Modified Gershgorin Disks for Companion Matrices*

Aaron Melman[†]

Abstract. All the zeros of a polynomial are contained in the union of Gershgorin disks derived from its companion matrix, a consequence of Gershgorin's theorem. However, this theorem does not exploit the structure of the companion matrix. We will use this structure to obtain smaller zero inclusion regions, thereby providing some nonstandard results to accompany and illustrate this frequently covered topic in numerical and matrix analysis.

Key words. Gershgorin, companion, matrix, polynomial, zero, root

AMS subject classifications. 15A18, 12D10, 15A42, 65H04

DOI. 10.1137/100797667

1. Introduction. A convenient way to obtain information on the location of the zeros of a polynomial is by locating the eigenvalues of its companion matrix. A well-known and easy tool to obtain such information is that of the Gershgorin disks, centered at the diagonal elements of the matrix, whose radii are simple functions of the matrix elements. Gershgorin's theorem states that the union of these disks is guaranteed to contain all the eigenvalues. Several classical bounds on the moduli of a polynomial's zeros such as, e.g., Cauchy's bound, follow easily from applying the Gershgorin disks to the companion matrix of the polynomial.

However, Gershgorin's theorem is almost always used as a black box, i.e., without taking into account any structure the matrix might have, and, as we will see in the next section, more detailed localization techniques can be developed.

We first review Gershgorin's theorem, its proof, and its application to the companion matrix of a polynomial, and then use the structure of this companion matrix to modify and improve the theorem. The modifications lead to the replacement of the disks by the intersection with another disk, with the exterior of a disk, with a plane, or, perhaps more interestingly, with an oval of Cassini. Such ovals also play a role in a different eigenvalue inclusion set, namely, the Brauer set. We thus demonstrate in an elementary way that more than just ordinary disks result from Gershgorin's idea for companion matrices. It allows us to present some nonstandard results to accompany and illustrate the general subject of eigenvalue and polynomial zero inclusion regions, which is often studied in a first numerical or matrix analysis course.

Eigenvalue inclusion regions have a rich and interesting history, starting with the work of Lucien Lévy, who obtained an equivalent formulation of Gershgorin's theorem for real matrices in 1881 [10]. The result was independently rediscovered many times

[14]. An in-depth study with many newer results can be found in [15], while Chapter 6 in [8] provides a good introduction and useful references.

Estimates for the zeros of polynomials usually take the form of bounds on their moduli. Well-known bounds were derived by Cauchy, as mentioned before, but also by Montel and Kojima, to name but a few (see [8, section 5.6]), while more advanced results can be found in, e.g., [11], [12], [13], and references therein. Our focus will be less on bounds and more on specific regions in the complex plane that contain the zeros. These regions will then also generate bounds that fall, generally speaking, into the same category as the aforementioned classical bounds. It is worth mentioning that modern software is able to compute polynomial zeros rapidly and to high precision by computing the eigenvalues of the companion matrix.

In the following section we will define the companion matrix mentioned above and explain how Gershgorin's theorem can be adapted to its specific structure. Two changes based on this adaptation are then carried out in subsequent sections.

2. Companion Matrices and Gershgorin's Theorem. We begin by formally stating Gershgorin's theorem and its application to companion matrices, after which we will informally delve into its proof.

THEOREM 2.1 (see Gershgorin [7]). *All the eigenvalues of the $n \times n$ complex matrix A with elements a_{ij} are located in the union of n disks,*

$$\bigcup_{i=1}^{n} \left\{ z \in \mathbb{C} \,:\, |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \right\} \,.$$

An analogous statement holds for the columns of the matrix, because the spectra of A and A^T are identical.

This means that all the eigenvalues must lie in a union of disks, each centered at a diagonal element of the matrix and having a radius equal to the corresponding deleted row sum, i.e., the sum of the moduli of the off-diagonal elements in the row corresponding to the diagonal element. We concentrate here on the Gershgorin set, but there exist more complicated eigenvalue inclusion sets, such as, e.g., the Brauer set (see [4], [8, Chapter 6] or more recent ones, a good survey of which can be found in [15].

A common way to obtain inclusion regions for the zeros of a polynomial is by finding eigenvalue inclusion regions for a companion matrix of the polynomial, whose eigenvalues are precisely the zeros of the polynomial. This idea is used in, e.g., [1], [3], [5], [11], [12], and [16], to name just a few references from the vast literature on this subject.

For a polynomial $p(z) = z^n + \alpha_{n-1} z^{n-1} + \cdots + \alpha_1 z + \alpha_0$ with complex coefficients and with $\alpha_0 \neq 0$, we will use the standard companion matrix $C(p)$ from [8, p. 146], defined as

$$C(p) = \begin{pmatrix} 0 & 0 & \ldots & 0 & -\alpha_0 \\ 1 & 0 & \ldots & 0 & -\alpha_1 \\ 0 & 1 & \ldots & 0 & -\alpha_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -\alpha_{n-1} \end{pmatrix} \,.$$

Its eigenvalues are the zeros of the polynomial p. This can easily be shown using determinants, but an elegant proof avoiding determinants can also be found in [8, p.

146]. Although we focus on $C(p)$, there exist other companion matrices (see, e.g., [2], [6]).

A direct application of Gershgorin's theorem to $C(p)$ guarantees that all the zeros of the polynomial p must lie in $\Gamma^{(0)}$, where

$$\Gamma^{(0)} = \bigcup_{j=1}^{n} \Gamma_j \,,$$

and

$$\Gamma_1 = \{z \in \mathbb{C} \,:\, |z| \le |\alpha_0|\} \,,$$
$$\Gamma_j = \{z \in \mathbb{C} \,:\, |z| \le 1 + |\alpha_{j-1}|\} \qquad (2 \le j \le n-1) \,,$$
$$\Gamma_n = \{z \in \mathbb{C} \,:\, |z + \alpha_{n-1}| \le 1\} \,.$$

Examples. Figure 2.1 shows $\Gamma^{(0)}$ for the polynomials $p_1(z) = z^4 - 3z^3 + 2iz^2 - 2z + 4$ (on the left) and $p_2(z) = z^4 + 2z^3 + z^2 + iz + 2$ (on the right). The individual disks are outlined inside the Gershgorin sets and the black dots represent the zeros.
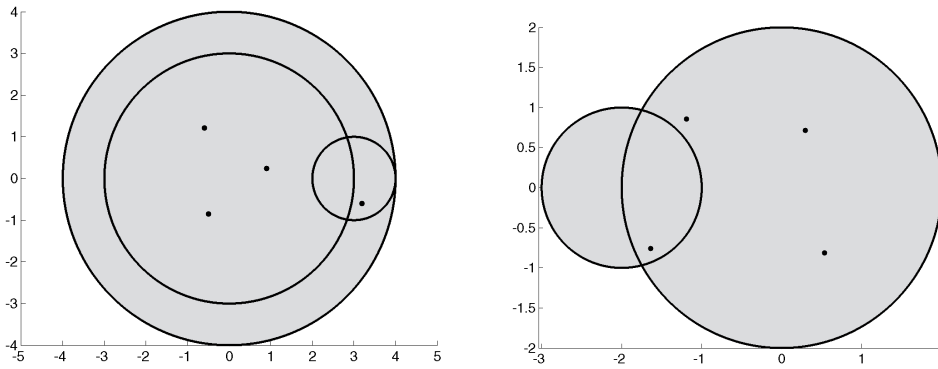
Fig. 2.1 *The $\Gamma^{(0)}$ sets for p_1 (left) and p_2 (right).*

The set $\Gamma^{(0)}$ provides the following upper bound on the modulus of any zero $\tilde{z}$ of the polynomial p:

(2.1)
$$|\tilde{z}| \le |\Gamma^{(0)}| = \max\{|\alpha_0|, 1 + |\alpha_1|, \dots, 1 + |\alpha_{n-1}|\} \,,$$

where $|S|$ denotes the largest modulus of any element in the set S of complex numbers. This bound implies Cauchy's bound, given by $|\tilde{z}| \le 1 + \max\{|\alpha_0|, |\alpha_1|, \dots, |\alpha_{n-1}|\}$.

The superscript (0) in $\Gamma^{(0)}$ denotes the unmodified Gershgorin set, i.e., the set obtained by a "black box" application of Gershgorin's theorem, with $C(p)$ being treated just like any other matrix. Of course, $C(p)$ is very much unlike any other matrix and its structure can be exploited to modify and improve $\Gamma^{(0)}$.

This is a good place to mention that, instead of $C(p)$, one could consider $QC(p)Q^{-1}$ for a nonsingular matrix Q. This new matrix has the same eigenvalues as $C(p)$ but may have smaller eigenvalue inclusion sets. A convenient choice for Q is a diagonal matrix with no zeros on the diagonal (see [8, Chapter 6]). When all the coefficients of the polynomial are nonzero, then an appropriately chosen diagonal matrix leads to

Kojima's bound (see (5.6.45) in [8]):

$$(2.2) \qquad |\tilde{z}| \leq \max \left\{ |\alpha_0|, 2\left|\frac{\alpha_1}{\alpha_2}\right|, \ldots, 2\left|\frac{\alpha_{n-2}}{\alpha_{n-1}}\right|, 2\,|\alpha_{n-1}| \right\} \; .$$

Although beyond our scope here, diagonal similarities can also be applied to the sets we are about to derive.

Our goal is now to find relatively simple modifications that can, at least sometimes, make a significant difference. To explain our strategy for improving $\Gamma^{(0)}$, we first need to understand the proof of Theorem 2.1 for a general complex matrix.

Proof of Theorem 1. Following the same standard procedure as in [8], assume that λ is an eigenvalue of a complex $n \times n$ matrix A with corresponding eigenvector x, i.e., $Ax = \lambda x$. Since x is an eigenvector, it has at least one nonzero component. Define x_ρ as a component of x with the largest absolute value, so that $|x_\rho| \geq |x_i|$ for all $i = 1, 2, \ldots, n$ and $x_\rho \neq 0$. Because $(Ax)_\rho = (\lambda x)_\rho$, we have

$$\lambda x_\rho = a_{\rho\rho} x_\rho + \sum_{\substack{j=1 \\ j \neq \rho}}^{n} a_{\rho j} x_j, \text{ from which it follows that } (\lambda - a_{\rho\rho}) x_\rho = \sum_{\substack{j=1 \\ j \neq \rho}}^{n} a_{\rho j} x_j \; .$$

Taking absolute values on both sides, using the triangle inequality, and dividing by $|x_\rho|$ yields

$$|\lambda - a_{\rho\rho}| \leq \sum_{\substack{j=1 \\ j \neq \rho}}^{n} |a_{\rho j}| \frac{|x_j|}{|x_\rho|} \leq \sum_{\substack{j=1 \\ j \neq \rho}}^{n} |a_{\rho j}| \; ,$$

because $|x_j|/|x_\rho| \leq 1$ for all $j \neq \rho$, i.e., λ must lie in a disk with center $a_{\rho\rho}$. Without knowing the eigenvectors, we do not know which ρ each eigenvalue corresponds to, so we must take the union of all such disks to obtain a region that is guaranteed to contain all eigenvalues, and that concludes the proof.

However, any structure the matrix A might exhibit is lost in the proof's uniform treatment of the elements of A and its premature use of absolute values, which wipes out many connections between the components of x. We will therefore revisit the proof while explicitly using the form of the companion matrix. Specifically, $C(p)x = \lambda x$ means that

$$(2.3) \qquad\qquad \lambda x_1 = -\alpha_0 x_n \; ,$$
$$(2.4) \qquad\qquad \lambda x_2 = x_1 - \alpha_1 x_n \; ,$$
$$(2.5) \qquad\qquad \lambda x_j = x_{j-1} - \alpha_{j-1} x_n \qquad (3 \leq j \leq n-2) \; ,$$
$$(2.6) \qquad\qquad \lambda x_{n-1} = x_{n-2} - \alpha_{n-2} x_n \; ,$$
$$(2.7) \qquad\qquad \lambda x_n = x_{n-1} - \alpha_{n-1} x_n \; ,$$

and it is this set of equations that we will use.

The modifications to the black box Gershgorin set $\Gamma^{(0)}$ will be carried out in two stages, modifying Γ_1 and Γ_2, respectively, with further possible modifications briefly mentioned later on. Each stage produces a new zero inclusion set $\Gamma^{(j)}$ ($j = 1, 2$) and these sets satisfy $\Gamma^{(2)} \subseteq \Gamma^{(1)} \subseteq \Gamma^{(0)}$. We will also compute an upper bound on the moduli of the zeros for each of the modified sets. The first modification is derived in the next section.

3. $\Gamma^{(1)}$: The First Modified Set. We now proceed as in the proof above, namely, we consider a component of x with the largest magnitude and take absolute values to arrive at inequalities that must be satisfied by λ. However, contrary to what we did before, we will first rewrite our equations by exploiting their specific structure and only then take absolute values.

Our first modification is very modest, namely, we just use the first equation (2.3) to eliminate x_n in the last equation (2.7), which turns it into $\lambda (\lambda + \alpha_{n-1}) x_1 = -\alpha_0 x_{n-1}$. This means that when x_1 has the largest magnitude of all the components of x, we obtain, in addition to $|\lambda| \le |\alpha_0|$, another inequality that must be satisfied by λ, namely, $|\lambda| \, |\lambda + \alpha_{n-1}| \le |\alpha_0|$.

We define

$$K_1 = \{ z \in \mathbb{C} \, : \, |z| \, |z + \alpha_{n-1}| \le |\alpha_0| \} \quad \text{and} \quad \Omega_1 = \Gamma_1 \cap K_1 \, .$$

The boundary of K_1 is a quartic curve known as an *oval of Cassini* with foci $-\alpha_{n-1}$ and the origin. It consists of either one or two loops. Ovals of Cassini also appear in a different and slightly more complicated eigenvalue inclusion set, namely, the Brauer set (see [4], [8, p. 380]), and it is interesting that we should encounter such an oval here as well.

We can therefore replace Γ_1 by Ω_1 in the Gershgorin set, which proves the following theorem.

THEOREM 3.1. *All the zeros of the polynomial $p(z) = z^n + \alpha_{n-1} z^{n-1} + \cdots + \alpha_1 z + \alpha_0$ with complex coefficients and with $\alpha_0 \ne 0$ can be found in $\Gamma^{(1)}$, where*

$$\Gamma^{(1)} = \Omega_1 \cup \left(\bigcup_{j=2}^{n} \Gamma_j \right) \, .$$

Since $\Omega_1 \subseteq \Gamma_1$, we clearly have that $\Gamma^{(1)} \subseteq \Gamma^{(0)}$, i.e., the new inclusion set is potentially smaller than $\Gamma^{(0)}$. We can expect this to happen when $|\alpha_0|$ is large enough for Γ_1 to dominate $\Gamma^{(0)}$. In that case, the intersection with the oval of Cassini will cut Γ_1 down to size, as in the following examples.

Examples. Consider the polynomials p_1 and p_2, defined as

$$p_1(z) = z^4 - 3z^3 + 2z^2 + 2z + 5 \qquad \text{and} \qquad p_2(z) = z^4 + 6z^3 + 2z^2 + 4z + 8 \, .$$

Their corresponding zero inclusion sets with the Gershgorin circles and ovals of Cassini are illustrated in Figure 3.1. The $\Gamma^{(0)}$ and $\Gamma^{(1)}$ sets are the areas shaded in dark and light gray, respectively. Those for p_1 are on the left, while those for p_2 are on the right. The Gershgorin circles are drawn slightly thicker than the ovals of Cassini. For p_1, the area of $\Gamma^{(1)}$ is approximately 40% of the area of $\Gamma^{(0)}$, whereas for p_2, this percentage is approximately 42%.

The Set K_1 and Ovals of Cassini. We need the basic properties of an oval of Cassini to describe the set Ω_1, and we summarize them here. They can be found in, e.g., [9, pp. 153–155].

The oval described by $|z| \, |z + \alpha_{n-1}| = |\alpha_0|$ consists of one loop containing both foci if $2\sqrt{|\alpha_0|} > |\alpha_{n-1}|$ and of two loops, each containing one of the foci, if $2\sqrt{|\alpha_0|} \le |\alpha_{n-1}|$. The oval is symmetric with respect to the line through its two focal points, and with respect to the perpendicular line that goes through the center of the segment connecting the foci. The set K_1 consists of the closed interior of the loop or loops.
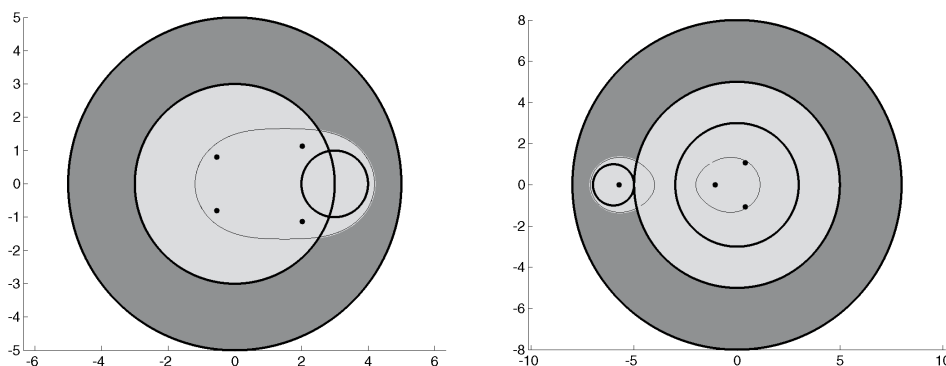
Fig. 3.1 *The $\Gamma^{(0)}$ and $\Gamma^{(1)}$ sets for the polynomials p_1 (left) and p_2 (right).*
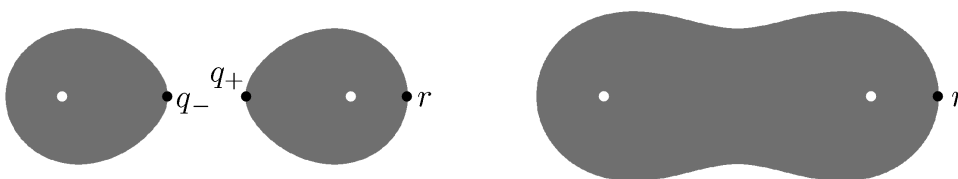
Fig. 3.2 *The set K_1 for $2\sqrt{|\alpha_0|} < |\alpha_{n-1}|$ (left) and $2\sqrt{|\alpha_0|} > |\alpha_{n-1}|$ (right).*

Figure 3.2 shows the set K_1 for $2\sqrt{|\alpha_0|} < |\alpha_{n-1}|$ on the left and $2\sqrt{|\alpha_0|} > |\alpha_{n-1}|$ on the right. The white dots are the foci: the origin on the left and $-\alpha_{n-1}$ on the right. The distances q_-, q_+, and r from the origin, which are defined below, are indicated by black dots.

The point on the boundary curve of K_1 that is furthest away from the origin lies at a distance of

$$r = \frac{1}{2}\left(|\alpha_{n-1}| + \sqrt{|\alpha_{n-1}|^2 + 4|\alpha_0|}\right)$$

from the origin in the direction of $-\alpha_{n-1}$ along the line connecting the foci. When there are two loops, they intersect this line between the origin and $-\alpha_{n-1}$ at two points, which lie at distances

$$q_\pm = \frac{1}{2}\left(|\alpha_{n-1}| \pm \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|}\right)$$

from the origin in the direction of $-\alpha_{n-1}$. The entire oval, whether it consists of one or two loops, is contained in the disk with radius r and centered at the origin. When there are two loops, then each loop is contained in a disk centered at the corresponding focus with radius q_-.

The Set $\Omega_1 = \Gamma_1 \cap K_1$. Let us now have a closer look at the set Ω_1, which was the intersection of the disk Γ_1 with the oval K_1. There are several possible scenarios. Observe, e.g., that the intersection is the entire oval when $|\alpha_0| \geq r$ (see the leftmost figure of Case Ia in Figure 3.3), or part of it when $|\alpha_0| < r$ (see the leftmost figure of
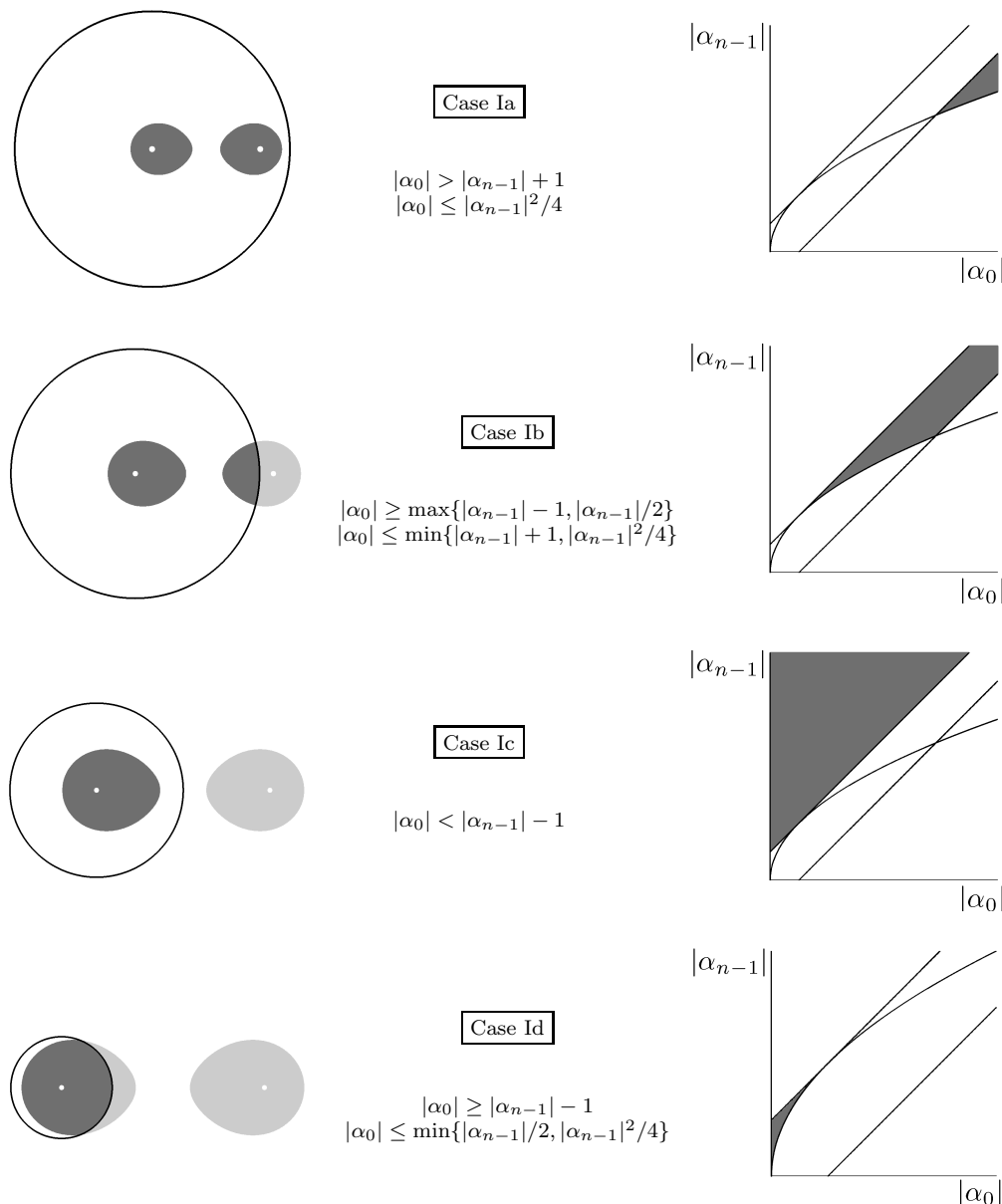
Fig. 3.3 *Scenarios for Ω_1 when $2\sqrt{|\alpha_0|} \leq |\alpha_{n-1}|$.*

Case Ib in Figure 3.3). In the latter case, the part of the oval that is included depends on the values of q_- and q_+. However, q_-, q_+, and r are all functions of $|\alpha_0|$ and $|\alpha_{n-1}|$, which makes the aforementioned conditions implicit and a little difficult to visualize. They have been made explicit in terms of $|\alpha_0|$ and $|\alpha_{n-1}|$ in Figures 3.3 and 3.4: for each case the figure on the left represents the particular scenario, the figure on the right shows the set of values in the $(|\alpha_0|, |\alpha_{n-1}|)$-plane for which it occurs, and in the middle is given the algebraic description of this same set, the derivation of which is
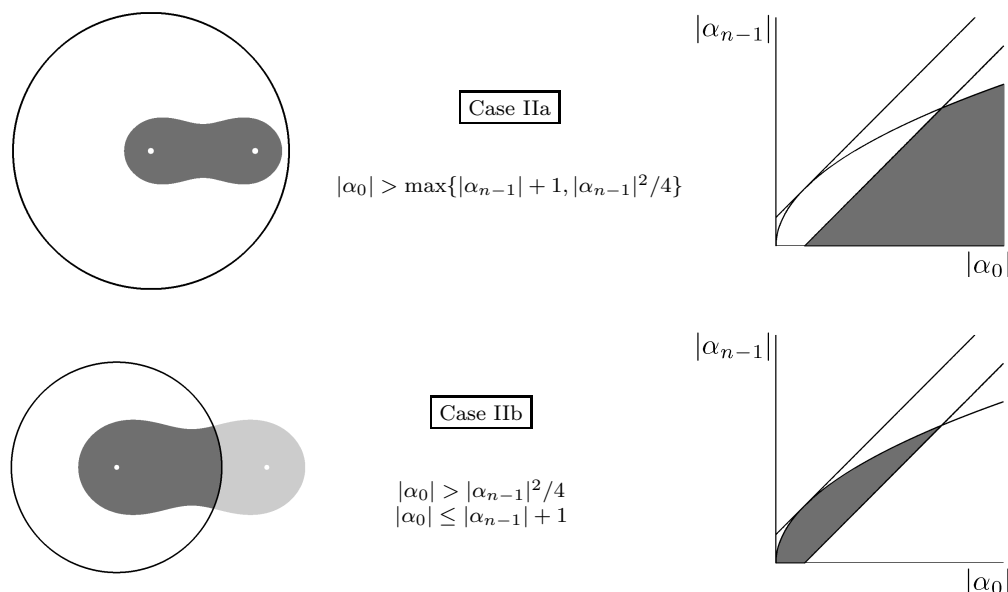
Fig. 3.4   *Scenarios for Ω_1 when $2\sqrt{|\alpha_0|} > |\alpha_{n-1}|$.*

explained in detail in Appendix A. The relative sizes of these sets give a sense of how likely the situations they describe are to occur. Eventually, they will be merged into three regions.

Bounds. The set $\Gamma^{(1)}$ can be used to obtain a bound on the largest modulus of the zeros of a polynomial just like $\Gamma^{(0)}$. The result is presented in the following theorem.

THEOREM 3.2.   *Any zero $\tilde{z}$ of the polynomial $p(z) = z^n + \alpha_{n-1}z^{n-1} + \cdots + \alpha_1 z + \alpha_0$ with complex coefficients and with $\alpha_0 \neq 0$ satisfies $|\tilde{z}| \leq |\Gamma^{(1)}|$, where*

$$|\Gamma^{(1)}| = \max\left\{\gamma, 1 + |\alpha_1|, \ldots, 1 + |\alpha_{n-1}|\right\},$$

and

$$\gamma = \begin{cases} \dfrac{1}{2}\left(|\alpha_{n-1}| + \sqrt{|\alpha_{n-1}|^2 + 4|\alpha_0|}\right) & (|\alpha_0| > 1 + |\alpha_{n-1}|), \\[2mm] 0 & (|\alpha_0| \leq 1 + |\alpha_{n-1}|). \end{cases}$$

Proof. We have that

(3.1) $$|\Gamma^{(1)}| = \max\{|\Omega_1|, 1 + |\alpha_1|, \ldots, 1 + |\alpha_{n-1}|\},$$

and, clearly, $|\Gamma^{(1)}|$ is an upper bound on the modulus of any zero of p.

From the bounds on the moduli of the elements in Ω_1 that were established in the description of the subcases of Case I and Case II, we conclude that in Case Ia and Case IIa, $|\Omega_1| = r$; in Case Ib, Case Id, and Case IIb, $|\Omega_1| = |\alpha_0|$; and in Case Ic, $|\Omega_1| = q_-$. See Appendix A. Combining the different regions from the third column in Figures 3.3
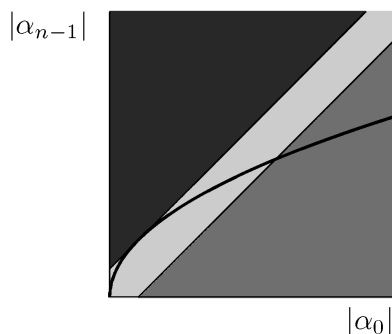
Fig. 3.5 *The regions corresponding to q_- (darker gray), r (lighter gray), and $|\alpha_0|$ (lightest gray).*

and 3.4, these three possibilities correspond to the regions of the $(|\alpha_0|, |\alpha_{n-1}|)$-plane in Figure 3.5 shaded in lighter gray, lightest gray, and darker gray, respectively. This means that

(3.2)
$$
|\Omega_1| = \begin{cases}
\frac{1}{2}\left(|\alpha_{n-1}| + \sqrt{|\alpha_{n-1}|^2 + 4|\alpha_0|}\right) & (|\alpha_0| > 1 + |\alpha_{n-1}|), \\[2mm]
|\alpha_0| & (|\alpha_{n-1}| - 1 \le |\alpha_0| \le |\alpha_{n-1}| + 1), \\[2mm]
\frac{1}{2}\left(|\alpha_{n-1}| - \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|}\right) & (|\alpha_0| < |\alpha_{n-1}| - 1).
\end{cases}
$$

However, for the computation of the maximum value the only important distinction is between $|\alpha_0| > |\alpha_{n-1}| + 1$ and $|\alpha_0| \le 1 + |\alpha_{n-1}|$. This is so because in the former case we have that

$$
\max\left\{|\Omega_1|, 1 + |\alpha_{n-1}|\right\} = \max\left\{\frac{1}{2}\left(|\alpha_{n-1}| + \sqrt{|\alpha_{n-1}|^2 + 4|\alpha_0|}\right), 1 + |\alpha_{n-1}|\right\},
$$

and in the latter case, $|\Omega_1| = |\alpha_0|$ or $|\Omega_1| = q_- = \min\{|\alpha_0|, q_-\}$, so that $|\Omega_1| \le |\alpha_0| \le 1 + |\alpha_{n-1}|$. In this case we obtain

$$
\max\{|\Omega_1|, 1 + |\alpha_{n-1}|\} = 1 + |\alpha_{n-1}| = \max\{0, 1 + |\alpha_{n-1}|\}.
$$

This completes the proof. □

Figure 3.5 shows, at a glance, the situation a given $(|\alpha_0|, |\alpha_{n-1}|)$-pair corresponds to: above the curve there are two loops, below it, just one. If the pair corresponds, e.g., to a point in the area shaded in darker gray, then α_0 is such that the intersection with the disk consists of the entire loop containing the origin, etc.

Note also that an alternative way to express $|\Gamma^{(1)}|$ is

$$
|\Gamma^{(1)}| = \max\left\{\min\left\{|\alpha_0|, \frac{1}{2}\left(|\alpha_{n-1}| + \sqrt{|\alpha_{n-1}|^2 + 4|\alpha_0|}\right)\right\}, 1 + |\alpha_1|, \ldots, 1 + |\alpha_{n-1}|\right\},
$$

because

$$
\min\left\{|\alpha_0|, \frac{1}{2}\left(|\alpha_{n-1}| + \sqrt{|\alpha_{n-1}|^2 + 4|\alpha_0|}\right)\right\}
$$

is not greater than $1 + |\alpha_{n-1}|$ when $|\alpha_0| \leq 1 + |\alpha_{n-1}|$ and when $|\alpha_0| > |\alpha_{n-1}| + 1$, we are in Case Ia or Case IIa and $|\Omega_1|$ is clearly determined by the minimum of $|\alpha_0|$ and r.

Since $\Omega_1 \subseteq \Gamma_1$, we have that $|\Gamma^{(1)}| \leq |\Gamma^{(0)}|$. As we mentioned before, the effect of our modification can be significant when $|\alpha_0|$ is large enough to dominate the computation of $\Gamma^{(0)}$. Let us illustrate this with a few examples.

Examples. Table 3.1 lists the values of $|\Gamma^{(0)}|$ and $|\Gamma^{(1)}|$, along with the modulus of the largest zero $|\tilde{z}_{max}|$, for five polynomials that are identical except for the constant term α_0. As $|\alpha_0|$ increases, the difference between the bounds becomes more pronounced. Here, Kojima's bound is the same as $|\Gamma^{(0)}|$, except for the top polynomial, for which it is equal to 4.

Table 3.1   *Comparison of $|\Gamma^{(0)}|$ and $|\Gamma^{(1)}|$.*

p	$|\Gamma^{(0)}|$	$|\Gamma^{(1)}|$	$|\tilde{z}_{max}|$
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 3z + 2$	5	5	2.7838
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 3z + 6$	6	5	2.7394
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 3z + 10$	10	5	2.6896
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 3z + 20$	20	5.5826	2.5274
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 3z + 40$	40	7.4031	2.3876

It is also sometimes possible to compute a lower bound on the moduli of the zeros, as in Case Ia and Case Ib, although we will not go into detail here.

4. $\Gamma^{(2)}$: The Second Modified Set. In this section, we continue what we started in the previous one. There, we used (2.3) to eliminate x_n in (2.7). Here, we will, in addition, use it to replace x_n by $-\lambda x_1/\alpha_0$ in (2.4). This yields

$$(4.1) \qquad \lambda x_2 = \left(1 + \frac{\alpha_1}{\alpha_0}\lambda\right) x_1 \ .$$

Now assume that $|x_2| \geq |x_j|$ for all j. Then from (4.1) we have

$$|\lambda||x_2| = \left|1 + \frac{\alpha_1}{\alpha_0}\lambda\right| |x_1| \ ,$$

which, after dividing both sides by $|x_2|$, becomes

$$(4.2) \qquad |\lambda| \leq \left|1 + \frac{\alpha_1}{\alpha_0}\lambda\right| \ .$$

Defining

$$K_2 = \left\{ z \in \mathbb{C} \ : \ |z| \leq \left|1 + \frac{\alpha_1}{\alpha_0}z\right| \right\} \ ,$$

we have that $\lambda \in K_2$. However, since we assumed that $|x_2| \geq |x_j|$ for all j, we also have from (2.4) that $\lambda \in \Gamma_2$, which was a disk centered at the origin with radius $1 + |\alpha_2|$. We conclude that, in this case, $\lambda \in \Gamma_2 \cap K_2$. Defining

$$\Omega_2 = \Gamma_2 \cap K_2 \ ,$$

this means that we have proved the following theorem.

THEOREM 4.1. *All the zeros of the polynomial* $p(z) = z^n + \alpha_{n-1} z^{n-1} + \cdots + \alpha_1 z + \alpha_0$ *with complex coefficients and with* $\alpha_0 \neq 0$ *can be found in* $\Gamma^{(2)}$, *where*

$$\Gamma^{(2)} = \Omega_1 \cup \Omega_2 \cup \left(\bigcup_{j=3}^{n} \Gamma_j \right) .$$

Since $\Omega_2 \subseteq \Gamma_2$, we have that $\Gamma^{(2)} \subseteq \Gamma^{(1)}$.

Examples. Figure 4.1 shows $\Gamma^{(2)}$ for the polynomials p_1 (on the left) and p_2 (on the right) defined by

$$p_1(z) = z^7 + 3z^6 - 4z^5 + z^4 + 6z^3 - 5z^2 + 8z + 8 \quad \text{and} \quad p_2(z) = z^5 + 2z^4 + iz^3 - 2z^2 + 6z + 8 .$$

For p_1, the zero inclusion regions $\Gamma^{(0)}$ and $\Gamma^{(1)}$ are identical. $\Gamma^{(2)}$ is shaded in lighter gray and is superimposed on on $\Gamma^{(0)}$, which is shaded in darker gray with its invisible part hiding behind $\Gamma^{(2)}$. The zeros of p_1 are indicated by the black dots. For p_2, $\Gamma^{(0)}$ is shaded in darker gray, $\Gamma^{(1)}$ in lighter gray, and $\Gamma^{(2)}$ in darkest gray. The white dots are the zeros of p_2.

For p_1, $\Gamma^{(1)}$ is identical to $\Gamma^{(0)}$ and the area of $\Gamma^{(2)}$ is approximately 82% of the area of $\Gamma^{(0)}$. For p_2 the areas of $\Gamma^{(1)}$ and $\Gamma^{(2)}$ are approximately 77% and 19% of the area of $\Gamma^{(0)}$, respectively.
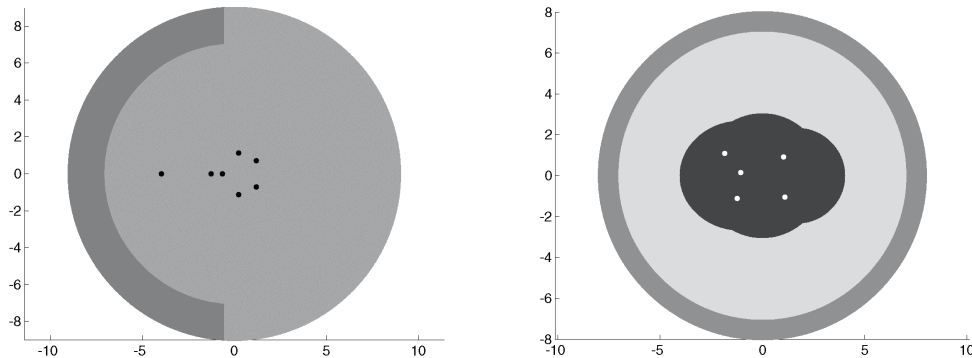
Fig. 4.1 *The* $\Gamma^{(2)}$ *sets for* p_1 *(left) and* p_2 *(right).*

The Set K_2. To understand what Ω_2 looks like, we first need to know what K_2 looks like, and that is what the next lemma is about.

LEMMA 4.2. *The set K_2 has the following properties.*

(1) *The set K_2 contains the origin.*

(2) *If $\alpha_1 = 0$, then K_2 is the closed unit disk.*

(3) *If $|\alpha_0| > |\alpha_1| > 0$, then K_2 is a closed disk with radius $|\alpha_0|^2 / \left(|\alpha_0|^2 - |\alpha_1|^2 \right)$ and center*

$$\frac{|\alpha_1|^2}{|\alpha_0|^2 - |\alpha_1|^2} \left(\frac{\alpha_0}{\alpha_1} \right) .$$

(4) *If $0 < |\alpha_0| < |\alpha_1|$, then K_2 is the closed exterior of a disk with radius $|\alpha_0|^2 / \left(|\alpha_1|^2 - |\alpha_0|^2 \right)$ and center*

$$\frac{|\alpha_1|^2}{|\alpha_0|^2 - |\alpha_1|^2} \left(\frac{\alpha_0}{\alpha_1} \right) .$$

(5) *If $0 < |\alpha_0| = |\alpha_1|$, then K_2 is the closed half-plane*

$$\mathrm{Re}\left(\frac{\alpha_0}{\alpha_1}\right) x + \mathrm{Im}\left(\frac{\alpha_0}{\alpha_1}\right) y + \frac{1}{2} \geq 0 ,$$

which contains α_0/α_1.

Proof. If $\alpha_1 = 0$, then statements (1) and (2) follow immediately from the definition of K_2. If $\alpha_1 \neq 0$, then the inequality in the definition of K_2 can be rewritten as

$$\left|\frac{\alpha_0}{\alpha_1}\right| |z| \leq \left|\frac{\alpha_0}{\alpha_1} + z\right| ,$$

which is of the form $|a||z| \leq |z + a|$, with $a = \alpha_0/\alpha_1$. Setting $z = x + iy$ and $a = a_1 + ia_2$ in this inequality and squaring both sides yields

$$|a|^2 \left(x^2 + y^2\right) \leq (x + a_1)^2 + (y + a_2)^2 ,$$

which can be rewritten as

$$(4.3) \qquad \left(|a|^2 - 1\right) x^2 + \left(|a|^2 - 1\right) y^2 - 2a_1 x - 2a_2 y - |a|^2 \leq 0 .$$

When $|a| > 1$ we can divide by $|a|^2 - 1$, preserving the direction of the inequality, and then complete the square to obtain

$$(4.4) \qquad \left(x - \frac{a_1}{|a|^2 - 1}\right)^2 + \left(y - \frac{a_2}{|a|^2 - 1}\right)^2 \leq \frac{|a|^4}{\left(|a|^2 - 1\right)^2} .$$

This represents the closed interior of a disk with center $(a_1/(|a|^2 - 1)$, $a_2/(|a|^2 - 1))$ and radius $|a|^2/(|a|^2 - 1)$, which, with the definition of a and after a little algebra, yields precisely statement (3).

When $|a| < 1$ we obtain, analogously, the closed exterior of a disk (the inequality's direction is reversed) with the same expression for the center and with radius $|a|^2/(1 - |a|^2)$, which leads directly to statement (4).

When $|a| = 1$, then we obtain from (4.3) that

$$a_1 x + a_2 y + \frac{1}{2} \geq 0 ,$$

which is a half-plane containing the point a, and this completes the proof. $\square$

The Set $\Omega_2 = \Gamma_2 \cap K_2$. With the previous lemma we now know that $\Gamma_2 \cap K_2$ is the intersection of two disks, or of a disk with the exterior of a disk, or of a disk with a half-plane. Figure 4.2 shows a few possible situations for $\Gamma_2 \cap K_2$, which is shaded in gray. The set Γ_2 is the larger of the two disks in the top left and middle subfigures, and in the bottom left subfigure. It is the left disk in the top right subfigure and the only disk in the bottom right subfigure.
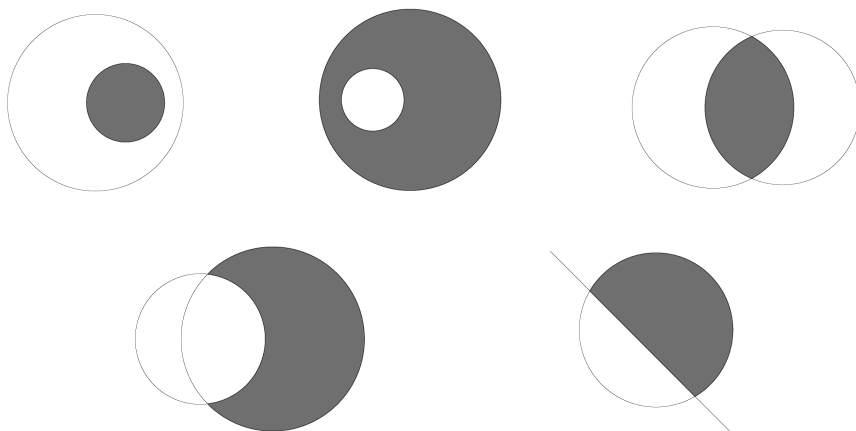
Fig. 4.2 *Examples of the set $\Gamma_2 \cap K_2$ (shaded areas).*

The set Ω_2 is never empty. This is shown in the next theorem, which also computes $|\Omega_2|$.

THEOREM 4.3. *The set Ω_2 is nonempty. Moreover,*

$$|\Omega_2| = \begin{cases} 1 + \dfrac{|\alpha_1|}{|\alpha_0| - |\alpha_1|} & (|\alpha_0| > 1 + |\alpha_1|), \\[3mm] 1 + |\alpha_1| & (|\alpha_0| \le 1 + |\alpha_1|) . \end{cases}$$

Proof. When $\alpha_1 = 0$, K_2 is the unit disk so that its intersection with Γ_2 is the very same nonempty unit disk. Consequently,

$$|\Omega_2| = 1 = 1 + |\alpha_1| = 1 + \frac{|\alpha_1|}{|\alpha_0| - |\alpha_1|} ,$$

which proves the theorem when $\alpha_1 = 0$. When $\alpha_1 \ne 0$, and with $a = \alpha_0/\alpha_1$, assume first that $|a| > 1$, i.e., $|\alpha_0| > |\alpha_1|$. From the previous lemma, we know that in this case, K_2 is a closed disk and $\Gamma_2 \cap K_2 \ne \emptyset$ because the distance d between the centers of Γ_2 and K_2 is not greater than the sum of the radii. To prove this, we need to show that

$$d = \frac{|a|}{|a|^2 - 1} \le \frac{|a|^2}{|a|^2 - 1} + 1 + |\alpha_1| .$$

Multiplying both sides of the inequality by $|a|^2 - 1$ turns this condition into

$$(2 + |\alpha_1|)|a|^2 - |a| - (1 + |\alpha_1|) \ge 0 ,$$

and this inequality is satisfied for $|a| > 1$. The intersection must therefore be nonempty. This scenario corresponds to the top left or the top right subfigure in Figure 4.2. Because we have the intersection of two disks, one of which is centered at the origin, it follows that

(4.5) $$|\Gamma_2 \cap K_2| = \min\left\{ 1 + |\alpha_1|, d + \frac{|a|^2}{|a|^2 - 1} \right\} .$$

Since

$$d + \frac{|a|^2}{|a|^2 - 1} = \frac{|a|}{|a|^2 - 1} + \frac{|a|^2}{|a|^2 - 1} = \frac{|a|(|a| + 1)}{(|a| - 1)(|a| + 1)} = \frac{|a|}{|a| - 1} ,$$

and recalling that $a = \alpha_0/\alpha_1$, (4.5) can be rewritten as

$$|\Gamma_2 \cap K_2| = \min \left\{ 1 + |\alpha_1|, \frac{|\alpha_0|}{|\alpha_0| - |\alpha_1|} \right\} = 1 + \min \left\{ |\alpha_1|, \frac{|\alpha_1|}{|\alpha_0| - |\alpha_1|} \right\} .$$

The minimum value is therefore given by $1 + |\alpha_1|$ when $|\alpha_0| - |\alpha_1| \leq 1$ and by $1 + (|\alpha_1|/(|\alpha_0| - |\alpha_1|))$ otherwise. This proves the theorem for $|\alpha_0| > |\alpha_1|$.

Next, assume that $|\alpha_0| < |\alpha_1|$, i.e., $|a| < 1$. The set $\Gamma_2 \cap K_2$ is nonempty in this case as well, because the radius of the boundary of K_2 is not greater than the sum of the distance d between the centers of Γ_2 and K_2 and the radius of Γ_2. That this is true follows from

$$\frac{|a|^2}{1 - |a|^2} \leq d + 1 + |\alpha_1| \iff \frac{|a|^2}{1 - |a|^2} \leq \frac{|a|}{1 - |a|^2} + 1 + |\alpha_1|$$
$$\iff (2 + |\alpha_1|)|a|^2 - |a| - (1 + |\alpha_1|) \leq 0 ,$$

and this inequality is satisfied for $|a| < 1$. This scenario corresponds to the top middle or the bottom left subfigure in Figure 4.2. Because we are intersecting Γ_2, which is centered at the origin, with the closed exterior of a disk, and because the intersection is nonempty, it follows that some part of the boundary of Γ_2 must belong to the intersection. This means that $|\Gamma_2 \cap K_2| = 1 + |\alpha_1|$. Since $|\alpha_0| < |\alpha_1|$ implies $|\alpha_0| \leq 1 + |\alpha_1|$, this proves the theorem for $|\alpha_0| < |\alpha_1|$.

Finally, we have the case $|\alpha_0| = |\alpha_1|$, or $|a| = 1$. This means that $a \in \Gamma_2$ and, since K_2 is a closed half-plane also containing a, we have, once again, that $\Gamma_2 \cap K_2 \neq \emptyset$. This scenario corresponds to the bottom right subfigure in Figure 4.2. Here, too, some part of the boundary of Γ_2 must belong to the intersection, which means that $|\Gamma_2 \cap K_2| = 1 + |\alpha_1|$, and that concludes the proof. $\square$

Although this theorem indicates that it is not uncommon that $|\Omega_2| = |\Gamma_2|$, the sets themselves can be quite different.

Bounds. Since

$$|\Gamma^{(2)}| = \max \left\{ |\Omega_1|, |\Omega_2|, 1 + |\alpha_2|, \ldots, 1 + |\alpha_{n-1}| \right\} ,$$

the bound on the moduli of the zeros in the following theorem is an immediate consequence of Theorems 3.2 and 4.3.

THEOREM 4.4. *Any zero $\tilde{z}$ of the polynomial $p(z) = z^n + \alpha_{n-1}z^{n-1} + \cdots + \alpha_1 z + \alpha_0$ with complex coefficients and with $\alpha_0 \neq 0$ satisfies $|\tilde{z}| \leq |\Gamma^{(2)}|$, with*

$$|\Gamma^{(2)}| = \max \left\{ \gamma, \delta, 1 + |\alpha_2|, \ldots, 1 + |\alpha_{n-1}| \right\} ,$$

where γ was defined in Theorem 3.2 *and*

$$\delta = \begin{cases} 1 + \dfrac{|\alpha_1|}{|\alpha_0| - |\alpha_1|} & (|\alpha_0| > 1 + |\alpha_1|), \\[2ex] 1 + |\alpha_1| & (|\alpha_0| \leq 1 + |\alpha_1|) . \end{cases}$$

Roughly speaking, the effect of the second modification is to allow $|\alpha_1|$ to become large when $|\alpha_0|$ is also large, as long as it does not get too close to $|\alpha_0|$. Table 4.1 lists the values of $|\Gamma^{(0)}|$, $|\Gamma^{(1)}|$, and $|\Gamma^{(2)}|$, along with the modulus of the largest zero $|\tilde{z}_{max}|$, for five polynomials that are identical except for the coefficient α_1. As $|\alpha_1|$ increases, the difference between the bounds becomes more pronounced, until it approaches $|\alpha_0|$ too closely and its advantage disappears. Kojima's bound is the same as $|\Gamma^{(0)}|$ for all polynomials.

Table 4.1 *Comparison of $|\Gamma^{(0)}|$, $|\Gamma^{(1)}|$, and $|\Gamma^{(2)}|$.*

p	$|\Gamma^{(0)}|$	$|\Gamma^{(1)}|$	$|\Gamma^{(2)}|$	$|\tilde{z}_{max}|$
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 3z + 20$	20	5.5826	5.5826	2.5274
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 6z + 20$	20	7	5.5826	2.6618
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 10z + 20$	20	11	5.5826	2.7938
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 18z + 20$	20	19	10	2.9868
$z^5 + 2z^4 + 3iz^3 - 4z^2 + 19z + 20$	20	20	20	3.0070

Summary of Results and Further Discussion. We have used (2.3)–(2.7) to derive two modifications of the Gershgorin set for a polynomial's companion matrix. They were obtained by using (2.3) to eliminate x_n in (2.7) and in (2.4), respectively. Although it would be beyond our scope here, there are many ways to continue modifying the Gershgorin set by further manipulating (2.3)–(2.7). To give just a few examples, one could use (2.3) to replace x_n by $-\lambda x_1/\alpha_0$ in the right-hand side of all the equations except the first two, or one could multiply (2.5) for $j = k$ by λ and then use the same equation for $j = k - 1$ to eliminate x_{k-1} and obtain new bounds on $|\lambda|$ when $|x_k|$ is maximal. The latter can also be achieved by considering $C^2(p)$ instead of $C(p)$, as it has the same eigenvectors while the eigenvalues are squared. Other powers of $C(p)$ could also be used.

Appendix A. In this appendix we explain in detail how the different cases in Figures 3.3 and 3.4 depend explicitly on $|\alpha_0|$ and $|\alpha_{n-1}|$.

Case I: $2\sqrt{|\alpha_0|} \leq |\alpha_{n-1}|$. In this case, the oval has two loops. We distinguish four subcases.

Case Ia: $|\alpha_0| > r$. The intersection consists of both loops. When $2|\alpha_0| > |\alpha_{n-1}|$, this condition can be rewritten as follows:

$$|\alpha_0| > r \iff 2|\alpha_0| > |\alpha_{n-1}| + \sqrt{|\alpha_{n-1}|^2 + 4|\alpha_0|}$$
$$\iff 2|\alpha_0| - |\alpha_{n-1}| > \sqrt{|\alpha_{n-1}|^2 + 4|\alpha_0|}$$
$$\iff (2|\alpha_0| - |\alpha_{n-1}|)^2 > |\alpha_{n-1}|^2 + 4|\alpha_0|$$
$$\text{(A.1)} \qquad \iff |\alpha_0| > |\alpha_{n-1}| + 1 .$$

Clearly, when $2|\alpha_0| \leq |\alpha_{n-1}|$, then $|\alpha_0| > r$ is impossible. We also note that inequality (A.1) implies that $2|\alpha_0| > |\alpha_{n-1}|$. We can summarize the conditions for this case as $|\alpha_{n-1}| + 1 < |\alpha_0| \leq |\alpha_{n-1}|^2/4$. This is the first scenario of Figure 3.3: on the left is the circle (the boundary of the disk) with the oval inside it, and in the middle is the algebraic description of the conditions, represented graphically by the shaded area on the right. The intersection with the disk is the part of the oval shaded in dark gray. The part shaded in light gray lies outside the intersection. In this case, the modulus of any $z \in \Omega_1$ must satisfy $0 \leq |z| \leq q_-$ or $q_+ \leq |z| \leq r$.

Case Ib: $q_+ \leq |\alpha_0| \leq r$. The intersection consists of the loop containing the origin and part (or all) of the loop containing $-\alpha_{n-1}$. When $2|\alpha_0| \geq |\alpha_{n-1}|$, the condition $|\alpha_0| \geq q_+$ can be rewritten as follows:

$$\begin{aligned} |\alpha_0| \geq q_+ &\Longleftrightarrow 2|\alpha_0| \geq |\alpha_{n-1}| + \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|} \\ &\Longleftrightarrow 2|\alpha_0| - |\alpha_{n-1}| \geq \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|} \\ &\Longleftrightarrow (2|\alpha_0| - |\alpha_{n-1}|)^2 \geq |\alpha_{n-1}|^2 - 4|\alpha_0| \\ &\Longleftrightarrow |\alpha_0| \geq |\alpha_{n-1}| - 1 \ . \end{aligned}$$

When $2|\alpha_0| < |\alpha_{n-1}|$, then $|\alpha_0| \geq q_+$ is obviously impossible. With $2|\alpha_0| \geq |\alpha_{n-1}|$, inequality (A.1) implies that $|\alpha_0| \leq r$ is equivalent to $|\alpha_0| \leq |\alpha_{n-1}| + 1$. Combining the two inequalities, we obtain that $q_+ \leq |\alpha_0| \leq r$ is equivalent to { $2|\alpha_0| \geq |\alpha_{n-1}|$ and $|\alpha_{n-1}| - 1 \leq |\alpha_0| \leq |\alpha_{n-1}| + 1$ }. The conditions for this case are summarized as

$$\max\{|\alpha_{n-1}| - 1, |\alpha_{n-1}|/2\} \leq |\alpha_0| \leq \min\{|\alpha_{n-1}| + 1, |\alpha_{n-1}|^2/4\} \ .$$

This situation is the second scenario in Figure 3.3. Here, the modulus of any $z \in \Omega_1$ satisfies $0 \leq |z| \leq q_-$ or $q_+ \leq |z| \leq |\alpha_0|$.

Case Ic: $q_- < |\alpha_0| < q_+$. The intersection consists of the loop containing the origin. Note that $2\sqrt{|\alpha_0|} = |\alpha_{n-1}|$ cannot occur in this case because that would mean $q_- = q_+$. When $2|\alpha_0| < |\alpha_{n-1}|$, the condition $|\alpha_0| > q_-$ can be simplified as follows:

$$\begin{aligned} |\alpha_0| > q_- &\Longleftrightarrow 2|\alpha_0| > |\alpha_{n-1}| - \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|} \\ &\Longleftrightarrow \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|} > |\alpha_{n-1}| - 2|\alpha_0| \\ &\Longleftrightarrow |\alpha_{n-1}|^2 - 4|\alpha_0| > (|\alpha_{n-1}| - 2|\alpha_0|)^2 \\ &\Longleftrightarrow |\alpha_0| < |\alpha_{n-1}| - 1 \ . \end{aligned}$$

Obviously this can happen only when $|\alpha_{n-1}| > 1$. Because $2|\alpha_0| < |\alpha_{n-1}|$, the expression for q_+ shows that $|\alpha_0| < q_+$ is automatically satisfied. When $2|\alpha_0| \geq |\alpha_{n-1}|$, then $|\alpha_0| > q_-$ is automatically satisfied, and

$$\begin{aligned} |\alpha_0| < q_+ &\Longleftrightarrow 2|\alpha_0| < |\alpha_{n-1}| + \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|} \\ &\Longleftrightarrow 2|\alpha_0| - |\alpha_{n-1}| < \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|} \\ &\Longleftrightarrow (2|\alpha_0| - |\alpha_{n-1}|)^2 < |\alpha_{n-1}|^2 - 4|\alpha_0| \\ &\Longleftrightarrow |\alpha_0| < |\alpha_{n-1}| - 1 \ . \end{aligned}$$

Once again, this can only happen when $|\alpha_{n-1}| > 1$. Combining the two inequalities, we obtain that $q_- < |\alpha_0| < q_+$ is equivalent to $2|\alpha_0| < |\alpha_{n-1}|$ and $|\alpha_0| < |\alpha_{n-1}| - 1$, or $2|\alpha_0| \geq |\alpha_{n-1}|$ and $|\alpha_0| < |\alpha_{n-1}| - 1$. This therefore reduces to the condition $|\alpha_0| < |\alpha_{n-1}| - 1$. Since Case Ic cannot occur when $2\sqrt{|\alpha_0|} = |\alpha_{n-1}|$ and since $\min\{|\alpha_{n-1}| - 1, |\alpha_{n-1}|^2/4\} = |\alpha_{n-1}| - 1$, we can summarize the conditions for this case as $|\alpha_0| < |\alpha_{n-1}| - 1$. This corresponds to the third scenario in Figure 3.3, where the modulus of any $z \in \Omega_1$ satisfies $0 \leq |z| \leq q_-$.

Case Id: $|\alpha_0| \leq q_-$. The intersection consists of part (or all) of the loop containing the origin. When $2|\alpha_0| \leq |\alpha_{n-1}|$, the condition $|\alpha_0| \leq q_-$ can be simplified as

follows:

$$|\alpha_0| \leq q_- \iff 2|\alpha_0| \leq |\alpha_{n-1}| - \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|}$$
$$\iff \sqrt{|\alpha_{n-1}|^2 - 4|\alpha_0|} \leq |\alpha_{n-1}| - 2|\alpha_0|$$
$$\iff |\alpha_{n-1}|^2 - 4|\alpha_0| \leq (|\alpha_{n-1}| - 2|\alpha_0|)^2$$
$$\iff |\alpha_0| \geq |\alpha_{n-1}| - 1 .$$

When $2|\alpha_0| > |\alpha_{n-1}|$, then $|\alpha_0| \leq q_-$ is impossible. These conditions can be condensed into

$$|\alpha_{n-1}| - 1 \leq |\alpha_0| \leq \min\{|\alpha_{n-1}|/2, |\alpha_{n-1}|^2/4\} .$$

This corresponds to the last scenario in Figure 3.3, where the modulus of any $z \in \Omega_1$ now satisfies $0 \leq |z| \leq |\alpha_0|$.

Case II: $2\sqrt{|\alpha_0|} > |\alpha_{n-1}|$. In this case, the oval has only one loop. We distinguish two subcases.

Case IIa: $|\alpha_0| > r$. The intersection consists of the entire oval. When $2|\alpha_0| > |\alpha_{n-1}|$, we obtain, as in Case Ia, that this condition is equivalent to $|\alpha_0| > |\alpha_{n-1}| + 1$. When $2|\alpha_0| \leq |\alpha_{n-1}|$, then $|\alpha_0| > r$ is impossible. Summarizing, we obtain $|\alpha_0| > \max\{|\alpha_{n-1}| + 1, |\alpha_{n-1}|^2/4\}$. This is the first scenario in Figure 3.4. In this case, the modulus of any $z \in \Omega_1$ satisfies $0 \leq |z| \leq r$.

Case IIb: $|\alpha_0| \leq r$. The intersection consists of part (or all) of the oval. When $2|\alpha_0| > |\alpha_{n-1}|$, then, as in Case Ib, we obtain from (A.1) that this condition is equivalent to $|\alpha_0| \leq |\alpha_{n-1}| + 1$. When $2|\alpha_0| \leq |\alpha_{n-1}|$, the condition is automatically satisfied. This is summarized as $|\alpha_{n-1}|^2/4 < |\alpha_0| \leq |\alpha_{n-1}| + 1$, and it is the second scenario of Figure 3.4. Here, the modulus of any $z \in \Omega_1$ satisfies $0 \leq |z| \leq |\alpha_0|$.

Appendix B. Below is the MATLAB 7 code used to produce a typical figure, in this case the right-hand side figure in Figure 4.1. It is similar enough to pseudo-code for it to be understandable with minimal knowledge of MATLAB. The code contains some redundancy in the interest of clarity. After the figure is created, its colors are converted to grayscale. Broadly speaking, the code creates a grid and then checks which grid points lie in the set of interest. These grid points are then plotted. The areas of the sets can be compared by comparing the number of grid points they contain. We have rounded throughout this work to the nearest larger integer percentage. All other figures were generated by very similar code.

MATLAB Code.
```
% v is the coefficient vector of the polynomial
% absv is the vector of absolute values of the coefficients in v
% m is the resolution of the figure
  v=[1 2 i -2 6 8];absv=abs(v);m=2000;
% n is the degree of the polynomial
  n=length(v)-1;
% comp_mat is the companion matrix of the polynomial
  comp_mat=(fliplr(flipud(compan(v))))';
% the vector of deleted row sums Rt is computed
  Rt=1+absv(n+1:-1:2);
  Rt(1)=absv(n+1);
  Rt(n)=1;
% the boundaries of the grid are established
  dd=diag(comp_mat);
```

```
  xmin=min(real(dd))-max(Rt);
  xmax=max(real(dd))+max(Rt);
  ymin=min(imag(dd))-max(Rt);
  ymax=max(imag(dd))+max(Rt);
% the horizontal and vertical distance between grid points is set
  del=min((xmax-xmin)/m,(ymax-ymin)/m);
% the real and imaginary parts of the grid points are generated
  [xx1,xx2]=meshgrid(xmin:del:xmax,ymin:del:ymax);
% the complex grid points are defined as z
  z=xx1+i*xx2;
% the indices of the grid points in the Gershgorin set
% are collected in gset
  gset=((abs(z-comp_mat(1,1)) <= Rt(1)));
  for jj=2:n
      gset=gset | ((abs(z-comp_mat(jj,jj)) <= Rt(jj)));
  end
% the Gershgorin set is plotted in magenta
  figure;hold on;axis equal;
  zz=z(gset);
  plot(real(zz),imag(zz),'m.')
% the indices of the grid points in the first modified Gershgorin set
% are collected in gset1
  gset1=(abs(z) <= absv(n+1));
  gset1=gset1 & (abs(z).*abs(z+v(2)) <= absv(n+1));
  for jj=1:n-2
      gset1=gset1 | ((abs(z) <= 1+absv(n+1-jj)));
  end
  gset1=gset1 | (abs(z+v(2)) <= 1);
% the first modified Gershgorin set is added to the figure in green
  zz1=z(gset1);
  plot(real(zz1),imag(zz1),'g.');
% the indices of the grid points in the second modified Gershgorin set
% are collected in gset2
  a=v(n+1)/v(n);
  gset2=(abs(z) <= absv(n+1));
  gset2=gset2 & (abs(z).*abs(z+v(2)) <= absv(n+1));
  gset2=gset2 | ((abs(z) <= 1 + absv(n)) & (abs(a)*abs(z) <= abs(a+z)));
  for jj=2:n-2
      gset2=gset2 | ((abs(z) <= 1+absv(n+1-jj)));
  end
  gset2=gset2 | (abs(z+v(2)) <= 1);
% the second modified Gershgorin set is added to the figure in blue
  zz2=z(gset2);
  plot(real(zz2),imag(zz2),'b.');
% the zeros of the polynomial are added to the figure as white dots
  eigv=eig(comp_mat);
  plot(real(eigv),imag(eigv),'w.','MarkerSize',20);
  hold off;
% computation of the approximate ratio of the area of the modified sets
% as a fraction of that of the Gershgorin set
  sz=size(zz);
  sz1=size(zz1);
  sz2=size(zz2);
  area1=sz1(1)/sz(1);
  area2=sz2(1)/sz(1);
  [area1 area2]
```

REFERENCES

[1] Yu. A. Alpin, M.-T. Chien, and L. Yeh, *The numerical radius and bounds for zeros of a polynomial*, Proc. Amer. Math. Soc., 131 (2003), pp. 725–730.
[2] S. Barnett, *Congenial matrices*, Linear Algebra Appl., 41 (1981), pp. 277–298.

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

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

[5] A. Edelman and H. Murakami, *Polynomial roots from companion matrix eigenvalues*, Math. Comp., 64 (1995), pp. 763–776.

[6] M. Fiedler, *A note on companion matrices*, Linear Algebra Appl., 372 (2003), pp. 325–331.

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

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

[9] J. D. Lawrence, *A Catalog of Special Plane Curves*, Dover, New York, 1972.

[10] L. Lévy, *Sur la possibilité de léquilibre électrique*, C. R. Acad. Sci. Paris, 93 (1881), pp. 706–708.

[11] H. Linden, *Bounds for the zeros of polynomials from eigenvalues and singular values of some companion matrices*, Linear Algebra Appl., 271 (1998), pp. 41–82.

[12] M. Marden, *Geometry of Polynomials*, Math. Surveys Monogr. 3, AMS, Providence, RI, 1966.

[13] Y.-J. Sun and J.-G. Hsieh, *A note on the circular bound of polynomial zeros*, IEEE Trans. Circuits Syst. I, 43 (1996), pp. 476–478.

[14] O. Taussky, *A recurring theorem on determinants*, Amer. Math. Monthly, 56 (1949), pp. 672–676.

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

[16] H. S. Wilf, *Perron-Frobenius theory and the zeros of polynomials*, Proc. Amer. Math. Soc., 12 (1961), pp. 247–250.