THE INSTRUCTOR'S GUIDE TO REAL INDUCTION

PETE L. CLARK

Contents

1. Introduction to Real Induction

1.1. "Induction is fundamentally discrete..."

Real induction is (my name for) a proof technique in elementary real analysis which is inspired by the principle of mathematical induction in the natural numbers. In broadest terms, the idea is to show that a statement holds for all real numbers in an interval by "pushing from left to right".

In many circles it seems to be a truism that such a thing is not possible: how many times have you heard – or said! – that induction is fundamentally "discrete" and that, alas, inductive methods are not available for "continuous variables"? I held this idea myself until rather recently.

1.2. ...is dead wrong!

Remarkably, this "induction is fundamentally discrete" idea has been refuted repeatedly in the literature, going back at least 93 years! The earliest instance I know of is Chao's 1919 "A note on 'continuous mathematical induction"' [Ch19].

Theorem 1. *(Chao) Let $a \in \mathbb{R}$, and let $S \subset \mathbb{R}$. Suppose that:*
(CI1) $a \in \mathbb{R}$.
(CI2) Suppose $\exists \Delta > 0$ such that $\forall x \in \mathbb{R}$, $x \in S \implies (x - \Delta, x + \Delta) \cap [a, \infty) \subset S$.
Then $[a, \infty) \subset S$.

Since this is not the formulation I want to discuss, I leave the proof to you, along with Chao's remark that the Archimedean nature of the ordering of $\mathbb{R}$ is being critically used here. This Archimedean feature is absent in later formulations.

But this is just the first – if it actually is the first – of many similar formulations of "continuous induction". A literature search turned up the following papers, each of which introduces some form of "continuous induction", in many cases with no reference to past precedent: [Kh23], [Pe26], [Kh49], [Du57], [Fo57], [MR68], [Sh72], [Be82], [Le82], [Sa84], [Do03], [Ka07], [Ha11].

1.3. Real Induction.

Consider for a moment "conventional" mathematical induction. To use it, one thinks in terms of *predicates* – i.e., statements $P(n)$ indexed by the natural numbers – but the cleanest enunication comes from thinking in terms of *subsets* of $\mathbb{N}$. The same goes for real induction.[1]

Let $a < b$ be real numbers. We define a subset $S \subset [a, b]$ to be **inductive** if:

(RI1) $a \in S$.
(RI2) If $a \leq x < b$, then $x \in S \implies [x, y] \subset S$ for some $y > x$.
(RI3) If $a < x \leq b$ and $[a, x) \subset S$, then $x \in S$.

Theorem 2. *(Real Induction) For a subset $S \subset [a, b]$, the following are equivalent:*
(i) S is inductive.
(ii) $S = [a, b]$.

Proof. (i) $\implies$ (ii): let $S \subset [a, b]$ be inductive. Seeking a contradiction, suppose $S' = [a, b] \setminus S$ is nonempty, so $\inf S'$ exists and is finite.
Case 1: $\inf S' = a$. Then by (RI1), $a \in S$, so by (RI2), there exists $y > a$ such that $[a, y] \subset S$, and thus y is a greater lower bound for S' then $a = \inf S'$: contradiction.
Case 2: $a < \inf S' \in S$. If $\inf S' = b$, then $S = [a, b]$. Otherwise, by (RI2) there exists $y > \inf S'$ such that $[\inf S', y'] \subset S$, contradicting the definition of $\inf S'$.
Case 3: $a < \inf S' \in S'$. Then $[a, \inf S') \subset S$, so by (RI3) $\inf S' \in S$: contradiction!
(ii) $\implies$ (i) is immediate. □

Theorem 2 is due to D. Hathaway [Ha11] and, independently, to me. But mathematically equivalent ideas have been around in the literature for a long time, some of which are *much closer* to our formulation than the one of Chao given above. Especially, I acknowledge my indebtedness to [Ka07]. I read this paper early in the

---

[1] Not everyone subscribes to this viewpoint: Chao phrases Theorem 1 in terms of predicates.

morning of Tuesday, September 7, 2010 and found it fascinating. Kalantari's formulation works with subsets $S \subset [a, b)$, replaces (RI2) and (RI3) by the single axiom

(RIK) For $x \in [a, b)$, if $[a, x) \subset S$, then there exists $y > x$ with $[a, y) \subset S$,[2]

and the conclusion is that a subset $S \subset [a, b)$ satisfying (RI1) and (RIK) must be equal to $[a, b)$. Unfortunately I was a bit confused by Kalantari's formulation, and I wrote to Professor Kalantari suggesting the "fix" of replacing (RIK) with (RI2) and (RI3). He wrote back later that morning to set me straight. I was scheduled to give a general interest talk for graduate students in the early afternoon, and I had planned to speak about binary quadratic forms. But I found real induction to be too intriguing to put down, and my talk at 2 pm that day was on real induction (in the formulation of Theorem 2). This was, perhaps, the best received non-research lecture I have ever given, and by Wednesday, September 8, 2010 I made and posted a written version, of which the present document is a direct descendant.

In 2011 D. Hathaway published a short note "Using Continuity Induction" [Ha11] giving an all but identical formulation: instead of (RI2), he takes

(RI2H) If $a \leq x < b$, then $x \in S \implies [x, x + \delta) \subset S$ for some $\delta > 0$.

This formulation is not only mathematically equivalent to mine but "pedagogically equivalent": if your students cannot readily see the equivalence of (RI2) and (RI2H), they are not ready for real induction in any form. It seems that Hathaway and I arrived at our formulations completely independently. Moreover, when first formulating real induction I too used (RI2H), but soon changed it to (RI2) with an eye to certain more general forms of induction.

The present document may be viewed in part as a more expansive version of Hathaway's note, which provides a pedgagogical push towards use of real induction in the classroom. He writes "Continuity induction has been rediscovered regularly (see the historical notes at the end of this Capsule). With each reappearance, the concept has been refined so that with this version we hope that it may finally be suitable for students in college-level analysis classes." I agree, and I want to lend my voice to this demonstration.

2. Variations on the theme of Real Induction

2.1. **Recovering Unbounded Interval Induction.**

Theorem 1 concerns unbounded above subsets, whereas Theorem 2 concerns subsets of bounded intervals. But it is easy to modify it to work with subsets of $[a, \infty)$: just omit reference to b in (RI2) and (RI3) above, i.e., replace them with:

(RI2′) If $a \leq x$, then $x \in S \implies [x, y] \subset S$ for some $y > x$.
(RI3′) If $a < x$ and $[a, x) \subset S$, then $x \in S$.

The proof that an inductive subset of $[a, \infty)$ must in fact be $[a, \infty)$ is unchanged.

[2] One also needs the convention $[x, x) = \{x\}$ here.

2.2. **Inductive subsets of $\mathbb{R}$.**

It is also natural to wonder whether there is an induction-like criterion for subsets of $\mathbb{R}$. This is indeed possible, in fact in several different ways.

First of all there is a **downward Real Induction** for intervals $[a, b]$ or $(-\infty, b]$ in which we "push from right to left" – its existence may be justified by the remark that $x \mapsto -x$ is an isomorphism from $\mathbb{R}$ to its order dual. In this way one can prove that a subset $S \subset \mathbb{R}$ is all of $\mathbb{R}$ by choosing $a \in \mathbb{R}$ and using "upward" induction to show that $S \cap [a, \infty) = [a, \infty)$ and $S \cap (-\infty, b] = (-\infty, b]$.

More ambitiously, let $(X, <)$ be an arbitrary linearly ordered set. We define a subset $S \subset X$ to be **inductive** if it satisfies all of the following:

(IS1) There exists $a \in X$ such that $(-\infty, a] \subset S$.
(IS2) For all $x \in S$, either $x = 1$ or there exists $y > x$ such that $[x, y] \subset S$.
(IS3) For all $x \in S$, if $(-\infty, x) \in S$, then $x \in S$.

Then essentially the same argument as in the proof of Theorem 2 shows that if X is any interval in $\mathbb{R}$, the only inductive subset of X is X itself.

2.3. **A Principle of (Linearly) Ordered Induction.**

Even more is true. A linearly ordered set $(X, <)$ is **Dedekind complete** if every nonempty subset $S \subset X$ which is bounded above has a least upper bound.

Theorem 3. *(Principle of Ordered Induction) For a linearly ordered set X, TFAE:*
(i) X is Dedekind complete.
(ii) The only inductive subset of X is X itself.

Proof. This is [Cl11, Thm. 1]. □

We have omitted the proof not because it is difficult (it isn't), but because it would be a digression from our intended topics of elementary real analysis and associated pedagogy. However, we will observe – with proof! – the following relationship between induction and Dedekind completeness in ordered fields.

Theorem 4. *In an ordered field F, the following are equivalent:*
(i) F is Dedekind complete: every nonempty bounded above subset has a supremum.
(ii) F satisfies the Principle of Real Induction: for all $a < b \in F$, a subset $S \subset [a, b]$ satisfying (RI1), (RI2) and (RI3) above must be $[a, b]$.

Proof. (i) $\implies$ (ii): This is simply a restatement of Theorem 2.
(ii) $\implies$ (i): Let $T \subset F$ be nonempty and bounded below by $a \in F$. We will show that T has an infimum. For this, let S be the set of lower bounds m of T with $a \leq m$. Let b be any element of T. Then $S \subset [a, b]$.
Step 1: Observe that $b \in S \iff b = \inf T$. In general the infimum could be smaller, so our strategy is not exactly to use real induction to prove $S = [a, b]$. Nevertheless we claim that S satisfies (RI1) and (RI3).
(RI1): Since a is a lower bound of T with $a \leq a$, we have $a \in S$.
(RI3): Suppose $x \in (a, b]$ and $[a, x) \subset S$, so every $y \in [a, x)$ is a lower bound for T. Then x is a lower bound for T: if not, there exists $t \in T$ such that $t < x$; taking any $y \in (t, x)$, we get that y is not a lower bound for T either, a contradiction.

Step 2: Since F satisfies the Principle of Real Induction, by Step 1 $S = [a, b]$ iff S satisfies (RI2). If $S = [a, b]$, then the element $b \in$ is a lower bound for T, so it must be the infimum of T. Now suppose that $S \neq [a, b]$, so by Step 1 S *does not* satisfy (RI2): there exists $x \in S$, $x < b$ such that for any $y > x$, there exists $z \in (x, y)$ such that $z \notin S$, i.e., z is not a lower bound for T. In other words x is a lower bound for T and no element larger than x is a lower bound for T, so $x = \inf T$. □

Problem 1. *Characterize the inductive subsets of* $I_{\mathbb{Q}} = [0, 1] \cap \mathbb{Q}$.

For instance, for any irrational $\alpha \in (0, 1)$, $[0, \alpha] \cap \mathbb{Q}$ is a proper inductive subset of $I_{\mathbb{Q}}$, suggesting perhaps some relationship to Dedekind cuts. It is clear, however, that there are many more inductive subsets than this.

2.4. **Strong Real Induction.**

Just as there exists a "strong" version of mathematical induction which replaces the inductive hypothesis

$$\forall n \in \mathbb{N}, n \in S \implies n + 1 \in S$$

with

$$\forall n \in \mathbb{N}, [0, n) \in S \implies n \in S,$$

there is a corresponding "strong" formulation of real induction. Let us set it up in terms of the following concepts: a subset S of a linearly ordered set X is **downward closed** if $\forall x, y \in X$, $x \in S$, $y \leq x \implies y \in S$. The intersection of any family of downward closed subsets is again downward closed, and thus every subset S of X contains a unique minimal downward closed set, namely the intersection of all downward closed subsets containing S. This subset will be called the **downward closure** of S and written $\underline{S}$. Note that we have

$$\underline{S} = \{x \in X \mid \exists s \in S \text{ such that } x \leq s\},$$

an often more useful characterization of $\underline{S}$ than the one involving intersections.

Theorem 5. *(Principle of Strong Real Induction) Let* $S \subset [a, b]$*, and suppose:*
(RI1) $a \in S$.
(SRI2) For all $x \in [a, b)$, $[a, x] \subset \underline{S} \implies \exists y \in \underline{S}$ *with* $x < y$.
(SRI3) For all $x \in (a, b]$, $[a, x) \subset \underline{S} \implies x \in S$.
Then $b \in S$.

Proof. The hypotheses imply that $\underline{S}$ is an inductive subset, and thus by Theorem 2, $\underline{S} = [a, b]$. Since b is the largest element of $[a, b]$, $b \in \underline{S} \iff b \in S$. □

Remark: The notion of a "principle of strong ordered induction" developed out of some conversations with Professor J.D. Hamkins concerning extensions of Theorem 4 to *partially* ordered sets in September of 2011. As an increasingly seasoned practitioner of real induction with a side interest in its order-theoretic ramifications and generalizations, I find Theorem 5 to be the most satisfactory and pleasant inductive principle enunciated thus far.

However, as someone who has taught real induction to real undergraduates, I *do not recommend* Theorem 5 for pedagogical use: the circle of ideas surrounding real induction and its uses is already sufficiently rich. *After* students have seen some real induction proofs in action, Theorem 5 could make a nice optional exercise.

I do think the *instructor* should beware of the fact that the many of the most standard Real Induction proofs are *de facto* employing Strong Real Induction. In particular, here are two features of many – but not all – Real Induction Proofs:

- The desired result follows from $b \in S$, making passage from $\underline{S}$ to S natural.
- The "natural inductive subset" S is *a priori* downward closed.

When either of these occurs, the instructor should be aware that showing $[x, x+\delta] \subset S$ is logically equivalent to showing $x+\delta \in S$. Conversely, sometimes the two statements are *not* logically equivalent, and one should address the minor complications necessary to show the former, e.g. by introducing $0 < \delta' \leq \delta$ and showing $x+\delta' \in S$. (Theorem 5 guarantees that these complications will be only notational.)

3. With our bare hands

Theorem 6. *(Intermediate Value Theorem (IVT)) Let $f : [a, b] \to \mathbb{R}$ be a continuous function, and let L be any number in between $f(a)$ and $f(b)$. Then there exists $c \in [a, b]$ such that $f(c) = L$.*

Proof. It is easy to reduce the theorem to the following special case: let $f : [a, b] \to \mathbb{R}$ be continuous and nowhere zero. If $f(a) > 0$, then $f(b) > 0$.

Let $S = \{x \in [a, b] \mid f(x) > 0\}$. Then $f(b) > 0$ iff $b \in S$. We will show $S = [a, b]$.
(RI1) By hypothesis, $f(a) > 0$, so $a \in S$.
(RI2) Let $x \in S$, $x < b$, so $f(x) > 0$. Since f is continuous at x, there exists $\delta > 0$ such that f is positive on $[x, x + \delta]$, and thus $[x, x + \delta] \subset S$.
(RI3) Let $x \in (a, b]$ be such that $[a, x) \subset S$, i.e., f is positive on $[a, x)$. We claim that $f(x) > 0$. Indeed, since $f(x) \neq 0$, the only other possibility is $f(x) < 0$, but if so, then by continuity there would exist $\delta > 0$ such that f is negative on $[x - \delta, x]$, i.e., f is both positive and negative at each point of $[x - \delta, x]$: contradiction! □

Now you may well be thinking that this proof is not so different from the standard proof of IVT using suprema. Let us recall how that goes: we may assume $f(a) < 0$, $f(b) > 0$, and our task is to find $c \in (a, b)$ with $f(c) = 0$. But let c be the least upper bound of the set of $x \in [a, b]$ such that $f(x) < 0$. Then using the key property that a continuous function which is positive (resp. negative) at a point must be positive (resp. negative) in some interval around that point, we see that $f(c) = 0$.

I agree completely: the Real Induction proof is not much more than a repackaging of the beautifully simple argument using suprema. In fact, in my recent Spivak Calculus class, I delayed discussion of least upper bounds (including Real Induction!) until the end of the first semester, and thus I stated and used IVT long before I proved it. Towards the end of the course I introduced the least upper bound axiom *and on that same day* used it to prove IVT. Then later I introduced Real Induction and gave the above proof. I think it's nice to introduce a new concept in a somewhat familiar context.

The proofs of the other three Interval Theorems are far from straightforward to construct...unless you've seen them before. The chief pedagogical merit of Real Induction is similar to the chief pedagogical merit of mathematical induction: it focuses students' efforts and gives them *general training* about how to get started on certain types of mathematical proofs. In the first examples of mathematical induction the statement itself is of the form "For all $n \in \mathbb{N}$, $P(n)$ holds", so it is clear what the induction hypothesis should be. However, mathematical induction is much more flexible and powerful than this once one learns to *try to find a statement* $P(n)$ whose truth for all n will give the desired result. She who develops skill at "finding the induction hypothesis" acquires a formidable mathematical weapon: for instance the Arithmetic-Geometric Mean Inequality, the Fundamental Theorem of Arithmetic, and the Law of Quadratic Reciprocity have all been proved in this way; in the last case, the first proof given (by Gauss) was by induction.

Similarly, to get a Real Induction proof properly underway, we need to find a subset $S \subset [a,b]$ for which the conclusion $S = [a,b]$ gives us the result we want, and for which our given hypotheses are suitable for "pushing from left to right". If we can find the right set S then we are, quite often, more than halfway there: the rest may take a little while to write out but is relatively straightforward to produce.

We now give Real Induction proofs of the remaining three interval theorems.

Theorem 7. *(Extreme Value Theorem (EVT))*
Let $f : [a,b] \to \mathbb{R}$ be continuous. Then:
a) f is bounded.
b) f attains a minimum and maximum value.

Proof. a) Let $S = \{x \in [a,b] \mid f : [a,x] \to \mathbb{R} \text{ is bounded}\}$.
(RI1): Evidently $a \in S$.
(RI2): Suppose $x \in S$, so that f is bounded on $[a,x]$. But then f is continuous at x, so is bounded near x: for instance, there exists $\delta > 0$ such that for all $y \in [x-\delta, x+\delta]$, $|f(y)| \leq |f(x)|+1$. So f is bounded on $[a,x]$ and also on $[x, x+\delta]$ and thus on $[a, x+\delta]$.
(RI3): Suppose $x \in (a,b]$ and $[a,x) \subset S$. Now **beware**: this *does not say* that f is bounded on $[a,x)$: rather it says that for all $a \leq y < x$, f is bounded on $[a,y]$. These are different statements: for instance, $f(x) = \frac{1}{x-2}$ is bounded on $[0,y]$ for all $y < 2$ but it is not bounded on $[0,2)$. But of course this f is not continuous at 2. So we can proceed almost exactly as we did above: since f is continuous at x, there exists $0 < \delta < x-a$ such that f is bounded on $[x-\delta, x]$. But since $a < x-\delta < x$ we know f is bounded on $[a, x-\delta]$, so f is bounded on $[a,x]$.
b) Let $m = \inf f([a,b])$ and $M = \sup f([a,b])$. By part a) we have

$$-\infty < m \leq M < \infty.$$

We want to show that there exist $x_m, x_M \in [a,b]$ such that $f(x_m) = m$, $f(x_M) = M$, i.e., that the infimum and supremum are actually attained as values of f. Suppose that there does not exist $x \in [a,b]$ with $f(x) = m$: then $f(x) > m$ for all $x \in [a,b]$ and the function $g_m : [a,b] \to \mathbb{R}$ by $g_m(x) = \frac{1}{f(x)-m}$ is defined and continuous. By the result of part a), g_m is bounded, but this is absurd: by definition of the infimum, $f(x) - m$ takes values less than $\frac{1}{n}$ for any $n \in \mathbb{Z}+$ and thus g_m takes values greater than n for any $n \in \mathbb{Z}^+$ and is accordingly unbounded. So indeed there must exist $x_m \in [a,b]$ such that $f(x_m) = m$. Similarly, assuming that $f(x) < M$ for all $x \in [a,b]$ gives rise to an unbounded continuous function $g_M : [a,b] \to \mathbb{R}$, $x \mapsto \frac{1}{M-f(x)}$, contradicting part a). So there exists $x_M \in [a,b]$ with $f(x_M) = M$. $\square$

Let $f : I \to \mathbb{R}$. For $\epsilon, \delta > 0$, let us say that f is (ϵ,δ)-UC on I if for all $x_1, x_2 \in I$, $|x_1 - x_2| < \delta \implies |f(x_1) - f(x_2)| < \epsilon$. This is a sort of halfway unpacking of the definition of uniform continuity. More precisely, $f : I \to \mathbb{R}$ is uniformly continuous iff for all $\epsilon > 0$, there exists $\delta > 0$ such that f is (ϵ, δ)-UC on I.

Lemma 8. *(Covering Lemma) Let $a < b < c < d$ be real numbers, and let $f : [a, d] \to \mathbb{R}$. Suppose that for real numbers $\epsilon_1, \delta_1, \delta_2 > 0$,*
- *f is (ϵ, δ_1)-UC on $[a, c]$ and*
- *f is (ϵ, δ_2)-UC on $[b, d]$.*

Then f is $(\epsilon, \min(\delta_1, \delta_2, c - b))$-UC on $[a, b]$.

Proof. Suppose $x_1 < x_2 \in I$ are such that $|x_1 - x_2| < \delta$. Then it cannot be the case that both $x_1 < b$ and $c < x_2$: if so, $x_2 - x_1 > c - b \geq \delta$. Thus we must have either that $b \leq x_1 < x_2$ or $x_1 < x_2 \leq c$. If $b \leq x_1 < x_2$, then $x_1, x_2 \in [b, d]$ and $|x_1 - x_2| < \delta \leq \delta_2$, so $|f(x_1) - f(x_2)| < \epsilon$. Similarly, if $x_1 < x_2 \leq c$, then $x_1, x_2 \in [a, c]$ and $|x_1 - x_2| < \delta \leq \delta_1$, so $|f(x_1) - f(x_2)| < \epsilon$. □

Theorem 9. *(Uniform Continuity Theorem) Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is uniformly continuous on $[a, b]$.*

Proof. For $\epsilon > 0$, let $S(\epsilon)$ be the set of $x \in [a, b]$ such that there exists $\delta > 0$ such that f is (ϵ, δ)-UC on $[a, x]$. To show that f is uniformly continuous on $[a, b]$, it suffices to show that $S(\epsilon) = [a, b]$ for all $\epsilon > 0$. We will show this by Real Induction.
(RI1): Trivially $a \in S(\epsilon)$: f is (ϵ, δ)-UC on $[a, a]$ for all $\delta > 0$!
(RI2): Suppose $x \in S(\epsilon)$, so there exists $\delta_1 > 0$ such that f is (ϵ, δ_1)-UC on $[a, x]$. Moreover, since f is continuous at x, there exists $\delta_2 > 0$ such that for all $c \in [x, x+\delta_2]$, $|f(c) - f(x)| < \frac{\epsilon}{2}$. Why $\frac{\epsilon}{2}$? Because then for all $c_1, c_2 \in [x - \delta_2, x + \delta_2]$,

$$|f(c_1) - f(c_2)| = |f(c_1) - f(x) + f(x) - f(c_2)| \leq |f(c_1) - f(x)| + |f(c_2) - f(x)| < \epsilon.$$

In other words, f is (ϵ, δ_2)-UC on $[x - \delta_2, x + \delta_2]$. We apply the Covering Lemma to f with $a < x - \delta_2 < x < x + \delta_2$ to conclude that f is $(\epsilon, \min(\delta, \delta_2, x - (x - \delta_2))) = (\epsilon, \min(\delta_1, \delta_2))$-UC on $[a, x + \delta_2]$. It follows that $[x, x + \delta_2] \subset S(\epsilon)$.
(RI3): Suppose $[a, x) \subset S(\epsilon)$. As above, since f is continuous at x, there exists $\delta_1 > 0$ such that f is (ϵ, δ_1)-UC on $[x - \delta_1, x]$. Since $x - \frac{\delta_1}{2} < x$, by hypothesis there exists δ_2 such that f is (ϵ, δ_2)-UC on $[a, x - \frac{\delta_1}{2}]$. We apply the Covering Lemma to f with $a < x - \delta_1 < x - \frac{\delta_1}{2} < x$ to conclude that f is $(\epsilon, \min(\delta_1, \delta_2, x - \frac{\delta_1}{2} - (x - \delta_1))) = (\epsilon, \min(\frac{\delta_1}{2}, \delta_2))$-UC on $[a, x]$. Thus $x \in S(\epsilon)$. □

Theorem 10. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function on a closed bounded interval. Then f is Darboux integrable.*

Proof. We will use Darboux's Integrability Criterion: we must show that for all $\epsilon > 0$, there exists a partition $\mathcal{P}$ of $[a, b]$ such that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon$. It is convenient to prove instead the following equivalent statement: for every $\epsilon > 0$, there exists a partion $\mathcal{P}$ of $[a, b]$ such that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < (b - a)\epsilon$.

Fix $\epsilon > 0$, and let $S(\epsilon)$ be the set of $x \in [a, b]$ such that there exists a partition $\mathcal{P}_x$ of $[a, b]$ with $U(f, \mathcal{P}_x) - L(f, \mathcal{P}_x) < \epsilon$. We want to show $b \in S(\epsilon)$, so it suffices to show $S(\epsilon) = [a, b]$. In fact it is necessary and sufficient: observe that if $x \in S(\epsilon)$ and $a \leq y \leq x$, then also $y \in S(\epsilon)$. We will show $S(\epsilon) = [a, b]$ by Real Induction.
(RI1) The only partition of $[a, a]$ is $\mathcal{P}_a = \{a\}$, and for this partition we have $U(f, \mathcal{P}_a) = L(f, \mathcal{P}_a) = f(a) \cdot 0 = 0$, so $U(f, \mathcal{P}_a) - L(f, \mathcal{P}_a) = 0 < \epsilon$.
(RI2) Suppose that for $x \in [a, b)$ we have $[a, x] \subset S(\epsilon)$. We must show that there is $\delta > 0$ such that $[a, x + \delta] \subset S(\epsilon)$, and by the above observation it is enough to find $\delta > 0$ such that $x + \delta \in S(\epsilon)$: we must find a partition $\mathcal{P}_{x+\delta}$ of $[a, x + \delta]$

such that $U(f, \mathcal{P}_{x+\delta}) - L(f, \mathcal{P}_{x+\delta}) < (x + \delta - a)\epsilon$). Since $x \in S(\epsilon)$, there is a partition $\mathcal{P}_x$ of $[a, x]$ with $U(f, \mathcal{P}_x) - L(f, \mathcal{P}_x) < (x - a)\epsilon$. Since f is continuous at x, we can make the difference between the maximum value and the minimum value of f as small as we want by taking a sufficiently small interval around x: i.e., there is $\delta > 0$ such that $\max(f, [x, x + \delta]) - \min(f, [x, x + \delta]) < \epsilon$. Now take the smallest partition of $[x, x + \delta]$, namely $\mathcal{P}' = \{x, x + \delta\}$. Then $U(f, \mathcal{P}') - L(f, \mathcal{P}') = (x+\delta-x)(\max(f, [x, x+\delta])-\min(f, [x, x+\delta])) < \delta\epsilon$. Thus if we put $\mathcal{P}_{x+\delta} = \mathcal{P}_x + \mathcal{P}'$ and use the fact that upper / lower sums add when split into subintervals, we have

$$U(f, \mathcal{P}_{x+\delta}) - L(f, \mathcal{P}_{x+\delta}) = U(f, \mathcal{P}_x) + U(f, \mathcal{P}') - L(f, \mathcal{P}_x) - L(f, \mathcal{P}')$$
$$= U(f, \mathcal{P}_x) - L(f, \mathcal{P}_x) + U(f, \mathcal{P}') - L(f, \mathcal{P}') < (x - a)\epsilon + \delta\epsilon = (x + \delta - a)\epsilon.$$

(RI3) Suppose that for $x \in (a, b]$ we have $[a, x) \subset S(\epsilon)$. We must show that $x \in S(\epsilon)$. The argument for this is the same as for (RI2) except we use the interval $[x - \delta, x]$ instead of $[x, x + \delta]$. Indeed: since f is continuous at x, there exists $\delta > 0$ such that $\max(f, [x-\delta, x])-\min(f, [x-\delta, x]) < \epsilon$. Since $x-\delta < x$, $x-\delta \in S(\epsilon)$ and thus there exists a partition $\mathcal{P}_{x-\delta}$ of $[a, x-\delta]$ such that $U(f, \mathcal{P}_{x-\delta}) = L(f, \mathcal{P}_{x-\delta}) = (x-\delta-a)\epsilon$. Let $\mathcal{P}' = \{x - \delta, x\}$ and let $\mathcal{P}_x = \mathcal{P}_{x-\delta} \cup \mathcal{P}'$. Then

$$U(f, \mathcal{P}_x) - L(f, \mathcal{P}_x) = U(f, \mathcal{P}_{x-\delta}) + U(f, \mathcal{P}') - (L(f, \mathcal{P}_{x-\delta}) + L(f, \mathcal{P}'))$$
$$= (U(f, \mathcal{P}_{x-\delta}) - L(f, \mathcal{P}_{x-\delta})) + \delta(\max(f, [x - \delta, x]) - \min(f, [x - \delta, x]))$$
$$< (x - \delta - a)\epsilon + \delta\epsilon = (x - a)\epsilon.$$

□

Theorem 11. *(Bolzano-Weierstrass) An infinite subset of* $[a, b]$ *has a limit point.*

Proof. Let $\mathcal{A}$ be an infinite subset of $[a, b]$, and let S be the set of x in $[a, b]$ such that *if* $\mathcal{A} \cap [a, x]$ is infinite, it has a limit point. It suffices to show that $S = [a, b]$, which we prove by Real Induction. As usual, (i) is trivial. Since $\mathcal{A} \cap [a, x)$ is finite iff $\mathcal{A} \cap [a, x]$ is finite, (iii) follows. As for (ii), suppose $x \in S$. If $\mathcal{A} \cap [a, x]$ is infinite, then by hypothesis it has a limit point and hence so does $[a, b]$. So we may assume that $\mathcal{A} \cap [a, x]$ is finite. Now either there exists $\delta > 0$ such that $\mathcal{A} \cap [a, x + \delta]$ is finite (in which case our inductive hypothesis is verified), or every interval $[x, x + \delta]$ contains infinitely many points of $\mathcal{A}$ in which case x itself is a limit point of $\mathcal{A}$. □

I have presented these proofs of Theorems 7, 9, 10, 11 in a "Spivak Calculus" class. Of these, the proof of Theorem 7 is the most palatable and provides a worthy showpiece of Real Induction.[3] The proof of Theorem 9 did not go nearly so well: it seems relentlessly technical from start to finish. In fact, the entire concept of uniform continuity is an exceptionally difficult one for students at this level to grasp.[4] I was warned of this by my colleague Ted Shifrin and also implicitly by Michael Spivak, in that Spivak's text relegates uniform continuity to an appendix, uses it only once later in the text – in the proof of Theorem 10 – and then later gives an alternate proof of Theorem 10 avoiding uniform continuity entirely.[5]

---

[3]In fact, now that you know about the method, please take a look at Spivak's proof of Theorem 7 [S, p. 135]. It is hard not to see it as a Real Induction proof in light disguise.

[4]I would welcome an extensive examination of students' difficulties with the uniform continuity concept. This is beyond the scope of both this article and my own pedagogical expertise, but I hope that others will take up this challenge.

[5]Spivak's uniform continuity-free proof of Theorem 10 [S, pp. 292-293] really is different from ours: inspired by the Fundamental Theorem of Calculus, he establishes equality of the upper and lower integrals by differentiation. This sort of proof goes back at least to M.J. Norris [No52].

The fact that we have provided a Real Induction proof for Theorem 10 is itself a recognition of the unpalatability of the proof of Theorem 9, for otherwise we would simply *use* Theorem 9 to prove Theorem 10: see Proposition 21 below. In fact the students found the "direct" inductive proof of Theorem 10 rather appealing. On the page it looks lengthy, because of the notation of partitions and subintervals. If students are comfortable with such things, then the proof flows swiftly, with the repetition between (RI2) and (RI3) building confidence in the main idea.

To summarize: we mounted a frontal assault on the four **Interval Theorems** of elementary real analysis. We feel that Real Induction provides both a streamlined approach and conceptual assistance in constructing and following these proofs. Nevertheless, the proof of Theorem 9 comes out as rather technical.

Are there other approaches to these and similar core theorems of elementary analysis? Of course, and an excellent instructor will take these into account and create her own "composite approach" which works best for her own tastes and teaching style, the specific course and even the particular cohort of students being taught. True mastery of the subject enables an almost limitless flexibility.

In the reminder of this article I want to discuss other approaches to these and closely related theorems. Will the reader be entirely surprised to hear that Real Induction can be as useful (or more!) in these other approaches as it was above?

4. With the help of connectedness

Proposition 12. *Let $f : X \to Y$ be a surjective continuous map of topological spaces. If X is connected, then so is Y.*

Proof. We prove the contrapositive: if A_1, A_2 are two nonempty, disjoint open and closed subsets with $A_1 \cup A_2 = Y$, then $f^{-1}(A_1)$, $f^{-1}(A_2)$ are two nonempty, disjoint open and closed subsets of X with $f^{-1}(A_1) \cup f^{-1}(A_2) = X$. □

Theorem 13. *The interval $[a, b]$ is connected.*

Proof. Suppose $[a, b] = A \cup B$, with A and B open and closed, and $A \cap B = \emptyset$. We assume $a \in A$ and prove by Real Induction that $A = [a, b]$: (RI1) is immediate, (RI2) holds because A is open, and (RI3) holds because A is closed. We're done! □

Combining Proposition 25a) and Theorem 13 yields IVT (Theorem 6).

Theorem 14. *Let $(F, +, \cdot, <)$ be an ordered field. The following are equivalent:*
(i) F is Dedekind complete – i.e., $F \cong \mathbb{R}$.
(ii) Every closed interval $[a, b]$ of F is connected in the order topology.

Proof. We have just seen (i) $\implies$ (ii).
(ii) $\implies$ (i): This is a special case of [Cl11, Thm. 4]. □

Thus the connectedness of closed, bounded intervals carries the full force of Dedekind completeness. However, unlike many other equivalents of Dedekind completeness, this result seems hard to apply directly.

Problem 2. *Give proofs of Theorems 7, 9 and 10 using only Theorem 13, i.e., without recourse to completeness, Real Induction, and so forth.*

5. With the help of compactness

Proposition 15. *Let $f : X \to Y$ be a surjective continuous map of topological spaces. If X is compact, so is Y.*

Proof. Let $\{V_i\}_{i\in I}$ be an open cover of Y. For $i \in I$, put $U_i = f^{-1}(V_i)$. Then $\{U_i\}_{i\in I}$ is an open cover of X. Since X is compact, there is a finite $J \subset I$ such that $\bigcup_{i\in J} U_i = X$, and then $Y = f(X) = f(\bigcup_{i\in J} U_i) = \bigcup_{i\in J} f(U_i) = \bigcup_{i\in J} V_i$. □

Theorem 16. *(Heine-Borel) The interval $[a, b]$ is compact.*

Proof. For an open covering $\mathcal{U} = \{U_i\}_{i\in I}$ of $[a, b]$, let

$$S = \{x \in [a, b] \mid \mathcal{U} \cap [a, x] \text{ has a finite subcovering}\}.$$

We prove $S = [a, b]$ by Real Induction. (RI1) is clear. (RI2): If $U_1, \ldots, U_n$ covers $[a, x]$, then some U_i contains $[x, x+\delta]$ for some $\delta > 0$. (RI3): if $[a, x) \subset S$, let $i_x \in I$ be such that $x \in U_{i_x}$, and let $\delta > 0$ be such that $[x - \delta, x] \in U_{i_x}$. Since $x - \delta \in S$, there is a finite $J \subset I$ with $\bigcup_{i\in J} U_i \supset [a, x - \delta]$, so $\{U_i\}_{i\in J} \cup U_{i_x}$ covers $[a, x]$. □

Combining Proposition 25b) and Theorem 16 yields EVT (Theorem 7).

Proposition 17. *Heine-Borel implies Bolzano-Weierstrass.*

Proof. Let A be an infinite subset of $[a, b]$. Then A contains a countably infinite subset, so we may as well assume A is countably infinite, with enumeration $\{a_n\}_{n=1}^{\infty}$. We define a sequence of closed subintervals I_k of $[a, b]$ as follows: put $I_0 = [a, b]$. Now divide I_0 up into $[a, \frac{a+b}{2}] \cup [\frac{a+b}{2}, b]$ and observe that since A is infinite, its intersection with at least one of these subintervals must be infinite: call that subinterval I_1. Having defined $I_n = [a_n, b_n]$, we define I_{n+1} in the same way – by splitting I_n into $[a_n, \frac{a_n+b_n}{2}] \cup [\frac{a_n+b_n}{2}, b_n]$. In this way we get a nested sequence $\{I_n\}$ of closed subintervals, and by compactness there is $x \in \bigcap_{n=1}^{\infty} I_n$. By construction, every neighborhood of x contains I_n for all sufficiently large n and thus infinitely many points of A, so x is a limit point of A. □

As a preliminary to the next result, recall that for a nonempty subset A of a metric space X, the **diameter** of A is $\sup_{x,y\in A} d(x, y)$.

Lemma 18. *Let (X, d) be a compact metric space, and let $\mathcal{U} = \{U_i\}_{i\in I}$ be an open cover of X. Then $\mathcal{U}$ admits a **Lebesgue number**: a $\delta > 0$ such that for every nonempty $A \subset X$ of diameter less than δ, $A \subset U_i$ for at least one $i \in I$.*

Proof. For each $x \in X$, there exists an element of the cover, say U_x, with $x \in U_x$. Since U_x is open, there exists $\delta_x > 0$ such that $B(x, \delta_x) \subset U_x$. Now consider the open cover $\mathcal{U}' = \{B(x, \frac{\delta_x}{2})\}_{x\in X}$; by compactness, there is a finite subcover $X = \bigcup_{i=1}^{n} B(x_i, \frac{\delta_{x_i}}{2})$. We claim that $\delta = \min_{i=1}^{n} \frac{\delta_{x_i}}{2}$ is a Lebesgue number for $\mathcal{U}$. Indeed, let $A \subset X$ be of diameter less than δ. Let $a \in A$, and choose an i, $1 \le i \le n$, such that $a \in B(x_i, \frac{\delta_{x_i}}{2})$. We claim that then $A \subset U_{x_i}$. Indeed, if $b \in A$, then $d(b, a) < \delta \le \frac{\delta_{x_i}}{2}$, so

$$d(b, x_i) \le d(b, a) + d(a, x_i) < \frac{\delta_{x_i}}{2} + \frac{\delta_{x_i}}{2} = \delta_{x_i},$$

and thus $b \in B(x_i, \delta_{x_i}) \subset U_{x_i}$. □

Remark: One can prove by Real Induction that every open cover of $[a, b]$ admits a Lebesgue number: §9.2.

Proposition 19. *Let $f : X \to Y$ be a continuous map between metric spaces. For $\epsilon > 0$, suppose that the open cover $\mathcal{U}_\epsilon = \{f^{-1}(B(y, \frac{\epsilon}{2})\}_{y \in Y}$ of X admits a Lebesgue number δ. Then f is (ϵ, δ)-UC.*

Proof. If $d(x, x') < \delta$, $x, x' \in B(x, \delta)$. Since δ is a Lebesgue number for $\mathcal{U}_\epsilon$, there is $y \in Y$ such that $f(B(x, \delta)) \subset B(y, \frac{\epsilon}{2})$ and thus

$$d(f(x), f(x')) < d(f(x), y)) + d(y, f(x')) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Theorem 20. *Let X be a compact metric space and $f : X \to \mathbb{R}$ a continuous function. Then f is uniformly continuous.*

Proof. Let $\epsilon > 0$. By Lemma 18, the covering $\{f^{-1}(B(y, \frac{\epsilon}{2}))\}_{y \in \mathbb{R}}$ of $[a, b]$ has a Lebesgue number $\delta > 0$, and then by Proposition 19, f is (ϵ, δ)-UC. Thus f is uniformly continuous. □

Proposition 21. *UCT (Theorem 9) implies DIT (Theorem 10).*

Proof. Let $f : [a, b] \to \mathbb{R}$ be continuous, hence, by Theorem 9, uniformly continuous: for all $\epsilon > 0$, there is $\delta > 0$ such that for all $x_1, x_2 \in [a, b]$, $|x_1 - x_2| < \delta \implies |f(x_1) - f(x_2)| < \frac{\epsilon}{b-a}$. Let $n \in \mathbb{Z}^+$ be such that $\frac{b-a}{n} < \delta$, and let $\mathcal{P}_n$ be the partition of $[a, b]$ into n subintervals of equal length $\frac{b-a}{n}$. Then

(1)
$$U(f, \mathcal{P}_n) - L(f, \mathcal{P}_n) = \sum_{i=0}^{n-1} (M_i(f) - m_i(f)) \left(\frac{b-a}{n}\right) \leq \left(\frac{b-a}{n}\right) \sum_{i=0}^{n-1} M_i(f) - m_i(f).$$

For $0 \leq i < n - 1$, $m_i(f) = f(c_i)$ and $M_i(f) = f(d_i)$ for some $c_i, d_i \in [x_i, x_{i+1}]$. Thus $|c_i - d_i| \leq x_{i+1} - x_i = \frac{b-a}{n} < \delta$, so

$$|M_i(f) - m_i(f)| = |f(d_i) - f(c_i)| < \frac{\epsilon}{b-a}. \quad (2)$$

Combining (1) and (2) gives

$$U(f, \mathcal{P}_n) - L(f, \mathcal{P}_n) \leq \left(\frac{b-a}{n}\right) \sum_{i=0}^{n-1} (M_i(f) - m_i(f)) \leq \left(\frac{b-a}{n}\right) \sum_{i=0}^{n-1} \frac{\epsilon}{b-a} = \epsilon.$$

□

Theorem 22. *(Darboux) If $f : [a, b] \to \mathbb{R}$ is differentiable, then f' satisfies the Intermediate Value Property.*

Proof. It is sufficient to establish the following: suppose $f'(a) < 0$ and $f'(b) > 0$; then there is $c \in (a, b)$ such that $f'(c) = 0$. Since f is differentiable on $[a, b]$, it is continuous on $[a, b]$ and thus by Theorem 7 attains a minimum value. Since $f'(a) < 0$, the minimum cannot occur at a. Similarly, since $f'(b) > 0$, the minimum cannot occur at b. So the minimum must occur at $c \in (a, b)$, and thus $f'(c) = 0$. □

Proposition 23. *Theorem 22 plus DIT (Theorem 10) implies IVT (Theorem 6).*

Proof. Let $f : [a, b] \to \mathbb{R}$. By Theorem 10, for each $x \in [a, b]$, we may define $F(x) = \int_a^x f$. We CLAIM that F is differentiable and $F' = f$.
PROOF OF CLAIM Since f is continuous at c, for all $\epsilon > 0$, there exists $\delta > 0$ such that $|x - c| < \delta \implies |f(x) - f(c)| < \epsilon$, or equivalently

$$f(c) - \epsilon < f(x) < f(c) + \epsilon.$$

Therefore

$$f(c) - \epsilon = \frac{\int_c^x f(c) - \epsilon}{x - c} \leq \frac{\int_c^x f}{x - c} \leq \frac{\int_c^x f(c) + \epsilon}{x - c} = f(c) + \epsilon,$$

and thus

$$|\frac{\int_c^x f}{x - c} - f(c)| \leq \epsilon.$$

Thus $F'(c)$ exists and is equal to $f(c)$. Thus f is a derivative, hence by Theorem 22 has the Intermediate Value Property. □

Proposition 24. *The Intermediate Value Theorem implies connectedness of* $[a, b]$.

Proof. Indeed, suppose $[a, b] = A \cup B$, where A and B are each open and closed, and $A \cap B = \varnothing$. Then function $f : [a, b] \to \mathbb{R}$ which sends $x \in A \mapsto 0$ and $x \in B \mapsto 1$ is continuous but does not have the Intermediate Value Property. □

Thus compactness has **big teeth**: everything we have discussed follows swiftly from Heine-Borel! This suggests that the most efficient approach to this material would be to prove Theorems 13 and 16 by Real Induction as above and then deduce everything else from Theorem 16.

I would not recommend this for a course at the Spivak Calculus level: first of all, open coverings and subcoverings are a level of abstraction above what most students are ready for at that point in their studies. Second, I believe that students *should* learn this much hard analysis: pushing around some ϵ's and δ's builds character, and although many of these particular ϵ-δ arguments can be "softened" by recasting them in topological terms, this is certainly not the case for all ϵ-δ arguments that will be encountered in one's further study of mathematics.

However, in a topology course that spends some time reviewing / recasting elementary real analysis, I think this approach is appropriate and enlightening.

6. With the help of sequences

Theorem 25. *Let* $\{x_n\}_{n=1}^{\infty}$ *be a sequence of real numbers.*
a) If $\{x_n\}$ *is bounded, it has a convergent subsequence.*
b) If $\{x_n\}$ *is Cauchy, it converges.*

Proof. a) We may suppose that the map $n \mapsto x_n$ is finite to one, for otherwise there exists a constant subsequence. Let X be the image of the sequence, i.e., the set $\{x_n \mid n \in \mathbb{Z}^+\}$. By assumption X is an infinite subset of some closed interval $[a, b]$, so by Theorem 11 there exists a limit point x of X. In particular, for all $n \in \mathbb{Z}^+$, there exists $k \in \mathbb{Z}^+$ such that $|x_{n_k} - x| < \frac{1}{n}$. Thus $x_{n_k} \to x$.
b) A Cauchy sequence is bounded, hence by part a) admits a convergent subsequence, say $x_{n_k} \to x$. We claim that x is the limit of the full sequence x_n. Indeed,

for any $\epsilon > 0$, choose N sufficiently large so that for all $n, m \geq N$, $|x_n - x_m| < \frac{\epsilon}{2}$, and choose k such that $n_k \geq N$ and $|x_{n_k} - x| < \frac{\epsilon}{2}$. Then for all $n \geq N$,

$$|x_n - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Theorem 26. *Theorem 25 implies EVT (Theorem 7).*

Proof. Seeking a contradiction, let $f : [a, b] \to \mathbb{R}$ be an *unbounded* continuous function. Then for each $n \in \mathbb{Z}^+$ we may choose $x_n \in [a, b]$ such that $|f(x_n)| \geq n$. By Theorem 4.1, after passing to a subsequence (which, as usual, we will suppress from our notation) we may suppose that x_n converges, say to $\alpha \in [a, b]$. Since f is continuous, $f(x_n) \to f(\alpha)$, so in particular $\{f(x_n)\}$ is bounded...contradiction!
(With regard to the attainment of extrema, we have no improvement to offer on the simple argument using suprema / infima given in the proof of Theorem 7. □

Theorem 27. *Theorem 25 implies UCT (Theorem 9).*

Proof. Seeking a contradiction, let $f : [a, b] \to \mathbb{R}$ be continuous but *not* uniformly continuous. Then there is $\epsilon > 0$ such that for all $n \in \mathbb{Z}^+$, there are $x_n, y_n \in [a, b]$ with $|x_n - y_n| < \frac{1}{n}$ and $|f(x_n) - f(y_n)| \geq \epsilon$. By Theorem 25, after passing to a subsequence (notationally suppressed!) x_n converges to some $\alpha \in [a, b]$, and thus also $y_n \to \alpha$. Since f is continuous $f(x_n)$ and $f(y_n)$ both converge to $f(\alpha)$, hence for sufficiently large n, $|f(x_n) - f(y_n)| < \epsilon$: contradiction! □

The proof of Theorem 27 is significantly shorter and sweeter than the direct proof of Theorem 9 given above (and any other direct proof I know). It is also *more direct* than the approach using Lebesgue numbers given above. In a "sequences last" approach to real analysis (as e.g. in [S]) it may be appropriate to give the direct proof in the first semester to "build character" and then give this proof later when sequences are discussed, as a "reward" for making it that far in the course.

Above we showed that Heine-Borel implies Bolzano-Weierstrass. It is natural to wonder about the converse. Really we are comparing several variant notions of compactness in topological spaces. In general these notions relate to each other in complicated ways – e.g. the space $[0, 1]^{[0,1]}$ is compact in the sense of Heine-Borel *but not* in the sense of Bolzano-Weierstrass, and conversely the real line endowed with the topology in which the nonempty open subsets are those with countable complement is compact in the sense of Bolzano-Weierstrass but not in the sense of Heine-Borel. However, for metric (or metrizable) spaces these notions – and several others – of compactness coincide. For completeness we give a treatment of this, in which the only possible novelty is a relatively streamlined approach.

Recall that a topological space X is compact iff it satisfies the **finite intersection property**: if $\{F_i\}_{i \in I}$ is a family of closed subsets of X such that for all finite subsets $J \subset I$, $\bigcap_{i \in J} F_i \neq \varnothing$, then $\bigcap_{i \in I} F_i = \varnothing$.

A topological space X is **sequentially compact** if every sequence admits a convergent subsequence.

A topological space X is **countably compact** if every countable open cover

$\{U_n\}_{n=1}^{\infty}$ admits a finite subcover. This is equivalent to the finite intersection property for countable families $\{F_n\}_{n=1}^{\infty}$ of closed subsets. By passing from F_n to $\mathcal{F}_n = \bigcap_{i=1}^{n} F_i$, we see that a space is countably compact iff every nested sequence of nonempty closed subsets has nonempty intersection.

A topological space X is **limit point compact** if every infinite subset $Y \subset X$ has an **accumulation point** in X, i.e., there exists $x \in X$ such that for every open neighborhood U of x, $U \setminus \{x\} \cap Y \neq \varnothing$.

Thus Bolzano-Weierstrass asserts that $[a, b]$ is limit point compact, whereas Theorem 25 asserts, in particular, that $[a, b]$ is sequentially compact.

Proposition 28. *Let X be a topological space.*
a) If X is countably compact, it is limit point compact.
b) In particular a compact space is limit point compact.
c) If X is sequentially compact, it is countably compact.
d) In particular a sequentially compact space is limit point compact.

Proof. a) We establish the contrapositive: suppose there exists an infinite subset of X with no accumulation point; then there exists a countably infinite subset $A \subset X$ with no accumulation point. Such a subset A must be closed, since any element of $\overline{A} \setminus A$ is an accumulation point of A. Moreover A must be discrete: for each $a \in A$, since a is not an accumulation point of A, there exists an open subset U such that $A \cap U = \{a\}$. Now write $A = \{a_n\}_{n=1}^{\infty}$, and define, for each $N \in \mathbb{Z}^+$, $F_N = \{a_n\}_{n=N}^{\infty}$. Then each F_N is closed, any finite intersection of F_N's is nonempty, but $\bigcap_{N=1}^{\infty} F_N = \varnothing$, so X is not countably compact.
b) Clearly a compact space is countably compact; now apply part a).
c) Let $\{F_n\}_{n=1}^{\infty}$ be a nested sequence of closed subsets of X, and choose for all $n \in \mathbb{Z}^+$ a point $x_n \in F_n$. By sequential compactness, after passing to a subsequence – let us suppose we have already done so and retain the current indexing – we get $x \in X$ such that $x_n \to x$. We claim $x \in \bigcap_{n=1}^{\infty} F_n$. Suppose not: then there is $N \in \mathbb{Z}^+$ such that $x \notin F_N$. But then $U = X \setminus F_N$ is an open neighborhood of x, so for all sufficiently large n, $x_n \in U$ and thus $x_n \notin F_N$. But as soon as $n \geq N$ we have $F_n \subset F_N$ and thus $x_n \notin F_n$, contradiction.
d) Apply part c) and then part a). □

Theorem 29.
a) A first countable, separated,[6] limit point compact space is sequentially compact.
b) In a metrizable space X, sequential compactness implies quasi-compactness, and hence quasi-compactness, sequential compactness, limit point compactness, and countable compactness are all equivalent properties.

Proof. a) Let a_n be a sequence in X. If the image of the sequence is finite, we may extract a constant, hence convergent, subsequence. Otherwise the image $A = \{a_n\}_{n=1}^{\infty}$ has a limit point a, and since X is separated every limit point is an ω-accumulation point: every neighborhood U of a contains infinitely many points of A. Let $\{N_n\}_{n=1}^{\infty}$ be a nested countable neighborhood base at x. Choose n_1 such that $x_{n_1} \in N_1$. For all $k > 1$, choose $n_k > n_{k-1}$ with $x_{n_k} \in N_k$. Then $x_{n_k} \to x$.
b) Since X is first countable and separated, sequential compactness, countable

[6] A separated space is one in which all singleton sets are closed.

compactness and limit point compactness of X are all equivalent. Finally, we will suppose that X is sequentially compact (and thus limit point compact) and show that it is compact. For each $n \in \mathbb{Z}^+$, choose by Zorn's Lemma a subset T_n which is maximal with respect to the property that the distance between any two elements is at least $\frac{1}{n}$.[7] Each set T_n admits no limit point, so must be finite. Since every point of X lies at a distance at most $\frac{1}{n}$ from some element of T_n, the set $T = \bigcup_{n=1}^{\infty} T_n$ is a countable dense subset. Since X is metrizable, from T we can build a countable base $\mathcal{B} = \{B_n\}_{n=1}^{\infty} = \{B(t, \frac{1}{n})\}_{t \in T, n \in \mathbb{Z}^+}$, and this in turn implies that every open cover $\mathcal{U} = \{U_i\}_{i \in I}$ has a countable subcover: for each $n \in \mathbb{Z}^+$, let i_n be an element of I such that $B_n \subset I_{i_n}$ if such an index exists; otherwise let i_n be any element of I. Then $\{I_{i_n}\}_{n=1}^{\infty}$ is a countable subcover: for $x \in X$, there exists $i \in I$ such that $x \in U_i$, hence $n \in \mathbb{Z}^+$ such that $x \in B_n \subset U_i$ and then $x \in B_n \subset I_{i_n}$. But since X is sequentially compact, it is countably compact, and thus the countable cover $\{I_{i_n}\}_{n=1}^{\infty}$ itself has a finite subcover, so X is compact. □

7. With the help of completeness

A metric space X is **complete** if all Cauchy sequences in X converge.

Proposition 30. *A sequentially compact metric space is complete.*

Proof. The argument given in the proof of Theorem 25b) applies verbatim here. □

Theorem 31. *For an Archimedean ordered field F, TFAE:*
(i) F is Dedekind complete.
(ii) Every Cauchy sequence in F converges.

Proof. (i) $\implies$ (ii): We have already established this in Theorem 25 above.
(ii) $\implies$ (i): Let $S \subset F$ be nonempty and bounded above, and write $\mathcal{U}(S)$ for the set of least upper bounds of S. Our strategy will be to construct a decreasing Cauchy sequence in $\mathcal{U}(S)$ and show that its limit is $\sup S$.

Let $a \in S$ and $b \in \mathcal{U}(S)$. Using the Archimedean property, we choose a negative integer $m < a$ and a positive integer $M > b$, so

$$m < a \leq b \leq M.$$

For each $n \in \mathbb{Z}^+$, we define

$$S_n = \{k \in \mathbb{Z} \mid \frac{k}{2^n} \in \mathcal{U}(A) \text{ and } k \leq 2^n M\}.$$

Every element of S_n lies in the interval $[2^n m, 2^n M]$ and $2^n M \in S_n$, so each S_n is finite and nonempty. Put $k_n = \min S_n$ and $a_n = \frac{k_n}{2^n}$, so $\frac{2k_n}{2^{n+1}} = \frac{k_n}{2^n} \in \mathcal{U}(S)$ wihle $\frac{2k_n - 2}{2^{n+1}} = \frac{k_n - 1}{2^n} \notin \mathcal{U}(S)$. It follows that we have either $k_{n+1} = 2k_n$ or $k_{n+1} = 2k_n - 1$ and thus either $a_{n+1} = a_n$ or $a_{n+1} = a_n - \frac{1}{2^{n+1}}$. In particular $\{a_n\}$ is decreasing. For all $1 \leq m < n$ we have

$$0 \leq a_m - a_n = (a_m - a_{m+1}) + (a_{m+1} - a_{m+2}) + \ldots + (a_{n-1} - a_n)$$
$$\leq 2^{-(m+1)} + \ldots + 2^{-n} = 2^{-m}.$$

[7] The appeal to Zorn's Lemma may not be to your taste. I like it because we have not yet needed to use the assumption that a subset without any limit points must be finite. But conversely, by making use of that assumption here, Zorn's Lemma can be avoided.

Thus $\{a_n\}$ is a Cauchy sequence, hence by our assumption on F $a_n \to L \in F$.

We CLAIM $L = \sup(S)$. Seeking a contradiction we suppose that $L \notin \mathcal{U}(S)$. Then there exists $x \in S$ such that $L < x$, and thus there exists $n \in \mathbb{Z}^+$ such that

$$a_n - L = |a_n - L| < x - L.$$

It follows that $a_n < x$, contradicting $a_n \in \mathcal{U}(S)$. So $L \in \mathcal{U}(S)$. Finally, if there exists $L' \in \mathcal{U}(S)$ with $L' < L$, then (using the Archimedean property) choose $n \in \mathbb{Z}^+$ with $\frac{1}{2^n} < L - L'$, and then

$$a_n - \frac{1}{2^n} \geq L - \frac{1}{2^n} > L',$$

so $a_n - \frac{1}{2^n} = \frac{k_n - 1}{2^n} \in \mathcal{U}(S)$, contradicting the minimality of k_n. □

Thus among Archimidean fields, completeness in the metric sense has the same force as Dedekind completeness.[8] Of course our derivation of the completeness of $\mathbb{R}$ *used* Dedekind completeness, Real Induction, and the Bolzano-Weierstrass Theorem. However, a tenable alternative is to take the metric completeness of $\mathbb{R}$ as its characteristic property (among Archimedean ordered fields): this is exactly the approach taken in many "sequences first" analysis courses. In this approach the key fact to establish is the sequential compactness of $[a, b]$. In this section and the next we present two further routes to these results.

The first begins by contemplating the following basic question about metric spaces: sequential compactness implies completeness but the converse does not hold: e.g. $\mathbb{R}$ is complete but not sequentially compact. So what is the additional ingredient necessary to get from completeness to sequential compactness?

Lemma 32. *For a metric space X, the following are equivalent:*
(i) For all $\epsilon > 0$, there exists a finite family $S_1, \ldots, S_N$ of subsets of X such that $\operatorname{diam} S_i < \epsilon$ for all i and $X = \bigcup_{i=1}^N S_i$.
(ii) For all $\epsilon > 0$, there are $x_1, \ldots, x_N \in X$ such that $\bigcup_{i=1}^N B(x_i, \epsilon) = X$.
*A metric space satisfying these equivalent conditions is **totally bounded**.*

Proof. (i) $\implies$ (ii): We may assume each S_i is nonempty, and choose $x_i \in S_i$. Since $\operatorname{diam} S_i < \epsilon$, $S_i \subset B(x_i, \epsilon)$ and thus $X = \bigcup_{i=1}^N B(x_i, \epsilon)$.
(ii) $\implies$ (i): For every $\epsilon > 0$, choose $x_1, \ldots, x_N$ such that $\bigcup_{i=1}^N B(x_i, \frac{\epsilon}{2}) = X$. We have covered X by finitely many sets each of diameter at most ϵ. □

Lemma 33. *Every subset of a totally bounded metric space is totally bounded.*

Proof. Suppose that X is totally bounded, and let $Y \subset X$. Since X is totally bounded, for each $\epsilon > 0$ there exist $S_1, \ldots, S_N \subset X$ such that $\operatorname{diam} S_i < \epsilon$ for all i and $X = \bigcup_{i=1}^N S_i$. Then $\operatorname{diam}(S_i \cap Y) < \epsilon$ for all i and $Y = \bigcup_{i=1}^N (S_i \cap Y)$. □

In any metric space a totally bounded set is bounded. The converse does not generally hold: for instance, an infinite set with the discrete metric is bounded but not totally bounded. Nevertheless every bounded subset of $\mathbb{R}$ is totally bounded.

Theorem 34. *a) A metric space is totally bounded iff every sequence admits a Cauchy subsequence.*
b) A metric space is sequentially compact iff it is complete and totally bounded.

[8] However, there are non-Archmiedean ordered fields in which every Cauchy sequence converges...and even ones in which every Cauchy sequence is ultimately constant!

Proof. a) Supppose X is totally bounded. Cover X by finitely many subsets $X_{1,1}, \ldots, X_{1,N(1)}$ of diameter at most 1. By the Pigeonhole Principle we may extract a subsequence x_{n_k} all of whose terms lie in $X_{1,i}$ for some fixed i. By Lemma 33 we may cover $X_{1,i}$ by finitely many subsets $X_{2,1}, \ldots, X_{2,N(2)}$ of diameter at most $\frac{1}{2}$ and again extract a sub(sub)sequence each of whose terms lies in $X_{2,i}$ for some fixed i. Continuing in this way, we may cover X_{k,i_k} by finitely many subsets each of diameter at most $\frac{1}{2^{k+1}}$ and extract a subsequence all of whse terms lie in one of these subsets. Finally, choose a *diagonal subsequence*, a sequence whose kth term lies in the kth extracted subsequence. This diagonal subsequence is Cauchy in X.

Now suppose that X is *not* totally bounded, and let $\epsilon > 0$ be such that X *does not* admit a finite covering by ϵ-balls. Choose $x_1 \in X$. Since $B(x_1, \epsilon) \neq X$, we may choose $x_2 \in X \setminus B(x_1, \epsilon)$. Since $X \neq B(x_1, \epsilon) \cup B(x_2, \epsilon)$, we may choose $x_3 \in X \setminus (B(x_1, \epsilon) \cup B(x_2, \epsilon)$. And so forth: having chosen $x_1, \ldots, x_n$, we may choose $x_{n+1} \in X \setminus \bigcup_{i=1}^n B(x_i, \epsilon)$. For the resulting sequence $\{x_n\}$ we have $d(x_n, x_{n+1}) \geq \epsilon$ for all $n \in \mathbb{Z}^+$, so there is no Cauchy subsequence.
b) Suppose X is sequentially compact. Since any Cauchy sequence admitting a convergent subsequence is itself convergent, X is complete. Moreover, since every convergent sequence is Cauchy, every sequence in X admits a Cauchy subsequence, so by part a) X is totally bounded. Conversely, suppose X is complete and totally bounded. Then by part a) any sequence in X admits a Cauchy subsequence, hence by completeness a convergent subsequence. □

8. With the help of monotonicity

Lemma 35. *(Monotone Sequence Lemma)*
a) In any ordered field F, an increasing (resp. decreasing) sequence $\{x_n\}$ is convergent iff its term set $\{x_n : n \in \mathbb{Z}^+\}$ admits a supremum (resp. infimum) in F.
b) Thus in a Dedekind complete ordered field, an increasing (resp. decreasing) sequence is convergent iff it is bounded above (resp. below).

Proof. a) It suffices to consider increasing sequences.

Suppose $x_n \to L$. If L were not an upper bound for $\{x_n\}_{n=1}^\infty$, there would be $N \in \mathbb{Z}^+$ such that $L < x_N$. Since the sequence is increasing, $L < x_N \leq x_n$ for all $n \geq N$. Taking $\epsilon = x_N - L$, for no $n \geq N$ do we have $|x_n - L| < \epsilon$, contradicting our assumption that $x_n \to L$. Similarly, suppose L is not the least upper bound: there is L' such that for all $n \in \mathbb{Z}^+$, $x_n \leq L' < L$. Let $\epsilon = L - L'$. Then for no n do we have $|x_n - L| < \epsilon$, contradicting our asumption that x_n converges to L.

Suppose $L = \sup\{x_n\}_{n=1}^\infty$. Then $L \geq x_n$ for all $n \in \mathbb{Z}^+$. Now fix $\epsilon > 0$ and suppose there were infinitely many n with $L \geq x_n + \epsilon$. Since the sequence is increasing, this inequality would then hold for all n: $L - \epsilon \geq x_n$ for all n, so that $L - \epsilon$ is a smaller upper bound for X than L, contradiction.
b) This follows immediately from part a). □

Lemma 36. *(Rising Sun Lemma) Every infinite sequence in a linearly ordered set has a monotone subsequence.*

Proof. Let us say that $n \in \mathbb{Z}^+$ is a **peak** of the sequence $\{a_n\}$ if for all $m < n$, $a_m < a_n$. Suppose first that there are infinitely many peaks. Then any sequence of peaks forms a strictly decreasing subsequence, hence we have found a monotone subsequence. So suppose on the contrary that there are only finitely many peaks, and let $N \in \mathbb{N}$ be such that there are no peaks $n \geq N$. Since $n_1 = N$ is not a peak, there exists $n_2 > n_1$ with $a_{n_2} \geq a_{n_1}$. Similarly, since n_2 is not a peak, there exists $n_3 > n_2$ with $a_{n_3} \geq a_{n_2}$. Continuing in this way we construct an infinite (not necessarily strictly) increasing subsequence $a_{n_1}, a_{n_2}, \ldots, a_{n_k}, \ldots$. Done! □

Theorem 37. *The Rising Sun Lemma implies the sequential compactness of* $[a, b]$.

Proof. Let $\{x_n\}$ be a sequence in $[a, b]$. By the Rising Sun Lemma, $\{x_n\}$ admits a monotone subsequence, which is then convergent by Lemma 35. □

9. Still More Real Induction

9.1. The Mean Value Inequality.

For an interval $I \subset \mathbb{R}$, let us denote by I° the interior of I.

Theorem 38. *(Mean Value Theorem) Let* $f : I \to \mathbb{R}$ *be continuous and differentiable on* I°. *For any* $a < b \in I$, *there exists* $c \in I^\circ$ *such that* $\frac{f(b)-f(a)}{b-a} = f'(c)$.

Let us briefly recall the standard proof: it is no loss of generality to assume that $I = [a, b]$. Moreover, by subtracting an appropriate linear function, we reduce to the case $f(a) = f(b) = 0$. Then the desired conclusion certainly holds if f is identically zero; otherwise, applying the Extreme Value Theorem, f assumes either its minimum or maximum value at an interior point c, at which $f'(c) = 0$.

Some mathematicians who teach calculus are not fans of Theorem 38. Many articles have been written on the prospect of replacing the Mean Value Theorem with some alternate (usually weaker) version that is (somehow) more appealing / intuitive to calculus students: e.g. [Be67], [Co67], [Tu97], [Ko09]. Most of these articles advocate replacing Theorem 38 with one or more of the following results.

Corollary 39. *(Mean Value Inequality) Let* $f : I \to \mathbb{R}$ *be differentiable. Suppose that there exists* $M \in \mathbb{R}$ *such that for all* $x \in I$, $f'(x) \geq M$. *Then for all* $x < y \in \mathbb{R}$, $f(y) - f(x) \geq M(y - x)$.

Corollary 40. *(Weakly Increasing Function Theorem) Let* $f : I \to \mathbb{R}$ *be differentiable, and suppose that for all* $x \in I$, $f'(x) \geq 0$. *Then* f *is* weakly increasing *on* I: *for all* $x, y \in I$, $x \leq y \implies f(x) \leq f(y)$.

Corollary 41. *(Increasing Function Theorem) Let* $f : I \to \mathbb{R}$ *be differentiable, and suppose that for all* $x \in I$, $f'(x) > 0$. *Then* f *is* increasing *on* I: *for all* $x, y \in I$, $x < y \implies f(x) < f(y)$.

Corollary 42. *(Racetrack Principle) Let* $f, g : [a, b] \to \mathbb{R}$ *be differentiable. If* $g'(x) \leq f'(x)$ *for all* $x \in [a, b]$, *then* $g(x) - g(a) \leq f(x) - f(a)$.

Corollary 43. *(Zero Velocity Theorem) Let* $f : I \to \mathbb{R}$ *be differentiable, and suppose that for all* $x \in I^\circ$, $f'(x) = 0$. *Then* f *is constant.*

Corollaries 39 through 43 are indeed all immediate consequences of the Mean Value Theorem, and Corollaries 39, 40 and 42 each immediately imply Corollary 43. In fact we claim that Corollaries 39, 40, 41 and 42 are all equivalent. It is immediate that Corollary 39 $\implies$ Corollary 40 $\implies$ 42, so it suffices to show Corollary 41 $\implies$ Corollary 40 and that Corollary 42 $\implies$ Corollary 39.

Assume that Corollary 41 holds, and seeking a contradiction suppose that there exist $a < b \in I$ with $f(a) > f(b)$ and $f'(x) \geq 0$ on $[a, b]$. For any $C > 0$, consider $f_C(x) = f(x) + Cx$. Then $f_C'(x) > 0$ on $[a,b]$ so by Corollary 41 f_C is strictly increasing: in particular $f_C(b) = f(b) + Cb > f(a) = f(a) + Ca$, or $f(a) - f(b) < C(b-a)$. Taking C to be sufficiently small gives a contradiction. Assume that Corollary 42 holds, suppose $f'(x) \geq M$ on $[x,y]$, and let $g(x) = Mx$. Then by Corollary 42 we have $f(y) - f(x) \geq g(y) - g(x) = M(y-x)$.

Remark: It is part of the folklore that the Mean Value Theorem is in some sense[9] strictly stronger than the Mean Value Inequality and its equivalents and that these four results are in turn strictly stronger than the Zero Velocity Theorem. The articles cited above contain a lot of information about which of these corollaries are needed for which calculus results. In brief, the Zero Velocity Theorem is all that is needed for integration theory (specifically, uniqueness of anti-derivatives), whereas for most of differential calculus the Mean Value Inequality is necessary and sufficient. A notable exception is L'Hôpital's Rule, which in its full strength seems to require Cauchy's form of the Mean Value Theorem (which is deduced from Rolle's Theorem and thus equivalent to the Mean Value Theorem).

I have no pedagogical misgivings about the Mean Value Theorem. On the contrary, I have presented the statement and even the proof in every differential calculus course I have ever taught. It seems to me that this result, which can be viewed as relating average velocity to instantaneous velocity and vividly applied in terms of issuing speeding tickets using cameras posted at highway checkpoints (I leave the details to you!), is both important and within the reach of freshman calculus students...certainly as much so as any of the corollaries given above.

But it seems there is a mathematical distinction to be made between Theorem 38 and Corollaries 39, 40, 41, 42: the former cannot – so far as I know! – be proved directly by Real Induction, while the latter four can. These proofs have a similar flavor, so we will content ourselves with a Real Induction proof of Corollary 39.

Proof. Let $a < b \in I$. Fix $\epsilon > 0$, and define

$$S_\epsilon = \{x \in [a,b] \mid f(x) - f(a) \geq (M-\epsilon)(x-a)\}.$$

We will prove by Real Induction that $S_\epsilon = [a,b]$. Then in particular $b \in S_\epsilon$, so $f(b) - f(a) \geq (M-\epsilon)(b-a)$. Since this holds for all $\epsilon > 0$, we deduce $f(b) - f(a) \geq M(b-a)$. (RI1) It is immediate that $a \in S_\epsilon$. (RI2) Since $f'(x) \geq M$, there is $\delta > 0$ such that $y \in [x, x+\delta] \implies \frac{f(y)-f(x)}{y-x} \geq M - \epsilon$. Thus $f(y) - f(a) = (f(y) - f(x)) + (f(x) - f(a)) \geq (M-\epsilon)(y-x) + (M-\epsilon)(x-a) = (y-a)(M-\epsilon)$, so $[x, x+\delta] \subset S_\epsilon$. (RI3) Suppose that for $x \in (a,b]$ we have $[a,x) \subset S_\epsilon$. As above, since $f'(x) \geq M$, there exists $\delta > 0$ such that $y \in [x-\delta, x] \implies \frac{f(x)-f(y)}{x-y} \geq M - \epsilon$. Thus $f(x) - f(a) = (f(x) - f(y)) + (f(y) - f(a)) \geq (M-\epsilon)(x-y) + (M-\epsilon)(y-a) = (x-a)(M-\epsilon)$, so $x \in S_\epsilon$. □

Problem 3. *Give either a direct proof of the Mean Value Theorem by Real Induction or a convincing explanation for why such a proof cannot exist.*

9.2. **Some Real Induction Proofs For the Reader.**

We have presented a number of Real Induction proofs. Our list of results of basic

[9]Alas, we will not attempt a precise formalization of this statement here.

real analysis that can be proved by Real Induction is not yet complete, but perhaps it is time to let the interested reader try to construct such proofs for herself. Here are four more results amenable to Real Induction...with the proofs omitted.

Theorem 44. *(Cantor Intersection Theorem) Let* $\{F_n\}_{n=1}^{\infty}$ *be a decreasing sequence of closed subsets of* $[a,b]$. *Put* $F = \bigcap_n F_n$. *Then* either $F \neq \varnothing$ *or there exists* $n \in \mathbb{Z}^+$ *such that* $F_n = \varnothing$.

Theorem 45. *Any open covering of* $[a,b]$ *admits a Lebesgue number.*

Theorem 46. *(Dini's Lemma) Let* $\{f_n\}_{n=1}^{\infty}$ *be a sequence of continuous real-valued functions on the interval* $[a,b]$ *which is* pointwise decreasing*: for all* $x \in [a,b]$ *and all* $n \in \mathbb{Z}^+$, $f_{n+1}(x) \leq f_n(x)$. *If* $f : [a,b] \to \mathbb{R}$ *is continuous and* $f_n \to f$ *pointwise, then* $f_n \to f$ *uniformly.*

Theorem 47. *(Arzelà-Ascoli)*
Let $\{f_n\}_{n=1}^{\infty}$ *be a sequence of continuous functions on* $[a,b]$ *such that:*
(a) There is $M \in \mathbb{R}$ *such that for all* $n \in \mathbb{Z}^+$ *and all* $x \in [a,b]$, $|f_n(x)| \leq M$ *and*
(b) For all $x \in [a,b]$ *and all* $\epsilon > 0$, *there exists* $\delta > 0$ *such that if* $|x - y| < \delta$, *then for all* $n \in \mathbb{Z}^+$, $|f_n(x) - f_n(y)| < \epsilon$.
Then there exists a subsequence $\{f_{n_k}\}$ *which is uniformly convergent on* $[a,b]$.

By now we have come close to exhausting the theorems the author (knows of and) considers "basic results in elementary real analysis", but surely it does not come close to exhausting the applicability of the method: as we saw, its scope is in principle as wide as that of the least upper bound axiom. Now that you are aware of the method, I invite you to keep an eye out to its use in the future.

Problem 4. *Find other theorems which can be proved via Real Induction.*

10. A Moral?

In fact, I invite you to keep an eye out for different approaches and proofs for basic results throughout undergraduate mathematics. The "American freshman calculus text" is a remarkable phenomenon of intellectual stultification: hundreds of specimens exist, almost all of which are small (infinitesimal?) deformations of one another. This contributes to a class environment where the instructor feels implicitly compelled to present the ideas of calculus as a liturgical recitation: the way it was presented to her is almost exactly the same way it is presented in the text, (almost) no matter which text is chosen! But in truth there are so many different approaches to this or any undergraduate course that a capable instructor could choose to present (some of) the material anew for each iteration of the course. Among these multiple routes, are perhaps some especially well adapted to the course goals, prerequisites and schedule as well as the background, abilities, tastes, needs, and interests of the current student cohort? It is worth thinking about.

References

[Be67] L. Bers, *Classroom Notes: On Avoiding the Mean Value Theorem.* Amer. Math. Monthly 74 (1967), 583.

[Be82] H. Bereková, *The principle of induction in continuum and related methods.* Acta Math. Univ. Comenian. 40(41) (1982), 97–100.

[Bo58] J. Bosák, *Generalization of the method of complete induction.* Acta Fac. Nat. Univ. Comenian. Math. 2 (1958), 255-256.

[Ch19] Y. R. Chao, *A note on "Continuous mathematical induction."* Bull. Amer. Math. Soc. Volume 26, (1919), 17–18.
[Cl11] P.L. Clark, *Induction and completeness in ordered sets*, submitted for publication.
[Co67] L.W. Cohen, *Classroom Notes: On Being Mean to the Mean Value Theorem.* Amer. Math. Monthly 74 (1967), 581-582.
[Do03] G. Dowek, *Preliminary investigations on induction over real numbers*, preprint, 2003.
[Du57] W.L. Duren, Jr., *Mathematical induction in sets.* Amer. Math. Monthly 64 1957 no. 8, part II, 19–22.
[Fo57] L.R. Ford, *Interval-additive propositions.* Amer. Math. Monthly 64 (1957), 106–108.
[Ha11] D. Hathaway, *Using Continuity Induction.* College Math. Journal 42 (2011), 229–231.
[Ka07] I. Kalantari, *Induction over the continuum.* (English summary) Induction, algorithmic learning theory, and philosophy, 145–154, Log. Epistemol. Unity Sci., 9, Springer, Dordrecht, 2007.
[Kh23] A. Khinchin, Fund. Math. 4 (1923), 164–166.
[Kh49] A. Khinchin, *The simplest linear continum.* Uspehi Matem. Nauk (N.S.) 4, (1949). no. 2 (30), 180–197.
[Ko09] J.J. Koliha, *Mean, meaner, and the meanest mean value theorem.* Amer. Math. Monthly 116 (2009), 356-361.
[Le82] H. Leinfelder, *A unifying principle in real analysis.* Real Anal. Exchange 8 (1982/83), no. 2, 511–518.
[MR68] R.M.F. Moss and G.T. Roberts, *A creeping lemma.* Amer. Math. Monthly 75 (1968), 649–652.
[Ne86] A. Neubrunnová, *On a unifying principle in real analysis.* Acta Math. Univ. Comenian. 48/49 (1986), 123–126 (1987).
[No52] M.J. Norris, *Integrability of continuous functions.* Amer. Math. Monthly 59 (1952), 244-245.
[Pe26] O. Perron, Jber. Deutsch. Math. Verein. 35 (1926), 194–203.
[Sa84] T. Šalát, *Remarks on unifying principles in real analysis.* Real Anal. Exchange 10 (1984/85), no. 2, 343–348.
[Sh72] P. Shanahan, *A unified proof of several basic theorems of real analysis.* Amer. Math. Monthly 79 (1972), 895–898.
[Sh74] P. Shanahan, *Addendum to: "A unified proof of several basic theorems of real analysis" (Amer. Math. Monthly 79 (1972), 895–898).* Amer. Math. Monthly 81 (1974), 890–891.
[S] M. Spivak, *Calculus.* Fourth edition.
[Tu97] T.W. Tucker, *Rethinking rigor in calculus: the role of the mean value theorem.* Amer. Math. Monthly 104 (1997), 231-240.

Department of Mathematics, Boyd Graduate Studies Research Center, University of Georgia, Athens, GA 30602-7403, USA
E-mail address: pete@math.uga.edu