MATH 2400 LECTURE NOTES: COMPLETENESS

PETE L. CLARK

Contents

1. Dedekind Completeness

1.1. Introducing (LUB) and (GLB).

Gather round, my friends: the time has come to tell what makes calculus work.

Recall that we began the course by considering the real numbers as a set endowed with two binary operations $+$ and $\cdot$ together with a relation $<$, and satisfying a longish list of familiar axioms (P0) through (P12), the **ordered field** axioms. We then showed that using these axioms we could deduce many other familiar properties of numbers and prove many other identities and inequalities.

However we did not claim that (P0) through (P12) was a *complete* list of axioms for $\mathbb{R}$. On the contrary, we saw that this could not be the case: for instance the rational numbers $\mathbb{Q}$ also satisfy the ordered field axioms but – as we have taken great pains to point out – most of the "big theorems" of calculus are meaningful but false when regarded as results applied to the system of rational numbers. So

Date: November 21, 2011.

there must be some further axiom, or property, of $\mathbb{R}$ which is needed to prove the three Interval Theorems, among others.

Here it is. Among structures F satisfying the ordered field axioms, consider the following further property:

(P14): **Least Upper Bound Axiom** (LUB): Let S be a nonempty subset of F which is bounded above. Then S admits a **least upper bound**.

This means exactly what it sounds like, but it is so important that we had better make sure. Recall a subset S of F is **bounded above** if there exists $M \in \mathbb{R}$ such that for all $x \in S$, $x \leq M$. (For future reference, a subset S of $\mathbb{R}$ is **bounded below** if there exists $m \in F$ such that for all $x \in S$, $m \leq x$.) By a **least upper bound** for a subset S of F, we mean an upper bound M which is less than any other upper bound: thus, M is a least upper bound for S if M is an upper bound for S and for any upper bound M' for S, $M \leq M'$.

There is a widely used synonym for "the least upper bound of S", namely the **supremum** of S. We also introduce the notation $\operatorname{lub} S = \sup S$ for the supremum of a subset S of an ordered field (when it exists).

The following is a useful alternate characterization of $\sup S$: the supremum of S is an upper bound M for S with the property that for any $M' < M$, M' is *not* an upper bound for S: explicitly, for all $M' < M$, there exists $x \in S$ with $M' < x$.

The definition of the least upper bound of a subset S makes sense for any set X endowed with an order relation $<$. Notice that the *uniqueness* of the supremum $\sup S$ is clear: we cannot have two different least upper bounds for a subset, because one of them will be larger than the other! Rather what is in question is the *existence* of least upper bounds, and (LUB) is an assertion about this.

Taking the risk of introducing even more terminology, we say that an ordered field $(F, +, \cdot, <)$ is **Dedekind complete**[1] if it satisfies the least upper bound axiom. Now here is the key fact lying at the foundations of calculus and real analysis.

Theorem 1. *a) The ordered field $\mathbb{R}$ is Dedekind complete.*
b) Conversely, any Dedekind complete ordered field is isomorphic to $\mathbb{R}$.

Part b) of Theorem 1 really means the following: if F is any Dedekind complete ordered field then there is a bijection $f : F \to \mathbb{R}$ which preserves the addition, multiplication and order structures in the following sense: for all $x, y \in F$,

- $f(x + y) = f(x) + f(y)$,
- $f(xy) = f(x)f(y)$, and
- If $x < y$, then $f(x) < f(y)$.

This concept of "isomorphism of structures" comes from a more advanced course –

[1] It is perhaps more common to say "complete" instead of "Dedekind complete". I have my reasons for preferring the lengthier terminology, but I won't trouble you with them.

abstract algebra – so it is probably best to let it go for now. One may take part b) to mean that there is *essentially* only one Dedekind complete ordered field: $\mathbb{R}$.

The proof of Theorem 1 involves *constructing* the real numbers in a mathematically rigorous way. This is something of a production, and although in some sense every serious student of mathematics should see a construction of $\mathbb{R}$ at some point of her career, this sense is similar to the one in which every serious student of computer science should build at least one working computer from scratch: in practice, one can probably get away with relying on the fact that many other people have performed this task in the past. Spivak does give a construction of $\mathbb{R}$ and a proof of Theorem 1 in the "Epilogue" of his text. And indeed, if we treat this material at all it will be at the very end of the course.

After discussing least upper bounds, it is only natural to consider the "dual" concept of greatest lower bounds. Again, this means exactly what it sounds like but it is so important that we spell it out explicitly: if S is a subset of an ordered field F, then a **greatest lower bound** for S, or an **infimum** of S, is an element $m \in F$ which is a lower bound for S – i.e., $m \leq x$ for all $x \in S$ – and is such that if m' is any lower bound for S then $m' \leq m$. Equivalently, $m = \inf S$ iff m is a lower bound for S and for any $m' > m$ there exists $x \in S$ with $x < m'$. Now consider:

(P14$'$): **Greatest Lower Bound Axiom** (GLB): Let S be a nonempty subset of F which is bounded below. Then S admits a greatest lower bound, or infimum.

Example 1.1: In any ordered field F, we may consider the subset

$$S_F = \{x \in F \mid x^2 < 2\}.$$

Then S_F is nonempty and bounded: indeed $0 \in S_F$ and if $x \in S_F$, then $|x| \leq 2$. Of course in the previous inequality we could do better: for instance, if $|x| > \frac{3}{2}$, then $x^2 > \frac{9}{4} > 2$, so also $\frac{-3}{2}$ is a lower bound for S_F and $\frac{3}{2}$ is an upper bound for S_F. Of course we could do better still...

Indeed the bounded set S_F will have an infimum and a supremum if and only if there are *best possible* inequalities $x \in S \implies m \leq x \leq M$, i.e., for which no improvement on either m or M is possible. Whether such best possible inequalities exist depends on the ordered field F. Indeed, it is clear that *if* $M = \sup S_F$ exists, then it must be a positive element of F with $M^2 = 2$: or in other words, what in precalculus mathematics one cavalierly writes as $M = \sqrt{2}$. similarly, *if* $m = \inf S_F$ exists, then it must be a negative element of F with $m^2 = 2$, or what we usually write as $-\sqrt{2}$. But here's the point: how do we know that our ordered field F contains such an element $\sqrt{2}$?

The answer of course is that depending on F such an element may or may not exist. As we saw at the beginning of the course, there is no *rational number* x with $x^2 = 2$, so if $F = \mathbb{Q}$ then our set $S_{\mathbb{Q}}$ has neither an infimum nor a supremum. Thus $\mathbb{Q}$ does not satisfy (LUB) or (GLB). On the other hand, we certainly believe that there is a real number whose square is 2. But...why do we believe this? As we have seen, the existence of a real square root of every non-negative real number is a consequence of the Intermediate Value Theorem...which is of course a theorem that we have exalted but not yet proved. A more fundamental answer is that we believe that $\sqrt{2}$ exists in $\mathbb{R}$ *because* of the Dedekind completeness of $\mathbb{R}$, i.e., according to Theorem 1 *every* nonempty bounded above subset of $\mathbb{R}$ has a supremum, so in particular $S_{\mathbb{R}}$ has a supremum, which must be $\sqrt{2}$.

An interesting feature of this example is that we can see that $\inf S_{\mathbb{R}}$ exists, even though we have not as yet addressed the issue of whether $\mathbb{R}$ satisfies (GLB). In general, $\inf S_F$ exists iff there is an element $y < 0$ in F with $y^2 = 2$. But okay: if in F we have a positive element x with $x^2 = 2$, we necessarily must also have a negative element y with $y^2 = 2$: namely, $y = -x$.

This turns out to be a very general phenomenon.

Theorem 2. *Let F be an ordered field.*
a) Then F satisfies (LUB) iff it satisfies (GLB).
b) In particular $\mathbb{R}$ satisfies both (LUB) and (GLB) and is (up to isomorphism) the only ordered field with this property.

Proof. a) I know two ways of showing that (LUB) $\iff$ (GLB). Both of these arguments is very nice in its own way, and I don't want to have to choose between them. So I will show you both, in the following way: I will use the first argument to show that (LUB) $\implies$ (GLB) and the second argument to show that (GLB) $\implies$ (LUB). (In Exercise 1.2 below, you are asked to do things the other way around.)
(LUB) $\implies$ (GLB): Let $S \subset F$ be nonempty and bounded below by m. Consider

$$-S = \{-x \mid x \in S\}.$$

Then $-S$ is nonempty and bounded above by $-m$. By (LUB), it has a least upper bound $\sup(-S)$. We claim that in fact $-\sup(-S)$ is a greatest lower bound for S, or more symbolically:

$$\inf S = -\sup -S.$$

You are asked to check this in Exercise 1.2 below.
(GLB) $\implies$ (LUB): Let S be nonempty and bounded above by M. Consider

$$\mathcal{U}(S) = \{x \in F \mid x \text{ is an upper bound for } S.\}.$$

Then $\mathcal{U}(S)$ is nonempty: indeed $M \in \mathcal{U}(S)$. Also $\mathcal{U}(S)$ is bounded below: indeed any $s \in S$ (there is at least one such s, since $S \neq \varnothing$!) is a lower bound for $\mathcal{U}(S)$. By (GLB) $\mathcal{U}(S)$ has a greatest lower bound $\inf \mathcal{U}(S)$. We claim that in fact $\inf \mathcal{U}(S)$ is a least upper bound for S, or more succinctly,

$$\sup S = \inf \mathcal{U}(S).$$

Once again, Exercise 1.2 asks you to check this.
b) By Theorem 1a), $\mathbb{R}$ satisfies (LUB), and thus by part a) it satisfies (GLB). By Theorem 1b) $\mathbb{R}$ is the only ordered field satisfying (LUB), so certainly it is the only ordered field satifying (LUB) *and* (GLB). □

Exercise 1.2: a) Fill in the details of the proof of Theorem 2a).
b) Let F be an ordered field, and let S be a subset of F. Suppose that $\inf S$ exists. Show that $\sup -S$ exists and

$$\sup -S = -\inf S.$$

c) Use part b) to give a second proof that (GLB) $\implies$ (LUB).
d) Let F be an ordered field, and let S be a subset of F. Define

$$\mathcal{L}(S) = \{x \in F \mid x \text{ is a lower bound for } S.\}.$$

Suppose that $\sup \mathcal{L}(S)$ exists. Show that $\inf S$ exists and

$$\inf S = \sup \mathcal{L}(S).$$

e) Use part d) to give a second proof that (LUB) $\implies$ (GLB).

The technique which was used to prove (LUB) $\implies$ (GLB) is very familiar: we multiply everything in sight by -1. It seems likely that by now we have used this type of argument more than any other single trick or technique. When this has come up we have usually used the phrase "and similarly one can show..." Perhaps this stalwart ally deserves better. Henceforth, when we wish to multiply by -1 to convert $\leq$ to $\geq$, max to min, sup to inf and so forth, we will say **by reflection**. This seems more appealing and also more specific than "similarly..."!

In view of Theorem 2 it is reasonable to use the term **Dedekind completeness** to refer to either or both of (LUB), (GLB), and we shall do so.

Theorem 3. *A Dedekind complete ordered field is Archimedean.*

Proof. We will prove the contrapositive: let F be a non-Archimedean ordered field: thus there exists $x \in F$ such that $n \leq x$ for all $n \in \mathbb{Z}^+$. Then the subset $\mathbb{Z}^+$ of F is bounded above by x, so in particular it is nonempty and bounded above. So, if F were Dedekind complete then $\sup \mathbb{Z}^+$ would exist.

But we claim that in no ordered field F does $\mathbb{Z}^+$ have a supremum. Indeed, suppose that $M = \sup \mathbb{Z}^+$. It follows that for all $n \in \mathbb{Z}^+$, $n \leq M$. But then it is equally true that for all $n \in \mathbb{Z}^+$, $n + 1 \leq M$, or equivalently, for all $n \in \mathbb{Z}^+$, $n \leq M - 1$, so $M - 1$ is a smaller upper bound for $\mathbb{Z}^+$ than $\sup \mathbb{Z}^+$: contradiction! $\square$

1.2. Calisthenics With Sup and Inf.

The material and presentation of this section is partly based on [A, §1.3.13].

CONVENTION: Whenever $\sup S$ appears in the conclusion of a result, the statement should be understood as including the assertion that $\sup S$ exists, i.e., that S is nonempty and bounded above. Similarly for $\inf S$: when it appears in the conclusion of a result then an implicit part of the conclusion is the assertion that $\inf S$ exists, i.e., that S is nonempty and bounded below.[2]

Proposition 4. *Let S be a nonempty subset of $\mathbb{R}$.*
a) Suppose S is bounded above. Then for every $\epsilon > 0$, there exists $x \in S$ such that $\sup S - \epsilon < x \leq \sup S$.
b) Conversely, suppose $M \in \mathbb{R}$ is an upper bound for S such that for all $\epsilon > 0$, there exists $x \in S$ with $M - \epsilon < x \leq M$. Then $M = \sup S$.
c) Suppose S is bounded below. Then for every $\epsilon > 0$, there exists $x \in S$ such that $\inf S \leq x < \inf S + \epsilon$.
d) Conversely, suppose $m \in \mathbb{R}$ is a lower bound for S such that for all $\epsilon > 0$, there exists $x \in S$ with $m \leq x \leq m + \epsilon$. Then $m = \inf S$.

Proof. a) Fix $\epsilon > 0$. Since $\sup S$ is the *least* upper bound of S and $\sup S - \epsilon < \sup S$, there exists $y \in S$ with $\sup S - \epsilon < y$. It follows that

$$\sup S - \epsilon < \min(y, \sup S) \leq \sup S,$$

so we may take $x = \min(y, \sup S)$.
b) By hypothesis, M is an upper bound for S and nothing smaller than M is an

[2] Notice that a similar convention governs the use of $\lim_{x \to c} f(x)$, so this is nothing new.

upper bound for S, so indeed $M = \sup S$.
c),d) These follow from parts a) and b) by reflection. $\square$

Exercise 1.3: Let $a, b \in \mathbb{R}$. Suppose that for all $\epsilon > 0$, $a \leq b + \epsilon$. Show that $a \leq b$.

Proposition 5. *Let X, Y be nonempty subsets of $\mathbb{R}$, and define*

$$X + Y = \{x + y \mid x \in X,\ y \in Y\}.$$

a) Suppose X and Y are bounded above. Then

$$\sup(X + Y) = \sup X + \sup Y.$$

b) Suppose X and Y are bounded below. Then

$$\inf(X + Y) = \inf X + \inf Y.$$

Proof. a) Let $x \in X$, $y \in Y$. Then $x \leq \sup X$ and $x \leq \sup Y$, so $x + y \leq \sup X + \sup Y$, and thus $\sup(X + Y) \leq \sup X + \sup Y$. Now fix $\epsilon > 0$. By Proposition 4 there are $x \in X$ and $y \in Y$ with $\sup X - \frac{\epsilon}{2} < x$, $\sup Y - \frac{\epsilon}{2} < y$, so

$$\sup X + \sup Y \leq x + y + \epsilon.$$

Since this folds for all $\epsilon > 0$, by Exercise 1.3 $\sup X + \sup Y \leq \sup(X + Y)$.
b) This follows from part a) by reflection. $\square$

Let X, Y be subsets of $\mathbb{R}$. We write $X \leq Y$ if for all $x \in X$ and all $y \in Y$, $x \leq y$. (In a similar way we define $X < Y, X \geq Y$, $X > Y$.)

Exercise 1.4: Let X, Y be subsets of $\mathbb{R}$. Give necessary and sufficient conditions for $X \leq Y$ and $Y \leq X$ both to hold. (Hint: in the case in which X and Y are both nonempty, $X = Y$ is necessary but not sufficient!)

Proposition 6. *Let X, Y be nonempty subsets of $\mathbb{R}$ with $X \leq Y$. Then*

$$\sup X \leq \inf Y.$$

Proof. Seeking a contradiction, we suppose that $\inf Y < \sup X$. Put

$$\epsilon = \frac{\sup X - \inf Y}{2}.$$

By Proposition 4 there are $x \in X$, $y \in Y$ with $\sup X - \epsilon < x$ and $y < \inf Y + \epsilon$. Since $X \leq Y$ this gives

$$\sup X - \epsilon < x \leq y < \inf Y + \epsilon$$

and thus

$$\sup X - \inf Y < 2\epsilon = \sup X - \inf Y,$$

a contradiction. $\square$

Proposition 7. *Let X, Y be nonempty subsets of $\mathbb{R}$ with $X \subseteq Y$. Then:*
a) If Y is bounded above, then $\sup X \leq \sup Y$.
b) If Y is bounded below, then $\inf Y \leq \inf X$.

Exercise 1.5: Prove Proposition 7.

1.3. **The Extended Real Numbers.**

As exciting and useful as this whole business with sup and inf is, there is one slightly annoying point: $\sup S$ and $\inf S$ are not defined for *every* subset of $\mathbb{R}$. Rather, for $\sup S$ to be defined, S must be nonempty and bounded above, and for $\inf S$ to be defined, S must be nonempty and bounded below.

Is there some way around this? There is. It involves bending the rules a bit, but in a very natural and useful way. Consider the subset $\mathbb{N}$ of $\mathbb{R}$. It is not bounded above, so it does not have a least upper bound in $\mathbb{R}$. Because $\mathbb{N}$ contains arbitrarily large elements of $\mathbb{R}$, it is not completely unreasonable to say that its elements approach *infinity* and thus to set $\sup \mathbb{N} = +\infty$. In other words, we are suggesting the following definition:

- If $S \subset \mathbb{R}$ is unbounded above, then we will say $\sup S = +\infty$.

Surely we also want to make the following definition ("by reflection"!):

- If $S \subset \mathbb{R}$ is unbounded below, then we will say $\inf S = -\infty$.

These definitions come with a **warning**: $\pm\infty$ **are not real numbers**! They are just symbols suggestively standing for a certain type of behavior of a subset of $\mathbb{R}$, in a similar (but, in fact, simpler) way as when we write $\lim_{x\to c} f(x) = \pm\infty$ and mean that the function has a certain type of behavior near the point c.

To give a name to what we have done, we define the **extended real numbers** $[-\infty, \infty] = \mathbb{R} \cup \{\pm\infty\}$ to be the real numbers together with these two formal symbols $-\infty$ and ∞. This extension is primarily *order-theoretic*: that is, we may extend the $\leq$ relation to the extended real numbers in the obvious way:

$$\forall x \in \mathbb{R}, -\infty < x < \infty.$$

Conversely much of the point of the extended real numbers is to give the real numbers, as an ordered set, the pleasant properties of a closed, bounded interval $[a, b]$: namely we have a largest and smallest element.

The extended real numbers $[-\infty, \infty]$ *are not* a field. In fact, we cannot even define the operations of $+$ and $\cdot$ unrestrictedly on them. However, it is useful to define some of these operations:

$$\forall x \in \mathbb{R},\ -\infty + x = -\infty,\ x + \infty = \infty.$$

$$\forall x \in (0, \infty),\ x \cdot \infty = \infty,\ x \cdot (-\infty) = -\infty.$$

$$\forall x \in (-\infty, 0),\ x \cdot \infty = -\infty,\ x \cdot (-\infty) = \infty.$$

$$\infty \cdot \infty = \infty,\ \infty \cdot (-\infty) = -\infty,\ (-\infty) \cdot (-\infty) = \infty.$$

$$\frac{1}{\infty} = \frac{1}{-\infty} = 0.$$

None of these definitions are really surprising, are they? If you think about it, they correspond to facts you have learned about manipulating infinite limits, e.g. if $\lim_{x\to c} f(x) = \infty$ and $\lim_{x\to c} g(x) = 17$, then $\lim_{x\to c} f(x) + g(x) = \infty$. However, certain other operations with the extended real numbers *are not defined*, for similar reasons. In particular we **do not define**

$$\infty - \infty,$$

$$0 \cdot \infty,$$

$$\frac{\pm\infty}{\pm\infty}.$$

Why not? Well, again we might think in terms of associated limits. The above are **indeterminate forms**: if I tell you that $\lim_{x \to c} f(x) = \infty$ and $\lim_{x \to c} g(x) = -\infty$, then what can you tell me about $\lim_{x \to c} f(x) + g(x)$? Answer: nothing, unless you know what specific functions f and g are. As a simple example, suppose

$$f(x) = \frac{1}{(x-c)^2} + 2011, \ g(x) = \frac{-1}{(x-c)^2}.$$

Then $\lim_{x \to c} f(x) = \infty$, $\lim_{x \to c} g(x) = -\infty$, but

$$\lim_{x \to c} f(x) + g(x) = \lim_{x \to c} 2011 = 2011.$$

So $\infty - \infty$ cannot have a universal definition independent of the chosen functions.[3] In a similar way, when evaluating limits $0 \cdot \infty$ is an indeterminate form: if $\lim_{x \to c} f(x) = 0$ and $\lim_{x \to c} g(x) = \infty$, then $\lim_{x \to c} f(x)g(x)$ depends on *how fast* f approaches zero compared to how fast g approaches infinity. Again, consider something like $f(x) = (x-c)^2$, $g(x) = \frac{2011}{(x-c)^2}$. And similarly for $\frac{\infty}{\infty}$.

These are good reasons. However, there are also more purely algebraic reasons: there is no way to define the above expressions in such a way to make the field axioms work out. For instance, let $a \in \mathbb{R}$. Then $a + \infty = \infty$. If therefore we were allowed to substract ∞ from ∞ we would deduce $a = \infty - \infty$, and thus $\infty - \infty$ could be any real number: that's not a well-defined operation.

Remark: Sometimes above we have alluded to the existence of ordered fields F which do not satisfy the Archimedean axiom, i.e., for which there exist elements x such that $x > n$ for all $n \in \mathbb{Z}^+$. In speaking about elements like x we sometimes call them *infinitely large*. This is a totally different use of "infinity" than the extended real numbers above. Indeed, no ordered field F can have a largest element x, because it follows easily from the field axioms that for any $x \in F$, $x + 1 > x$. The moral: although we call $\pm\infty$ "extended real numbers", one should not think of them as being elements of a number system at all, but rather limiting cases of such things.

One of the merits of this extended definition of $\sup S$ and $\inf S$ is that it works nicely with calculations: in particular, all of the "calisthenics" of the previous section have nice analogues for unbounded sets. We leave it to the reader to investigate this phenomenon on her own. In particular though, let's look back at Proposition 7: it says that, under conditions ensuring that the sets are nonempty and bounded above / below, that if $X \subset Y \subset \mathbb{R}$, then

$$\sup X \leq \sup Y,$$

$$\inf Y \leq \inf X.$$

[3] In the unlikely event you think that perhaps $\infty - \infty = 2011$ always, try constructing another example...or wait until next semester and ask me again.

This definition could have motivated our definition of sup and inf for unbounded sets, as follows: for $n \in \mathbb{Z}$ and $X \subset \mathbb{R}$, put

$$X^n = \{x \in X \mid x \leq n\}, \ X_n = \{x \in X \mid x \geq n\}.$$

The idea here is that in defining X^n we are cutting it off at n in order to force it to be bounded above, but in increasingly generous ways. We have

$$X^0 \subset X^1 \subset \ldots \subset X$$

and also

$$X = \bigcup_{n=0}^{\infty} X^n;$$

in other words, every element of X is a subset of X^n for some n (this is precisely the Archimedean property). Applying Proposition 7, we get that for every nonempty subset X of $\mathbb{R}$,

$$\sup X^0 \leq \sup X^1 \leq \sup X^2 \leq \ldots \sup X^n \leq \ldots.$$

Suppose moreover that X is bounded above. Then some $N \in \mathbb{Z}^+$ is an upper bound for X, i.e., $X = X^N = X^{N+1} = \ldots$, so the sequence $\sup X^n$ is eventually constant, and in particular $\lim_{n\to\infty} \sup X^n = \sup X$. On the other hand, if X is bounded above, then the sequence $\sup X^n$ is not eventually constant; in fact it takes increasingly large values, and thus

$$\lim_{n\to\infty} \sup X^n = \infty.$$

Thus if we take as our definition for $\sup X$, $\lim_{n\to\infty} \sup X^n$, then for X which is unbounded above, we get $\sup X = \lim_{n\to\infty} \sup X^n = \infty$. By reflection, a similar discussion holds for $\inf X$.

There is, however, one last piece of business to attend to: we said we wanted $\sup S$ and $\inf S$ to be defined for *all* subsets of $\mathbb{R}$: what if $S = \varnothing$? There is an answer for this as well, but many people find it confusing and counterintuitive at first, so let me approach it again using Proposition 7. For each $n \in \mathbb{Z}$, consider the set $P_n = \{n\}$: i.e., P_n has a single element, the integer n. Certainly then $\inf P_n = \sup P_n = n$. So what? Well, I claim we can use these sets P_n along with Proposition 7 to see what $\inf \varnothing$ and $\sup \varnothing$ should be. Namely, to define these quantities in such a way as to obey Proposition 7, then for all $n \in \mathbb{Z}$, because $\varnothing \subset \{n\}$, we must have

$$\sup \varnothing \leq \sup\{n\} = n$$

and

$$\inf \varnothing \geq \inf\{n\} = n.$$

There is exactly one extended real number which is less than or equal to every integer: $-\infty$. Similarly, there is exactly one extended real number which is greater than or equal to every integer: ∞. Therefore the inexorable conclusion is

$$\sup \varnothing = -\infty, \ \inf \varnothing = \infty.$$

Other reasonable thought leads to this conclusion: for instance, in class I had a lot of success with the "pushing" conception of suprema and infima. Namely, if your set S is bounded above, then you start out to the right of every element of your set – i.e., at some upper bound of S – and keep pushing to the left until you can't push any farther without passing by some element of S. What happens if you try this with $\varnothing$? Well, every real number is an upper bound for $\varnothing$, so start anywhere and push to the left: you can keep pushing as far as you want, because you will never hit an element of the set. Thus you can push all the way to $-\infty$, so to speak. Similarly for infima, by reflection.

2. Intervals and the Intermediate Value Theorem

2.1. Convex subsets of $\mathbb{R}$.

We say that a subset S of $\mathbb{R}$ is **convex** if for all $x < y \in S$, the entire interval $[x, y]$ lies in S. In other words, a convex set is one that whenever two points are in it, all in between points are also in it.

Example 2.1: The empty set $\varnothing$ is convex. For any $x \in \mathbb{R}$, the singleton set $\{x\}$ is convex. In both cases the definition applies *vacuously*: until we have two distinct points of S, there is nothing to check!

Example 2.2: We claim any interval is convex. This is immediate – or it would be, if we didn't have so many different kinds of intervals to write down and check. One needs to see that the definition applies to invervals of all of the following forms:

$$(a, b),\ [a, b),\ (a, b],\ [a, b],\ (-\infty, b), (-\infty, b],\ (a, \infty), [a, \infty), (-\infty, \infty).$$

All these verifications are trivial appeals to things like the transitivity of $\leq$ and $\geq$.

Are there any nonempty convex sets other than intervals? (Just to be sure, we count $\{x\} = [x, x]$ as an interval.[4]) A little thought suggests that the answer should be *no*. But more thought shows that if so we had better use the Dedekind completeness of $\mathbb{R}$, because if we work over $\mathbb{Q}$ with all of the corresponding definitions then there are nonempty convex sets which are not intervals, e.g.

$$S = \{x \in \mathbb{Q} \mid x^2 < 2\}.$$

This has a familiar theme: replacing $\mathbb{Q}$ by $\mathbb{R}$ we would get an interval, namely $(-\sqrt{2}, \sqrt{2})$, but once again $\pm\sqrt{2} \notin \mathbb{Q}$. When one looks carefully at the definitions it is no trouble to check that *working solely in the rational numbers S is a convex set but is not an interval.*

Remark: Perhaps the above example seems legalistic, or maybe even a little silly. It really isn't: one may surmise that contemplation of such examples led Dedekind to his *construction* of the real numbers via **Dedekind cuts**. This construction may be discussed at the end of this course. Most contemporary analysts prefer a rival construction of $\mathbb{R}$ due to Cauchy using **Cauchy sequences**. I agree that Cauchy's construction is simpler. However, both are important in later mathematics: Cauchy's construction works in the context of a general **metric space** (and, with certain modifications, in a general **uniform space**) to construct an associated **complete space**. Dedekind's construction works in the context of a general **linearly ordered set** to construct an associated Dedekind-complete ordered set.

Theorem 8. *Any nonempty convex subset D of $\mathbb{R}$ is an interval.*

[4]However, we do not wish to say whether the empty set is an interval. Throughout these notes the reader may notice minor linguistic contortions to ensure that this issue never arises.

Proof. We have already seen the most important insight for the proof: we *must* use the Dedekind-completeness of $\mathbb{R}$ in our argument. With this in mind the only remaining challenge is one of organization: we are given a nonempty convex subset D of $\mathbb{R}$ and we want to show it is an interval, but as above an interval can have any one of nine basic shapes. It may be quite tedious to argue that one of nine things must occur!

So we just need to set things up a bit carefully: here goes: let $a \in [-\infty, \infty)$ be the infimum of D, and let $b \in (-\infty, \infty]$ be the supremum of D. Let $I = (a, b)$, and let $\overline{I}$ be the **closure** of I, i.e., if a is finite, we include a; if b is finite, we include b.
Step 1: We claim that $I \subset D \subset \overline{I}$. Let $x \in I$.

Case 1: Suppose $I = (a, b)$ with $a, b \in \mathbb{R}$. Let $z \in (a, b)$. Then, since $z > a = \inf D$, there exists $c \in D$ with $c < z$. Similarly, since $z < b = \sup D$, then there exists $d \in D$ with $z < d$. Since D is convex, $z \in D$. Now suppose $z \in D$. We must have $\inf D = a \leq z \leq b = \sup D$.

Case 2: Suppose $I = (-\infty, b)$, and let $z \in I$. Since D is unbounded below, there exists $a \in D$ with $a < z$. Moreover, since $z < \sup D$, there exists $b \in D$ such that $z < b$. Since D is convex, $z \in D$. Next, let $z \in D$. We wish to show that $z \in \overline{I} = (-\infty, b]$; in other words, we want $z \leq b$. But since $z \in D$ and $b = \sup D$, this is immediate. Thus $I \subset D \subset \overline{I}$.

Case 3: Suppose $I = (a, \infty)$. This is similar to Case 2 and is left to the reader.

Case 4: Suppose $I = (-\infty, \infty) = \mathbb{R}$. Let $z \in \mathbb{R}$. Since D is unbounded below, there exists $a \in D$ with $a < z$, and since D is unbounded above there exists $b \in D$ with $z < b$. Since D is convex, $z \in D$. Thus $I = D = \overline{I} = \mathbb{R}$.
Step 2: We claim that any subset D which contains I and is contained in $\overline{I}$ is an interval. Indeed I and $\overline{I}$ are both intervals, and the only case in which there is any subset D strictly in between them is $I = (a, b)$ with $a, b \in \mathbb{R}$ – in this case D could also be $[a, b)$ or $(a, b]$, and both are intervals. $\square$

Recall that a function $f : D \to \mathbb{R}$ satisfies the **Intermediate Value Property** (IVP) if for all $[a, b] \subset D$, for all L in between $f(a)$ and $f(b)$ is of the form $f(c)$ for some $c \in (a, b)$. As you may well have noticed, the IVP is closely related to the notion of a convex subset. The following result clarifies this connection.

Theorem 9. *For $f : D \subset \mathbb{R} \to \mathbb{R}$, the following are equivalent:*
(i) For all $[a, b] \subset D$, $f([a, b])$ is a convex subset of $\mathbb{R}$.
(ii) f satisfies the Intermediate Value Property.
(iii) For any interval $I \subset D$, $f(I)$ is an interval.

Proof. (i) $\implies$ (ii): For all $[a, b] \subset D$, $f([a, b])$ is a convex subset containing $f(a)$ and $f(b)$, hence it contains all numbers in between $f(a)$ and $f(b)$.
(ii) $\implies$ (iii): Suppose that f satisfies IVP, and let $I \subset D$ be an interval. We want to show that $f(I)$ is an interval. By Theorem 8 it suffices to show that $f(I)$ is convex. Assume not: then there exists $a < b \in I$ and some L in between $f(a)$ and $f(b)$ such that $L \neq f(c)$ for any $c \in I$. In particular $L \neq f(c)$ for any $c \in [a, b]$, contradicting the Intermediate Value Property.
(iii) $\implies$ (i): This is immediate: $[a, b]$ is an interval, so by assumption $f([a, b])$ is an interval, hence a convex subset. $\square$

2.2. **The (Strong) Intermediate Value Theorem.**

Theorem 10. *(Strong Intermediate Value Theorem) If $f : I \to \mathbb{R}$ is continuous, then f satisfies the Intermediate Value Property. In particular, $f(I)$ is an interval.*

Proof. Step 1: We make the following CLAIM: if $f : [a, b] \to \mathbb{R}$ is continuous, $f(a) < 0$ and $f(b) > 0$, then there exists $c \in (a, b)$ such that $f(c) = 0$.
PROOF OF CLAIM: Let $S = \{x \in [a, b] \mid f(x) < 0\}$. Since $a \in S$, S is nonempty. Moreover S is bounded above by b. Therefore S has a least upper bound $c = \sup S$. It is easy to see that we must have $f(c) = 0$. Indeed, if $f(c) < 0$, then – as we have seen several times – there exists $\delta > 0$ such that $f(x) < 0$ for all $x \in (c - \delta, c + \delta)$, and thus there are elements of S larger than c, contradicting $c = \sup S$. Similarly, if $f(c) > 0$, then there exists $\delta > 0$ such that $f(x) > 0$ for all $x \in (c - \delta, c + \delta)$, in which case any element of $(c - \delta, c)$ gives a smaller upper bound for S than c. By the process of elimination we must have $f(c) = 0$!
Step 2: We will show that f satisfies the Intermediate Value Property: for all $[a, b] \subset I$, and any L in between $f(a)$ and $f(b)$, we must find $c \in (a, b)$ such that $f(c) = L$. If $f(a) = f(b)$ there is nothing to show. If $f(a) > f(b)$, then we may replace f by $-f$ (this is still a continuous function), so it is enough to treat the case $f(a) < L < f(b)$. Now consider the function $g(x) = f(x) - L$. Since f is continuous, so is g; moreover $g(a) = f(a) - L < 0$ and $g(b) = f(b) - L > 0$. Therefore by Step 1 there exists $c \in (a, b)$ such that $0 = g(c) = f(c) - L$, i.e., such that $f(c) = L$.
Step 3: Finally, since $f : I \to \mathbb{R}$ satisfies the Intermediate Value Property, by Theorem 9 it maps every subinterval of I to an interval of $\mathbb{R}$. In particular $f(I)$ itself is an interval in $\mathbb{R}$. $\square$

Remark: Theorem 10 is in fact a mild improvement of the Intermediate Value Theorem we stated earlier in these notes. This version of IVT applies to continuous functions with domain *any* interval, not just an interval of the form $[a, b]$, and includes a result that we previously called the **Interval Image Theorem**.

2.3. **The Intermediate Value Theorem Implies Dedekind Completeness.**

Theorem 11. *Let F be an ordered field such that every continuous function $f : F \to F$ satisfies the Intermediate Value Property. Then F is Dedekind complete.*

Proof. We will prove the contrapositive: suppose F is not Dedekind complete, and let $S \subset F$ be nonempty and bounded above but without a least upper bound in F. Let $\mathcal{U}(S)$ be the set of upper bounds of S. We define a function $f : F \to F$ by:
• $f(x) = -1$, if $x \notin \mathcal{U}(S)$,
• $f(x) = 1$, $x \in \mathcal{U}(S)$.
Then f is continuous on F – indeed, a point of discontinuity would occur only at the least upper bound of S, which is assumed not to exist. Moreover f takes the value -1 – at any element $s \in S$, which cannot be an upper bound for S because then it would be the *maximum* element of S – and the value 1 at any upper bound for S (we have assumed that S is bounded above so such elements exist), but it never takes the value zero, so f does not satisfy IVP. $\square$

Exercise 2.3: Show in detail that the function $f : F \to F$ constructed in the proof of Theorem 11 is continuous at every element of F.

3. The Monotone Jump Theorem

Theorem 12. *(Monotone Jump) Let $f : I \to \mathbb{R}$ be weakly monotone, and let $c \in I$.*
a) Suppose c is an interior point of I. Then $\lim_{x \to c^-} f(x)$ and $\lim_{x \to c^+} f(x)$ both

exist, and

$$\lim_{x\to c^-} f(x) \leq f(c) \leq \lim_{x\to c^+} f(x).$$

b) Suppose c is the left endpoint of I. Then $\lim_{x\to c^+} f(x)$ *exists and is greater than or equal to $f(c)$.*
c) Suppose c is the right endpoint of I. Then $\lim_{x\to c^-} f(x)$ *exists and is less than or equal to $f(c)$.*

Proof. a) Step 0: As usual, we may f is weakly increasing. We define

$$L = \{f(x) \mid x \in I, x < c\},\ R = \{f(x) \mid x \in I, x > c\}.$$

Since f is weakly increasing, L is bounded above by $f(c)$ and U is bounded below by $f(c)$. Therefore we may define

$$\mathfrak{l} = \sup L, \mathfrak{r} = \inf R.$$

Step 1: For all $x < c$, $f(x) \leq f(c)$, $f(c)$ is an upper bound for L, so $\mathfrak{l} \leq f(c)$. For all $c < x$, $f(c) \leq f(x)$, so $f(c)$ is a lower bound for R, so $f(c) \leq \mathfrak{r}$. Thus

$$\mathfrak{l} \leq f(c) \leq \mathfrak{r}. \tag{1}$$

Step 2: We claim $\lim_{x\to c^-} f(x) = \mathfrak{l}$. To see this, let $\epsilon > 0$. Since $\mathfrak{l}$ is the least upper bound of L and $\mathfrak{l} - \epsilon < \mathfrak{l}$, $\mathfrak{l} - \epsilon$ is not an upper bound for L: there exists $x_0 < c$ such that $f(x_0) > \mathfrak{l} - \epsilon$. Since f is weakly increasing, for all $x_0 < x < c$ we have

$$\mathfrak{l} - \epsilon < f(x_0) \leq f(x) \leq \mathfrak{l} < \mathfrak{l} + \epsilon.$$

Thus we may take $\delta = c - x_0$.
Step 3: We claim $\lim_{x\to c^+} f(x) = \mathfrak{r}$: this is shown as above and is left to the reader.
Step 4: Substituting the results of Steps 2 and 3 into (1) gives the desired result.
b) and c): The arguments at an endpoint are routine modifications of those of part a) above and are left to the reader as an opportunity to check her understanding. $\square$

Theorem 13. *For a weakly monotone function $f : I \to \mathbb{R}$, TFAE:*
(i) $f(I)$ is an interval.
(ii) f is continuous.

Proof. As usual, it is no loss of generality to assume f is weakly increasing.
(i) $\implies$ (ii): If f is *not* continuous on all of I, then by the Monotone Jump Theorem $f(I)$ fails to be convex. In more detail: suppose f is discontinuous at c. If c is an interior point then either $\lim_{x\to c^-} f(x) < f(c)$ or $f(c) < \lim_{x\to c^+} f(x)$. In the former case, choose any $b \in I$, $b < c$. Then $f(I)$ contains $f(b) < f(c)$ but not the in-between point $\lim_{x\to c^-} f(x)$. In the latter case, choose any $d \in I$, $c < d$. Then $f(I)$ contains $f(c) < f(d)$ but not the in-between point $\lim_{x\to c^+} f(x)$. Similar arguments hold if c is the left or right endpoint of I: these are left to the reader. Thus in all cases $f(I)$ is not convex hence is not an interval.
(ii) $\implies$ (i): This follows immediately from Theorem 10. $\square$

With Theorems 10 and 13 in hand, we get an especially snappy proof of the Continuous Inverse Function Theorem. Let $f : I \to \mathbb{R}$ be continuous and injective. By Theorem 10, $f(I) = J$ is an interval. Moreover $f : I \to J$ is a bijection, with inverse function $f^{-1} : J \to I$. Since f is monotone, so is f^{-1}. Moreover $f^{-1}(J) = I$ is an interval, so by Theorem 13, f^{-1} is continuous!

4. Real Induction

Theorem 14. *(Principle of Real Induction) Let $a < b$ be real numbers, let $S \subset [a, b]$, and suppose:*
(RI1) $a \in S$,
(RI2) for all $x \in S$, if $x \neq b$ there exists $y > x$ such that $[x, y] \subset S$.
(RI3) For all $x \in \mathbb{R}$, if $[a, x) \in S$, then $x \in S$.
Then $S = [a, b]$.

Proof. Seeking a contradiction we suppose not: $S' = [a, b] \setminus S$ is nonempty. It is bounded below by a, so has a (finite!) greatest lower bound $\inf S'$. However:
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! □

Example 4.1: Let us reprove the Intermediate Value Theorem. Recall that the key special case of IVT, from which the full theorem easily follows, is this: if $f : [a, b] \to \mathbb{R}$ is continuous, $f(a) < 0$ and $f(b) > 0$, then there exists $c \in (a, b)$ with $f(c) = 0$. We prove this by real induction, as follows. Let $S = \{x \in [a, b] \mid f(x) \geq 0\}$. We know that S is proper in $[a, b]$, so applying real induction shows that one of (RI1), (RI2) and (RI3) must fail. We have $a \in S$ – so (RI1) holds – and if a continuous function is non-negative on $[a, c)$, then it is also non-negative at c: (RI3). So (RI2) must fail: there exists $y \in (a, b]$ such that $f(y) \geq 0$ but there is no $\epsilon > 0$ such that f is non-negative on $[y, y + \epsilon)$. This implies $f(y) = 0$.

Example 4.1, redux: In class I handled the proof of IVT by Real Induction differently, and in a way which I think gives a better first example of the method (most Real Induction proofs are *not* by contradiction). This strategy follows [Ka07]. Namely, IVT is equivalent to: let $f : [a, b] \to \mathbb{R}$ be continuous and nowhere zero. If $f(a) > 0$, then $f(b) > 0$. We prove this by Real Induction. Let

$$S = \{x \in [a, b] \mid f(x) > 0\}.$$

Then $f(b) > 0$ iff $b \in S$. We will show $S = [a, b]$ by real induction, which suffices.
(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!

The following result shows that Real Induction does not only uses the Dedekind completeness of $\mathbb{R}$ but actually carries the full force of it.

Theorem 15. *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) through (RI3) above must be all of $[a, b]$.

Proof. (i) $\Longrightarrow$ (ii): This is simply a restatement of Theorem 14.
(ii) $\Longrightarrow$ (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 $x = \inf T$. $\square$

Remark: Like Dedekind completeness, the notion of "Real Induction" depends only on the ordering relation $<$ and not on the field operations $+$ and $\cdot$. In fact, given an arbitrary ordered set $(F,<)$ – i.e., we need not have operations $+$ or $\cdot$ at all – it makes sense to speak of Dedekind completeness and also of whether the Principle of Real Induction holds. In a recent note [Cl11], I proved that Theorem 15 holds in this general context: an ordered set F is Dedekind complete iff the only it satisfies a "Principle of Ordered Induction".

5. The Extreme Value Theorem

Theorem 16. *(Extreme Value Theorem)*
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 that $x \in (a,b]$ and $[a,x) \subset S$. Now **beware**:[5] 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 really 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, as usual, the key feature of this counterexample is a lack of continuity: this f is not continuous at 2. Having said this, it becomes clear that 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

[5]I am embarrassed to admit that the previous version of my lecture notes fell into exactly this trap. These notes were taken from a piece I wrote last year after giving a talk for math graduate students at UGA. They have never been formally published, but they are available on the web and have been read by several (dozen?) people, none of whom pointed out this mistake. Oh, well...

$[x-\delta, x]$. But since $a < x - \delta < x$ we know also that 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 $f(x_M) = M$. $\square$

As with the Intermediate Value Theorem, one can show that Dedekind completeness is not just sufficient but *necessary* in order for the Extreme Value Theorem to hold.

Theorem 17.
Let F be an ordered field which is not *Dedekind complete, and let $a < b \in F$.*
a) There is a continuous function $f : [a,b] \to F$ which is unbounded.
b) There is a bounded continuous function $f : [a,b] \to F$ which does not attain either a minimum or a maximum value.

Exercise 5.1: Prove Theorem 17. (Note: this is is a challenging exercise. In fact, I have not myself written out a complete proof of Theorem 17, which is not a standard result. I do hope it is actually true...)

6. Uniform Continuity

6.1. The Definition; Key Examples.

noindent Let I be an interval and $f : I \to \mathbb{R}$. Then f is **uniformly continuous on I** if for every $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x_1, x_2 \in I$, if $|x_1 - x_2| < \delta$ then $|f(x_1) - f(x_2)| < \epsilon$.

In order to show what the difference is between uniform continuty on I and "mere" continuity on I – i.e., continuity at every point of I – let us rephrase the standard ϵ-δ definition of continuity using the notation above. Namely:

A function $f : I \to \mathbb{R}$ is **continuous on I** if for every $\epsilon > 0$ and every $x_1 \in I$, there exists $\delta > 0$ such that for all $x_2 \in I$, if $|x_1 - x_2| < \delta$ then $|f(x_1) - f(x_2)| < \epsilon$.

These two definitions are eerily (and let's admit it: confusingly, at first) similar: they use all the same words and symbols. The only difference is in the *ordering of the quantifiers*: in the definition of continuity, player two gets to hear the value of ϵ and also the value of x_1 before choosing her value of δ. In the definition of uniform continuity, player two only gets to hear the value of ϵ: thus, her choice of δ must work simultaneously – or, in the lingo of this subject, **uniformly** – across all values of $x_1 \in I$. That's the only difference. Of course, switching the order of quantifiers in general makes a big diffference in the meaning and truth of mathematical statements, and this is no exception. Let's look at some simple examples.

Example 6.1: Let $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = mx + b$, $m \neq 0$. We claim that f is **uniformly continuous** on $\mathbb{R}$. In fact the argument that we gave for continuity long ago shows this, because for every $\epsilon > 0$ we took $\delta = \frac{\epsilon}{|m|}$. Although we used this δ to show that f is continuous at some arbitrary point $c \in \mathbb{R}$, evidently the choice of δ does not depend on the point c: it works uniformly across all values of c. Thus f is uniformly continuous on $\mathbb{R}$.

Example 6.2: Let $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. This time I claim that our usual proof *did not* show uniform continuity. Let's see it in action. To show that f is continuous at c, we factored $x^2 - c$ into $(x - c)(x + c)$ and saw that to get some control on the other factor $x + c$ we needed to restrict x to some bounded interval around c, say $[c-1, c+1]$. On this interval $|x+c| \leq |x|+|c| \leq |c|+1+|c| \leq 2|c|+1$. So by taking $\delta = \min(1, \frac{\epsilon}{2|c|+1})$ we found that if $|x - c| < \delta$ then

$$|f(x) - f(c)| = |x - c||x + c| \leq \frac{\epsilon}{2|c| + 1} \cdot (2|c| + 1) = \epsilon.$$

But the above choice of δ *depends on* c. So it doesn't show that f is uniformly continuous on $\mathbb{R}$. In fact the function $f(x) = x^2$ is *not* uniformly continuous on $\mathbb{R}$. For instance, take $\epsilon = 1$. If it were uniformly continuous, there would have to be some $\delta > 0$ such that for all $x_1, x_2 \in \mathbb{R}$ with $|x_1 - x_2| < \delta$, $|x_1^2 - x_2^2| < \epsilon$. But this is not possible: take any $\delta > 0$. Then for any $x \in \mathbb{R}$, x and $x + \frac{\delta}{2}$ are less than δ apart, and $|x^2 - (x + \frac{\delta}{2})^2| = |x\delta + \frac{\delta^2}{4}|$. But if I get to choose x after you choose δ, this expression can be made arbitrarily large. In particular, if $x = \frac{1}{\delta}$, then it is strictly greater than 1. So f is not uniformly continuous on $\mathbb{R}$.

Remark: In fact a polynomial function $f : \mathbb{R} \to \mathbb{R}$ is uniformly continuous on $\mathbb{R}$ if and only if it has degree at most one. The reasoning is similar to the above.

So that's sad: uniform continuity is apparently quite rare. But wait! What if the domain is a closed, bounded interval I? For instance, by restricting $f(x) = x^2$ to any such interval, it *is* uniformly continuous. Indeed, we may as well assume $I = [-M, M]$, because any I is contained in such an interval, and uniform continuity on $[-M, M]$ implies uniform continuity on I. Now we need only use the fact that we are assuming $|c| \leq M$ to remove the dependence of δ on c: since $|c| \leq M$ we have $\frac{\epsilon}{2|c|+1}) \geq \frac{1}{2M+1}$, so for $\epsilon > 0$ we may take $\delta = \min(1, \frac{1}{2M+1})$. This shows that $f(x) = x^2$ is uniformly continuous on $[-M, M]$.

It turns out that one can always recover uniform continuity from continuity by restricting to a closed bounded interval: this is the last of our Interval Theorems.

6.2. The Uniform Continuity Theorem.

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.

The following small technical argument will be applied twice in the proof of the Uniform Continuity Theorem, so advance treatment of this argument should make the proof of the Uniform Continuity Theorem more palatable.

Lemma 18. *(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 19. *(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 Patching 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 Patching 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 $x \in S(\epsilon)$. □

7. The Bolzano-Weierstrass Theorem

Let $S \subset \mathbb{R}$. We say that $x \in \mathbb{R}$ is a **limit point** of S if for every $\delta > 0$, there exists $s \in S$ with $0 < |s - x| < \delta$. Equivalently, x is a limit point of S if every open interval I containing x also contains an element s of S which is not equal to x.

Proposition 20. *For $S \subset \mathbb{R}$ and $x \in \mathbb{R}$, the following are equivalent:*
(i) Every open interval I containing x also contains infinitely many *points of S.*
(ii) x is a limit point of S.

Example 7.1: If $S = \mathbb{R}$, then every $x \in \mathbb{R}$ is a limit point. More generally, if $S \subset \mathbb{R}$ is **dense** – i.e., if every nonempty open interval I contains an element of S – then every point of $\mathbb{R}$ is a limit point of S. In particular this holds when $S = \mathbb{Q}$ and when $S = \mathbb{R} \setminus \mathbb{Q}$. Note that these examples show that a limit point x of S may or may not be an element of S: both cases can occur.

Example 7.2: If $S \subset T$ and x is a limit point of S, then x is a limit point of T.

Nonexample 7.3: No finite subset S of $\mathbb{R}$ has a limit point.

Nonexample 7.4: The subset $\mathbb{Z}$ has no limit points: indeed, for any $x \in \mathbb{R}$, take $I = (x-1, x+1)$. Then I is bounded so contains only finitely many integers. More generally, let S be a subset such that for all $M > 0$, $S \cap [-M, M]$ is finite. Then S has no limit points.

In fact this is the most general possible nonexample: if for some $M > 0$ $S \cap [-M, M]$ is infinite, then S must have a limit point. In other words:

Theorem 21. *(Bolzano-Weierstrass) Every infinite subset A of a closed bounded interval $[a, b]$ has a limit point.*

Proof. Let S be the set of $x \in [a, b]$ such that *if* $A \cap [a, x]$ is infinite, then $A \cap [a, x]$ has a limit point. It suffices to show that $S = [a, b]$: then $A = A \cap [a, b]$ is infinite and thus has a limit point. We will use Real Induction.

(RI1): Since $A \cap [a, a] \subset \{a\}$ is finite, the "induction hypothesis" holds vacuously. (RI2): suppose $x \in S$. If $A \cap [a, x]$ is infinite, then by hypothesis $A \cap [a, x]$ has a limit point and hence so does $[a, b]$. So we may assume $A \cap [a, x]$ is finite. Now either there exists $\delta > 0$ such that $A \cap [a, x + \delta]$ is finite – which verifies our induction hypothesis – or every interval $[x, x + \delta]$ contains infinitely many points of A, in which case x is a limit point of A. (RI3): Suppose $[a, x) \subset S$. Then, if there exists some $y < x$ such that $A \cap [a, y]$ is infinite, then by hypothesis $A \cap [a, y]$ has a limit point and thus so does A. So we may assume that $A \cap [a, y]$ is finite for all $y < x$. As above, this means either that $A \cap [a, x]$ is finite, or that every interval $(x - \delta, x)$ intersects A, in which case x is a limit point of A. □

Remark: Theorem 21 is the "set version" of Bolzano-Weierstrass. There is a more common "sequence version" of Bolzano-Weierstrass: **every bounded sequence admits a convergent subsequence**. Sadly, we have not yet spoken of sequences, their convergence, and subsequences, but we will: this will be a major theme of the second half of the course. Once we acquire enough vocabulary to understand the above boldfaced statement, we will see almost immediately that it is equivalent to Theorem 21 above. We can (and will!) then use the Bolzano-Weierstrass Theorem for sequences to give new proofs of the Extreme Value Theorem and the Uniform Continuity Theorem which are quicker, cleaner and more conceptual.

References

[A] T.M. Apostol, *Calculus. Volume One.*

[Cl10] P.L. Clark, *Real induction.* `http://math.uga.edu/~pete/realinduction.pdf`

[Cl11] P.L. Clark, *Induction and completeness in ordered sets.* `http://math.uga.edu/~pete/induction_completeness_brief.pdf`

[Ha11] D. Hathaway, *Using Continuity Induction.* College Math. Journal 42 (2011), 229–231.

[Ka07] I. Kalantari, *Induction over the continuum.* Induction, algorithmic learning theory, and philosophy, 145–154, Log. Epistemol. Unity Sci., 9, Springer, Dordrecht, 2007.

[S] M. Spivak, *Calculus.* Fourth edition.