THE TEACHING OF MATHEMATICS

2014, Vol. XVII, 1, pp. 39–50

CAUCHY-TYPE INCLUSION AND EXCLUSION REGIONS FOR POLYNOMIAL ZEROS

Aaron Melman

Abstract. A classical result by Cauchy defines a disk containg all the zeros of a polynomial. We derive several related results by using similarity transformations of a polynomial's companion matrix, together with Gershgorin's theorem. We thus show that Cauchy's original result can be seen as but one member of a family of related results.

MathEduc Subject Classification: H35

MSC Subject Classification: 97H30, 12D10, 15A18

Key words and phrases: Cauchy; Gershgorin; zero location; root location; polynomial.

1. Introduction

Polynomials are ubiquitous in mathematics and applications, and very often it is their zeros that matter most. Although it can be difficult to accurately compute them, there are many relatively easy ways to describe their location. One such well-known result by Cauchy ([4], [7, Th.(27,1)]) states that all the zeros of a polynomial $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ with complex coefficients and $a_0 \neq 0$ are contained in a disk in the complex plane, centered at the origin, whose radius is the unique positive root of the real polynomial $f(x) = x^n - |a_{n-1}|x^{n-1} - \cdots - |a_1|x - |a_0|$. An analogous result defines a disk excluding the zeros. This can be proven by straightforward algebraic manipulation, but it can also be explained with simple linear algebra tools which has the advantage of providing a very natural way to obtain this result (see [1]) as well as unifying it with several related ones (see [8]).

Our plan here is to derive several more such results that take the form of disks that contain some or all zeros and disks that contain no zeros. The radii of those disks are determined by the positive roots of polynomials similar to f, showing that Cauchy's result can be seen as but one member of a family of similar results.

Along the way, we will combine basic linear algebra concepts such as eigenvalues, similarity transformations, companion matrices, and Gershgorin disks, as well as a little analytic geometry, to provide an introduction to topics that, although accessible to students, are not always covered in undergraduate classes.

Let us begin by introducing the few definitions and preliminary results we need, along with some notation. We denote the closure of a set G by $\bar{G}$ and its complement by G'. The open disk in the complex plane with center a and radius r

is denoted by $D(a;r)$, while $E(a;r)$ denotes the open exterior of a disk with center a and radius r. We therefore have, e.g., that $D'(a;r) = \overline{E}(a;r)$. We will use $H(a,\delta)$ for the open halfplane, orthogonal to a, of the form $\mathrm{Re}(a)x + \mathrm{Im}(a)y + \delta < 0$.

The *companion matrix* of the aforementioned polynomial p is defined by (see, e.g., [6, p. 146]):

$$C(p) = \begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & \vdots \\ & & \ddots & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{pmatrix}$$

where blank spaces represent zeros. Its characteristic polynomial is p and its eigenvalues are therefore the zeros of p. Those eigenvalues do not change if we apply a similarity transformation to $C(p)$, so that the zeros of p are also the eigenvalues of the matrix $C_x(p) = \Delta_x^{-1} C(p) \Delta_x$, where Δ_x is the diagonal matrix with diagonal $[x^n, x^{n-1}, \ldots, x]$ and $x > 0$. It is a straightforward exercise to show that

$$C_x(p) = \begin{pmatrix} 0 & & & & -a_0/x^{n-1} \\ x & 0 & & & -a_1/x^{n-2} \\ & x & \ddots & & \vdots \\ & & \ddots & 0 & -a_{n-2}/x \\ & & & x & -a_{n-1} \end{pmatrix} .$$

We now use the column version of Gershgorin's theorem ([5], [6, Section 6.1]), which states that all the eigenvalues of a matrix lie in the union of the disks centered at the diagonal elements of the matrix with radii equal to the corresponding deleted column sums. This means that all the eigenvalues of a complex $n \times n$ matrix A with elements a_{ij} are contained in the union

$$\bigcup_{i=1}^{n} \overline{D}(a_{ii};\, K_i'(A)),$$

where $K_i'(A) = \sum_{j=1, j\neq i}^{n} |a_{ji}|$. Moreover, if this union is composed of disjoint sets, then each set contains as many eigenvalues as the number of disks in that set. Applying the theorem to A^T, which has the same eigenvalues as A, leads to an analogous result for the rows.

The matrix $C_x(p)$ has only two different deleted column sums, which makes it easy to compute its Gershgorin column set. Doing so, we find that all the zeros of p lie in the union of two closed disks, D_1 and D_2, with

$$D_1 \equiv \overline{D}(0,x),\ D_2 \equiv \overline{D}(-a_{n-1};\rho(x)),\ \text{and}$$
$$\rho(x) = \frac{|a_{n-2}|}{x} + \frac{|a_{n-3}|}{x^2} + \cdots + \frac{|a_1|}{x^{n-2}} + \frac{|a_0|}{x^{n-1}}.$$

As x increases, D_1 expands, while D_2 shrinks, until D_1 completely envelops D_2 at the precise moment that

$$x = |a_{n-1}| + \rho(x) = |a_{n-1}| + \frac{|a_{n-2}|}{x} + \frac{|a_{n-3}|}{x^2} + \cdots + \frac{|a_1|}{x^{n-2}} + \frac{|a_0|}{x^{n-1}}.$$

All the zeros of p then lie in the single disk $\overline{D}(0;s)$, where s is the unique positive solution of

$$f(x) = x^n - |a_{n-1}|x^{n-1} - \cdots - |a_1|x - |a_0| = 0,$$

which is precisely Cauchy's result. This was shown in [1] and [8], but in [8] this idea was taken one step further by also letting x decrease, so that now D_2 expands to eventually encompass D_1 when

$$x + |a_{n-1}| = \frac{|a_{n-2}|}{x} + \frac{|a_{n-3}|}{x^2} + \cdots + \frac{|a_1|}{x^{n-2}} + \frac{|a_0|}{x^{n-1}}.$$

Therefore, all the zeros of p also lie in $\overline{D}(-a_{n-1}; |a_{n-1}| + t)$, where t is the unique positive solution of

$$g(x) = x^n + |a_{n-1}|x^{n-1} - |a_{n-2}|x^{n-2} - \cdots - |a_1|x - |a_0| = 0.$$

Consequently, all the zeros of p lie in $\overline{D}(0;s) \cap \overline{D}(-a_{n-1}; |a_{n-1}|+t)$. When $a_{n-1} = 0$, the two disks are identical. In that case a different type of inclusion region can be derived.

Before we continue, we note that solving the auxiliary *real* polynomial equations that we will encounter is much easier than computing *all* the (complex) zeros of a polynomial, typically requiring no more than a few iterations with a simple numerical method. It is not the focus of this paper and we will not dwell on it. Although figures will always include the actual zeros of polynomials, this is for illustrative purposes only.

We are now ready to explore many more related Cauchy-like results (we will assume throughout that $a_0 \neq 0$, otherwise the polynomial can be trivially simplified.) In Section 2, we consider a few more zero inclusion regions for special values of the parameter x and then apply the same techniques to the reciprocal polynomial in Section 3 and Section 4, which generates exclusion regions as well, all of which are based on disks. We conclude with a summary of the inclusion and exclusion regions that we have derived. For more information we refer to monographs [9, 10, 12], as well as a recent paper [3].

2. Additional inclusion disks

Throughout this section we will assume that $a_{n-1} \neq 0$ to avoid trivial situations. Previously, we used $C_x(p)$ for special values of the parameter x that generated a particularly simple Gershgorin set, namely, a single disk. However, this simplicity may sometimes come at the price of a large radius. Instead, we can force the two disks comprising the Gershgorin set to be of the same size. The resulting region is still simple and, although the disks are now no longer included in each other, they are each necessarily smaller than either $\overline{D}(0;s)$ or $\overline{D}(-a_{n-1};t)$ mentioned in the introduction. The two disks forming the Gershgorin column set for $C_x(p)$ have the same radius when

$$x = \rho(x) = \frac{|a_{n-2}|}{x} + \frac{|a_{n-3}|}{x^2} + \cdots + \frac{|a_1|}{x^{n-2}} + \frac{|a_0|}{x^{n-1}}.$$

All the zeros of p are then contained in $\overline{D}(0;u)\cup\overline{D}(-a_{n-1};u)$, where u is the unique positive solution of

$$h(x) = x^n - |a_{n-2}|x^{n-2} - \cdots - |a_1|x - |a_0| = 0.$$

The following example compares the three inclusion regions we have found so far in the introduction and the present section.

Example 1. Consider the polynomial $q_1(z) = z^8 + (2+2i)z^7 - 6z^6 + 5z^5 - 3z^4 + 2iz^3 - z^2 + 8z + 2$, for which $a_{n-1} = 2+2i$, $s = 4.46$, $t + |a_{n-1}| = 4.85$, and $u = 2.87$, where s, t, and u are as defined above. The corresponding zero inclusion regions are shown on the left in Figure 1: the two circles in thick line in Figure 1 are the boundaries of $\overline{D}(0;s)$ and $\overline{D}(-a_{n-1},|a_{n-1}|+t)$, whereas the shaded area is the Gershgorin set when both its constituent disks have the same radius u. The black dots represent the zeros of q_1.

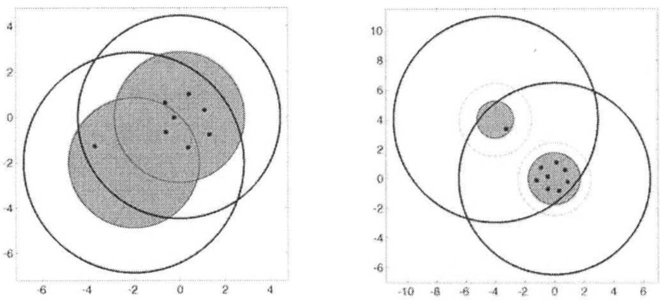

Figure 1. Disjoint zero inclusion disks for q_1 (left) and q_2 (right)

Another natural special value for the parameter x is one for which the two disks in the Gershgorin column set of $C_x(p)$ are disjoint. We recall that in such a case $n-1$ zeros of p will be contained in $\overline{D}(0;x)$, while the remaining zero lies in $\overline{D}(-a_{n-1};\rho(x))$. Alas, this is unfortunately not always possible. If there exists such a value for x, then it needs to be such that $x + \rho(x) < |a_{n-1}|$, which is equivalent to

$$\phi(x) = x^n - |a_{n-1}|x^{n-1} + |a_{n-2}|x^{n-2} + \cdots + |a_1|x + |a_0| < 0. \tag{1}$$

By Descartes' rule of signs, the equation $\phi(x) = 0$ has either two positive roots or none. In the latter case, we cannot separate the disks, but if the equation has two roots x_1 and x_2 with $0 < x_1 < x_2$, then the disks $\overline{D}(0;x)$ and $\overline{D}(-a_{n-1};\rho(x))$ are disjoint for any x such that $x_1 < x < x_2$, which implies that $n-1$ zeros of p lie in $\overline{D}(0;x_1)$ (since it is the union of $n-1$ disks) while the remaining zero lies in $\overline{D}(-a_{n-1};\rho(x_2))$, where $\rho(x_2) = |a_{n-1}| - x_2$. The resulting bounds on the moduli of the zeros are, in fact, a special case of Pellet's theorem ([11], [7, Th. (28,1)]). The following is an example of a polynomial for which a Gershgorin column set with disjoint disks can be obtained.

EXAMPLE 2. For $q_2(z) = z^9 + (4-4i)z^8 - 5iz^7 - 2z^6 + z^5 + 2iz^3 + (4+i)z + 2$, we have that $a_{n-1} = 4 - 4i$, $s = 6.48$, $t + |a_{n-1}| = 6.95$, $x_1 = 1.76$, $x_2 = 4.41$, $|a_{n-1}| - x_2 = 1.25$, and $u = 2.46$, where s, t, and u have the same meaning as before. The disjoint disks with radii x_1 and $|a_{n-1}| - x_2$ can be seen on the right in Figure 1, shaded in dark grey. Since $x_1 < u < x_2$, the two disks of equal radius are also disjoint. Their boundaries are the dashed circles in the figure. The two circles in thick line are, as before, the boundaries of $\overline{D}(0; s)$ and $\overline{D}(-a_{n-1}, |a_{n-1}| + t)$. The black dots are the zeros of q_2.

3. Exclusion disks

As we are about to show, applying the same ideas as before to the reciprocal polynomial leads to similarly simple *exclusion* regions for the zeros. We define the reciprocal polynomial p_r of p as

$$p_r(z) = a_0^{-1} z^n p(z^{-1}) = z^n + \frac{a_1}{a_0} z^{n-1} + \frac{a_2}{a_0} z^{n-2} + \cdots + \frac{a_{n-1}}{a_0} z + \frac{1}{a_0}.$$

Its zeros are the reciprocals of those of p. In this section and the next we assume that $a_1 \neq 0$ to avoid trivial cases.

By the same arguments as before, all zeros of p_r lie in a disk centered at the origin, whose radius s_r is the unique positive root of

$$f_r(x) = x^n - \left|\frac{a_1}{a_0}\right| x^{n-1} - \left|\frac{a_2}{a_0}\right| x^{n-2} - \cdots - \left|\frac{a_{n-1}}{a_0}\right| x - \left|\frac{1}{a_0}\right|. \tag{2}$$

For any zero ζ of p, this implies that $1/\zeta \in \overline{D}(0; s_r)$, or $\zeta \in \overline{E}(0; 1/s_r)$, i.e., $|\zeta| \geq 1/s_r$. The open disk $D(0; 1/s_r)$ is therefore an *exclusion* disk for the zeros of p.

This can also be obtained by observing that $f_r(1/x) = 0$ is equivalent to

$$x^n + |a_{n-1}|x^{n-1} + |a_{n-2}|x^{n-2} + \cdots + |a_1|x - |a_0| = 0,$$

and therefore no zero of p has a modulus less than the unique positive root $1/s_r$ of this equation. This result, also ascribed to Cauchy, is often mentioned alongside the first (classical) one. It is the "analogous result" referred to in the first paragraph of the introduction.

A less well-known exclusion disk is obtained from the counterpart for p_r of the inclusion disk $\overline{D}(-a_{n-1}; |a_{n-1}| + t)$ for p. Applying the exact same reasoning as before implies that all the zeros of p_r are contained in $\overline{D}(-a_1/a_0; |a_1/a_0| + t_r)$, where t_r is the unique positive root of

$$g_r(x) = x^n + \left|\frac{a_1}{a_0}\right| x^{n-1} - \left|\frac{a_2}{a_0}\right| x^{n-2} - \cdots - \left|\frac{a_{n-1}}{a_0}\right| x - \left|\frac{1}{a_0}\right|. \tag{3}$$

For any zero ζ of p this means that $|1/\zeta + a_1/a_0| \leq |a_1/a_0| + t_r$. Since $a_1 \neq 0$, this is the same as saying that ζ belongs to the set

$$S = \{\, z \in \mathbf{C} \;:\; |z + b| \leq \gamma |b||z| \,\},$$

with $b = a_0/a_1$ and $\gamma = |a_1/a_0| + t_r$. Since the boundary of the set S is the pre-image of a circle under the Möbius transformation $z \mapsto 1/z$, it will be either a circle or a line, but to find out exactly what S looks like requires a little more work. It is facilitated by the following technical lemma, which will also be used later on.

Lemma 1. *Let* $S = \{\, z \in \mathbf{C} \,:\, |z+b| \leq \gamma |b||z| \,\}$, *with* $\gamma > 0$ *and* $b \neq 0$.

(1) *If* $\gamma|b| > 1$, *then* $S = \overline{E}\left(\dfrac{b}{\gamma^2|b|^2 - 1} \;;\; \dfrac{\gamma|b|^2}{\gamma^2|b|^2 - 1} \right)$.

(2) *If* $\gamma|b| < 1$, *then* $S = \overline{D}\left(\dfrac{b}{\gamma^2|b|^2 - 1} \;;\; \dfrac{\gamma|b|^2}{1 - \gamma^2|b|^2} \right)$.

(3) *If* $\gamma|b| = 1$, *then* $S = \overline{H}\left(b \;;\; \dfrac{|b|^2}{2} \right)$.

Proof. Setting $z = x + iy$ and $b = b_1 + ib_2$ and squaring both sides of the inequality defining S yields

$$(x+b_1)^2 + (y+b_2)^2 \leq \gamma^2 |b|^2 (x^2+y^2),$$

which is equivalent to

$$(\gamma^2|b|^2 - 1)x^2 + (\gamma^2|b|^2 - 1)y^2 - 2b_1 x - 2b_2 y - |b|^2 \geq 0. \tag{4}$$

If $\gamma|b| > 1$, we divide by $\gamma^2|b|^2 - 1$ and complete the square to obtain

$$\left(x - \frac{b_1}{\gamma^2|b|^2 - 1} \right)^2 + \left(y - \frac{b_2}{\gamma^2|b|^2 - 1} \right)^2 \geqslant \frac{\gamma^2|b|^4}{(\gamma^2|b|^2 - 1)^2},$$

which represents the closed *exterior* of a disk with center $(b_1/(\gamma^2|b|^2 - 1),\, b_2/(\gamma^2|b|^2 - 1))$ and radius $\gamma|b|^2/(\gamma^2|b|^2 - 1)$. If $\gamma|b| < 1$ we obtain analogously the closed *interior* of a disk with center $(b_1/(\gamma^2|b|^2 - 1)\,,\, b_2/(\gamma^2|b|^2 - 1))$ and radius $\gamma|b|^2/(1 - \gamma^2|b|^2)$.

Finally, if $\gamma|b| = 1$ we obtain from (4) the closed halfplane

$$\mathrm{Re}(b)x + \mathrm{Im}(b)y + \frac{|b|^2}{2} \leq 0. \qquad \blacksquare$$

Remarks. (1) The radius of the sets in case (1) and case (2) is bounded from below by $1/\gamma$ and $\gamma|b|^2$, respectively. For case (1), this can be seen from

$$\frac{\gamma|b|^2}{\gamma^2|b|^2 - 1} = \left(\frac{\gamma^2|b|^2}{\gamma^2|b|^2 - 1} \right) \frac{1}{\gamma} > \frac{1}{\gamma}. \tag{5}$$

The bound for case (2) is immediate.

(2) The boundary of the halfplane in case (3) is orthogonal to b.

(3) It follows from the definition of S that it always contains $-b$ and never the origin.

(4) Although not difficult to do, the set S in Lemma 1 is often not derived in a typical undergraduate text, in favor of a simpler but much cruder set.

We recall that we found that all the zeros of p lie in the set S, defined in Lemma 1, with $b = a_0/a_1$ and $\gamma = |a_1/a_0| + t_r$. For these values we obtain

$$\gamma|b| = \left(\left|\frac{a_1}{a_0}\right| + t_r\right)\left|\frac{a_0}{a_1}\right| = 1 + \left|\frac{a_0}{a_1}\right| t_r > 1,$$

which means that we are in case (1) of Lemma 1. The radius there is given by

$$\frac{\gamma|b|^2}{\gamma^2|b|^2 - 1} = \frac{(|a_1/a_0| + t_r)|a_0/a_1|^2}{(|a_1/a_0| + t_r|)^2|a_0/a_1|^2 - 1} = \frac{|a_1/a_0| + t_r}{2|a_1/a_0|t_r + t_r^2}. \tag{6}$$

Since $b\bar{b} = |b|^2$ implies $b = |b|^2/\bar{b}$, we have that $b = |a_0/a_1|^2\overline{a_1/a_0}$, and

$$\frac{b}{\gamma^2|b|^2 - 1} = \frac{|a_0/a_1|^2\,\overline{a_1/a_0}}{(|a_1/a_0| + t_r|)^2|a_0/a_1|^2 - 1} = \frac{\overline{a_1/a_0}}{2|a_1/a_0|t_r + t_r^2}.$$

We have obtained that all the zeros of p are contained in the set

$$\overline{E}\left(\frac{\overline{a_1/a_0}}{2|a_1/a_0|t_r + t_r^2}\,;\,\frac{|a_1/a_0| + t_r}{2|a_1/a_0|t_r + t_r^2}\right), \tag{7}$$

the closed exterior of a disk. In other words, they are excluded from an open disk. To compare the size of this exclusion disk with its better known sibling $D(0 : 1/s_r)$, we first observe from their definitions in (2) and (3) that $f_r(x) \le g_r(x)$ for $x > 0$, implying that $t_r \le s_r$, so that $1/t_r \ge 1/s_r$. Furthermore, from (6) we see that

$$\frac{\gamma|b|^2}{\gamma^2|b|^2 - 1} = \frac{|a_1/a_0| + t_r}{2|a_1/a_0|t_r + t_r^2} = \left(\frac{1 + |a_0/a_1|t_r}{2 + |a_0/a_1|t_r}\right)\frac{1}{t_r},$$

demonstrating that the radius in the set defined by (7) lies between $1/(2t_r)$ and $1/t_r$, from which we conclude that it is at least of the same order of magnitude as $1/s_r$ and potentially larger. The latter can be expected to happen when $|a_1/a_0|$ becomes large since that will make t_r smaller and s_r larger. From (5), a simple lower bound on the radius is given by $(|a_1/a_0| + t_r)^{-1}$.

Finally, we note that the zeros of p are necessarily excluded from the union of these two exclusion disks. When the *inclusion* region for a given value of x is $D_1 \cup D_2$, then the corresponding *exclusion* region is $(D_1 \cup D_2)' = D_1' \cap D_2'$. The following example illustrates the exclusion disks we just derived.

Example 3. For the polynomial $q_3(z) = z^8 + (1+i)z^7 - (2-i)z^5 - 3z^4 + 2iz^3 - z^2 + 8iz + 2$ one obtains $s_r = 4.20$ and $t_r = 1.02$. The corresponding exclusion disks are given by $D(0; 0.24)$ and $D(-0.44i; 0.55)$, respectively. We see that the second disk is markedly larger than the first (classical) exclusion disk because $|a_1|$ is relatively large compared to $|a_0|$. In Figure 2 these disks are shaded in dark and light gray, respectively. The large circle and the arc are the boundary and part of the boundary, respectively, of the inclusion disks from the introduction, which are drawn here to provide a sense of scale. The black dots are the zeros of q_3.

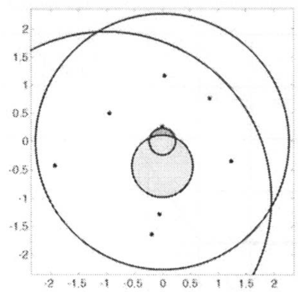

Figure 2. Zero exclusion disks for q_3

4. Additional exclusion regions

We can create more exclusion regions, just as we created more inclusion regions, by considering additional special values for the parameter x in $C_x(p_r)$. Proceeding analogously as in Section 2, we obtain that the Gershgorin column set of $C_x(p_r)$ consists of two disks of the same radius u_r if

$$h_r(u_r) = u_r^n - \left|\frac{a_2}{a_0}\right| u_r^{n-2} - \left|\frac{a_3}{a_0}\right| u_r^{n-3} - \cdots - \left|\frac{a_{n-1}}{a_0}\right| u_r - \left|\frac{1}{a_0}\right| = 0.$$

The corresponding *exclusion* region for the zeros of p depends, as was shown in Lemma 1, on the value of $u_r|a_0/a_1|$ and will be a disk, exterior of a disk, or a half-plane, when that value is greater than, less than, or equal to one, respectively. Following are two examples: Example 4 where the aforementioned exclusion region is a disk and Example 5 where it is the exterior of a disk.

Example 4. Consider $q_4(z) = z^8 + z^7 + (2-i)z^5 - 4z^4 + 2iz^3 - z^2 + (4+2i)z + 5$, for which $s_r = 1.53$, $t_r = 1.04$, and $u_r = 1.21$. On the left in Figure 3 are the corresponding exclusion disks, which are given by $D(0; 1/s_r) \equiv D(0; 0.65)$, shaded in dark gray, and

$$D\left(\frac{\overline{a_1/a_0}}{2|a_1/a_0|t_r + t_r^2} \, ; \, \frac{|a_1/a_0| + t_r}{2|a_1/a_0|t_r + t_r^2}\right) \equiv D(0.27 - 0.14i; 0.66),$$

shaded in light gray, respectively. Both disks' boundaries are also outlined. On the right in Figure 3 we have, shaded in dark gray, the intersection of the exclusion disks $D(0; 1/u_r) \equiv D(0; 0.83)$ and

$$D\left(\frac{\overline{a_1/a_0}}{u_r^2 - |a_1/a_0|^2} \, ; \, \frac{u_r}{u_r^2 - |a_1/a_0|^2}\right) \equiv D(1.20 - 0.60i; 1.82).$$

Because of its size, only the relevant part of this disk is shown in the figure. We note that $u_r|a_0/a_1| = 1.35 > 1$, corresponding to case (1) of Lemma 1, so that the

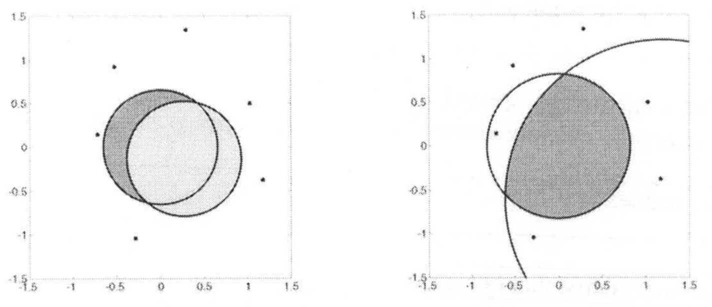

Figure 3. Zero exclusion regions for q_4

corresponding excluded region is a disk. Only six zeros of q_4, which are the black dots, can be seen.

EXAMPLE 5. Let us look at the polynomial $q_5(z) = z^8 + iz^7 + (2-i)z^5 - z^4 + 2iz^3 - z^2 + (1+3i)z + 1$, for which $s_r = 3.62$, $t_r = 1.21$, and $u_r = 1.74$. The corresponding exclusion disks, shown on the left in Figure 4, are given by $D(0; 1/s_r) \equiv D(0; 0.28)$, shaded in dark gray, and

$$D\left(\frac{\overline{a_1/a_0}}{2|a_1/a_0|t_r + t_r^2}\,;\ \frac{|a_1/a_0| + t_r}{2|a_1/a_0|t_r + t_r^2}\right) \equiv D(0.11 - 0.33i; 0.48),$$

shaded in light gray, respectively, with their boundaries outlined. We also have $u_r = 1.74$, so that $u_r|a_0/a_1| = 0.55 < 1$, in which case Lemma 1 implies that the corresponding *excluded* region is the exterior of a closed disk. Accordingly, the right side of Figure 4 shows, shaded in dark gray, the intersection of the exclusion sets $D(0; 1/u_r) \equiv D(0; 0.58)$ and

$$E\left(\frac{\overline{a_1/a_0}}{u_r^2 - |a_1/a_0|^2}\,;\ \frac{u_r}{|a_1/a_0|^2 - u_r^2}\right) \equiv E(-0.14 + 0.43i; 0.25).$$

Only five zeros of q_5 (the black dots) can be seen.

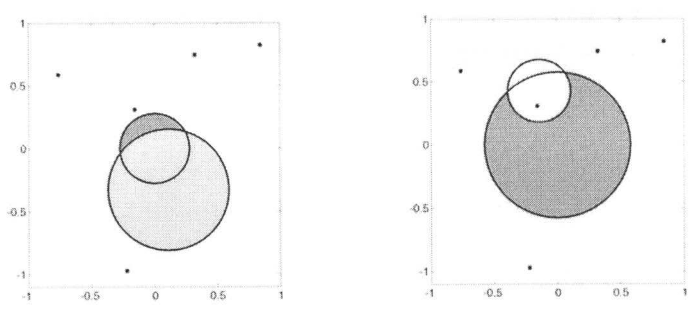

Figure 4. Zero exclusion regions for q_5

As in Section 2, we can try to find values for the parameter x causing the two disks of the Gershgorin column set of $C_x(p_r)$ to be disjoint. The disk centered at the origin then necessarily contains $n-1$ zeros of p_r, while the disk centered at $-a_1/a_0$ contains the remaining one. Again proceeding as before, we find that this is possible for any x such that $y_1 < x < y_2$, where y_1 and y_2 are the unique positive solutions of $\phi_r(x) = 0$, with

$$(8) \qquad \phi_r(x) = x^n - \left|\frac{a_1}{a_0}\right| x^{n-1} + \left|\frac{a_2}{a_0}\right| x^{n-2} + \cdots + \left|\frac{a_{n-1}}{a_0}\right| x + \left|\frac{1}{a_0}\right| .$$

The zeros of p_r are then contained in the disjoint union of $\overline{D}(0; y_1)$ and $\overline{D}(-a_1/a_0; |a_1/a_0| - y_2)$. For the zeros of p, this means that $D(0; 1/y_1)$ is an exclusion disk for $n-1$ of its zeros, while, by Lemma 1,

$$\overline{D}\left(\frac{\overline{a_1/a_0}}{(|a_1/a_0| - y_2)^2 - (a_1/a_0)^2} \, ; \, \frac{|a_1/a_0| - y_2}{(a_1/a_0)^2 - (|a_1/a_0| - y_2)^2} \right)$$
$$= \overline{D}\left(\frac{-\,\overline{a_1/a_0}}{(2|a_1/a_0| - y_2)y_2} \, ; \, \frac{|a_1/a_0| - y_2}{(2|a_1/a_0| - y_2)y_2} \right)$$

is an inclusion disk containing the remaining zero. The latter follows from case (2) in Lemma 1 because $(|a_1/a_0| - y_2)|a_0/a_1| < 1$. The images of two disjoint sets under a continuous transformation remain disjoint, so that this disk must be disjoint from $\overline{E}(0; 1/y_1)$, meaning that it lies entirely in $D(0; 1/y_1)$. This can be verified directly by computing the sum of its radius and midpoint modulus:

$$\frac{|a_1/a_0|}{(2|a_1/a_0| - y_2)y_2} + \frac{|a_1/a_0| - y_2}{(2|a_1/a_0| - y_2)y_2} = \frac{1}{y_2} < \frac{1}{y_1}.$$

We illustrate the above with the following example.

Example 6. Figure 5 shows inclusion and exclusion sets for the zeros of $q_6(z) = z^8 + 3z^7 + (4-i)z^5 - 4z^4 + 2iz^3 - z^2 + (2+3i)z + 1$, for which $s_r = 4.05$, $t_r = 1.38$, $y_1 = 2.33$, and $y_2 = 2.58$. For comparison, the two exclusion disks from Section 3 are shown on the left, namely, $D(0; 0.25)$ and $D(0.17 - 0.25i; 0.42)$, shaded in dark and light gray, respectively. On the right we have the *exclusion* disk $D(0, 1/y_1) \equiv D(0; 0.43)$ for five zeros of q_6, shaded in dark gray, with inside it the *inclusion* disk

$$\overline{D}\left(\frac{-\,\overline{a_1/a_0}}{(2|a_1/a_0| - y_2)y_2} \, ; \, \frac{|a_1/a_0| - y_2}{(2|a_1/a_0| - y_2)y_2} \right) \equiv \overline{D}(-0.17 + 0.25i; 0.09),$$

containing the remaining sixth zero. As expected, $(|a_1/a_0| - y_2)|a_0/a_1| = 0.28 < 1$. Like in all the other examples, the origin lies in an exclusion region, as it should. The black dots are the zeros of q_6, only five of which are shown.

We conclude this section by mentioning that additional interesting values for x can be obtained by manipulating the properties of the inclusion and exclusion

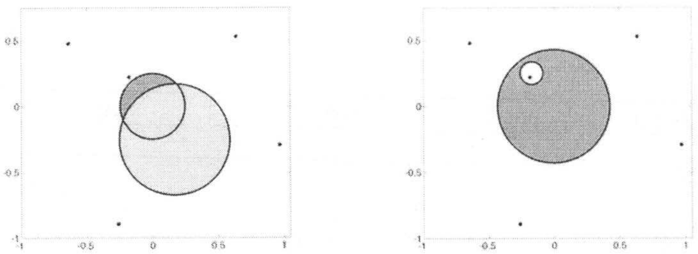

Figure 5. Zero exclusion regions for q_6

regions for the zeros of p, resulting from those for the zeros of p_r. Although we will not pursue this here, one idea could be, e.g., to replace the requirement that the disks containing the zeros of p_r have equal radius by the requirement that the resulting exclusion regions for the zeros of p have equal radius. When doing so, one needs to bear in mind that unions for inclusion regions correspond to intersections for exclusion regions.

In what follows, we have summarized the inclusion and exclusion regions that we derived for all zeros of the polynomial p, not including the regions that can only be obtained when the functions ϕ and ϕ_r, defined by (1) and (8), respectively, have positive roots.

Summary of inclusion and exclusion regions for all zeros of $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$

$$\begin{aligned}
f(x) &= x^n - |a_{n-1}|x^{n-1} - |a_{n-2}|x^{n-2} - \cdots - |a_1|x - |a_0| \\
g(x) &= x^n + |a_{n-1}|x^{n-1} - |a_{n-2}|x^{n-2} - \cdots - |a_1|x - |a_0| \\
h(x) &= x^n - |a_{n-2}|x^{n-2} - |a_{n-3}|x^{n-3} - \cdots - |a_1|x - |a_0| \\
f_r(x) &= x^n - \left|\frac{a_1}{a_0}\right| x^{n-1} - \left|\frac{a_2}{a_0}\right| x^{n-2} - \cdots - \left|\frac{a_{n-1}}{a_0}\right| x - \left|\frac{1}{a_0}\right| \\
g_r(x) &= x^n + \left|\frac{a_1}{a_0}\right| x^{n-1} - \left|\frac{a_2}{a_0}\right| x^{n-2} - \cdots - \left|\frac{a_{n-1}}{a_0}\right| x - \left|\frac{1}{a_0}\right| \\
h_r(x) &= x^n - \left|\frac{a_2}{a_0}\right| x^{n-2} - \left|\frac{a_3}{a_0}\right| x^{n-3} - \cdots - \left|\frac{a_{n-1}}{a_0}\right| x - \left|\frac{1}{a_0}\right|
\end{aligned}$$

Equation	**Root**	**Zero inclusion region**
$f(x) = 0$	s	$\overline{D}(0\,;\ s)$
$g(x) = 0$	t	$\overline{D}(-a_{n-1}\,;\ \|a_{n-1}\| + t)$
$h(x) = 0$	u	$\overline{D}(0\,;\ u) \cup \overline{D}(-a_{n-1}\,;\ u)$

Equation	**Root**	**Zero exclusion region**
$f_r(x) = 0$	s_r	$D\left(0\,;\,\dfrac{1}{s_r}\right)$
$g_r(x) = 0$	t_r	$D\left(\dfrac{\overline{a_1/a_0}}{2\lvert a_1/a_0\rvert t_r + t_r^2}\,;\,\dfrac{\lvert a_1/a_0\rvert + t_r}{2\lvert a_1/a_0\rvert t_r + t_r^2}\right)$
$h_r(x) = 0$	$u_r > \left\lvert\dfrac{a_1}{a_0}\right\rvert$	$D\left(0\,;\,\dfrac{1}{u_r}\right) \cap D\left(\dfrac{\overline{a_1/a_0}}{u_r^2 - \lvert a_1/a_0\rvert^2}\,;\,\dfrac{u_r}{u_r^2 - \lvert a_1/a_0\rvert^2}\right)$
	$u_r < \left\lvert\dfrac{a_1}{a_0}\right\rvert$	$D\left(0\,;\,\dfrac{1}{u_r}\right) \cap E\left(\dfrac{\overline{a_1/a_0}}{u_r^2 - \lvert a_1/a_0\rvert^2}\,;\,\dfrac{u_r}{\lvert a_1/a_0\rvert^2 - u_r^2}\right)$
	$u_r = \left\lvert\dfrac{a_1}{a_0}\right\rvert$	$D\left(0\,;\,\dfrac{1}{u_r}\right) \cap H\left(-\overline{a_1/a_0}\,;\,-\dfrac{1}{2}\right)$

Afterthoughts. There exist many other eigenvalue inclusion sets (see, e.g., [13], which also provides interesting historical background material). The simplest of these is the Brauer set ([2], [6, Theorem 6.4.7], [13, Theorem 2.2]), composed of a union of ovals of Cassini. Applying these other sets to the companion matrix of a polynomial and its reciprocal will then lead to more zero inclusion and exclusion regions, although they quickly become complicated, perhaps too complicated.

REFERENCES

[1] H.E. Bell, *Gershgorin's theorem and the zeros of polynomials*, Amer. Math. Monthly, **72** (1965), 292–295.

[2] A. Brauer, *Limits for the characteristic roots of a matrix II*, Duke Math. J., **14** (1947), 21–26.

[3] C. Carstensen, *A note on matrix methods for location of the zeros of polynomials*, Facta Univ. Ser. Math. Inform. **28**, 4 (2013), 373–378.

[4] 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, p. 86–161. Gauthiers-Villars et fils, Paris, 1891.

[5] S. Gerschgorin, *Über die Abgrenzung der Eigenwerte einer Matrix*, Izv. Akad. Nauk SSSR, Ser. Fiz.-Mat., **6** (1931), 749–754.

[6] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1988.

[7] M. Marden, *Geometry of Polynomials*, Mathematical Surveys, No. 3, American Mathematical Society, Providence, R.I., 1966.

[8] A. Melman, *The twin of a theorem by Cauchy,* Amer. Math. Monthly, **120** (2013), 164–168.

[9] G.V. Milovanović, D.S. Mitrinović, Th.M. Rassias, *Topics in Polynomials: Extremal Problems, Inequalities, Zeros*, World Scientific, Singapore, 1994.

[10] N. Obreškov, *Zeros of Polynomials* (Transl. from Bulgarian), Bulgarian Academic Monographs **7**, Drinov Academic Publishing House, Sofia, 2003.

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

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

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

Department of Applied Mathematics, School of Engineering, Santa Clara University, Santa Clara, CA 95053, U.S.A.

E-mail: amelman@scu.edu