Numer. Math. 59, 349–360 (1991)

Numerische
Mathematik

Inclusion of the roots of a polynomial based on Gerschgorin's theorem*

C. Carstensen
Institut für Angewandte Mathematik, Universität Hannover, Welfengarten 1, W-3000 Hannover 1, Federal Republic of Germany

Received July 13, 1989

Summary. In this note a new companion matrix is presented which can be interpreted as a product of Werner's companion matrices [13]. Gerschgorin's theorem yields an inclusion of the roots of a polynomial which is best in the sense of [4] and generalizes a result of L. Elsner [5]. This inclusion is better than the one due to W. Börsch–Supan in [1].

Mathematics Subject Classification: 65H05, 12D10

1 Introduction

Throughout this note a matrix $B \in \mathbb{C}^{n \times n}$ is called a companion matrix of a monic polynomial f iff the characteristic polynomial χ_B of B is equal to $(-1)^n f$. The most common example is the well-known Frobenius companion matrix of f-often simply called companion matrix- which has various applications in numerical analysis. For instance, Gerschgorin's theorem can be used to bound the roots of f; see e.g. [11]. The main disadvantage of these bounds is that the centre of these Gerschgorin discs cannot be chosen arbitrarily. This is possible to a certain extent using Werner's companion matrix [13]. As in the case of the Frobenius's companion matrix the problem remains that the superdiagonal elements cannot vanish so that the Gerschgorin discs cannot get arbitrarily small without a suitable similarity transformation. L. Elsner considered another companion matrix in [5] which becomes a diagonal matrix if the roots of f are chosen as diagonal elements. By continuity, the use of good approximations to the roots for the diagonal elements leads to small Gerschgorin discs and error estimates are easily obtained. But this works only if the roots are distinct.

In Sect. 3 of this note a new companion matrix is presented which generalizes both the companion matrices from [5] and [13]. With its flexibility multiple roots can be approximated while the superdiagonal elements may vanish if distinct roots are approximated.

* Dedicated to Professor E. Stein on the occasion of his 60th birthday.

In Sect. 4 an algorithm is formulated which computes the new companion matrix of a given polynomial f using only particular lower order divided differences of f.

Before applying Gerschgorin's theorem to a companion matrix B one may transform B into $T^{-1}BT$ with a regular matrix T. In the simplest case T is a diagonal matrix. Of course the diagonal matrix T should be chosen such that the Gerschgorin discs become as small as possible. The optimal entries of T can be characterised as an eigenvector of a certain eigenvalue problem; see [4, 9, and 10] and the literature given there.

In Sect. 5 this optimal transformation matrix T is given explicitly and Gerschgorin's theorem yields an inclusion for the roots of f. The proof is rather elementary because one has only to compute $T^{-1}BT$ and add some absolute values of non-diagonal elements. The inclusion generalizes a result from L. Elsner [5].

In Sect. 6 a particular case of the inclusion is compared with a well-known error estimate due to W. Börsch–Supan which is proved analytically in [1]. It is shown that the new inclusion is slightly better in both flexibility and quality. This is illustrated by two examples at the end of this note.

Of course Gerschgorin's theorem is not the only way to get inclusions for the eigenvalues of a matrix B; see [11, 6, and 7] for instance. But in this way very sharp estimates are obtained by using "best Gerschgorin discs".

2 Notation

Let $\mathbb{N}$ or $\mathbb{N}_0$ denote the set of the positive or the non-negative integers, respectively. I_n denotes the n-dimensional unit matrix in $\mathbb{C}^{n\times n}$. For $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$, let $\operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ or $\operatorname{bidiag}(\alpha_1, \ldots, \alpha_n)$ be the matrix in $\mathbb{C}^{n\times n}$ having the diagonal elements $\alpha_1, \ldots, \alpha_n$ and the $(n-1)$ superdiagonal elements $0, \ldots, 0$ or $1, \ldots, 1$, respectively, so that all the other entries vanish. Similarly, $\operatorname{diag}(A_1, \ldots, A_m) \in \mathbb{C}^{n\times n}$ is referred to as the blockdiagonal matrix having the m-matrices $A_1 \in \mathbb{C}^{k_1\times k_1}, \ldots, A_m \in \mathbb{C}^{k_m\times k_m}$ as diagonal blocks, $n := \sum_{i=1}^m k_i$.

$\mathbb{P}_n$ denotes the set of all polynomials in $\mathbb{C}$ of degree less than or equal to n, while $\mathbb{P}_n^{\text{monic}}$ is the subset of all monic (i.e., the leading coefficient is 1) polynomials of degree n. By $[x_1, \ldots, x_k]f$ we denote the divided difference of $f\colon \mathbb{C} \to \mathbb{C}$, f assumed sufficiently differentiable, with respect to the knots $x_1, \ldots, x_k \in \mathbb{C}$. Let $\Pi_{(x_1, \ldots, x_k)} \in \mathbb{P}_k^{\text{monic}}$ be the polynomial defined by $\Pi_{(x_1, \ldots, x_k)}(x) := \prod_{i=1}^k (x - x_i)$, $x \in \mathbb{C}$.

Let $m \in \mathbb{N}$, $k_1, \ldots, k_m \in \mathbb{N}$, $n := \sum_{i=1}^m k_i$. For fixed $\alpha_1^1, \ldots \alpha_1^{k_1}, \ldots, \alpha_m^1, \ldots \alpha_m^{k_m} \in \mathbb{C}$ and all $i \in \{1, \ldots, m\}$ define the polynomials Π_α, $\Pi_{\alpha_i^c}$ as

$$\Pi_\alpha := \prod_{\nu=1}^m \Pi_{(\alpha_\nu^1, \ldots, \alpha_\nu^{k_\nu})}, \qquad \Pi_{\alpha_i^c} := \prod_{\substack{\nu=1\\ \nu \neq i}}^m \Pi_{(\alpha_\nu^1, \ldots, \alpha_\nu^{k_\nu})},$$

which coincides with the definition above taking

$$\alpha := (\alpha_1^1, \ldots, \alpha_m^{k_m}) \quad \text{and} \quad \alpha_i^c := (\alpha_1^1, \ldots, \alpha_{i-1}^{k_{i-1}}, \alpha_{i-1}^1, \ldots, \alpha_m^{k_m})\,.$$

3 A new companion matrix

We start with some useful lemmas. The proof of the first lemma can be found in [8].

Lemma 3.1. *Let $A \in \mathbb{C}^{n \times n}$ and $a, b \in \mathbb{C}^n$. Let $x \in \mathbb{C}$ be no eigenvalue of A. Then*

$$\det(A - ab^{\mathrm{T}} - xI_n) = (1 - b^{\mathrm{T}}(A - xI_n)^{-1}a) \cdot \det(A - xI_n)\,. \tag{1}$$

The simple proof of the next lemma will be omitted.

Lemma 3.2. *Let $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ and $x \in \mathbb{C} \backslash \{\alpha_1, \ldots, \alpha_n\}$. Then*

$$\operatorname{bidiag}(\alpha_1 - x, \ldots, \alpha_n - x)^{-1}_{i,j} = \begin{cases} 0 & \text{if } i > j \\ \dfrac{(-1)^{i+j}}{\prod_{\nu=i}^{j} (\alpha_\nu - x)} & \text{if } i \leqq j \end{cases}. \tag{2}$$

Theorem 3.3. *For all $i \in \{1, \ldots, m\}$, let $a_i = (a_i^1, \ldots, a_i^{k_i})^{\mathrm{T}}$, $b_i = (b_i^1, \ldots, b_i^{k_i})^{\mathrm{T}} \in \mathbb{C}^{k_i}$, and set*

$$A := \operatorname{diag}(\operatorname{bidiag}(\alpha_1^1, \ldots, \alpha_1^{k_1}), \ldots, \operatorname{bidiag}(\alpha_m^1, \ldots, \alpha_m^{k_m})) \in \mathbb{C}^{n \times n},$$
$$a := (a_1^{\mathrm{T}}, \ldots, a_m^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{C}^n, \qquad b := (b_1^{\mathrm{T}}, \ldots, b_m^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{C}^n\,.$$

Then

$$(-1)^n \chi_{A - ab^{\mathrm{T}}} = \Pi_\alpha + \sum_{i=1}^{m} \sum_{1 \leqq \nu \leqq \mu \leqq k_i} a_i^\mu b_i^\nu \, \Pi_{(\alpha_i^1, \ldots, \alpha_i^{\nu-1}, \alpha_i^{\mu+1}, \ldots, \alpha_i^{k_i})} \Pi_{\alpha_i^c}\,. \tag{3}$$

Proof. Since (3) is an identity between polynomials it suffices to prove it for $x \in \mathbb{C} \backslash \{\alpha_1^1, \ldots, \alpha_1^{k_1}, \ldots, \alpha_m^1, \ldots, \alpha_m^{k_m}\}$. Then, by Lemmas 3.1 and 3.2,

$$\begin{aligned}
&(-1)^n \chi_{A - ab^{\mathrm{T}}}(x) \\
&= \left(1 - \sum_{i=1}^{m} b_i^{\mathrm{T}} \operatorname{bidiag}(\alpha_i^1 - x, \ldots, \alpha_i^{k_i} - x)^{-1} a_i\right) \prod_{i=1}^{m} \prod_{j=1}^{k_i} (x - \alpha_i^j) \\
&= \left(1 - \sum_{i=1}^{m} \sum_{1 \leqq \nu \leqq \mu \leqq k_i} b_i^\nu \frac{(-1)^{\nu+\mu}}{\prod_{\kappa=\nu}^{\mu} (\alpha_i^\kappa - x)} a_i^\mu \right) \prod_{i=1}^{m} \prod_{j=1}^{k_i} (x - \alpha_i^j)\,.
\end{aligned}$$

This implies (3). □

Corollary 3.4. *If $a_i = (0, \ldots, 0, 1)^{\mathrm{T}} \in \mathbb{C}^{k_i}$ for all $i \in \{1, \ldots, m\}$ Theorem 3.3 reduces to*

$$(-1)^n \chi_{A - ab^{\mathrm{T}}} = \Pi_\alpha + \sum_{i=1}^{m} \sum_{1 \leqq \nu \leqq k_i} b_i^\nu \, \Pi_{(\alpha_i^1, \ldots, \alpha_i^{\nu-1})} \Pi_{\alpha_i^c}\,. \tag{4}$$

Remark 3.5. Remarks on particular cases.

1. Theorem 3.3 was presented for the first time in [2] with the particular case $\alpha_i^1 = \ldots = \alpha_i^{k_i}$ for all $i \in \{1, \ldots, m\}$.

2. For $m = 1$, $n = k_1$ and $0 = \alpha_1^1 = \ldots = \alpha_1^n$ the matrix $A - ab^{\mathrm{T}}$ in Corollary 3.4 is well-known as Frobenius's companion matrix. In the case of arbitrary $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ the matrix is presented by W. Werner in [13].

3. For $m = n$, $k_1 = \ldots = k_n = 1$ and $a_1 = \ldots = a_n = 1$ the matrix $A - ab^{\mathrm{T}}$ in Corollary 3.4 can be found in [12, 5, 9, 14, 2].

4. In [12] another companion matrix is presented, which is also a particular case of Theorem 3.3; for a proof see [3].

5. Theorem 3.3 can be generalized to a product rule for companion matrices; see [3].

4 Computation of the companion matrix

Given a matrix A and a vector a as in Theorem 3.3 and a monic polynomial f of degree n, the following algorithm computes a vector b such that $A - ab^{\mathrm{T}}$ is a companion matrix of f.

Algorithm 4.1

Input: α_i^j and $[\alpha_i^1, \ldots, \alpha_i^j]f$ for all $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k_i\}$.

$$i = 1, \ldots, m$$

$$j = 1, \ldots, k_i$$

$$b_i^j := \frac{[\alpha_i^i, \ldots, \alpha_i^j]f - \sum_{k=1}^{j-1} b_i^k \left\{ \sum_{\kappa=k}^{k_i} a_i^\kappa \, [\alpha_i^k, \ldots, \alpha_i^j](\Pi_{(\alpha_i^{\kappa+1}, \ldots, \alpha_i^{k_i})} \Pi_{\alpha_i^c}) \right\}}{\Pi_{\alpha_i^c}(\alpha_i^j) \sum_{\kappa=j}^{k_i} \alpha_i^\kappa \, \Pi_{(\alpha_i^{\kappa+1}, \ldots, \alpha_i^{k_i})}(\alpha_i^j)}$$

Output: b_i^j for all $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k_i\}$.

Theorem 4.2. *If $f \in \mathbb{P}_n^{\text{monic}}$ and if none of the denominators in Algorithm* 4.1 *vanish then $A - ab^{\mathrm{T}}$ is a companion matrix of f.*

Proof. Let $g := (-1)^n \chi_{A - ab^{\mathrm{T}}}$, where b is defined by the output of Algorithm 4.1. The condition on the denominators implies for all $i, j \in \{1, \ldots, m\}$

$$i \neq j \Rightarrow \{\alpha_i^1, \ldots, \alpha_i^{k_i}\} \cap \{\alpha_j^1, \ldots, \alpha_j^{k_j}\} = \emptyset \; . \tag{5}$$

$g, f \in \mathbb{P}_n^{\text{monic}}$ leads to $r := f - g \in \mathbb{P}_{n-1}$. Therefore and because of (5), $r = 0$ if and only if for all $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k_i\}$ $[\alpha_i^1, \ldots, \alpha_i^j]\, r = 0$ (in both cases r has n zeros $\alpha_1^1, \ldots, \alpha_1^{k_1}, \ldots, \alpha_m^1, \ldots, \alpha_m^{k_m}$).

For $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k_i\}$ Theorem 3.3 shows

$$\begin{aligned}[\alpha_i^1, \ldots, \alpha_i^j]g &= \sum_{1 \leqq \nu \leqq \mu \leqq k_i} a_i^\mu b_i^\nu [\alpha_i^1, \ldots, \alpha_i^j] \, (\Pi_{(\alpha_i^1, \ldots, \alpha_i^{\nu-1}, \alpha_i^{\mu+1}, \ldots, \alpha_i^{k_i})} \Pi_{\alpha_i^c}) \\ &= \sum_{\nu=1}^{j} \sum_{\mu=\nu}^{k_i} a_i^\mu b_i^\nu [\alpha_i^\nu, \ldots, \alpha_i^j] \, (\Pi_{(\alpha_i^{\mu+1}, \ldots, \alpha_i^{k_i})} \Pi_{\alpha_i^c}) \; ,\end{aligned}$$

by Leibniz's rule for divided differences. On the other hand a simple calculation with the rule in Algorithm 4.1 shows that the last expression is equal to $[\alpha_i^1, \ldots, \alpha_i^j]f$. □

Remark 4.3

1. In general the condition on the denominators in Algorithm 4.1 cannot be dropped. Indeed, if some denominator vanishes then there exists a monic polynomial f which has no companion matrix $A - ab^T$ for any $b \in \mathbb{C}^n$.

Proof. Let $i \in \{1, \ldots, m\}, j \in \{1, \ldots, k_i\}$ such that

$$(6) \qquad 0 = \Pi_{\alpha_i^c}(\alpha_i^j) \sum_{\mu=j}^{k_i} a_i^\mu \, \Pi_{(\alpha_i^{\mu+1}, \ldots, \alpha_i^{k_i})}(\alpha_i^j) \,.$$

For any $b \in \mathbb{C}^n$, by (3),

$$(-1)^n \cdot \chi_{A-ab^t}(\alpha_i^j) = \sum_{1 \leqq \nu \leqq k_i} b_i^\nu \, \Pi_{(\alpha_i^1, \ldots, \alpha_i^{\nu-1})}(\alpha_i^j)$$
$$\times \sum_{\nu \leqq \mu \leqq k_i} a_i^\mu \, \Pi_{(\alpha_i^{\mu+1}, \ldots, \alpha_i^{k_i})}(\alpha_i^j) \, \Pi_{\alpha_i^c}(\alpha_i^j) \,.$$

Clearly, the indices in the first or second sum can be restricted to $1 \leqq \nu \leqq j$ or $j \leqq \mu \leqq k_i$, respectively. Therefore (6) implies that the second sum vanishes. Thus $\chi_{A-ab^T}(\alpha_i^j) = 0$, so we cannot get a companion matrix for any f not having the root α_i^j. □

2. In the particular case of Corollary 3.4 Algorithm 4.1 reduces to

$$\left[\begin{array}{l} i = 1, \ldots, m \\ \left[\begin{array}{l} j = 1, \ldots, k_i \\ b_i^j := \dfrac{[\alpha_i^i, \ldots, \alpha_i^j]f - \sum_{k=1}^{j-1} b_i^k \, [\alpha_i^k, \ldots, \alpha_i^j] \, \Pi_{\alpha_i^c}}{\Pi_{\alpha_i^c}(\alpha_i^j)} \end{array}\right. \end{array}\right.$$

Hence the only restriction is (5).

3. In Algorithm 4.1 several divided differences of polynomials with known zeros $\alpha_1^1, \ldots, \alpha_1^{k_1}, \ldots, \alpha_m^1, \ldots, \alpha_m^{k_m}$ are needed. These can be easily computed using the well-known general Horner Algorithm.

5 Inclusion of zeros of polynomials

In this section we will apply Gerschgorin's theorem to the companion matrix of Sect. 3.

Let $f \in \mathbb{P}_n^{\text{monic}}$ be a polynomial and $\alpha_1^1, \ldots, \alpha_1^{k_1}, \ldots, \alpha_m^1, \ldots, \alpha_m^{k_m}$ n approximations of the n zeros of f. They have to be indexed so that (5) is fulfilled. Our aim is to give simultaneous error estimates for approximations with the same lower index. Therefore we should index the n approximations so that the members of

$\{\alpha_i^1, \ldots, \alpha_i^{k_i}\}$ are close to each other while the members of any two different sets $\{\alpha_i^1, \ldots, \alpha_i^{k_i}\}$ and $\{\alpha_j^1, \ldots, \alpha_j^{k_j}\}$ $(i \neq j)$ are quite different. We assume $m \geqq 2$.

Next we compute the divided differences $[\alpha_i^1, \ldots, \alpha_i^j]\, f$ for all $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k_i\}$. Then we can compute a companion matrix $A - ab^{\mathrm{T}}$ of f using Algorithm 4.1. This yields a representation (3) of f.

For simplification and because of its efficiency we only consider the particular case of Corollary 3.4 and the representation (4) of f; see [3] for the general case.

For the next theorem we need some further abbreviations. Let $I \subseteqq \{1, \ldots, m\}$ and $I^{\mathrm{C}} := \{1, \ldots, m\} | I$ be fixed and non-empty and let $\zeta \in \mathbb{C}$ be the centre of in- and exclusion-circles. Define

$$\underline{d} := \max_{i \in I} \max \{|\alpha_i^1 - \zeta|, \ldots |\alpha_i^{k_i - 1} - \zeta|, |\alpha_i^{k_i} - \zeta - b_i^{k_i}| - |b_i^{k_i}|\}$$

$$\bar{d} := \min_{j \in I^{\mathrm{C}}} \min \{|\alpha_j^1 - \zeta|, \ldots |\alpha_j^{k_j - 1} - \zeta|, |\alpha_j^{k_j} - \zeta - b_j^{k_j}| + |b_j^{k_j}|\} .$$

If $\underline{d} < \bar{d}$ we consider the function $h\colon]\underline{d}, \bar{d}[\to \mathbb{R}$, defined for $r \in]\underline{d}, \bar{d}[$ by

$$h(r) := \sum_{i \in I} \sum_{k=1}^{k_i} \frac{|b_i^k|}{[r - |\alpha_i^{k_i} - \zeta - b_i^{k_i}| + |b_i^{k_i}|] \sum_{\mu = k}^{k_i - 1} [r - |\alpha_i^{\mu} - \zeta|]} + \sum_{j \in I^{\mathrm{C}}} \sum_{k=1}^{k_j} \frac{|b_j^k|}{[|\alpha_j^{k_j} - \zeta - b_j^{k_j}| + |b_j^{k_j}| - r] \prod_{\mu = k}^{k_j - 1} [|\alpha_j^{\mu} - \zeta| - r]} . \tag{7}$$

Theorem 5.1. *Let f be a polynomial equal to (4). If there exists a $r \in]\underline{d}, \bar{d}[$ with $h(r) \leqq 1$, then in the closed disc with center ζ and radius r lie at least $\sum_{i \in I} k_i$ zeros while outside the open disc with center ζ and radius r lie at least $\sum_{j \in I^{\mathrm{C}}} k_j$ zeros of f.*

Proof. Without loss of generality we can assume $\zeta = 0$ and $I = \{1, \ldots, p\}$ with $p \in \{1, \ldots, n-1\}$. Define

$$\gamma_i^j := \begin{cases} \dfrac{1}{[r - |\alpha_i^{k_i} - b_i^{k_i}| + |b_i^{k_i}|] \prod_{\kappa = j}^{k_i - 1} [r - |\alpha_i^{\kappa}|]}, & \text{if } i \leqq p \\[2ex] \dfrac{1}{[|\alpha_i^{k_i} - \mathrm{b}_i^{k_i}| + |b_i^{k_i}| - r] \prod_{\kappa = j}^{k_i - 1} [|\alpha_i^{\kappa}| - r]}, & \text{if } i > p \end{cases} \tag{8}$$

for any $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k_i\}$. Since $\underline{d} < r < \bar{d}$ these coefficients are positive so that

$$T := \operatorname{diag}(\gamma_1^1, \ldots, \gamma_1^{k_1}, \ldots, \gamma_m^1, \ldots, \gamma_m^{k_m})$$

is regular. We apply the well-known Gerschgorin theorem (see, for example, [7]) on $T^{-1}(A - ab^{\mathrm{T}})\, T$.

For any $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k_i - 1\}$ the centre of the Gerschgorin disc belonging to an index (i, j) is α_i^j and the radius is equal to

$$\frac{\gamma_i^{j+1}}{\gamma_i^j} = \begin{cases} r - |\alpha_i^j| , & \text{if } i \leqq p \\ |\alpha_i^j| - r , & \text{if } i > p \end{cases} .$$

The center of the Gerschgorin disc subordinated to an index (i, k_i) is $\alpha_i^{k_i} - b_i^{k_i}$ and by (7) the radius is equal to

$$\frac{1}{\gamma_i^{k_i}}\left(\sum_{j=1}^{m}\sum_{k=1}^{k_j}|b_j^k|\gamma_j^k\right) - |b_i^{k_i}| = \frac{h(r)}{\gamma_i^{k_i}} - |b_i^{k_i}| .$$

Using $h(r) \leqq 1$ the last radius is less or equal $r - |\alpha_i^{k_i} - b_i^{k_i}|$ or $|\alpha_i^{k_i} - b_i^{k_i}| - r$ if $i \leqq p$ or $i > p$, respectively.

Thus any Gerschgorin disc belonging to an index (i, j) lies in the closed disc with center $\zeta = 0$ and radius r while any other Gerschgorin disc lies outside this open disc. Therefore Gerschgorin's theorem proves the assertion. □

Remark 5.2

1. Let $A - ab^{\mathrm{T}}$ be the companion matrix of f from Corollary 3.4. If $A - ab^{\mathrm{T}}$ is irreducible then h, defined in (7), is a smooth strict convex function with $\lim_{r\to \underline{d}+0} h(r) = \infty = \lim_{r-\bar{d}-0} h(r)$. If $\min h[\,]\underline{d}, \bar{d}[\,] < 1$, then there are exactly two roots of $h(r) = 1$, denoted by $\underline{r}$ and $\bar{r}$; $\underline{d} < \underline{r} < \bar{r} < \bar{d}$.

Then Theorem 5.1 states that there are exactly $\sum_{i\in I} k_i$ zeros of f in the closed disc with center ζ and radius $\underline{r}$ while the other zeros of f lie outside the open disc with centre ζ and radius $\bar{r}$.

2. If $h(r) < 1$ then upper or lower bounds for $\underline{r}$ or $\bar{r}$, respectively, can be computed as the roots of $h(r) = 1$ using the regular falsi for instance, because the convexity of h yields monotonous convergence.

In general, the computation of (b_i^j) using complex arithmetic is more expensive than an easy application of the regular falsi in real arithmetic. Therefore a short iteration of the regular falsi for h could be efficient in view of the reduction of the inclusion discs. Also, this is correct if we want exact inclusions and use a suitable arithmetic where round-off errors should be taken into consideration.

3. The error estimate of Theorem 5.1 is best in the sense of minimal Gerschgorin discs [4]. To see this use the notation of the proof of Theorem 5.1 for $r \in \{\underline{r}, \bar{r}\}$ and define $H \in \mathbb{R}^{n\times n}$ by

$$H_{\nu,\mu} := \begin{cases} -|(A - ab^{\mathrm{T}})_{\nu,\mu}| & \text{if } \nu = \mu > \sum_{\kappa=1}^{p} k_\kappa \\ |(A - ab^{\mathrm{T}})_{\nu,\mu}| & \text{otherwise} \end{cases},$$

$$E_p := \operatorname{diag}(-I_{k_1}, \ldots, -I_{k_p}, I_{k_{p+1}}, \ldots, I_{k_m}) \in \mathbb{R}^{n\times n} .$$

Then by easy calculations using $h(r) = 1$ as in the proof of Theorem 5.1

$$(H - rE_p)(\gamma_1^1, \ldots, \gamma_1^{k_1}, \ldots, \gamma_m^1, \ldots, \gamma_m^{k_m})^{\mathrm{T}} = 0 \in \mathbb{R}^n .$$

Therefore $\underline{r}$ and $\bar{r}$ are eigenvalues of $E_p^{-1} H$ with a positive eigenvector. Satz 4 from [4] states that any choice of a diagonal matrix T will produce a weaker estimate.

4. Theorem 5.1 generalizes a result of L. Elsner for the particular case $m = n$ and $1 = k_1 = \ldots = k_m$ in [5].

5. A direct application of Gerschgorin's theorem to $T^{-1}(A - ab^{\mathrm{T}})T$ (T defined in the proof of Theorem 5.1) will give slightly better inclusion regions. Other inclusion theorems like Bauer's Theorem on the ovales of Cassini (see, for example [7]) can also be used to estimate the eigenvalues of $T^{-1}(A - ab^{\mathrm{T}})T$.

6. Obviously, to obtain good estimates we need good approximations. Therefore Theorem 5.1 is especially suitable for a posteriori error estimates which are often inevitable: for instance, if we compute the zeros of a given polynomial f using deflations particularly in the presence of cluster of zeros.

7. Before we can apply any locally convergent complex interval arithmetic to find the roots of f we need initial discs (or rectangles etc.) containing the zeros of f. Again, this inclusions may be achieved using the inclusions proposed in Theorem 5.1. Because of the great computational costs of most interval methods it could be efficient to obtain and reduce the initial discs using Theorem 5.1 in spite of additional expenditure.

6 A particular case

In this section we will compare the inclusion of Theorem 5.1 with a well-known analytic result due to W. Börsch-Supan in [1].

We use the notation of the last section, only considering the case

$$I = \{1\} \quad \text{and} \quad \alpha_i := \alpha_i^1 = \ldots = \alpha_i^{k_i} \qquad \text{for all } i \in \{1, \ldots, m\} .$$

To avoid uninteresting and trivial particular cases let us assume that for all $i \in \{1, \ldots, m\}\, b_i^1 \neq 0$; otherwise $A - ab^{\mathrm{T}}$ will be reducible (the vector a is still chosen as in Corollary 3.4).

Let $\zeta = \alpha_1$, let $R > 0$ be the (unique) positive root of

$$1 = \sum_{j=1}^{k_1} \frac{|b_1^j|}{R^{k_1 - j + 1}} \tag{9}$$

and define

$$s := \sum_{i=2}^{m} \sum_{j=1}^{k_i} \frac{|b_i^j|}{|\alpha_1 - \alpha_i|^{k_i - j + 1}}, \qquad d := R \max\left\{ \frac{k_j}{|\alpha_1 - \alpha_j|} \,\middle|\, j \in \{2, \ldots, m\} \right\}.$$

The following lemma states that the separating condition in Satz 2 in [1] implies the separation condition of Theorem 5.1.

Lemma 6.1 *If*

$$\sqrt{s} + \sqrt{d} < 1 ,$$

then, for

$$\underline{R} := R \cdot \left\{ 1 + \frac{2s}{1 - s - d + \sqrt{[1 - s - d]^2 - 4ds}} \right\}$$

$$\bar{R} := R \cdot \left\{ 1 + \frac{2s}{1 - s - d - \sqrt{[1 - s - d]^2 - 4ds}} \right\},$$

$$0 = \underline{d} < \underline{R} < \bar{R} < \bar{d} \qquad \textit{and} \qquad h(\underline{R}) \leqq 1, \qquad h(\bar{R}) \leqq 1 .$$

Proof. Define $\underline{\kappa}, \bar{\kappa} > 0$ by $R(1 + \underline{\kappa}) = \underline{R}$ and $R(1 + \bar{\kappa}) = \bar{R}$. Then

$$\underline{\kappa} + \bar{\kappa} = \frac{1 - s - d}{d} \quad \text{and} \quad \underline{\kappa}\bar{\kappa} = \frac{s}{d}$$

and for any $\kappa \in [\underline{\kappa}, \bar{\kappa}]$

$$0 \geqq (\kappa - \underline{\kappa})(\kappa - \bar{\kappa}) = \kappa^2 - \frac{\kappa}{d}(1 - s - d) + \frac{s}{d}.$$

Multiplying the last inequality with $d/\kappa > 0$ leads to

$$d(1 + \kappa) + s\frac{\kappa + 1}{\kappa} \leqq 1. \tag{10}$$

Since $0 < s$ and $0 < \kappa$,

$$d(1 + \kappa) < 1.$$

The definition of d implies

$$\frac{R}{\min\{|\alpha_1 - \alpha_j| \mid j \in \{2, \ldots, m\}\}} \leqq d,$$

so that

$$R(1 + \kappa) < \min\{|\alpha_1 - \alpha_j| \mid j \in \{2, \ldots, m\}\} \leqq \bar{d}. \tag{11}$$

Therefore, $0 = d \leqq R < \underline{R} < \bar{R} < \bar{d}$.

For $\kappa \in [\underline{k}, \bar{\kappa}]$ we consider

$$h(R(1 + \kappa)) \leqq \sum_{j=1}^{k_1} |b_1^j|(R(1 + \kappa))^{j - k_1 - 1} \tag{12}$$
$$+ \sum_{i=2}^{m} \sum_{j=1}^{k_i} \frac{|b_i^j|}{[|\alpha_1 - \alpha_i| - (R(1 + \kappa))]^{k_i + 1 - j}}.$$

By (9) the first sum is not greater than $(1 + \kappa)^{-1}$, while each member in the second sum can be estimated using Bernoulli's inequality. Indeed, for all $i \in \{2, \ldots, m\}$, $j \in \{1, \ldots, k_i\}$, (11) implies $\dfrac{R(1 + \kappa)}{|\alpha_i - \alpha_1|} < 1$ so that

$$[|\alpha_1 - \alpha_i| - (R(1 + \kappa))]^{k_i + 1 - j} \geqq [|\alpha_1 - \alpha_i|]^{k_i + 1 - j}\left(1 - (k_i + 1 - j)\frac{R(1 + \kappa)}{|\alpha_i - \alpha_1|}\right).$$

Therefore, by the definitions of s and d the second sum in (12) is bounded by $s/(1 - d(1 + \kappa))$ which is not greater than $\kappa/(1 + \kappa)$ because of (10).

Altogether,

$$h(R(1 + \kappa)) \leqq \frac{1}{1 + \kappa} + \frac{\kappa}{1 + \kappa} = 1$$

for $\kappa \in [\underline{\kappa}, \bar{\kappa}]$. □

Remark 6.2.

1. Lemma 6.1 and Theorem 5.1 give a new elementary and quite different proof of Satz 2 in [1]. Moreover, if $h(\underline{R}) < 1$ the estimate $\underline{R}$ can easily be improved. For instance, by the strict convexity of h,

$$R' := \underline{R} - (\underline{R} - R)\frac{1 - h(\underline{R})}{h(R) - h(\underline{R})}$$

fulfills $h(R') < 1$ and is a better bound than $\underline{R}$.

2. The case $h(\underline{R}) = 1$ while $\sqrt{s} + \sqrt{d} < 1$ is only possible in the quite uninteresting situation

$$k_1 = k_2 = \ldots = k_m = 1 \qquad \text{and} \qquad |\alpha_i - \alpha_1 - b_i^{k_i}| + |b_j^{k_j}| = |\alpha_i - \alpha_1| = \frac{R}{d}$$

for all $i \in \{2, \ldots, m\}$.

For a proof one can discuss the equality in each estimate in the proof of Lemma 6.1. For example, the first member in the first sum of (12) does not vanish, so $(1 + \kappa) = (1 + \kappa)^{k_1}$ implies $k_1 = 1$. The equality in Bernoulli's inequality leads to $k_1 = k_2 = \ldots = k_m$, while the other assertions arise from the equality in the triangle inequality.

We conclude this note with two examples. We will compare the best estimate $R^* =: R(1 + \kappa^*)$ of Theorem 5.1 defined as the lower zero of $h - 1$ and the estimate $\underline{R} =: R(1 + \underline{\kappa})$ from Satz 2 in [1] defined in Lemma 6.1.

The first example represents the situation in which each multiple root of f is very well approximated by a multiple knot α_i with the correct multiplicity k_i. Indeed, this general context yields asymptotically the choice of the coefficients.

Example 6.3. Let $\Delta_1, \ldots, \Delta_m > 0$ and

$$|b_i^j| = \binom{k_i}{j-1} \Delta_i^{k_i - j + 1}$$

for all $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, k_i\}$. Let f be defined by (4). We assume that $\Delta_1, \ldots, \Delta_m$ are sufficiently small so that $\sqrt{s} + \sqrt{d} < 1$.

Since $h(R^*) = 1$,

$$1 = \left(\frac{\Delta_1}{R(1 + \kappa^*)} + 1\right)^{k_1} - 1 + s^*,$$

where s^* denotes the second sum in (7) which is bounded as in the proof of Lemma 6.1,

$$s^* \leqq \frac{s}{1 + \kappa^*}.$$

On the other hand the definition of R yields

$$\left(1 + \frac{\Delta_1}{R}\right)^{k_1} = 2.$$

Hence

$$\kappa^* = \frac{\sqrt[k_1]{2} - 1}{\sqrt[k_1]{2 - s^*} - 1} - 1.$$

Since $\underline{\kappa} \geqq s \geqq s^*(1 + \kappa^*)$,

$$\frac{\kappa^*}{\underline{\kappa}} \leqq \frac{\sqrt[k_1]{2} - \sqrt[k_1]{2 - s^*}}{s^*(1 + \kappa^*)\left(\sqrt[k_1]{2 - s^*}\right)}.$$

Table 1

i	ζ_i	α_i	k_i f_1	f_2	f_3	f_4
1	$0.0 + 5.1i$	$0.0 + 5.0i$	1	1	1	1
2	$0.2 + 4.3i$	$0.1 + 4.3i$	1	1	1	1
3	$2.1 + 2.3i$	$2.0 + 2.3i$	1	1	1	1
4	$4.7 + 3.8i$	$4.7 + 3.9i$	1	2	3	4
5	$6.7 + 6.7i$	$6.7 + 6.6i$	1	1	1	1
6	$9.1 + 6.6i$	$9.0 + 6.6i$	1	1	1	1

Table 2

Bounds for f_1				Bounds for f_2			
i	k_i	R_i^*	[1]	i	k_i	R_i^*	[1]
1	1	0.1411975601	0.151	1	1	0.1459495725	0.158
2	1	0.1530164828	0.166	2	1	0.1579425322	0.173
3	1	0.1161687711	0.117	3	1	0.1189119669	0.120
4	1	0.1145788178	0.115	4	2	0.2731549505	0.276
5	1	0.1179523269	0.119	5	1	0.1246988916	0.126
6	1	0.1100754760	0.110	6	1	0.1137394577	0.144

Bounds for f_3				Bounds for f_4			
i	k_i	R_i^*	[1]	i	k_i	R_i^*	[1]
1	1	0.1508867988	0.165	1	1	0.1560229298	0.175
2	1	0.1630868591	0.181	2	1	0.1684670831	0.191
3	1	0.1219205791	0.124	3	1	0.1252226462	0.128
4	3	0.4356022139	0.44	4	4	0.6010582943	0.62
5	1	0.1317086293	0.135	5	1	0.1390076249	0.145
6	1	0.1174985422	0.119	6	1	0.1213583066	0.124

The last bound tends to

$$\frac{\sqrt[k_1]{2}}{2k_1(\sqrt[k_1]{2} - 1)}, \tag{13}$$

if the errors $(\Delta_1, \ldots, \Delta_m)$ tends to $(0, \ldots, 0)$.

The following table shows the values of the bound (13) for different k_1.

k_1	1	2	3	4	5	10	∞
(13)	1	0.85355	0.80789	0.78565	0.77250	0.74664	0.72135

This example shows that the estimates in Theorem 5.1 are slightly better, but not asymptotically equal to those of Satz 2 in [1] if $k_1 > 1$. However, for $k_1 = 1$ one should take $\zeta = \alpha_1 - b_1^1$; then it is easily proved that the bound in Theorem 5.1 is much better than that of Satz 1 in [1].

Example 6.4. Numerical Example from [1]. Let $k = 1, 2, 3, 4$ and $f_k \in \mathbb{P}_{5+k}^{\text{monic}}$ with the zeros $\zeta_1, \ldots, \zeta_6$ and let the approximations $\alpha_1, \ldots, \alpha_6$ with the multiplicities $k_1, \ldots, k_6$ be given by Table 1. The polynomials differ only in the multiplicity of the fourth root.

The approximations $\alpha_1, \ldots, \alpha_6$ are used in Theorem 5.1 with the multiplicities of the zeros. The best bounds R^* for the approximation α_i are computed and listed in Table 2. These estimates are slightly better than those of [1, Table 2].

The examples show that the inclusion regions of Theorem 5.1 are slightly smaller than those of Satz 2 in [1]. On the other hand the computational work may be slightly more laborious because the computation of the zero of $h - 1$ may be more expensive than the computation of R.

References

1. Börsch-Supan, W.: Residuenabschätzung für Polynom-Nullstellen mittels Lagrange-Interpolation. Numer. Math. **14**, 287–296 (1970)
2. Carstensen, C. Stein, E.: Über die Falksche ECP-Transformation und Verallgemeinerungen. ZAMM **69**, 375–391 (1989)
3. Carstensen, C.: On a linear construction of companion matrices. Linear Algebra Appl. (1990) (accepted for publication)
4. Elsner, L.: Über Eigenwerteinschließungen mit Hilfe von Gerschgorin–Kreisen. ZAMM **50**, 381–384 (1970)
5. Elsner, L.: A Remark on Simultaneous Inclusions of the Zeros of a Polynomial by Gershgorin's Theorem. Numer. Math. **21**, 425–427 (1973)
6. Elsner, L.: On the Variation of the Spectra of Matrices. Linear Algebra Appl. **47**, 127–138 (1982)
7. Horn, R.A., Johnson, C.A.: Matrix Analysis. Cambridge: Cambridge University Press, 1985
8. Householder, A.S.: The Theory of Matrices in Numerical Analysis. New York: Blaisdell 1964
9. Klein, P.P.: Über eine Eigenwertaufgabe, die bei der Einschließung von Matrixeigenwerten und Polynomnullstellen durch kleinste isolierte Gerschgorinkreise auftritt. Ph.D. Thesis, TU Clausthal, 1978
10. Medley, H.I., Varga, R.S.: On Smallest Isolated Gerschgorin–Disks for Eigenvalues III. Numer. Math. **11**, 361–369 (1968)
11. Parodi, M.: La Localisation des Valeurs Charactéristiques des Matrices et ses Applications. Paris: Gauthier-Villars 1959
12. Smith, B.T.: Error Bounds for Zeros of a Polynomial Based Upon Gerschgorin's Theorems. JACM **17**, 661–674 (1970)
13. Werner, W.: A Generalized Companion Matrix of a Polynomial and Some Applications. Linear Algebra Appl. **55**, 19–36 (1983)
14. Zurmühl, R., Falk, S.: Matrizen und ihre Anwendungen. Vols 1, 2, 5th Edn. Berlin Heidelberg New York: Springer 1984, 1986