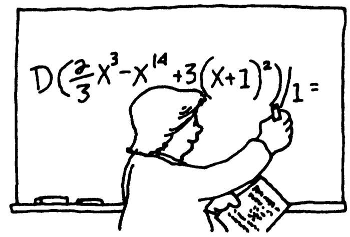
CLASSROOM CAPSULES

EDITOR

Thomas A. Farmer
Department of Mathematics and Statistics
Miami University
Oxford, OH 45056-1641



A Classroom Capsule is a short article that contains a new insight on a topic taught in the earlier years of undergraduate mathematics. Please submit manuscripts prepared according to the guidelines on the inside front cover to Tom Farmer.

Order Relations and a Proof of l'Hôpital's Rule

Leonard Gillman (gillman@mail.utexas.edu), 1606 The High Road, Austin, TX, 78746-2236

This note stems from my having read and enjoyed the recent MAA book of selected works of R. P. Boas [**1**].

The "order" definition of limit. The ε, δ definition of the limit or continuity of a real-valued function f defined on a subset of $\mathbb{R}$ refers to metric properties of $\mathbb{R}$. But in most circumstances—particularly in freshman calculus—*order* concepts are sufficient, while metric details are irrelevant and a visual and mental distraction. Moreover, a series of ε's and δ's in the definition prepares the mind for computation even when none may be forthcoming. To insist that the required intervals in the definition of limit be *centered* at a or $f(a)$ is usually unnecessary and often ridiculous.

Trivially, the family of *all* open intervals about a point constitutes a base for the neighborhood system at the point. Thus, continuity of f at a means that, for any open interval J about $f(a)$, there is an open interval about a that f takes into J. This formulation is simply the specialization to real functions on $\mathbb{R}$ of the definition (in terms of basic open sets) for arbitrary topological spaces. Likewise,

$$\lim_{x \to a} f(x) = L \tag{1}$$

means that, for any open interval J about L, there is a *punctured* open interval $I \setminus \{a\}$ that f takes into J. (Not "*deleted* interval," for gosh sakes: that's like "escaped prison.") Starting thus with an interval about L focuses directly on the goal and is certainly more natural than starting with a number ε having no visible relation to the problem; it also reduces the number and variety of symbols. For limits at infinity and infinite limits, this is essentially the definition we are all accustomed to.

At the working level we often refer to the actual endpoints of the intervals. Thus (1) means (for L finite) that if A and B are any numbers satisfying $A < L < B$, then

$$A < f(x) < B$$

near a—that is, on (= throughout) some punctured interval $I \setminus \{a\}$.

We may also consider the challenges from each side independently. Note that the elementary "proximity" theorem,

$$\textit{If} \quad \lim_{x \to a} F(x) > A, \quad \textit{then} \quad F(x) > A \quad \textit{near} \quad a, \tag{2}$$

is now just a quotation from the order definition of limit.

An advantage in using the order definition is shown clearly in the limit theorems for powers of a function. For example, let us show that if $f(x) \to L \neq 0$ as $x \to a$, then $1/f(x) \to 1/L$. Say $L > 0$. Given a challenge $A < 1/L < B$, we may assume $A > 0$. Then $1/B < L < 1/A$, and because $f(x) \to L$, $1/B < f(x) < 1/A$ near a. Then $A < 1/f(x) < B$ near a. □

Textbooks that adopted the order definition of limit include [**6**], [**2**], and [**5**], in increasing level of commitment, the commitment in [**5**] being total.

Proof of l'Hôpital's rule. The rule as stated in l'Hôpital's own calculus book was an elementary special case of the modern version, for which the usual proof depends on Cauchy's extended mean value theorem. But the proof of Cauchy's theorem starts out with a monster such as

$$\text{Consider the function} \quad F(t) = [f(b) - f(a)]g(t) - [g(b) - g(a)]f(t),$$

a *deus ex machina* that frightens students and from which they learn nothing, though professional mathematicians are charmed by its elegance. The earliest attacks on a proof free of mean value theorems seems to be the 1877 paper by Victor Rouquet [**8**], which treats the special case $f(t)/t$, and the 1889 calculus book by Otto Stolz [**10,** p. 82].

The proofs that follow for the form $0/0$ and ∞/∞ are based, resepectively, on Boas [**4**] and [**3**], which are reproduced in [**1**]. Both bypass the Cauchy theorem but use ε's. In contrast, Rudin [**9**] does without ε's but uses Cauchy.

L'Hôpital's rule is best formulated as a theorem about one-sided limits. I will consider the limits to be all from the right, and for convenience I shorten the symbol $\lim_{t \to a^+}$ to lim. The term *near* a^+ means of course *on an interval* (a, u); I think it is due to Redheffer [**7**]. Likewise, *near* ∞ means *for sufficiently large* t.

L'Hôpital's rule. *Let f and g be differentiable on an interval (a, v) ($v = \infty$ being permitted), with g' being continuous on the interval, and $g'(t) \neq 0$ near a^+. If* $\lim f(t)/g(t)$ *assumes the indeterminate form* $0/0$ *or if* $\lim g(t) = \infty$, *and if* $\lim f'(t)/g'(t)$ *exists, finite or infinite, then*

$$\lim \frac{f(t)}{g(t)} = \lim \frac{f'(t)}{g'(t)}.$$

In the proofs that follow, let

$$L = \lim \frac{f'(t)}{g'(t)}. \tag{3}$$

The discussion assumes that L is finite; for L infinite, just ignore the condition $L < B$ and the inequalities that ensue from it. Since g' is never zero near a^+, it is of one sign there (intermediate value theorem).

Proof for the form 0/0. Given a challenge

$$A < L < B, \tag{4}$$

we will show that

$$A < \frac{f(t)}{g(t)} < B \quad \text{near} \quad a^+. \tag{5}$$

From (3) and (4) we have, by definition of limit,

$$A < \frac{f'(t)}{g'(t)} < B \quad \text{near} \quad a^+. \tag{6}$$

For convenience, define $f(a) = g(a) = 0$; then f and g are continuous at 0. Say $g'(t) > 0$; then $g(t) > 0$ for $t > a$; also, multiplication by $g'(t)$ in (7) preserves order:

$$Ag'(t) < f'(t) < Bg'(t) \quad \text{near} \quad a^+. \tag{7}$$

Thus $(f - Ag)'(t) > 0$ near a^+, so $f - Ag$ is increasing on an interval $[a, v_1)$; that is,

$$(f - Ag)(t) > (f - Ag)(a) = 0 \quad \text{near} \quad a^+.$$

Consequently, $f(t) > Ag(t)$. Similarly, $f(t) < Bg(t)$, so

$$Ag(t) < f(t) < Bg(t) \quad \text{near} \quad a^+. \tag{8}$$

(Alternatively, obtain these inequalities by integrating (7) on $[a, t]$, remembering to use a different variable of integration). Finally, dividing in (8) by $g(t)$ yields (5). □

Proof for the case g(t) → ∞. This proof is a little more complicated. Given a challenge $A < L < B$, we will show that

$$A < \frac{f(t)}{g(t)} < B \quad \text{near} \quad a^+.$$

Choose A^* and B^* such that $A < A^* < L < B^* < B$. Since $\lim f'(t)/g'(t) = L$, we have by definition of limit

$$A^* < \frac{f'(t)}{g'(t)} < B^* \quad \text{near} \quad a^+. \tag{9}$$

Since $g(t) \to \infty$ as $t \to a^+$, $g(t) > 0$ and $g'(t) < 0$ near a^+. Multiplying by $g'(t)$ in (9) then reverses order:

$$A^*g'(t) > f'(t) > B^*g'(t) \quad \text{near} \quad a^+.$$

Looking at the second inequality, we see that $(f - B^*g)'(t) > 0$, so $f - B^*g$ is increasing. Thus for $x < y$, $(f - B^*g)(x) < (f - B^*g)(y)$, which implies

$$f(x) - f(y) < B^*g(x) - B^*g(y).$$

Transposing $f(y)$ and dividing by $g(x)$ yields

$$\frac{f(x)}{g(x)} < B^* + \frac{f(y) - B^* g(y)}{g(x)}.$$

Fix y; as $x \to a^+$, the second term on the right goes to 0, so the right-hand side approaches B^* and is therefore eventually less than B; consequently

$$\frac{f(x)}{g(x)} < B \quad \text{near} \quad a^+.$$

Similarly, $f(x)/g(x) > A$ near a^+. □

Technical comment. The hypothesis that g' be continuous is included so that we can apply the intermediate value theorem for continuous functions. But it is redundant, as *all* derivatives enjoy the intermediate value property. This *intermediate value theorem for derivatives* is held in disfavor by many teachers, however, on the grounds that derivatives that arise naturally in standard calculus courses are always continuous. Nevertheless, I like the theorem, particularly because its proof is so simple and instructive.

Intermediate value theorem for derivatives. *If f' is defined on $[a, b]$ and has opposite signs at a and b, then it is zero at some point in between.*

The proof rests on a lemma that every calculus student should understand:

Lemma. (Functions increasing at a point.) *If $f'(a) > 0$, then $f(x) > f(a)$ for x near a^+ and $f(x) < f(a)$ for x near a^-.*

Note: It does not follow that there is an interval about a on which f is increasing—though the simplest counterexample I can cook up is $f(x) = x + x^3 \sin(1/x)$.

Proof of the lemma. Consider the difference quotient $F(x) = [f(x) - f(a)]/(x - a)$; then $\lim_{x \to a} F(x) = f'(a) > 0$. By the proximity theorem (2), $F(x) > 0$ near a. Hence for x near a, the fraction has the same sign upstairs as downstairs, and the desired conclusion follows. □

Proof of the theorem. Say $f'(a) > 0 > f'(b)$. By the lemma, there are points x near a^+ for which $f(x) > f(a)$. Similarly, there are points x near b^- for which $f(x) > f(b)$. Consequently, the maximum of the continuous function f on the closed interval $[a, b]$ does not occur at either endpoint. It therefore occurs at an interior point; and at such a point the derivative is zero. □

Acknowledgment. I wish to thank the referees for several helpful suggestions.

References

1. Gerald L. Alexanderson and Dale H. Mugler, eds., *Lion-Hunting & Other Mathematical Pursuits*, Mathematical Association of America, Washington, DC, 1955.
2. Lipman Bers, *Calculus*, Holt, Rinehart & Winston, New York, 1969.

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.

———o———

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 $f(z)=z^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0$, where $a_0,a_1,\ldots,a_n\in\mathbb{C}$, and n a positive integer. If z is a zero of f, then*

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

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