3. R. P. Boas, L'hospital's rule without mean-value theorems, *American Mathematical Monthly* 76:9 (1969) 1051–1053.
4. ———, Indeterminate forms revisited, *Mathematics Magazine* 63:3 (1990) 155–159.
5. Leonard Gillman and Robert H. McDowell, *Calculus*, Norton, New York, 1973.
6. Richard E. Johnson and Fred L. Kiokemeister, *Calculus with Analytic Geometry*, Allyn and Bacon, Boston, 1957.
7. Ray Redheffer, Some thoughts about limits, *Mathematics Magazine* 62:3 (1989) 176–184.
8. Victor Rouquet, Note sur les vraies valeurs des expressions de la forme ∞/∞, *Nouvelles Annales de Mathématique* 16:2 (1877) 113–116.
9. Walter Rudin, *Principles of Mathematical Analysis*, 2nd ed., McGraw-Hill, New York, 1953.
10. Otto Stolz, *Grundzüge der Differential- und Integral-rechnung*, vol. 1, Teubner, Leipzig, 1893.

Bounding the Roots of Polynomials

Holly P. Hirst (hph@math.appstate.edu) and Wade T. Macey, Appalachian State University, Boone, NC 28608

In these days of ubiquitous graphing devices, a standard problem in mathematics courses at all levels asks the student to generate a graph of a polynomial function on an interval that contains all the real roots. In this article we will discuss some simple bounds on the roots of a polynomial function based upon its coefficients. The results actually give disks in the complex plane that are guaranteed to contain all of the roots, real or complex, of the polynomial.

The bounds we describe are not new. The novelty of our presentation lies in the simplicity of the proof of the first theorem, which uses only elementary properties of absolute values and thus is easy to understand and apply even for pre-calculus students. One of the bounds on the roots that we will present was first reported by Cauchy in 1829. After Cauchy's work was published, bounding roots of polynomials remained a popular topic of study for over a century; many people produced related results using widely differing techniques from areas such as linear algebra and complex analysis. Thus the study of bounds for the roots of polynomials in terms of the coefficients convincingly demonstrates the interconnections between different fields of mathematics.

We found an added bonus when we looked into the history of this topic—a well documented historical record of the development of an idea that is accessible to undergraduates. Many results about polynomial roots are described in detail in one convenient source [**3**], which gives an excellent account of the activity in this area over the past two centuries. We recommend it for all who study polynomials, regardless of their particular interest.

We begin with our main result.

Theorem 1. *Given $f : \mathbb{C} \to \mathbb{C}$ defined by $f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, where $a_0, a_1, \ldots, a_n \in \mathbb{C}$, and n a positive integer. If z is a zero of f, then*

$$|z| \leq \max\left\{1, \sum_{i=0}^{n-1} |a_i|\right\}. \tag{1}$$

Proof. Let z be a zero of f. If $a_0 = a_1 = \cdots = a_{n-1} = 0$, so $f(z) = z^n$, then $|z| = 0 < 1$.

Now suppose that at least one of $a_0, a_1, a_2, \ldots a_{n-1}$ is not zero. If the modulus of our root z is less than or equal to 1 then (1) is satisfied, so we need only show that if $|z| > 1$ then

$$1 < |z| \leq |a_{n-1}| + \cdots + |a_1| + |a_0| . \tag{2}$$

Factoring out the term of highest degree gives

$$f(z) = z^n \left[1 + \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \cdots + \frac{a_1}{z^{n-1}} + \frac{a_0}{z^n}\right].$$

Since $z \neq 0$ and $f(z) = 0$, the second factor must vanish, so

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

The triangle inequality then gives

$$1 \leq \left|\frac{a_{n-1}}{z}\right| + \left|\frac{a_{n-2}}{z^2}\right| + \cdots + \left|\frac{a_1}{z^{n-1}}\right| + \left|\frac{a_0}{z^n}\right|.$$

But $|z| > 1$ implies that $\left|z^k\right| > |z|$, and so $\dfrac{1}{|z^k|} < \dfrac{1}{|z|}$, for $2 \leq k \leq n$. Thus

$$1 \leq \frac{|a_{n-1}|}{|z|} + \frac{|a_{n-2}|}{|z|} + \cdots + \frac{|a_1|}{|z|} + \frac{|a_0|}{|z|},$$

and multiplying through by $|z|$ yields (2). $\square$

Notice that the last inequality is strict unless $a_0 = a_1 = \cdots = a_{n-2} = 0$. Thus inequality (2) is strict unless f has the form $f(z) = z^n - az^{n-1}$, $a \neq 0$. This polynomial has roots 0 and a, so when $|a| > 1$ this polynomial gives the only case where equality is achieved in (2).

This theorem can be used to help pre-calculus students find an interval containing all the roots, which can be a useful step before plotting the graph as part of an investigation of a given polynomial.

Example. Find an interval that is guaranteed to contain all of the real roots of the polynomial $3x^5 + 5x^3 - 9x^2 + 4x + 12$.

Solution. To apply Theorem 1, we must use the corresponding monic polynomial $x^5 + \frac{5}{3}x^3 - 3x^2 + \frac{4}{3}x + 4$. Since $\sum |a_i| = \frac{5}{3} + 3 + \frac{4}{3} + 4 = 10$, we conclude that the real roots will lie in the interval $(-10, 10)$.

Theorem 1 is similar to many results that have been discovered during the last two centuries. The following theorem was published by Cauchy in 1829 [**2**].

Theorem 2. (Cauchy) *All zeros of the polynomial* $f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ *lie in the disk*

$$|z| < 1 + \max_{0 \leq k \leq n-1} |a_k| .$$

We may assume $z \neq 0$.

Proof. Cauchy's bound requires for its proof the triangle inequality and the formula for the sum of an infinite geometric series. By the triangle inequality,

$$\begin{aligned}|f(z)| &\geq |z^n| - \left(\left|a_{n-1}z^{n-1}\right| + \left|a_{n-2}z^{n-2}\right| + \cdots + |a_1 z| + |a_0|\right) \\ &= |z^n| \left(1 - \left|\frac{a_{n-1}}{z}\right| - \left|\frac{a_{n-2}}{z^2}\right| - \cdots - \left|\frac{a_1}{z^{n-1}}\right| - \left|\frac{a_0}{z^n}\right|\right).\end{aligned}$$

If $M = \max\limits_{0 \leq k \leq n-1} |a_k|$, it follows that

$$\begin{aligned}|f(z)| &\geq |z|^n \left(1 - M \sum_{j=1}^{n} \frac{1}{|z|^j}\right) \\ &> |z|^n \left(1 - M \sum_{j=1}^{\infty} |z|^{-j}\right).\end{aligned}$$

For $|z| > 1$, the geometric series converges with sum $1/(|z| - 1)$, giving

$$\begin{aligned}|f(z)| &> |z|^n \left(1 - \frac{M}{|z| - 1}\right) \\ &= |z|^n \left(\frac{|z| - 1 - M}{|z| - 1}\right).\end{aligned}$$

From this inequality we see that if $|z| \geq 1 + M$, then $|f(z)| > 0$. Thus, if z is a zero of f, we must have $|z| < 1 + M$. $\square$

Cauchy's bound is a bit simpler than that of Theorem 1; however, the proof relies upon knowledge of infinite series. Also, Cauchy's disk is larger for many polynomials than the disk from Theorem 1. For example, if $f(z) = z^n - az^{n-1}$, where $|a| > 1$, Theorem 1 gives $|z| \leq |a|$, while Cauchy's bound is $|z| < 1 + |a|$. As another interesting case, consider $f(z) = z^n - a$, which has n roots, all of modulus $\sqrt[n]{|a|}$. If $|a| < 1$, Theorem 1 gives the bound $|z| \leq 1$ and if $|a| > 1$, it gives the bound $|z| < |a|$. In either case Cauchy's disk is larger: $|z| < 1 + |a|$. On the other hand, for the polynomial in the example above, Cauchy's bound is sharper: $|x| \leq 5$.

Students who have had an introduction to eigenvalues and eigenvectors may be surprised by the following proof of Theorem 1. We should not overlook such opportunities to demonstrate relationships between different areas of mathematics. Recall that the *companion matrix* of the monic polynomial

$$f(z) = z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_1 z + a_0$$

is the $n \times n$ matrix

$$\mathrm{C} = \begin{bmatrix} -a_{n-1} & -a_{n-2} & \cdots & -a_1 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

The proof by induction that the characteristic polynomial $\det(xI - \mathrm{C})$ is $f(x)$ makes an excellent exercise on the cofactor expansion of determinants [**5**]. Thus the

roots of any polynomial are the eigenvalues of its companion matrix, and we can bring the machinery of linear algebra to bear to find properties of the eigenvalues.

The next theorem is simple to prove directly from the definition of eigenvalue. The proof can be found in many linear algebra texts, such as [**4**].

Theorem. (Gershgorin) *The eigenvalues of an $n \times n$ matrix A are contained in the union of the disks in the complex plane given by*

$$D_i = \left\{ z \ : \ |z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \right\}, \qquad i = 1, \ldots, n.$$

When Gershgorin's theorem is applied to the companion matrix of the polynomial f, the resulting disks are

$$D_1 = \left\{ z : \ |z + a_{n-1}| \leq \sum_{i=0}^{n-2} |a_i| \right\}, \quad D_i = \{ z : |z| \leq 1 \}, \qquad i > 1.$$

Since

$$|z| = |(0 - a_{n-1}) + (z + a_{n-1})| \leq |a_{n-1}| + |z + a_{n-1}|,$$

the points in D_1 satisfy

$$|z| \leq \sum_{i=0}^{n-1} |a_i|,$$

and so Gershgorin's theorem implies Theorem 1. (Again, for $|z| > 1$ we get strict inequality here except for $f(z) = z^n - az^{n-1}$.)

It is possible to look at these theorems from yet another interesting point of view—that of complex analysis. Rouché's theorem can be used to derive bounds on the roots of polynomials similar to those in Theorems 1 and 2 [**1**]. The real strength of this approach is that Rouché's theorem can be used to determine the exact number of zeros in a given region of the complex plane.

It was intriguing to us that Theorems 1 and 2 have simple proofs, one of which has been known since 1829, and yet most recent references to these results cite more sophisticated proofs using Gerschgorin's theorem from linear algebra or Rouché's theorem from complex analysis.

References

1. L. V. Alfors, *Complex Analysis*, 3rd ed., McGraw-Hill, New York, 1979,152–153.
2. A. L. Cauchy, Exercises de mathématique,*Oeuvres* **2**, vol. 9, 1829, 122.
3. M. Marden, *Geometry of Polynomials*, American Mathematics Society, Providence, RI, 1966.
4. G. W. Stewart, *Introduction to Matrix Computations*, Academic Press, Orlando, FL, 1973, 273.
5. G. Strang, *Linear Algebra and Its Applications*, 2nd ed., Academic Press, New York, 1980, 304.