PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 146, Number 2, February 2018, Pages 613–624
http://dx.doi.org/10.1090/proc/13826
Article electronically published on October 18, 2017

IMPROVED CAUCHY RADIUS FOR SCALAR AND MATRIX POLYNOMIALS

A. MELMAN

(Communicated by Walter Van Assche)

ABSTRACT. We improve the Cauchy radius of both scalar and matrix polynomials, which is an upper bound on the moduli of the zeros and eigenvalues, respectively, by using appropriate polynomial multipliers.

1. INTRODUCTION

A simple but classical result from 1829 due to Cauchy ([2], [5, Th.(27,1), p. 122 and Exercise 1, p. 126]) states that the zeros of a polynomial $p(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, with complex coefficients and $a_n \neq 0$, lie in $|z| \leq \rho[p]$, where $\rho[p]$ is the *Cauchy radius* of p, namely, the unique positive solution of

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

A smaller Cauchy radius was obtained much more recently in [8, Theorem 8.3.1] by Rahman and Schmeisser, who showed that $\rho[(a_n z^k - a_{n-k})p(z)] \leq \rho[p]$, where k is the smallest positive integer such that $a_{n-k} \neq 0$, i.e., a better bound can be found by using a polynomial multiplier.

A generalization to matrix polynomials of Cauchy's classical bound for scalar polynomials was derived in [1], [4], and [6]. It states that all the eigenvalues of the regular matrix polynomial $P(z) = A_n z^n + A_{n-1}z^{n-1} + \cdots + A_1 z + A_0$, with complex coefficient matrices and A_n nonsingular, lie in $|z| \leq \rho[P]$, where, as in the scalar case, $\rho[P]$ is called the *Cauchy radius* of P, which is the unique positive solution of

$$\|A_n^{-1}\|^{-1} z^n - \|A_{n-1}\| z^{n-1} - \cdots - \|A_1\| z - \|A_0\| = 0$$

for any matrix norm. The eigenvalues of P are the complex numbers z for which a nonzero complex vector v exists such that $P(z)v = 0$. If A_n is nonsingular, they are the solutions of $\det P(z) = 0$. A matrix polynomial P is *regular* if $\det P$ is not identically zero. When P is linear and monic, i.e., $P(z) = Iz + A_0$, one obtains the standard eigenvalue problem.

In [7], the improved Cauchy radius from [8] was also generalized to matrix polynomials. It was shown there that, under mild conditions on A_n and with k the smallest positive integer such that A_{n-k} is not the null matrix, both

$$\rho\left[\left(A_n z^k - A_{n-k}\right) P(z)\right] \leq \rho\,[P] \quad \text{and} \quad \rho\left[P(z)\left(A_n z^k - A_{n-k}\right)\right] \leq \rho\,[P] \; .$$

Received by the editors November 17, 2016 and, in revised form, March 15, 2017.
2010 *Mathematics Subject Classification.* Primary 30C15, 47A56, 65F15.
Key words and phrases. Bound, zero, root, polynomial, eigenvalue, multiplier, matrix polynomial, Cauchy radius.

Other multipliers with these properties do not seem to exist in the literature, and our purpose here is to derive different multipliers that also improve the Cauchy radius for both scalar and matrix polynomials and that, in general, perform better than the improvements from [7] and [8].

In Section 2 we present such polynomial multipliers first for matrix polynomials, while we consider scalar polynomials as a special case in Section 3.

2. Improved Cauchy radius for matrix polynomials

The following theorem presents three matrix polynomials, obtained by multiplying a given matrix polynomial P by another matrix polynomial, that have a smaller Cauchy radius than that of P. Clearly, for any matrix polynomial T, a region in the complex plane containing all the eigenvalues of TP or PT also contains those of P.

Theorem 2.1. *Let $P(z) = \sum_{j=0}^{n} A_j z^j$ be a regular matrix polynomial of degree n that is at least a trinomial, with square complex matrices A_j $(0 \leq j \leq n)$ and A_n nonsingular, and let k and ℓ be the smallest positive integers such that A_{n-k} and $A_{n-k-\ell}$ are not the null matrix. Define*

$$Q_1^{(L)}(z) = \left(A_n z^{k+\ell} - A_{n-k} z^{\ell} - A_{n-k-\ell}\right) P(z),$$
$$Q_2^{(L)}(z) = \left(A_n z^{2k} - A_{n-k} z^{k} + A_{n-k}^2 A_n^{-1}\right) P(z),$$

and, when $\ell = k$,

$$Q_3^{(L)}(z) = \left(A_n z^{2k} - A_{n-k} z^{k} - A_{n-2k} + A_{n-k}^2 A_n^{-1}\right) P(z).$$

Furthermore, define

$$Q_1^{(R)}(z) = P(z) \left(A_n z^{k+\ell} - A_{n-k} z^{\ell} - A_{n-k-\ell}\right),$$
$$Q_2^{(R)}(z) = P(z) \left(A_n z^{2k} - A_{n-k} z^{k} + A_{n-k}^2 A_n^{-1}\right),$$

and, when $\ell = k$,

$$Q_3^{(R)}(z) = P(z) \left(A_n z^{2k} - A_{n-k} z^{k} - A_{n-2k} + A_{n-k}^2 A_n^{-1}\right).$$

For any matrix norm $\|.\|$, if $\|A_n^{-2}\|^{-1} = \|A_n\| \|A_n^{-1}\|^{-1}$ and if $A_n A_{n-k} = A_{n-k} A_n$ and $A_n A_{n-k-\ell} = A_{n-k-\ell} A_n$, then, for any admissible values of k and ℓ, it follows that $\rho[Q_j^{(L)}] \leq \rho[P]$ for $j = 1, 2$. If $\ell = k$, then $\rho[Q_3^{(L)}] \leq \rho[P]$. Analogous results hold for $Q_j^{(R)}$ $(j = 1, 2, 3)$.

Proof. We prove the theorem for $Q_j^{(L)}$ $(j = 1, 2, 3)$; the proof for $Q_j^{(R)}$ $(j = 1, 2, 3)$ is analogous. Let k and ℓ be as in the statement of the theorem and, if it exists, let s be the first positive integer such that $A_{n-k-\ell-s}$ is not the null matrix (when P is a trinomial, then no such s exists). We will make use of the following expression, where m is any positive integer and M is any square complex matrix:

$$\begin{aligned} &\left(A_n z^{k+m} - A_{n-k} z^{m} - M\right) P(z) \\ &= A_n^2 z^{n+k+m} + A_n A_{n-k-\ell} z^{n-\ell+m} - A_{n-k}^2 z^{n-k+m} - M A_n z^n \\ &\qquad + A_n z^{k+m} \sum_{j=0}^{n-k-\ell-s} A_j z^j - A_{n-k} z^m \sum_{j=0}^{n-k-\ell} A_j z^j - M \sum_{j=0}^{n-k} A_j z^j \,. \end{aligned} \tag{2.1}$$

If P is a trinomial, then the summation in (2.1) with upper index limit $n-k-\ell-s$ is set equal to zero.

We begin with $Q_1^{(L)}$, which is obtained by setting $m = \ell$ and $M = A_{n-k-\ell}$ in (2.1):

$$(2.2)\qquad \left(A_n z^{k+\ell} - A_{n-k} z^{\ell} - A_{n-k-\ell}\right) P(z) = Q_1^{(L)}(z) = A_n^2 z^{n+k+\ell} + S(z),$$

where

$$(2.3)\quad S(z) = -A_{n-k}^2 z^{n-k+\ell} + A_n z^{k+\ell} \sum_{j=0}^{n-k-\ell-s} A_j z^j - A_{n-k} z^{\ell} \sum_{j=0}^{n-k-\ell} A_j z^j - A_{n-k-\ell} \sum_{j=0}^{n-k} A_j z^j = \sum_{j=0}^{\nu} B_j z^j,$$

$\nu \le n - \min\{s, k-\ell\}$ (when P is a trinomial, $\nu \le 2(n-k) = 2\ell$, since in this case $k+\ell = n$ so that $k - \ell = 2k - n$), and each matrix B_j is a sum of terms of the form A_j or $A_i A_j$. If we define

$$\Phi(z) = \sum_{j=0}^{\nu} \|B_j\| z^j,$$

then the Cauchy radius of $Q_1^{(L)}$ is the unique positive solution of $\|A_n^{-2}\|^{-1} z^{n+k+\ell} - \Phi(z) = 0$. We now set $x = \rho[P]$, i.e., x satisfies

$$(2.4)\qquad \|A_n^{-1}\|^{-1} x^n - \|A_{n-k}\| x^{n-k} - \|A_{n-k-\ell}\| x^{n-k-\ell} - \sum_{j=0}^{n-k-\ell-s} \|A_j\| x^j = 0.$$

Using (2.4) and the basic properties $\|A+B\| \le \|A\| + \|B\|$ and $\|AB\| \le \|A\|\|B\|$ of matrix norms, we have that

$$\begin{aligned}(2.5)\quad \Phi(x) &\le \|A_{n-k}\|^2 x^{n-k+\ell} + \|A_n\| x^{k+\ell} \sum_{j=0}^{n-k-\ell-s} \|A_j\| x^j \\ &\quad + \|A_{n-k}\| x^{\ell} \sum_{j=0}^{n-k-\ell} \|A_j\| x^j + \|A_{n-k-\ell}\| \sum_{j=0}^{n-k} \|A_j\| x^j \\ &= \|A_{n-k}\|^2 x^{n-k+\ell} + \|A_n\| x^{k+\ell} \left(\|A_n^{-1}\|^{-1} x^n - \|A_{n-k}\| x^{n-k} - \|A_{n-k-\ell}\| x^{n-k-\ell}\right) \\ &\quad + \|A_{n-k}\| x^{\ell} \left(\|A_n^{-1}\|^{-1} x^n - \|A_{n-k}\| x^{n-k}\right) + \|A_{n-k-\ell}\| \left(\|A_n^{-1}\|^{-1} x^n\right) \\ &\le \|A_n\| \|A_n^{-1}\|^{-1} x^{n+k+\ell} = \|A_n^{-2}\|^{-1} x^{n+k+\ell}.\end{aligned}$$

We have used both the fact that $\|A_n^{-1}\|^{-1} \le \|A_n\|$ and our assumption that $\|A_n^{-2}\|^{-1} = \|A_n\|\|A_n^{-1}\|^{-1}$. This means that $\|A_n^{-2}\|^{-1} x^{n+k+\ell} - \Phi(x) \ge 0$ and, therefore, that x must lie to the right of $\rho[Q_1^{(L)}]$, i.e., $\rho[Q_1^{(L)}] \le \rho[P]$.

When P is a trinomial, the second term in the right-hand side of (2.5) is absent, and the result follows analogously.

For $Q_2^{(L)}$ the proof is similar and we will omit unnecessary details. Here we set $m = k$ and $M = -A_{n-k}^2 A_n^{-1}$ in (2.1) to obtain

$$(2.6)\qquad \left(A_n z^{2k} - A_{n-k} z^k + A_{n-k}^2 A_n^{-1}\right) P(z) = Q_2^{(L)}(z) = A_n^2 z^{n+2k} + S(z),$$

where

$$(2.7)\quad S(z) = A_n A_{n-k-\ell} z^{n-\ell+k} + A_n z^{2k} \sum_{j=0}^{n-k-\ell-s} A_j z^j - A_{n-k} z^k \sum_{j=0}^{n-k-\ell} A_j z^j + A_{n-k}^2 A_n^{-1} \sum_{j=0}^{n-k} A_j z^j = \sum_{j=0}^{\nu} B_j z^j,$$

and $\nu = n - \min\{k, \ell - k\}$ (when P is a trinomial, $\nu = n - \min\{k, n-2k\}$). The Cauchy radius of $Q_2^{(L)}$ is the unique positive solution of $\|A_n^{-2}\|^{-1} z^{n+2k} - \Phi(z) = 0$, where $\Phi(z) = \sum_{j=0}^{\nu} \|B_j\| z^j$. With $x = \rho[P]$ and (2.4), we have

$$\begin{aligned}\Phi(x) \quad \leq \quad & \|A_n\|\|A_{n-k-\ell}\| x^{n-\ell+k} + \|A_n\| x^{2k} \sum_{j=0}^{n-k-\ell-s} \|A_j\| x^j \\ & + \|A_{n-k}\| x^k \sum_{j=0}^{n-k-\ell} \|A_j\| x^j + \|A_{n-k}\|^2 \|A_n^{-1}\| \sum_{j=0}^{n-k} \|A_j\| x^j \\ = \; & \|A_n\|\|A_{n-k-\ell}\| x^{n-\ell+k} \\ & + \|A_n\| x^{2k} \left(\|A_n^{-1}\|^{-1} x^n - \|A_{n-k}\| x^{n-k} - \|A_{n-k-\ell}\| x^{n-k-\ell} \right) \\ & + \|A_{n-k}\| x^k \left(\|A_n^{-1}\|^{-1} x^n - \|A_{n-k}\| x^{n-k} \right) + \|A_{n-k}\|^2 \|A_n^{-1}\| \left(\|A_n^{-1}\|^{-1} x^n \right) \\ \leq \; & \|A_n\| \|A_n^{-1}\|^{-1} x^{n+2k} = \|A_n^{-2}\|^{-1} x^{n+2k}.\end{aligned}$$

Therefore, $\|A_n^{-2}\|^{-1} x^{n+2k} - \Phi(x) \geq 0$, implying that the Cauchy radius of $Q_2^{(L)}$ is smaller than that of P. When P is a trinomial, the same result follows analogously as before.

For $Q_3^{(L)}$, with $\ell = k$, we set $m = k = \ell$ and $M = A_{n-2k} - A_{n-k}^2 A_n^{-1}$ in (2.1), which gives

$$(2.8)\quad \left(A_n z^{2k} - A_{n-k} z^k - A_{n-2k} + A_{n-k}^2 A_n^{-1} \right) P(z) = Q_3^{(L)}(z) = A_n^2 z^{n+2k} + S(z),$$

where

$$(2.9)\quad S(z) = A_n z^{2k} \sum_{j=0}^{n-2k-s} A_j z^j - A_{n-k} z^k \sum_{j=0}^{n-2k} A_j z^j - \left(A_{n-2k} - A_{n-k}^2 A_n^{-1} \right) \sum_{j=0}^{n-k} A_j z^j = \sum_{j=0}^{\nu} B_j z^j,$$

and $\nu \leq n - \min\{k, s\} \leq n - 1$ ($\nu \leq n - k \leq n - 1$ when P is a trinomial). With $\Phi(z) = \sum_{j=0}^{\nu} \|B_j\| z^j$, the Cauchy radius of $Q_3^{(L)}$ is the unique positive solution of $\|A_n^{-2}\|^{-1} \; z^{n+2k} - \Phi(z) = 0$. Setting $x = \rho[P]$, which satisfies equation (2.4), we have

$$\begin{aligned}
\Phi(x) &\leq \|A_n\|x^{2k}\sum_{j=0}^{n-2k-s}\|A_j\|x^j + \|A_{n-k}\|x^k\sum_{j=0}^{n-2k}\|A_j\|x^j \\
&\qquad\qquad +\|A_{n-2k}-A_{n-k}^2A_n^{-1}\|\sum_{j=0}^{n-k}\|A_j\|x^j \\
&= \|A_n\|x^{2k}\left(\|A_n^{-1}\|^{-1}x^n - \|A_{n-k}\|x^{n-k} - \|A_{n-2k}\|x^{n-2k}\right) \\
&\qquad +\|A_{n-k}\|x^k\left(\|A_n^{-1}\|^{-1}x^n - \|A_{n-k}\|x^{n-k}\right) \\
&\qquad\qquad +\|A_{n-2k}-A_{n-k}^2A_n^{-1}\|\left(\|A_n^{-1}\|^{-1}x^n\right) \\
&\leq \|A_n\|\|A_n^{-1}\|^{-1}x^{n+2k} \\
&\qquad +\left(\|A_{n-2k}-A_{n-k}^2A_n^{-1}\|\|A_n^{-1}\|^{-1} - \|A_n\|\|A_{n-2k}\| - \|A_{n-k}\|^2\right)x^n \\
&\leq \|A_n\|\|A_n^{-1}\|^{-1}x^{n+2k} \\
&\qquad +\left(\|A_{n-2k}\|\|A_n^{-1}\|^{-1} + \|A_{n-k}\|^2\|A_n^{-1}\|\|A_n^{-1}\|^{-1}\right. \\
&\qquad\qquad \left.-\|A_n\|\|A_{n-2k}\| - \|A_{n-k}\|^2\right)x^n \\
&\leq \|A_n\|\|A_n^{-1}\|^{-1}x^{n+2k} + \left(\|A_{n-2k}\|\|A_n^{-1}\|^{-1} - \|A_n\|\|A_{n-2k}\|\right)x^n \\
&\leq \|A_n\|\|A_n^{-1}\|^{-1}x^{n+2k} = \|A_n^{-2}\|^{-1}x^{n+2k}.
\end{aligned}$$

We have obtained that $\|A_n^{-2}\|^{-1}x^{n+2k} - \Phi(x) \geq 0$, with an analogous result when P is a trinomial, which means that the Cauchy radius of $Q_3^{(L)}$ is smaller than that of P. This completes the proof. □

Remarks.

- The matrix $A_{n-k}^2A_n^{-1}$ in the definitions of $Q_j^{(L)}$ and $Q_j^{(R)}$ for $j = 2, 3$ could be replaced by $A_n^{-1}A_{n-k}^2$ since positive and negative powers (if they exist) of commuting matrices also commute.
- If $A_{n-2k} = A_{n-k}^2A_n^{-1}$, then $\rho[Q_3^{(L)}]$ and $\rho[Q_3^{(R)}]$ are both equal to the improved Cauchy radius of Theorem 8.3.1 in [8].
- The conditions $A_nA_{n-k} = A_{n-k}A_n$, $A_nA_{n-k-\ell} = A_{n-k-\ell}A_n$, and $\|A_n^{-2}\|^{-1} = \|A_n\|\|A_n^{-1}\|^{-1}$ may appear restrictive, but they are always satisfied if $A_n = I$ and $\|I\| = 1$. The former can be achieved by multiplying P by A_n^{-1}, which needs to be computed anyway to obtain the Cauchy radius.
- In general, the multipliers of P are different from the ones obtained by repeatedly using Theorem 2.2 in [7], as their degrees can easily be seen to be different.

The more zero coefficients a polynomial has, all else being the same, the smaller its Cauchy radius will be. Although the matrix polynomials $Q_j^{(L)}$ for $j = 1, 2, 3$ may have additional zero coefficients (null matrices), the ones that we have some control over are the leading zeros immediately following the highest coefficient.

The following lemma allows us to compare their number, thereby indicating which multiplier might be preferable for given values of k and ℓ.

Lemma 2.2. *Let P, $Q_j^{(L)}$ and $Q_j^{(R)}$ $(j = 1, 2, 3)$ be as in Theorem 2.1. Then, with the same notation as before, the following holds:*

- *When $\ell < k$, the leading powers of z in $Q_1^{(L)}$ are $n + k + \ell$ and $\nu \leq n - 1$, whereas for $Q_2^{(L)}$ they are $n + 2k$ and $n + k - \ell \geq n + 1$.*
- *When $\ell > k$, the leading powers of z in $Q_1^{(L)}$ are $n+k+\ell$ and $n+\ell-k \geq n+1$, whereas for $Q_2^{(L)}$ they are $n + 2k$ and $\nu \leq n - 1$.*
- *When $\ell = k$, the leading powers of z in $Q_1^{(L)}$ and $Q_2^{(L)}$ are $n + 2k$ and n, whereas for $Q_3^{(L)}$ they are $n + 2k$ and $\nu \leq n - 1$.*
- *All of the above results also hold true for $Q_j^{(R)}$ $(j = 1, 2, 3)$.*

Proof. From (2.2) and (2.3) we have that, when $\ell < k$, then the leading powers of $Q_1^{(L)}$ are $n + k + \ell$ and $\nu \leq n - \min\{k - \ell, s\} \leq n - 1$ (as before, when P is a trinomial, we obtain that $\nu \leq 2(n - k) = 2\ell \leq n - 1$). Here, n, k, ℓ, s, and ν are as before. When $\ell > k$, then those powers become $n + k + \ell$ and $n + \ell - k \geq n + 1$, and when $\ell = k$, they are $n + 2k$ and n.

Similarly, we observe from (2.6) and (2.7) that, when $\ell < k$, the leading powers of $Q_2^{(L)}$ are $n + 2k$ and $n + k - \ell \geq n + 1$, whereas for $\ell > k$, they are $n + 2k$ and $\nu = n - \min\{k, \ell - k\} \leq n - 1$. When $\ell = k$, those powers become $n + 2k$ and n, as for $Q_1^{(L)}$.

When $\ell = k$, equations (2.8) and (2.9) show that the highest powers of $Q_3^{(L)}$ are $n + 2k$ and $\nu \leq n - \min\{k, s\} \leq n - 1$ (when P is a trinomial, $\nu \leq n - k \leq n - 1$). The proof for $Q_j^{(R)}$ $(j = 1, 2, 3)$ is analogous. □

The number of leading zero coefficients is now easily determined with Lemma 2.2 from the leading powers of $Q_j^{(L)}$ for $j = 1, 2, 3$. They can be found on the left in Table 1 for the worst case (i.e., smallest number of zeros), namely, when $\nu = n-1$, where ν is as in Lemma 2.2, while the degrees of $Q_j^{(L)}$ for $j = 1, 2, 3$ can be found on the right.

Table 1. Number of zero coefficients (left) and degrees of $Q_j^{(L)}$ for $j = 1, 2, 3$ (right).

	$Q_1^{(L)}$	$Q_2^{(L)}$	$Q_3^{(L)}$
$\ell < k$	$k + \ell$	$k + \ell - 1$	-
$\ell > k$	$2k - 1$	$2k$	-
$\ell = k$	$2k - 1$	$2k - 1$	$2k$

	$Q_1^{(L)}$	$Q_2^{(L)}$	$Q_3^{(L)}$
$\ell < k$	$n + k + \ell$	$n + 2k$	-
$\ell > k$	$n + k + \ell$	$n + 2k$	-
$\ell = k$	$n + 2k$	$n + 2k$	$n + 2k$

Table 1 shows that, when $\ell < k$, $Q_1^{(L)}$ has a higher number of leading zero coefficients than $Q_2^{(L)}$, while its degree is lower. When $\ell > k$, the same conclusion holds with $Q_1^{(L)}$ and $Q_2^{(L)}$ trading places, and when $\ell = k$, then $Q_3^{(L)}$ has more such zero coefficients than both $Q_1^{(L)}$ and $Q_2^{(L)}$, while they all have the same degree.

Analogous results are obtained for $Q_j^{(R)}$ for $j = 1, 2, 3$. We thus arrive at the following choice to improve the Cauchy radius of P:

$$Q^{(L)}(z) = \begin{cases} \left(A_n z^{k+\ell} - A_{n-k} z^{\ell} - A_{n-k-\ell}\right) P(z) & \text{if } \ell < k, \\ \left(A_n z^{2k} - A_{n-k} z^{k} + A_{n-k}^2 A_n^{-1}\right) P(z) & \text{if } \ell > k, \\ \left(A_n z^{2k} - A_{n-k} z^{k} - A_{n-2k} + A_{n-k}^2 A_n^{-1}\right) P(z) & \text{if } \ell = k, \end{cases} \tag{2.10}$$

and we choose $Q^{(R)}$ analogously.

Remarks.

- Theorem 2.1 can be applied recursively to improve the Cauchy radius further. One could also alternate between (L) and (R) versions, although, in general, there does not seem to be a large difference between the two.
- The improved Cauchy radii require additional matrix multiplications, while a real scalar polynomial equation of a degree higher than that of P needs to be solved. The latter can be dealt with very efficiently so that, as the matrix size increases, the cost tends to be dominated by the matrix multiplications. It therefore depends on the application if this additional computational cost is justified.
- The choice of $Q^{(L)}$ or $Q^{(R)}$, which was based on the number of leading zeros, is not guaranteed to produce better results than other choices, although the numerical examples below seem to indicate that it performs well.
- It is, in general, difficult to predict which norm provides the best result, but in many applications the size of the matrix coefficients limits that choice to the 1-norm or the ∞-norm.

We illustrate the usefulness of Theorem 2.1 and our choice of $Q^{(L)}$, defined by (2.10), and compare it to Theorem 2.2 from [7] (the generalization to matrix polynomials of Theorem 8.3.1 in [8]) at the hand of the following two examples. In the first, we generate random matrix polynomials, whereas the second one is taken from the engineering literature.

Example 1. Here we generated 1000 matrix polynomials with complex elements, whose real and complex parts are uniformly randomly distributed on the interval $[-10, 10]$. We then premultiplied each matrix polynomial by the inverse of its leading coefficient to make its leading coefficient the identity matrix. We examined four cases with $n = 20$ and 25×25 coefficients: $k = 3, \ell = 5$, $k = 5, \ell = 3$, $k = \ell = 5$, and $k = \ell = 1$, and one case with $n = 4$, 250×250 coefficients, and $k = \ell = 1$. Tables 2 and 3 list the averages of the ratios of the Cauchy radii to the modulus of the largest eigenvalue, i.e., the closer this number is to 1, the better it is. This was done for the Cauchy radius of the given matrix polynomial with the 1-norm and five consecutive applications of Theorem 2.1, labeled as level 1-5, using $Q^{(L)}$ defined by (2.10) for each application. In each column, the numbers on the left are the ratios obtained by Theorem 2.1, while the ones on the right are the ratios from Theorem 2.2 in [7]. Clearly, significant improvements can be obtained from Theorem 2.1. Moreover, the advantage of having another multiplier in addition to the one from [7] is that it can sometimes accelerate an otherwise slowly progressing recursion.

TABLE 2. Comparison of Cauchy radii for Example 1.

Level	$n = 20$, $m = 25$ $k = 3$, $\ell = 5$		$n = 20$, $m = 25$ $k = 5$, $\ell = 3$		$n = 20$, $m = 25$ $k = \ell = 5$	
Cauchy	1.991		1.482		1.492	
1	1.257	1.404	1.236	1.264	1.165	1.231
2	1.135	1.198	1.155	1.235	1.151	1.145
3	1.127	1.190	1.145	1.217	1.146	1.358
4	1.123	1.186	1.117	1.152	1.093	1.130
5	1.118	1.184	1.070	1.145	1.087	1.126

TABLE 3. Comparison of Cauchy radii for Example 1.

Level	$n = 20$, $m = 25$ $k = \ell = 1$		$n = 4$, $m = 250$ $k = \ell = 1$	
Cauchy	8.442		33.963	
1	2.003	2.880	3.154	5.725
2	1.419	1.770	1.763	2.419
3	1.237	1.681	1.361	2.350
4	1.195	1.366	1.326	1.574
5	1.194	1.328	1.326	1.543

Example 2. This example is taken from [3], where a structural dynamics model representing a reinforced concrete machine foundation is formulated as a sparse quadratic 3627×3627 eigenvalue problem with $k = \ell = 1$. Of the many bounds on the eigenvalues that were examined in [4] for the 1-norm and ∞-norm for this problem (the 2-norm is too costly here), the Cauchy radius was among the best. Theorem 2.2 in [7] improves those bounds significantly, but Theorem 2.1 improves them even more. Table 4 shows the actual Cauchy radius and its improvements from five successive applications of Theorem 2.2 in [7] and Theorem 2.1 for the 1-norm on the left and the ∞-norm on the right. In each column, the numbers on the left are obtained from Theorem 2.1, while those on the right are from Theorem 2.2 in [7]. Here too, we have used $Q^{(L)}$ defined in (2.10). The modulus of the largest eigenvalue is 2.120×10^4, and in the table all bounds were divided by 10^4.

TABLE 4. Comparison of Cauchy radii for Example 2 with the 1-norm (left) and the ∞-norm (right).

Cauchy	3.532	
Level 1	2.762	3.349
Level 2	2.427	2.737
Level 3	2.413	2.722
Level 4	2.272	2.425
Level 5	2.271	2.419

Cauchy	3.173	
Level 1	2.658	3.064
Level 2	2.380	2.652
Level 3	2.363	2.598
Level 4	2.260	2.380
Level 5	2.260	2.374

3. Improved Cauchy radius for scalar polynomials

Since scalar polynomials are 1×1 matrix polynomials, Theorem 2.1 can be applied to them as a special case. Moreover, because of their scalar nature, the theorem can be slightly refined, as stated in the following theorem.

Theorem 3.1. *Let $p(z) = \sum_{j=0}^{n} a_j z^j$ be a polynomial of degree n with complex coefficients that is at least a trinomial, and let k and ℓ be the smallest positive integers such that a_{n-k} and $a_{n-k-\ell}$ are not zero. Define*

$$q_1(z) = \left(a_n z^{k+\ell} - a_{n-k} z^{\ell} - a_{n-k-\ell}\right) p(z),$$
$$q_2(z) = \left(a_n z^{2k} - a_{n-k} z^k + \frac{a_{n-k}^2}{a_n}\right) p(z),$$

and, when $\ell = k$,

$$q_3(z) = \left(a_n z^{2k} - a_{n-k} z^k - a_{n-2k} + \frac{a_{n-k}^2}{a_n}\right) p(z).$$

Then the following holds:

(1) *For any admissible values of k and ℓ, $\rho[q_j] \leq \rho[p]$ for $j = 1, 2$, and if $\ell = k$, then $\rho[q_3] \leq \rho[p]$.*

(2) *If all the coefficients of p are nonzero, then the inequalities in part (1) are strict, unless p has a zero of modulus $\rho[p]$.*

Proof. The first part of the theorem follows immediately from Theorem 2.1 as a special case because complex numbers are 1×1 complex matrices. The second part requires some elaboration. To avoid tedious repetition, we present a detailed proof only for q_1, and sketch the proof for q_2 and q_3. Throughout, if the index of a quantity is inadmissible, then that quantity is set equal to zero.

We now assume that all the coefficients of p are nonzero, so that $k = \ell = 1$, and we begin with q_1. The expression corresponding to $S(z)$, defined by (2.3) in the proof of Theorem 2.1, is given by

$$\begin{aligned} S(z) &= -a_{n-1}^2 z^n + a_n z^2 \sum_{j=1}^{n-3} a_j z^j - a_{n-1} z \sum_{j=0}^{n-2} a_j z^j - a_{n-2} \sum_{j=0}^{n-1} a_j z^j \\ &= -a_{n-1}^2 z^n + \sum_{j=2}^{n-1} a_n a_{j-2} z^j - \sum_{j=1}^{n-1} a_{n-1} a_{j-1} z^j - \sum_{j=0}^{n-1} a_{n-2} a_j z^j \\ &= -a_{n-1}^2 z^n + \sum_{j=2}^{n-1} \left(a_n a_{j-2} - a_{n-1} a_{j-1} - a_{n-2} a_j\right) z^j \\ &\qquad - \left(a_{n-1} a_0 + a_{n-2} a_1\right) z - a_{n-2} a_0, \end{aligned}$$

while the expression corresponding to Φ becomes

$$\begin{aligned} \Phi(z) = |a_{n-1}|^2 z^n + \sum_{j=2}^{n-1} &\left|a_n a_{j-2} - a_{n-1} a_{j-1} - a_{n-2} a_j\right| z^j \\ &+ \left|a_{n-1} a_0 + a_{n-2} a_1\right| z + |a_{n-2} a_0|. \end{aligned}$$

For $x = \rho[p]$, the inequality corresponding to (2.5) is

$$\text{(3.1)} \quad \Phi(x) \leq |a_{n-1}|^2 x^n + \sum_{j=2}^{n-1} \left(|a_n a_{j-2}| + |a_{n-1} a_{j-1}| + |a_{n-2} a_j| \right) x^j + \left(|a_{n-1} a_0| + |a_{n-2} a_1| \right) x + |a_{n-2} a_0|.$$

The inequality in (3.1) is strict, unless

$$\text{(3.2)} \quad |a_n a_{j-2} - a_{n-1} a_{j-1} - a_{n-2} a_j| = |a_n a_{j-2}| + |a_{n-1} a_{j-1}| + |a_{n-2} a_j| \quad (j = 2, \ldots, n-1)$$

and

$$\text{(3.3)} \quad |a_{n-1} a_0 + a_{n-2} a_1| = |a_{n-1} a_0| + |a_{n-2} a_1|.$$

We now define $\varphi_j = \arg a_j$ and use $\varphi \cong \psi$ to indicate that φ and ψ only differ by an integer multiple of 2π, so that $e^{i\varphi} = e^{i\psi}$. If (3.2) and (3.3) hold, then we have from (3.2) for $j = 2, \ldots, n-1$, that

$$\text{(3.4)} \quad \varphi_n + \varphi_{j-2} \cong \varphi_{n-1} + \varphi_{j-1} + \pi \text{ , or } \varphi_{j-2} \cong \varphi_{j-1} + \varphi_{n-1} - \varphi_n + \pi,$$

and

$$\text{(3.5)} \quad \varphi_{n-1} + \varphi_{j-1} \cong \varphi_{n-2} + \varphi_j,$$

while from (3.3) we have

$$\text{(3.6)} \quad \varphi_{n-1} + \varphi_0 \cong \varphi_{n-2} + \varphi_1.$$

Combining (3.4) with the substitution $j = j-1$ in (3.5), we obtain for $j = 3, \ldots, n-1$ that

$$\varphi_{j-2} \cong \varphi_{j-1} + \varphi_{n-1} - \varphi_n + \pi \cong \varphi_{j-1} + \varphi_{n-2} - \varphi_{n-1},$$

which implies that

$$\text{(3.7)} \quad \varphi_{n-2} \cong 2\varphi_{n-1} - \varphi_n + \pi.$$

Substituting this in (3.6) shows that (3.6) is covered by (3.4). The expression in (3.4) is equivalent to

$$\text{(3.8)} \quad \varphi_{j-1} \cong \varphi_j + \varphi_{n-1} - \varphi_n + \pi \quad (j = 1, \ldots, n-2),$$

and we have obtained from (3.7) that (3.8) also holds for $j = n-1$. From here on, the proof follows that of Theorem 8.3.1.in [8]. As in that proof, the equations in (3.8), used recursively for $j = n-1, \ldots, 1$, yield

$$\varphi_{n-j} \cong \varphi_{n-1} + (j-1)\Delta \quad (j = 1, \ldots, n),$$

where $\Delta = \varphi_{n-1} - \varphi_n + \pi$, which is equivalent to

$$\text{(3.9)} \quad \varphi_j \cong (n-j)\Delta + \varphi_n - \pi \quad (j = 0, \ldots, n-1).$$

Using (3.9), we now show that, under these conditions, $xe^{i\Delta}$, where $x = \rho[p]$, is a zero of p:

$$\begin{aligned}\sum_{j=0}^{n} a_j (xe^{i\Delta})^j &= \sum_{j=0}^{n-1} |a_j| e^{i\varphi_j} x^j e^{ij\Delta} + |a_n| e^{i\varphi_n} x^n e^{in\Delta} \\ &= e^{i(\varphi_n + n\Delta)} \left(\sum_{j=0}^{n-1} |a_j| e^{i(\varphi_j - (n-j)\Delta - \varphi_n)} x^j + |a_n| x^n \right) \\ &= e^{i(\varphi_n + n\Delta)} \left(\sum_{j=0}^{n-1} e^{-i\pi} |a_j| x^j + |a_n| x^n \right) = 0,\end{aligned}$$

since $e^{-i\pi} = -1$, so that $xe^{i\Delta}$ is indeed a zero of p.

For q_2, we obtain for $S(z)$, defined in (2.7),

$$\begin{aligned}S(z) &= a_n a_{n-2} z^n + a_n z^2 \sum_{j=1}^{n-3} a_j z^j - a_{n-1} z \sum_{j=0}^{n-2} a_j z^j + \frac{a_{n-1}^2}{a_n} \sum_{j=0}^{n-1} a_j z^j \\ &= a_n a_{n-2} z^n + \sum_{j=2}^{n-1} a_n a_{j-2} z^j - \sum_{j=1}^{n-1} a_{n-1} a_{j-1} z^j + \sum_{j=0}^{n-1} \frac{a_{n-1}^2 a_j}{a_n} z^j \\ &= a_n a_{n-2} z^n + \sum_{j=2}^{n-1} \left(a_n a_{j-2} - a_{n-1} a_{j-1} + \frac{a_{n-1}^2 a_j}{a_n} \right) z^j \\ &\qquad + \left(-a_{n-1} a_0 + \frac{a_{n-1}^2 a_1}{a_n} \right) z + \frac{a_{n-1}^2 a_0}{a_n},\end{aligned}$$

and for q_3, we obtain, as in (2.9),

$$\begin{aligned}S(z) &= a_n z^2 \sum_{j=1}^{n-3} a_j z^j - a_{n-1} z \sum_{j=0}^{n-2} a_j z^j - \left(a_{n-2} - \frac{a_{n-1}^2}{a_n} \right) \sum_{j=0}^{n-1} a_j z^j \\ &= \sum_{j=2}^{n-1} a_n a_{j-2} z^j - \sum_{j=1}^{n-1} a_{n-1} a_{j-1} z^j - \sum_{j=0}^{n-1} \left(a_{n-2} - \frac{a_{n-1}^2}{a_n} \right) a_j z^j \\ &= \sum_{j=2}^{n-1} \left(a_n a_{j-2} - a_{n-1} a_{j-1} - a_{n-2} a_j + \frac{a_{n-1}^2 a_j}{a_n} \right) z^j \\ &\qquad + \left(-a_{n-1} a_0 - a_{n-2} a_1 + \frac{a_{n-1}^2 a_1}{a_n} \right) z + \left(-a_{n-2} + \frac{a_{n-1}^2}{a_n} \right) a_0.\end{aligned}$$

Analogously to the proof for q_1, we now obtain the same equations (3.9) for both q_2 and q_3, from which the proof follows for these polynomials as well. □

Here too, and for the same reasons as in the matrix case, we make the following choice to improve the Cauchy radius of p:

$$q(z) = \begin{cases} \left(a_n z^{k+\ell} - a_{n-k} z^{\ell} - a_{n-k-\ell}\right) p(z) & \text{if } \ell < k, \\ \left(a_n z^{2k} - a_{n-k} z^{k} + a_{n-k}^2 a_n^{-1}\right) p(z) & \text{if } \ell > k, \\ \left(a_n z^{2k} - a_{n-k} z^{k} - a_{n-2k} + \dfrac{a_{n-k}^2}{a_n}\right) p(z) & \text{if } \ell = k. \end{cases}$$

Acknowledgement

The author thanks Françoise Tisseur for providing the matrices in Example 2.

References

[1] Dario A. Bini, Vanni Noferini, and Meisam Sharify, *Locating the eigenvalues of matrix polynomials*, SIAM J. Matrix Anal. Appl. **34** (2013), no. 4, 1708–1727, DOI 10.1137/120886741. MR3144796

[2] A. L. Cauchy, *Sur la résolution des équations numériques et sur la théorie de l'élimination*, Exercices de Mathématiques, Quatrième Année, 65–128. de Bure frères, Paris, 1829. Also in: Oeuvres Complètes, Série 2, Tome 9, 86–161. Gauthiers-Villars et fils, Paris, 1891.

[3] A. Feriani, F. Perotti, and V. Simoncini, *Iterative system solvers for the frequency analysis of linear mechanical systems*, Comput. Methods Appl. Mech, Eng., **19** (2000), 1719–1739.

[4] Nicholas J. Higham and Françoise Tisseur, *Bounds for eigenvalues of matrix polynomials*, Linear Algebra Appl. **358** (2003), 5–22, DOI 10.1016/S0024-3795(01)00316-0. Special issue on accurate solution of eigenvalue problems (Hagen, 2000). MR1942721

[5] Morris Marden, *Geometry of polynomials*, Second edition. Mathematical Surveys, No. 3, American Mathematical Society, Providence, R.I., 1966. MR0225972

[6] A. Melman, *Generalization and variations of Pellet's theorem for matrix polynomials*, Linear Algebra Appl. **439** (2013), no. 5, 1550–1567, DOI 10.1016/j.laa.2013.05.003. MR3067822

[7] A. Melman, *Bounds for eigenvalues of matrix polynomials with applications to scalar polynomials*, Linear Algebra Appl. **504** (2016), 190–203, DOI 10.1016/j.laa.2016.04.005. MR3502534

[8] Q. I. Rahman and G. Schmeisser, *Analytic theory of polynomials*, London Mathematical Society Monographs. New Series, vol. 26, The Clarendon Press, Oxford University Press, Oxford, 2002. MR1954841

Department of Applied Mathematics, School of Engineering, Santa Clara University, Santa Clara, California 95053

E-mail address: amelman@scu.edu