C. R. Acad. Sci. Paris, Ser. I 354 (2016) 859–863

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Numerical analysis

Improvement of Pellet's theorem for scalar and matrix polynomials

Amélioration du théorème de Pellet pour polynômes scalaires et matriciels

Aaron Melman

Department of Applied Mathematics, School of Engineering, Santa Clara University, Santa Clara, CA 95053, USA

ARTICLE INFO

Article history:
Received 14 December 2015
Accepted after revision 11 April 2016
Available online 10 May 2016

Presented by Olivier Pironneau

ABSTRACT

We improve Pellet's theorem for both scalar and matrix polynomials by using polynomial multipliers.

RÉSUMÉ

Nous améliorons le théorème de Pellet pour les polynômes scalaires et matriciels en utilisant des multiplicateurs polynomiaux.

1. Introduction

The following theorem (Pellet, 1881) provides, when applicable, inclusions for subsets of zeros of a polynomial. It is a direct consequence of Rouché's theorem.

Theorem 1.1. *([8], [5, Th. (28,1), p. 128]) Let $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ be a polynomial of degree $n \geq 2$ with complex coefficients and $a_\ell \neq 0$ for some ℓ with $1 \leq \ell \leq n-1$, and let the polynomial $|a_n|z^n + |a_{n-1}|z^{n-1} + \cdots + |a_{\ell+1}|z^{\ell+1} - |a_\ell|z^\ell + |a_{\ell-1}|z^{\ell-1} + \cdots + |a_0|$ have two distinct positive roots ρ_1 and ρ_2 with $\rho_1 < \rho_2$. Then p has exactly ℓ zeros in or on the circle $|z| = \rho_1$ and no zeros in the open annular ring $\rho_1 < |z| < \rho_2$.*

The quantities ρ_1 and ρ_2 in the statement of Pellet's theorem will be called the *Pellet ℓ-radii* of the polynomial p. We note that, by Descartes' rule of signs, the real polynomial determining the Pellet radii has either two or no positive zeros. A limit case of Pellet's theorem, ascribed to Cauchy ([2], [5, Th. (27,1), p. 122 and Exercise 1, p. 126]), states that all the zeros of the polynomial $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ with complex coefficients and $n \geq 2$, lie in $|z| \leq r$, where r is the unique positive solution to $|a_n|x^n - |a_{n-1}|x^{n-1} - \cdots - |a_1|x - |a_0| = 0$. The bound r, called the *Cauchy radius* of p, is the best possible of bounds depending only on the moduli of the coefficients. A similar upper bound for the moduli of the

E-mail address: amelman@scu.edu.

http://dx.doi.org/10.1016/j.crma.2016.04.006
1631-073X/

eigenvalues of a matrix polynomial was derived in [1,3,6]. Pellet's theorem also has a matrix version that was derived in [1] and [6]. We state it next.

Theorem 1.2. *([1,6]) Let $P(z) = A_n z^n + A_{n-1}z^{n-1} + \cdots + A_1 z + A_0$ be a matrix polynomial of degree $n \geq 2$, $A_j \in \mathbb{C}^{m\times m}$ for $j = 0, \dots, n$. Let A_ℓ be nonsingular for some ℓ with $1 \leq \ell \leq n-1$, and let the polynomial $\|A_n\|x^n + \|A_{n-1}\|x^{n-1} + \cdots + \|A_{\ell+1}\|x^{\ell+1} - \|A_\ell^{-1}\|^{-1}x^\ell + \|A_{\ell-1}\|x^{\ell-1} + \cdots + \|A_1\|x + \|A_0\|$ have two distinct positive roots ρ_1 and ρ_2 with $\rho_1 < \rho_2$. Then $\det(P)$ has exactly ℓm zeros in or on the circle $|z| = \rho_1$ and no zeros in the open annular ring $\rho_1 < |z| < \rho_2$.*

The matrix norms are assumed to be subordinate (induced by a vector norm). Analogously to the scalar case, we call the quantities ρ_1 and ρ_2 the *Pellet ℓ-radii* of P.

For scalar polynomials, the Cauchy radius was improved only relatively recently in Theorem 8.3.1 of [9] using the common technique of multiplying the given polynomial by an appropriately chosen multiplier, i.e., by an appropriately chosen polynomial. The contribution of this theorem lies in identifying the correct multiplier.

In Section 2 of this note we show that the same technique – with a different multiplier – also improves Pellet's theorem for scalar polynomials, and we generalize this result to matrix polynomials in Section 3. The numerical solution of the real equations we will encounter is an irrelevant matter here. Efficient methods for their solution can be found in, e.g., [7].

2. Improved Pellet radii for scalar polynomials

The following theorem improves Pellet's theorem for scalar polynomials.

Theorem 2.1. *Let $p(z) = \sum_{j=0}^n a_j z^j$ be a polynomial of degree n, other than a monomial, with complex coefficients and $a_\ell \neq 0$, $1 \leq \ell \leq n-1$, and with Pellet ℓ-radii ρ_1 and ρ_2, for which $0 < \rho_1 < \rho_2$. Denote by k the smallest positive integer such that $a_{\ell-k} \neq 0$, and define $q(z) = \left(a_\ell z^k - a_{\ell-k}\right) p(z)$. Then the following holds.*

(1) *The polynomial q has Pellet $(\ell+k)$-radii σ_1 and σ_2 that satisfy $0 < \sigma_1 \leq \rho_1 < \rho_2 \leq \sigma_2$, and p has exactly ℓ zeros in or on the circle $|z| = \sigma_1$ and no zeros in the open annular ring $\sigma_1 < |z| < \sigma_2$.*

(2) *If all the coefficients of p are nonzero, then $0 < \sigma_1 < \rho_1 < \rho_2 < \sigma_2$ unless p has zeros of modulus ρ_1 and ρ_2.*

Proof. Throughout the proof, we use the convention that $a_j = 0$ when $j > n$. We begin by examining q. Since $a_{\ell-1} = a_{\ell-2} = \cdots = a_{\ell-k+1} = 0$, we can write

$$q(z) = \left(a_\ell z^k - a_{\ell-k}\right)\left(\sum_{j=\ell}^n a_j z^j + \sum_{j=0}^{\ell-k} a_j z^j\right) = a_\ell z^k \sum_{j=\ell}^n a_j z^j - a_{\ell-k}\sum_{j=\ell}^n a_j z^j + a_\ell z^k \sum_{j=0}^{\ell-k} a_j z^j - a_{\ell-k}\sum_{j=0}^{\ell-k} a_j z^j$$

$$= a_\ell^2 z^{\ell+k} + a_\ell z^k \sum_{j=\ell+1}^n a_j z^j - a_{\ell-k}\sum_{j=\ell+1}^n a_j z^j + a_\ell z^k \sum_{j=0}^{\ell-k-1} a_j z^j - a_{\ell-k}\sum_{j=0}^{\ell-k} a_j z^j . \tag{1}$$

Note that the coefficient of z^ℓ in (1) is zero. Depending on whether $\ell+k \leq n$ or $\ell+k > n$, the coefficient of $z^{\ell+k}$ in (1) is $a_\ell^2 - a_{\ell-k}a_{\ell+k}$ or a_ℓ^2, respectively. By convention, $a_{\ell+k} = 0$ when $\ell+k > n$, so that $q(z) = \left(a_\ell^2 - a_{\ell-k}a_{\ell+k}\right) z^{\ell+k} + \sum_{j=0, j\neq\ell,\ell+k}^{n+k} b_j z^j$, where the coefficients b_j are of the form $a_\ell a_j$, $a_{\ell-k}a_j$, or $a_\ell a_j - a_{\ell-k}a_{j+k}$. If we define

$$\varphi(z) = \sum_{\substack{j=0 \\ j\neq\ell,\ell+k}}^{n+k} \left|b_j\right| z^j , \tag{2}$$

then the Pellet $(\ell+k)$-radii of q, if they exist, are the positive zeros of $\left|a_\ell^2 - a_{\ell-k}a_{\ell+k}\right| z^{\ell+k} = \varphi(z)$. We first show that $a_\ell^2 - a_{\ell-k}a_{\ell+k} \neq 0$. If $\ell+k > n$, so that $a_{\ell+k} = 0$, then this follows immediately from $a_\ell \neq 0$. Assume therefore that $a_{\ell+k} \neq 0$. For $\rho = \rho_1$ or $\rho = \rho_2$, where ρ_1 and ρ_2 are the Pellet ℓ-radii of p, we have:

$$|a_\ell|\rho^\ell = \sum_{\substack{j=0 \\ j\neq\ell}}^n |a_j|\rho^j = |a_{\ell-k}|\rho^{\ell-k} + |a_{\ell+k}|\rho^{\ell+k} + \sum_{\substack{j=0 \\ j\neq\ell-k,\ell,\ell+k}}^n |a_j|\rho^j . \tag{3}$$

Since $a_{\ell-k} \neq 0$, the inequalities $|a_{\ell-k}|\rho^{\ell-k} < |a_\ell|\rho^\ell$ and $|a_{\ell+k}|\rho^{\ell+k} < |a_\ell|\rho^\ell$ hold. Consequently, one obtains that $|a_{\ell-k}|\rho^{\ell-k}|a_{\ell+k}|\rho^{\ell+k} < |a_\ell|^2\rho^{2\ell}$, and, therefore, $|a_{\ell-k}||a_{\ell+k}| < |a_\ell|^2$ so that $\left|a_\ell^2 - a_{\ell-k}a_{\ell+k}\right| \geq |a_\ell|^2 - \left|a_{\ell-k}a_{\ell+k}\right| > 0$. We now compute an upper bound on $\varphi(\rho)$, defined in (2). With $\left|a_\ell a_j - a_{\ell-k}a_{j+k}\right| \leq \left|a_\ell a_j\right| + \left|a_{\ell-k}a_{j+k}\right|$, we have:

$$\varphi(\rho) \leq |a_\ell|\rho^k \sum_{j=\ell+1}^n |a_j|\rho^j + |a_{\ell-k}| \sum_{\substack{j=\ell+1 \\ j\neq\ell+k}}^n |a_j|\rho^j + |a_\ell|\rho^k \sum_{j=0}^{\ell-k-1} |a_j|\rho^j + |a_{\ell-k}|\sum_{j=0}^{\ell-k} |a_j|\rho^j . \tag{4}$$

With (3) we then obtain from (4) that

$$
\begin{aligned}
&= |a_\ell|\rho^k\left(|a_\ell|\rho^\ell - \sum_{j=0}^{\ell-k}|a_j|\rho^j\right) + |a_{\ell-k}|\left(|a_\ell|\rho^\ell - \sum_{j=0}^{\ell-k}|a_j|\rho^j - |a_{\ell+k}|\rho^{\ell+k}\right) \\
&\quad + |a_\ell|\rho^k\left(\sum_{j=0}^{\ell-k}|a_j|\rho^j - |a_{\ell-k}|\rho^{\ell-k}\right) + |a_{\ell-k}|\left(\sum_{j=0}^{\ell-k}|a_j|\rho^j\right) \\
&= \left(|a_\ell|^2 - |a_{\ell-k}a_{\ell+k}|\right)\rho^{\ell+k} \le |a_\ell^2 - a_{\ell-k}a_{\ell+k}|\rho^{\ell+k}\,. \qquad (5)
\end{aligned}
$$

Therefore, $|a_\ell^2 - a_{\ell-k}a_{\ell+k}|\rho^{\ell+k} - \varphi(\rho) \ge 0$. Because this is true for $\rho = \rho_1$ and for $\rho = \rho_2$, we conclude that $\varphi(z) - |a_\ell^2 - a_{\ell-k}a_{\ell+k}|z^{\ell+k}$ has two positive zeros σ_1 and σ_2 with $\sigma_1 \le \rho_1$ and $\sigma_2 \ge \rho_2$. Consequently, q does not have zeros with moduli in the interval (σ_1, σ_2), and, therefore, neither does p. Since p has ℓ zeros with modulus at most ρ_1, this concludes the proof of part (1).

For part (2) we have that all the coefficients of p are nonzero, i.e., $k = 1$, so that

$$
\varphi(z) = |a_\ell a_n|z^{n+1} + \sum_{j=\ell+2}^{n}|a_\ell a_{j-1} - a_{\ell-1}a_j|z^j + \sum_{j=1}^{\ell-1}|a_\ell a_{j-1} - a_{\ell-1}a_j|z^j + |a_{\ell-1}a_0|\,.
$$

If $\sigma_1 = \rho_1$ or $\sigma_2 = \rho_2$, then $|a_\ell^2 - a_{\ell-k}a_{\ell+k}|\rho^{\ell+1} = \varphi(\rho)$ for either $\rho = \rho_1$ or $\rho = \rho_2$, respectively. For this to be true, inequalities (4) and (5) with $k = 1$ must hold as equalities, implying that $|a_\ell a_{j-1} - a_{\ell-1}a_j| = |a_\ell a_{j-1}| + |a_{\ell-1}a_j|$ for $j \ne \ell, \ell+1$, and $|a_\ell^2 - a_{\ell-1}a_{\ell+1}| = |a_\ell|^2 - |a_{\ell-1}a_{\ell+1}|$. We remark that, since these conditions are independent of ρ, $|a_\ell^2 - a_{\ell-k}a_{\ell+k}|\rho^{\ell+1} = \varphi(\rho)$ either holds for both $\rho = \rho_1$ and $\rho = \rho_2$, or does not hold. Bearing in mind that the coefficients are nonzero and denoting the arguments of the complex numbers a_j by θ_j, i.e., $a_j = |a_j|e^{i\theta_j}$, we obtain:

$$
|a_\ell a_{j-1} - a_{\ell-1}a_j| = |a_\ell a_{j-1}| + |a_{\ell-1}a_j| \Longrightarrow \theta_\ell + \theta_{j-1} = \pi + \theta_{\ell-1} + \theta_j\,, \qquad (6)
$$

$$
|a_\ell^2 - a_{\ell-1}a_{\ell+1}| = |a_\ell^2| - |a_{\ell-1}a_{\ell+1}| \Longrightarrow 2\theta_\ell = \theta_{\ell-1} + \theta_{\ell+1}\,. \qquad (7)
$$

Defining $\Delta = \pi + \theta_{\ell-1} - \theta_\ell$, the second equation in (6) is equivalent to $\theta_j = \theta_{j-1} - \Delta$. Applying this recursively for $j = \ell+2, \ldots, n$ yields

$$
\theta_j = \theta_{\ell+1} - (j - \ell - 1)\Delta \quad (j = \ell+2, ..., n)\,, \qquad (8)
$$

and, likewise, when the recursion runs as $j = \ell-1, \ldots, 1$,

$$
\theta_j = \theta_{\ell-1} + (\ell - j - 1)\Delta \quad (j = 0, ..., \ell-2)\,. \qquad (9)
$$

Assuming that the conditions in (6) and (7) are satisfied, we claim that both $\rho_1 e^{i\Delta}$ and $\rho_2 e^{i\Delta}$ are zeros of p. To show this, we evaluate $p(\rho e^{i\Delta})$, where $\rho = \rho_1$ or $\rho = \rho_2$, using (8) and (9):

$$
\begin{aligned}
p\left(\rho e^{i\Delta}\right) &= \sum_{j=\ell+2}^{n}|a_j|e^{i\theta_j}e^{ij\Delta}\rho^j + |a_{\ell+1}|e^{i\theta_{\ell+1}}e^{i(\ell+1)\Delta}\rho^{\ell+1} + |a_\ell|e^{i\theta_\ell}e^{i\ell\Delta}\rho^\ell \\
&\qquad + |a_{\ell-1}|e^{i\theta_{\ell-1}}e^{i(\ell-1)\Delta}\rho^{\ell-1} + \sum_{j=0}^{\ell-2}|a_j|e^{i\theta_j}e^{ij\Delta}\rho^j \\
&= e^{i(\theta_{\ell+1}+(\ell+1)\Delta)}\left(\sum_{j=\ell+1}^{n}|a_j|\rho^j\right) + |a_\ell|e^{i(\theta_\ell+\ell\Delta)}\rho^\ell + e^{i(\theta_{\ell-1}+(\ell-1)\Delta)}\left(\sum_{j=0}^{\ell-1}|a_j|\rho^j\right). \qquad (10)
\end{aligned}
$$

With the definition of Δ and the second equation in (7), we have

$$
\theta_\ell + \ell\Delta = \theta_\ell + \ell\theta_{\ell-1} - \ell\theta_\ell + \ell\pi = \theta_{\ell-1} + (\ell-1)\Delta + \pi\,, \qquad (11)
$$

$$
\theta_{\ell+1} = 2\theta_\ell - \theta_{\ell-1} = \theta_{\ell-1} - 2(\theta_{\ell-1} - \theta_\ell) = \theta_{\ell-1} - 2\Delta + 2\pi\,, \qquad (12)
$$

so that

$$
\theta_{\ell+1} + (\ell+1)\Delta = \theta_{\ell-1} - 2\Delta + 2\pi + (\ell+1)\Delta = \theta_{\ell-1} + (\ell-1)\Delta + 2\pi\,. \qquad (13)
$$

Using (11), (12), and (13) in (10) yields

$$
p\left(\rho e^{i\Delta}\right) = e^{i(\theta_{\ell-1}+(\ell-1)\Delta)}\left(\sum_{j=\ell+1}^{n}|a_j|\rho^j - |a_\ell|\rho^\ell + \sum_{j=0}^{\ell-1}|a_j|\rho^j\right) = 0\,.
$$

This concludes the proof. □

Theorem 2.1 can be applied repeatedly to further improve Pellet's theorem. It can even be used to find Pellet radii when no such radii can otherwise be computed. The following example shows two successive applications of Theorem 2.1 for a simple quartic polynomial. We use the same notation as in Theorem 2.1.

Example. Consider the polynomial $p(z) = 2z^4 - z^3 + 10z^2 - z - 4$ and set $\ell = 2$, which means that $k = 1$. The moduli of the zeros of p are given by 0.5553, 0.6758, 2.3086, and 2.3086, while its Pellet 2-radii form the interval $\Lambda_1 = [0.7701, 1.7892]$, separating the moduli of the two smallest and the two largest zeros. Applying Theorem 2.1, we obtain $q_1(z) = (10z + 1)p(z) = 20z^5 - 8z^4 + 99z^3 - 41z - 4$, for which the Pellet 3-radii yield an interval $\Lambda_2 = [0.7532, 1.8952]$. Applying the theorem once more with $\ell = 3$ and $k = 2$ produces $q_2(z) = (99z^2 + 41)q_1(z) = 1980z^7 - 792z^6 + 10621z^5 - 328z^4 - 396z^2 - 1681z - 164$, for which the Pellet 5-radii yield an interval $\Lambda_3 = [0.7107, 2.0928]$. Clearly, $\Lambda_1 \subseteq \Lambda_2 \subseteq \Lambda_3$.

3. Improved Pellet radii for matrix polynomials

The following theorem improves the matrix version of Pellet's theorem by generalizing Theorem 2.1 to matrix polynomials. All matrix norms are assumed to be subordinate (induced).

Theorem 3.1. *Let $P(z) = \sum_{j=0}^{n} A_j z^j$ be a matrix polynomial of degree n, other than a matrix monomial, with square complex matrix coefficients and A_ℓ nonsingular, and with Pellet ℓ-radii ρ_1 and ρ_2, where $1 \le \ell \le n-1$ and $0 < \rho_1 < \rho_2$. Denote by k the smallest positive integer such that $A_{\ell-k}$ is not the null matrix, let $A_\ell A_{\ell-k} = A_{\ell-k} A_\ell$, and define $Q^{(L)}(z) = \left(A_\ell z^k - A_{\ell-k}\right) P(z)$ and $Q^{(R)}(z) = P(z)\left(A_\ell z^k - A_{\ell-k}\right)$. If $\|A_\ell^{-2}\| = \|A_\ell^{-1}\| \|A_\ell\|^{-1}$, then $Q^{(L)}$ has Pellet $(\ell+k)$-radii $\sigma_1^{(L)}$ and $\sigma_2^{(L)}$, satisfying $0 < \sigma_1^{(L)} \le \rho_1 < \rho_2 \le \sigma_2^{(L)}$, and $\det(P)$ has exactly ℓm zeros in or on the circle $|z| = \sigma_1^{(L)}$, and no zeros in the open annular ring $\sigma_1^{(L)} < |z| < \sigma_2^{(L)}$. An analogous result holds for $Q^{(R)}$.*

Proof. We prove the theorem for $Q^{(L)}$, the proof for $Q^{(R)}$ being analogous. We use the convention that $A_{\ell+k}$ is the null matrix if $\ell + k > n$. Since $A_\ell A_{\ell-k} = A_{\ell-k} A_\ell$ and $A_{\ell-1} = A_{\ell-2} = \cdots = A_{\ell-k+1} = 0$, we obtain similarly as in the scalar case that $Q^{(L)}$ can be written as

$$Q^{(L)}(z) = A_\ell^2 z^{\ell+k} + A_\ell z^k \sum_{j=\ell+1}^{n} A_j z^j - A_{\ell-k} \sum_{j=\ell+1}^{n} A_j z^j + A_\ell z^k \sum_{j=0}^{\ell-k-1} A_j z^j - A_{\ell-k} \sum_{j=0}^{\ell-k} A_j z^j . \tag{14}$$

The coefficient of z^ℓ in (14) vanishes, and the coefficient of $z^{\ell+k}$ is $A_\ell^2 - A_{\ell-k}A_{\ell+k}$, so that $Q^{(L)}(z) = \left(A_\ell^2 - A_{\ell-k}A_{\ell+k}\right) z^{\ell+k} + \sum_{j=0, j\neq \ell, \ell+k}^{n+k} B_j z^j$, where the coefficients B_j are of the form $A_\ell A_j$, $A_{\ell-k}A_j$, or $A_\ell A_j - A_{\ell-k}A_{j+k}$. If we define

$$\Phi(z) = \sum_{\substack{j=0 \\ j\neq \ell, \ell+k}}^{n+k} \|B_j\| z^j \tag{15}$$

for any subordinate (induced) matrix norm, then, if they exist, the $(\ell+k)$-Pellet radii of $Q^{(L)}$ are the positive zeros of $\left\|\left(A_\ell^2 - A_{\ell-k}A_{\ell+k}\right)^{-1}\right\|^{-1} z^{\ell+k} = \Phi(z)$. We first establish that $A_\ell^2 - A_{\ell-k}A_{\ell+k}$ is nonsingular. If $\ell + k > n$, so that $A_{\ell+k}$ is the null matrix, then this follows from the nonsingularity of A_ℓ. Assume therefore that $A_{\ell+k}$ is not the null matrix. For $\rho = \rho_1$ or $\rho = \rho_2$, where ρ_1 and ρ_2 are the Pellet ℓ-radii of P, we have

$$\|A_\ell^{-1}\|^{-1} \rho^\ell = \sum_{\substack{j=0 \\ j\neq \ell}}^{n} \|A_j\| \rho^j = \|A_{\ell-k}\| \rho^{\ell-k} + \|A_{\ell+k}\| \rho^{\ell+k} + \sum_{\substack{j=0 \\ j\neq \ell-k, \ell, \ell+k}}^{n} \|A_j\| \rho^j . \tag{16}$$

Since $A_{\ell-k}$ is not the null matrix,

$$\|A_{\ell-k}\| \rho^{\ell-k} \|A_{\ell+k}\| \rho^{\ell+k} < \|A_\ell^{-1}\|^{-2} \rho^{2\ell} \implies \|A_\ell^{-1}\|^2 \|A_{\ell-k}\| \|A_{\ell+k}\| < 1 .$$

Consequently, since $\|A_\ell^{-2} A_{\ell-k} A_{\ell+k}\| \le \|A_\ell^{-2}\| \|A_{\ell-k}\| \|A_{\ell+k}\| \le \|A_\ell^{-1}\|^2 \|A_{\ell-k}\| \|A_{\ell+k}\| < 1$, the matrix $I - A_\ell^{-2} A_{\ell-k} A_\ell$ is nonsingular [4, p. 351], and because

$$I - A_\ell^{-2} A_{\ell-k} A_{\ell+k} = A_\ell^{-2} \left(A_\ell^2 - A_{\ell-k} A_{\ell+k}\right) ,$$

$A_\ell^2 - A_{\ell-k}A_{\ell+k}$ is also nonsingular. Let us examine the norm of its inverse. Since

$$\left\|\left(A_\ell^2 - A_{\ell-k}A_{\ell+k}\right)^{-1}\right\| = \left\|\left(I - A_\ell^{-2} A_{\ell-k} A_{\ell+k}\right)^{-1} A_\ell^{-2}\right\| \le \left\|\left(I - A_\ell^{-2} A_{\ell-k} A_{\ell+k}\right)^{-1}\right\| \|A_\ell^{-2}\| ,$$

we have:

$$\left\| \left(A_\ell^2 - A_{\ell-k}A_{\ell+k}\right)^{-1} \right\|^{-1} \geq \left\| \left(I - A_\ell^{-2}A_{\ell-k}A_{\ell+k}\right)^{-1} \right\|^{-1} \left\| A_\ell^{-2} \right\|^{-1}$$

$$\geq \left(1 - \left\| A_\ell^{-2}A_{\ell-k}A_{\ell+k} \right\|\right) \left\| A_\ell^{-2} \right\|^{-1} \tag{17}$$

$$\geq \left\| A_\ell^{-2} \right\|^{-1} - \left\| A_{\ell-k} \right\| \left\| A_{\ell+k} \right\| . \tag{18}$$

The inequality in (17) is a consequence of the fact that $\left\| A_\ell^{-2}A_{\ell-k}A_{\ell+k} \right\| < 1$ [4, p. 351].

We now analyze the value of $\Phi(\rho)$, defined in (15), where $\rho = \rho_1$ or $\rho = \rho_2$. Because $\|A_\ell A_j - A_{\ell-k}A_{j+k}\| \leq \|A_\ell\|\|A_j\| + \|A_{\ell-k}\|\|A_{j+k}\|$, we obtain from (14) and (15) that

$$\Phi(\rho) \leq \|A_\ell\|\rho^k \sum_{j=\ell+1}^{n} \|A_j\|\rho^j + \|A_{\ell-k}\| \sum_{\substack{j=\ell+1 \\ j\neq \ell+k}}^{n} \|A_j\|\rho^j + \|A_\ell\|\rho^k \sum_{j=0}^{\ell-k-1} \|A_j\|\rho^j + \|A_{\ell-k}\| \sum_{j=0}^{\ell-k} \|A_j\|\rho^j .$$

Using (16) then yields:

$$\begin{aligned}\Phi(\rho) &\leq \|A_\ell\|\rho^k \left(\|A_\ell^{-1}\|^{-1}\rho^\ell - \sum_{j=0}^{\ell-k} \|A_j\|\rho^j \right) + \|A_{\ell-k}\| \left(\|A_\ell^{-1}\|^{-1}\rho^\ell - \sum_{j=0}^{\ell-k} \|A_j\|\rho^j - \|A_{\ell+k}\|\rho^{\ell+k} \right) \\ &\quad + \|A_\ell\|\rho^k \left(\sum_{j=0}^{\ell-k} \|A_j\|\rho^j - \|A_{\ell-k}\|\rho^{\ell-k} \right) + \|A_{\ell-k}\| \left(\sum_{j=0}^{\ell-k} \|A_j\|\rho^j \right) \\ &= \left(\|A_\ell\| \|A_\ell^{-1}\|^{-1} - \|A_{\ell-k}\|\|A_{\ell+k}\| \right) \rho^{\ell+k} + \|A_{\ell-k}\| \left(\|A_\ell^{-1}\|^{-1} - \|A_\ell\| \right) \rho^\ell \\ &\leq \left(\|A_\ell\| \|A_\ell^{-1}\|^{-1} - \|A_{\ell-k}\|\|A_{\ell+k}\| \right) \rho^{\ell+k} . \end{aligned} \tag{19}$$

The last inequality follows from the fact that $1 = \|I\| = \|A_\ell A_\ell^{-1}\| \leq \|A_\ell\|\|A_\ell^{-1}\|$. Since we assumed that $\|A_\ell^{-2}\| = \|A_\ell^{-1}\|\|A_\ell\|^{-1}$, we obtain from (18) and (19) for both $\rho = \rho_1$ and $\rho = \rho_2$ that

$$\begin{aligned}\left\| \left(A_\ell^2 - A_{\ell-k}A_{\ell+k}\right)^{-1} \right\|^{-1} \rho^{\ell+k} - \Phi(\rho) &\geq \left(\|A_\ell^{-2}\|^{-1} - \|A_{\ell-k}\|\|A_{\ell+k}\| \right) \rho^{\ell+k} - \Phi(\rho) \\ &= \left(\|A_\ell\| \|A_\ell^{-1}\|^{-1} - \|A_{\ell-k}\|\|A_{\ell+k}\| \right) \rho^{\ell+k} - \Phi(\rho) \\ &\geq 0 .\end{aligned}$$

We conclude that $\Phi(z) - \left\| \left(A_\ell^2 - A_{\ell-k}A_{\ell+k}\right)^{-1} \right\|^{-1} z^{\ell+k}$ has two positive zeros σ_1 and σ_2 with $\sigma_1 \leq \rho_1$ and $\sigma_2 \geq \rho_2$. As a result, $\det(Q^{(L)})$ does not have zeros with moduli in the interval (σ_1, σ_2), and, therefore, neither does $\det(P)$. Since $\det(P)$ has ℓm zeros with modulus at most ρ_1, this concludes the proof. □

Like Theorem 2.1, one can apply Theorem 3.1 repeatedly to further improve the matrix version of Pellet's theorem and it can also sometimes be used to find Pellet radii when no such radii can otherwise be computed. Although the conditions $\|A_\ell^{-2}\| = \|A_\ell^{-1}\|\|A_\ell\|^{-1}$ and $A_\ell A_{\ell-k} = A_{\ell-k}A_\ell$ are restrictive, they are always satisfied when $A_\ell = I$, which can be obtained by pre- or postmultiplication by A_ℓ^{-1}, the computation of which is required anyway to apply the theorem. In general, there does not seem to be a large difference between the "left" and "right" versions of the theorem, although there could be exceptions. In the case of successive applications of the theorem, it is possible to alternate between left and right versions.

References

[1] D.A. Bini, V. Noferini, M. Sharify, Locating the eigenvalues of matrix polynomials, SIAM J. Matrix Anal. Appl. 34 (2013) 1708–1727.

[2] A.L. Cauchy, Sur la résolution des équations numériques et sur la théorie de l'élimination, in: Exercices de mathématiques, Quatrième année, de Bure frères, Paris, 1829, pp. 65–128. Also in: Œuvres complètes, Série 2, Tome 9, Gauthier-Villars et fils, Paris, 1891, pp. 86–161.

[3] N.J. Higham, F. Tisseur, Bounds for eigenvalues of matrix polynomials, Linear Algebra Appl. 358 (2003) 5–22.

[4] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, UK, 2013.

[5] M. Marden, Geometry of Polynomials, Mathematical Surveys, vol. 3, American Mathematical Society, Providence, RI, USA, 1966.

[6] A. Melman, Generalization and variations of Pellet's theorem for matrix polynomials, Linear Algebra Appl. 439 (2013) 1550–1567.

[7] A. Melman, Implementation of Pellet's theorem, Numer. Algorithms 65 (2014) 293–304.

[8] M.A. Pellet, Sur un mode de séparation des racines des équations et la formule de Lagrange, Bull. Sci. Math. 5 (1881) 393–395.

[9] Q.I. Rahman, G. Schmeisser, Analytic Theory of Polynomials, London Mathematical Society Monographs. New Series, vol. 26, The Clarendon Press, Oxford University Press, Oxford, UK, 2002.