On Householder Sets for Matrix Polynomials

Thomas R. Cameron[1], Panayiotis J. Psarrakos[2]

Abstract

We present a generalization of Householder sets for matrix polynomials. After defining these sets, we analyze their topological and algebraic properties, which include containing all of the eigenvalues of a given matrix polynomial. Then, we use instances of these sets to derive the Geršgorin set, weighted Geršgorin set, and weighted pseudospectra of a matrix polynomial. Finally, we show that Householder sets are intimately connected to the Bauer-Fike theorem by using these sets to derive Bauer-Fike-type bounds for matrix polynomials.

Keywords: `Householder sets, Geršgorin sets, pseudospectra, eigenvalue, matrix polynomials`
2010 MSC: 47J10, 65F15, 65F35

1. Introduction

In 1964, Alston S. Householder presented an elegant norm derivation of the Geršgorin set of a matrix [7]. Later, in 2004, Richard S. Varga labeled the normed defined sets used in this derivation as Householder sets [19].

Throughout this article, we are interested in *matrix polynomials* of size n and degree m:

$$P(\lambda) = A_m\lambda^m + \cdots + A_1\lambda + A_0, \qquad A_m \neq 0, \tag{1}$$

where the coefficients satisfy $A_i \in \mathbb{C}^{n\times n}$, for $i = 0, 1, \ldots, m$, and λ is a complex variable. We assume that the matrix polynomial is *regular*, that is, $\det P(\lambda)$ is not identically zero. A *finite eigenvalue* of $P(\lambda)$ is any scalar $\mu \in \mathbb{C}$ such that $\det P(\mu) = 0$. A nonzero vector $v \in \mathbb{C}^n$ is an *eigenvector* associated with the eigenvalue μ provided that

$$P(\mu)v = 0.$$

We refer to (μ, v) as an *eigenpair* of the matrix polynomial $P(\lambda)$. Furthermore, the *geometric multiplicity* of μ is the dimension of the null space of $P(\mu)$ and its *algebraic multiplicity* is the multiplicity of μ as a root of $\det P(\lambda)$.

[1]Department of Mathematics and Computer Science, Davidson College, USA (thcameron@davidson.edu)
[2]Department of Mathematics, National Technical University of Athens, Greece (ppsarr@math.ntua.gr)

Associated with each $P(\lambda)$ is the *reversal* matrix polynomial:

$$P_R(\lambda) = \lambda^m P(\lambda^{-1}) = A_m + \cdots + A_1\lambda^{m-1} + A_0\lambda^m.$$

We say that $\mu = \infty$ is an eigenvalue of $P(\lambda)$ if and only if $\mu = 0$ is an eigenvalue of $P_R(\lambda)$. The geometric and algebraic multiplicity of $\mu = \infty$ as an eigenvalue of $P(\lambda)$ are defined by the geometric and algebraic multiplicity of $\mu = 0$ as an eigenvalue of $P_R(\lambda)$, respectively. In addition, the eigenvectors corresponding to $\mu = \infty$ for $P(\lambda)$ are defined by the eigenvectors corresponding to $\mu = 0$ for $P_R(\lambda)$. Finally, the *spectrum* of $P(\lambda)$ is the set of all its eigenvalues (finite and infinite), denoted by $\sigma(P)$. If $P(\lambda) = \lambda I - A$, then the spectrum of $P(\lambda)$ reduces to the spectrum of the matrix A, which we denote by $\sigma(A)$.

The outline of this article follows: In Section 2, we define Householder sets for matrix polynomials and analyze their topological and algebraic properties. Throughout this section, we reference properties of subharmonic functions that are given in Appendix A. Then, in Section 3, we use Householder sets to derive other inclusion sets, such as the Geršgorin set, the weighted Geršgorin set, and the weighted pseudospectra for matrix polynomials; illustrative examples are provided. Finally, in Section 4, we apply Householder sets to derive Bauer-Fike-type bounds for matrix polynomials; examples of how these bounds can be applied to perturbation theory are provided.

2. Householder Sets

Suppose that (μ, v) is an eigenpair of the matrix polynomial $P(\lambda)$ as defined in (1). Then, for any matrix polynomial $Q(\lambda)$ of size n, the following holds:

$$(P(\mu) - Q(\mu))\, v = -Q(\mu)v.$$

If $Q(\mu)$ is invertible, then we have

$$Q(\mu)^{-1}\,(P(\mu) - Q(\mu))\, v = -v.$$

Thus, any induced matrix norm $\|\cdot\|$ satisfies

$$\left\|Q(\mu)^{-1}\,(P(\mu) - Q(\mu))\right\| \geq 1.$$

This equation motivates the definition of the generalized *Householder set* of $P(\lambda)$ with respect to $Q(\lambda)$ and $\|\cdot\|$:

$$H(P, Q) = \sigma(Q) \cup S(P, Q),$$

where

$$S(P, Q) = \{\mu \in \mathbb{C} \setminus \sigma(Q) \colon \ \left\|Q(\mu)^{-1}\,(P(\mu) - Q(\mu))\right\| \geq 1\}.$$

A similar derivation is done in Section 1.4 of [19] for matrices $A, B \in \mathbb{C}^{n\times n}$. The resulting set is known as the Householder set for A with respect to B and $\|\cdot\|$, and is defined as follows:

$$H(A, B) = \sigma(B) \cup S(A, B),$$

where

$$S(A,B) = \{\mu \in \mathbb{C} \setminus \sigma(B) \colon \ \left\| (\mu I - B)^{-1} (A - B) \right\| \geq 1\}.$$

When $P(\lambda) = \lambda I - A$ and $Q(\lambda) = \lambda I - B$, the generalized Householder set reduces to the Householder set for A with respect to B and $\|\cdot\|$.

Throughout this article, we assume that $Q(\lambda)$ is a matrix polynomial of degree less than or equal to the degree of $P(\lambda)$. Furthermore, we consider the matrix polynomial

$$\hat{Q}(\lambda) = \lambda^m Q(\lambda^{-1}),$$

where m is the degree of $P(\lambda)$. Note that $\hat{Q}(\lambda)$ has a zero constant coefficient if and only if the degree of $Q(\lambda)$ is strictly less than m. Otherwise, $\hat{Q}(\lambda)$ is the reversal matrix polynomial associated with $Q(\lambda)$. In addition, we have

$$\begin{aligned} H(P_R, \hat{Q}) \setminus \{0\} &= \left\{ \mu \in \mathbb{C} \setminus \{0\} \colon \ \mu \in \left(\sigma(\hat{Q}) \cup S(P_R, \hat{Q}) \right) \right\} \\ &= \left\{ \mu \in \mathbb{C} \setminus \{0\} \colon \ \mu^{-1} \in (\sigma(Q) \cup S(P,Q)) \right\} \\ &= \left\{ \mu \in \mathbb{C} \setminus \{0\} \colon \ \mu^{-1} \in H(P,Q) \right\}. \end{aligned}$$

As with the eigenvalues of $P(\lambda)$, we say that $\mu = \infty$ lies in $H(P,Q)$ if and only if 0 lies in $H(P_R, \hat{Q})$.

Theorem 2.1. *All eigenvalues (finite and infinite) of the matrix polynomial $P(\lambda)$ lie in the Householder set $H(P,Q)$.*

Proof. Suppose that $\mu \in \sigma(P)$ is a finite eigenvalue. If $\mu \in \sigma(Q)$, then it is clear that $\mu \in H(P,Q)$.

Otherwise, $Q(\mu)$ is invertible and

$$Q(\mu)^{-1} (P(\mu) - Q(\mu))\, v = -v$$

for some nonzero $v \in \mathbb{C}^n$. Therefore, any induced operator norm satisfies

$$\left\| Q(\mu)^{-1} (P(\mu) - Q(\mu)) \right\| \geq 1,$$

and it follows that $\mu \in S(P,Q) \subseteq H(P,Q)$.

Now, suppose that $\mu = \infty$ is an eigenvalue of $P(\lambda)$. Then 0 is an eigenvalue of $P_R(\lambda)$. Furthermore, by the previous argument, it follows that 0 lies in $H(P_R, \hat{Q})$ and, therefore, ∞ lies in $H(P,Q)$. □

In what follows, we discuss the properties of Householder sets. We begin with some basic properties which we use to determine topological properties such as necessary and sufficient conditions on the boundedness of Householder sets. Finally, we prove that under suitable conditions the number of connected components of a Householder set is no greater than the number of eigenvalues of $P(\lambda)$. Throughout, we assume that $P(\lambda)$ is a matrix polynomial as defined in (1) and $Q(\lambda)$ is a matrix polynomial of degree less than or equal to the degree of $P(\lambda)$.

Proposition 2.2. *The following properties hold for all Householder sets of $P(\lambda)$.*

I. *$H(P,Q)$ is a closed subset of $\mathbb{C}$.*

II. *For any $\alpha \in \mathbb{C} \setminus \{0\}$, the following hold:*
   i. *$H(P_\alpha, Q_\alpha) = \alpha^{-1} H(P,Q)$, where $P_\alpha(\lambda) = P(\alpha\lambda)$;*
   ii. *$H(\alpha P, \alpha Q)) = H(P,Q)$, where $\alpha P(\lambda)$ results in the coefficients of $P(\lambda)$ being multiplied by α;*
   iii. *$H({}^\alpha P, {}^\alpha Q) = H(P,Q) - \alpha$, where ${}^\alpha P(\lambda) = P(\lambda + \alpha)$.*

III. *If $\det Q(\lambda)$ is (identically) zero, then $H(P,Q) = \mathbb{C}$.*

IV. *Suppose that $P(\lambda) = \beta Q(\lambda)$ for some $\beta \in \mathbb{C} \setminus \{0\}$. If $|\beta - 1| < 1$, then $H(P,Q) = \sigma(P)$, and if $|\beta - 1| \geq 1$, then $H(P,Q) = \mathbb{C}$.*

V. *If the coefficients of $Q(\lambda)$ and $P(\lambda)$ are real and $\|\cdot\|$ is invariant under the complex conjugation of matrix entries, then the set $H(P,Q)$ is symmetric with respect to the real axis.*

Proof.

I. It is clear that $\sigma(Q)$ is closed as it is either a finite set or all of $\mathbb{C}$. Thus, we only need to show that the set $S(P,Q)$ is closed. To this end, suppose that $\mu \in \mathbb{C} \setminus \sigma(Q)$ satisfies

$$\left\| Q(\mu)^{-1} \left(P(\mu) - Q(\mu) \right) \right\| < 1.$$

Since $Q(\mu)$ is invertible, it follows that there exists a neighborhood about μ for which $Q(\lambda)^{-1} \left(P(\lambda) - Q(\lambda) \right)$ is continuous. Therefore, for $z \in \mathbb{C}$ that are sufficiently close to μ we have

$$\left\| Q(z)^{-1} \left(P(z) - Q(z) \right) \right\| < 1.$$

Hence, the set $\mathbb{C} \setminus S(P,Q)$ is open, and it follows that $S(P,Q)$ is closed.

II. Let $\alpha \in \mathbb{C} \setminus \{0\}$.
   i. Note that

$$\begin{aligned} H(P_\alpha, Q_\alpha) &= \sigma(Q_\alpha) \cup S(P_\alpha, Q_\alpha) \\ &= \alpha^{-1}\sigma(Q) \cup \alpha^{-1} S(P,Q) \\ &= \alpha^{-1} H(P,Q). \end{aligned}$$

   ii. Similarly, we have

$$\begin{aligned} H(\alpha P, \alpha Q) &= \sigma(\alpha Q) \cup S(\alpha P, \alpha Q) \\ &= \sigma(Q) \cup S(P,Q) \\ &= H(P,Q). \end{aligned}$$

   iii. Finally,

$$\begin{aligned} H({}^\alpha P, {}^\alpha Q) &= \sigma({}^\alpha Q) \cup S({}^\alpha P, {}^\alpha Q) \\ &= (\sigma(Q) - \alpha) \cup (S(P,Q) - \alpha) \\ &= H(P,Q) - \alpha. \end{aligned}$$

III. If $\det Q(\lambda)$ is identically zero, then

$$\mathbb{C} = \sigma(Q) \subseteq H(P,Q) \subseteq \mathbb{C}.$$

Hence, $H(P,Q) = \mathbb{C}$.

IV. If $P(\lambda) = \beta Q(\lambda)$ for some $\beta \in \mathbb{C} \setminus \{0\}$, then

$$\begin{aligned} H(P,Q) &= \sigma(Q) \cup S(\beta Q, Q) \\ &= \sigma(Q) \cup \{\mu \in \mathbb{C} \setminus \sigma(Q) \colon \ \ |\beta - 1| \geq 1\}. \end{aligned}$$

Hence, $H(P,Q) = \mathbb{C}$ if $|\beta - 1| \geq 1$ and $H(P,Q) = \sigma(Q) = \sigma(P)$ otherwise.

V. If the coefficients of a matrix polynomial are real, then its spectrum is symmetric with respect to the real axis. Thus, all that remains is to show that $S(P,Q)$ is symmetric with respect to the real axis when $\|\cdot\|$ is invariant under the complex conjugate of matrix entries. To that end, let $\mu \in S(P,Q)$. Then,

$$\begin{aligned} \left\| Q(\overline{\mu})^{-1} \left(P(\overline{\mu}) - Q(\overline{\mu}) \right) \right\| &= \left\| \overline{Q(\mu)^{-1} \left(P(\mu) - Q(\mu) \right)} \right\| \\ &= \left\| Q(\mu)^{-1} \left(P(\mu) - Q(\mu) \right) \right\| \geq 1, \end{aligned}$$

and it follows that $\overline{\mu} \in S(P,Q)$.

□

Proposition 2.3. *The set* $\{\mu \in \mathbb{C} \setminus \sigma(Q) \colon \ \ \left\| Q(\mu)^{-1} \left(P(\mu) - Q(\mu) \right) \right\| > 1\}$ *lies in the interior of* $S(P,Q) \subseteq H(P,Q)$*. As a result, the boundary of the Householder set is contained in the following union:*

$$\sigma(Q) \cup \left\{ \mu \in \mathbb{C} \setminus \sigma(Q) \colon \ \ \left\| Q(\mu)^{-1} \left(P(\mu) - Q(\mu) \right) \right\| = 1 \right\}.$$

Proof. Suppose that $\mu \in \mathbb{C} \setminus \sigma(Q)$ satisfies

$$\left\| Q(\mu)^{-1} \left(P(\mu) - Q(\mu) \right) \right\| > 1.$$

Then, by continuity, there exists an $\varepsilon > 0$ such that for every $z \in \mathbb{C}$ with $|\mu - z| < \varepsilon$, we have $z \notin \sigma(Q)$ and

$$\left\| Q(z)^{-1} \left(P(z) - Q(z) \right) \right\| > 1.$$

Thus, μ is an interior point of $S(P,Q)$. Therefore, if μ is a boundary point of $H(P,Q)$, then $\mu \in \sigma(Q)$, or $\mu \notin \sigma(Q)$ and

$$\left\| Q(\mu)^{-1} \left(P(\mu) - Q(\mu) \right) \right\| = 1.$$

□

In the following results, we make use of properties of subharmonic functions that are given in Appendix A.

Theorem 2.4. *If μ is an isolated point of $H(P,Q)$, then $\mu \in \sigma(Q)$.*

Proof. For the sake of contradiction, assume that $\mu \notin \sigma(Q)$ and μ is an isolated point of $H(P,Q)$. Then, it follows from Proposition 2.3 that

$$\left\| Q(\mu)^{-1} \left(P(\mu) - Q(\mu) \right) \right\| = 1.$$

Furthermore, there exists an $\varepsilon > 0$ such that the closed disk

$$D(\mu, \varepsilon) = \{ \lambda \in \mathbb{C} \colon \ |\lambda - \mu| \leq \varepsilon \} \tag{2}$$

contains no other points of $H(P,Q)$.

Define $\phi(\lambda) = \left\| Q(\lambda)^{-1} \left(P(\lambda) - Q(\lambda) \right) \right\|$ on $D(\mu, \varepsilon)$. Since $D(\mu, \varepsilon)$ contains no points of $\sigma(Q)$, it follows that $Q(\lambda)^{-1} \left(P(\lambda) - Q(\lambda) \right)$ is a non-zero analytic matrix-valued function on $D(\mu, \varepsilon)$. Therefore, by Theorem A.3, it follows that $\phi(\lambda)$ is a subharmonic function on $D(\mu, \varepsilon)$. As a subharmonic function, $\phi(\lambda)$ should obtain its maximum on $\partial D(\mu, \varepsilon)$. However, $\phi(\mu) = 1$ and $\phi(\lambda) < 1$ for all $\lambda \in D(\mu, \varepsilon) \setminus \{\mu\}$, which is a contradiction, and it follows that $\mu \in \sigma(Q)$. □

Let Ω be a closed subset of $\mathbb{C}$ and let $\mu \in \Omega$. The *local dimension* of μ is defined by

$$\lim_{h \to 0^+} \dim\{\Omega \cap D(\mu, h)\},$$

where $h \in \mathbb{R}_+$ and $\dim\{\cdot\}$ denotes the topological dimension [8]. Any isolated point in Ω has local dimension 0. Furthermore, any non-isolated point in Ω has local dimension 2 if and only if its belongs to the closure of the interior of Ω, and 1 otherwise.

Theorem 2.5. *Any point of $S(P,Q)$ has local dimension 2.*

Proof. Let $\mu \in S(P,Q)$. It is immediately clear from Theorem 2.4 that μ cannot have local dimension 0. Now, for the sake of contradiction, suppose that μ has local dimension 1. Then, since μ is not an interior point of $S(P,Q)$, there exists an $\varepsilon > 0$ such that

$$S(P,Q) \cap D(\mu, \varepsilon) \subseteq \left\{ \mu \in \mathbb{C} \setminus \sigma(Q) \colon \ \left\| Q(\mu)^{-1} \left(P(\mu) - Q(\mu) \right) \right\| = 1 \right\},$$

where $D(\mu, \varepsilon)$ is the closed disk defined in (2).

Thus, by Theorem A.3, $\phi(\lambda) = \left\| Q(\lambda)^{-1} \left(P(\lambda) - Q(\lambda) \right) \right\|$ is a subharmonic function on $D(\mu, \varepsilon)$, and $\phi(\lambda)$ takes on its maximum value of 1 at (infinitely many) interior points of $D(\mu, \varepsilon)$. However, $\phi(\lambda)$ is non-constant on $D(\mu, \varepsilon)$ since the disk must contain a point that is not in $S(P,Q)$, otherwise μ would be an interior point of $S(P,Q)$. Therefore, $\phi(\lambda)$ taking on its maximum value of 1 in the interior of $D(\mu, \varepsilon)$ contradicts the maximum principle, and it follows that no $\mu \in S(P,Q)$ has local dimension 1. □

As a corollary of Theorem 2.5, note that any point of $H(P,Q)$ has local dimension either 2 or 0, that is, Householder sets of $P(\lambda)$ cannot have parts which are curves.

By Theorem 2.4, if the origin is an isolated point of $H(P, Q)$ it follows that $0 \in \sigma(Q)$. Therefore, if $0 \notin \sigma(\hat{Q})$, then the origin is not an isolated point of $H(P_R, \hat{Q})$ and it follows that $H(P, Q)$ is not the union of a bounded set and ∞. Furthermore, if $0 \notin \sigma(\hat{Q})$, it follows that $Q(\lambda)$ is a matrix polynomial of degree m, which we denote by

$$Q(\lambda) = B_m\lambda^m + \cdots + B_1\lambda + B_0,$$

where B_m is invertible. We are now ready to prove necessary and sufficient conditions on the boundedness of Householder sets.

Theorem 2.6. *Suppose that $0 \notin \sigma(\hat{Q})$. Then, $H(P, Q)$ is unbounded if and only if $0 \in S(A_m, B_m)$.*

Proof. Suppose that $0 \in S(A_m, B_m)$, that is, $\left\|B_m^{-1}(A_m - B_m)\right\| \geq 1$. Then, $0 \in S(P_R, \hat{Q}) \subseteq H(P_R, \hat{Q})$ and, therefore, $\infty \in H(P, Q)$. By hypothesis, 0 is not an isolated point of $H(P_R, \hat{Q})$ and, hence, ∞ is not an isolated point of $H(P, Q)$.

Conversely, suppose that $H(P, Q)$ is unbounded. Then, there is a sequence $\{u_l\}_{l\in\mathbb{N}}$ in $H(P, Q)$ such that $|u_l| \to \infty$ as $l \to \infty$. For every $l \in \mathbb{N}$, $u_l \in \sigma(Q)$ or $u_l \in S(P, Q)$. Since $0 \notin \sigma(\hat{Q})$, $Q(\lambda)$ is regular and it follows that the cardinality of $\sigma(Q)$ is finite. Therefore, there exists an $N > 0$ such that $u_l \notin \sigma(Q)$ and, hence, $u_l \in S(P, Q)$ for all $l > N$. So, for $l > N$, we have

$$\left\|\hat{Q}(u_l^{-1})^{-1}\left(P_R(u_l^{-1}) - \hat{Q}(u_l^{-1})\right)\right\| \geq 1.$$

Taking the limit as $l \to \infty$ we find that

$$\left\|\hat{Q}(0)^{-1}\left(P_R(0) - \hat{Q}(0)\right)\right\| \geq 1,$$

which implies that $0 \in S(A_m, B_m)$. □

It is well-known, see Corollary 6.1.6 of [6], that the number of Geršgorin disks corresponding to $A \in \mathbb{C}^{n\times n}$ is bounded above by n, that is, the number of eigenvalues of A. Similarly, under suitable conditions, the number of connected components of the Geršgorin set of a matrix polynomial $P(\lambda)$ is bounded above by the number of eigenvalues of $P(\lambda)$, see Theorem 2.9 of [13]. The following theorem generalizes these results for Householder sets.

Theorem 2.7. *Suppose that $0 \notin \sigma(\hat{Q})$ and $H(P, Q)$ is bounded. Then, the number of connected components of $H(P, Q)$ is bounded above by nm. Moreover, each connected component has the same number of eigenvalues of $Q(\lambda)$ and $P(\lambda)$, counting multiplicities.*

Proof. Let $F(\lambda) = P(\lambda) - Q(\lambda)$, and define the family of matrix polynomials

$$P_t(\lambda) = Q(\lambda) + tF(\lambda),$$

for $t \in [0,1]$. Let $0 \le t_1 \le t_2 \le 1$ and suppose that $\mu \in H(P_{t_1}, Q)$. Then, either $\mu \in \sigma(Q)$ or $t_1 \left\| Q(\mu)^{-1} F(\mu) \right\| \ge 1$. In either case, we have $\mu \in H(P_{t_2}, Q)$. Therefore, $H(P_t, Q)$ is a nondecreasing family of compact sets for $t \in [0,1]$. Since $0 \notin \sigma(\hat{Q})$, it follows that $Q(\lambda)$ has degree m and B_m is invertible.

In addition, since $H(P,Q)$ is bounded, the scalar polynomial $\det P_t(\lambda)$ has degree nm for all $t \in [0,1]$. The continuity of the roots of $\det P_t(\lambda)$ with respect to t implies that every eigenvalue of $Q(\lambda)$ is connected to an eigenvalue of $P(\lambda)$ by a continuous curve in $H(P,Q)$. Since every such curve must lie in a connected component of $H(P,Q)$, the result follows. □

3. More Inclusion Sets

Householder sets were originally used to give elegant derivations of the Geršgorin and weighted Geršgorin set of a matrix [7, 19]. It is in this spirit that we use the generalized Householder sets defined in Section 2 to formulate the Geršgorin set, the weighted Geršgorin set, and the weighted pseudospectra of a matrix polynomial. Here and throughout the remainder of the article, we denote by $\operatorname{diag} P(\lambda)$ the diagonal matrix polynomial whose diagonal entries are equal to the diagonal entries of $P(\lambda)$, by $P(\lambda)_{i,j}$ the (i,j)-th entry of $P(\lambda)$, and by $\operatorname{diag}(x_1, \ldots, x_n)$ a diagonal matrix whose diagonal entries are $x_1, \ldots, x_n$.

3.1. The Geršgorin Set

Recently, in [2, 10, 13], the authors developed a generalized Geršgorin set for matrix polynomials and analytic matrix-valued functions. Specifically, the *Geršgorin set* of the matrix polynomial $P(\lambda)$ is defined by

$$G(P) = \bigcup_{i=1}^{n} G_i(P), \tag{3}$$

where

$$G_i(P) = \left\{ \mu \in \mathbb{C} \colon \ |P(\mu)_{i,i}| \le \sum_{\substack{j=1 \\ j \ne i}}^{n} |P(\mu)_{i,j}| \right\}.$$

When $P(\lambda) = \lambda I - A$, $G(P)$ reduces to the classic Geršgorin set of the matrix A [6], which we denote by $G(A)$. Also, if $P(\lambda) = \lambda B - A$, then $G(P)$ reduces to the Geršgorin-type set of the matrix pencil (A, B) [9, 16]. In this section, we show that the Geršgorin set for matrix polynomials is an instance of the Householder sets defined in Section 2.

Theorem 3.1. *Let $P(\lambda)$ be a matrix polynomial as defined in* (1)*. Then, the Householder set of $P(\lambda)$ with respect to* $\operatorname{diag} P(\lambda)$ *and* $\|\cdot\| = \|\cdot\|_\infty$ *coincides with the Geršgorin set of $P(\lambda)$.*

Proof. Suppose that $\mu \in H(P, \operatorname{diag} P)$. If $\mu \in \sigma(\operatorname{diag} P)$, then there is some $i \in \{1, \ldots, n\}$ such that $P(\mu)_{i,i} = 0$ and it follows that $\mu \in G_i(P) \subseteq G(P)$. Otherwise, $\mu \in S(P, \operatorname{diag} P)$ and it follows that

$$\left\| \operatorname{diag} P(\mu)^{-1} \left(P(\mu) - \operatorname{diag} P(\mu)\right) \right\|_{\infty} \geq 1.$$

Note that the entries of the matrix $\operatorname{diag} P(\mu)^{-1} \left(P(\mu) - \operatorname{diag} P(\mu)\right) = [x_{i,j}]$ can be written as $x_{i,j} = 0$ if $i = j$ and

$$x_{i,j} = \frac{P(\mu)_{i,j}}{P(\mu)_{i,i}}$$

otherwise. Therefore,

$$1 \leq \left\| \operatorname{diag} P(\mu)^{-1} \left(P(\mu) - \operatorname{diag} P(\mu)\right) \right\|_{\infty} = \max_{1 \leq i \leq n} \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} \left| \frac{P(\mu)_{i,j}}{P(\mu)_{i,i}} \right| \right).$$

Hence, there is an $i \in \{1, \ldots, n\}$ such that

$$|P(\mu)_{i,i}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |P(\mu)_{i,j}|$$

and it follows that $\mu \in G_i(P) \subseteq G(P)$.

Conversely, suppose that $\mu \in G(P)$. Then, there exists an $i \in \{1, \ldots, n\}$ such that

$$|P(\mu)_{i,i}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |P(\mu)_{i,j}|.$$

If $P(\mu)_{i,i} = 0$, then it follows that $\mu \in \sigma(\operatorname{diag} P) \subseteq H(P, \operatorname{diag} P)$. Otherwise,

$$1 \leq \max_{1 \leq i \leq n} \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} \left| \frac{P(\mu)_{i,j}}{P(\mu)_{i,i}} \right| \right) = \left\| \operatorname{diag} P(\mu)^{-1} \left(P(\mu) - \operatorname{diag} P(\mu)\right) \right\|_{\infty}$$

and, hence, $\mu \in S(P, \operatorname{diag} P) \subseteq H(P, \operatorname{diag} P)$. □

3.2. The Weighted Geršgorin Set

Geršgorin was the first to recognize the use of similarity transformations $X^{-1}AX$, where $X = \operatorname{diag}(x_1, \ldots, x_n)$ and $x_i > 0$, for $i = 1, \ldots, n$, to sharpen the bounds on the Geršgorin set of $A \in \mathbb{C}^{n \times n}$ [4]. Any similarity transformation can be used; however, a diagonal similarity transformation with positive entries has the advantage of an easily understood impact on the inclusion set. The set

$G^X(A) = G(X^{-1}AX)$ is known as the *weighted Geršgorin set* of A with respect to X and is equal to the following union

$$G^X(A) = \bigcup_{i=1}^{n} G_i^X(A), \tag{4}$$

where

$$G_i^X(A) = \left\{ \mu \in \mathbb{C}\colon \quad |\mu - a_{i,i}| \le \sum_{\substack{j=1\\ j\neq i}}^{n} \frac{|a_{i,j}|\, x_j}{x_i} \right\}.$$

Now, define the set of all positive diagonal matrices

$$\mathbb{D} = \{\operatorname{diag}(x_1, \ldots, x_n)\colon\ x_i > 0,\ i = 1, \ldots, n\}.$$

For each $X \in \mathbb{D}$, we define

$$\nu_X(u) = \left\|X^{-1}u\right\|_{\infty}, \tag{5}$$

for all $u \in \mathbb{C}^n$. It is clear that $\nu_X(\cdot)$ is a norm on $\mathbb{C}^n$. Furthermore, the matrix norm induced by $\nu_X(\cdot)$ is defined by

$$\nu_X(A) = \left\|X^{-1}AX\right\|_{\infty},$$

for all $A \in \mathbb{C}^{n\times n}$. The *weighted Geršgorin set* of the matrix polynomial $P(\lambda)$, denoted by $G^X(P)$, is the Householder set of $P(\lambda)$ with respect to $\operatorname{diag} P(\lambda)$ and $\|\cdot\| = \nu_X(\cdot)$.

Theorem 3.2. *Let $P(\lambda) = \lambda I - A$. Then, $G^X(P)$ reduces to the weighted Geršgorin set with respect to X of the matrix A as defined in* (4).

Proof. Suppose that $\mu \in H(P, \operatorname{diag} P)$. If $\mu \in \sigma(\operatorname{diag} P)$, then there is some $i \in \{1, \ldots, n\}$ such that $\mu - a_{i,i} = 0$ and it follows that $\mu \in G_i^X(A) \subseteq G^X(A)$. Otherwise, $\mu \in S(P, \operatorname{diag} P)$, and it follows that

$$\nu_X\left(\operatorname{diag} P(\mu)^{-1}\left(P(\mu) - \operatorname{diag} P(\mu)\right)\right) \ge 1,$$

that is,

$$\left\|X^{-1}\operatorname{diag}(\mu I - A)^{-1}(\operatorname{diag} A - A)\, X\right\|_{\infty} \ge 1.$$

Therefore, we have

$$1 \le \max_{1\le i\le n}\left(\sum_{\substack{j=1\\ j\neq i}}^{n}\left|\frac{a_{ij}x_j}{(\mu - a_{ii})x_i}\right|\right).$$

So, there is an $i \in \{1, \ldots, n\}$ such that

$$|\mu - a_{ii}| \le \sum_{\substack{j=1\\ j\neq i}}^{n} \frac{|a_{ij}|\, x_j}{x_i},$$

which implies that $\mu \in G_i^X(A) \subseteq G^X(A)$.

Conversely, suppose that $\mu \in G^X(A)$. Then, there exists an $i \in \{1, \ldots, n\}$ such that

$$|\mu - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{|a_{ij}|\, x_j}{x_i}.$$

If $\mu - a_{ii} = 0$, then it follows that $\mu \in \sigma(\operatorname{diag} P) \subseteq H(P, \operatorname{diag} P)$. Otherwise,

$$\begin{aligned} 1 &\leq \max_{1 \leq i \leq n} \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} \left| \frac{a_{ij} x_j}{(\mu - a_{ii}) x_i} \right| \right) \\ &= \nu_X \left(\operatorname{diag} P(\mu)^{-1} \left(P(\mu) - \operatorname{diag} P(\mu) \right) \right) \end{aligned}$$

and, hence, $\mu \in S(P, \operatorname{diag} P) \subseteq H(P, \operatorname{diag} P)$. □

It follows from Theorem 3.2 that $G^X(P)$ is a generalization of the weighted Geršgorin set of a matrix. Furthermore, the intersection

$$\bigcap_{X \in \mathbb{D}} G^X(P)$$

is equal to the minimal Geršgorin set of $P(\lambda)$ as defined in [10].

3.3. The Weighted Pseudospectrum

Finally, we consider the weighted pseudospectra of a matrix polynomial, which is an established tool for gaining insight into the sensitivity of eigenvalues to perturbations in the coefficients of the matrix polynomial [1, 3, 11, 14, 17]. In this section, we show that the weighted pseudospectra can be represented as a union of Householder sets.

Let $P(\lambda)$ be a matrix polynomial as defined in (1) and consider (additive) perturbations of $P(\lambda)$ of the form

$$P_\Delta(\lambda) = (A_m + \Delta_m)\lambda^m + \cdots + (A_1 + \Delta_1)\lambda + (A_0 + \Delta_0),$$

where the matrices $\Delta_0, \Delta_1, \ldots, \Delta_m \in \mathbb{C}^{n \times n}$ are arbitrary. For a given $\varepsilon > 0$ and a set of nonnegative weights $\mathbf{w} = \{w_0, w_1, \ldots, w_m\}$, the *$\varepsilon$-pseudospectrum* of $P(\lambda)$ with respect to $\mathbf{w}$ is defined by

$$\sigma_{\varepsilon,\mathbf{w}}(P) = \{\mu \in \mathbb{C}\colon \ \det P_\Delta(\mu) = 0, \ \|\Delta_j\| \leq \varepsilon w_j, \ j = 0, 1, \ldots, m\},$$

where $\|\cdot\|$ is any induced matrix norm.

It is useful to define the associated compact set of perturbations of $P(\lambda)$

$$\mathcal{B}(P, \varepsilon, \mathbf{w}) = \{P_\Delta(\lambda)\colon \ \|\Delta_j\| \leq \varepsilon w_j, \ j = 0, 1, \ldots, m\}. \tag{6}$$

Then, the ε-pseudospectrum of $P(\lambda)$ can also be expressed in the form

$$\sigma_{\varepsilon,\mathbf{w}}(P) = \{\mu \in \mathbb{C}: \ \det P_\Delta(\mu) = 0, \ P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathbf{w})\}.$$

Furthermore, by Lemma 2.1 of [17], we have

$$\sigma_{\varepsilon,\mathbf{w}}(P) = \left\{\mu \in \mathbb{C}: \ \left\|P(\mu)^{-1}\right\|^{-1} \leq \varepsilon q_{\mathbf{w}}(|\mu|)\right\},$$

where

$$q_{\mathbf{w}}(\lambda) = w_m\lambda^m + \cdots + w_1\lambda + w_0.$$

Theorem 3.3. *Let $P(\lambda)$ be a matrix polynomial as defined in (1), and let $\varepsilon > 0$ and $\mathbf{w}$ be given. Then, for any perturbation $P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathbf{w})$, we have*

$$H(P_\Delta, P) \subseteq \sigma_{\varepsilon,\mathbf{w}}(P).$$

Proof. Suppose that $\mu \in H(P_\Delta, P)$. If $\mu \in \sigma(P)$, then $\mu \in \sigma_{\varepsilon,\mathbf{w}}(P)$. Otherwise, $\mu \in S(P_\Delta, P)$, and it follows that

$$\left\|P(\mu)^{-1}\left(\Delta_m\mu^m + \cdots + \Delta_1\mu + \Delta_0\right)\right\| \geq 1.$$

Therefore,

$$\begin{aligned} 1 &\leq \left\|P(\mu)^{-1}\left(\Delta_m\mu^m + \cdots + \Delta_1\mu + \Delta_0\right)\right\| \\ &\leq \varepsilon \left\|P(\mu)^{-1}\right\| q_{\mathbf{w}}(|\mu|), \end{aligned}$$

which implies that

$$\left\|P(\mu)^{-1}\right\|^{-1} \leq \varepsilon q_{\mathbf{w}}(|\mu|),$$

and $\mu \in \sigma_{\varepsilon,\mathbf{w}}(P)$. □

Corollary 3.4. *Let $P(\lambda)$ be a matrix polynomial as defined in (1), and let $\varepsilon > 0$ and $\mathbf{w}$ be given. Then, the union of $H(P_\Delta, P)$ over all $P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathbf{w})$ is equal to $\sigma_{\varepsilon,\mathbf{w}}(P)$.*

Proof. By Theorem 2.1, the eigenvalues of $P_\Delta(\lambda)$ are contained in $H(P_\Delta, P)$. Thus, $\sigma_{\varepsilon,\mathbf{w}}(P)$ is contained in the union of $H(P_\Delta, P)$ over all $P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathbf{w})$, and the result follows from Theorem 3.3. □

Corollary 3.5. *Let $P(\lambda)$ be a matrix polynomial as defined in (1), $Q(\lambda)$ a matrix polynomial of degree m, with coefficients $B_i \in \mathbb{C}^{n\times n}$ for $i = 0, 1, \ldots, m$, and $\mathbf{w}$ a given set of weights. If*

$$\varepsilon \geq \max\{\|A_j - B_j\|/w_j \ : \ w_j > 0, \ j = 0, 1, \ldots, m\},$$

then for every $\mu \in H(P, Q)$ there is a $Q_\Delta(\lambda) \in \mathcal{B}(Q, \varepsilon, \mathbf{w})$ such that $\mu \in \sigma(Q_\Delta)$.

Proof. For $j = 0, 1, \ldots, m$, define $\Delta_j = A_j - B_j$. Then, for $j = 0, 1, \ldots, m$, we have $\|\Delta_j\| \leq \varepsilon w_j$. Therefore,

$$\begin{aligned} P(\lambda) &= A_m\lambda^m + \cdots + A_1\lambda + A_0 \\ &= (B_m + \Delta_m)\lambda^m + \cdots + (B_1 + \Delta_1)\lambda + (B_0 + \Delta_0) \end{aligned}$$

is an element of $\mathcal{B}(Q, \varepsilon, \mathbf{w})$. By Theorem 3.3, it follows that $H(P, Q) \subseteq \sigma_{\varepsilon,\mathbf{w}}(Q)$. □

3.4. Examples

In this section, we present three examples to illustrate the properties derived in Section 2. For each example, we plot three Householder sets that correspond to the inclusion sets from Sections 3.1–3.3. Furthermore, the shaded (light blue) region represents the interior of the Householder set, the darker curve (orange) is the boundary of the set[3], and the asterisks (black) are the eigenvalues of the matrix polynomial. It is worth mentioning that we compute the interior and boundary of these Householder sets with a brute force method in Mathematica, and we are interested in developing more efficient methods for computing the boundary of Householder sets.

Example 3.6. Let

$$P(\lambda) = \begin{bmatrix} 5\lambda^6 + \lambda^3 + 7 & 2\lambda^6 + 4 \\ \lambda^6 + 1 & 2\lambda^6 + 3\lambda \end{bmatrix}$$

and consider the three Householder sets in Figure 1.

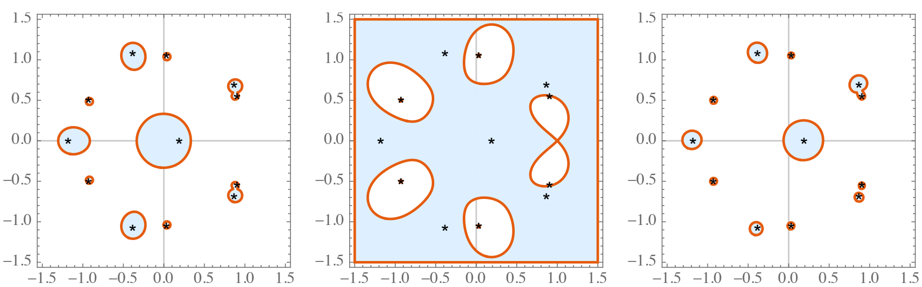

Figure 1: Householder sets of Example 3.6

The Householder set on the left corresponds to the Geršgorin set of $P(\lambda)$; the Householder set in the center corresponds to the weighted Geršgorin set of $P(\lambda)$, where $X = \operatorname{diag}(5, 2)$; the Householder set on the right corresponds to $H(P_\Delta, P)$, where $\|\cdot\| = \|\cdot\|_\infty$ and $P_\Delta(\lambda)$ is an element of $\mathcal{B}(P, \varepsilon, \mathbf{w})$ as defined in (6), randomly selected with $\varepsilon = 0.075$ and $w_j = \|A_j\|_\infty$, for $j = 0, 1, \ldots, m$.

Note that the Householder sets on the left and in the center are symmetric with respect to the real axis, verifying Proposition 2.2 since the coefficients of $P(\lambda)$ are real. In contrast, the Householder set on the right is not symmetric, which can be attributed to $P_\Delta(\lambda)$ having complex coefficients. Also, since none of the Householder sets make up the whole complex plane, we know that both $P(\lambda)$ and $\operatorname{diag} P(\lambda)$ are regular. However, the Householder set in the center is unbounded, by Theorem 2.6, since $0 \in H(A_m, B_m)$. Finally, by Theorem 3.3, the Householder set on the right is a subset of the ε-pseudospectrum of $P(\lambda)$.

[3]Note that the darker curve (orange) surrounding the central plot of Figure 1 is not the boundary of the set, but rather indicates that the set is unbounded.

Example 3.7. Let

$$P(\lambda) = \begin{bmatrix} \lambda^{10} + 3.5i & 3i & 0 \\ 2 & \lambda^{10} - 2i & 0 \\ 0.5 + 2i & -7i & \lambda^{10} + 5 - 3i \end{bmatrix}$$

and consider the three Householder sets in Figure 2.

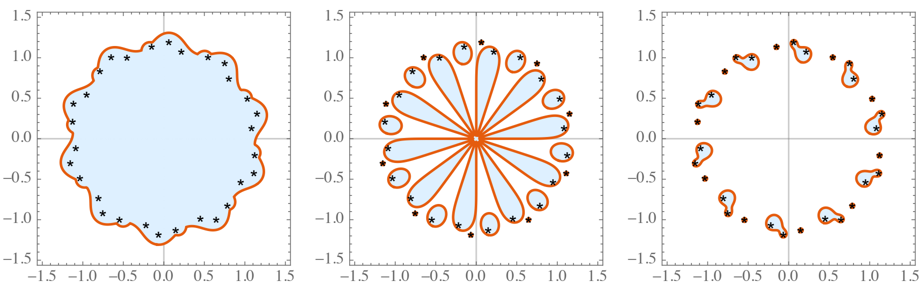

Figure 2: Householder sets of Example 3.7

The Householder set on the left corresponds to the Geršgorin set of $P(\lambda)$; the Householder set in the center corresponds to the weighted Geršgorin set of $P(\lambda)$, where $X = \operatorname{diag}(1, 1, 10)$; the Householder set on the right corresponds to $H(P_\Delta, P)$, where $\|\cdot\| = \|\cdot\|_\infty$ and $P_\Delta(\lambda)$ is an element of $\mathcal{B}(P, \varepsilon, \mathbf{w})$ as defined in (6), randomly selected with $\varepsilon = 0.075$ and $w_j = \|A_j\|_\infty$, for $j = 0, 1, \ldots, m$.

Note that the Householder set in the center illustrates the potential of the weighted Geršgorin set to sharpen the bounds on the original Geršgorin set, whereas the weights chosen in Example 3.6 made the bounds worse.

Example 3.8. Let

$$P(\lambda) = \begin{bmatrix} \lambda^2 - 2\lambda + 1 & 0 & \lambda \\ 0 & \lambda^2 - 1 & 0 \\ 0 & 0 & \lambda^2 + 1 \end{bmatrix}$$

and consider the three Householder sets in Figure 3.

The Householder set on the left corresponds to the Geršgorin set of $P(\lambda)$; the Householder set in the center corresponds to the weighted Geršgorin set of $P(\lambda)$, where $X = \operatorname{diag}(10, 1, 1)$; the Householder set on the right corresponds to $H(P_\Delta, P)$, where $\|\cdot\| = \|\cdot\|_\infty$ and $P_\Delta(\lambda)$ is an element of $\mathcal{B}(P, \varepsilon, \mathbf{w})$ as defined in (6), randomly selected with $\varepsilon = 0.075$ and $w_j = \|A_j\|_\infty$, for $j = 0, 1, \ldots, m$.

Note that the Householder set on the left and in the center has three isolated points. These three points are eigenvalues of $\operatorname{diag} P(\lambda)$, which we expected from the necessary condition in Theorem 2.4. However, there is an eigenvalue of $\operatorname{diag} P(\lambda)$ that is not isolated, thus, this condition is not sufficient. Finally, since $H(P_\Delta, P)$ is a subset the ε-pseudospectrum, the right part of the figure indicates that the eigenvalue 1 is more sensitive to perturbations than the other

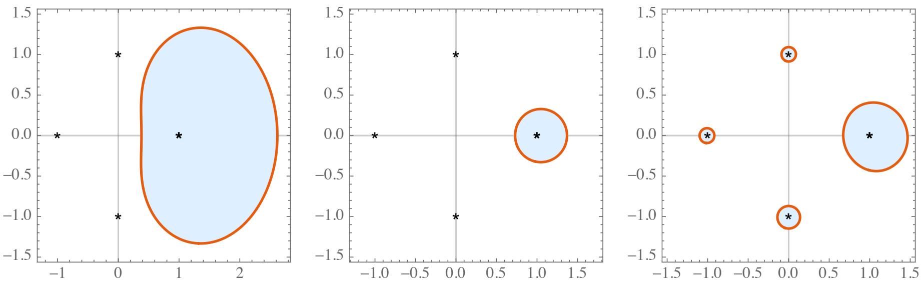

Figure 3: Householder sets of Example 3.8

eigenvalues of $P(\lambda)$. This increased sensitivity is expected from Theorem 2 of [14]; indeed, 1 is the only multiple eigenvalue of $P(\lambda)$ and it has algebraic multiplicity 3 and geometric multiplicity 2.

4. Bauer-Fike-Type Bounds

In Section 3, we showed that Householder sets can be used to give elegant derivations of other inclusion sets. In addition, these normed derived sets are intimately connected with the Bauer-Fike theorem, see Theorem IV.1.6 of [16], which in turn is deeply connected to the perturbation theory of matrices [18, 20]. In particular, the following proposition holds for any matrix polynomial $P(\lambda)$ as defined in (1).

Proposition 4.1. *Let $\Delta P(\lambda)$ be a matrix polynomial of size n and degree less than or equal to m, also let $P_\Delta(\lambda) = P(\lambda) + \Delta P(\lambda)$. If μ_Δ is an eigenvalue of $P_\Delta(\lambda)$ that is not an eigenvalue of $P(\lambda)$, then for any invertible $M \in \mathbb{C}^{n\times n}$ we have*

$$\left\| M^{-1} P(\mu_\Delta)^{-1} M \right\|^{-1} \leq \left\| M^{-1} \Delta P(\mu) M \right\|,$$

where $\|\cdot\|$ is any induced matrix norm.

Proof. Since $\mu_\Delta \in \sigma(P_\Delta)$ and $\mu \notin \sigma(P)$, it follows that

$$\mu_\Delta \in S(M^{-1} P_\Delta(\lambda) M, M^{-1} P(\lambda) M).$$

Therefore, $\left\| (M^{-1} P(\mu_\Delta)^{-1} M)(M^{-1} \Delta P(\mu_\Delta) M) \right\| \geq 1$, and the result follows. □

Note that the Bauer-Fike theorem is a special case of Proposition 4.1. To this end, let $P(\lambda) = A - \lambda I$ and $\Delta P(\lambda) = E$; then, the result in Proposition 4.1 is equivalent to the result in Theorem IV.1.6 of [16]. Moreover, this result immediately implies the Bauer-Fike theorem for diagonalizable matrices, see Theorem 6.3.2 of [6], provided that $\|\cdot\|$ is a matrix norm induced by an absolute norm on $\mathbb{C}^n$ as defined in (5.4.18) of [6].

Prior to extending the Bauer-Fike theorem for diagonalizable matrices to matrix polynomials, we note that some matrix polynomials can be diagonalized by congruence or strict equivalence, see [12]. While this requires the coefficient matrices satisfy a strong commutativity condition, there are many applications in engineering where this condition arises naturally. In this case, it is easy to see how Proposition 4.1 can be used to derive a generalized Bauer-Fike theorem for simultaneously diagonalizable matrix polynomials.

More generally, every matrix polynomial can be diagonalized as follows:

$$E(\lambda)P(\lambda)F(\lambda) = D(\lambda), \tag{7}$$

where $D(\lambda) = \operatorname{diag}(d_1(\lambda), \ldots, d_n(\lambda))$ and $E(\lambda)$ and $F(\lambda)$ are unimodular, that is, matrix polynomials with constant nonzero determinant [5]. In this more general setting, the following theorem holds.

Theorem 4.2. *Let $\Delta P(\lambda)$ be a matrix polynomial of size n and degree less than or equal to m, also let $P_\Delta(\lambda) = P(\lambda) + \Delta P(\lambda)$. Then, for each eigenvalue $\mu_\Delta \in \sigma(P_\Delta)$, there is a $d_i(\lambda)$ as in (7) such that*

$$|d_i(\mu_\Delta)| \leq \|E(\mu_\Delta)\| \, \|F(\mu_\Delta)\| \, \|\Delta P(\mu_\Delta)\| \,, \tag{8}$$

where $\|\cdot\|$ is any absolute induced matrix norm.

Proof. Let $\mu_\Delta \in \sigma(P_\Delta)$ and consider the Householder set of $E(\lambda)P_\Delta(\lambda)F(\lambda)$ with respect to $D(\lambda)$ and $\|\cdot\|$. By Theorem 2.1, it follows that $\mu_\Delta \in \sigma(D)$ or

$$\left\| D(\mu_\Delta)^{-1} E(\mu_\Delta) \Delta P(\mu_\Delta) F(\mu_\Delta) \right\| \geq 1.$$

In the former case, the result is trivial. Assuming the latter we have

$$\begin{aligned} 1 &\leq \left\| D(\mu_\Delta)^{-1} E(\mu_\Delta) \Delta P(\mu_\Delta) F(\mu_\Delta) \right\| \\ &\leq \|E(\mu_\Delta) \Delta P(\mu_\Delta) F(\mu_\Delta)\| \, \left\| D(\mu_\Delta)^{-1} \right\| \\ &= \|E(\mu_\Delta) \Delta P(\mu_\Delta) F(\mu_\Delta)\| \max_{1 \leq i \leq n} \left| d_i(\mu_\Delta)^{-1} \right|, \end{aligned}$$

where the last line follows from Theorem 5.6.36 of [6]. Therefore,

$$\begin{aligned} \min_{1 \leq i \leq n} |d_i(\mu_\Delta)| &\leq \|E(\mu_\Delta) \Delta P(\mu_\Delta) F(\mu_\Delta)\| \\ &\leq \|E(\mu_\Delta)\| \, \|F(\mu_\Delta)\| \, \|\Delta P(\mu_\Delta)\| \,. \end{aligned}$$

□

Note that the Bauer-Fike theorem for diagonalizable matrices is a special instance of Theorem 4.2. Indeed, let $P(\lambda) = \lambda I - A$ and $\Delta P(\lambda) = E$, where $A, E \in \mathbb{C}^{n \times n}$. If A is diagonalizable, then there exists an invertible $M \in \mathbb{C}^{n \times n}$ such that $M^{-1}(\lambda I - A) M = \operatorname{diag}(\lambda - \mu_1, \ldots, \lambda - \mu_n)$, where $\mu_1, \ldots, \mu_n$ are the eigenvalues of A. By Theorem 4.2, for each $\mu_\Delta \in \sigma(A + E)$, there exists a $\mu_i \in \sigma(A)$ such that

$$|\mu_\Delta - \mu_i| \leq \left\| M^{-1} \right\| \, \|M\| \, \|E\| \,.$$

Example 4.3. The following matrix polynomial is from [12]:

$$P(\lambda) = \lambda^2 \begin{bmatrix} 41 & 12 \\ 12 & 34 \end{bmatrix} + \lambda \begin{bmatrix} -73 & -36 \\ -36 & -52 \end{bmatrix} + \begin{bmatrix} 32 & 24 \\ 24 & 18 \end{bmatrix}.$$

Note that the coefficient matrices of $P(\lambda)$ are simultaneously diagonalizable by congruence. In particular, there exists a unitary matrix U such that

$$\begin{aligned} U^*P(\lambda)U &= \mathrm{diag}(50\lambda^2 - 100\lambda + 50, 25\lambda^2 - 25\lambda) \\ &:= \mathrm{diag}(d_1(\lambda), d_2(\lambda)). \end{aligned}$$

Let $\varepsilon > 0$ and $\mathbf{w} = \{w_1, w_2, w_3\}$ be nonnegative weights. Keeping in mind the notation of Section 3.3, for any $\mu_\Delta \in \sigma(P_\Delta)$, where $P_\Delta(\lambda) \in \mathcal{B}(P, \varepsilon, \mathbf{w})$, Theorem 4.2 implies that there is a $d_i(\lambda)$, $i \in \{1, 2\}$, such that

$$|d_i(\mu_\Delta)| \le \varepsilon q_{\mathbf{w}}(|\mu_\Delta|).$$

Next, we consider the case where the $D(\lambda)$ in (7) is the Smith form of $P(\lambda)$. Then, the diagonal entries $d_i(\lambda)$ are known as the invariant polynomials of $P(\lambda)$. Furthermore, each invariant polynomial can be represented as a product of linear factors

$$d_i(\lambda) = (\lambda - \mu_{i,1})^{\alpha_{i,1}} \cdots (\lambda - \mu_{i,k_i})^{\alpha_{i,k_i}}, \qquad i = 1, \ldots, n, \tag{9}$$

where $\mu_{i,1}, \ldots, \mu_{i,k_i}$ are distinct complex numbers and $\alpha_{i,1}, \ldots, \alpha_{i,k_i}$ are positive integers. The factors $(\lambda - \mu_{i,j})^{\alpha_{i,j}}$, $j = 1, \ldots, k_i$, $i = 1, \ldots, n$, are called the elementary divisors of $P(\lambda)$. Furthermore, an elementary divisor is called linear if $a_{i,j} = 1$, and nonlinear otherwise. Finally, it is clear that the complex numbers $\mu_{i,1}, \ldots, \mu_{i,k_i}$ are eigenvalues of $P(\lambda)$.

Now, let $\mu_\Delta \in \sigma(P_\Delta)$ as defined the hypothesis of Theorem 4.2. Then, there exists an invariant polynomial $d_i(\lambda)$ of $P(\lambda)$ that satisfies the bound in (8). Let $d_i(\lambda)$ be written as in (9) and select $\mu_{i,l}$, $l \in \{1, \ldots, k_i\}$, that minimizes $|\mu_\Delta - \mu_{i,l}|$. Furthermore, define

$$f(\lambda) = \prod_{\substack{j=1 \\ j \neq l}}^{k_i} (\lambda - \mu_{i,j})^{\alpha_{i,j}},$$

so that $d_i(\lambda) = f(\lambda)(\lambda - \mu_{i,l})^{\alpha_{i,l}}$. Applying the bounds in (8), it follows that

$$|\mu_\Delta - \mu_{i,l}| \le \left(\frac{\|E(\mu_\Delta)\| \, \|F(\mu_\Delta)\| \, \|\Delta P(\mu_\Delta)\|}{|f(\mu_\Delta)|} \right)^{1/\alpha_{i,l}}, \tag{10}$$

where $\alpha_{i,l}$ is no greater than the algebraic multiplicity of $\mu_{i,l}$ and $f(\lambda)$ is a polynomial that is nonzero in a neighborhood of μ_Δ.

Example 4.4. The following matrix polynomial is from [12]:

$$P(\lambda) = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

It can be shown, see [12], that $P(\lambda)$ cannot be diagonalized via congruence or strict equivalence. However, we can still use the Smith form of $P(\lambda)$ and (10) in order to analyze the affects of a perturbation on the eigenvalues of $P(\lambda)$.

To this end, consider the perturbed matrix polynomial

$$P_\Delta(\lambda) = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1+\varepsilon \\ 1-\varepsilon & 2 \end{bmatrix},$$

where $0 < \varepsilon < 1$. The eigenvalues of $P_\Delta(\lambda)$ are

$$-1 + \sqrt{\varepsilon} e^{k\pi i/4},$$

for $k = 1, 3, 5, 7$. Furthermore,

$$E(\lambda) = \begin{bmatrix} 0 & 1 \\ 1 & \lambda(2 + 3\lambda + \lambda^2) \end{bmatrix} \text{ and } F(\lambda) = \begin{bmatrix} -(\lambda+1) & (\lambda+1)^2 \\ 1 & -(\lambda+1) \end{bmatrix}$$

are unimodular matrix polynomials such that $E(\lambda)P(\lambda)F(\lambda) = \operatorname{diag}(1, (\lambda+1)^4)$. For any $\mu_\Delta \in \sigma(P_\Delta)$, it can easily be verified that

$$\|E(\mu_\Delta)\|_\infty \le 1 + 2\sqrt{\varepsilon} \text{ and } \|F(\mu_\Delta)\|_\infty \le 1 + 2\sqrt{\varepsilon}.$$

Keeping in mind the notation of Section 3.3, it follows that

$$\|P_\Delta(\mu_\Delta)\| \le \varepsilon q_{\mathbf{w}}(|\mu_\Delta|),$$

where $\mathbf{w} = \{1, 0, 0\}$. Applying the bounds in (10), we verify that

$$\begin{aligned} \sqrt{\varepsilon} = |\mu_\Delta - \mu| &\le \left(\varepsilon(1 + 4\sqrt{\varepsilon} + 4\varepsilon)\right)^{1/4} \\ &\le (9\varepsilon)^{1/4}. \end{aligned}$$

5. Conclusion

Householder sets for matrix polynomials have a wide range of interesting properties and applications. In Section 2, we introduced the Householder sets for matrix polynomials and analyzed their topological and algebraic properties. Then, in Section 3, we showed that instances of Householder sets could be used to derive other well-known inclusion sets for matrix polynomials. Specifically, we derived the Geršgorin set, weighted Geršgorin set, and weighted pseudospectra of a matrix polynomial. Finally, in Section 4, we showed that Householder sets are intimately connected to the Bauer-Fike theorem by using these sets to derive Bauer-Fike-type bounds for matrix polynomials. We note that our definition of Householder sets can easily be extended for analytic matrix-valued functions, as well as for consistent and compatible matrix norms; many of our results will be applicable in these more general settings. Future research includes the application of Householder sets to the analysis of additive and multiplicative perturbation theory for matrix polynomials, as well as an efficient method for computing the boundary of the Householder sets.

Appendix A. Subharmonic Functions

In this appendix, we give our working definition of subharmonic functions and prove some basic properties of subharmonic functions which are used in this article.

Let Ω be an open set in $\mathbb{C}$. The continuous function $f\colon \Omega \to \mathbb{R}$ is said to be *subharmonic* in Ω provided that for any closed disk $D(z,r) \subset \Omega$ of center z and radius r,

$$f(z) \leq \frac{1}{2\pi}\int_0^{2\pi} f(z+re^{i\theta})d\theta.$$

Proposition A.1. *Suppose $\{f_\alpha\colon\ \alpha \in A\}$ is a family of subharmonic functions in Ω that are bounded above. Then,*

$$f(\lambda) = \sup_{\alpha\in A} f_\alpha(\lambda)$$

is subharmonic on Ω.

Proof. The function f is clearly continuous. Now, consider the closed disk $D(z,r) \subset \Omega$. For any $\alpha \in A$ we have

$$\begin{aligned} f_\alpha(z) &\leq \frac{1}{2\pi}\int_0^{2\pi} f_\alpha(z+re^{i\theta})d\theta \\ &\leq \frac{1}{2\pi}\int_0^{2\pi} f(z+re^{i\theta})d\theta. \end{aligned}$$

Therefore,

$$f(z) = \sup_{\alpha\in A} f_\alpha(z) \leq \frac{1}{2\pi}\int_0^{2\pi} f(z+re^{i\theta})d\theta$$

and it follows that f is subharmonic in Ω. □

Proposition A.2. *Suppose that $f\colon \Omega \to \mathbb{R}$ is subharmonic in Ω and $\phi\colon \mathbb{R} \to \mathbb{R}$ is continuous, convex, and non-decreasing. Then, $\phi(f(\lambda))$ is subharmonic in Ω.*

Proof. Again, it is clear that $\phi(f(\lambda))$ is continuous. Furthermore, for any closed disk $D(z,r) \subset \Omega$, we have

$$\frac{1}{2\pi}\int_0^{2\pi} \phi(f(z+re^{i\theta}))d\theta \geq \phi\left(\frac{1}{2\pi}\int_0^{2\pi} f(z+re^{i\theta})d\theta\right) \geq \phi(f(z)).$$

The first inequality is a result of Jensen's inequality, see Theorem 3.3 of [15], and the second inequality is a result of f being subharmonic and ϕ being non-decreasing. □

Theorem A.3. *Let Ω be a region and suppose that $A\colon \Omega \to \mathbb{C}^{n\times n}$ is a non-zero analytic matrix-valued function. Furthermore, let $\|\cdot\|$ be any induced matrix norm. Then, $\log\|A(\lambda)\|$ and $\|A(\lambda)\|$ are both subharmonic functions in Ω.*

Proof. We need only to show that the former assertion is true since the latter will then follow from Proposition A.2. Let $\|\cdot\|$ denote the norm on $\mathbb{C}^n$ from which our matrix norm is induced. Furthermore, let $\|\cdot\|_D$ denote the dual of $\|\cdot\|$. By Theorem 5.5.9 and Theorem 5.6.2 of [6], it follows that

$$\|A(\lambda)\| = \max_{\|x\|=\|y\|_D=1} |y^* A(\lambda)x|.$$

For each $x, y \in \mathbb{C}^n \setminus \{0\}$, $y^* A(\lambda)x$ is a non-zero analytic function in Ω. Also, by Theorem 17.3 of [15], the function $\log |y^* A(\lambda)x|$ is subharmonic in Ω. Therefore, by Proposition A.1,

$$\log \|A(\lambda)\| = \max_{\|x\|=\|y\|_D=1} \log |y^* A(\lambda)x|$$

is subharmonic in Ω $\square$

References

[1] Ahmad, S.S., Alam, R., 2009. Pseudospectra, crtical points, and multiple eigenvalues of matrix polynomials. Linear Algebra Appl. 430, 1171–1195.

[2] Bindel, D., Hood, A., 2013. Localization theorems for nonlinear eigenvalue problems. SIAM J. Matrix Anal. Appl. 34, 1728–1749.

[3] Boulton, L., Lancaster, P., Psarrakos, P., 2008. On pseudospectra of matrix polynomials and their boundaries. Math. Comp. 77, 313–334.

[4] Geršgorin, S., 1931. Über die Abgrenzung der Eigenwerte einer matrix. Izv. Akad. Nauk SSR Ser. Mat. 1, 749–754.

[5] Gohberg, I., Lancaster, P., Rodman, L., 2009. Matrix Polynomials. SIAM, Philadelphia, PA.

[6] Horn, R.A., Johnson, C.R., 2013. Matrix Analysis. 2nd ed., Cambridge University Press, New York, New York.

[7] Householder, A.S., 1964. Theory of Matrices in Numerical Analysis. Blaisell, New York, New York.

[8] Hurewicz, W., Wallman, H., 1948. Dimension Theory (PMS-4). Princeton University Press, Princeton, NJ.

[9] Kostić, V., Cvetković, L.J., Varga, R.S., 2009. Geršgorin-type localizations of generalized eigenvalues. Numer. Linear Algebr. Appl. 16, 883–898.

[10] Kostić, V., Gardašević, D., 2018. On the Geršgorin-type localizations for nonlinear eigenvalue problems. Appl. Math. Comput. 37, 179–189.

[11] Lancaster, P., Psarrakos, P., 2005. On the pseudospectra of matrix polynomials. SIAM J. Matrix Anal. Appl. 27, 115–129.

[12] Lancaster, P., Zaballa, I., 2009. Diagonalizable quadratic eigenvalue problems. Mech. Systems Signal Process. 23, 1134–1144.

[13] Michailidou, C., Psarrakos, P., 2018. Gershgorin type sets for eigenvalues of matrix polynomials. Electron. J. Linear Algebra 34, 652–674.

[14] Papathanasiou, N., Psarrakos, P., 2010. On condition numbers of polynomial eigenvalue problems. Appl. Math. Comput. 216, 1194–1205.

[15] Rudin, W., 1987. Real and Complex Analysis. WCB McGraw-Hill, Boston, Massachusetts.

[16] Stewart, G.W., Sun, J.G., 1990. Matrix Perturbation Theory. Academic Press, Cambridge, MA.

[17] Tisseur, F., Higham, N., 2001. Structured pseudospectra for polynomial eigenvalue problems. SIAM J. Matrix Anal. Appl. 23, 187–208.

[18] Trefethen, L.N., Embree, M., 2005. Spectra and pseudospectra: the behavior of nonnormal matrices and operators. Princeton University Press, Princeton, New Jersey.

[19] Varga, R.S., 2004. Geršgorin and His Circles. Springer-Verlag, Berlin and Heidelberg.

[20] Wilkinson, J.H., 1965. The algebraic eigenvalue problem. numerical mathematics and scientific computation, Oxford University Press.