB (with and/or without balancing, see comments below); then, if $\{\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_{20}\}$ denote the computed eigenvalues, we compute the polynomial $\widetilde{p}(z) = \prod_{k=1}^{20}(z - \widetilde{\lambda}_k) = z^{20} + \sum_{k=0}^{19} \widetilde{a}_k z^k$ using the function `vpa` (variable precision arithmetic) followed by the command `poly` on a diagonal matrix whose diagonal entries are $\{\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_{20}\}$, in MATLAB with 32 decimal digits of accuracy.

4.1 *Numerical experiments that show the dependence of the normwise backward error with $\|p\|_\infty$*

In this subsection, we perform numerical experiments to determine whether or not the largest normwise backward errors that may be obtained if the roots of monic polynomials are computed as the eigenvalues of a Fiedler matrix M_σ with a backward stable eigenvalue algorithm, behave like $\|\widetilde{p}-p\|_\infty/\|p\|_\infty = O(u)\|p\|_\infty^2$, when M_σ is a Fiedler matrix other than the Frobenius ones, or like $\|\widetilde{p}-p\|_\infty/\|p\|_\infty = O(u)\|p\|_\infty$, when M_σ is one of the Frobenius companion matrices, as it is predicted by Corollary 3.2. We perform numerical experiments with and without balancing the Fiedler matrices. Our results show that if we do not balance the Fiedler matrices the bound in Corollary 3.2, although in a lot of cases is very pessimistic, predicts well the dependence with $\|p\|_\infty$ of the largest backward errors. If the Fiedler matrices are balanced, our results show that there is still a dependence with $\|p\|_\infty$ of the largest normwise backward errors, and that this dependence is similar for all Fiedler matrices. Also we show that the backward errors that are usually obtained when the Fiedler matrices are balanced are almost independent of the norm of the polynomials, and that polynomial root-finding algorithms using Fiedler matrices are usually normwise backward stable.

In order to see the dependence of the backward error with $\|p\|_\infty$ we proceed as follows. For each $k = 0, 1, \ldots, 10$ we generate 500 random degree-20 polynomials with coefficients of the form $a \cdot 10^c$, where a is drawn from the uniform distribution on the interval $[-1, 1]$ and c is drawn from the uniform distribution on the interval $[-k, k]$, also we set $a_0 = 10^k$. The reasons to set $a_0 = 10^k$ is to fix the infinite norm of the 500 random polynomials to be 10^k. For each of these 11 samples of 500 random polynomials, we compute the normwise backward errors, as it is explained at the beginning of Section 4, when their roots are computed as the eigenvalues of the four Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, with and without balancing them.

In Figures 1-(a), 1-(b), 1-(c), and 1-(d) we plot the decimal logarithms of the maximum and the minimum normwise backward errors obtained for each of the 11 samples of 500 random polynomials against the logarithms of the norm of the polynomials, when their roots are computed as the eigenvalues of $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, respectively, without balancing them. We also plot a linear fitting for the logarithms of the maximum normwise backward errors in order to get the dependence with $\|p\|_\infty$.



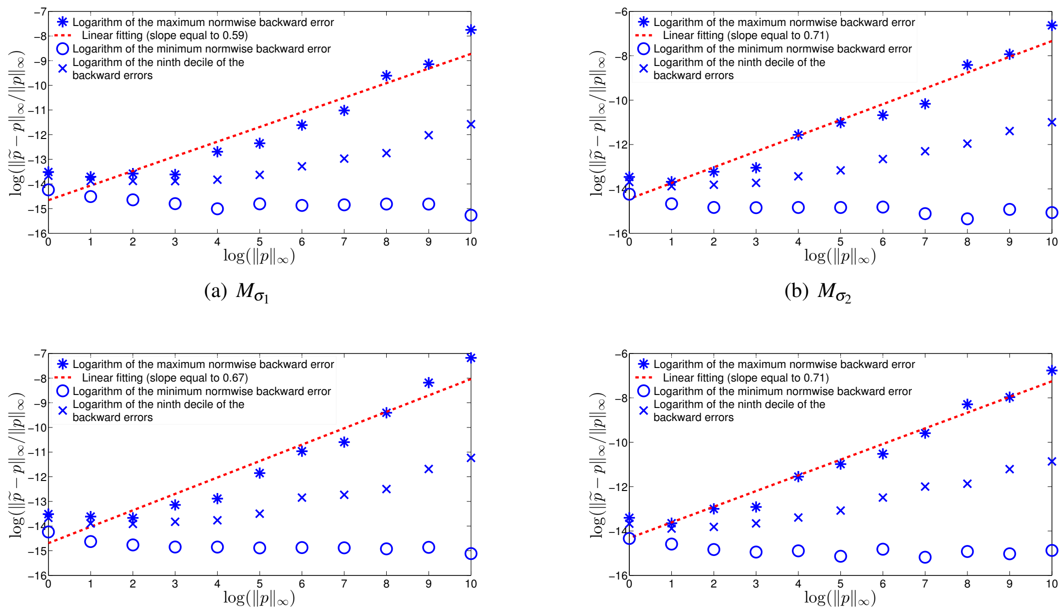
(a) M_{σ_1} (b) M_{σ_2} (c) M_{σ_3} (d) M_{σ_4}

FIG. 1. Decimal logarithms of the maximum and minimum normwise backward errors obtained for each of the 11 samples of 500 random degree-20 polynomials, for $k = 0, 1, \ldots, 10$, with a fixed infinite norm equal to 10^k and with coefficients of the form $a \cdot 10^c$, where a is drawn from the uniform distribution on $[-1, 1]$ and c is drawn from the uniform distribution on $[-k, k]$, and where we set $a_0 = 10^k$, when their roots are computed as the eigenvalues of the Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, without balancing them.

As may be seen in Figures 1-(a), 1-(b), 1-(c), and 1-(d), there is a dependence with $\|p\|_\infty$ of the largest normwise backward errors of the form $\|p\|_\infty^\alpha$. From the linear fittings we obtain $\alpha = 0.85$ for $M_{\sigma_1} = C_2$, $\alpha = 1.9$ for M_{σ_2},

$\alpha = 1.7$ for M_{σ_3}, and $\alpha = 1.8$ for M_{σ_4}. This is consistent with the bound in Corollary 3.2, which predicts $\alpha = 1$ for the Frobenius companion matrices C_1 and C_2, and $\alpha = 2$ for Fiedler matrices other than the Frobenius ones. Also note that in Figures 1-(a),1-(b),1-(c) and 1-(d) it may be seen that the bound in Corollary 3.2 is in some cases very pessimistic, since there are polynomials for which we get small normwise backward errors, regardless of their norms.

Next, we investigate the effect of balancing the Fiedler matrices in the backward errors. In Figures 2-(a), 2-(b), 2-(c), and 2-(d), we plot the decimal logarithms of the maximum and the minimum normwise backward errors obtained for each of the 11 samples of 500 random polynomials against the logarithms of the norm of the polynomials, when their roots are computed as the eigenvalues of $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, respectively, but in this case the Fiedler matrices are balanced before we compute their eigenvalues. As in the previous experiment, we plot a linear fitting for the logarithms of the maximum normwise backward errors in order to get the dependence with $\|p\|_\infty$. We also plot the ninth decile of the normwise backward error for each of the 11 samples.



(a) M_{σ_1} (b) M_{σ_2}

(c) M_{σ_3} (d) M_{σ_4}

FIG. 2. Decimal logarithms of the maximum and minimum normwise backward errors obtained for the 11 samples of 500 random degree-20 polynomials with, for $k = 0, 1, \ldots, 10$, a fixed infinite norm equal to 10^k and with coefficients of the form $a \cdot 10^c$, where a is drawn from the uniform distribution on $[-1, 1]$ and c is drawn from the uniform distribution on $[-k, k]$, and where we set $a_0 = 10^k$, when their roots are computed as the eigenvalues of the Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, balancing them before computing their eigenvalues.

As may be seen in Figures 2-(a), 2-(b), 2-(c), and 2-(d), there is a dependence of the largest backward errors with the norm of the polynomials of the form $\|p\|_\infty^\alpha$, but this dependence is more or less similar for all four Fiedler matrices. In particular, from the linear fittings, we get $\alpha = 0.59$ for $M_{\sigma_1} = C_2$, $\alpha = 0.71$ for M_{σ_2}, $\alpha = 0.67$ for M_{σ_3}, and $\alpha = 0.71$ for M_{σ_4}. Also notice that 90% of the backward errors obtained when the roots of the polynomials are computed as the roots of $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$ are excellent, since they are more or less between 10^{-12} and 10^{-16}, even for polynomials with norms as large as 10^{10}.

4.2 *Numerical experiments with polynomials of moderate coefficients*

In this subsection we show that, from the point of view of backward errors, when the coefficients of $p(z)$ are bounded in absolute value by a moderate number, any Fiedler matrix may be used for the root-finding problem with the same reliability as the Frobenius companion matrices. In particular, we provide numerical evidence that supports what we claim in Section 3.5, namely, that computing the roots of a monic polynomial $p(z)$ as in (1.1), with $|a_i|$ moderate, for $i = 0, 1, \ldots, n-1$, as the eigenvalues of a Fiedler matrix using a backward stable eigenvalue algorithm is normwise backward stable, regardless of the Fiedler matrix that is used. In addition, we show that to have $|a_i|$ moderate, for $i = 0, 1, \ldots, n-1$, it is not enough to guarantee coefficientwise backward stability. Finally, we provide numerical evidence that supports the last sentence in Section 3.5, namely, that (3.24) holds when $|a_i| = \Theta(1)$, for $i = 0, 1, \ldots, n-1$, regardless of the Fiedler matrix that is used.

In the first set of numerical experiments, we consider a random sample of 1000 degree-20 polynomials with coefficients drawn from the uniform distribution on the interval $[-100, 100]$, but we set $a_{19} = 10^{-10}$. The reason for setting $a_{19} = 10^{-10}$ is to show that we may have a small normwise backward error but, at the same time, we may have a big coefficientwise backward error. In Table 1, we give the mean, the maximum and the minimum of the decimal logarithms of the normwise and coefficientwise backward errors (Log-Mean NBE, Log-Maximum NBE, Log-Minimum NBE, Log-Mean CBE, Log-Maximum CBE and Log-Minimum CBE, respectively) obtained when the roots of the polynomials are computed as the eigenvalues of the Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, without balancing them.

	M_{σ_1}	M_{σ_2}	M_{σ_3}	M_{σ_4}
Log-Mean NBE	-13.6	-12.6	-13.6	-12.5
Log-Maximum NBE	-12.9	-11.7	-12.1	-11.6
Log-Minimum NBE	-14.5	-13.4	-14.3	-13.4
Log-Mean CBE	-3.5	-3.8	-3.6	-3.8
Log-Maximum CBE	-2.7	-2.9	-2.7	-2.9
Log-Minimum CBE	-6.7	-6.9	-7.1	-6.3

Table 1. Mean, maximum and minimum of the decimal logarithms of the normwise (NBE) and coefficientwise (CBE) backward errors obtained for 1000 random degree-20 polynomials, with coefficients drawn from the uniform distribution on $[-100, 100]$ and setting $a_{19} = 10^{-10}$, when their roots are computed as the eigenvalues of the Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, without balancing them.

As may be seen in Table 1, the normwise backward errors obtained when the roots of the polynomials are computed as the eigenvalues of $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$ are excellent for all four Fiedler matrices. But also note that this is not true for the coefficientwise backward errors. These results are consistent with the claims in Section 3.5.

Next, we consider a random sample of 1000 degree-20 polynomials with coefficients of the form 10^{c_1} where c_1 is drawn from the uniform distribution on the interval $[-2, 2]$. In Table 2 we give the mean, the maximum and the minimum of the decimal logarithms of the normwise and coefficientwise backward errors (Log-Mean NBE, Log-Maximum NBE, Log-Minimum NBE, Log-Mean CBE, Log-Maximum CBE and Log-Minimum CBE, respectively) obtained when the roots of the polynomials are computed as the eigenvalues of the Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, without balancing them.

	M_{σ_1}	M_{σ_2}	M_{σ_3}	M_{σ_4}
Log-Mean NBE	-14.1	-13.2	-14.1	-13.3
Log-Maximum NBE	-13.4	-11.8	-12.5	-11.7
Log-Minimum NBE	-14.7	-14.5	-14.8	-14.8
Log-Mean CBE	-11.0	-10.2	-11.0	-10.2
Log-Maximum CBE	-10.0	-8.3	-9.1	-8.4
Log-Minimum CBE	-12.4	-12.2	-12.6	-12.7

Table 2. Mean, maximum and minimum of the decimal logarithms of the normwise (NBE) and coefficientwise (CBE) backward errors obtained for 1000 random degree-20 polynomials, with coefficients of the form 10^{c_1}, where c_1 is drawn from the uniform distribution on $[-2, 2]$, when their roots are computed as the eigenvalues of the Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, without balancing them.

As may be seen in Table 2, the normwise backward errors obtained when the roots of the polynomials are computed as the eigenvalues of $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$ are excellent, as in Table 1. The coefficientwise backward errors are not so small as the normwise ones, but they are still excellent since we are dealing with polynomials whose coefficients may have absolute values that differ in four orders of magnitude.

4.3 *Numerical experiments balancing Fiedler matrices*

In this subsection we perform numerical experiments to study, from the point of view of backward errors, the effect of balancing Fiedler matrices. We show that, when a Fiedler matrix M_σ is balanced before computing its eigenvalues, the backward error obtained if we compute the roots of $p(z)$ as the eigenvalues of M_σ may be much smaller than the backward error that is obtained when M_σ is not balanced, regardless of the Fiedler matrix that is used. We show also that balancing a Fiedler matrix is usually enough to guarantee that the process of computing the roots of a polynomial as the eigenvalues of a Fiedler matrix is normwise backward stable, even if the polynomial has large coefficients. Finally, we investigate the effect of the size of the coefficient a_{n-1}, since Proposition 3.11 suggests that it plays a key role in getting or not backward stability after balancing Fiedler matrices. To be precise, Proposition 3.11 shows that, for large values of $|a_{n-1}|$, the condition number of any coefficient of the characteristic polynomial of any Fiedler matrix will be large, regardless of the balancing. This leads us to expect large backward errors when $|a_{n-1}|$ is large.

We consider a random sample of 1000 degree-20 polynomials with coefficients of the form

$$a_1 \cdot 10^{c_1} + i a_2 \cdot 10^{c_2}, \tag{4.1}$$

where i denotes the imaginary unit, and a_1, a_2 are drawn from the uniform distribution on the interval $[-1,1]$ and c_1 and c_2 are drawn from the uniform distribution on the interval $[-10,10]$. These polynomials, considered in Toh & Trefethen (1994), allow us to measure the normwise backward errors with varying orders of magnitude in the coefficients of $p(z)$. We also consider a second sample of 1000 degree-20 polynomials with coefficients of the form (4.1), but we fix $a_{19} = 1$.

For the first sample of random polynomials, in Tables 3-(a) and 3-(b) we give the mean, the maximum and the minimum of the decimal logarithms of the normwise backward errors (Log-Mean NBE, Log-Maximum NBE, Log-Minimum NBE, respectively) obtained when the roots of the polynomials are computed as the eigenvalues of $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, when these Fiedler matrices are not or are balanced, respectively.

(a) The Fiedler matrices are not balanced.

	M_{σ_1}	M_{σ_2}	M_{σ_3}	M_{σ_4}
Log-Mean NBE	-10.5	-2.4	-9.9	-3.0
Log-Maximum NBE	-5.8	3.2	0.1	3.5
Log-Minimum NBE	-14.7	-8.9	-14.7	-10.0

(b) The Fiedler matrices are balanced.

	M_{σ_1}	M_{σ_2}	M_{σ_3}	M_{σ_4}
Log-Mean NBE	-13.1	-13.1	-13.1	-12.9
Log-Maximum NBE	-8.1	-7.5	-8.0	-7.8
Log-Minimum NBE	-14.7	-14.9	-15.1	-14.8

Table 3. Mean, maximum, and minimum of the decimal logarithms of the normwise backward errors obtained for a sample of 1000 random degree-20 polynomials, with coefficients of the form (4.1), when their roots are computed as the eigenvalues of the Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, without balancing and with balancing.

Several observations may be drawn from the data in Tables 3-(a) and 3-(b). First note, from the data in Log-Maximum NBE in Table 3-(a), that if the Fiedler matrices are not balanced, the backward errors obtained may be very large. Note also that the largest of these backward errors is consistent with (3.7) for the Frobenius companion matrices, and with (3.8) for Fiedler matrices other than the Frobenius ones. Second, note that the process of balancing the Fiedler matrices makes that the backward errors obtained after balancing may be much smaller than the backward errors obtained when the Fiedler matrices are not balanced (this is especially evident for M_{σ_2} and M_{σ_3}). Finally, note, from the data in Log-Maximum NBE in Table 3-(b), that there are polynomials for which balancing the Fiedler matrices does not guarantee that the process of computing their roots as the eigenvalues of Fiedler matrices is normwise backward stable.

In Tables 4-(a) and 4-(b) we display the mean, the maximum and the minimum of the decimal logarithms of the normwise backward errors (Log-Mean NBE, Log-Maximum NBE, Log-Minimum NBE, respectively) that are obtained when the roots of the polynomials of the second sample are computed as the eigenvalues of the four Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, when the Fiedler matrices are not or are balanced, respectively. Recall that for this sample of degree-20 random polynomials we set $a_{19} = 1$.

As in the first sample of random polynomials, we may see in Tables 4-(a) and 4-(b) that the backward errors obtained when the Fiedler matrices are not balanced may be very large. Also, we may see that the backward errors may be much smaller when the Fiedler matrices are balanced. Finally note that for this second sample the largest

(a) The Fiedler matrices are not balanced.

	M_{σ_1}	M_{σ_2}	M_{σ_3}	M_{σ_4}
Log-Mean NBE	-6.9	-3.2	-6.9	-3.4
Log-Maximum NBE	-5.6	3.0	-3.4	3.0
Log-Minimum NBE	-9.8	-10.6	-9.9	-11.1

(b) The Fiedler matrices are balanced.

	M_{σ_1}	M_{σ_2}	M_{σ_3}	M_{σ_4}
Log-Mean NBE	-13.9	-13.9	-13.9	-13.7
Log-Maximum NBE	-11.6	-11.1	-11.6	-10.4
Log-Minimum NBE	-15.1	-14.8	-15.0	-15.0

Table 4. Mean, maximum, and minimum of the decimal logarithms of the normwise backward errors obtained when the roots of the polynomials of the second sample of random polynomials (i. e., coefficients from (4.1) and $a_{19} = 1$) are computed as the eigenvalues of the four Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, without balancing and with balancing.

backward errors obtained when the Fiedler matrices are balanced are smaller than the largest ones obtained for the first sample.

4.4 *Using Theorem 3.3 to predict the coefficientwise backward error*

In this subsection we show that Theorem 3.3 can be used to predict the coefficientwise backward error, without computing explicitly the polynomial $\widetilde{p}(z)$ (something that may not be possible for high degree polynomials, since using `vpa` makes this process very slow), and that this backward error is usually small for all Fiedler matrices if the process of balancing is used. Of course, the normwise backward error can be also predicted from Theorem 3.3, but we omit it for brevity. As in Edelman & Murakami (1995) and Toh & Trefethen (1994), we explore the following degree-20 monic polynomials:

($p1$) the Wilkinson polynomial: $p(z) = \prod_{k=1}^{20}(z-k)$,

($p2$) the monic polynomial with zeros: $-2, -1.8, -1.6, \ldots, 1.6, 1.8$,

($p3$) $p(z) = (20!)\sum_{k=0}^{20} z^k/k!$,

($p4$) the Bernoulli polynomial of degree 20,

($p5$) $p(z) = \sum_{k=0}^{20} z^k$,

($p6$) the monic polynomial with zeros $2^{-10}, 2^{-9}, \ldots, 2^8, 2^9$,

($p7$) the Chebyshev polynomial of degree 20,

($p8$) the monic polynomial with zeros equally spaced on a sine curve, that is,

$$p(z) = \prod_{k=-10}^{9}\left(z - \frac{2\pi}{19}(k+0.5) - i \cdot \sin\frac{2\pi}{19}(k+0.5)\right).$$

As in Edelman & Murakami (1995), we first compute the coefficients exactly or with high precision using Mathematica. We then read these coefficients into MATLAB and take the rounded coefficients stored in MATLAB as our official test cases. Also, we consider again the four Fiedler companion matrices associated with degree-20 polynomials introduced at the beginning of Section 4, namely $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$.

We repeat the numerical experiments performed in Edelman & Murakami (1995). Our results show that Theorem 3.3 always predicts a small componentwise backward error, regardless of the Fiedler matrix that is used, and that this predicted backward error is usually pessimistic by at most one, two or three orders of magnitude, except for the polynomial $p6$, where the predicted backward error is pessimistic by 6 orders of magnitude. Note that in this case the ratio $(|a_{19}| \cdot |a_1|)/|a_0|$ is of order 2^{19}, so Proposition 3.11 ensures that the condition number for the coefficient a_0 is large. However, the perturbations in the numerical experiments does not seem to affect this coefficient in such a severe way.

In order to use Theorem 3.3 to predict the coefficientwise backward error, we need to model the backward error introduced by the algorithm for computing the eigenvalues of a Fiedler matrix. Since standard eigenvalue algorithms first balance the matrix, if we set $B_\sigma := DM_\sigma D^{-1}$, where D is the diagonal matrix that balances M_σ, then a backward stable eigenvalue algorithm applied to a Fiedler matrix M_σ computes the exact eigenvalues of the matrix $B_\sigma + \widetilde{E}$, with $\|\widetilde{E}\| = O(u)\|B_\sigma\|$. Due to these considerations, we model the backward error introduced by a backward stable eigenvalue algorithm applied to M_σ by means of an error matrix $\widetilde{E} = (\widetilde{E}_{ij})$, with

$$\widetilde{E}_{ij} = 2^{-52} \cdot \|B_\sigma\|_2 \cdot \varepsilon_{ij} \quad \text{for } i,j = 1,2,\ldots,20, \tag{4.2}$$

where ε_{ij} is drawn from the uniform distribution on the interval $[-1,1]$. If we denote by $\{\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_n\}$ the eigenvalues of $B_\sigma + \widetilde{E}$ then, since a similarity transformation does not change the characteristic polynomial of a matrix, these eigenvalues are the roots of the characteristic polynomial of $D^{-1}(B_\sigma + \widetilde{E})D = M_\sigma + E$, where $E = D^{-1}\widetilde{E}D$, that is, they are the roots of the polynomial $\det(zI - M_\sigma - E) = \widetilde{p}(z) = z^n + \sum_{k=0}^{n-1} \widetilde{a}_n z^k$. Then, we can use Theorem 3.3 to compute the coefficients of $\widetilde{p}(z)$, up to first order in E, and use $\max_{a_k \neq 0} |\widetilde{a}_k - a_k|/|a_k|$ as a prediction of the coefficientwise backward error. Finally we can compare this predicted backward error with the observed one, computed as explained at the beginning of Section 4.

In Table 5, we display the decimal logarithms of the predicted and the observed coefficientwise backward error (Log Predicted CBE and Log Observed CBE, respectively), when the roots of the polynomials $p1$-$p8$ are computed as the eigenvalues of the Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$.

(a) M_{σ_1}

	$p1$	$p2$	$p3$	$p4$	$p5$	$p6$	$p7$	$p8$
Log Predicted CBE	-12.0	-10.4	-13.4	-13.3	-13.7	-8.8	-12.3	-13.6
Log Observed CBE	-13.9	-13.5	-14.2	-13.8	-13.9	-13.8	-14.7	-14.6

(b) M_{σ_2}

	$p1$	$p2$	$p3$	$p4$	$p5$	$p6$	$p7$	$p8$
Log Predicted CBE	-12.3	-12.1	-10.3	-13.3	-13.4	-9.3	-13.2	-13.3
Log Observed CBE	-13.8	-14.0	-12.0	-13.8	-13.6	-14.1	-13.7	-13.9

(c) M_{σ_3}

	$p1$	$p2$	$p3$	$p4$	$p5$	$p6$	$p7$	$p8$
Log Predicted CBE	-12.1	-12.9	-13.5	-13.0	-13.7	-8.8	-12.3	-13.5
Log Observed CBE	-14.0	-13.8	-13.7	-14.1	-13.9	-13.8	-14.7	-14.3

(d) M_{σ_4}

	$p1$	$p2$	$p3$	$p4$	$p5$	$p6$	$p7$	$p8$
Log Predicted CBE	-12.3	-12.8	-12.8	-13.5	-13.3	-9.6	-13.9	-13.7
Log Observed CBE	-14.0	-14.2	-13.4	-14.1	-13.9	-14.0	-15.1	-14.1

Table 5. Decimal logarithms of the predicted and observed coefficientwise backward error for the Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$ of the eight polynomials $p1$-$p8$.

As may be seen in Table 5, the coefficientwise backward errors are well predicted by Theorem 3.3, with the exception of the polynomial $p6$. In Edelman & Murakami (1995) it was also observed that the coefficientwise backward error for $p6$, when the Frobenius companion matrix is used to compute its roots, was far more favorable than the predicted one. For this polynomial and for the four Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}, M_{\sigma_4}$, the coefficientwise backward error comes from $|\widetilde{a}_0 - a_0|/|a_0|$. The most important conclusion to be extracted from Table 5 for our purposes is that the four Fiedler matrices $M_{\sigma_1}, M_{\sigma_2}, M_{\sigma_3}$, and M_{σ_4} behave equally well from the point of view of backward errors in polynomials $p1 - p8$.

5. The Sylvester space of Fiedler matrices

The study of the geometry of matrix spaces sheds light on the explanation of numerical processes involving matrices or matrix pencils. In particular, the theory of orbits has been used in the analysis of errors of the algorithms for computing eigenvalues and canonical forms (see Arnold (1971), Edelman *et al.* (1997, 1999) and Edelman & Murakami (1995)). In this section, and inspired by the motivating paper by Edelman & Murakami (1995), we analyze from a geometrical point of view the polynomial root-finding problem solved as an eigenvalue problem with Fiedler companion matrices. Our main result is Theorem 5.3, where we prove that the space of Sylvester matrices associated with a given Fiedler matrix M_σ is transversal to the similarity orbit of M_σ. This result extends the corresponding one for Frobenius companion matrices (Edelman & Murakami, 1995, Prop. 2.1).

Let $p(z)$ be a monic polynomial as in (1.1) and let M_σ be a Fiedler matrix of $p(z)$. Let us consider the Euclidean matrix space $\mathbb{C}^{n\times n}$ with the usual Frobenius inner product

$$(A,B) = \operatorname{tr}(AB^*),$$

where M^* denotes the conjugate transpose of $M \in \mathbb{C}^{n\times n}$. In this space, the set of matrices similar to a given matrix $A \in \mathbb{C}^{n\times n}$ is a differentiable manifold in $\mathbb{C}^{n\times n}$. This manifold is the orbit of A under the action of similarity:

$$\mathscr{O}(A) := \{SAS^{-1} \,:\, \det(S) \neq 0\}.$$

We will refer to the elements of a manifold as *points*, even though all manifolds considered in this paper are manifolds whose points are matrices.

It is known that the tangent space of $\mathscr{O}(A)$ at A is the set

$$T_A\mathscr{O}(A) := \{AX - XA \text{ for some } X \in \mathbb{C}^{n\times n}\}.$$

The *normal space* of $\mathscr{O}(A)$ at A, denoted by $N_A\mathscr{O}(A)$, is the set of matrices orthogonal to any matrix in $T_A\mathscr{O}(A)$:

$$N_A\mathscr{O}(A) := \{Y \in \mathbb{C}^{n\times n} \text{ such that } (Y,V) = 0, \ \text{ for all } V \in T_A\mathscr{O}(A)\},$$

and the *centralizer of* A is the set of matrices commuting with A:

$$C(A) := \{X \in \mathbb{C}^{n\times n} \text{ such that } AX - XA = 0\}$$

The following facts are already known:

(a) $C(A^*) = N_A\mathscr{O}(A)$ (see (Arnold, 1971, Lemma, p. 34)).

(b) If A is a non-derogatory matrix, then:

- (b1) $C(A) = \{q(A) \,:\, q \text{ is a polynomial}\}$ (see (Horn & Johnson, 1985, Th. 3.2.4.2)).
- (b2) $\dim C(A) = n$ (see (Arnold, 1971, Corollary, p. 35)).

(c) M_σ is a non-derogatory matrix, for all σ.

For claim (c), just recall that M_σ is similar to C_1, and that C_1 is non-derogatory (see (Horn & Johnson, 1985, p. 147)).

As a consequence of claims (a)–(c) above, we have that $\dim N_{M_\sigma}\mathscr{O}(M_\sigma) = n$, for all σ, so there is a basis of $N_{M_\sigma}\mathscr{O}(M_\sigma)$ consisting of n matrices which are polynomials in M_σ^*. This is stated in Proposition 5.1.

PROPOSITION 5.1 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial, $\sigma : \{0,1,\ldots,n-1\} \to \{1,\ldots,n\}$ be a bijection, M_σ be the Fiedler matrix of $p(z)$ associated with the bijection σ, and let $p_d(z)$ be the dth Horner shift of $p(z)$, for $d = 0,1,\ldots,n$. Set $p_0(M_\sigma) = I_n$ and

$$p_{n-k}(M_\sigma) = M_\sigma^{n-k} + a_{n-1}M_\sigma^{n-k-1} + \cdots + a_{k+1}M_\sigma + a_k I, \quad \text{for } k = 1,\ldots,n-1.$$

Then $\{p_k(M_\sigma)^*\}_{k=0}^{n-1}$ is a basis for $N_{M_\sigma}\mathscr{O}(M_\sigma)$.

Note that the set $\{p_k(M_\sigma)^*\}_{k=0}^{n-1}$ is linearly independent because, since M_σ is non-derogatory, its minimal polynomial coincides with its characteristic polynomial. Any n linearly independent polynomials in M_σ^* would serve as a basis for $N_{M_\sigma}\mathscr{O}(M_\sigma)$, but in Section 3.2.1 we have seen that the matrices $p_k(M_\sigma)$ play an important role in determining how the coefficients of the characteristic polynomial of M_σ change when the matrix is perturbed (see (3.10)).

First order perturbations of the coefficients of $p(z)$, with $p(z) = \det(zI - C_1)$, have been studied in Edelman & Murakami (1995). To do so, the authors decompose the perturbation matrix E as

$$E = E^{\tan} + E^{\mathrm{syl}}, \tag{5.1}$$

where $E^{\tan}$ belongs to the tangent space to $\mathscr{O}(C_1)$ at C_1 and E^{syl} is of the form

$$E^{\mathrm{syl}} = \begin{bmatrix} E_{11} & \dots & E_{1n} \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}.$$

The matrix E^{syl} belongs to the tangent space (at any point) to the *Sylvester space* of C_1. We recall that the (affine) Sylvester space of C_1 is the set of all matrices of the form

$$\begin{bmatrix} E_{11} & E_{12} & \dots & E_{1n} \\ 1 & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{bmatrix},$$

that is, the set of "all first Frobenius companion matrices"[2]. It may be proved that, to first order in E, the matrix $E^{\tan}$ does not affect the coefficients of $p(z)$. Below, we prove an equivalent result for any Fiedler matrix M_σ. For this, we first define the Sylvester space of any Fiedler matrix, which is a natural generalization of the Sylvester space of C_1.

DEFINITION 5.2 (Sylvester space of a Fiedler matrix) Let $\sigma : \{0,1,\dots,n-1\} \to \{1,\dots,n\}$ be a bijection. Then, the (affine) Sylvester space associated with the bijection σ, denoted by $\mathrm{Syl}(\sigma)$, is the set of Fiedler matrices associated with σ, that is,

$$\mathrm{Syl}(\sigma) := \left\{ M_\sigma(p) : \; p(z) = z^n + \sum_{k=0}^{n-1} c_k z^k, \quad c_k \in \mathbb{C} \right\},$$

where $M_\sigma(p)$ is the matrix in (2.2).

For example, the Sylvester space associated with the bijection σ, such that $\mathrm{PCIS}(\sigma) = (1,1,1,0,0,0)$, is the set of matrices of the form

$$\begin{bmatrix} c_6 & c_5 & c_4 & c_3 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_2 & 0 & 1 & 0 \\ 0 & 0 & 0 & c_1 & 0 & 0 & 1 \\ 0 & 0 & 0 & c_0 & 0 & 0 & 0 \end{bmatrix},$$

where $c_k \in \mathbb{C}$, for $k = 0,1,\dots,6$, may take any value. The tangent space of $\mathrm{Syl}(\sigma)$ at a given point, denoted by $\mathrm{TSyl}(\sigma)$, is the set of matrices that we get if we remove the entries identically equal to 1 in the matrix above. In other words, the underlying vector space to the affine space. For example, for the previous bijection σ, the tangent space of

[2] We note that the companion matrix considered in Edelman & Murakami (1995) is not exactly C_1, but the companion matrix obtained from C_1 in (1.2) after performing a symmetry through the main anti-diagonal, and accordingly with the Sylvester space.

$\mathrm{Syl}(\sigma)$ is the set of matrices of the form

$$\begin{bmatrix} c_6 & c_5 & c_4 & c_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_0 & 0 & 0 & 0 \end{bmatrix},$$

where $c_k \in \mathbb{C}$, for $k = 0, 1, \ldots, 6$, may take any value. Observe that the tangent space of $\mathrm{Syl}(\sigma)$ in any matrix $M \in \mathrm{Syl}(\sigma)$ is independent of M. This is the reason why we just write $\mathrm{TSyl}(\sigma)$ without specifying the base point.

In order to extend the transversality identity (5.1) to the Sylvester space of any Fiedler matrix, we first need the following result, which is in turn an extension of (Edelman & Murakami, 1995, Eq. (5), p. 768).

LEMMA 5.1 Let E^{syl} be a matrix in $\mathrm{TSyl}(\sigma)$ with nonzero entries equal to $E_0^{\mathrm{syl}}, E_1^{\mathrm{syl}}, \ldots, E_{n-1}^{\mathrm{syl}}$, where the entry E_k^{syl}, for $k = 0, 1, \ldots, n-1$, is in the same position as the coefficient $-a_k$ in M_σ. Then, for $k = 0, 1, \ldots, n-1$,

$$\mathrm{tr}(E^{\mathrm{syl}}\, p_{n-k-1}(M_\sigma)) = -E_k^{\mathrm{syl}}. \tag{5.2}$$

Proof. Let $\widetilde{p}(z) = z^n + \sum_{k=0}^{n-1} \widetilde{a}_k z^k$ be the characteristic polynomial of $M_\sigma + E^{\mathrm{syl}}$. We know, by Propositions 3.1 and 3.4, that $\widetilde{a}_k = a_k - \mathrm{tr}(E^{\mathrm{syl}}\, p_{n-k-1}(M_\sigma)) + O(\|E^{\mathrm{syl}}\|^2)$. But $M_\sigma + E^{\mathrm{syl}}$ is a Fiedler matrix of the polynomial $z^n + \sum_{k=0}^{n-1}(a_k + E_k^{\mathrm{syl}})z^k$, therefore we have $\widetilde{a}_k = a_k + E_k^{\mathrm{syl}}$. From these two formulas we get

$$\mathrm{tr}(E^{\mathrm{syl}}\, p_{n-k-1}(M_\sigma)) + O(\|E^{\mathrm{syl}}\|^2) = -E_k^{\mathrm{syl}}.$$

Since this last equation is true regardless of the value of $E_0^{\mathrm{syl}}, E_1^{\mathrm{syl}}, \ldots, E_{n-1}^{\mathrm{syl}}$, (5.2) follows. □

THEOREM 5.3 Let $p(z) = z^n + \sum_{k=0}^{n-1} a_k z^k$ be a monic polynomial, $\sigma : \{0, 1, \ldots, n-1\} \to \{1, \ldots, n\}$ be a bijection, and let M_σ be the Fiedler matrix of $p(z)$ associated to the bijection σ. Then $\mathrm{Syl}(\sigma)$ is transversal to $\mathscr{O}(M_\sigma)$ at M_σ, i.e., every matrix $E \in \mathbb{C}^{n\times n}$ can be expressed as

$$E = E^{\tan} + E^{\mathrm{syl}}, \tag{5.3}$$

where $E^{\mathrm{syl}} \in \mathrm{TSyl}(\sigma)$ and $E^{\tan} \in T_{M_\sigma}\mathscr{O}(M_\sigma)$.

Proof. Let E^{syl} be a matrix in $\mathrm{TSyl}(\sigma)$ with nonzero entries $E_k^{\mathrm{syl}} := -\mathrm{tr}(E\, p_{n-k-1}(M_\sigma))$, for $k = 0, 1, \ldots, n-1$, where the entry E_k^{syl} is in the same position as $-a_k$ in M_σ. We may write the matrix E as $E^{\mathrm{syl}} + E^{\tan}$, where $E^{\tan} = E - E^{\mathrm{syl}}$. We have to check that $E^{\tan} \in T_{M_\sigma}\mathscr{O}(M_\sigma)$. Indeed, using Lemma 5.1,

$$\mathrm{tr}(E\, p_{n-k-1}(M_\sigma)) = \mathrm{tr}(E^{\mathrm{syl}}\, p_{n-k-1}(M_\sigma)) + \mathrm{tr}(E^{\tan}\, p_{n-k-1}(M_\sigma)) = \mathrm{tr}(E\, p_{n-k-1}(M_\sigma)) + \mathrm{tr}(E^{\tan}\, p_{n-k-1}(M_\sigma)).$$

From this, we deduce that $\mathrm{tr}(E^{\tan}\, p_{n-k-1}(M_\sigma)) = 0$, for $k = 0, 1, 2, \ldots, n-1$. But, from Proposition 5.1, we have that $\{p_k(M_\sigma)^*\}_{k=0}^{n-1}$ is a basis for $N_{M_\sigma}\mathscr{O}(M_\sigma)$, therefore $E^{\tan} \in T_{M_\sigma}\mathscr{O}(M_\sigma)$. □

Theorem 5.3, together with (5.2) show us that the component $E^{\tan}$ of the perturbation matrix E does not contribute to the first order term of $a_k(M_\sigma + E)$, so that only the "transversal complement" E^{syl} contributes to first order. In other words:

$$a_k(M_\sigma + E) = a_k - \mathrm{tr}(p_{n-k-1}(M_\sigma)E)) + O(\|E\|^2) = a_k - \mathrm{tr}(p_{n-k-1}(M_\sigma)E^{\mathrm{syl}}) + O(\|E\|^2) = a_k(M_\sigma + E^{\mathrm{syl}}) + O(\|E\|^2).$$

Also, from the considerations above, if E_k^{syl} denotes, as in Lemma 5.1, the entry of E^{syl} which is located in the same position as the coefficient $-a_k$ in M_σ, then we have, up to first order in E,

$$E_k^{\mathrm{syl}} = a_k(M_\sigma + E) - a_k = -\sum_{i,j=1}^{n} p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})E_{ij}, \tag{5.4}$$

as in (3.6), with $p_{ij}^{(\sigma,k)}(a_0, a_1, \ldots, a_{n-1})$ given by Theorem 3.3. Recall that the remaining entries of E^{syl} are zero. Hence, from (5.3) and (5.4) we may get explicit expressions for the entries of $E^{\tan} = E - E^{\mathrm{syl}}$ in terms of the entries of E and the coefficients $a_0, a_1, \ldots, a_{n-1}$.

We want to emphasize that in the approach followed by Edelman & Murakami (1995), the fact that E^{syl} is transversal to the tangent space of $\mathscr{O}(C_1)$ at C_1 is key to get the first order expression for $a_k(C_1+E)$. More precisely: using this transversality (namely, equation (5.3) with E^{syl} being the Sylvester space for C_1), together with the identity $\text{tr}(p_{n-k}(C_1)E^{\text{tan}}) = 0$, and the explicit expression $-E^{\text{syl}}_{k-1} = \text{tr}(p_{n-k}(C_1)E)$, both them valid for $k = 1,\ldots,n$, they get an explicit expression for $\text{tr}(p_{n-k}(C_1)E)$, which is the first order term of $a_{k-1}(C_1+E)$. This can be done because the matrices $p_{n-k}(C_1)$, for $k = 1,\ldots,n$, have a simple structure that allows to compute $\text{tr}(p_{n-k}(C_1)E^{\text{syl}})$ easily and explicitly, for all $k = 1,\ldots,n$. Unfortunately, for arbitrary Fiedler matrices, to get explicit expressions of $\text{tr}(p_{n-k}(M_\sigma)E)$ by hand is quite involved. Hence, we have obtained the first-order term of $a_k(M_\sigma+E)$ directly from $\text{adj}(zI-M_\sigma)$. This approach is completely independent of the transversality of E^{syl} and the tangent space, though, as we have seen in Theorem 5.3, this fact is still true for arbitrary Fiedler matrices.

6. Conclusions

In this paper, we have analyzed some numerical features of the polynomial root-finding problem when considered as a standard eigenvalue problem by means of Fiedler companion matrices. In particular, we have described the first-order change of the characteristic polynomial of any Fiedler matrix under small perturbations of the matrix. This description has led us to conclude that polynomial root-finding algorithms based on backward stable eigenvalue algorithms using Fiedler companion matrices, are backward stable only if $\|p\|_\infty$ is moderate. More precisely, given a monic polynomial $p(z)$, if $\widetilde{p}(z)$ denotes the monic polynomial whose roots are the computed eigenvalues of a Fiedler companion matrix of $p(z)$, obtained with a backward stable eigenvalue algorithm, then it is not possible to guarantee, in general, that

$$\frac{\|\widetilde{p}-p\|_\infty}{\|p\|_\infty} = O(u),$$

where u is the machine epsilon of the computer. Namely, the computed roots of $p(z)$ are not necessarily the roots of a nearby polynomial. We have seen, however, that

$$\frac{\|\widetilde{p}-p\|_\infty}{\|p\|_\infty} = O(u)\|p\|_\infty^2,$$

for any Fiedler companion matrix other than the first and second Frobenius companion matrices, and that

$$\frac{\|\widetilde{p}-p\|_\infty}{\|p\|_\infty} = O(u)\|p\|_\infty,$$

for the first and second Frobenius companion matrices (which are particular cases of Fiedler matrices). Extensive numerical experiments have been included to illustrate these theoretical results.

Acknowledgements. This work has been supported by the Ministerio de Economía y Competitividad of Spain through grant MTM2012-32542. The authors wish to thank Yuji Nakatsukasa and Vanni Noferini for reading a draft of this manuscript, and for their insightful comments, which helped us to improve the paper.

REFERENCES

ARNOLD, V. I. (1971) Matrices depending on parameters. *Uspehi Mat. Nauk*, **26**, 101–114. English translation: Russian Math. Surveys 26 (1971), no. 2, 29–43.

AURENTZ, J. L., VANDEBRIL, R. & WATKINS, D. S. (2013) Fast computation of the zeros of a polynomial via factorization of the companion matrix. *SIAM J. Sci. Comput.*, **35**, A255–A269.

BERNSTEIN, D. S. (2009) *Matrix mathematics. Theory, facts, and formulas*, 2nd edn. Princeton University Press, Princeton, NJ.

BHATIA, R. & JAIN, T. (2009) Higher order derivatives and perturbation bounds for determinants. *Linear Algebra Appl.*, **431**, 2102–2108.

BINI, D. A., GEMIGNANI, L. & PAN, V. Y. (2004) Improved initialization of the accelerated and robust QR-like polynomial root-finding. *Electron. Trans. Numer. Anal.*, **17**, 195–205.

BINI, D. A., GEMIGNANI, L. & PAN, V. Y. (2005) Fast and stable QR eigenvalue algorithms for generalized companion matrices and secular equations. *Numer. Math.*, **100**, 373–408.

BINI, D. A., BOITO, P., EIDELMAN, Y., GEMIGNANI, L. & GOHBERG, I. (2010) A fast implicit QR eigenvalue algorithm for companion matrices. *Linear Algebra Appl.*, **432**, 2006–2031.

BRUGNANO, L. & TRIGIANTE, D. (1995) Polynomial roots: the ultimate answer? *Linear Algebra Appl.*, **225**, 207–219.

CALVETTI, D., KIM, S.-M. & REICHEL, L. (2002) The restarted *QR*-algorithm for eigenvalue computation of structured matrices. *J. Comput. Appl. Math.*, **149**, 415–422.

CHANDRASEKARAN, S., GU, M., XIA, J. & ZHU, J. (2008) A fast QR algorithm for companion matrices. *Recent advances in matrix and operator theory*. Oper. Theory Adv. Appl., vol. 179. Birkhäuser, Basel, pp. 111–143.

DE TERÁN, F., DOPICO, F. M. & MACKEY, D. S. (2010) Fiedler companion linearizations and the recovery of minimal indices. *SIAM J. Matrix Anal. Appl.*, **31**, 2181–2204.

DE TERÁN, F., DOPICO, F. M. & PÉREZ, J. (2013) Condition numbers for inversion of Fiedler companion matrices. *Linear Algebra Appl.*, **439**, 944–981.

DE TERÁN, F., DOPICO, F. M. & PÉREZ, J. (2014) New bounds for roots of polynomials based on Fiedler companion matrices. *Linear Algebra Appl.*, **To appear**.

EDELMAN, A., ELMROTH, E. & KÅGSTRÖM, B. (1997) A geometric approach to perturbation theory of matrices and matrix pencils. I. Versal deformations. *SIAM J. Matrix Anal. Appl.*, **18**, 653–692.

EDELMAN, A., ELMROTH, E. & KÅGSTRÖM, B. (1999) A geometric approach to perturbation theory of matrices and matrix pencils. II. A stratification-enhanced staircase algorithm. *SIAM J. Matrix Anal. Appl.*, **20**, 667–699 (electronic).

EDELMAN, A. & MURAKAMI, H. (1995) Polynomial roots from companion matrix eigenvalues. *Math. Comp.*, **64**, 763–776.

FIEDLER, M. (2003) A note on companion matrices. *Linear Algebra Appl.*, **372**, 325–331.

GANTMACHER, F. R. (1959) *The theory of matrices. Vols. 1, 2*. Chelsea Publishing Co., New York.

GEMIGNANI, L. (2007) Structured matrix methods for polynomial root-finding. *ISSAC 2007*. ACM, New York, pp. 175–180.

GOLUB, G. H. & VAN LOAN, C. F. (1996) *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences, 3rd edn. Johns Hopkins University Press, Baltimore, MD.

GRAUERT, H. & FRITZSCHE, K. (1976) *Several complex variables*. Springer-Verlag, New York-Heidelberg. Translated from the German, Graduate Texts in Mathematics, Vol. 38.

HIGHAM, N. J. (2002) *Accuracy and stability of numerical algorithms*, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

HORN, R. A. & JOHNSON, C. R. (1985) *Matrix analysis*. Cambridge University Press, Cambridge.

IPSEN, I. C. F. & REHMAN, R. (2008) Perturbation bounds for determinants and characteristic polynomials. *SIAM J. Matrix Anal. Appl.*, **30**, 762–776.

LEMMONIER, D. & VAN DOOREN, P. (2003) Optimal scaling of companion pencils for the *qz*-algorithm. *Proceedings SIAM Appl. Lin. Alg. Conference*, **Paper CP7–4**.

MOLER, C. (1991) Cleve's corner: Roots–of polynomials, that is. *The Mathworks Newsletter*, **5**, 8–9.

NIU, X.-M. & SAKURAI, T. (2003) A method for finding the zeros of polynomials using a companion matrix. *Japan J. Indust. Appl. Math.*, **20**, 239–256.

PARLETT, B. N. & REINSCH, C. (1969) Balancing a matrix for calculation of eigenvalues and eigenvectors. *Numer. Math.*, **13**, 293–304.

TISSEUR, F. (2000) Backward error and condition of polynomial eigenvalue problems. *Linear Algebra Appl.*, **309**, 339–361.

TOH, K.-C. & TREFETHEN, L. N. (1994) Pseudozeros of polynomials and pseudospectra of companion matrices. *Numer. Math.*, **68**, 403–425.

VAN BAREL, M., VANDEBRIL, R., VAN DOOREN, P. & FREDERIX, K. (2010) Implicit double shift *QR*-algorithm for companion matrices. *Numer. Math.*, **116**, 177–212.

VAN DOOREN, P. & DEWILDE, P. (1983) The eigenstructure of an arbitrary polynomial matrix: computational aspects. *Linear Algebra Appl.*, **50**, 545–579.

ZHLOBICH, P. (2012) Differential qd algorithm with shifts for rank-structured matrices. *SIAM J. Matrix Anal. Appl.*, **33**, 1153–1171.