

Explicit Solutions for Root Optimization of a Polynomial Family with One Affine Constraint

Vincent D. Blondel[‡], Mert Gurbuzbalaban[‡], Alexander Megretski[‡]
and Michael L. Overton[‡]

Abstract

Given a family of real or complex monic polynomials of fixed degree with one affine constraint on their coefficients, consider the problem of minimizing the root radius (largest modulus of the roots) or root abscissa (largest real part of the roots). We give constructive methods for efficiently computing the globally optimal value as well as an optimal polynomial when the optimal value is attained and an approximation when it is not. An optimal polynomial can always be chosen to have at most two distinct roots in the real case and just one distinct root in the complex case. Examples are presented illustrating the results, including several fixed-order controller optimal design problems.

I. INTRODUCTION

A fundamental general class of problems is as follows: given a set of monic polynomials of degree n whose coefficients depend on parameters, determine a choice for these parameters for which the polynomial is stable, or show that no such stabilization is possible. Variations on this

V.D. Blondel, Department of Mathematical Engineering, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. Email: vincent.blondel@uclouvain.be.

M. Gurbuzbalaban, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA. Supported in part by the National Science Foundation under grant DMS-1016325. Email: mert@courant.nyu.edu.

A. Megretski, Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. Email: ameg@mit.edu.

M.L. Overton, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA. Supported in part by the National Science Foundation under grant DMS-1016325 and in part by the Belgian Network DYSCO during a visit to Université catholique de Louvain. Email: overton@cs.nyu.edu.

[‡] These authors contributed equally to this work.

stabilization problem have been studied for more than half a century and several were mentioned in [BGL95] as being among the “major open problems in control systems theory”.

In this paper, we show that there is one important special case of the polynomial stabilization problem which is explicitly solvable: when the dependence on parameters is affine and the number of parameters is $n - 1$, or equivalently, when there is a single affine constraint on the coefficients. In this setting, regardless of whether the coefficients are allowed to be complex or restricted to be real, the problem of globally minimizing the root radius (defined as the maximum of the moduli of the roots) or root abscissa (maximum of the real parts) may be solved efficiently, even though the minimization objective is nonconvex and not Lipschitz continuous at minimizers. The polynomial is Schur (respectively Hurwitz) stabilizable if and only if the globally minimal value of the root radius (abscissa) is less than one (zero). This particular class of polynomial stabilization problems includes at least two interesting control problems. The first is the classical static output feedback stabilization problem in state space with one input and $m - 1$ independent outputs, where m is the system order [Che79a]. The second is a frequency-domain stabilization problem for a controller of order $m - 2$ [Ran89, p. 651]. In the second case, if stabilization is not possible, then the minimal order required for stabilization is $m - 1$. How to compute the minimal such order in general is a long-standing open question.

As a specific continuous-time example, consider the classical two-mass-spring dynamical system. It was shown in [HO06] that the minimal order required for stabilization is 2 and that the problem of maximizing the closed-loop asymptotic decay rate in this case is equivalent to the optimization problem

$$\min_{p \in P} \max_{z \in \mathbb{C}} \{\operatorname{Re} z \mid p(z) = 0\}$$

where

$$P = \{(z^4 + 2z^2)(x_0 + x_1z + z^2) + y_0 + y_1z + y_2z^2 \mid x_0, x_1, y_0, y_1, y_2 \in \mathbb{R}\}.$$

Thus P is a set of monic polynomials with degree 6 whose coefficients depend affinely on 5 parameters. A construction was given in [HO06] of a polynomial with one distinct root with multiplicity 6 and its local optimality was proved using techniques from nonsmooth analysis. Theorem 7 below validates this construction in a more general setting and proves global optimality.

The global minimization methods just mentioned are explained in a sequence of theorems that we present below. Theorem 1 shows that in the discrete-time case with real coefficients, the optimal polynomial can always be chosen to have at most two distinct roots, regardless of n , while Theorem 6 shows that in the discrete-time case with complex coefficients, the optimal polynomial can always be chosen to have just one distinct root. The continuous-time case is more subtle, because the globally infimal value of the root abscissa may not be attained. Theorem 7 shows that if it is attained, the corresponding optimal polynomial may be chosen to have just one distinct root, while Theorem 13 treats the case in which the optimal value is not attained. As in the discrete-time case, two roots play a role, but now one of them may not be finite. More precisely, the globally optimal value of the root abscissa may be arbitrarily well approximated by a polynomial with two distinct roots, only one of which is bounded. Finally, Theorem 14 shows that in the continuous-time case with complex coefficients, the optimal value is always attained by a polynomial with just one distinct root.

Our work was originally inspired by a combination of numerical experiments and mathematical analysis of special cases reported in [BLO01], [BHLO06], [HO06]. As we began investigating a more general theory, A. Rantzer drew our attention to a remarkable 1979 Ph.D. thesis of Raymond Chen [Che79b], which in fact derived a method to compute the globally infimal value of the abscissa in the continuous-time case with real coefficients. Chen also obtained some key related results for the discrete-time case with real coefficients, as explained in detail below. However, he did *not* provide generally applicable methods for constructing globally optimal or approximately optimal solutions, indeed remarking that he was lacking such methods [Che79b, p. 29 and p. 71]. Neither did he consider the complex case, for which it is a curious fact that our theorems are easier to state but apparently harder to prove than in the real case when the globally optimal value is attained.

The theorems just summarized are presented in Sections II and III for the discrete-time and continuous-time cases, respectively. The algorithms implicit in the theorems are implemented in a publicly available MATLAB code. Examples illustrating various cases, including the subtleties involved when the globally optimal abscissa is not attained, are presented in Section IV. We make some concluding remarks about possible generalizations in Section V.

II. DISCRETE-TIME STABILITY

Let $\rho(p)$ denote the *root radius* of a polynomial p ,

$$\rho(p) = \max \{|z| \mid p(z) = 0, z \in \mathbb{C}\}.$$

The following result shows that when the root radius is minimized over monic polynomials with real coefficients subject to a single affine constraint, the optimal polynomial can be chosen to have at most two distinct roots (zeros), and hence at least one multiple root when $n > 2$.

Theorem 1: Let B_0, B_1, \dots, B_n be real scalars (with B_1, \dots, B_n not all zero) and consider the affine family of monic polynomials

$$P = \{z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n \mid B_0 + \sum_{j=1}^n B_j a_j = 0, a_i \in \mathbb{R}\}.$$

The optimization problem

$$\rho^* := \inf_{p \in P} \rho(p)$$

has a globally optimal solution of the form

$$p^*(z) = (z - \gamma)^{n-k} (z + \gamma)^k \in P$$

for some integer k with $0 \leq k \leq n$, where $\gamma = \rho^*$.

Proof: Existence of an optimal solution is easy. Take any $p_0 \in P$ and define $P_0 = \{p \in P \mid \rho(p) \leq \rho(p_0)\}$. The set P_0 is bounded and closed. Since $\inf_{p \in P} \rho(p) = \inf_{p \in P_0} \rho(p)$, optimality is attained for some $p \in P_0 \subseteq P$.

We now prove the existence of an optimal solution that has the claimed structure. Let

$$p(z) = \prod_{i=1}^{n_1} (z + c_i) \prod_{i=n_1+1}^{n_2} (z^2 + 2d_i z + e_i)$$

be an optimal solution with $n_1 + 2(n_2 - n_1) = n$, $c_i, d_i, e_i \in \mathbb{R}$, $e_i > |d_i|$ and $\rho(p) = r$. We first show that there is an optimal solution whose roots all have magnitude r . Consider therefore the perturbed polynomial

$$\begin{aligned} p_\Delta(z) &= \prod_{i=1}^{n_1} (z + c_i(1 + \Delta_i)) \prod_{i=n_1+1}^{n_2} (z^2 + 2d_i z + e_i(1 + \Delta_i)) \\ &= z^n + a_1(\Delta) z^{n-1} + \dots + a_{n-1}(\Delta) z + a_n(\Delta), \end{aligned}$$

with $p_\Delta \in P$. The function

$$L(\Delta) = B_0 + B_1 a_1(\Delta) + \dots + B_{n-1} a_{n-1}(\Delta) + B_n a_n(\Delta)$$

is a multilinear function from \mathbb{R}^{n_2} to \mathbb{R} and it satisfies $L(0) = 0$. Observe that the case $n_2 = 1$ can occur only if $n = 1$ or $n = 2$ and in that case the result is easy to verify, so assume that $n_2 \geq 2$. Consider now a perturbation Δ_j associated with a root or a conjugate pair of roots that do not have maximal magnitude (i.e., $1 \leq j \leq n_1$ and $|c_j| < r$, or $n_1 + 1 \leq j \leq n_2$ and $e_j < r^2$), and define

$$\mu_j := \frac{\partial L}{\partial \Delta_j}(0).$$

If $\mu_j \neq 0$ then by the implicit function theorem one can find some Δ in a neighborhood of the origin for which $\Delta_i < 0$ for $i \neq j$ with $L(\Delta) = 0$ and therefore for which $\rho(q_\Delta) < \rho^*$, contradicting the optimality of q . On the other hand, if $\mu_j = 0$, then, since L is linear in Δ_j , we have $L(0, \dots, 0, \Delta_j, 0, \dots, 0) = L(0) = 0$ for all Δ_j , and so Δ_j can be chosen so that the corresponding root or conjugate pair of roots has magnitude exactly equal to r . Thus, an optimal polynomial whose roots have equal magnitudes can always be found.

If $r = 0$, the result is established, so in what follows suppose that $r > 0$. We need to show that all roots can be chosen to be real. We start from some optimal solution whose roots have magnitude $r > 0$, say

$$p(z) = \prod_{i=1}^{n_1} (z^2 + 2d_i z + r^2) \prod_{i=1}^{n_2} (z + r) \prod_{i=1}^{n_3} (z - r),$$

with $d_i \in \mathbb{R}$. Consider the perturbed polynomial

$$\begin{aligned} p_\Delta(z) &= \prod_{i=1}^{n_1} (z^2 + 2d_i(1 + \Delta_{2i})z + r^2(1 + \Delta_{2i-1})) \times \\ &\quad \prod_{i=1}^{n_2} (z + r(1 + \Delta_{2n_1+i})) \prod_{i=1}^{n_3} (z - r(1 + \Delta_{2n_1+n_2+i})) \\ &= z^n + a_1(\Delta)z^{n-1} + \dots + a_{n-1}(\Delta)z + a_n(\Delta), \end{aligned}$$

now including a perturbation to d_i , so the function

$$L(\Delta) = B_0 + B_1 a_1(\Delta) + \dots + B_{n-1} a_{n-1}(\Delta) + B_n a_n(\Delta)$$

is now a multilinear function from \mathbb{R}^n to \mathbb{R} that satisfies $L(0) = 0$. Let j be an index $1 \leq j \leq n_1$ for which $d_j \neq \pm r$ and define

$$\mu_j := \frac{\partial L}{\partial \Delta_{2j}}(0).$$

If $\mu_j \neq 0$ then by the same argument as above one can find a value of Δ in the neighborhood of the origin for which $\Delta_i < 0$ for $i \neq 2j$ with $L(\Delta) = 0$ and therefore for which $\rho(p_\Delta) < r$,

which contradicts the optimality of p . So we must have $\mu_j = 0$. But then Δ_{2j} can be modified as desired while preserving the condition $L(\Delta) = 0$ and so in particular it may be chosen so that $d_i(1 + \Delta_{2j}) = \pm r$. Repeated application of this argument leads to a polynomial $p^*(z)$ whose roots are all $\pm r$. ■

Notice that $p^*(z) \in P$ if and only if γ satisfies a certain polynomial equality once k is fixed. The following corollary is a direct consequence of this fact, showing that γ in Theorem 1 can be computed explicitly.

Corollary 2: Let γ be the globally optimal value whose existence is asserted in Theorem 1, and consider the set

$$\Xi = \{r \in \mathbb{R} \mid g_k(r) = 0 \text{ for some } k \in \{0, 1, \dots, n\}\}$$

where

$$g_k(z) = B_0 v_0 + B_1 v_1 z + \dots + B_{n-1} v_{n-1} z^{n-1} + B_n v_n z^n$$

and (v_0, \dots, v_n) is the convolution of the vectors

$$\left(\binom{n-k}{0}, \binom{n-k}{1}, \dots, \binom{n-k}{n-k} \right) \text{ and } \left(\binom{k}{0}, -\binom{k}{1}, \dots, (-1)^k \binom{k}{k} \right)$$

for $k = 0, \dots, n$. Then, $-\gamma$ is an element of Ξ with smallest magnitude.

Although Theorem 1 and Corollary 2 are both new, they are related to results in [Che79b], as we now explain. Let

$$H_P = \{(a_1, a_2, \dots, a_n) \in \mathbb{R}^n \mid z^n + a_1 z^{n-1} + \dots + a_n \in P\} \quad (1)$$

be the set of coefficients of polynomials in P . The set H_P is a hyperplane, by which we mean an $n - 1$ dimensional affine subspace of \mathbb{R}^n . Let

$$C_r^n = \{(a_1, a_2, \dots, a_n) \in \mathbb{R}^n \mid p(z) = z^n + a_1 z^{n-1} \dots + a_n \text{ and } \rho(p) < r\}$$

be the set of coefficients of monic polynomials with root radius smaller than r . Clearly, $\rho^* < r$ if and only if $H_P \cap C_r^n \neq \emptyset$. The root optimization problem is then equivalent to finding the infimum of r such that the hyperplane H_P intersects the set C_r^n . The latter set is known to be nonconvex, characterized by several algebraic inequalities, so this would appear to be difficult.

However, since C_r^n is open and connected, it intersects a given hyperplane if and only if its convex hull intersects the hyperplane:

Lemma 3: (Chen [Che79b, Lemma 2.1.2]; see also [Che79a, Lemma 2.1]) Let H be a hyperplane in \mathbb{R}^n , that is an $n - 1$ dimensional affine subspace of \mathbb{R}^n , and let $S \subset \mathbb{R}^n$ be an open connected set. Then $H \cap S \neq \emptyset$ if and only if $H \cap \text{conv}(S) \neq \emptyset$.

The set $\text{conv}(C_r^n)$ is an open simplex so it is easy to characterize its intersection with H_P :

Theorem 4: (Chen, special case of [Che79b, Prop. 3.1.7] and also Fam and Meditch [FM78], for the case $r = 1$; see also [HP05, Prop. 4.1.26].) We have

$$\text{conv}(C_r^n) = \text{conv}(\nu_1, \nu_2, \dots, \nu_{n+1})$$

where the vertices

$$\nu_k = \{(a_1, a_2, \dots, a_n) \in \mathbb{R}^n \mid (z - r)^{n-k}(z + r)^k = z^n + \sum_{j=1}^n a_j z^j\}$$

are the coefficients of the polynomials $(z - r)^{n-k}(z + r)^k$.

Since the optimum ρ^* is attained, the closure of $\text{conv}(C_{\rho^*}^n)$ and the hyperplane H_P must have a non-empty intersection. Theorem 1 says that, in fact, the intersection of H_P with $C_{\rho^*}^n$ must contain at least one vertex of $\text{conv}(C_{\rho^*}^n)$, and Corollary 2 explains how to find it. In contrast, Chen uses Theorem 4 to derive a procedure (his Theorem 3.2.2) for testing whether the minimal value ρ^* of Theorem 1 is greater or less than a given value r (see also [Che79a, Theorem 2.6]). This could be used to define a bisection method for approximating ρ^* , but it would not yield the optimal polynomial $p^*(z)$. Note that the main tool used in the proof of Theorem 1 is the implicit function theorem, in contrast to the sequence of algebraic results leading to Theorem 4.

Remark 5: The techniques used in Theorem 1 are all local. Thus, any locally optimal minimizer can be perturbed to yield a locally optimal minimizer of the form $(z - \beta)^{n-k}(z + \beta)^k \in P$ for some integer k , where β is the root radius attained at the local minimizer. Furthermore, all real roots $-\beta$ of the polynomials g_k in Corollary 2 define candidates for local minimizers, and while not all of them are guaranteed to be local minimizers, those with smallest magnitude (usually there will only be one) are guaranteed to be global minimizers. See Example 8 in Section IV.

The work of Chen [Che79b] was limited to polynomials with real coefficients. A complex analogue of Theorem 1 is simpler to state because the optimal polynomial may be chosen to

have only one distinct root, a multiple root if $n > 1$. However, the proof is substantially more complicated than for the real case and is deferred to Appendix A.

Theorem 6: Let B_0, B_1, \dots, B_n be complex scalars (with B_1, \dots, B_n not all zero) and consider the affine family of polynomials

$$P = \{z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n \mid B_0 + \sum_{j=1}^n B_j a_j = 0, a_i \in \mathbb{C}\}.$$

The optimization problem

$$\rho^* := \inf_{p \in P} \rho(p)$$

has an optimal solution of the form

$$p^*(z) = (z - \gamma)^n \in P$$

with $-\gamma$ given by a root of smallest magnitude of the polynomial

$$h(z) = B_n z^n + B_{n-1} \binom{n}{n-1} z^{n-1} + \dots + B_1 \binom{n}{1} z + B_0.$$

III. CONTINUOUS-TIME STABILITY

Let $\alpha(p)$ denote the *root abscissa* of a polynomial p ,

$$\alpha(p) = \max \{\operatorname{Re}(z) \mid p(z) = 0, z \in \mathbb{C}\}.$$

We now consider minimization of the root abscissa of a monic polynomial with real coefficients subject to a single affine constraint. In this case, the infimum may not be attained.

Theorem 7: Let B_0, B_1, \dots, B_n be real scalars (with B_1, \dots, B_n not all zero) and consider the affine family of polynomials

$$P = \{z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n \mid B_0 + \sum_{j=1}^n B_j a_j = 0, a_i \in \mathbb{R}\}.$$

Let $k = \max\{j : B_j \neq 0\}$. Define the polynomial of degree k

$$h(z) = B_n z^n + B_{n-1} \binom{n}{n-1} z^{n-1} + \dots + B_1 \binom{n}{1} z + B_0.$$

Consider the optimization problem

$$\alpha^* := \inf_{p \in P} \alpha(p).$$

Then

$$\alpha^* = \min \{\beta \in \mathbb{R} \mid h^{(i)}(-\beta) = 0 \text{ for some } i \in \{0, \dots, k-1\}\},$$

where $h^{(i)}$ is the i -th derivative of h . Furthermore, the optimal value is attained by a minimizing polynomial p^* if and only if $-\alpha^*$ is a root of h , that is $i = 0$, and in this case we can take

$$p^*(z) = (z - \gamma)^n \in P$$

with $\gamma = \alpha^*$.

The first part of this result, the characterization of the infimal value, is due to Chen [Che79b, Theorem 2.3.1]. Furthermore, Chen also observed the “if” part of the second statement, showing [Che79b, p.29] that if $-\alpha^*$ is a root of h (as opposed to one of its derivatives), the optimal value α^* is attained by the polynomial with a single distinct root α^* . However, he noted on the same page that he did not have a general method to construct a polynomial with an abscissa equal to a given value $\tilde{\alpha} > \alpha^*$. Nor did he characterize the case when the infimum is attained. We now address both these issues.

Because the infimum may not be attained, we cannot prove Theorem 7 using a variant of the proof of Theorem 1. Instead, we follow Chen’s development. Define the hyperplane of feasible coefficients as previously (see equation (1)). Let

$$S_\zeta^n := \{(a_1, a_2, \dots, a_n) \in \mathbb{R}^n \mid z^n + a_1 z^{n-1} + \dots + a_n = 0 \text{ implies } \operatorname{Re}(z) < \zeta\}$$

denote the set of coefficients of monic polynomials with root abscissa less than ζ , where $\zeta \in \mathbb{R}$ is a given parameter.

Definition 8: (S_ζ^n -stabilizability) A hyperplane $H_P \subset \mathbb{R}^n$ is said to be S_ζ^n -stabilizable if $H_P \cap S_\zeta^n \neq \emptyset$.

As in the root radius case, Lemma 3 shows that although S_ζ^n is a complicated nonconvex set, a hyperplane H_P is S_ζ^n -stabilizable if and only if H_P intersects $\operatorname{conv} S_\zeta^n$, a polyhedral convex cone which can be characterized as follows:

Theorem 9: (Chen [Che79b, Theorem 2.1.8]) We have

$$\operatorname{conv}(S_\zeta^n) = \nu + \operatorname{pos}(\{\tilde{e}_i\}) = \left\{ \nu + \sum_{i=1}^n r_i \tilde{e}_i \mid r_i \geq 0 \right\},$$

an open polyhedral convex cone with vertex

$$\nu = \sum_{j=1}^n \binom{n}{j} (-\zeta)^j e_j$$

and extreme rays

$$\tilde{e}_i = \sum_{j=i}^n \binom{n-i}{j-i} (-\zeta)^{j-i} e_j,$$

where $\{e_j\}_{j=1}^n$ is the standard basis of \mathbb{R}^n .

This leads to the following characterization of S_ζ^n -stabilizability:

Theorem 10: (Chen, a variant of [Che79b, Theorem 2.2.2]; see also [Che79a, Theorem 2.4])

Define the hyperplane H_P as in equation (1), the polynomial h and the integer k as in Theorem

7. Then the following statements are equivalent:

- 1) H_P is S_ζ^n -stabilizable
- 2) There exist nonnegative integers j, \tilde{j} with $0 \leq j < \tilde{j} \leq k$ such that

$$h^{(j)}(-\zeta)h^{(\tilde{j})}(-\zeta) < 0$$

where $h^{(j)}(-\zeta)$ denotes the j -th derivative of $h(z)$ at $z = -\zeta$.

To prove the last part of Theorem 7, we need the following lemma.

Lemma 11: We have $h(-\zeta) = 0$ if and only if $(z - \zeta)^n \in P$. Furthermore, for $i \in \{1, 2, \dots, k-1\}$, $h^{(i)}(-\zeta) = 0$ if and only if exactly one of the following two conditions hold:

- 1) $L_i \cap H_P = \emptyset$ and $h(-\zeta) \neq 0$
- 2) $L_i \in H_P$ and $h(-\zeta) = 0$

where

$$L_i = \{\nu + r_i \tilde{e}_i \mid r_i \geq 0\}, \quad i = 1, 2, \dots, n.$$

is the i -th extreme ray of the cone $\text{conv}(S_\zeta^n)$ given in Theorem 9.

Proof: We have

$$h(-\zeta) = \sum_{j=0}^n B_j \binom{n}{j} (-\zeta)^j = B_0 + (B_1, B_2, \dots, B_n) \cdot \nu$$

where \cdot denotes the usual dot product in \mathbb{R}^n . Therefore,

$$\begin{aligned} h(-\zeta) = 0 &\iff B_0 + (B_1, B_2, \dots, B_n) \cdot \nu = 0 \\ &\iff \nu \in H_P \\ &\iff z^n + \sum_{i=1}^n \nu_i z^{n-i} = (z - \zeta)^n \in P \end{aligned} \tag{2}$$

proves the first part of the lemma. Now, let $i \in \{1, 2, \dots, k-1\}$. A straightforward calculation gives

$$\begin{aligned} h^{(i)}(-\zeta) &= \frac{n!}{(n-i)!} \sum_{j=i}^n B_j \binom{n-i}{j-i} (-\zeta)^{j-i} \\ &= \frac{n!}{(n-i)!} (B_1, B_2, \dots, B_n) \cdot \tilde{e}_i \end{aligned}$$

Hence,

$$\begin{aligned} h^{(i)}(-\zeta) = 0 &\iff (B_1, B_2, \dots, B_n) \cdot \tilde{e}_i = 0 \\ &\iff L_i \in H := \{(a_1, a_2, \dots, a_n) \mid - (B_1, B_2, \dots, B_n) \cdot \nu \\ &\quad + \sum_{j=1}^n B_j a_j = 0\} \end{aligned}$$

If $B_0 = -(B_1, B_2, \dots, B_n) \cdot \nu$, then $H = H_P$, $\nu \in H_P$ and from (2), we get $h(-\zeta) = 0$ (case (1)). Otherwise, the hyperplane H is parallel to H_P and $H \cap H_P = \emptyset$, so that $L_i \cap H_P = \emptyset$, and also $h(-\zeta) \neq 0$ (otherwise by (2), $\nu \in L_i \cap H_P$ which would be a contradiction); this is case (2). ■

Now we are ready to complete the proof of Theorem 7.

Proof: Chen's theorem [Che79b, Theorem 2.3.1] establishes the characterization of the optimal value,

$$\inf_{p \in P} \alpha(p) = \alpha^* = \min\{\beta \mid \prod_{i=0}^{k-1} h^{(i)}(-\beta) = 0\}.$$

Let $l \in \{0, 1, \dots, k-1\}$ be the smallest integer such that $h^{(l)}(-\alpha^*) = 0$. If $l = 0$, then $-\alpha^*$ is a root of h and by Lemma 11, $p^*(z) = (z - \gamma)^n \in P$ is an optimizer with $\gamma = \alpha^*$.

Suppose now that $l > 0$. We will show that the infimum is not attained. Suppose the contrary, that is $H_P \cap \text{cl}(S_{\alpha^*}^n) \neq \emptyset$ so that $H_P \cap \text{cl}(\text{conv} S_{\alpha^*}^n) \neq \emptyset$. Without loss of generality, assume $B_k > 0$ so that $h^{(k)}$ is the constant function $k!B_k > 0$ and the derivatives $h^{(j)}$, $j = 1, 2, \dots, k-1$ each have leading coefficient (coefficient of z^{k-j}) also having positive sign. By Theorem 10, $h^{(j)}(-\tilde{\alpha}) > 0$ for any $j = 1, 2, \dots, k$ and $\tilde{\alpha} < \alpha^*$ and, in addition, $h^{(j)}(-\alpha^*) > 0$ for $0 \leq j < l$.

By continuity of $h^{(j)}$, we have

$$h^{(j)}(-\alpha^*) \begin{cases} > 0 & \text{if } 0 \leq j < l \\ = 0 & \text{if } j = l \\ \geq 0 & \text{if } l < j < k \\ > 0 & \text{if } j = k \end{cases}$$

It thus follows from Theorem 10 that H_P is not $S_{\alpha^*}^n$ -stabilizable, which means $H_P \cap S_{\alpha^*}^n = \emptyset$, or equivalently, by Lemma 3, that $H_P \cap \text{conv} S_{\alpha^*}^n = \emptyset$. Since $\text{conv} S_{\alpha^*}^n$ is an open set, it follows from the assumption made above that its boundary intersects H_P . Pick a point $y \in H_P \cap \text{bd}(\text{conv} S_{\alpha^*}^n)$. It is easy to show that H_P is a supporting hyperplane to the convex cone $\text{conv} S_{\alpha^*}^n$ at the boundary point y . Since every hyperplane supporting a convex cone must pass through the vertex of the cone [HUL93, A.4.2], it follows that $\nu \in H_P$. On the other hand, since $l > 0$, Lemma 11 implies $L_l \cap H_P = \emptyset$. This is a contradiction. ■

Remark 12: If $-\beta$ is a real root of $h(z)$, then $(z - \beta)^n \in P$. Such a polynomial is often, though not always, a local minimizer of $\alpha(p)$, but it is a global minimizer if and only if $-\beta$ is the largest such real root *and* no other roots of derivatives of h are larger than $-\beta$.

We now address the case where the infimum is not attained.

Theorem 13: Assume that $-\alpha^*$ is not a root of h . Let ℓ be the smallest integer $i \in \{1, \dots, k-1\}$ for which $-\alpha^*$ is a root of $h^{(i)}$. Then, for all sufficiently small $\epsilon > 0$ there exists a real scalar M_ϵ for which

$$p_\epsilon(z) := (z - M_\epsilon)^m (z - (\alpha^* + \epsilon))^{n-m} \in P$$

where $m = \ell$ or $\ell + 1$, and $M_\epsilon \rightarrow -\infty$ as $\epsilon \rightarrow 0$.

Proof: By Theorem 7, the optimal abscissa value α^* is not attained. Without loss of generality, assume $\alpha^* = 0$. Otherwise, write $z = \tilde{z} + \alpha^*$ and rewrite P as the set of monic polynomials in \tilde{z} with an affine constraint.

For $0 < m \leq n$, we have $p_\epsilon(z) = (z + K)^m (z - \epsilon)^{n-m} \in P$ if and only if its coefficients are real and

$$\begin{aligned}
0 &= \left(B_0 + B_1 \binom{n-m}{1} (-\epsilon) + B_2 \binom{n-m}{2} (-\epsilon)^2 + \cdots + B_{n-m} (-\epsilon)^{n-m} \right) \\
&+ \binom{m}{1} \left(B_1 + B_2 \binom{n-m}{1} (-\epsilon) + B_3 \binom{n-m}{2} (-\epsilon)^2 + \cdots + B_{n-m+1} (-\epsilon)^{n-m} \right) K \\
&+ \binom{m}{2} \left(B_2 + B_3 \binom{n-m}{1} (-\epsilon) + \cdots + B_{n-m+2} (-\epsilon)^{n-m} \right) K^2 \\
&+ \cdots + \left(B_m + B_{m+1} \binom{n-m}{1} (-\epsilon) + \cdots + B_n (-\epsilon)^{n-m} \right) K^m \\
&= \eta_0(\epsilon) + \eta_1(\epsilon)K + \cdots + \eta_m(\epsilon)K^m =: f_\epsilon(K).
\end{aligned}$$

Thus, $p_\epsilon \in P$ if and only if K is a real root of f_ϵ , a polynomial of degree m whose coefficients depend on ϵ . By Theorem 10, the $h^{(j)}(\epsilon)$ have the same sign for all $\epsilon > 0$ and for all $j \in \{0, 1, \dots, k\}$, which we take to be positive. By the definition of ℓ , $h^{(j)}(0) > 0$ for $j < \ell$ and $h^{(\ell)}(0) = 0$ which gives $\eta_j(0) = \binom{m}{j} \frac{h^{(j)}(0)}{n!(n-j)!} > 0$ for $j < \ell$ and similarly $\eta_\ell(0) = 0$. We have also

$$\eta_m(\epsilon) = \sum_{j=m}^n B_j \binom{n-m}{j-m} (-\epsilon)^{j-m} = \frac{(n-m)!}{n!} h^{(m)}(-\epsilon) \quad (3)$$

and

$$\eta_{m-1}(\epsilon) = m \sum_{j=m}^n B_{j-1} \binom{n-m}{j-m} (-\epsilon)^{j-m} \quad (4)$$

$$= m \frac{(n-m)!}{n!} ((n-m+1)h^{(m-1)}(-\epsilon) + \epsilon h^{(m)}(-\epsilon)). \quad (5)$$

Let $m = \ell$. We have $\eta_\ell(\epsilon) > 0$ for $\epsilon < 0$ and $\eta_\ell(0) = 0$. The polynomial η_ℓ might change sign around 0, depending on the multiplicity of 0 as a root. If 0 is a root of η_ℓ with an odd multiplicity, $\eta_\ell(\epsilon) < 0$ for $\epsilon > 0$ small enough and so the coefficients of f_ϵ have one and only one sign change. By Descartes' rule of signs, f_ϵ has one and only one root K with positive real part which must therefore be real. Setting $M_\epsilon = -K$, we have $p_\epsilon(z) = (z - \epsilon)^{n-m} (z + K)^m \in P$ as desired. If the multiplicity is even, then the multiplicity of 0 as a root of $h^{(\ell)}$ is also even by (3). Then, $h^{(\ell+1)}$ must have 0 as a root with odd multiplicity and $h^{(\ell+1)}$ changes sign around 0. Set $m = \ell + 1$ in this case and repeat a similar argument: By (3), η_m changes sign around 0, i.e. $\eta_m < 0$ for $\epsilon > 0$ small enough. Furthermore, from (5), $\eta_{m-1} > 0$ for $\epsilon > 0$, ϵ small enough. As

a result, the coefficients of f_ϵ have one and only one sign change, for $\epsilon > 0$, ϵ small enough. We again get the existence of p_ϵ in P with the desired structure.

Finally, let us show that $M_\epsilon \rightarrow -\infty$. Suppose this is not the case. Then, there exists a sequence $\epsilon_\kappa \downarrow 0$ and a positive number R such that $\sup_\kappa \rho(p_{\epsilon_\kappa}) \leq R$. Since $\text{cl}(C_R^n)$ is compact by Theorem 4, there exists a positive constant \tilde{R} such that all of the coefficients of the polynomial p_{ϵ_κ} are bounded by \tilde{R} , uniformly over κ . By compactness, there exists a subsequence $p_{\epsilon_{\kappa_\iota}}$ converging to a limit p_* pointwise. Furthermore, $p_* \in P$ since P is closed. By continuity of the abscissa mapping, $\alpha(p_*) = \lim_{\iota \rightarrow \infty} \alpha(p_{\epsilon_{\kappa_\iota}}) = 0$. This implies that the optimal abscissa is attained on P , which is a contradiction. ■

Theorem 7 showed that in the real case the infimal value is not attained if and only if the polynomial h has a derivative of any order between 1 and $k - 1$ with a real root to the right of the rightmost real root of h . However, it is not possible that a derivative of h has a complex root to the right of the rightmost complex root of h . This follows immediately from the Gauss-Lucas theorem, which states that the roots of the derivative of a polynomial p must lie in the convex hull of the roots of p [BLO04], [Mar66]. This suggests that the infimal value of the optimal abscissa problem with complex coefficients is always attained at a polynomial with a single distinct root, namely a rightmost root of h . Indeed, this is established in the following theorem, whose proof can be found in Appendix B.

Theorem 14: Let B_0, B_1, \dots, B_n be complex scalars (with B_1, \dots, B_n not all zero) and consider the affine family of polynomials

$$P = \{z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n \mid B_0 + \sum_{j=1}^n B_j a_j = 0, a_i \in \mathbb{C}\}.$$

The optimization problem

$$\alpha^* := \inf_{p \in P} \alpha(p)$$

has an optimal solution of the form

$$p^*(z) = (z - \gamma)^n \in P$$

with $-\gamma$ given by a root with largest real part of the polynomial h where

$$h(z) = B_n z^n + B_{n-1} \binom{n}{n-1} z^{n-1} + \dots + B_1 \binom{n}{1} z + B_0.$$

IV. EXAMPLES

Example 1. The following simple example is from [BLO01], where it was proved using the Gauss-Lucas theorem that $p_*(z) = z^n$ is a global optimizer of the abscissa over the set of polynomials

$$P = \{z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n \mid a_1 + a_2 = 0, a_i \in \mathbb{C}\}.$$

We calculate $h(z) = \binom{n}{2} z(z + \frac{2}{n-1})$. Theorem 7 proves global optimality over $a_i \in \mathbb{R}$ and Theorem 14 proves global optimality over $a_i \in \mathbb{C}$.

Example 2. As mentioned in Section 1, Henrion and Overton [HO06] showed that the problem of finding a second-order linear controller that maximizes the closed-loop asymptotic decay rate for the classical two-mass-spring system is equivalent to an abscissa minimization problem for a monic polynomial of degree 6 whose coefficients depend affinely on 5 parameters, or equivalently with a single affine constraint on the coefficients. Theorem 7 (as well as Theorem 14) establishes global optimality of the locally optimal polynomial constructed in [HO06].

Example 3. This is derived from a ‘‘Belgian chocolate’’ stabilization challenge problem of Blondel [Blo94]: given $a(z) = z^2 - 2\delta z + 1$ and $b(z) = z^2 - 1$, find the range of real values of δ for which there exist polynomials x and y such that $\deg(x) \geq \deg(y)$ and $\alpha(xy(ax + by)) < 0$. This problem remains unsolved. However, inspired by numerical experiments, [BHLO06] gave a solution for $\delta < \bar{\delta} = (1/2)\sqrt{2 + \sqrt{2}} \approx 0.924$. When x is constrained to be a monic polynomial with degree 3 and y to be a constant, the minimization of $\alpha(xy(ax + by))$ reduces to

$$\inf_{p \in P} \alpha(p)$$

where

$$P = \{(z^2 - 2\delta z + 1)(z^3 + \sum_{k=0}^2 w_k z^k) + (z^2 - 1)v \mid w_0, w_1, w_2, v \in \mathbb{C}\}.$$

For nonzero fixed δ , P is a set of monic polynomials with degree 5 whose coefficients depend affinely on 4 parameters, or equivalently with a single affine constraint on the coefficients. In [BHLO06] a polynomial in P with one distinct root of multiplicity 5 was constructed and proved to be locally optimal using nonsmooth analysis. Theorems 7 and 14 prove its global optimality. They also apply to the case when x is constrained to be monic with degree 4; then, as shown in [BHLO06], stabilization is possible for $\delta < \tilde{\delta} = (1/4)\sqrt{10 + 2\sqrt{5}} \approx 0.951$.

Example 4. The polynomial achieving the minimal root radius may not be unique. Let $P = \{z^2 + a_1z + a_2 \mid 1 + a_1 + a_2 = 0, a_i \in \mathbb{R}\}$. We have

$$\rho^* := \inf_{p \in P} \rho(p) = \inf_{a_2 \in \mathbb{R}} \rho(z^2 - (a_2 + 1)z + a_2) = \inf_{a_2 \in \mathbb{R}} \rho((z - a_2)(z - 1)) = 1.$$

The minimal value is attained on a continuum of polynomials of the form $(z - a_2)(z - 1)$ for any $-1 \leq a_2 \leq 1$ and hence minimizers are not unique. The existence of the minimizers $(z - 1)^2$ and $(z + 1)(z - 1)$ is consistent with Theorem 1. The same example shows that the minimizer for the radius optimization problem with complex coefficients may not be unique.

Example 5. Likewise, a polynomial achieving the minimal root abscissa may not be unique. Let $P = \{z^2 + a_1z + a_2 \mid a_1 = 0, a_2 \in \mathbb{R}\}$. We have

$$\alpha^* = \inf_{p \in P} \alpha(p) = \inf_{a_2 \in \mathbb{R}} \alpha(z^2 + a_2) = 0.$$

Here $B_0 = B_2 = 0$, $B_1 = 1$. The optimum is attained at $p^*(z) = z^2$, where $-\alpha^* = 0$ is a root of the polynomial $h(z) = z$, as claimed in Theorem 7. However, the optimum is attained at a continuum of polynomials of the form $z^2 + a_2$ for any $a_2 > 0$.

Example 6. In this example, the infimal root abscissa is not attained. Let $P = \{z^2 + a_1z + a_2 \mid a_1 \in \mathbb{R} \text{ and } a_2 = -1\}$. We have $h(z) = z^2 + 1$, so $-\alpha^* = 0$ is a root of $h^{(1)}$ but not of h . Thus, Theorem 13 applies with $\ell = 1$. Indeed

$$\begin{aligned} \alpha^* = \inf_{p \in P} \alpha(p) &= \inf_{a_1 \in \mathbb{R}} \alpha(z^2 + a_1z - 1) \\ &= \inf_{a_1 \in \mathbb{R}} \max \left\{ \frac{-a_1 - \sqrt{a_1^2 + 4}}{2}, \frac{-a_1 + \sqrt{a_1^2 + 4}}{2} \right\} = 0. \end{aligned}$$

This infimum is not attained, but as $a_1 \rightarrow \infty$, setting $\epsilon = \frac{-a_1 + \sqrt{a_1^2 + 4}}{2} \rightarrow 0$ and $M_\epsilon = \frac{-a_1 - \sqrt{a_1^2 + 4}}{2} \rightarrow -\infty$ gives $(z - M_\epsilon)(z - \epsilon) \in P$ as claimed in Theorem 13.

Example 7. Consider the family $P = \{z^3 + a_1z^2 + a_2z + a_3 \mid a_1, a_2 \in \mathbb{R} \text{ and } a_3 = -1\}$. We have $h(z) = z^3 + 1$, so $-\alpha^* = 0$ is a root of both $h^{(1)}$ and $h^{(2)}$. Thus, the assumptions of Theorem 13 are again satisfied with $\ell = 1$. However, this example shows the necessity of setting $m = \ell + 1$ when $h^{(\ell)}$ has a root of even multiplicity at $-\alpha^*$. Setting $m = \ell = 1$ is impossible since then $(z - M_\epsilon)^m(z - \epsilon)^{n-m} \in P$ implies $M_\epsilon = \frac{1}{\epsilon^2} \rightarrow +\infty$ as $\epsilon \rightarrow 0$. On the other hand, when $m = \ell + 1 = 2$, we have $(z - M_\epsilon)^m(z - \epsilon)^{n-m} \in P$ with $M_\epsilon = -\frac{1}{\sqrt{\epsilon}} \rightarrow -\infty$ as $\epsilon \downarrow 0$.

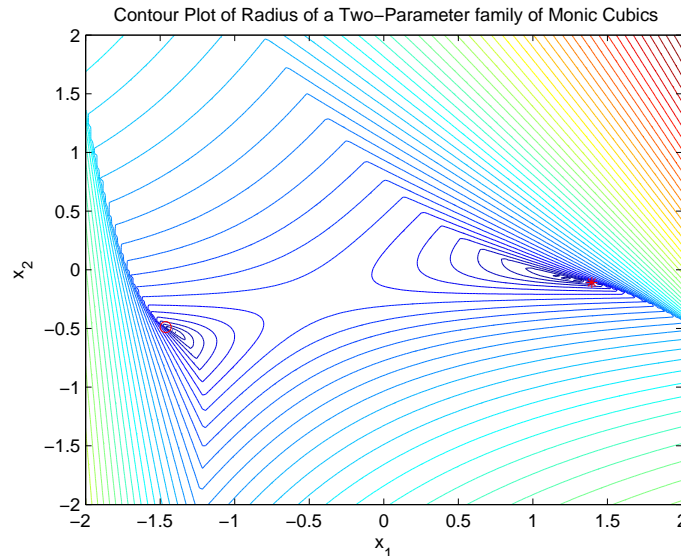


Fig. 1. Contour plot $\rho(p)$ for a randomly generated family of monic cubics with coefficients depending affinely on two real parameters. The asterisk shows the global minimizer and the circle shows a local minimizer.

Example 8. Figure 1 shows a contour plot of the root radius of a randomly generated family of monic cubics with coefficients depending affinely on two real parameters x_1 and x_2 (for details of the parametrization, see the web page referenced below). The asterisk indicates a global minimizer defining the cubic $(z - \gamma_1)^3$ and the circle indicates a local minimizer defining the cubic $(z - \gamma_2)^3$. Both $-\gamma_1 \approx -0.541$ and $-\gamma_2 \approx 0.567$ are roots of g_0 (see Corollary 2 and Remark 5). Note the steep contours near the global and local minimizers, indicating the non-Lipschitz behavior of the root radius ρ as the triple root splits [BLO05].

Example 9. Consider the monic quintic polynomials subject to the affine constraint $B_0 + \sum_{i=1}^5 a_i = 0$. Figure 2 plots the optimal abscissa, over $a_i \in \mathbb{R}$ and $a_i \in \mathbb{C}$ respectively, as a function of $B_0 \in [-2, 4]$. For $B_0 \leq 1$, the largest root of the polynomial h is real so the optimal abscissa is attained with the same optimal value for both the real and complex cases. For $B_0 = 1$, $h(z) = (1+z)^5$ so all of the roots of h and its derivatives are identical. Furthermore, when $B_0 = 1$, the hyperplane $H_P = \{(a_1, a_2, \dots, a_n) \mid B_0 + \sum_{i=1}^5 a_i = 0\}$ contains the extreme lines L_i of the cone $\text{conv}(S_1^n)$ for $i = 1, 2, 3, 4$ as well as the boundary of $\text{conv}(S_1^n)$ which is the positive span of the vectors $L_i, i = 1, 2, 3, 4$, so we have $\text{aff}(S_1^n) = H_P$. As B_0 is increased, the hyperplane H_P is translated away from $\text{conv}(S_1^n)$, and so for $B_0 > 1$, H_P does not intersect

$\text{conv}(S_1^n)$ but it is parallel to it. Thus, for the real case, the infimum is not attained for $B_0 > 1$ and has the constant value 1. In contrast, in the complex case, the optimal value decays as B_0 is increased.

More examples may be explored by downloading a publicly available¹ MATLAB code implementing the constructive algorithms implicit in Theorems 1, 6, 7 and 14 as well as Corollary 2 and Theorem 13. A code generating all the examples of this section (along with Figures 1 and 2) is also available at the same website. In general, there does not seem to be any difficulty obtaining an accurate globally optimal value for the root abscissa or root radius in the real or complex case. However, even in the cases where an optimal solution exists, the coefficients may be large, so that rounding errors in the computed coefficients result in a large constraint residual, and the difficulty is compounded when the optimal abscissa value is not attained and a polynomial with an approximately optimal abscissa value is computed: hence, it is inadvisable to choose ϵ in Theorem 13 too small. Furthermore, the multiple roots of the optimal polynomials are not robust with respect to small perturbations in the coefficients. Optimizing a more robust objective such as the so-called complex stability “radius” (in the data-perturbation sense) of the polynomial may be of more practical use; see [BHLO06, Section II].

V. CONCLUDING REMARKS

Suppose there are κ constraints on the coefficients. In this case, we conjecture, based on numerical experiments, that there always exists an optimal polynomial with at most $\kappa - 1$ roots having modulus less than ρ^* or having real part less than α^* respectively. However, there does not seem to be a useful bound on the number of possible distinct roots. Thus, computing global optimizers appears to be difficult.

When there are κ constraints, we can obtain upper and lower bounds on the optimal value as follows. Lower bounds can be obtained by solving many problems with only *one* constraint, each of which is obtained from random linear combinations of the prescribed κ constraints. Upper bounds can be obtained by local optimization of the relevant objective ρ or α over an affine parametrization which is obtained from computing the null space of the given constraints. However, the gap between these bounds cannot be expected to be small.

¹www.cs.nyu.edu/overton/software/affpoly/

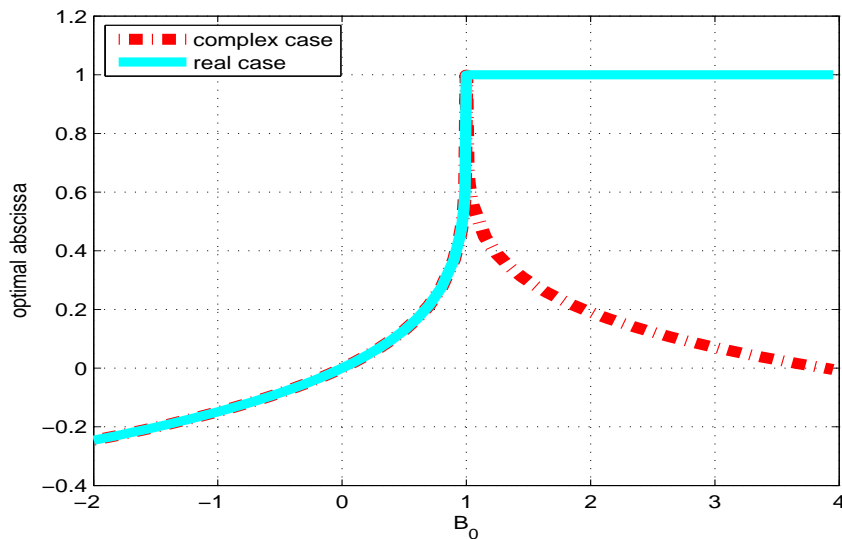


Fig. 2. The optimal root abscissa for a polynomial family $z^5 + a_1z^4 + \dots + a_5$ with constraint $B_0 + \sum_{i=1}^5 a_i = 0$, with B_0 varied from -2 to 4 . The solid curve shows the optimal abscissa for the case $a_i \in \mathbb{R}$ and the dotted curve for the case $a_i \in \mathbb{C}$.

The results do not extend to the more general case of an affine family of $n \times n$ matrices depending on $n - 1$ parameters. For example, consider the matrix family

$$A(\xi) = \begin{bmatrix} \xi & 1 \\ -1 & \xi \end{bmatrix}.$$

This matrix depends affinely on a single parameter ξ , but its characteristic polynomial, a monic polynomial of degree 2, does not, so the results given here do not apply. The minimal spectral radius (maximum of the moduli of the eigenvalues) of $A(\xi)$ is attained by $\xi = 0$, for which the eigenvalues are $\pm j$. Nonetheless, experiments show that it is often the case that optimizing the spectral radius or spectral abscissa of a matrix depending affinely on parameters yields a matrix with multiple eigenvalues, or several multiple eigenvalues with the same radius or abscissa value; an interesting example is analyzed in [GO07].

APPENDIX A
PROOF OF THEOREM 6

We begin with some notation. For a positive integer n , let \mathcal{P}_n denote the complex vector space of all polynomials

$$p(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n$$

with complex coefficients $a_i \in \mathbb{C}$. Let \mathcal{P}_n^1 be the affine subset of \mathcal{P}_n consisting of all polynomials with $a_0 = 1$ (the monic polynomials).

Definition 15: For $p \in \mathcal{P}_n^1$ and $k \in \{1, 2, \dots, n\}$, let $r_k(p)$ denote the k -th largest absolute value of a root of p , i.e. $r_k(p) = |z_k|$ when

$$p(z) = (z - z_1)(z - z_2) \cdots (z - z_n), \quad |z_1| \geq |z_2| \geq \cdots \geq |z_n|.$$

For $p \in \mathcal{P}_n^1$ define

$$\phi(p) = \max\{|z_1 - z_2| \mid |z_1| = |z_2| = r_1(p), p(z_1) = p(z_2) = 0\},$$

the diameter of the set of roots with maximal modulus (zero if p has only one distinct root with maximal modulus). Given $g \in \mathcal{P}_n^1$ and a linear functional $L : \mathcal{P}_n \mapsto \mathbb{C}$, let X_0 be the set of all $p \in \mathcal{P}_n^1$ such that $L(p) = L(g)$. For $k = 1, 2, \dots, n$ define X_k as the set of all $p \in X_{k-1}$ for which $r_k(p)$ equals the minimum of r_k on X_{k-1} .

We will need the following preliminary result.

Lemma 16: For $g \in \mathcal{P}_2^1$, let X_0, X_1, X_2 be defined as above. Then one of the following statements is true:

- (a) the functional $r_1(\cdot)$ has a unique minimizer p_* on X_0 , and there exists $a \in \mathbb{C}$ such that $p_*(z) = (z - a)^2$, or equivalently $X_0 \supseteq X_1 = X_2 = \{(z - a)^2\}$.
- (b) the functional $r_2(\cdot)$ has a unique minimizer p_* on X_1 , and there exists $a \in \mathbb{C}$ such that $p_*(z) = (z - a)z$, or equivalently $X_0 \supseteq X_1 \supset X_2 = \{(z - a)z\}$.
- (c) there exist $a \in \mathbb{C}$, $a \neq 0$, $\theta \in (0, \pi)$, and a continuous strictly decreasing function $\psi : [0, 2\pi] \mapsto [-\pi, \pi]$ satisfying interpolation constraints

$$\psi(0) = \pi, \quad \psi(\theta) = \theta, \quad \psi(\pi) = 0, \quad \psi(2\pi - \theta) = -\theta, \quad \psi(2\pi) = -\pi,$$

such that

$$X_0 \supseteq X_1 = X_2 = \{p(z) = (z - ae^{jt})(z - ae^{j\psi(t)}) \mid t \in [0, 2\pi]\}.$$

Proof: Let the complex numbers q_0, q_1, q_2 be defined by

$$L(z^2 + a_1z + a_2) - L(g) = q_2 + q_1a_1 + q_0a_2.$$

Then a polynomial $p(z) = (z - z_1)(z - z_2)$ belongs to X_0 if and only if

$$q_2 - q_1(z_1 + z_2) + q_0z_1z_2 = 0. \quad (6)$$

When $q_0 \neq 0$ define b, c such that

$$c = \frac{q_1}{q_0}, \quad b^2 = c^2 - \frac{q_2}{q_0}, \quad \operatorname{Re}(\bar{c}b) \geq 0,$$

i.e. (6) is equivalent to

$$(z_1 - c)(z_2 - c) = b^2,$$

and $|c + b| \geq |c - b|$. One of the following situations must occur.

- 1) $q_0 = q_1 = 0$: the fact that (6) must be feasible yields $q_2 = 0$, hence $X_0 = \mathcal{P}_2^1$, and the minimal value 0 of $r_1(p)$ for $p \in X_0$ is attained at a single point $p(z) = z^2$ (case (a)).
- 2) $q_0 = 0, q_1 \neq 0$: condition (6) is equivalent to

$$\frac{z_1 + z_2}{2} = \frac{q_2}{2q_1} \stackrel{\text{def}}{=} a.$$

Since the inequality

$$\left| \frac{z_1 + z_2}{2} \right| \geq \max\{|z_1|, |z_2|\}$$

implies $z_1 = z_2$, the minimal value $|a|$ of $r_1(p)$ for $p \in X_0$ is attained at a single point $p(z) = (z - a)^2$ (case (a)).

- 3) $q_0 \neq 0, b = 0$: condition (6) holds when either $z_1 = c$ or $z_2 = c$, hence the minimal value $|c|$ of $r_1(p)$ for $p \in X_0$ is attained on polynomials of the form $p(z) = (z - c)(z - u)$, where $|u| \leq |c|$, and the minimal value 0 of $r_2(p)$ for $p \in X_1$ is attained at a single point $p(z) = (z - c)z$ (case (b)).
- 4) $q_0 \neq 0, b \neq 0, |c + b| > |c - b|$: the minimal value $|c - b|$ of $r_1(p)$ for $p \in X_0$ is attained at a single point $p(z) = (z - c + b)^2$ (case (a)). The statement is obvious when $b = c$. To see that no other polynomial $p \in X_0$ achieves the value of $|b - c|$ (or better) when $b \neq c$, it suffices to show that the disc $D_0 = \{z \mid |z| \leq |c - b|\}$ and its image $f(D_0)$ under the map $f : z \mapsto c + b^2/(z - c)$ have a unique common point $z = c - b$. It is well known

that f is a bijection of the extended complex plane $C \cup \{\infty\}$ to itself which maps discs to discs, complements of discs, or half-planes, and also maps boundaries to boundaries. Since $z_0 = c - b = f(c - b)$ is a fixed point of f , and $f'(z_0) = -1$ is negative real, $f(D_0)$ is tangential to D_0 at z_0 . By the negativity of $f'(z_0)$, z_0 will be the only intersection of D_0 and $f(D_0)$.

- 5) $q_0 \neq 0$, $b \neq 0$, $|c + b| = |c - b|$: the minimal value $|c - b| = |c + b|$ of $r_1(p)$ for $p \in X_0$ is attained on the set of polynomials of the form $p(z) = (z - z_1)(z - z_2)$, where z_1 is an arbitrary complex number such that $|z_1| = |c - b|$, and

$$z_2 = \frac{b^2}{z_1 - c} + c$$

(which automatically implies $|z_2| = |c - b|$). This is case (c), where $a = jb$ and $\theta = \pi/2$ when $c = 0$, and otherwise a, θ are defined by

$$a = \frac{|c - b|}{|c|}c, \quad e^{j\theta} = \frac{c \pm b}{|c \pm b|} \frac{|c|}{c}$$

(the plus sign is to be used when the imaginary part of $b\bar{c}$ is positive), and ψ is a “phase” representation of the map $z_1 \mapsto z_2$.

This completes the proof. ■

In order to establish Theorem 6, we first state and prove a related “super-optimization” problem, again using Definition 15.

Theorem 17: The functional ϕ achieves its minimum on X_n , and for every minimizer $p_* \in X_n$ of ϕ , there exists $a \in \mathbb{C}$ and nonnegative integers κ, λ such that $p_*(z) = (z - a)^\kappa z^\lambda$.

Proof: We use induction with respect to n . When $n = 2$, the statement follows by Lemma 16. Assume the statement is true for $n = m > 1$. Consider the case $n = m + 1$. Note that ϕ is continuous on X_n , and hence achieves its minimum at a polynomial

$$p_*(z) = (z - z_1) \dots (z - z_m)(z - z_{m+1}), \quad |z_1| \geq |z_2| \geq \dots \geq |z_m| \geq |z_{m+1}|.$$

One of the following three situations must take place.

- 1) $|z_1| = 0$. Then $p_*(z) = z^n$.
- 2) $|z_{m+1}| < |z_1|$. Then, according to Lemma 16, $p_* \in X_n$ implies $z_{m+1} = 0$. Moreover, the polynomial

$$q(z) = (z - z_1) \dots (z - z_m)$$

must be optimal in the sense of Theorem 17 with $n = m$, and hence, by the inductive hypothesis, $q(z) = (z - a)^\kappa z^\lambda$, which implies $p_*(z) = (z - a)^\kappa z^{\lambda+1}$.

- 3) $|z_{m+1}| = |z_1| = d > 0$. According to the inductive hypothesis, the set X_n must contain a polynomial of the form $p(z) = (z - a)^m(z - z_1)$, where $|a| = d$. Let Ω be the shortest arc of the circle $|z| = d$ connecting the points a and z_1 (if $a = -z_1$, take one of the two arcs of equal length). Among all polynomials $p \in X_n$ with roots in Ω , take the one with the minimal radius of the root set, and denote it by

$$q_* : q_*(z) = (z - w_1)(z - w_2) \dots (z - w_n).$$

Since $q_* \in X_n$, we have $|w_i| = |a|$ for all i .

Let us show that all roots w_i of q_* are equal, i.e. that $q_*(z) = (z - b)^n$ for some $b \in \mathbb{C}$. By construction, $w_i \in \Omega$ lie within an arc of angular length not larger than π . Let u_1 and u_2 be the two most distant values among w_i . Applying Lemma 16 to the polynomial $p(z) = (z - u_1)(z - u_2)$, shows that the case (c) takes place (otherwise $q_* \notin X_n$), hence for every pair $w_i = u_1, w_k = u_2$ of the roots of q_* , it is possible to replace w_i and w_k with a pair of equal roots $\tilde{w}_i = \tilde{w}_k = \tilde{w} \in \Omega$, where \tilde{w} lies strictly between u_1 and u_2 . If u_1 or u_2 , as the roots of q_* , have multiplicity 1, this immediately leads to a polynomial $\tilde{q}_* \in X_n$ with the root set contained in Ω and having a smaller diameter. If the multiplicities of u_1, u_2 are greater than 1, this process can be repeated until lack of optimality of q_* is proved.

This completes the proof. ■

Corollary 18: Let $L : \mathcal{P}_n \mapsto \mathbb{C}$ be a linear functional. Then for every $g \in \mathcal{P}_n^1$ there exists a polynomial $f(z) = (z - a)^n$ (where $a \in \mathbb{C}$) such that $L(f) = L(g)$ and

$$r_1(f) = \min\{r_1(p) \mid p \in \mathcal{P}_n^1, L(p) = L(g)\}.$$

Proof: Taking into account the statement of Theorem 17, it is sufficient to show that if the set X_1 contains a polynomial of the form $p(z) = (z - a)^\kappa z^\lambda$ with $|a| > 0$ and $\lambda > 0$ then it also contains the polynomial $q(z) = (z - a)^{\kappa+1} z^{\lambda-1}$. Indeed, for $\Delta, \delta \in \mathbb{C}$ let

$$p_{\Delta, \delta}(z) = (z - a - \Delta)^\kappa z^{\lambda-1}(z - \delta), \quad f(\Delta, \delta) = L(p_{\Delta, \delta}) - L(g).$$

Note that $f(\Delta, \delta)$ is a polynomial, linear with respect to δ , i.e

$$f(\Delta, \delta) = f_0(\Delta) + f_1(\Delta)\delta.$$

By construction $f_0(0) = 0$. Moreover, $q \in X_1$ implies $f_1(0) = 0$, as otherwise $p_{\Delta,\delta} \in X_0$ and $r_1(p_{\Delta,\delta}) < r_1(q)$ for

$$\Delta = -at, \quad \delta = -\frac{f_0(\Delta)}{f_1(\Delta)},$$

where $t > 0$ is sufficiently small. Since $f_0(0) = f_1(0) = 0$, we have $p_{0,a} \in X_1$, which completes the proof. ■

Now the stage is set for the proof of Theorem 6.

Proof: Each choice of B_0, B_1, \dots, B_n corresponds to a linear functional $L : \mathcal{P}_n \mapsto \mathbb{C}$ of the form $L(p) = \sum_{i=0}^n B_i a_i$. Thus, we wish to prove that given a linear functional $L : \mathcal{P}_n \mapsto \mathbb{C}$ and a polynomial $g \in \mathcal{P}_n^1$, the minimum of $r_1(p)$ over all polynomials $p \in \mathcal{P}_n^1$ satisfying the constraint $L(p) = L(g)$ can be attained on a polynomial $p = f$ of the form $f(z) = (z - a)^n$ for some $a \in \mathbb{C}$. But this is exactly the statement of Corollary 18 proved above. With the existence of the minimizer with the claimed structure established, the property that $-\gamma$ is a root of the polynomial h now follows from the fact that $(z - \gamma)^n$ is in P if and only if $h(-\gamma) = 0$. ■

APPENDIX B

PROOF OF THEOREM 14

We will derive Theorem 14 from Theorem 6, proved in Appendix A. Let the linear functional $L_n : \mathcal{P}_n \mapsto \mathbb{C}$ be defined by

$$L_n p = \lim_{z \rightarrow \infty} z^{-n} p(z).$$

We have $p \in \mathcal{P}_n^1$ if and only if $L_n p = 1$. Theorem 6 is equivalent to saying that given a linear functional $F : \mathcal{P}_n \mapsto \mathbb{C}$, $F \neq 0$, $F \neq L_n$, the minimal root radius over the set

$$\{q \in \mathcal{P}_n \mid Fq = 0, L_n q = 1\}$$

is attained at a polynomial q^* with only one distinct root. It is easy to see that replacing the constraint $L_n q = 1$ with $L_n q = c$ for any $c \in \mathbb{C}, c \neq 0$ would not change the minimal root radius, as the optimizer would simply become cq^* . As a consequence, we have the following theorem, equivalent to Theorem 6:

Theorem 19: Given a linear functional $F : \mathcal{P}_n \mapsto \mathbb{C}$ such that $F \neq cL_n$ for every $c \in \mathbb{C}$, the root radius ρ achieves its minimum on

$$\Omega_F = \{q \in \mathcal{P}_n \mid Fq = 0, L_n q \neq 0\}$$

at a polynomial $q^* = g_a$ of the form $g_a(z) = (z - a)^n$ for some $a \in \mathbb{C}$.

To prove Theorem 14, it suffices to show that the minimal abscissa is attained at a polynomial with only one distinct root γ , since then $-\gamma$ would have to be the rightmost root (root with the largest real part) of h . We will prove the following equivalent statement:

Theorem 20: Given a linear functional $G : \mathcal{P}_n \mapsto \mathbb{C}$ such that $G \neq cL_n$ for every $c \in \mathbb{C}$, the root abscissa α achieves its minimum on

$$\Omega_G = \{p \in \mathcal{P}_n : Gp = 0, L_n p \neq 0\}$$

at a polynomial $p^* = g_\gamma$ of the form $g_\gamma(s) = (s - \gamma)^n$ for some $\gamma \in \mathbb{C}$.

Proof: The proof follows from Theorem 19 as we now explain. It is sufficient to demonstrate that

(*) *if $p \in \Omega_G$ and $\alpha(p) < \sigma$ for some $\sigma \in \mathbb{R}$, then there exists $d \in \mathbb{C}$ such that the polynomial $r(s) = g_d(s) = (s - d)^n$ satisfies $r \in \Omega_G$ and $\alpha(r) < \sigma$.*

This is because (*) implies that the optimizer can be chosen to have one (distinct) root. Notice that (*) states implicitly that the optimum is attained, because every such d is a root of the polynomial $\tilde{p}(w) = Gg_w$, and so there is a finite number of choices of $d \in \mathbb{C}$ such that $g_d \in \Omega_G$. The optimal abscissa would then be one of such d 's with the smallest real part.

To prove (*), for $\sigma \in \mathbb{R}$, let A_σ be the function $A_\sigma : \mathcal{P}_n \mapsto \mathcal{P}_n$ mapping $p \in \mathcal{P}_n$ to $q = A_\sigma p \in \mathcal{P}_n$ defined by the identity

$$q(z) = (z - 1)^n p \left(\sigma + \frac{z + 1}{z - 1} \right) \quad (z \neq 1).$$

Note that A_σ is linear and invertible. If $p \in \Omega_G$, $q = A_\sigma p$, and $\alpha(p) < \sigma$, then $q(z) \neq 0$ for $|z| \geq 1$, because

$$\operatorname{Re}(s) \geq \sigma \quad \text{for } s = \sigma + \frac{z + 1}{z - 1}, \quad z \neq 1, \quad |z| \geq 1,$$

and

$$q(1) = \lim_{z \rightarrow 1} (z - 1)^n p \left(\sigma + \frac{z + 1}{z - 1} \right) = 2^n L_n p \neq 0,$$

so $\rho(q) < 1$. In addition,

$$L_n q = \lim_{z \rightarrow \infty} \frac{(z - 1)^n}{z^n} p \left(\sigma + \frac{z + 1}{z - 1} \right) = p(\sigma + 1) \neq 0,$$

which implies that $q \in \Omega_F$ for $F = GA_\sigma^{-1}$, and that $F \neq cL_n$ for every $c \in \mathbb{C}$.

By Theorem 19, there exists $a \in \mathbb{C}$ such that $|a| < 1$ and the polynomial $q_0(z) = (z - a)^n$ is in Ω_F . Let $r_0 = A_\sigma^{-1}q_0$. By definition,

$$(z - 1)^n r_0 \left(\sigma + \frac{z + 1}{z - 1} \right) = (z - a)^n,$$

which means that

$$r_0 \left(\sigma + \frac{z + 1}{z - 1} \right) = \left(\frac{z - a}{z - 1} \right)^n = \left(\frac{1 + a}{2} + \frac{1 - a}{2} \frac{z + 1}{z - 1} \right)^n,$$

i.e.

$$r_0(s) = \left(\frac{1 - a}{2} \right)^n \left(s - \sigma - \frac{a + 1}{a - 1} \right)^n.$$

Since $Gr_0 = FA_\sigma r_0 = Fq_0 = 0$, we can set $r(s) = (s - d)^n$ with

$$d = \sigma + \frac{a + 1}{a - 1}.$$

and we have $\operatorname{Re}(d) < \sigma$ since $|a| < 1$. Therefore, $\alpha(r) < \sigma$. ■

REFERENCES

- [BGL95] V. Blondel, M. Gevers, and A. Lindquist. Survey on the state of systems and control. *European Journal of Control*, 1:5–23, 1995.
- [BHLO06] J.V. Burke, D. Henrion, A.S. Lewis, and M.L. Overton. Stabilization via nonsmooth, nonconvex optimization. *IEEE Transactions on Automatics Control*, 51:1760–1769, 2006.
- [Blo94] V. Blondel. *Simultaneous Stabilization of Linear Systems*. Lecture Notes in Control and Information Sciences 191. Springer, Berlin, 1994.
- [BLO01] J.V. Burke, A.S. Lewis, and M.L. Overton. Optimizing matrix stability. *Proceedings of the American Mathematical Society*, 129:1635–1642, 2001.
- [BLO04] J.V. Burke, A.S. Lewis, and M.L. Overton. Variational analysis of the abscissa mapping for polynomials via the Gauss-Lucas theorem. *Journal of Global Optimization*, 28:259–268, 2004.
- [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.
- [Che79a] R. Chen. On the problem of direct output feedback stabilization. In *Proceedings of MTNS*, pages 412–414, 1979.
- [Che79b] R. Chen. *Output Feedback Stabilization of Linear Systems*. PhD thesis, University of Florida, 1979.
- [FM78] A. T. Fam and J. S. Meditch. A canonical parameter space for linear systems design. *IEEE Trans. Automat. Control*, 23(3):454–458, 1978.
- [GO07] K. K. Gade and M. L. Overton. Optimizing the asymptotic convergence rate of the Diaconis-Holmes-Neal sampler. *Adv. in Appl. Math.*, 38(3):382–403, 2007.
- [HO06] D. Henrion and M. L. Overton. Maximizing the closed loop asymptotic decay rate for the two-mass-spring control problem. Technical Report 06342, LAAS-CNRS, March 2006. <http://homepages.laas.fr/henrion/Papers/massspring.pdf>.

- [HP05] D. Hinrichsen and A.J. Pritchard. *Mathematical Systems Theory I: Modelling, State Space Analysis, Stability and Robustness*. Springer, Berlin, Heidelberg and New York, 2005.
- [HUL93] J.B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*. Springer-Verlag, New York, 1993. Two volumes.
- [Mar66] M. Marden. *Geometry of Polynomials*. American Mathematical Society, 1966.
- [Ran89] A. Rantzer. Equivalence between stability of partial realizations and feedback stabilization—applications to reduced order stabilization. *Linear Algebra Appl.*, 122/123/124:641–653, 1989.