ALTERNATIVE PROOFS OF SOME RESULTS FROM ELEMENTARY ANALYSIS

Richard E. Bayne, James E. Joseph, Myung H. Kwack, and Thomas H. Lawson

In an endeavor to encourage mathematics students to search for and study various methods of proof we have embarked on a program designed to develop and present approaches different from those found in the textbooks used by the students. As a result of this effort proofs of several well–known results from elementary analysis have been developed and presented. In this article we offer some of these proofs. Although we do not know if the proofs are new they are elegant, different from those usually found in books in analysis, and may be of use to students and teachers of the subject. First we present two proofs that a real–valued function which is 1–1 and continuous on an interval I in the reals is either increasing or decreasing on I. These proofs are different from and should be compared with that given in [9]. It is commented in [9] that it is possible to give a straightforward, but cumbersome, proof that involves keeping track of a lot of cases [4]. The proof given there dispenses with those unpleasant details, but is described by the author as "rather tricky." We then apply the so–called "creeping" method to establish three classical results: Dini's Lemma, the Bolzano–Weierstrass Theorem for the reals, and the Heine–Borel Theorem for Euclidean n–space. This method is abstracted and discussed in [6]. Proofs of Dini's Lemma and the Bolzano–Weierstrass Theorem are given in [6]. Those proofs should be compared with the proofs offered here. The proof of the Heine–Borel Theorem joins the method with induction on the dimension of the space and should be compared with that given in [8].

The article concludes with a nice induction proof of the Cauchy–Schwarz Inequality for complex numbers, a proof of the Lagrange Identity for complex numbers obtained by expressing the expansion of

$$\left| \sum_{k=1}^{n} \overline{z}_k w_k \right|^2$$

in a different form, and a generalization of the Lagrange Identity to inner–product spaces. These proofs should be compared with proofs found in books such as [1, 2, 5, 8], and with a proof in [3] using the upwards–downwards form of the principle of mathematical induction. We come to our results.

<u>Theorem 1</u>. A real–valued function f which is 1–1 and continuous on an interval I in the reals is either increasing or decreasing on I.

Proof. First, if $a, b \in I$ and $a < b$ then

$$\min\{f(a), f(b)\} \leq f(x) \leq \max\{f(a), f(b)\}$$

for all $x \in [a, b]$. To see this, if $v \in (a, b)$ satisfies

$$f(v) < \min\{f(a), f(b)\} \quad \text{or} \quad f(v) > \max\{f(a), f(b)\}$$

choose y satisfying

$$\begin{aligned} &f(v) < y < \min\{f(a), f(b)\} < \max\{f(a), f(b)\} \quad \text{or} \\ &f(v) > y > \max\{f(a), f(b)\} > \min\{f(a), f(b)\}. \end{aligned}$$

Then, from the Intermediate Value Theorem, there exists $p \in (a, v)$, $q \in (v, b)$ such that $f(p) = f(q) = y$, contradicting the hypothesis that f is 1–1. This says that f achieves absolute extrema on any closed subinterval and at the endpoints of the interval. Now suppose $a, b \in I$ with $a < b$ and $f(a) < f(b)$. Let $a \leq x < y \leq b$. It follows that $f(x) < f(y)$ so f is increasing on $[a, b]$. Finally, let $x, y \in I$ with $x \leq a < b \leq y$. Then $f(x) < f(y)$ so f is increasing on $[x, y]$. The proof is complete.

Another Proof of Theorem 1. Let $a, b \in I$ with $a < b$ and define F on $[a, b]$ by

$$F(x) = (f(b) - f(a))(f(x) - f(a)).$$

Then F is 1–1 and continuous on $[a, b]$, $F(a) = 0$, $F(b) > 0$. Hence, $F(x) > 0$ for all $x \in (a, b]$. Now fix $y \in (a, b]$ and define G on $[a, b]$ by

$$G(x) = (f(y) - f(a))(f(y) - f(x)).$$

Then G is 1–1 and continuous on $[a, y]$, $G(y) = 0$, $G(a) > 0$. So $G(x) > 0$ for all $x \in [a, y)$. Hence, for all x, y such that $a \leq x < y \leq b$,

$$(f(b) - f(a))(f(y) - f(a)) = F(y) > 0, \tag{$*$}$$

$$(f(y) - f(a))(f(y) - f(x)) = G(x) > 0, \tag{$**$}$$

and consequently

$$(f(b) - f(a))(f(y) - f(x)) > 0. \tag{$***$}$$

If $a, b, x, y, \in I$ with $a < b$, $x < y$ then, from inequality $(***)$

$$\begin{aligned} &(f(\max\{b, y\}) - f(\min\{a, x\}))(f(y) - f(x)) > 0 \quad \text{and} \\ &(f(\max\{b, y\}) - f(\min\{a, x\}))(f(b) - f(a)) > 0. \end{aligned}$$

Therefore
$$(f(b) - f(a))(f(y) - f(x)) > 0.$$
The proof is complete.

We observe from the proofs above that the next result holds.

Theorem 2. The following statements are equivalent for a continuous real–valued function f defined on an interval I in the reals.

(a) The function f is 1–1 on I.
(b) The function f is either strictly increasing or strictly decreasing on I.
(c) The function f does not take a local extrema on the interior of I.

The proofs of Dini's Lemma, the Bolzano–Weierstrass Theorem, and the Heine–Borel Theorem, given below, utilize a so–called "creeping" method. Again these proofs should be compared with those in [2, 4, 6, 7, 8]. A real–valued function defined on a subset of the reals is upper semicontinuous at x if for each y satisfying $f(x) < y$ there is an open interval I such that $x \in I$ and each $v \in I$ which is in the domain of f satisfies $f(v) < y$.

Dini's Lemma. If f_n is a sequence of upper semicontinuous real–valued functions defined on $[0, 1]$ such that, at each point x, the sequence $f_n(x)$ is a nonincreasing sequence converging to 0, then f_n converges uniformly to the constant function 0 on $[0, 1]$.

Proof. Let $\epsilon > 0$. Choose m such that $f_m(0) < \epsilon$ and $z \in (0, 1]$ such that $f_m([0, z]) \subset [0, \epsilon)$. For each $n \geq m$ we have $f_n([0, z]) \subset [0, \epsilon)$. Let $c = \sup\{z \in (0, 1] : f_n([0, z]) \subset [0, \epsilon)$, ultimately$\}$. Choose $a \in (0, c)$ such that $f_n([a, c]) \subset [0, \epsilon)$, ultimately and $b \in [c, 1]$ such that $f_n([0, b]) \subset [0, \epsilon)$, ultimately. It follows that $f_n([0, c]) \subset [0, \epsilon)$, ultimately. If $c < 1$ we may choose b satisfying $c < b < 1$ and $f_n([0, b]) \subset [0, \epsilon)$, ultimately. This would contradict a property of supremum. Hence, $f_n([0, 1]) \subset [0, \epsilon)$, ultimately.

Let $\mathbb{N}$ be the set of natural numbers. A sequence in a set X is a function from $\mathbb{N}$ to X. If f is a sequence in X, a subsequence of f is a sequence of the form $f \circ \mu$ where $\mu \colon \mathbb{N} \to \mathbb{N}$ is strictly increasing. It can be shown by induction that

(i) If $n \in \mathbb{N}$ and $\mu_1, \ldots, \mu_n$ are strictly increasing from $\mathbb{N}$ to $\mathbb{N}$, then $\mu_1 \circ \cdots \circ \mu_n$ is strictly increasing.
(ii) If $\mu \colon \mathbb{N} \to \mathbb{N}$ is strictly increasing then $\mu(n) \geq n$.

The following proposition is used often in mathematics.

Proposition. If $\{F_n : n \in \mathbb{N}\}$ is a family of subsets of a set X such that $F_{n+1} \subset F_n$ and f is a sequence in X such that, for each n, some subsequence of f is in F_n, there is a subsequence h of f such that, for each n, h is ultimately in F_n.

Proof. Let $f \circ \mu_n$ be a subsequence of f in F_n. Define $\psi : \mathbb{N} \to \mathbb{N}$ by $\psi(n) = \mu_n \circ \cdots \circ \mu_1(n)$. Then

$$\psi(n+1) = \mu_{n+1} \circ \mu_n \circ \cdots \circ \mu_1(n+1) \geq \mu_n \circ \cdots \circ \mu_1(n+1) > \psi(n).$$

Hence, $h = f \circ \psi$ is a subsequence of f. If $m \in \mathbb{N}$ and $n > m$, then $h(n) = f \circ \psi(n) \in F_n \subset F_m$. The proof is complete.

Definition. A sequence g in the reals is bounded if $g(\mathbb{N})$ is a bounded subset of the reals.

Theorem 3 (Bolzano–Weierstrass). Each bounded sequence in the reals has a convergent subsequence.

Proof. Let x_n be a sequence in $[a, b]$ and let c be the supremum of all z such that no subsequence of x_n is in $(-\infty, z)$. Note that a is such a z and that b is an upper bound for the set of such z. It is now quite evident from the proposition immediately above that some subsequence of x_n converges to c and that $c = \liminf x_n$.

Remark 1. The proof of Theorem 3 given in [6] uses, as a statement of the Bolzano–Weierstrass Theorem, the contrapositive of the statement

Every bounded infinite subset of the reals has an accumulation point.

To facilitate our proof of the Heine–Borel Theorem, we give preliminary definitions, remarks, and notations. The notation $\mathbb{R}^n$ will represent Euclidean n–space.

Definition. An *n–dimensional symmetric open interval I in* $\mathbb{R}^n$ is defined by

$$I = \{(x_1, \ldots, x_n) \mid a_k < x_k < b_k\},$$

where $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$, $b_k - a_k = b_j - a_j$ for all k, j. The center of I is the point $((a_1 + b_1)/2, \ldots, (a_n + b_n)/2)$.

Definition. A subset Q of $\mathbb{R}^n$ is said to be *open* if for each $x \in Q$ there is a symmetric open interval I with center at x such that $I \subset Q$. A subset Q of $\mathbb{R}^n$ is *closed* if $\mathbb{R}^n - Q$ is open.

For a positive real number v,

$$S_n(v) = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : -v \leq x_k \leq v \text{ for each } k\}.$$

Definition. A subset Q of $\mathbb{R}^n$ is *bounded* if $Q \subset S_n(v)$ for some v.

It is left to the reader to show that $S_n(v)$ is a closed subset of $\mathbb{R}^n$.

Definition. A collection of sets Λ *covers a set* K if $K \subset \bigcup_\Lambda V$.

Definition. A subset K of $\mathbb{R}^n$ is *compact* if for each collection Λ of open sets which covers K, there is a finite $\Lambda^* \subset \Lambda$ which covers K.

Remark 2. Each compact subset of $\mathbb{R}^n$ is closed and bounded.

Remark 3. Each closed subset of a compact subset of $\mathbb{R}^n$ is compact.

Remark 4. If K is a compact subset of $\mathbb{R}^n$ and g is a continuous function from K to $\mathbb{R}^{n+1}$, then $g(K)$ is a compact subset of $\mathbb{R}^{n+1}$.

Theorem 4. If $S_n(v)$ is a compact subset of $\mathbb{R}^n$ for each positive real number v, then $S_{n+1}(v)$ is a compact subset of $\mathbb{R}^{n+1}$ for each positive real number v.

Proof. Let Λ be a collection of open subsets of $\mathbb{R}^{n+1}$ which covers $S_{n+1}(v)$, and for each $y \in [-v, v]$ let $L(y) = \{(x_1, \dots, x_n, x_{n+1}) \in S_{n+1}(v) : x_{n+1} = y\}$. Then

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : \{(x_1, \dots, x_n, x_{n+1}) \in L(y)\} = S_n(v)$$

and it is easily seen that the function $g\colon S_n(v) \to L(y)$ defined by

$$g(x_1, \dots, x_n) = (x_1, \dots, x_n, y)$$

is onto and continuous. Let

$$A = \{r \in (-v, v] : \text{some finite } \Gamma \subset \Lambda \text{ covers } \bigcup_{y \le r} L(y)\}.$$

Since $L(-v)$ is compact choose a finite collection of $n+1$–dimensional symmetric open intervals which covers $L(-v)$ such that the centers are in $L(-v)$ and such that each is a subset of some member of Λ. Let r be one–half of the smallest edge length among the intervals. Then $r \in A$. Let $c = \sup A$. Again we observe that $L(c)$ is compact and, as with the argument used with $L(-v)$, there is an r such that $-v < r < c$ and such that some finite collection of elements of Λ covers $\bigcup_{r \le y \le c} L(y)$. Moreover, some finite subset of Λ covers $\bigcup_{y \le r} L(y)$ so $c \in A$. If $c < v$, then any finite collection of elements of Λ which covers $\bigcup_{y \le c} L(y)$ also covers $\bigcup_{y \le r} L(y)$ for some $c < r < v$, contradicting a property of supremum.

Theorem 5 (Heine–Borel). Every closed and bounded subset of $\mathbb{R}^n$ is compact.

Proof. Every closed and bounded subset of $\mathbb{R}$ is compact [4]. Suppose every closed and bounded subset of $\mathbb{R}^n$ is compact. Then $S_n(v)$ is a compact subset of $\mathbb{R}^n$ for each positive real number v. From the last theorem, $S_{n+1}(v)$ is a compact subset of $\mathbb{R}^{n+1}$ for each real number v. If $K \subset \mathbb{R}^{n+1}$ is bounded, then $K \subset S_{n+1}(v)$ for some v. If K is also closed, then it is a closed subset of a compact set and is therefore compact.

Induction Proof of the Cauchy–Schwarz Inequality for Complex Numbers.
The inequality

$$\left|\sum_{k=1}^{n} \overline{z}_k w_k\right| \le \left(\sum_{k=1}^{n} |z_k|^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} |w_k|^2\right)^{\frac{1}{2}}$$

holds for all complex numbers $z_1, \dots, z_n, w_1, \dots, w_n$. Equality holds if and only if $z_j w_m = z_m w_j$ for $1 \le m < j \le n$.

Proof. The proofs for the cases $n = 1$ and $n = 2$ are left to the reader. Suppose $n > 2$ is an integer and that the statement holds for $1 \le j \le n$. Then for all sequences of complex numbers

$$z_1, z_2, \dots, z_n, z_{n+1}, w_1, w_2, \dots, w_n, w_{n+1},$$

it follows for $1 \le m < j \le n+1$ that

$$\left|\sum_{k=1}^{n+1} \overline{z}_k w_k\right| \le \left|\sum_{k \notin \{j,m\}} \overline{z}_k w_k\right| + |\overline{z}_j w_j + \overline{z}_m w_m|$$

$$\le \left|\sum_{k \notin \{j,m\}} \overline{z}_k w_k\right| + \sqrt{|z_j|^2 + |z_m|^2}\sqrt{|w_j|^2 + |w_m|^2}$$

$$\le \sum_{k \notin \{j,m\}} |z_k||w_k| + \sqrt{|z_j|^2 + |z_m|^2}\sqrt{|w_j|^2 + |w_m|^2}$$

$$\le \left(\sum_{k \notin \{j,m\}} |z_k|^2 + \left(\sqrt{|z_j|^2 + |z_m|^2}\right)^2\right)^{\frac{1}{2}} \left(\sum_{k \notin \{j,m\}} |w_k|^2 + \left(\sqrt{|w_j|^2 + |w_m|^2}\right)^2\right)^{\frac{1}{2}}$$

$$= \left(\sum_{k=1}^{n+1} |z_k|^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n+1} |w_k|^2\right)^{\frac{1}{2}}.$$

If

$$\left|\sum_{k=1}^{n+1} \overline{z}_k w_k\right| = \left(\sum_{k=1}^{n+1} |z_k|^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n+1} |w_k|^2\right)^{\frac{1}{2}}$$

it follows from the above inequalities that

$$|\overline{z}_j w_j + \overline{z}_m w_m| = \sqrt{|z_j|^2 + |z_m|^2}\sqrt{|w_j|^2 + |w_m|^2}$$

and from the case for $n = 2$ that $z_j w_m = z_m w_j$.

If $n > 2$ and $z_1, z_2, \ldots, z_n, w_1, w_2, \ldots, w_n$, satisfy $z_m w_k = z_k w_m$ for $k \neq m$, then

$$\begin{aligned}\sum_{k=1}^{n} |z_k|^2 \sum_{k=1}^{n} |w_k|^2 - \left|\sum_{k=1}^{n} \overline{z}_k w_k\right|^2 &= \sum_{k\neq m} |z_k|^2 |w_m|^2 - \sum_{k\neq m} \overline{z}_k w_k z_m \overline{w}_m \\ &= \sum_{k\neq m} |z_k|^2 |w_m|^2 - \sum_{k\neq m} |z_k|^2 |w_m|^2 = 0.\end{aligned}$$

<u>The Lagrange Identity (complex numbers)</u>. If $z_1, z_2, \ldots, z_n, w_1, w_2, \ldots, w_n$ are complex numbers then

$$\left|\sum_{k=1}^{n} \overline{z}_k w_k\right|^2 = \sum_{k=1}^{n} |z_k|^2 \sum_{k=1}^{n} |w_k|^2 - \sum_{1\le k<j\le n} |z_k w_j - z_j w_k|^2.$$

<u>Proof</u>.

$$\begin{aligned}\left|\sum_{k=1}^{n} \overline{z}_k w_k\right|^2 &= \sum_{k=1}^{n} \overline{z}_k w_k \overline{\sum_{k=1}^{n} \overline{z}_k w_k} \\ &= \sum_{j=k} \overline{z}_j w_j z_k \overline{w}_k + \sum_{k<j} \overline{z}_j w_j z_k \overline{w}_k + \sum_{k>j} \overline{z}_j w_j z_k \overline{w}_k \\ &= \sum_{k=1}^{n} |z_k|^2 |w_k|^2 + \sum_{k<j} \overline{z}_j w_j z_k \overline{w}_k + \sum_{k<j} \overline{z}_k w_k z_j \overline{w}_j \\ &= \sum_{k=1}^{n} |z_k|^2 \sum_{k=1}^{n} |w_k|^2 - \sum_{1\le k<j\le n} |z_k w_j - z_j w_k|^2.\end{aligned}$$

The Cauchy–Schwarz Inequality for inner–product spaces is usually established by considering the sign of a special quadratic polynomial, "a clever trick that is not easy to motivate" [1]. The foundation for an interesting and easily motivated proof of the inequality for inner–product spaces may be found in [5] and derives from an equation that may be viewed as a generalization of the Lagrange Identity. The generalization of this identity itself derives from a generalization of the Pythagorean Theorem. If $\vec{x}, \vec{y}$ are vectors in an inner–product space, represent the *inner–product* of $\vec{x}$ and $\vec{y}$ by $\vec{x} \cdot \vec{y}$ and the *norm* of $\vec{x}$ by $\|\vec{x}\|$. Then $\vec{x} \perp \vec{y}$ if and only if $\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$. It is easy to see that for $\vec{x}, \vec{y}$

$$(\|\vec{y}\|^2\vec{x} - (\vec{x}\cdot\vec{y})\vec{y}) \perp (\vec{x}\cdot\vec{y})\vec{y}$$

and that

$$(\|\vec{y}\|^2\vec{x} - (\vec{x}\cdot\vec{y})\vec{y}) + (\vec{x}\cdot\vec{y})\vec{y} = \|\vec{y}\|^2\vec{x}.$$

Hence,

$$\big\|\|\vec{y}\|^2\vec{x}\big\|^2 = \big\|\|\vec{y}\|^2\vec{x} - (\vec{x}\cdot\vec{y})\vec{y}\big\|^2 + \|(\vec{x}\cdot\vec{y})\vec{y}\|^2$$

$$|\vec{x}\cdot\vec{y}|^2\|\vec{y}\|^2 = \|\vec{y}\|^4\|\vec{x}\|^2 - \big\|\|\vec{y}\|^2\vec{x} - (\vec{x}\cdot\vec{y})\vec{y}\big\|^2.$$

References

1. M. Anton and C. Rorres, *Linear Algebra with Applications*, Wiley, New York, 1987.
2. T. A. Apostol, *Mathematical Analysis*, 2nd ed., Addison–Wesley, London, 1974.
3. F. Dubeau, "Cauchy–Bunyakowski–Schwarz Inequality Revisited," *American Mathematical Monthly*, 97 (1990), 419–421.
4. E. Gaughan, *Introduction to Analysis*, 5th ed., Brooks/Cole, Pacific Grove, California, 1998.
5. S. Lang, *A First Course in Calculus*, 5th ed., Springer–Verlag, New York, 1986.
6. R. M. F. Moss and G. T. Roberts, "A Creeping Lemma," *American Mathematical Monthly*, 75 (1968), 649–652.
7. W. Rudin, *Principles of Mathematical Analysis*, 2nd ed., McGraw–Hill, New York, 1996.

8. S. So, "Sequences, Mathematical Induction, and the Heine–Borel Theorem," *Missouri Journal of Mathematical Sciences*, 3 (1991), 84–90.

9. M. Spivak, *Calculus*, 2nd ed., Publish or Perish, Inc., Houston, Texas, 1980.

Richard E. Bayne
Department of Mathematics
Howard University
Washington, D.C. 20059
email: bayne@scs.howard.edu

James E. Joseph
Department of Mathematics
Howard University
Washington, D.C. 20059
email: jjoseph@fac.howard.edu

Myung H. Kwack
Department of Mathematics
Howard University
Washington, D.C. 20059
email: mkwack@howard.edu

Thomas H. Lawson
Center for Academic Reinforcement
Howard University
Washington, D.C. 20059
email: tlawson@howard.edu