Pure and Applied Functional Analysis
Volume 3, Number 4, 2018, 639–651

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

OPTIMIZATION OF A CAUCHY RADIUS IMPROVEMENT

AARON MELMAN

Abstract. The Cauchy radius of a polynomial is a well-known upper bound on the magnitude of the largest zero of a polynomial. What is less well-known is that it was relatively recently improved by Rahman and Schmeisser who showed that a better Cauchy radius can be obtained by multiplying the polynomial by another one. Although better, it is not necessarily optimal, and what we have set out to do here is to find the optimal multiplier from a class of multipliers for polynomials with real coefficients. It requires the minimization of an nonsmooth nonconvex implicitly defined nonlinear function. Surprisingly, the optimal multiplier can be narrowed down, in the worst case, to one of two candidates in a finite number of steps without ever solving a single nonlinear equation.

1. Introduction

A classical result by Cauchy from 1829 ([1], [3, Th.(27,1), p.122 and Exercise 1, p.126]) states that all the zeros of the polynomial $p(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ with complex coefficients, lie in the disk defined by $|z| \leq \rho[p]$, where $\rho[p]$ is the *Cauchy radius* of p. It is the unique positive solution of the real nonlinear equation $f(x) = 0$, where

$$f(x) = |a_n|x^n - |a_{n-1}|x^{n-1} - \cdots - |a_1|x - |a_0| \,. \tag{1.1}$$

A lower bound on the magnitude of the zeros can be obtained by applying this result to the reverse polynomial (when $a_0 \neq 0$), whose zeros are the reciprocals of those of p. This bound is the best of all bounds that depend only on the moduli of the coefficients, and is therefore not easy to improve. A less well-known improvement of this result was published in a book from 2002 by Rahman and Schmeisser (Theorem 8.3.1 in [5]), where, if a_{n-k} is the first nonzero coefficient after a_n, it was shown that the Cauchy radius of the polynomial $(a_n z^k - a_{n-k})p(z)$ (whose zeros include those of p) is not larger than that of p. In practice, it is almost always smaller. In fact, the theorem remains true even when a_{n-k} is not the first nonzero coefficient.

Whether there exist optimal multipliers, and how to find them if they do, are difficult questions to answer, but we can provide a partial answer. Since we already know from [5] that a multiplier exists that guarantees a smaller (or, at least, not larger) Cauchy radius, it is natural to try and find an optimal multiplier that, of all multipliers of the form $a_n z^k - t$, provides the smallest Cauchy radius.

2010 *Mathematics Subject Classification.* 12D10, 30C15, 65H04, 65K10.
Key words and phrases. bound, zero, root, real polynomial, polynomial multiplier, Cauchy.

We will solve this problem when the coefficients of the polynomial p are real, which requires the minimization of a nonsmooth nonconvex implicitly defined nonlinear function.

Our main motivation is to show that, in spite of such adverse properties, such a problem can be solved efficiently. Specifically, we found - surprisingly - that the optimal multiplier can be narrowed down, in the worst case, to one of two candidates in a finite number of steps without ever solving a single nonlinear equation. Although we are less interested in practical considerations, we will address these as well.

In addition to [5], the interested reader can find many results about polynomial bounds in [3] and [4].

The paper is organized as follows. In Section 2 we collect preliminary results that we will need in Section 3 to obtain our main results. At the end of that section we provide a few numerical examples. The appendix contains the proof of a lemma.

2. Preliminaries

Throughout, we consider a polynomial $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ such that $n \geq 3$, $a_n a_0 \neq 0$, and p is not a monomial or of the form $a_n z^n - a_0$. This avoids either trivial situations or situations that can be simplified. We will also frequently use g_t instead of $\partial g / \partial t$.

We begin by examining the function f defined in (1.1) and its derivatives. The conditions on p imply that $f(0) < 0$. They also imply that $f'(\varepsilon)$ is strictly negative for a sufficiently small positive value of ε, and the same is true for f'', unless it is a monomial, in which case it is positive for positive values of x. Since f, f', and f'' all become unbounded as $x \to +\infty$, f and f' must have at least one positive root, and the same is true for f'' unless it is a monomial. From Descartes' rule of signs, f and f' each have a unique positive root with multiplicity one, since their coefficients change sign only once, and the same is true for f'' unless it is a monomial (when $f''(x) > 0$ for $x > 0$).

Denoting the positive root of f by r, these considerations imply that f must have a minimum at a point s with $0 < s < r$, $f'(s) = 0$, and $f''(s) > 0$. As a result, we have that $f(r) > 0$, $f'(r) > 0$, and $f''(r) > 0$. Figure 1 illustrates the behavior of f.

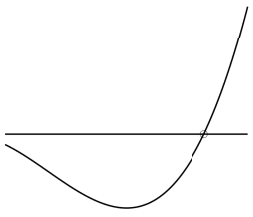

Figure 1. Graph of f.

Next, and to avoid digressions later on, we make a few basic observations about the solutions of equations of the form $\sum_{j=1}^{m} \gamma_j |t - b_j| = \beta$ where $t, b_j \in \mathbb{R}$ and $\gamma_j, \beta > 0$, that will be described in Lemma 2.1. Its proof is technical and will be deferred to the appendix; instead, we present the following simple example that exhibits all the important properties of the solutions.

Consider the function

$$\psi_1(t) := 2|t+1| + \frac{1}{2}|t-1| + |t-2| \ ,$$

which is continuous and piecewise linear as its expression on the intervals $(-\infty, -1]$, $[-1, 1]$, $[1, 2]$, and $[2, +\infty)$ shows:

$$\psi_1(t) = \begin{cases} -2(t+1) - (1/2)(t-1) - (t-2) & (-\infty < t \le -1) \\ 2(t+1) - (1/2)(t-1) - (t-2) & (-1 \le t \le 1) \\ 2(t+1) + (1/2)(t-1) - (t-2) & (1 \le t \le 2) \\ 2(t+1) + (1/2)(t-1) + (t-2) & (2 \le t \le +\infty) \ , \end{cases}$$

which becomes

$$\psi_1(t) = \begin{cases} -(7/2)t + 1/2 & (-\infty < t \le -1) \\ (1/2)t + 9/2 & (-1 \le t \le 1) \\ (3/2)t + 7/2 & (1 \le t \le 2) \\ (7/2)t - 1/2 & (2 \le t \le +\infty) \ . \end{cases}$$

The slope increases monotonically from one interval to the next, as can be seen from the graph of ψ_1 on the left in Figure 2. Clearly, $\psi_1(t) = \beta$ for $\beta > 0$ can only have two, one, or no solutions, as illustrated by the dashed horizontal lines in the figure. If there is one solution, then it is obtained at one of the points $\{-1, 1, 2\}$, which in this case is -1. If there are two solutions, then they must lie in different intervals. If one of the slopes had been zero on an interval and β would have coincided with the function value there, then all values of t in that interval would have been solutions. A function that exhibits such behavior is given by

$$\psi_2(t) := 2|t+1| + |t-1| + |t-2| \ ,$$

and its graph is shown on the right in Figure 2.

The possible number of solutions is not a surprise since they are the intersection of a line (the real axis) with the boundary of the convex set (e.g., for ψ_1) $\left\{ z \in \mathbb{C} \ : \ 2|z+1| + \frac{1}{2}|z-1| + |z-2| \le \beta \right\}$, which is a so-called "*weighted 3-ellipse*" (see, e.g., [7]). A weighted k-ellipse is the set of all points such that the weighted sum of their distances to k foci is a constant; it is a generalization of an ellipse. Although it lacks some of the basic properties of an ellipse, e.g., it does not necessarily contain its foci, it does share the important property that the region it encloses is convex.

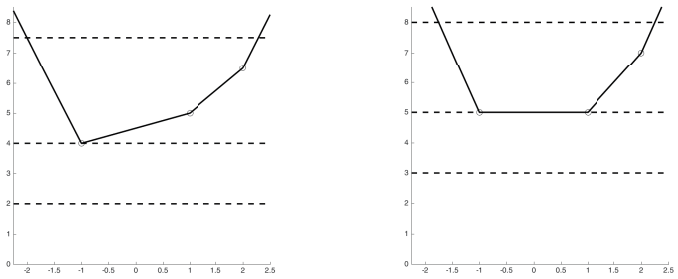

FIGURE 2. Graph of ψ_1 (left) and ψ_2 (right).

The following lemma formalizes the conclusions from our example. Its proof can be found in the appendix.

Lemma 2.1. *Let*

$$\varphi(t) = \sum_{j=1}^{m} \gamma_j |t - b_j| \ ,$$

where $\gamma_j > 0$ *and* $b_j \in \mathbb{R}$ *with* $b_1 < b_2 < \cdots < b_m$. *Define* $I_0 = (-\infty, b_1]$, $I_m = [b_m, +\infty)$, *and* $I_j = [b_j, b_{j+1}]$ *for* $1 \leq j \leq m-1$. *Denote the interior of* I_j *by* I_j'.

Then the equation $\varphi(t) = \beta$ *with* $\beta > 0$ *can only have infinitely many, two, one, or no solutions with the following properties.*

- *If there is only one solution, then it is obtained at one of the nodes* b_j *with* $1 \leq j \leq m$.
- *If there are two solutions* t_1 *and* t_2, *with* $t_1 < t_2$, *then* $t_1 \in I_{j_1}$ *and* $t_2 \in I_{j_2}$, *where* $j_1 < j_2$ *and* $0 \leq j_1, j_2 \leq m$.
- *If there are infinitely many solutions, then all solutions lie in one interval* I_j *with* $1 \leq j \leq m-1$.

3. Multiplier optimization

We are now ready to tackle the minimization problem

$$\min_{-\infty \leq t \leq +\infty} \left\{ \rho \left[\left(a_n z^k - t \right) p(z) \right] \right\} \ . \tag{3.1}$$

Throughout this section, we consider a polynomial $p(z) = a_n z^n + a_{n-k} z^{n-k} + \cdots + a_1 z + a_0$ with real coefficients, where a_{n-k} is the first nonzero coefficient after a_n, i.e., $a_{n-1} = a_{n-2} = ... = a_{n-k+1} = 0$ and $a_{n-k} \neq 0$. From Theorem 8.3.1 in [5], we know that there exists a multiplier of the form $a_n z^k - a_{n-k}$ so that

$$\rho \left[\left(a_n z^k - a_{n-k} \right) p \right] \leq \rho \left[p \right] \ ,$$

which motivates the form of the multiplier in (3.1).

We begin by finding an explicit expression for $q(t,z) := \left(a_n z^k - t\right) p(z)$. Straightforward polynomial multiplication yields

$$\begin{aligned}
q(t,z) &= \left(a_n z^k - t\right) \sum_{j=0}^{n} a_j z^j \\
&= \sum_{j=0}^{n} a_n a_j z^{j+k} - \sum_{j=0}^{n} t a_j z^j \\
&= a_n^2 z^{n+k} + \left(a_n a_{n-k} - t a_n\right) z^n + \sum_{j=0}^{n-k-1} a_n a_j z^{j+k} - \sum_{j=0}^{n-k} t a_j z^j \\
&= a_n^2 z^{n+k} + \left(a_n a_{n-k} - t a_n\right) z^n + \sum_{j=k}^{n-1} a_n a_{j-k} z^j - \sum_{j=0}^{n-k} t a_j z^j \\
&= a_n^2 z^{n+k} + \left(a_n a_{n-k} - t a_n\right) z^n + \sum_{j=k}^{n-1} a_n a_{j-k} z^j \\
&\qquad\qquad - \sum_{j=k}^{n-k} t a_j z^j - \sum_{j=0}^{k-1} t a_j z^j \\
&= a_n^2 z^{n+k} + \left(a_n a_{n-k} - t a_n\right) z^n + \sum_{j=n-k+1}^{n-1} a_n a_{j-k} z^j \\
&\qquad\qquad + \sum_{j=k}^{n-k} \left(a_n a_{j-k} - t a_j\right) z^j - \sum_{j=0}^{k-1} t a_j z^j \ .
\end{aligned}$$

If the set of admissible indices j in a summation is empty or if the coefficient index is inadmissible, the summation or coefficient is replaced by zero. If we define the real polynomial g as

$$\begin{aligned}
g(t,x) = |a_n|^2 x^{n+k} - |a_n a_{n-k} - t a_n| x^n - \sum_{j=n-k+1}^{n-1} |a_n a_{j-k}| x^j \\
- \sum_{j=k}^{n-k} |a_n a_{j-k} - t a_j| x^j - |t| \sum_{j=0}^{k-1} |a_j| x^j \ .
\end{aligned} \tag{3.2}$$

then, for a given t, the Cauchy radius $\rho[q(t,z)]$ is the unique positive root of $g(t,x) = 0$. To write $g(t,x)$ in a more suitable way, we need the following definitions:

$$\begin{aligned}
S &= \{k, k+1, ..., n-k\} \ , \\
S_1 &= \{j : j \in S \,\&\, a_j = 0\} \ , \\
S_2 &= \{j : j \in S \,\&\, a_j \neq 0\} \ ,
\end{aligned}$$

where $S = \emptyset$ if $k > n-k$, so that $S = S_1 \cup S_2$ and $S_1 \cap S_2 = \emptyset$. Note that, when $S \neq \emptyset$, $n-k$ is always in S_2. With these definitions, $g(t,x)$ becomes

$$g(t,x) = -|a_n||t - a_{n-k}|x^n - \sum_{j\in S_2} |a_j||t - a_n a_{j-k}/a_j|x^j - |t|\sum_{j=0}^{k-1}|a_j|x^j + |a_n|^2 x^{n+k} - \sum_{j=n-k+1}^{n-1} |a_n a_{j-k}|x^j - \sum_{j\in S_1} |a_n a_{j-k}|x^j \ . \tag{3.3}$$

Clearly, for a given value of x, $g(t,x)$ is a continuous piecewise linear function of t with nodes at 0, a_{n-k}, and $a_n a_{j-k}/a_j$ for $j \in S_2$. Without loss of generality, we will assume that all these nodes are distinct and that their number is m. For convenience, we relabel them $b_1, b_2, \ldots, b_m$, where $b_1 < b_2 < \cdots < b_m$. As in Lemma 2.1, these nodes divide the real line into $m+1$ intervals I_j and the equation $g(t,x) = 0$ for a fixed value of x will have one, two, or no solutions for t, or an interval of solutions. With this notation, we have

$$g(t,x) = -\sum_{j=1}^{m} \alpha_j(x)|t - b_j| + |a_n|^2 x^{n+k} - \sum_{j=n-k+1}^{n-1} |a_n a_{j-k}|x^j - \sum_{j\in S_1} |a_n a_{j-k}|x^j \ , \tag{3.4}$$

where the $\alpha_j(x)$ are functions of x only, of the form $|a_j|x^j$ or a sum of such expressions. When $t \in I_\ell$, $1 \le j \le m-1$, then

$$\begin{cases} |t - b_j| = t - b_j & (j \le \ell) \\ |t - b_j| = b_j - t & (j > \ell) \ . \end{cases}$$

On the interval I_ℓ, $0 \le \ell \le m$, $g(t,x)$ is therefore given by

$$g(t,x) = -\sum_{j=1}^{\ell} \alpha_j(x)(t - b_j) + \sum_{j=\ell+1}^{m} \alpha_j(x)(t - b_j) + |a_n|^2 x^{n+k} - \sum_{j=n-k+1}^{n-1} |a_n a_{j-k}|x^j - \sum_{j\in S_1} |a_n a_{j-k}|x^j \ , \tag{3.5}$$

where we have once again used the convention that a summation with an inadmissible range is replaced by zero.

As a result, for any $t \in I'_\ell$ (the interior of I_ℓ), one obtains

$$\frac{\partial g}{\partial t}(t,x) = -\sum_{j=1}^{\ell} \alpha_j(x) + \sum_{j=\ell+1}^{m} \alpha_j(x) \ \ , (0 \le \ell \le m) \ ,$$

i.e., for given x, g_t is constant and independent of t on the interior of each interval. Since $\alpha_j(x) > 0$ for $x > 0$, we obtain for a given value $x = \bar{x} > 0$ that

$$\begin{cases} g_t(t_1, \bar{x}) > g_t(t_2, \bar{x}) & (t_1 \in I'_i, t_2 \in I'_j \text{ and } i < j) \\ g_t(t_1, \bar{x}) = g_t(t_2, \bar{x}) & (t_1, t_2 \in I'_j) \ , \end{cases} \tag{3.6}$$

i.e., $g_t(t, \bar{x})$ decreases as t increases from I'_j to I'_{j+1}.

For a fixed value $t = \bar{t}$, the function $g(\bar{t}, x)$ is of the same form as f in (1.1) and therefore has the same properties that were listed at the beginning of Section 2. If

$\bar{x} = \rho[q(\bar{t}, z)]$, then these properties imply that

$$\frac{\partial g}{\partial x}(\bar{t}, \bar{x}) > 0 \quad \text{and} \quad \frac{\partial^2 g}{\partial x^2}(\bar{t}, \bar{x}) > 0 \, . \tag{3.7}$$

We now use all of the information we collected on $g(t, x)$ to describe the properties of $\rho[q(t, z)]$ in the following theorem.

Theorem 3.1. *Let $p(z) = a_n z^n + a_{n-k} z^{n-k} + \cdots + a_1 z + a_0$ with real coefficients be such that $n \geq 3$, $a_n a_0 \neq 0$, $a_{n-1} = a_{n-2} = ... = a_{n-k+1} = 0$, and $a_{n-k} \neq 0$ for $k \in \{1, 2, ..., n-1\}$. Then the continuous function $h(t) = \rho\left[\left(a_n z^k - t\right) p\right]$ is continuously differentiable everywhere, except at a finite number of nodes. Its global minimum is obtained either at one of these nodes, or on an interval whose endpoints are adjacent nodes. Any horizontal line above the minimum value intersects the graph of h in exactly two points. Strictly in between adjacent nodes, h' does not change its sign.*

Proof. For a given value of t, $h(t) = \rho\left[\left(a_n z^k - t\right) p\right]$ is the unique positive root of $g(t, x)$, defined in (3.2). To each t value corresponds exactly one x value and, since a polynomial's roots are continuous functions of its coefficients, which here are themselves continuous functions of t, we conclude that $h(t)$ is a continuous function for any t. In addition, as $t \to \pm\infty$, $h(t) \to +\infty$, which can be seen from (3.2): setting $g(t, x) = 0$ as $t \to \pm\infty$ shows that the equation can only be balanced by letting $x \to +\infty$. We also saw that $g(t, x)$ can be written as in (3.4), where the m nodes b_j, which only depend on the coefficients of p, were determined by (3.3). These nodes, ordered in increasing order, determine $m+1$ intervals I_j spanning the entire real axis.

Now assume that $x_0 = h(t_0) = \rho[q(t_0, x)]$ for a value of t_0 in I'_ℓ (the interior of I_ℓ) for some $0 \leq \ell \leq m$. Then x_0 is the unique positive root of $g(t_0, x) = 0$, where $g(t, x)$ is as in (3.5). In addition, for any $\ell \in \{0, 1, \ldots, m\}$, $g(t, x) : I'_\ell \times \mathbb{R}^+ \to \mathbb{R}^+$ is a continuously differentiable function with, by (3.7), $\partial g / \partial x > 0$ at (t_0, x_0). From the implicit function theorem (e.g., [6, Theorem 9.28]), we then know that the equation $g(t, x) = 0$ can be solved for x in terms of t in a neighborhood of (t_0, x_0) on which $x(t)$ is also continuously differentiable. We have from its definition that $h(t) = x(t)$. To compute $x'(t)$ we can use the equation $g(t, x) = 0$ to obtain

$$\frac{dg}{dt} = \frac{\partial g}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial g}{\partial t} = 0 \, ,$$

which yields

$$\frac{dx}{dt} = -\left(\frac{\partial g}{\partial x}\right)^{-1} \cdot \frac{\partial g}{\partial t} \, . \tag{3.8}$$

From Lemma 2.1 and (3.4) we know that, for fixed $\bar{x} > 0$, an equation of the form $h(t) = \bar{x}$, whose solutions are those of $g(t, \bar{x}) = 0$, either has no solutions or it has one, two, or an interval of solutions. This means that, as t goes from $-\infty$ to $+\infty$ and bearing in mind that $h(t)$ becomes unbounded for $t \to \pm\infty$, its graph has to start with a negative slope and end with a positive one, with or without a zero slope in between. It is a continuous function, so it must have a global minimum, but it cannot have other extrema. If the minimum x^* is obtained at a point $t^* \in I'_\ell$, where

$x(t)$ (and therefore $h(t)$) is differentiable, then, with (3.8), $g_t(t^*, x^*) = x'(t^*) = 0$. Because $g_t(t, x^*)$ is piecewise linear in t, $g_t(t^*, x^*) = 0$ implies that its slope on I_ℓ vanishes. Consequently, $x(t) = x^*$ on I_ℓ, and we obtain a single interval - from one node to an adjacent one - of values for t at which h achieves its minimum.

This also means that when the minimum is obtained at a single point, then this point cannot be in the interior of one of the intervals I_j, so that it must necessarily be a node.

Finally, h' does not change its sign on the interior of any interval because this would require it to vanish at an interior point. As we just saw, that can only happen when the slope is zero over the entire interval and this can be the case for at most one interval, on which, being equal to zero, h' also does not change its sign. This concludes the proof. □

Remark. Although it follows from the general shape of the graph of h that when a horizontal line intersects it in exactly two points, the slopes at these points must have opposite signs, it can also be demonstrated analytically, as follows. If there are exactly two values of t corresponding to a specific value of $\bar{x}$, namely, t_1 and t_2 with $t_1 < t_2$, $t_1 \in I'_{j_1}$, and $t_2 \in I'_{j_2}$, and $j_1 < j_2$, and if $h'(t_1) = x'(t_1) > 0$, then (3.6) would imply that $x'(t_2) > 0$ as well because it implies that $-\partial g/\partial t$ is less on I'_{j_1} than on I'_{j_2}. This is impossible because of the continuity of $h(t)$ and the fact that there are only two solutions. Thus $x'(t_1) < 0$ since it cannot vanish as that would imply an interval of solutions. Likewise, $x'(t_2) > 0$.

Examples. To illustrate the graph of h, we consider the polynomials

$$p_1(z) = z^5 - 4z^4 + z^3 - z^2 + 2z - 1 \quad \text{and} \quad p_2(z) = z^5 - 4z^4 + 3z^3 - z^2 + 4z + 1\,.$$

Figure 3 shows the graphs of the associated h_1 and h_2 functions on the left and right, respectively. The small circles indicate the nodes. There are six nodes for p_1, but only five for p_2 as two of them coincide. The bottom horizontal dashed line indicates the largest modulus of the zeros of the polynomial in question (3.7729 for p_1 and 2.9117 for p_2), the middle dashed line indicates the value of the bound obtained from Theorem 8.3.1 in [5] (4.0000 for p_1 and 4.0035 for p_2), whereas the top dashed line corresponds to $h(0)$, which is precisely the Cauchy radius (4.3134 for p_1 and 4.7205 for p_2).

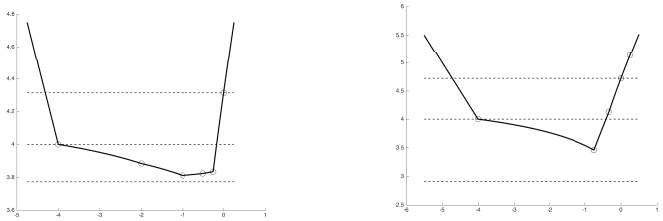

FIGURE 3. Graph of h_1 (left) and h_2 (right).

Computation of the minimum

We now propose a method, based on the properties of h, that, at least in theory, will

narrow down the optimal t among all polynomial multipliers of the form $a_n z^k - t$ to one of two nodes in a finite number of steps. Its approach is, in a sense, the opposite of how one proceeds to compute the Cauchy radius: rather than compute the value of x that corresponds to a given value of t, it starts with a given value of x and finds the two corresponding t values that bracket the optimal value(s). The process starts with upper and lower bounds on the minimum value and then bisects these bounds: if no t values correspond to the bisected value then this becomes the new lower bound; otherwise it becomes the new upper bound with two new values of t providing a better bracket for the optimal value(s). The process is then repeated until the two values of t bracket either a single node, or a pair of adjacent nodes. If a single node is bracketed, then the minimum of h is attained at this node and the process stops. If a pair of adjacent nodes is bracketed, then the minimum is attained at one of the nodes, or it is attained at all points of the closed interval determined by the nodes. In theory (i.e., if all values can be computed exactly), such a procedure will terminate in a finite number of steps, for each of which a piecewise linear equation needs to be solved, a simple matter. Not a single nonlinear equation needs to be solved. The reason one must stop when two nodes are bracketed is that, when there is an interval of optimal values, the process will continue indefinitely, unless one is extraordinarily lucky to precisely hit the minimum value.

If so desired, one can then compute the Cauchy radius obtained from the optimal node, or, if there are two, with one of them, or both, or a weighted average of the two. One could also try to predict the number of solutions by computing derivative values.

In this regard, we note that, as a practical matter, one would naturally be more interested in the bound than in the value of t for which that bound is obtained, in which case one might forgo the computation of both t-values and simply stop when the upper and lower bounds are sufficiently close.

Example

For a concrete explanation of the above procedure, let us consider $p_1(z) = z^5 - 4z^4 + z^3 - z^2 + 2z - 1$ from the previous example. Here, $k = 1$, and the class of multipliers that we want to optimize over is of the form $z - t$. We have that

$$(z-t)p_1(z) = z^6 - (t+4)z^5 + 4\left(t + \frac{1}{4}\right)z^4 - (t+1)z^3 + (t+2)z^2 - 2\left(t + \frac{1}{2}\right)z + t\,.$$

The Cauchy radius of $(z-t)p_1(z)$ for fixed t is given by the unique positve solution x of $g(t,x) = 0$, where

$$g(t,x) = x^6 - |t+4|x^5 - 4\left|t + \frac{1}{4}\right|x^4 - |t+1|x^3 - |t+2|x^2 - 2\left|t + \frac{1}{2}\right|x - |t|\,.$$

Using the same terminology as before, the nodes of the piecewise linear function $g(t,x)$ of t for fixed x are $-4, -2, -1, -1/2, -1/4, 0$, which can be seen on the left in Figure 3.

To start, we use the somewhat crude but easily and explicitly computable upper bound on the moduli of the zeros of p_1 from (8.1.9) in [5], which, for a polynomial

$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$ with $a_0 a_n \neq 0$, is given by

$$1 + \max\left\{\left|\frac{a_0}{a_n}\right|, \left|\frac{a_1}{a_n}\right|, \ldots, \left|\frac{a_{n-1}}{a_n}\right|\right\} \geq \rho[p] \; .$$

To obtain a lower bound, we apply this bound to the reverse polynomial $p^{\#}(z) = z^n p(1/z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n$, whose zeros are the reciprocals of the zeros of p. For any zero ζ of p we then obtain

$$1 + \max\left\{\left|\frac{a_0}{a_n}\right|, \left|\frac{a_1}{a_n}\right|, \ldots, \left|\frac{a_{n-1}}{a_n}\right|\right\} \leq |\zeta| \leq \frac{1}{1 + \max\left\{\left|\frac{a_n}{a_0}\right|, \left|\frac{a_{n-1}}{a_0}\right|, \ldots, \left|\frac{a_1}{a_0}\right|\right\}} \; .$$

If we apply these simple bounds to p_1, one easily obtains that any zero ζ of p_1 satisfies $0.2 \leq |\zeta| \leq 5$. Therefore, for any $t \in \mathbb{R}$, $0.2 \leq \rho[(a_n z - t)p_1]$, and we also have that $\rho[p_1] < 5$, which corresponds to $t = 0$. This means that $g(t, 5) = 0$ must have two solutions. We find that they are the endpoints of the interval $[-5, 0.401]$, which contains all six nodes. For the next upper bound we try $(0.2 + 5)/2 = 2.6$. However, $g(t, 2.6) = 0$ has no solutions, implying that 2.6 replaces 0.2 as the new lower bound on the Cauchy radius of $(z - t)p_1(z)$. Next, we try $(2.6 + 5)/2 = 3.8$ as the upper bound. Again, $g(t, 3.8) = 0$ has no solutions and 3.8 replaces 2.6 as the next lower bound, leading to the next upper bound $(3.8 + 5)/2 = 4.4$. This time $g(t, 4.4) = 0$ has two solutions, namely, the endpoints of $[-4.4, 0.048]$, an interval that now contains five nodes. Continuing in this fashion, we obtain successive nested intervals (from right to left)

$$[-1.3130, -0.2486] \subseteq [-2.9718, -0.1292] \subseteq [-4.1, -0.1148]$$

containing, respectively, five, four and three nodes, until the final interval $[-1.0852, -0.6742]$, containing exactly one node, -1, which must therefore be the node that produces the optimal multiplier, i.e., $z + 1$. This last interval corresponds to an upper bound of 3.8187 on the moduli of the zeros of p_1, obtained by solving simple piecewise linear equations. We note that this bound is already better than the Cauchy radius (4.3134), which requires the solution of a nonlinear equation.

Numerical considerations

Although it was not our main purpose here, we briefly address a few numerical issues that a more sophisticated algorithm would have to take into account. The function h may have a slope that is almost zero over an interval that includes several nodes, so that the number of steps, while remaining finite, may increase considerably. However, even if that were not the case, numerical cancellation problems, due to finite precision calculations, may make the slope appear to be zero, in which case the process would generally never terminate. One also needs to consider that we are merely computing a bound, and this should not entail more computation than to compute the roots themselves. Some safeguards should therefore be put in place that keep the computational cost at a reasonable level, e.g., no more than the computational cost of computing a single Cauchy radius. In addition, the bisection method described above is rather crude and can be improved.

Numerical examples

To illustrate the aforementioned method, we have generated 1000 polynomials with coefficients that are uniformly randomly distributed on $[-4, 4]$ with k=1 and

$n = 20, 60, 100$. We then applied the aforementioned method and compared the results to the Cauchy radius of p and its improvement from [5]. In Table 1 we have listed the average of their ratios to the modulus of the largest zero, i.e., the closer this number is to unity, the better it is. These numbers can be found in the second, third, fourth, and fifth columns for, respectively, the Cauchy radius of p, the improvement from [5], the upper bound obtained at the moment the two candidate optimal nodes are identified, i.e., without attempting any further refinement, and, finally, the optimal bound. For the starting upper bounds, we chose the minimum of the explicit and easily computed bounds (8.1.9) from [5] (a standard classical bound) and (3.1.10) from [4] (originally from [2]), while the lower bounds were obtained by applying those same bounds to the reverse polynomial. Of course, better bounds will lead to fewer steps. The first column lists the degree of the polynomial, while the last one lists the average number of steps the method needed with the standard deviation in parentheses. These results appear quite satisfactory: the number of steps varies little with the degree of the polynomial, and even stopping the method early produced on average a better bound. Moreover, the numbers reported below are quite close because they are averages. Individual cases sometimes exhibit signifcant improvements, without the need to solve a nonlinear equation.

deg(p)	Cauchy	Rahman-Schmeisser	2 nodes	Optimal	No. Iterations
20	1.265	1.145	1.135	1.125	8.2 (4.4)
60	1.258	1.143	1.133	1.127	10.1 (4.7)
100	1.255	1.143	1.128	1.126	10.8 (5.1)

TABLE 1. Comparison of bounds.

We have also run an experiment where we did not try to optimize but, rather, simply tried to reduce the upper bound. In it, only 5 steps were carried out, requiring significantly less work than solving a single nonlinear equation of the form $f(x) = 0$, with f as in (1.1). The average ratios we obtained for $n = 20, 60, 100$ were $1.168, 1.166, 1.166$, respectively. This means that, on average, we obtained results that are better than for the Cauchy radius of p and only slightly worse than its Rahman-Schmeisser improvement, and all this without solving any nonlinear equations.

Extensions

Theorem 8.3.1 in [5] can be applied repeatedly to obtain successively smaller Cauchy radii. Here, we concentrated on the optimization of a single application of the theorem. The optimization can also be repeated for successive applications.

The results we obtained were for a polynomial with real coefficients, and the main stumbling block in extending them to the complex case is that an equation of the form $\varphi(t) = \beta$, with φ as in Lemma 2.1, needs to be solved or at least checked for a solution when the nodes b_j are complex numbers. In other words, we need to at least be able to check if the interior of a weighted k-ellipse is empty or not. This may require too much work, compared to the actual computation of the zeros,

although it may be possible to consider special cases where the number of nodes is small.

4. Appendix

Proof of Lemma 2.1. The possible number of solutions (0, 1, 2, $+\infty$) follows from the fact that the set

$$D = \left\{ z \in \mathbb{C} \ : \ \sum_{j=1}^{m} \gamma_j |z - a_j| \leq \beta \right\} ,$$

with $\gamma_j, \beta > 0$ and $a_j \in \mathbb{C}$ is a bounded convex set. Therefore, the intersection of a line with its boundary, which is the weighted m-ellipse $\partial D = \left\{ z \in \mathbb{C} \ : \ \sum_{j=1}^{m} \gamma_j |z - a_j| = \beta \right\}$, can have no solutions, one solution, two solutions, or a line segment of solutions (when the m-ellipse is degenerate). Here, this line is the real line.

Let us now consider the function φ. It is piecewise linear and its slope increases monotonically from $-\sum_{j=1}^{m} \gamma_j < 0$ on I'_0 to $\sum_{j=1}^{m} \gamma_j > 0$ on I'_m, with the changes in slope occurring at the nodes. To show this, let $t \in I'_\ell$ for $1 \leq \ell \leq m-1$. Then

$$\varphi(t) = \left(\sum_{j=1}^{\ell} \gamma_j - \sum_{j=\ell+1}^{m} \gamma_j \right) t - \left(\sum_{j=1}^{\ell} \gamma_j a_j - \sum_{j=\ell+1}^{m} \gamma_j a_j \right) ,$$

i.e., at the transition of t from I'_ℓ to $I'_{\ell+1}$, the quantity $2\gamma_\ell$ is added to the slope, increasing it. As t increases from $-\infty$ to $+\infty$, the slope will at some point increase from negative to positive while either becoming zero or not. We now turn to the equation $\varphi(t) = \beta$. If the slope of φ is never zero, then there will be a node a_ℓ $(1 \leq \ell \leq m)$ such that the slope is negative to its left and positive to its right, i.e., $\varphi(a_\ell)$ is the global minimum value of φ. If $\varphi(a_\ell) < \beta$, there is no solution, if $\varphi(a_\ell) = \beta$, there is a unique solution, namely, a_ℓ, and if $\varphi(a_\ell) > \beta$, there will be two solutions in different intervals I_{j_1} and I_{j_2}, one on each side of a_ℓ. If the slope becomes zero, then it will be zero on some I'_ℓ $(1 \leq \ell \leq m-1)$, and $\varphi(t)$ will be the minimum value of φ for any $t \in I_\ell$. If this minimum value is equal to β, then all t in I_ℓ will be solutions. Similar conclusions as before hold for no solutions or two solutions. □

References

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

[2] A. Joyal, G. Labelle and Q. I. Rahman, *On the location of zeros of polynomials*, Canad. Math. Bull. **10** (1967), 53–63.

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

[4] G. V. Milovanović, D. S. Mitrinović and Th. M. Rassias, *Topics in Polynomials: Extremal Problems, Inequalities, Zeros*, World Scientific, Singapore, 1999.

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

[6] W. Rudin, *Principles of mathematical analysis*, Third edition. McGraw-Hill Kogakusha, Tokyo, 1976.
[7] J. Sekino, *n-Ellipses and the minimum distance sum problem*, The Monthly **106** (1999), 193–202.

Manuscript received December 5 2017
revised March 28 2018

A. Melman
Department of Applied Mathematics, School of Engineering, Santa Clara University, CA 95053, USA
E-mail address: amelman@scu.edu