

Research Article

Open Access

Special Issue Dedicated to Charles R. Johnson

A. Melman*

Extensions of the Eneström-Kakeya theorem for matrix polynomials

<https://doi.org/10.1515/spma-2019-0024>

Received August 19, 2019; accepted December 3, 2019

Abstract: The classical Eneström-Kakeya theorem establishes explicit upper and lower bounds on the zeros of a polynomial with positive coefficients and has been generalized for positive definite matrix polynomials by several authors. Recently, extensions that improve the (scalar) Eneström-Kakeya theorem were obtained with a transparent and unified approach using just a few tools. Here, the same tools are used to generalize these extensions to positive definite matrix polynomials, while at the same time generalizing the tools themselves. In the process, a framework is developed that can naturally generate additional similar results.

Keywords: matrix polynomial, positive definite, polynomial eigenvalue, Cauchy, Eneström-Kakeya

MSC: 12D10, 30C15, 65H05

1 Introduction

We examine the generalization of the Eneström-Kakeya theorem ([5], [6], and [9]) for *scalar* polynomials with positive coefficients to *matrix* polynomials with positive definite matrix coefficients. In the scalar case, this elegant theorem establishes upper and lower bounds on the moduli of polynomial zeros that are simple explicit functions of the coefficients. It has been generalized to positive definite matrix polynomials in, e.g., [4] and [10], but the scalar version was also recently extended and improved in [15], where different zero inclusion regions were established in a transparent and unified way using just a few tools. We generalize some of these extended results to positive definite matrix polynomials and, at the same time, generalize the tools as well. In doing so, we establish a framework to naturally generate additional similar results. Throughout, we concentrate on upper bounds or inclusion regions for the zeros or eigenvalues since corresponding lower bounds or exclusion regions can then be obtained from the reverse polynomial.

The paper is organized as follows. In Section 2, we briefly review the scalar Eneström-Kakeya theorem and informally present a different way than is found elsewhere to extend it to matrix polynomials. This then sets the stage for the generalization of the aforementioned extensions of the theorem in Section 3 and Section 4. Some numerical illustrations are presented in Section 5.

2 The Eneström-Kakeya theorem

We begin by stating the scalar version of the Eneström-Kakeya theorem ([5], [6], [9]).

*Corresponding Author: A. Melman: Department of Applied Mathematics, School of Engineering, Santa Clara University, Santa Clara, CA 95053, E-mail : amelman@scu.edu

Theorem 2.1 (Eneström-Kakeya). *All the zeros of the real polynomial $p(z) = \sum_{j=0}^n a_j z^j$ with positive coefficients lie in the disk defined by $|z| \leq \gamma$, where*

$$\gamma = \max_{0 \leq j \leq n-1} \frac{a_j}{a_{j+1}}.$$

A simple proof is obtained from the classical observation by Cauchy from 1829 ([3], see also, e.g., [11, Th.(27,1), p.122 and Exercise 1, p.126], [16, Theorem 3.1.1], [17, Theorem 8.1.3]), which is the following theorem.

Theorem 2.2 (Cauchy). *All the zeros of the polynomial $p(z) = \sum_{j=0}^n a_j z^j$ of degree n with complex coefficients lie in the disk $|z| \leq \rho$, where ρ is the largest positive solution of*

$$|a_n|z^n - \sum_{j=0}^{n-1} |a_j|z^j = 0.$$

The quantity ρ is called the *Cauchy radius* of p .

To prove the Eneström-Kakeya theorem, consider, for $\gamma \in \mathbb{R}$,

$$(z - \gamma)p(z) = a_n z^{n+1} + \sum_{j=1}^n (a_{j-1} - \gamma a_j) z^j - \gamma a_0. \quad (1)$$

Clearly, any upper bound on the zeros of $(z - \gamma)p(z)$ will also be an upper bound on the zeros of p . From (1) we observe that the coefficient of z^{n+1} in (1) is positive, while the remaining coefficients will be nonpositive if

$$\gamma = \max_{0 \leq j \leq n-1} \frac{a_j}{a_{j+1}}.$$

With this value of γ , the Cauchy radius of $(z - \gamma)p(z)$ is the unique positive solution of $(z - \gamma)p(z) = 0$ and, since p cannot have positive zeros, γ must be that Cauchy radius, which proves the theorem.

The Cauchy radius requires the solution of a nonlinear equation, but the Eneström-Kakeya bound is easily computed as the maximum of a set of numbers. It can be better or worse than the Cauchy radius.

Let us now turn our attention to matrix polynomials of the form $P(z) = \sum_{j=0}^n A_j z^j$, with square complex coefficient matrices. These occur in polynomial eigenvalue problems, where eigenvalues $\zeta \in \mathbb{C}$ and eigenvectors $u \in \mathbb{C}^n$, $u \neq 0$, are sought such that $P(\zeta)u = 0$. If A_n is singular, there are infinite eigenvalues and if A_0 is singular there are zero eigenvalues. If A_n is nonsingular, the eigenvalues are the solutions of $\det(P(z)) = 0$. The eigenvalues of MPN , where M and N are nonsingular constant matrices are the same as those of P . A matrix polynomial P is *regular* if $\det(P)$ is not identically zero. We will only consider regular matrix polynomials. Note that when P is linear and monic, i.e., $P(z) = Iz + A_0$, one obtains the standard eigenvalue problem. Polynomial eigenvalue problems are encountered in many engineering applications (e.g., [1]).

From here on, we consider matrix polynomials with positive definite coefficients, i.e., *Hermitian* matrices with positive eigenvalues. Such matrix polynomials are regular. When the eigenvalues of a matrix are non-negative, but some are zero, the matrix is positive semidefinite. Negative definiteness is defined analogously. We assume familiarity with the basic properties of positive definite matrices that can be found in, e.g., [8, Chapter 7]. We will also often use the fact that, for positive definite matrices A and B , $A^{1/2}$ and $B^{1/2}$ are well-defined and the matrices $A^{1/2}BA^{1/2}$, $B^{1/2}AB^{1/2}$, AB , and BA all have the same eigenvalues. We denote by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ the smallest and largest eigenvalues of the matrix A , respectively. When A and B are positive definite and $\mu, \nu \in \mathbb{R}$, then, for any matrix norm,

$$\lambda_{\min}(A) = \frac{1}{\lambda_{\max}(A^{-1})} \geq \|A^{-1}\|^{-1},$$

$$\mu \leq \nu \Rightarrow \lambda_{\max}(A - \mu B) \geq \lambda_{\max}(A - \nu B).$$

We denote by A^* the Hermitian conjugate of A and vector norms will be considered to be Euclidean, i.e., $\|u\|^2 = u^*u$.

The Cauchy radius of a scalar polynomial is the best upper bound on the magnitudes of its zeros among all bounds that depend only on the magnitudes of the coefficients. Since the latter is the case for most commonly used bounds, it makes the Cauchy radius an excellent bound for benchmarking purposes. We now reexamine it in the context of matrix polynomials. Let P therefore be a matrix polynomial with positive definite coefficients, and consider an eigenvector $u \in \mathbb{C}^n$, $\|u\| = 1$, and eigenvalue ζ of $A_n^{-1/2}PA_n^{-1/2}$, which is also a positive definite matrix polynomial that has the same eigenvalues as P . Since $A_n^{-1/2}P(\zeta)A_n^{-1/2}u = 0$, we have

$$u^* A_n^{-1/2} P(\zeta) A_n^{-1/2} u = \zeta^n + \sum_{j=0}^{n-1} \left(u^* A_n^{-1/2} A_j A_n^{-1/2} u \right) \zeta^j = 0,$$

which implies by Theorem 2.2 that $|\zeta|$ is bounded by the unique positive solution of

$$z^n - \sum_{j=0}^{n-1} \left(u^* A_n^{-1/2} A_j A_n^{-1/2} u \right) z^j = 0. \tag{2}$$

However, the eigenvectors are unknown, so that we need to consider the largest possible positive solution over all unit vectors u . From the positive definiteness of $A_n^{-1/2}A_jA_n^{-1/2}$ we have that

$$u^* A_n^{-1/2} A_j A_n^{-1/2} u \leq \lambda_{\max} \left(A_n^{-1/2} A_j A_n^{-1/2} \right) = \lambda_{\max} \left(A_j A_n^{-1} \right) = \lambda_{\max} \left(A_n^{-1} A_j \right),$$

which leads to the following theorem.

Theorem 2.3. *The eigenvalues of the matrix polynomial $P(z) = \sum_{j=0}^n A_j z^j$ with positive definite coefficients lie in the disk $|z| \leq \rho$, where ρ is the unique positive solution of*

$$z^n - \sum_{j=0}^{n-1} \lambda_{\max} \left(A_n^{-1} A_j \right) z^j = 0. \tag{3}$$

Moreover, increasing the coefficients in (3) increases its positive solution, with the following corollary as an immediate consequence.

Corollary 2.1. *The eigenvalues of the matrix polynomial $P(z) = \sum_{j=0}^n A_j z^j$ with positive definite coefficients lie in the disk $|z| \leq \rho$, where ρ is the unique positive solution of*

$$z^n - \sum_{j=0}^{n-1} \|A_n^{-1} A_j\| z^j = 0.$$

This is essentially a generalization to matrix polynomials of the Cauchy radius that was derived in [2], [7], and [13] for general square complex matrix polynomials. For a given norm, it is optimal among all bounds that depend only on $\|A_n^{-1}A_0\|, \dots, \|A_n^{-1}A_{n-1}\|$.

One can avoid multiplying P by $A_n^{-1/2}$ and obtain, instead of (2), the equation

$$(u^* A_n u) z^n - \sum_{j=0}^{n-1} (u^* A_j u) z^j = 0,$$

so that the equivalent of (3) becomes

$$\lambda_{\min} (A_n) z^n - \sum_{j=0}^{n-1} \lambda_{\max} (A_j) z^j = 0.$$

It is not hard to show that this leads to a worse bound than the one obtained from (3) and we will not use it.

We remark that if only monic matrix polynomials are considered, then multiplying P by $A_n^{-1/2}$ can be avoided altogether.

Generalizations to matrix polynomials of the Eneström-Kakeya theorem were obtained in, e.g., [4] and [10]. However, we shall now informally rederive it to introduce a different approach, needed to provide a transparent framework for the generalization of some of the aforementioned extensions of the Eneström-Kakeya theorem in [15], which is the main focus of this work.

With u and ζ as before, consider the scalar polynomial $q(z) := (z - \gamma) u^* A_n^{-1/2} P(z) A_n^{-1/2} u$ with $\gamma \in \mathbb{R}$, for which $q(\zeta) = 0$. We have

$$q(z) = z^{n+1} + \sum_{j=1}^n \left(u^* A_n^{-1/2} (A_{j-1} - \gamma A_j) A_n^{-1/2} u \right) z^j - \gamma u^* A_n^{-1/2} A_0 A_n^{-1/2} u . \tag{4}$$

The coefficient of z^{n+1} in (4) is positive, the constant term is nonpositive for $\gamma \geq 0$, and the remaining coefficients will be nonpositive if, for $1 \leq j \leq n$, $A_n^{-1/2} (A_{j-1} - \gamma A_j) A_n^{-1/2}$ is negative semidefinite, which is equivalent to $A_{j-1} - \gamma A_j$ being negative semidefinite. This will be true if $\gamma \geq \lambda_{\max} (A_j^{-1} A_{j-1})$ for $1 \leq j \leq n$. In that case, the Cauchy radius of q , which is its own unique positive zero, must be equal to γ since $u^* A_n^{-1/2} P(z) A_n^{-1/2} u$ has no positive zeros. We have arrived at the following theorem, which is the generalization of the scalar Eneström-Kakeya theorem to positive definite matrix polynomials.

Theorem 2.4. *The eigenvalues of the matrix polynomial $P(z) = \sum_{j=0}^n A_j z^j$ with positive definite coefficients lie in the disk $|z| \leq \gamma$, where*

$$\gamma = \max_{0 \leq j \leq n-1} \left\{ \lambda_{\max} \left(A_{j+1}^{-1} A_j \right) \right\} .$$

We also have the following immediate corollary, which follows from the inequalities $\lambda_{\max} (A) \leq \|A\|$ and $\lambda_{\max} (AB) \leq \lambda_{\max} (A) \lambda_{\max} (B)$, when A and B are positive definite.

Corollary 2.2. *The eigenvalues of the matrix polynomial $P(z) = \sum_{j=0}^n A_j z^j$ with positive definite coefficients lie in the disk $|z| \leq \gamma$, where, for any matrix norm,*

$$\gamma = \max_{0 \leq j \leq n-1} \frac{\lambda_{\max} (A_j)}{\lambda_{\min} (A_{j+1})} , \quad \gamma = \max_{0 \leq j \leq n-1} \left\| A_j A_{j+1}^{-1} \right\| , \quad \text{or} \quad \gamma = \max_{0 \leq j \leq n-1} \left\| A_{j+1}^{-1} A_j \right\| .$$

Theorem 2.4 was obtained as Theorem 2.6 in [4] for the special case of positive definite matrices, while Corollary 2.2 with the first value of γ , which is larger than in Theorem 2.4, was obtained in [10].

It is of theoretical interest that one can derive a matrix version of the Eneström-Kakeya theorem, but it is only practical if the eigenvalue computations and the inversions it requires are not too cumbersome. This may be the case for a high degree polynomial with small sized coefficients, or when the coefficients are highly structured, although this discussion falls outside our scope. Generally speaking, the computation of the Cauchy radius, even if it involves solving a nonlinear equation (which the Eneström-Kakeya theorem avoids), requires a computational effort that is several orders of magnitude less than the computation of the eigenvalues it bounds. It therefore remains a useful alternative to the Eneström-Kakeya theorem.

The structure of the next two sections reflects the approach we have followed: a Cauchy-like result for scalar polynomials is first generalized into a Cauchy-like result for matrix polynomials and then the same is done for a corresponding Eneström-Kakeya-like theorem. Each is organized in its own subsection.

3 Extensions of the Eneström-Kakeya theorem (single disk)

In [12, Theorem 1], another disk that also contains all the zeros of a scalar polynomial was defined, which was used in [15] to derive an Eneström-Kakeya related theorem, just as we used Theorem 2.2 to derive the Eneström-Kakeya theorem. We generalize it in this section to matrix polynomials.

3.1 Cauchy related results

The following theorem is Theorem 1 in [12].

Theorem 3.1. *All the zeros of the polynomial $p(z) = \sum_{j=0}^n a_j z^j$ of degree $n \geq 2$ with complex coefficients lie in the disk $|z + a_{n-1}/a_n| \leq |a_{n-1}/a_n| + \sigma$, where σ is the largest nonnegative solution of*

$$|a_n|z^n + |a_{n-1}|z^{n-1} - \sum_{j=0}^{n-2} |a_j|z^j = 0.$$

We define σ in Theorem 3.1 as the *Cauchy radius of the second kind*. As we did for Theorem 2.2, we first generalize this theorem for positive definite matrix polynomials.

Theorem 3.2. *The eigenvalues of the matrix polynomial $P(z) = \sum_{j=0}^n A_j z^j$, $n \geq 2$, with positive definite coefficients, lie in the disk $|z + \lambda_{\min}(A_n^{-1}A_{n-1})| \leq \lambda_{\max}(A_n^{-1}A_{n-1}) + \sigma$, where σ is the unique positive solution of*

$$z^n + \lambda_{\min}(A_n^{-1}A_{n-1})z^{n-1} - \sum_{j=0}^{n-2} \lambda_{\max}(A_n^{-1}A_j)z^j = 0. \quad (5)$$

Proof. As for Theorem 2.3, we consider $A_n^{-1/2}PA_n^{-1/2}$ with an eigenvector $u \in \mathbb{C}^n$, $\|u\| = 1$, corresponding to an eigenvalue ζ , so that $u^*A_n^{-1/2}P(\zeta)A_n^{-1/2}u = 0$. Theorem 3.1 applied to this scalar polynomial then implies that all its zeros lie in the disk

$$\left| z + u^*A_n^{-1/2}A_{n-1}A_n^{-1/2}u \right| \leq u^*A_n^{-1/2}A_{n-1}A_n^{-1/2}u + \sigma_u,$$

where σ_u is the unique positive solution of

$$z^n + \left(u^*A_n^{-1/2}A_{n-1}A_n^{-1/2}u \right) z^{n-1} - \sum_{j=0}^{n-2} \left(u^*A_n^{-1/2}A_jA_n^{-1/2}u \right) z^j = 0.$$

Because of the positive definiteness of the coefficient matrices, an upper bound on σ_u that does not depend on u is provided by the unique positive solution ρ of the equation

$$z^n + \lambda_{\min}(A_n^{-1}A_{n-1})z^{n-1} - \sum_{j=0}^{n-2} \lambda_{\max}(A_n^{-1}A_j)z^j = 0.$$

However, we do not know u and therefore must assume that $u^*A_n^{-1/2}A_{n-1}A_n^{-1/2}u$ can be any number in the interval $[\lambda_{\min}(A_n^{-1}A_{n-1}), \lambda_{\max}(A_n^{-1}A_{n-1})]$. Each such number θ corresponds to a zero inclusion disk of the form $|z + \theta| \leq \theta + \rho$, but all these disks are contained in the one defined in the statement of the theorem. \square

The following corollary follows from the inequalities $\lambda_{\max}(A) \leq \|A\|$ and $\lambda_{\min}(A) \geq \|A^{-1}\|^{-1}$ that hold for positive definite matrices.

Corollary 3.1. *The eigenvalues of the matrix polynomial $P(z) = \sum_{j=0}^n A_j z^j$, $n \geq 2$, with positive definite coefficients, lie in the disk $|z + \|A_n^{-1}A_{n-1}\|^{-1}| \leq \|A_n^{-1}A_{n-1}\| + \sigma$, where σ is the unique positive solution of*

$$z^n + \|A_n^{-1}A_{n-1}\|^{-1}z^{n-1} - \sum_{j=0}^{n-2} \|A_n^{-1}A_j\|z^j = 0.$$

We note that, as was the case for Cauchy's observation, we could refrain from multiplying by $A_n^{-1/2}$, resulting in the equation

$$\lambda_{\min}(A_n)z^n + \lambda_{\min}(A_{n-1})z^{n-1} - \sum_{j=0}^{n-2} \lambda_{\max}(A_j)z^j = 0,$$

whose unique positive solution is never smaller than the one of equation (5) in Theorem 3.2.

3.2 Eneström-Kakeya related results

We now derive a matrix version of Theorem 4.2 in [15], which is an extension of the Eneström-Kakeya theorem.

Theorem 3.3. *The eigenvalues of the matrix polynomial $P(z) = \sum_{j=0}^n A_j z^j$, $n \geq 2$, with positive definite coefficients, lie in the disk $|z + \lambda_{\max}(A_n^{-1}A_{n-1})| \leq \lambda_{\max}(A_n^{-1}A_{n-1}) + \sqrt{\gamma}$, where*

$$\gamma = \max_{0 \leq j \leq n-2} \left\{ \lambda_{\max}(A_{j+2}^{-1}A_j) \right\}.$$

Proof. We rely once more on the fact that $A_n^{-1/2}PA_n^{-1/2}$ and P have the same eigenvalues. Let $u \in \mathbb{C}^n$, $\|u\| = 1$, be an eigenvector of $A_n^{-1/2}PA_n^{-1/2}$ with corresponding eigenvalue ζ , so that $u^*A_n^{-1/2}P(\zeta)A_n^{-1/2}u = 0$. Define a polynomial $q(z) := (z^2 - \gamma)u^*A_n^{-1/2}P(z)A_n^{-1/2}u$ with $\gamma \in \mathbb{R}$, so that $q(\zeta) = 0$. Then we have

$$\begin{aligned} (z^2 - \gamma)u^*A_n^{-1/2}P(z)A_n^{-1/2}u &= z^{n+2} + \left(u^*A_n^{-1/2}A_{n-1}A_n^{-1/2}u\right)z^{n+1} + \sum_{j=2}^n \left(u^*A_n^{-1/2}(A_{j-2} - \gamma A_j)A_n^{-1/2}u\right)z^j \\ &\quad - \gamma \left(u^*A_n^{-1/2}A_1A_n^{-1/2}u\right)z - \gamma u^*A_n^{-1/2}A_0A_n^{-1/2}u. \end{aligned}$$

The coefficients of z^{n+2} and z^{n+1} are positive and the coefficients of the linear and constant term are non-positive if $\gamma \geq 0$. The remaining coefficients will be nonpositive if, for $j = 2, \dots, n$, $A_n^{-1/2}(A_{j-2} - \gamma A_j)A_n^{-1/2}$ is negative semidefinite, which is equivalent to $A_{j-2} - \gamma A_j$ being negative semidefinite. This will be true if we choose

$$\gamma = \max_{0 \leq j \leq n-2} \left\{ \lambda_{\max}(A_j A_{j+2}^{-1}) \right\} = \max_{0 \leq j \leq n-2} \left\{ \lambda_{\max}(A_{j+2}^{-1}A_j) \right\}.$$

Reasoning similarly as in the proof of the scalar Eneström-Kakeya theorem, the Cauchy radius of the second kind of q , which is its unique positive zero, must be equal to $\sqrt{\gamma}$ since $u^*A_n^{-1/2}P(z)A_n^{-1/2}u$ has no positive zeros. From Theorem 3.1, we know that the zeros of q , and therefore those of $u^*A_n^{-1/2}PA_n^{-1/2}u$ must lie in the disk $|z + u^*A_n^{-1/2}A_{n-1}A_n^{-1/2}u| \leq u^*A_n^{-1/2}A_{n-1}A_n^{-1/2}u + \sqrt{\gamma}$. The rest of the proof follows exactly as in the proof of Theorem 3.2. \square

The following corollary is immediate.

Corollary 3.2. *The eigenvalues of the matrix polynomial $P(z) = \sum_{j=0}^n A_j z^j$, $n \geq 2$, with positive definite coefficients, lie in the disk $|z + \|A_n^{-1}A_{n-1}\| \leq \|A_n^{-1}A_{n-1}\| + \sqrt{\gamma}$, where*

$$\gamma = \max_{0 \leq j \leq n-2} \left\| A_{j+2}^{-1}A_j \right\| \quad \text{or} \quad \gamma = \max_{0 \leq j \leq n-2} \left\| A_j A_{j+2}^{-1} \right\|.$$

4 Extensions of the Eneström-Kakeya theorem (multiple disks)

Both Theorems 2.2 and 3.1 are special cases of more general zero inclusion regions that were derived in [14]. As in the previous section, we first examine, in Section 4.1, how these results can be generalized for positive definite matrix polynomials in the spirit of Theorems 2.3 and 3.2, before deriving extensions of the Eneström-Kakeya theorem in Section 4.2. Instead of being composed of a single disk, as before, the eigenvalue inclusion regions we shall encounter in this section will consist of multiple (usually smaller) disks.

4.1 Cauchy related results for scalar polynomials

In the following, we consider a polynomial $p(z) = \sum_{j=0}^n a_j z^j$ with complex coefficients, unless specified otherwise. We first state the basic definitions and results from [14] that we will refer to later on.

Definition 4.1. The Cauchy radius of the k th kind of p , with $k = 1, \dots, n$, is defined as the unique positive solution s_k of

$$|a_n|z^n + |a_{n-1}|z^{n-1} + \dots + |a_{n-k+1}|z^{n-k+1} - |a_{n-k}|z^{n-k} - \dots - |a_1|z - |a_0| = 0 .$$

When $k = 1$, s_1 is simply called the Cauchy radius of p .

Definition 4.2. The sets $\Gamma_1(k)$ and $\Gamma_2(k)$ for a polynomial p , with $k = 1, \dots, n$, are defined as

$$\Gamma_1(k) = \left\{ z \in \mathbb{C} : |a_n z^k + a_{n-1} z^{k-1} + \dots + a_{n-k+1} z| \leq |a_n|s_k^k + |a_{n-1}|s_k^{k-1} + \dots + |a_{n-k+1}|s_k \right\}$$

and

$$\Gamma_2(k) = \left\{ z \in \mathbb{C} : |a_n z^k + a_{n-1} z^{k-1} + \dots + a_{n-k+1} z + a_{n-k}| \leq |a_n|s_{k+1}^k + |a_{n-1}|s_{k+1}^{k-1} + \dots + |a_{n-k+1}|s_{k+1} + |a_{n-k}| \right\} .$$

The boundaries of $\Gamma_1(k)$ and $\Gamma_2(k)$ are lemniscates. With these definitions, we state the following theorem, which is Theorem 3.2 in [14], where k here corresponds to $n - k$ in that theorem.

Theorem 4.1. All the zeros of the complex polynomial $p(z) = \sum_{j=0}^n a_j z^j$ lie in the sets $\Gamma_1(k)$ and $\Gamma_2(k)$. If $\Gamma_1(k)$ or $\Gamma_2(k)$ consists of disjoint regions whose boundaries are simple closed (Jordan) curves and ℓ is the number of foci of its bounding lemniscate that are contained in any such region, then that region contains ℓ zeros of p when it does not contain the origin, and $\ell + n - k$ zeros of p when it does contain the origin.

The inclusion disk from Theorem 2.2 is the set $\Gamma_1(1)$, while the one from Theorem 3.1 is the set $\Gamma_2(1)$. As k increases, $\Gamma_1(k)$ and $\Gamma_2(k)$ become too complicated. Instead, we will approximate a region Γ , bounded by a lemniscate of the form $|q(z)| = R$, where q is a polynomial of degree m with zeros c_j , by a union of disks centered at the zeros of q with radius $R^{1/m}$, for which

$$\Gamma \subseteq \bigcup_{j=1}^m \left\{ z \in \mathbb{C} : |z - c_j| \leq R^{1/m} \right\} .$$

Although larger than Γ , this union of disks is easier to work with, and is, as we shall see, still useful.

The following theorem uses Theorem 4.1 to derive an eigenvalue inclusion region for positive definite matrix polynomials composed of a continuum of disks.

Theorem 4.2. Let $P(z) = \sum_{j=0}^n A_j z^j$ be a matrix polynomial of degree $n \geq 2$ with positive definite coefficients, let σ be the unique positive solution of

$$z^n + \lambda_{\min} \left(A_n^{-1} A_{n-1} \right) z^{n-1} - \sum_{j=0}^{n-2} \lambda_{\max} \left(A_n^{-1} A_j \right) z^j = 0 ,$$

and let $R = \sigma^2 + \sigma \lambda_{\max} \left(A_n^{-1} A_{n-1} \right)$. Then all the eigenvalues of P lie in the set

$$\left\{ z \in \mathbb{C} : |z| \leq R^{1/2} \right\} \cup \left(\bigcup_{\lambda_{\min} \left(A_n^{-1} A_{n-1} \right) \leq t \leq \lambda_{\max} \left(A_n^{-1} A_{n-1} \right)} \left\{ z \in \mathbb{C} : |z + t| \leq R^{1/2} \right\} \right) .$$

Proof. As before, consider $A_n^{-1/2} P A_n^{-1/2}$ with an eigenvector $u \in \mathbb{C}^n$, $\|u\| = 1$, corresponding to an eigenvalue ζ , so that $u^* A_n^{-1/2} P(\zeta) A_n^{-1/2} u = 0$. Theorem 4.1 applied with the set $\Gamma_1(2)$ to this scalar polynomial then implies that all its zeros lie in the set $\Delta(u)$, defined by

$$\Delta(u) = \left\{ z \in \mathbb{C} : \left| z \left(z + u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right) \right| \leq \sigma_u^2 + \sigma_u u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right\} ,$$

where σ_u is the unique positive solution of

$$z^n + \left(u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right) z^{n-1} - \sum_{j=0}^{n-2} \left(u^* A_n^{-1/2} A_j A_n^{-1/2} u \right) z^j = 0 .$$

Because of the positive definiteness of the coefficient matrices, an upper bound on σ_u that does not depend on u is provided by the unique positive solution σ of the equation

$$z^n + \lambda_{\min} \left(A_n^{-1} A_{n-1} \right) z^{n-1} - \sum_{j=0}^{n-2} \lambda_{\max} \left(A_n^{-1} A_j \right) z^j = 0 ,$$

so that

$$\Delta(u) \subseteq \left\{ z \in \mathbb{C} : \left| z \left(z + u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right) \right| \leq \sigma^2 + \sigma u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right\} .$$

The set $\Delta(u)$ was obtained for one particular (unknown) eigenvector u . Therefore, all the eigenvalues of P must lie in the set $\cup_{u \in \mathbb{C}^n, \|u\|=1} \Delta(u)$. Since $u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u$ can lie anywhere in the interval $[\lambda_{\min} (A_n^{-1} A_{n-1}), \lambda_{\max} (A_n^{-1} A_{n-1})]$ and $u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \leq \lambda_{\max} (A_n^{-1} A_{n-1})$, we obtain, with R as in the statement of the theorem, that

$$\bigcup_{u \in \mathbb{C}^n, \|u\|=1} \Delta(u) \subseteq \left\{ z \in \mathbb{C} : |z| \leq R^{1/2} \right\} \cup \left(\bigcup_{\lambda_{\min}(A_n^{-1}A_{n-1}) \leq t \leq \lambda_{\max}(A_n^{-1}A_{n-1})} \left\{ z \in \mathbb{C} : |z+t| \leq R^{1/2} \right\} \right) .$$

□

Corollary 4.1. Let $P(z) = \sum_{j=0}^n A_j z^j$ be a matrix polynomial of degree $n \geq 2$ with positive definite coefficients, let σ be the unique positive solution of

$$z^n + \left\| A_{n-1}^{-1} A_n \right\|^{-1} z^{n-1} - \sum_{j=0}^{n-2} \left\| A_n^{-1} A_j \right\| z^j = 0 ,$$

and let $R = \sigma^2 + \sigma \left\| A_n^{-1} A_{n-1} \right\|$. Then all the eigenvalues of P lie in the set

$$\left\{ z \in \mathbb{C} : |z| \leq R^{1/2} \right\} \cup \left(\bigcup_{\left\| A_{n-1}^{-1} A_n \right\|^{-1} \leq t \leq \left\| A_n^{-1} A_{n-1} \right\|} \left\{ z \in \mathbb{C} : |z+t| \leq R^{1/2} \right\} \right) .$$

The inclusion sets in the previous theorem and corollary are composed of a disk and a region bounded by a rectangle with semicircles attached to its left and right ends. The proof of these results clearly shows that it would not be an easy matter to use Theorem 4.1 with $\Gamma_1(k)$ or $\Gamma_2(k)$ when $k > 2$ because the locations of the lemniscates' foci in that theorem become too cumbersome to estimate. In the following section we will see that such problems can be circumvented with an Eneström-Keakeya approach.

4.2 Eneström-Keakeya related results for matrix polynomials

In this section we derive an eigenvalue inclusion region that is a generalization to positive definite matrix polynomials of another extension of the Eneström-Keakeya theorem for scalar polynomials in [15], this time involving more than one inclusion disk. To avoid tedious repetition, we will work out this one inclusion region, deliberately chosen to be a little more complicated, as this is sufficient to demonstrate how the same techniques can then be applied to derive more such results.

Theorem 4.3. Let $P(z) = \sum_{j=0}^n A_j z^j$ be a matrix polynomial of degree $n \geq 3$ with positive definite coefficients, define

$$\alpha := \lambda_{\max} \left(A_n^{-1} A_{n-2} \right) ,$$

$$\beta := \max \left\{ 0, \max_{0 \leq j \leq n-3} \lambda_{\max} \left(\left(A_j - \lambda_{\min} \left(A_n^{-1} A_{n-2} \right) A_{j+2} \right) A_{j+3}^{-1} \right) \right\} ,$$

and let R be the unique positive solution of $z^3 - az - \beta = 0$. Then all the eigenvalues of P lie in the set

$$\left\{ z \in \mathbb{C} : |z| \leq R^{1/3} \right\} \cup \left(\bigcup_{\lambda_{\min}(A_n^{-1}A_{n-1}) \leq t \leq \lambda_{\max}(A_n^{-1}A_{n-1})} \left\{ z \in \mathbb{C} : |z+t| \leq R^{1/3} \right\} \right).$$

Proof. The proof is reminiscent of that of Theorem 3.3, except that here we use a cubic multiplier: let $u \in \mathbb{C}^n$, $\|u\| = 1$, be an eigenvector of $A_n^{-1/2}PA_n^{-1/2}$ with corresponding eigenvalue ζ , and define a new polynomial $q(z) := (z^3 - \gamma_1 z - \gamma_0) u^* A_n^{-1/2} P(z) A_n^{-1/2} u$, so that $q(\zeta) = 0$ and

$$\begin{aligned} q(z) &= z^{n+3} + \left(u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right) z^{n+2} + \left(u^* A_n^{-1/2} (A_{n-2} - \gamma_1 A_n) A_n^{-1/2} u \right) z^{n+1} \\ &\quad + \sum_{j=3}^n \left(u^* A_n^{-1/2} (A_{j-3} - \gamma_1 A_{j-1} - \gamma_0 A_j) A_n^{-1/2} u \right) z^j \\ &\quad - \left(u^* A_n^{-1/2} (\gamma_1 A_1 + \gamma_0 A_2) A_n^{-1/2} u \right) z^2 - \left(u^* A_n^{-1/2} (\gamma_1 A_0 + \gamma_0 A_1) A_n^{-1/2} u \right) z \\ &\quad - \gamma_0 A_n^{-1/2} A_0 A_n^{-1/2}. \end{aligned}$$

The coefficients of the two leading powers of z are positive, the coefficient of z^{n+1} vanishes if we set $\gamma_1 = u^* A_n^{-1/2} A_{n-2} A_n^{-1/2} u$, and the coefficients of z^j for $3 \leq j \leq n$ are nonpositive if the matrices $A_{j-3} - \gamma_1 A_{j-1} - \gamma_0 A_j$ are negative semidefinite. The latter will be achieved if, for all such j , $\gamma_0 \geq \lambda_{\max} \left((A_{j-3} - \gamma_1 A_{j-1}) A_j^{-1} \right)$. We therefore set

$$\gamma_0 = \max \left\{ 0, \max_{0 \leq j \leq n-3} \lambda_{\max} \left((A_j - \gamma_1 A_{j+2}) A_{j+3}^{-1} \right) \right\},$$

which ensures that the constant term and the coefficients of z and z^2 are also nonpositive.

Under these conditions, the unique positive zero r_u of $z^3 - \gamma_1 z - \gamma_0$ is the Cauchy radius of the third kind of q , so that, by Theorem 4.1 with $\Gamma_1(3)$, all the zeros of q lie in the set $\Delta(u)$, which is given by

$$\Delta(u) = \left\{ z \in \mathbb{C} : \left| z^2 \left(z + u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right) \right| \leq Q_u \right\},$$

where $Q_u = r_u^3 + \left(u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right) r_u^2$, and we have

$$\Delta(u) \subseteq \left\{ z \in \mathbb{C} : |z| \leq Q_u^{1/3} \right\} \cup \left\{ z \in \mathbb{C} : \left| z + u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right| \leq Q_u^{1/3} \right\}.$$

Moreover, $\lambda_{\min}(A_n^{-1}A_{n-2}) \leq \gamma_1 \leq \lambda_{\max}(A_n^{-1}A_{n-2}) = \alpha$, implying that

$$\gamma_0 \leq \max \left\{ 0, \max_{0 \leq j \leq n-3} \lambda_{\max} \left((A_j - \lambda_{\min}(A_n^{-1}A_{n-2}) A_{j+2}) A_{j+3}^{-1} \right) \right\} = \beta,$$

which in turn implies that $Q_u \leq R$, where R is the unique positive solution of $z^3 - az - \beta = 0$, so that

$$\Delta(u) \subseteq \left\{ z \in \mathbb{C} : |z| \leq R^{1/3} \right\} \cup \left\{ z \in \mathbb{C} : \left| z + u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \right| \leq R^{1/3} \right\}.$$

All the eigenvalues of P must lie in the set $\cup_{u \in \mathbb{C}^n, \|u\|=1} \Delta(u)$ and, because

$$\lambda_{\min}(A_n^{-1}A_{n-1}) \leq u^* A_n^{-1/2} A_{n-1} A_n^{-1/2} u \leq \lambda_{\max}(A_n^{-1}A_{n-1}),$$

we can include this set in a larger one that no longer depends on u , namely,

$$\bigcup_{u \in \mathbb{C}^n, \|u\|=1} \Delta(u) \subseteq \left\{ z \in \mathbb{C} : |z| \leq R^{1/3} \right\} \cup \left(\bigcup_{\lambda_{\min}(A_n^{-1}A_{n-1}) \leq t \leq \lambda_{\max}(A_n^{-1}A_{n-1})} \left\{ z \in \mathbb{C} : |z+t| \leq R^{1/3} \right\} \right),$$

which is precisely the eigenvalue inclusion set defined in the statement of the theorem. \square

As before, we immediately have the following corollary.

Corollary 4.2. Let $P(z) = \sum_{j=0}^n A_j z^j$ be a matrix polynomial of degree $n \geq 3$ with positive definite coefficients, define

$$\alpha := \|A_n^{-1} A_{n-2}\|$$

$$\beta := \max \left\{ 0, \max_{0 \leq j \leq n-3} \left\| \left(A_j - \|A_{n-2}^{-1} A_n\|^{-1} A_{j+2} \right) A_{j+3}^{-1} \right\| \right\},$$

and let R be the unique positive solution of $z^3 - \alpha z - \beta = 0$. Then all the eigenvalues of P lie in the set

$$\left\{ z \in \mathbb{C} : |z| \leq R^{1/3} \right\} \cup \left(\bigcup_{\|A_{n-1}^{-1} A_n\|^{-1} \leq t \leq \|A_n^{-1} A_{n-1}\|} \left\{ z \in \mathbb{C} : |z + t| \leq R^{1/3} \right\} \right).$$

We note that, compared to the other computations necessary to obtain the inclusion sets, the computational effort to find the unique positive zero of the scalar cubic in Theorem 4.3 and Corollary 4.2 is negligible.

It is by now clear that additional results similar to the ones we have obtained can be produced by choosing an appropriate multiplier for $u^* A^{-1/2} P(z) A^{-1/2} u$ and combining it with one of the sets from Theorem 4.1.

5 Numerical results

Although our main focus was to obtain extended versions of the Eneström-Kakeya theorem for matrix polynomials without worrying too much about their implementation, it is nonetheless worthwhile to consider some numerical examples, especially given their rarity in the literature.

Figure 1 shows what typical inclusion regions look like for a random positive definite matrix polynomial of degree $n = 6$ and coefficient size $m = 50$, where the ratio of the largest to the smallest eigenvalue for each coefficient matrix was approximately 100. The figure shows the inclusion disks from Theorems 2.4 and 3.3 in solid and dashed line, respectively, and the one from Theorem 4.3 (the smallest) in solid line. The eigenvalues are indicated by black asterisks. A scalar or matrix polynomial of degree n with positive or positive definite coefficients, respectively, has no zeros or eigenvalues, respectively, in the wedge

$$\left\{ z \in \mathbb{C} : -\frac{\pi}{n} \leq \text{Arg}(z) \leq \frac{\pi}{n} \right\}.$$

To avoid clutter, this wedge is not shown in the figure.

It is, in general, difficult to predict which eigenvalue inclusion regions from the previous sections are preferable over others, and the the numerical examples here are merely meant to illustrate those results. It would not be possible to compare all possible combinations of all possible variations of the inclusion regions that can be obtained. Instead, we have chosen a few key comparisons that should be sufficient to show the strengths and weaknesses of the various eigenvalue inclusion regions.

To that end, we generated 1000 random positive definite matrix polynomials of two types: TYPE 1 with coefficients A for which $\lambda_{\max}(A)/\lambda_{\min}(A)$ is typically of order 10 and TYPE 2 for which this ratio is typically of order 10^5 . We achieved this by generating matrices with elements whose real and complex parts are uniformly randomly distributed on the interval $[-4, 4]$, and then forming $A^* A$ for each such matrix A . Finally, we added a multiple of the identity matrix for TYPE 1, but not for TYPE 2. The degree of the polynomials was arbitrarily chosen to be $n = 10$ and the size of the coefficient matrices $m = 20$, as the trends we observed did not vary much with those values.

The following tables compare the radii of the Cauchy-related and Eneström-Kakeya -related disks, obtained from the 1000 randomly generated matrix polynomials: Table 1 lists the averages of the ratios of the radii of the disks that define the inclusion regions in Theorem 2.3 versus those in Theorem 2.4 for the first column (denoted by $\Gamma_1(1)$), in Theorem 3.2 versus Theorem 3.3 for the second column (denoted by $\Gamma_2(1)$), and in Theorem 4.2 versus Theorem 4.3 for the third column (denoted by $\Gamma_1(2)/\Gamma_1(3)$), respectively. Table 2 is constructed in the same way for the corresponding corollaries that are based on norms, rather than eigenvalues. We used the easily computed 1-norm for all matrices.

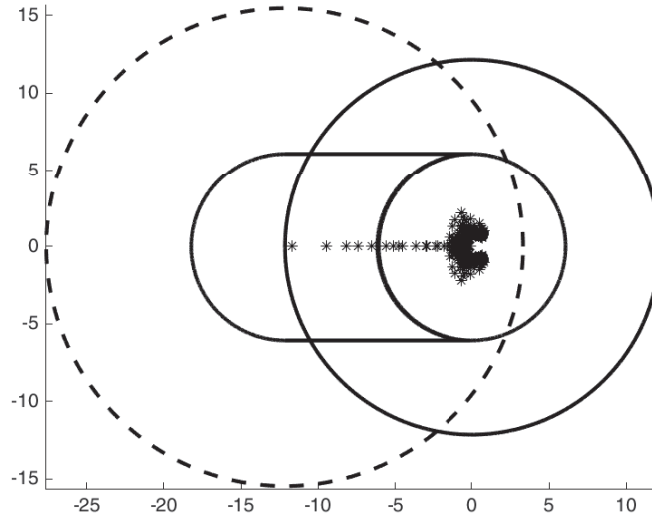


Figure 1: Eigenvalue inclusion regions from Theorems 2.4, 3.3, and 4.3.

Table 1: Comparison of inclusion regions based on eigenvalues.

	$\Gamma_1(1)$	$\Gamma_2(1)$	$\Gamma_1(2)/\Gamma_1(3)$
TYPE 1	0.98	1.03	1.23
TYPE 2	0.27	0.86	1.56

Table 2: Comparison of inclusion regions based on norms.

	$\Gamma_1(1)$	$\Gamma_2(1)$	$\Gamma_1(2)/\Gamma_1(3)$
TYPE 1	0.88	1.01	1.32
TYPE 2	0.26	0.89	1.58

A number smaller than one indicates that the Cauchy-related result is better. One observes that Cauchy-related results are better when the eigenvalues of the coefficient matrices exhibit considerable variance, and the last column most clearly shows the advantage of the Eneström-Kakeya -related Theorem 4.3 and Corollary 4.2, where such variance has less of an impact.

Conclusion. We have generalized improved versions of the scalar Eneström-Kakeya theorem to positive definite matrix polynomials, by combining polynomial multipliers with Cauchy-like zero inclusion regions for scalar polynomials, which establishes a clear framework for the derivation of additional results. In the process we have also generalized several Cauchy-like zero inclusion regions from scalar to matrix polynomials.

References

- [1] T. Betcke, N.J. Higham, V. Mehrmann, C. Schröder, C., F. Tisseur, *NLEVP: a collection of nonlinear eigenvalue problems*. ACM Trans. Math. Software, 39 (2013), no. 2, Art. 7, 28 pp.
- [2] D.A. Bini, V. Noferini, M. Sharify, *Locating the eigenvalues of matrix polynomials*. SIAM J. Matrix Anal. Appl., 34 (2013), 1708–1727.

- [3] A.L. Cauchy, *Sur la résolution des équations numériques et sur la théorie de l'élimination*. Exercices de Mathématiques, Quatrième Année, p.65–128. de Bure frères, Paris, 1829. Also in: Oeuvres Complètes, Série 2, Tome 9, 86–161. Gauthiers-Villars et fils, Paris, 1891.
- [4] G. Dirr, H.K. Wimmer, *An Eneström-Kekeya theorem for Hermitian polynomial matrices*. IEEE Trans. Automat. Control 52 (2007), 2151–2153.
- [5] G. Eneström, *Härledning af en allmän formel for antalet pensionärer vid en godtycklig tidpunkt förefinnas inom en sluten pensionskassa*. Öfversigt af Kungl. Vetenskap-Akademiens Förhandlingar (Stockholm), 50 (1893), 405–415.
- [6] G. Eneström, *Remarque sur un théorème relatif aux racines de l'équation $a_n x^n + \dots + a_0$ où tous les coefficients sont réels et positifs*. Tôhoku Math. J., 18 (1920), 34–36. Translation of reference [5].
- [7] N.J. Higham, F. Tisseur, *Bounds for eigenvalues of matrix polynomials*. Linear Algebra Appl., 358 (2003), 5–22.
- [8] R.A. Horn, C.R. Johnson, *Matrix Analysis*. Cambridge University Press, Cambridge, 2013.
- [9] S. Kekeya, *On the limits of the roots of an algebraic equation with positive coefficients*. Tôhoku Math. J., 2 (1912), 140–142.
- [10] C.-T. Le, T.-H.-B. Du, T.-D. Nguyen *On the location of eigenvalues of matrix polynomials*. Operators and Matrices, to appear. Available as arXiv:1703.00747 (2017).
- [11] M. Marden, *Geometry of polynomials*. Second edition. Mathematical Surveys, No. 3, American Mathematical Society, Providence, R.I., 1966.
- [12] A. Melman, *The twin of a theorem by Cauchy*. Amer. Math. Monthly, 120 (2013), 164–168.
- [13] A. Melman, *Generalization and variations of Pellet's theorem for matrix polynomials*. Linear Algebra Appl., 439 (2013), 1550–1567.
- [14] A. Melman, *Geometric aspects of Pellet's and related theorems*. Rocky Mountain J. Math., 45 (2015), 603–621.
- [15] A. Melman, *Extensions of the Eneström-Kekeya theorem*. Elemente der Mathematik, to appear. Available as arXiv:1711.09517v6
- [16] G.V. Milovanović, D.S. Mitrinović, Th.M. Rassias, *Topics in polynomials: extremal problems, inequalities, zeros*. World Scientific Publishing Co., Inc., River Edge, NJ, 1994.
- [17] Q.I. Rahman, G. Schmeisser, *Analytic Theory of Polynomials* London Mathematical Society Monographs. New Series, 26. The Clarendon Press, Oxford University Press, Oxford, 2002.