

POLYNOMIAL ROOT RADIUS OPTIMIZATION WITH AFFINE CONSTRAINTS

JULIA EATON*, SARA GRUNDEL, MERT GÜRBÜZBALABAN, AND MICHAEL L. OVERTON

ABSTRACT. The root radius of a polynomial is the maximum of the moduli of its roots (zeros). We consider the following optimization problem: minimize the root radius over monic polynomials of degree n , with either real or complex coefficients, subject to k consistent affine constraints on the coefficients. We show that there always exists an optimal polynomial with at most $k - 1$ inactive roots, that is, whose modulus is strictly less than the optimal root radius. We illustrate our results using some examples arising in feedback control.

1. INTRODUCTION

The root radius of a polynomial is the maximum of the moduli of its roots (zeros). This is a fundamental function that arises in many applications. Perhaps one of the most important application areas is feedback control in the frequency domain, which is modeled by rational system transfer functions. For a transfer function to be discrete-time stable, its denominator—a polynomial—must have its roots in the open unit disk, or equivalently, its root radius must be less than one.

When the coefficients of a polynomial family depend on parameters, it is natural to consider minimizing the polynomial root radius to obtain or enhance stability. However, such an optimization problem is nontrivial because the root radius is not convex, and furthermore it is not locally Lipschitz at points in polynomial space with multiple roots—more specifically, multiple *active* roots, that is, roots whose modulus equals the maximum modulus. Nonetheless, as we show in this paper, the global minimizers of these optimization problems have a remarkable structural property when dependence on the parameters is *affine*, or, equivalently, affine constraints are imposed on the polynomial coefficients. Specifically, for the polynomial root radius optimization problem over monic polynomials of degree n with either real or complex coefficients, if k consistent affine constraints are imposed on the coefficients,

Date: February 5, 2015.

2010 *Mathematics Subject Classification.* 93D15, 49K30, 90C26.

* Corresponding author.

Julia Eaton. Interdisciplinary Arts & Sciences, University of Washington Tacoma, Tacoma, WA 98402 (jreaton@uw.edu). Supported in part by the NSF-AWM Mentoring Travel Grant and the School of Interdisciplinary Arts & Sciences Research Fund.

Sara Grundel. Max Planck Institute for Dynamics of Complex Technical Systems, 39106 Magdeburg, Germany (grundel@mpi-magdeburg.mpg.de).

Mert Gürbüzbalaban, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Boston, MA 02139 (mertg@mit.edu).

Michael L. Overton. Courant Institute of Mathematical Sciences, New York University, New York, NY 10012 (overton@cims.nyu.edu). Supported in part by National Science Foundation grant DMS-1317205.

then there always exists an optimal polynomial with at most $k - 1$ inactive roots, that is, with at least $n - k + 1$ active roots.

As far as we know, such results have not been obtained or even contemplated previously, except that such a conjecture was made, based on experimental results, by Blondel, Gürbüzbalaban, Overton and Megretski [BGMO12, section V]. That paper, building on results of Chen in [Che79b] and [Che79a], obtained stronger results for the special case $k = 1$, concerning not only root *activity*, but also root *multiplicity*. Specifically, in the case of minimizing the root radius over polynomials with complex coefficients subject to just one affine constraint, it was shown in [BGMO12] that there always exists an optimal polynomial of the form $p(z) = (z - \gamma)^n$ for some $\gamma \in \mathbb{C}$, meaning not only that there are n active roots, but that they actually coincide. When the coefficients are restricted to be real, it was shown that there always exists an optimal polynomial of the form $p(z) = (z - \gamma)^m(z + \gamma)^{n-m}$ for some $\gamma \in \mathbb{R}$ and $m \in \{0, \dots, n\}$, implying the existence of an optimal polynomial with multiple roots when $n \geq 3$. In contrast, when $k > 1$, there may not exist an optimal polynomial with multiple roots.

As a concrete example, consider minimizing the root radius of the cubic polynomial $p(z) = z^3 + az + 1$ over $a \in \mathbb{R}$, or equivalently, imposing $k = 2$ affine constraints: the quadratic coefficient is 0 and the constant coefficient is 1. The minimal value 1 is attained by the unique optimal polynomial $p_*(z) = z^3 + 1$ with three active roots, one more than required by our general theorem, but these are the three cube roots of unity and hence are distinct; multiple active roots do not occur. In contrast, consider minimizing the root radius of the polynomial $p(z) = z^3 + bz^2 + az + 1$ over $a \in \mathbb{R}, b \in \mathbb{R}$, allowing both the linear and quadratic coefficients to vary, or equivalently, imposing just one constraint: the constant coefficient is 1. Then the minimal value remains 1, but the optimal polynomial is no longer unique; any polynomial with $a = b \in [-1, 3]$ has root radius 1, and in particular, when $a = b = -1$, there are two active roots at 1 and one at -1 , while for $a = b = 3$, all three active roots coincide at -1 .

The proofs of the results given in [BGMO12] for a single constraint were completely different for the cases of real and complex coefficients, with the latter being far more complicated than the former. Our proof of root activity when there are several affine constraints uses a more general approach that applies to both real and complex coefficients and recovers the activity (but not the multiplicity) results for the case where there is just one constraint.

The root radius is one of a class of functions known as max root functions for polynomials. A great deal is known about variational analysis of polynomial max root functions [BLO05, BE12]. That analysis focuses on the challenges presented by the non-Lipschitz behavior of these functions around polynomials with multiple active roots, a phenomenon studied in one variable by Newton using his “Newton diagram” or “Newton polygon” (see [MBO97]). However, the focus there is on describing the variational properties of max root functions when the active root multiplicities are known. The present work is concerned with what active root structure is possible at minimizers.

The paper is organized as follows. We start in Section 2 by proving the simplest case of our main result, namely, when the number of affine constraints k is just one less than the polynomial degree n and the coefficients are restricted to be real. Section 3 lays the groundwork for the general case, establishing some key lemmas for optimization of polynomial max root functions. Section 4 then presents our main results for optimization of the root

radius, applicable to both real and complex coefficients. Section 5 presents some examples that arise in feedback control, illustrating our results. Section 6 makes some concluding remarks.

2. THE SIMPLEST CASE

Let ρ denote the root radius of a polynomial, that is, the largest of the moduli of its roots. Our result is easily obtained in the case that k , the number of affine constraints, is $n - 1$, particularly when the coefficients are restricted to be real, so we treat this case first. Imposing $n - 1$ consistent, linearly independent affine constraints on a monic real polynomial of degree $n \geq 2$ is equivalent to parametrizing the coefficients with just one variable, that is our polynomial p is in the set

$$\{p(z, \lambda) = p_0 + \lambda p_1 : \lambda \in \mathbb{R}\}$$

where p_0 is a fixed monic polynomial of degree n and p_1 is a fixed nonzero polynomial with degree at most $n - 1$, not necessarily monic. For each λ , let $\gamma_1(\lambda), \dots, \gamma_n(\lambda)$ denote the roots of $p(z, \lambda)$, ordered first by highest-to-lowest modulus and second by lowest-to-highest angle, where angles are measured from the positive real axis, and repeated according to multiplicity. Thus, we have the factorization

$$(1) \quad p(z, \lambda) = p_0(z) + \lambda p_1(z) = \prod_{i=1}^n (z - \gamma_i(\lambda))$$

and the minimization problem becomes

$$\inf_{\lambda \in \mathbb{R}} \rho(p(z, \lambda)) \equiv \inf_{\lambda \in \mathbb{R}} |\gamma_1(\lambda)|.$$

We claim that there exists an optimal polynomial with at least $n - k + 1 = 2$ active roots, that is, at least two roots attaining the maximum modulus value. Since the lower level sets of ρ are compact, an optimal polynomial $p_* = p(z, \lambda^*)$ exists, and by definition, p_* has at least one active root, $\gamma_1(\lambda^*)$. If $|\gamma_1(\lambda^*)| = 0$, then all of the roots of p are zero and there is nothing more to show. So suppose $|\gamma_1(\lambda^*)| > 0$. If the multiplicity of $\gamma_1(\lambda^*)$ is at least two, then there are at least two active roots, and if $\gamma_1(\lambda^*)$ is complex, then its complex conjugate must also be a root of $p(z, \lambda)$, which implies there are at least two active roots. In either case there is nothing more to show. So we may assume that $\gamma_1(\lambda^*)$ is a simple real root, which is therefore differentiable with respect to λ at λ^* . We have $\gamma_1(\lambda^*) = |\gamma_1(\lambda^*)|$ or $\gamma_1(\lambda^*) = -|\gamma_1(\lambda^*)|$, so that the first-order optimality condition $\partial_{\lambda|\lambda=\lambda^*}(|\gamma_1(\lambda)|) = 0$ implies that

$$(2) \quad \partial_{\lambda|\lambda=\lambda^*}(\gamma_1(\lambda)) = 0.$$

Differentiating (1) with respect to λ on both sides, evaluating the result at $\lambda = \lambda^*$, and applying (2) leads to

$$p_1(z) = (z - \gamma_1(\lambda^*))r_1(z),$$

where

$$r_1(z) := \partial_{\lambda|\lambda=\lambda^*} \prod_{i=2}^n (z - \gamma_i(\lambda)).$$

Therefore,

$$p(z, \lambda^*) = p_0(z) + \lambda^* p_1(z) = p_0(z) + \lambda^* (z - \gamma_1(\lambda^*)) r_1(z).$$

Substituting $z = \gamma_1(\lambda^*)$ yields $p_0(\gamma_1(\lambda^*)) = 0$, which implies that $p_0(z)$ also has a root at $\gamma_1(\lambda^*)$. Therefore,

$$p(z, \lambda) = (z - \gamma_1(\lambda^*)) (r_0(z) + \lambda r_1(z)),$$

where $r_0(z)$ is such that $p_0(z) = (z - \gamma_1(\lambda^*))r_0(z)$. Note that $\rho(r_0(z) + \lambda^* r_1(z)) < |\gamma_1(\lambda^*)|$, since otherwise p_* would already have two active roots. We also have

$$\liminf_{|\lambda| \rightarrow \infty} \rho(r_0(z) + \lambda r_1(z)) = \infty,$$

since the lower level sets of ρ are compact and r_1 cannot be zero as p_1 is not zero. Therefore, by continuity there must exist some $\tilde{\lambda} \in \mathbb{R}$ for which

$$\rho(r_0(z) + \tilde{\lambda} r_1(z)) = |\gamma_1(\lambda^*)|.$$

Then the polynomial

$$(z - \gamma_1(\lambda^*)) (r_0(z) + \tilde{\lambda} r_1(z)) = p_0(z) + \tilde{\lambda} p_1(z)$$

attains the optimal radius $|\gamma_1(\lambda^*)| = \rho(p_*)$ and has at least two active roots.

This result can be similarly derived for polynomials with complex coefficients, but we omit the details for brevity and move on to the general case.

3. POLYNOMIAL MAX ROOT FUNCTIONS ON AFFINE SETS

Let $n \in \mathbb{N}$ be fixed. Let $\mathcal{P}_n(\mathbb{F})$ denote the vector space of polynomials of degree n or less, and let $\mathcal{P}_n^1(\mathbb{F}) \subset \mathcal{P}_n(\mathbb{F})$ denote the set of monic polynomials of degree n over \mathbb{F} , where \mathbb{F} is the field \mathbb{C} or \mathbb{R} .

Let $f : \mathbb{C} \rightarrow \mathbb{R}$. The function f induces a function $\varphi_f : \mathcal{P}_n(\mathbb{F}) \rightarrow \mathbb{R}$ in the following way:

$$(3) \quad \varphi_f(p) = \max\{f(\zeta) : p(\zeta) = 0\}.$$

We call φ_f a *polynomial max root function*, or simply *max root function*. If f is continuous, so is φ_f . Two polynomial max root functions that arise in applications are the polynomial root radius $\rho : \mathcal{P}_n^1(\mathbb{F}) \rightarrow \mathbb{R}$, given by

$$(4) \quad \rho(p) = \max\{|\zeta| : p(\zeta) = 0\};$$

and the polynomial root abscissa $\alpha : \mathcal{P}_n^1(\mathbb{F}) \rightarrow \mathbb{R}$, given by

$$\alpha(p) = \max\{\operatorname{Re}(\zeta) : p(\zeta) = 0\}.$$

Polynomial max root functions are, in general, non-convex and non-locally-Lipschitz, as seen by the following example: consider the polynomial family $\{p_\varepsilon(z) : p_\varepsilon(z) = z^n - \varepsilon, \varepsilon > 0\}$. Then $\rho(p_\varepsilon) = \alpha(p_\varepsilon) = \sqrt[n]{\varepsilon}$, so the root radius and abscissa are non-convex and non-locally-Lipschitz for $n > 1$. We say a root ζ of p is *active* relative to φ_f if

$$f(\zeta) = \varphi_f(p).$$

Otherwise ζ is *inactive* relative to φ_f , which corresponds to the case where $f(\zeta) < \varphi_f(p)$.

Let $k \in \{1, \dots, n\}$ be a fixed number of affine constraints on the coefficients of monic polynomials of degree n . We represent the affine constraints as follows. Let

$$(5) \quad \mathcal{A}_{\mathbb{F}} = \left\{ z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0 : \begin{array}{l} B_n^{(\ell)} + \sum_{i=0}^{n-1} B_i^{(\ell)} a_i = 0, \text{ for } \ell = 1, \dots, k, \\ \text{where } a_i \in \mathbb{F} \text{ for } i = 0, \dots, n-1. \end{array} \right\}$$

denote the affine subspace of $\mathcal{P}_n^1(\mathbb{F})$ satisfying the k constraints. We assume the k constraints are consistent, so that $\mathcal{A}_{\mathbb{F}} \neq \emptyset$, and that they are linearly independent, so that the dimension of $\mathcal{A}_{\mathbb{F}}$ is $n - k$.

An alternative representation of $\mathcal{A}_{\mathbb{F}}$ is in the coefficient space \mathbb{F}^n as follows. Let $p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0$. Then $p \in \mathcal{A}_{\mathbb{F}}$ if and only if $Ba = b$, where $a = (a_0, \dots, a_{n-1})^T$,

$$(6) \quad B = \begin{bmatrix} B_0^{(1)} & B_1^{(1)} & \dots & B_{n-1}^{(1)} \\ B_0^{(2)} & B_1^{(2)} & \dots & B_{n-1}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ B_0^{(k)} & B_1^{(k)} & \dots & B_{n-1}^{(k)} \end{bmatrix}, \quad \text{and} \quad b = \begin{bmatrix} -B_n^{(1)} \\ -B_n^{(2)} \\ \vdots \\ -B_n^{(k)} \end{bmatrix}.$$

Then $\dim(\mathcal{A}_{\mathbb{F}}) = n - k$ requires that $\text{rank}(B) = k$. Let $\mathbb{A} \subset \mathbb{F}^n$ denote the set of feasible coefficients:

$$(7) \quad \mathbb{A} = \{a : Ba - b = 0\}.$$

We use both representations of the feasible set.

We are interested in problems of the form

$$\inf_{p \in \mathcal{A}_{\mathbb{F}}} \varphi_f(p),$$

where \mathbb{F} is \mathbb{R} or \mathbb{C} . Our results concern the number of active and inactive roots of optimal solutions. They depend on the fundamental perturbation result developed in Section 3.2 below.

3.1. Preliminary result for linear transformations.

Lemma 3.1 (Linear transformation rank lemma). *Let $k, \ell, n \in \mathbb{N}$ and $r \in \mathbb{N} \cup \{0\}$ with $r \leq k \leq \ell \leq n$. Let $F : \mathbb{F}^{\ell} \rightarrow \mathbb{F}^n$ and $G : \mathbb{F}^n \rightarrow \mathbb{F}^k$ be linear transformations with $\text{rank}(F) = \text{rank}(G) = k$. Then $\dim(\text{range}(F) \cap \text{null}(G)) = r$ if and only if $\text{rank}(G \circ F) = k - r$.*

This is proved in Appendix A.1.

Corollary 3.2. *Let $k, \ell, n \in \mathbb{N}$ and $r \in \mathbb{N} \cup \{0\}$ with $r \leq k \leq \ell \leq n$. Let $F : \mathbb{F}^{\ell} \rightarrow \mathbb{F}^n$ and $G : \mathbb{F}^n \rightarrow \mathbb{F}^k$ be linear transformations with $\text{rank}(F) = \text{rank}(G) = k$. Let $A_F \subset \mathbb{F}^n$ be an affine space that is a translate of $\text{range}(F)$ by a fixed vector and let $A_G \subset \mathbb{F}^n$ be an affine space that is a translate of $\text{null}(G)$ by the same vector. Then*

$$\dim(A_F \cap A_G) = r,$$

if and only if

$$\text{rank}(G \circ F) = k - r,$$

where the dimension of an affine set $A \subset \mathbb{F}^n$ is taken to be $\dim(A - \{a\})$ for any $a \in A$.

3.2. Results for inactive root perturbations. In this section we establish a lemma for inactive root perturbations that is utilized by the main result of this paper. First we lay the groundwork with the necessary definitions. Let $f : \mathbb{C} \rightarrow \mathbb{R}$ be a continuous function, and let $\varphi_f : \mathcal{P}_n^1(\mathbb{F}) \rightarrow \mathbb{R}$ be its associated max root function, as defined in (3). Let $\mathcal{A}_{\mathbb{F}} \subset \mathcal{P}_n^1(\mathbb{F})$ be as defined in (5), so that $\dim(\mathcal{A}_{\mathbb{F}}) = n - k$, where the number of constraints k satisfies $1 \leq k < n$. Note that the other cases $k = n$ or $k > n$ are trivial since the feasible set reduces to a singleton or to the empty set.

Let $p_* \in \mathcal{A}_{\mathbb{F}}$ where

$$(8) \quad p_*(z) = z^n + a_{n-1}^* z^{n-1} + \cdots + a_0^*$$

with function value

$$(9) \quad \gamma_* := \varphi_f(p_*).$$

Since $p_* \in \mathcal{A}_{\mathbb{F}}$, we have $Ba^* = b$, where B and b are defined in (6). Suppose further that p_* has

$$(10) \quad n_I \geq k$$

inactive roots and

$$(11) \quad n_A := n - n_I$$

active roots. Factor p_* according to root activity as

$$(12) \quad p_*(z) = q^A(z)q^I(z),$$

where $q^A(z)$ is a polynomial of degree n_A whose roots are all active, and $q^I(z)$ is a polynomial of degree $n_I \geq k$ whose roots are all inactive. Express $q^I(z)$ by

$$(13) \quad q^I(z) = z^{n_I} + q_{n_I-1} z^{n_I-1} + \cdots + q_0.$$

Consider linear perturbations $\Delta \in \mathbb{F}^k$ of the first k coefficients of $q^I(z)$, where the indexing goes from 0 to $k-1$:

$$(14) \quad q_{\Delta}^I(z) = z^{n_I} + q_{n_I-1} z^{n_I-1} + \cdots + q_k z^k + (q_{k-1} + \Delta_{k-1}) z^{k-1} + \cdots + (q_0 + \Delta_0).$$

Multiplying $q_{\Delta}^I(z)$ by $q^A(z)$ yields linear perturbations of the coefficients of $p_*(z)$:

$$(15) \quad p_{\Delta}(z) = q^A(z)q_{\Delta}^I(z) = z^n + a_{n-1}(\Delta)z^{n-1} + \cdots + a_0(\Delta).$$

We can therefore express the *inactive-root coefficient perturbation map* $C^I : \mathbb{F}^k \rightarrow \mathbb{F}^n$ by

$$(16) \quad C^I(\Delta) = a^* + M\Delta,$$

so that each coefficient is perturbed as

$$(17) \quad a_i(\Delta) = a_i^* + (M\Delta)_i,$$

for $i = 0, \dots, n-1$, where $M \in \mathbb{F}^{n \times k}$ (whose rows and columns are indexed from 0 to $n-1$ and 0 to $k-1$, respectively), and $C^I(0) = (a_0^*, a_1^*, \dots, a_{n-1}^*)^T$ are the coefficients of $p_*(z)$ as in (8). Since f is continuous, small perturbations of q^I do not affect the active part, that is,

$$\varphi_f(p_*) = \varphi_f(p_{\Delta})$$

for small Δ . Let $H \subset \mathbb{F}^n$ denote the image of C^I , that is,

$$(18) \quad H := \{a^* + Mv : v \in \mathbb{F}^k\},$$

which is an affine subspace of \mathbb{F}^n of dimension less than or equal k .

Lemma 3.3 (Inactive perturbation, I). *Let f and φ_f be as in (3), with f continuous. Let $p_* \in \mathcal{A}_{\mathbb{F}}$, and let γ_* , q^I , q_{Δ}^I , C^I , and H be as in (8)–(18). Then*

$$(19) \quad \text{rank}(M) = k,$$

so that

$$\dim(H) = k.$$

Proof. Note that we consider dimension over \mathbb{F} . Let $\Delta \in \mathbb{F}^k$ be fixed. Straightforward computations show

$$\begin{aligned} M\Delta = 0 &\iff M\Delta + a^* = a^* \\ &\iff p_\Delta = p_* \\ &\iff q_\Delta^I(z) - q^I(z) = 0 \quad \forall z \\ &\iff \Delta_{k-1}z^{k-1} + \dots + \Delta_0 = 0 \quad \forall z \\ &\iff \Delta = 0, \end{aligned}$$

which implies that $\dim \text{null}(M) = 0$, that is, $\text{rank}(M) = k$, which establishes (19). \square

Corollary 3.4 (Inactive perturbation, II). *Let f and φ_f be as in (3), with $f : \mathbb{C} \rightarrow \mathbb{R}$ continuous. Let $p_* \in \mathcal{A}_{\mathbb{F}}$, and let γ_* , q^I , q_Δ^I , C^I , and H be as in (8)–(18). Then*

$$\dim(\mathbb{A} \cap H) = r \iff \text{rank}(BM) = k - r,$$

where B and \mathbb{A} are given in (6) and (7), and where $\dim(\mathbb{A} \cap H)$ means the dimension of the subspace $\mathbb{A} \cap H - \{a^*\}$ where $a^* = (a_0^*, a_1^*, \dots, a_{n-1}^*)^T$ is given in (8) and $0 \leq r \leq k$.

Proof. This follows from Corollary 3.2 and Lemma 3.3. \square

Our primary interest in Corollary 3.4 is in the cases $r = 0$ and $r = k$, which will be used to prove the results in the following section.

4. OPTIMIZING THE ROOT RADIUS

We now confine our attention to the root radius, one of the most important polynomial max root functions, which arises, for example, in the analysis of stability of discrete-time dynamical systems. The root radius $\rho : \mathcal{P}_n(\mathbb{F}) \rightarrow \mathbb{R}$ is defined by

$$(20) \quad \rho(p) = \max\{|\zeta| : p(\zeta) = 0\}.$$

Consider the problem

$$(21) \quad \inf_{p \in \mathcal{A}_{\mathbb{F}}} \rho(p),$$

where $\mathcal{A}_{\mathbb{F}} \subset \mathcal{P}_n^1(\mathbb{F})$ is defined in (5). Define the closed σ -Schur stable set in $\mathcal{P}_n^1(\mathbb{F})$ by

$$S_\sigma^{\mathbb{F}, \rho} = \{p \in \mathcal{P}_n^1(\mathbb{F}) : \rho(p) \leq \sigma\},$$

and denote its boundary by

$$\text{bdry}(S_\sigma^{\mathbb{F}, \rho}) = \{p \in \mathcal{P}_n^1(\mathbb{F}) : \rho(p) = \sigma\}.$$

Note that $S_\sigma^{\mathbb{F}, \rho}$ is compact for all $\sigma \in \mathbb{R}$ and that it is nonempty if and only if $\sigma \geq 0$. It is known that the boundary of the Schur stable set consists of ruled surfaces [FM78],[HP05, Section 4.1.3], as illustrated by Figure 1 for the case $\mathbb{F} = \mathbb{R}$, $n = 3$ and $\sigma = 1/2$. (A similar illustration appears in [HP05] for $\sigma = 1$.)

Theorem 4.1 (Root radius over real and complex coefficients). *Let $n \in \mathbb{N}$ and $k \in \{1, \dots, n\}$. Consider the root radius minimization problem (21) with the feasible set $\mathcal{A}_{\mathbb{F}}$ (defined in (5)), where \mathbb{F} is \mathbb{R} or \mathbb{C} and k is the number of consistent, linearly independent affine constraints, so that $\dim(\mathcal{A}_{\mathbb{F}}) = n - k$. Then there exists an optimal solution with at least $n - k + 1$ active roots, or equivalently, with at most $k - 1$ inactive roots.*

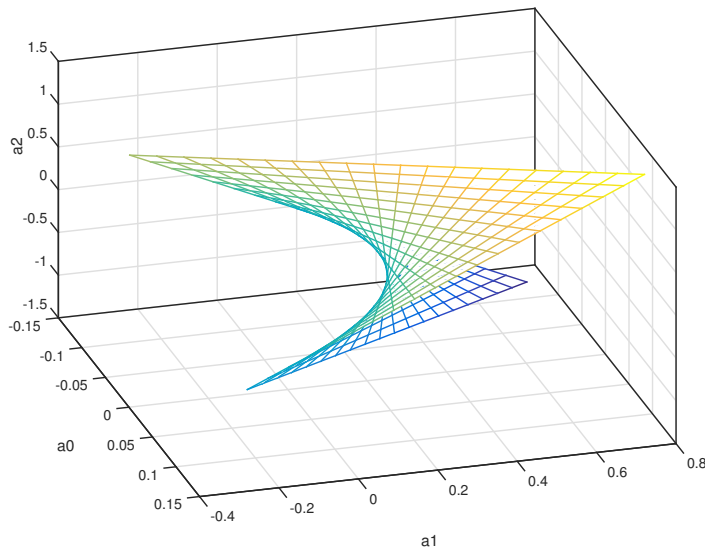


FIGURE 1. An illustration of the boundary of the Schur stable set $S_\sigma^{\mathbb{R}, \rho}$ for $n = 3$ and $\sigma = 1/2$, identifying $\mathcal{P}_3^1(\mathbb{R})$ with \mathbb{R}^3 .

Proof. The hypothesis that $\mathcal{A}_{\mathbb{F}}$ is nonempty implies a feasible polynomial exists. By the compactness of $S_\sigma^{\mathbb{F}, \rho}$ for each $\sigma \in \mathbb{R}$ and the continuity of ρ , an optimal solution exists. Let $p_* = z^n + a_{n-1}^* z^{n-1} + \dots + a_0^*$ be an optimizer with optimal value γ_* . Since $p_* \in \mathcal{A}_{\mathbb{F}}$, we have $Ba^* = b$. Note that p_* has at least one active root and at most $n - 1$ inactive roots. In the case where $k = n$, we have that $\mathcal{A}_{\mathbb{F}} = \{p_*\}$ and there is nothing more to show. So suppose $k < n$.

If $\rho(p_*) = 0$, then $p_* = z^n$, which implies that every root is active, establishing the result in this case. Suppose, therefore, that $\rho(p_*) > 0$ and that there are $n_I \geq k$ inactive roots and $n_A = n - n_I$ active roots (as in (11)). Factor p_* according to root activity as in (12): $p_*(z) = q^A(z)q^I(z)$, and express $q^I(z)$ as in (13) by $q^I(z) = z^{n_I} + q_{n_I-1}z^{n_I-1} + \dots + q_0$. In the real case, where $\mathbb{F} = \mathbb{R}$, express the active part in terms of its real roots and complex conjugate pairs by

$$q^A(z) = \prod_{i=1}^{n_1} (z - h_i) \prod_{i=n_1+1}^{n_2} (z^2 - 2d_i z + \gamma_*^2),$$

where

$$(22) \quad h_i \in \mathbb{R}, |h_i| = \gamma_* \quad \text{for } i = 1, \dots, n_1, \quad \text{and} \quad d_i \in \mathbb{R}, |d_i| < \gamma_* \quad \text{for } i = n_1 + 1, \dots, n_2;$$

note that n_1 and n_2 satisfy $n_A = n_1 + 2(n_2 - n_1)$. In the complex case, where $\mathbb{F} = \mathbb{C}$, express the active part in terms of its linear factors by

$$q^A(z) = \prod_{i=1}^{n_2} (z - h_i)$$

where

$$(23) \quad h_i \in \mathbb{C}, \quad |h_i| = \gamma_* \quad \text{for } i = 1, \dots, n_2,$$

where we define $n_2 := n_A$ for notational consistency. Consider affine perturbations $\Delta \in \mathbb{F}^k$ of the coefficients of $q^I(z)$ as in (14): $q_\Delta^I(z) = z^{n_I} + q_{n_I-1}z^{n_I-1} + \dots + q_k z^k + (q_{k-1} + \Delta_{k-1})z^{k-1} + \dots + (q_0 + \Delta_0)$. As in (15), multiplying $q_\Delta^I(z)$ by $q^A(z)$ yields $p_\Delta(z)$ —a family of affine perturbations of the coefficients of $p_*(z)$ that influence only the inactive roots when Δ is small. We utilize the inactive-root coefficient perturbation map $C^I : \mathbb{F}^k \rightarrow \mathbb{F}^n$ as given in (16), where $C^I(\Delta) = a^* + M\Delta$ and $a_i(\Delta) = a_i^* + (M\Delta)_i$, so that $C^I(0) = (a_0^*, a_1^*, \dots, a_{n-1}^*)^T$ are the coefficients of $p_*(z)$ as defined in (8). Let $H \subset \mathbb{F}^n$ be the image of C^I as defined in (18).

By Corollary 3.4,

$$(24) \quad \dim(\mathbb{A} \cap H) = 0 \iff \text{rank}(BM) = k.$$

In the real case, where $\mathbb{F} = \mathbb{R}$, consider perturbations $\tilde{\Delta} \in \mathbb{R}^{n_2}$ of the active part

$$q_\Delta^A(z) := \prod_{i=1}^{n_1} (z - h_i(1 + \tilde{\Delta}_i)) \prod_{i=n_1+1}^{n_2} (z^2 - 2d_i z + \gamma_*^2(1 + \tilde{\Delta}_i)),$$

where h_i and d_i are given in (22). In the complex case, where $\mathbb{F} = \mathbb{C}$, consider perturbations $\tilde{\Delta} \in \mathbb{C}^{n_2}$ of the active part

$$q_\Delta^A(z) := \prod_{i=1}^{n_2} (z - h_i(1 + \tilde{\Delta}_i))$$

where h_i is given in (23).¹

Let

$$p_{\tilde{\Delta}, \Delta}(z) = q_\Delta^A(z)q_\Delta^I(z) = z^n + a_{n-1}(\tilde{\Delta}, \Delta)z^{n-1} + \dots + a_0(\tilde{\Delta}, \Delta),$$

the perturbation of p_* obtained by perturbing both $q^I(z)$ and the active part $q^A(z)$. Define the expanded perturbed coefficient map $C : \mathbb{F}^{n_2} \times \mathbb{F}^k \rightarrow \mathbb{F}^n$ by

$$C(\tilde{\Delta}, \Delta) = \left(a_0(\tilde{\Delta}, \Delta), \dots, a_{n-1}(\tilde{\Delta}, \Delta) \right)^T,$$

where each perturbed coefficient $a_i(\tilde{\Delta}, \Delta)$ is multilinear in the components of $\tilde{\Delta}$ and Δ . Note that $C(0, \Delta) = C^I(\Delta) = a^* + M\Delta$, where a^* is given in (8), and C^I and M are given in (16). Also note that $C(0, 0) = a^*$. Define the maps $L : \mathbb{F}^{n_2} \times \mathbb{F}^k \rightarrow \mathbb{F}^k$ and $L^I : \mathbb{F}^k \rightarrow \mathbb{F}^k$ by

$$L(\tilde{\Delta}, \Delta) = BC(\tilde{\Delta}, \Delta) - b \quad \text{and} \quad L^I(\Delta) = BC^I(\Delta) - b.$$

Then $L^I(0) = BC^I(0) - b = Ba^* - b = 0$. Note that

$$(25) \quad D_{L^I}(0) = BM,$$

¹It is worth noting that in this proof, we perturb the *roots* of the active part, as opposed to the *coefficients* of the active part, as was done in [BGM012, Theorem 1].

the real or complex² derivative of L^I evaluated at 0. The proof proceeds by making a rank argument about this derivative. The component Δ does not affect the value of $\rho(p_{\tilde{\Delta}, \Delta})$ provided Δ and $\tilde{\Delta}$ are small, since Δ corresponds to perturbations of the inactive part only as the roots are continuous in the coefficients and the complex modulus function is continuous. Suppose that $\dim(\mathbb{A} \cap H) = 0$ (in which case $\mathbb{A} \cap H = \{a^*\}$). By (24), $\text{rank}(BM) = k$. This implies $\text{rank}(D_{L^I}(0)) = k$. By the implicit function theorem³ there exists $(\tilde{\Delta}, \Delta) \in \mathbb{F}^{n_2} \times \mathbb{F}^k$ near the origin satisfying $L(\tilde{\Delta}, \Delta) = 0$, so that $p_{\tilde{\Delta}, \Delta}$ is feasible. Furthermore, in the real case, we can choose $\tilde{\Delta}_i < 0$ for all $i \in \{1, 2, \dots, n_2\}$, so that the perturbed factors $(z - h_i(1 + \tilde{\Delta}_i))$ for $i = 1, \dots, n_1$, yield roots of modulus strictly less than γ_* , as do the roots corresponding to the perturbed conjugate pairs $(z^2 - 2d_i z + \gamma_*^2(1 + \tilde{\Delta}_i))$, for $i = n_1 + 1, \dots, n_2$; in the complex case, we can choose $\tilde{\Delta}_i$ to be real and strictly negative for $i \in \{1, 2, \dots, n_2\}$, so that the perturbed factors $(z - h_i(1 + \tilde{\Delta}_i))$ for $i = 1, \dots, n_2$, yield roots of modulus strictly less than γ_* . Therefore $\rho(p_{\tilde{\Delta}, \Delta}) < \gamma_*$, which contradicts the optimality of p_* . Hence $\dim(\mathbb{A} \cap H) \geq 1$. So there exists $v \neq 0$ such that $a^* + Mv \in \mathbb{A} \cap H$. In fact, $a^* + M(tv) \in \mathbb{A} \cap H$ for all $t \in \mathbb{R}$, i.e. $C^I(tv) \in \mathbb{A} \cap H$. Furthermore, for $|t|$ near 0, $C^I(tv) \in \mathbb{A} \cap H \cap \text{bdry}(S_{\gamma_*}^{\mathbb{R}, \rho})$.

Consider the one-parameter polynomial family

$$r_t(z) := q^A(z)q_{\Delta}^I(z) \quad \text{with } \Delta = tv, t \in \mathbb{R}.$$

By construction, r_t is feasible for all t since its coefficients lie in \mathbb{A} . Let $f(t) = \rho(q_{\Delta}^I)$, the root radius of the perturbed inactive part. Then f is clearly continuous and $f(0) < \gamma_*$. The lower level sets of ρ are bounded, and so f is unbounded as $|t| \rightarrow \infty$. By continuity of f , there exists $t_* \in \mathbb{R}$ such that $f(t_*) = \gamma_*$. This implies that the perturbed polynomial $q_{\Delta_*}^I$, where $\Delta_* = t_*v$, has at least one active root so the polynomial

$$p_{\text{new}}(z) = q^A(z)q_{\Delta_*}^I(z)$$

has at least $n_A + 1$ active roots, i.e., at least one more active root than p_* , while having the same root radius as p_* . Let s denote the increase in the number of active roots. Then $s \geq 1$, and the number of inactive roots of p_{new} equals $n_I - s$. If $n_I - s \geq k$, we can apply the same procedure to the polynomial p_{new} which is re-factored to reflect the change in activity as

$$p_{\text{new}} = q_{\text{new}}^A q_{\text{new}}^I$$

with $\deg(q_{\text{new}}^A) = n_A + s$ and $\deg(q_{\text{new}}^I) = n - (n_A + s)$. This process is repeated as many times as necessary and must terminate with an optimal polynomial with at most $k - 1$ inactive roots. This completes the proof. \square

Remark 4.2 (local optimality). *Although our primary interest is in globally optimal polynomials, the activity result of Theorem 4.1 holds for locally optimal polynomials as well.*

Remark 4.3. *If $\dim \text{null}(D_{L^I}(0)) = k$ where D_{L^I} is given in (25), then applying Corollary 3.4 with $r = k$ leads to $H \subseteq \mathbb{A}$, so that any perturbation $\Delta \in \mathbb{F}^k$ is feasible, where \mathbb{A} and H*

²The components of L^I are affine with respect to each complex variable and therefore their derivatives satisfy the Cauchy-Riemann equations.

³For a complex version see [FG02, Theorem 7.6].

are defined in (7) and (18), respectively. Thus, there exists an optimizer such that all roots are active.

Remark 4.4. *Though we only perturbed a particular subset of k coefficients of q^l , we could have perturbed any k of the coefficients. We chose to perturb only the first k for notational ease. In the complex case, another approach is to factor the inactive part into two monic polynomials—one with exactly degree k —and perturb only the coefficients of that polynomial.*

5. EXAMPLES FROM FEEDBACK CONTROL

Let $a \in \mathcal{P}_{d_a}^1(\mathbb{R})$ and $b \in \mathcal{P}_{d_b}(\mathbb{R})$ be given, with $d_a > d_b$. In control one refers to the rational function b/a as a *plant* or *open loop system* that is *strictly proper* as $d_a > d_b$. In feedback control design, one wishes to construct another rational function y/x , called a *controller*, so that the *closed loop system*, modeled by the rational function $q/p \equiv bx/(ax+by)$ [BG04, HP05, Ran89] has desired properties, most notably with poles, that is roots of $p = ax + by$, in the relevant stability region. Let $x \in \mathcal{P}_{d_x}^1(\mathbb{R})$ and $y \in \mathcal{P}_{d_y}(\mathbb{R})$ with y/x *proper*, that is $d_x \geq d_y$, and write

$$\begin{aligned} a(z) &= z^{d_a} + a_{d_a-1}z^{d_a-1} + \cdots + a_0, & b(z) &= b_{d_b}z^{d_b} + b_{d_b-1}z^{d_b-1} + \cdots + b_0, \\ x(z) &= z^{d_x} + x_{d_x-1}z^{d_x-1} + \cdots + x_0, & y(z) &= y_{d_y}z^{d_y} + y_{d_y-1}z^{d_y-1} + \cdots + y_0. \end{aligned}$$

Then ax is the monic polynomial

$$z^{d_p} + (a_{d_a-1} + x_{d_x-1})z^{d_p-1} + (a_{d_a-2} + a_{d_a-1}x_{d_x-1} + x_{d_x-2})z^{d_p-2} + \cdots + a_0x_0$$

with degree $d_p = d_a + d_x$, and with coefficients depending affinely on the d_x variables x_{d_x-1}, \dots, x_0 . Clearly, the polynomial $p = ax + by$ also has degree d_p and depends affinely on the $d_x + d_y + 1$ variables $x_{d_x-1}, \dots, x_0, y_{d_y}, \dots, y_0$. If we now minimize the root radius of p over these variables, we have an optimization problem of the form (21), with $n = d_p$ and **at most**

$$k = (d_a + d_x) - (d_x + d_y + 1) = d_a - d_y - 1$$

affine constraints on the coefficients of p . **Note that if a and b have no nonconstant common factors, then this optimization problem has exactly k affine constraints.**

We do not have a method to find global minimizers of max root optimization problems so we approximated them using a local optimization method, as was done in [Ove14] for matrix eigenvalue optimization problems. As explained in [LO13], the quasi-Newton method known as BFGS, which originated in 1970 to minimize differentiable functions [NW00], is also extremely effective for finding local minimizers of nonsmooth functions, particularly in the locally Lipschitz case, but also including non-Lipschitz functions such as the root radius, although the same accuracy cannot be expected in the latter case. So, to search for minimizers of $\rho(p)$, where p is given above, we ran BFGS from 1000 randomly generated starting points for each problem instance, with the hope that for small problem instances, global minima will be found, although as noted in Remark 4.2, our result on root activity also applies to local minima. The experiments were conducted using MATLAB [Mat14].

Let $a(z) = z^5 + z^3 - z + 2$ and let $b(z) = z^2 + 1$, and let us restrict attention to the case that the degrees of x and y are the same, say $d = d_x = d_y$. We consider problem instances with

d	$n = d + 5$	$\#\text{vars} = 2d + 1$	$\#\text{constr} = 4 - d$	$\text{min active} = \min(2d + 2, n)$	actual active
0	5	1	4	2	3
1	6	3	3	4	6
2	7	5	2	6	6
3	8	7	1	8	8
4	9	9	0	9	9

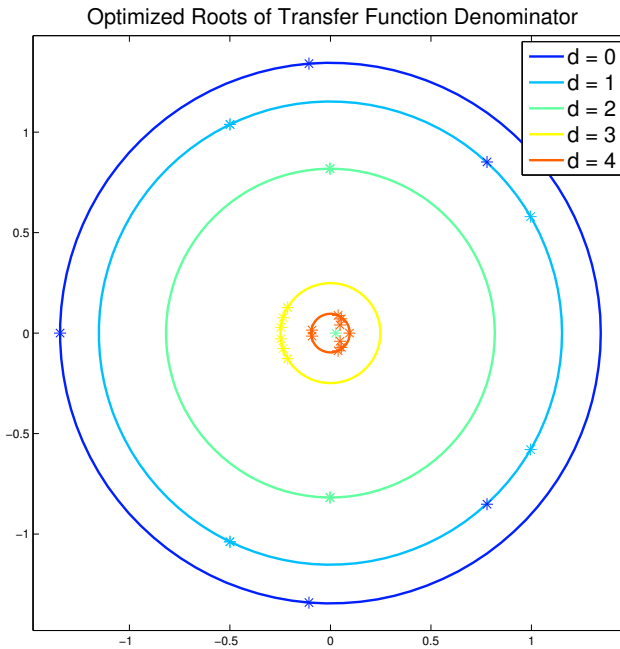
TABLE 1. Experiment with $a(z) = z^5 + z^3 - z + 2$ and $b(z) = z^2 + 1$ 

FIGURE 2. Optimized Roots of Transfer Function Denominator

d ranging from 0 to 4. Table 1 shows the corresponding values for $n = d_p = d_a + d = d + 5$, the number of variables $2d + 1$, the number of constraints $k = (d + 5) - (2d + 1) = 4 - d$, the minimum number of active roots of the optimal polynomial⁴ according to Theorem 4.1, namely $n - k + 1 = 2d + 2$, and the actual number of active roots (within a tolerance) of the polynomial with least root radius found by BFGS. We see that in each case, the computational results are consistent with the theory.

For each value of d , Figure 5 shows the actual roots of the computed optimal polynomial \tilde{p} together with the associated circle of radius $\rho(\tilde{p})$. We see that for $d = 0$, there are two inactive roots, and for $d = 2$, there is one inactive root; otherwise all roots are active within a small tolerance. Note that for each $d > 0$, some of the optimal roots are coincident, but

⁴Assuming that it is unique. The instance with $d = 4$ reported in the last row of the table has no constraints, so the optimal root radius is zero; hence in this case the minimal number of active roots is $n = 9$.

the number varies, and when $d = 0$, all five optimal roots are distinct. In the case $d = 3$, with just one constraint, the optimal root radius found by BFGS agrees well with the known globally optimal value obtained using the algorithm described in [BGMO12]. In this case, all the computed roots are clustered near one real number, although one cannot expect them to be exactly coincident, as they are in theory, given the known sensitivity of multiple roots (in this case with multiplicity 8) to small numerical errors. In the case $d = 4$, when there are no constraints, the optimal polynomial is z^9 with all roots at zero, but again, but we cannot expect to be able to compute this very accurately.

Finally, note from Figure 5 that when d is increased, the optimal root radius decreases. This must always be the case, as a simple argument shows. Suppose x, y are optimal for a given degree d . If we now increase d by one, then the polynomial $z(a(z)x(z) + b(z)y(z))$ has the same roots as the previous optimal polynomial $a(z)x(z) + b(z)y(z)$ along with one additional root at zero. So, it has the same radius, and since $z(x(z))$ is monic, optimizing over $x(z)$ and $y(z)$ with increased degree can only reduce the radius further, not increase it.

6. CONCLUDING REMARKS

In this paper, we have presented a remarkable structural property of solutions of polynomial root radius optimization problems subject to affine constraints. An equally important polynomial max root function is the root abscissa, defined as the maximum of the real parts of the roots; this arises in stability analysis of continuous-time dynamical systems. Optimization of the root abscissa of a polynomial subject to affine constraints is a more difficult problem to analyze because, in the case of real coefficients, the infimal abscissa value may not be attained by any optimal polynomial. It was shown in [BGMO12] that in the single-constraint case, the infimal value may be approximated arbitrarily accurately by a sequence of polynomials with one set of coincident real roots converging to the optimal value and another set of coincident real roots diverging to $-\infty$. Because of the difficulties presented by this issue, we did not consider the polynomial root abscissa in this paper, but leave analysis of this function to future work. More generally, it would be interesting to try to extend our results to all polynomial max root functions.

Another natural idea is to consider extensions from polynomial roots to matrix eigenvalues, but this seems difficult. As with polynomial max root functions, a lot is known about variational analysis of spectral max functions such as the spectral radius and spectral abscissa [BO01, BLO01]. However, simple examples indicate that results analogous to Theorem 4.1 do not hold for affine matrix families. The reason for this is that although the eigenvalues of a matrix are the roots of the characteristic polynomial, a matrix family depending affinely on a set of parameters does not correspond to an affine characteristic polynomial family, except in special cases such as a companion matrix family. For example, let

$$A_0 = \begin{bmatrix} 0.5000 & 0 \\ 0 & -1.0000 \end{bmatrix} \quad \text{and} \quad A_1 = \begin{bmatrix} -1.0000 & -2.0000 \\ -0.5000 & 1.0000 \end{bmatrix}.$$

The unique global minimizer of the spectral radius of $A_0 + \lambda A_1$ over $\lambda \in \mathbb{R}$ has just one active eigenvalue, so even the simplest case described in Section 2 does not extend from polynomial roots to matrix eigenvalues. The most important question for future work is whether our results could lead to an efficient algorithm for global minimization of polynomial max root functions subject to affine constraints. In the case of a single constraint, the results

in [BGMO12] guarantee that a minimizer of the form $(z - \gamma)^n$ (in the complex case) or $(z - \gamma)^m(z + \gamma)^{n-m}$ (in the real case) exists, and combining this knowledge with the given constraint leads immediately to efficient computation of γ (and m , in the real case).

However, when there is more than one constraint, we have no such formula for an optimal solution. Indeed, it is quite possible that global polynomial max root optimization with more than one constraint is hard in a theoretical sense, as is known to be the case for a number of related problems in control [BT97, BT00].

APPENDIX

A.1. Proof of Lemma 3.1.

Proof. The proof has two parts: the case where $r = 0$ and $r > 0$. Note that the condition that $\ell \geq k$ is only necessary so that $\text{rank}(F) = k$ is possible. Beyond that, ℓ is not used in the argument below.

Suppose $\dim(\text{range}(F) \cap \text{null}(G)) = r = 0$. Let $\{v_1, \dots, v_k\} \subset \mathbb{F}^n$ be a basis for $\text{range}(F)$. Then $\text{rank}(G \circ F) = \dim(\text{span}\{G(v_1), \dots, G(v_k)\}) \leq k$. Suppose $\text{rank}(G \circ F) < k$. Then there exist $c_1, \dots, c_k \in \mathbb{F}$ with

$$(26) \quad \sum_{i=1}^k |c_i| > 0$$

such that $\sum_{i=1}^k c_i G(v_i) = 0$, which implies that $G(\sum_{i=1}^k c_i v_i) = 0$ by the linearity of G , which implies that $\sum_{i=1}^k c_i v_i \in \text{null}(G)$. But $\text{range}(F) \cap \text{null}(G) = \{0\}$, which implies $\sum_{i=1}^k |c_i| = 0$ since $\{v_1, \dots, v_k\} \subset \mathbb{F}^n$ is a linearly independent set, which contradicts (26). Therefore $\text{rank}(G \circ F) = k$.

Next suppose $\text{rank}(G \circ F) = k$. Let $m = \dim(\text{range}(F) \cap \text{null}(G))$. Note that $m \leq k$. If $m \geq 1$, there exists a basis $\{u_1, \dots, u_m\} \subset \mathbb{F}^n$ of $\text{range}(F) \cap \text{null}(G)$. Let $\{v_1, \dots, v_{k-m}\} \subset \mathbb{F}^n$ be either the empty set or a linearly independent set such that $\{u_1, \dots, u_m, v_1, \dots, v_{k-m}\}$ is a basis for $\text{range}(F)$. Then, since $G(u_1) = \dots = G(u_m) = 0$,

$$k = \text{rank}(G \circ F) = \dim(\text{span}\{G(u_1), \dots, G(u_m), G(v_1), \dots, G(v_{k-m})\}) \leq k - m,$$

which implies that $m \leq 0$. Therefore $m = 0$, and so $\dim(\text{range}(F) \cap \text{null}(G)) = 0$.

Next, suppose $\dim(\text{range}(F) \cap \text{null}(G)) = r \geq 1$. Note that $r \leq k$. Let $u_1, \dots, u_r \in \mathbb{F}^n$ be a basis for $\text{range}(F) \cap \text{null}(G)$ and let $\{v_1, \dots, v_{k-r}\} \subset \mathbb{F}^n$ be either the empty set or linearly independent such that $\{u_1, \dots, u_r, v_1, \dots, v_{k-r}\}$ is a basis for $\text{range}(F)$. Then, since $G(u_1) = \dots = G(u_r) = 0$,

$$\text{rank}(G \circ F) = \dim(\text{span}\{G(u_1), \dots, G(u_r), G(v_1), \dots, G(v_{k-r})\}) \leq k - r.$$

Suppose $\text{rank}(G \circ F) < k - r$. Then there exist $c_1, \dots, c_{k-r} \in \mathbb{F}$ with

$$(27) \quad \sum_{i=1}^{k-r} |c_i| > 0$$

such that $\sum_{i=1}^{k-r} c_i G(v_i) = 0$, which implies that $G(\sum_{i=1}^{k-r} c_i v_i) = 0$ by the linearity of G , which implies that $\sum_{i=1}^{k-r} c_i v_i \in \text{range}(F) \cap \text{null}(G)$. This further implies that $\sum_{i=1}^{k-r} c_i v_i \in$

$\text{span}\{u_1, \dots, u_r\}$, which can only occur if $\sum_{i=1}^{k-r} |c_i| = 0$ since $\{u_1, \dots, u_r, v_1, \dots, v_{k-r}\} \subset \mathbb{F}^n$ is a linearly independent set, which contradicts (27). Therefore $\text{rank}(G \circ F) = k - r$.

Next suppose $\text{rank}(G \circ F) = k - r$ for some $r \in \{1, 2, \dots, k\}$. Let $m = \dim(\text{range}(F) \cap \text{null}(G))$. If $m \geq 1$, there exists a basis $u_1, \dots, u_m \in \mathbb{F}^n$ for $\text{range}(F) \cap \text{null}(G)$. Let $\{v_1, \dots, v_{k-m}\} \subset \mathbb{F}^n$ be either the empty set or a linearly independent set such that $\{u_1, \dots, u_m, v_1, \dots, v_{k-m}\}$ is a basis for $\text{range}(F)$ which is of dimension k by the assumption. Then, since $G(u_1) = \dots = G(u_m) = 0$,

$$k - r = \text{rank}(G \circ F) = \dim(\text{span}\{G(u_1), \dots, G(u_m), G(v_1), \dots, G(v_{k-m})\}) \leq k - m,$$

which implies that $m \leq r$. Suppose $m < r$. Since $\text{rank}(G \circ F) = k - r$, $\{G(v_1), \dots, G(v_{k-m})\} \subset \mathbb{F}^k$ is a linearly dependent set. So there exists $c_1, \dots, c_{k-m} \in \mathbb{F}$ with

$$(28) \quad \sum_{i=1}^{k-m} |c_i| > 0$$

such that $\sum_{i=1}^{k-m} c_i G(v_i) = 0$, which implies that $G(\sum_{i=1}^{k-m} c_i v_i) = 0$ by the linearity of G . This implies that $\sum_{i=1}^{k-m} c_i v_i \in \text{range}(F) \cap \text{null}(G) = \text{span}\{u_1, \dots, u_m\}$, which can only occur if $\sum_{i=1}^{k-m} |c_i| = 0$ since $\{u_1, \dots, u_r, v_1, \dots, v_{k-m}\} \subset \mathbb{F}^n$ is a linearly independent set, which contradicts (28). Therefore $m = r$, that is, $\dim(\text{range}(F) \cap \text{null}(G)) = r$. \square

REFERENCES

- [BE12] J.V. Burke and J. Eaton. On the subdifferential regularity of max root functions for polynomials. *Nonlinear Analysis: Theory, Methods and Applications*, 75(3):1168–1187, February 2012.
- [BG04] D.N. Burghes and A. Graham. *Control and optimal control theories with applications*. Elsevier, 2004.
- [BGM012] V.D. Blondel, M. Gürbüzbalaban, A. Megretski, and M.L. Overton. Explicit solutions for root optimization of a polynomial family with one affine constraint. *Automatic Control, IEEE Transactions on*, 57(12):3078–3089, Dec 2012.
- [BLO01] J.V. Burke, A.S. Lewis, and M.L. Overton. Optimal stability and eigenvalue multiplicity. *Found. Comput. Math.*, 1(2):205–225, 2001.
- [BLO05] J.V. Burke, A.S. Lewis, and M.L. Overton. Variational analysis of functions of the roots of polynomials. *Mathematical Programming*, 104:263–292, 2005.
- [BO01] J.V. Burke and M.L. Overton. Variational analysis of non-Lipschitz spectral functions. *Math. Program.*, 90(2, Ser. A):317–351, 2001.
- [BT97] V.D. Blondel and J.N. Tsitsiklis. NP-hardness of some linear control design problems. *SIAM Journal on Control and Optimization*, 35:2118–2127, 1997.
- [BT00] V.D. Blondel and J.N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica*, 36(9):1249–1274, 2000.
- [Che79a] R. Chen. On the problem of direct output feedback stabilization. *Proceedings of MTNS*, pages 412–414, 1979.
- [Che79b] R. Chen. *Output feedback stabilization of linear systems*. PhD thesis, University of Florida, Gainesville, 1979.
- [FG02] K. Fritzsche and H. Grauert. *From Holomorphic Functions to Complex Manifolds*. Graduate Texts in Mathematics. Springer, New York, 2002.
- [FM78] A.T. Fam and J.S. Meditch. A canonical parameter space for linear systems design. *Automatic Control, IEEE Transactions on*, 23(3):454–458, 1978.
- [HP05] D. Hinrichsen and A.J. Pritchard. *Mathematical systems theory I: modelling, state space analysis, stability and robustness*, volume 1. Springer, 2005.

- [LO13] A.S. Lewis and M.L. Overton. Nonsmooth optimization via quasi-newton methods. *Mathematical Programming*, 141:135–163, 2013.
- [Mat14] MathWorks Inc. MATLAB, 2014. www.mathworks.com/products/matlab/.
- [MBO97] J. Moro, J.V. Burke, and M.L. Overton. On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *SIAM J. Matrix Anal. Appl.*, 18(4):793–817, 1997.
- [NW00] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, 2nd edition, 2000.
- [Ove14] M.L. Overton. Stability optimization for polynomials and matrices. In O. Kirillov and D. Pelinovsky, editors, *Nonlinear Physical Systems: Spectral Analysis, Stability and Bifurcations*, pages 351–375. John Wiley and Sons, Inc., 2014.
- [Ran89] A. Rantzer. Equivalence between stability of partial realizations and feedback stabilization—applications to reduced order stabilization. *Linear Algebra and its Applications*, 122:641–653, 1989.