1. Analysis (cont)

1.6. Two Preliminary Results

**Theorem:** (Theorem A for future ref): Given two subsets, $P$ and $Q$, of the real line, if $P \subset Q$ then $\sup(P) \leq \sup(Q)$ and $\inf(P) \geq \inf(Q)$.

Obvious just by looking at a picture: if $P$ is smaller than $Q$ then then every upper bound to $Q$ must also be an upper bound to $P$ and there may be upper bounds to $P$ that aren’t upper bounds to $Q$. Doing the proof is just a matter of saying the obvious in symbols.

**Proof:** We’ll just prove the first half about supers. Pick $P, Q \subset \mathbb{R}$ such that $P \subset Q$. There are two cases to consider:

*Case A:* $\sup(Q) = \infty$; in this case, there’s nothing to prove because *any* possible sup for $P$ has to be less than or equal to infinity.
Case B: \( \sup(Q) = \bar{b} \). We’ll consider \( b > \bar{b} \) and prove that \( b \) cannot be a least upper bound for \( P \). This will establish that \( \sup(P) \leq \sup(Q) \). Well, \( \bar{b} \geq q \), for every \( q \in Q \). Since \( P \subset Q \), \( \bar{b} \geq p \), for every \( p \in P \). Hence \( \bar{b} \) is an upper bound for \( P \). Since \( b > \bar{b} \), \( b \) cannot be a least upper bound for \( P \). □

Notice that if the statement of Theorem A had omitted any mention of the real line, then the theorem would have been false. Why?

**Theorem:** (Theorem B for future ref): Given \( S \subset \mathbb{R} \), \( b \in \mathbb{R} \) is the least upper bound (supremum) of \( S \) iff \( b \) is an upper bound for \( S \) and if for every \( \epsilon > 0 \), there exists \( s \in S \) such that \( b - s < \epsilon \).

**Proof:** Need to prove this in both directions, i.e., (a) if \( b \) satisfies the requirements of the theorem then it is a supremum; (b) if it *doesn’t* satisfy the requirements of the theorem, then it *isn’t* a supremum.

Proof of (a): Assume that for every \( \epsilon > 0 \), there exists \( s \in S \) such that \( b - s < \epsilon \). We need to show that \( b \leq b' \) for every upper bound \( b' \) for \( S \). We’ll do this by showing that if \( b' < b \) then \( b' \) cannot be an upper bound for \( S \). Pick \( b' = b - \epsilon \), for some \( \epsilon > 0 \). By assumption, \( \exists s \in S \) s.t. \( b - s < \epsilon = b - b' \), i.e., \( -s < -b' \) or \( s > b' \) proving that \( b' \) isn’t an upper bound for \( S \).

Proof of (b): Now assume that one of the requirements of the theorem isn’t satisfied. The first requirement is that \( b \) is an upper bound for \( S \). If this requirement is violated, then trivially \( b \) can’t be a supremum; Now suppose that the \( \forall \epsilon \) condition fails, i.e., \( \exists \epsilon > 0 \) such that \( \forall s \in S \), \( b - s \geq \epsilon \) (see Fig. 1). We’ll show that \( b \) cannot be a *least* upper bound for \( S \), i.e., we’ll show that there exists \( b' < b \) such that \( b' \) is an upper bound for \( S \). Pick \( b' = b - \epsilon \). Since \( b - s \geq \epsilon \), for all \( s \in S \), or, equivalently, \( b - \epsilon \geq s \), for all \( s \in S \), then \( b' = b - \epsilon \geq s \), for all \( s \in S \). Hence \( b' \) is an upper bound for \( S \) and is also smaller than \( b \). So we’ve proved that \( b \) isn’t a least upper bound. □
1.7. Cauchy Sequences

The notion of a Cauchy sequence is another tool that we will need later.

Definition: A sequence \( \{x_n\}, x_n \in \mathbb{R} \) is called a Cauchy sequence with respect to a metric \( d \) if \( \forall \epsilon > 0 \exists N \in \mathbb{N} \) such that for \( n, m \geq N \), \( d(x_n, x_m) < \epsilon \).

That is, the sequence bunches up as \( n \) progresses. More specifically, if you have the graph of a Cauchy sequence, you can draw a line above the sequence (i.e., which bounds the sequence from above) and a line below the sequence (i.e., which bounds the sequence from below) with the property that the two lines get closer and closer together.

Example: The sequence \( \{x_n\} \) defined by \( x_n = (-1)^n/n \).

Notice the difference between a Cauchy sequence and a convergent sequence: a convergent sequence converges to a point; a Cauchy sequence bunches up, but there may be nothing that it converges to. For example, consider the sequence \( 1 - 1/2, ..., 1 - 1/n, ... \) in the set \((0, 1)\). The sequence satisfies the defn of a Cauchy sequence, but it is not a convergent sequence. The thing it is trying to converge to just isn’t in the set.

That is, sequences that are trying to converge to something, whether or not the something they are trying to converge to doesn’t exist, are called Cauchy sequences.
A better example of a Cauchy sequence that isn’t a convergent sequence is the sequence of **continuous functions** we defined in a previous lecture, \( \{f_1, f_2, \ldots f_n\ldots \} \), where

\[
f_n = \begin{cases} 
-1 & \text{if } x \leq -1/n \\
x & \text{if } -1/n < x < 1/n \\
1 & \text{if } x \geq 1/n
\end{cases}
\]

While we don’t know very much yet about metrics on the space of functions\(^1\), you can see intuitively that the \( f_n \)'s are *in a certain sense* getting closer and closer together, i.e., their graphs are becoming indistinguishable from each other. (That is, this sequence is Cauchy with respect to *some* metrics on the space of continuous functions—such as the metric-like function defined in footnote 1—and isn’t Cauchy with respect to others.) However, the sequence doesn’t have a limit in the space of continuous functions: if \( f \) were a limit function, it would *have to* have the property that \( f(x) = 1 \), for all \( x > 0 \) and \( f(x) = -1 \) for all \( x < 0 \). But what happens at zero?? Clearly the function can’t be continuous. This is a typical example of a sequence of objects that all belong to the same set (i.e., the set of continuous functions), in which the elements get closer and closer together, but don’t converge to anything in that original set.

The real number system has the following very special property:

**Theorem:** Every sequence in \( \mathbb{R} \) that is Cauchy w.r.t. the Pythagorian metric converges with respect to that metric to a number \( x \in \mathbb{R} \).

This theorem is true much more generally, but the *proof* I give below depends on a property of the Pythagorian metric that does not hold for all metrics. In particular, it uses the fact that under the Pythagorian metric, Cauchy sequences are necessarily *bounded*. This is property does not hold for all metrics. Specifically, consider the metric \( \rho \) (known to generations of ARE students as “Leo’s

---

\(^1\) In class I gave one example of a metric on functions: given \( f : X \to \mathbb{R} \) and \( g : X \to \mathbb{R} \), define \( \delta(f, g) = \sup_{x \in X} |(f(x) - g(x))| \). You should check that \( \delta \) indeed satisfies the four requirements of a metric.
favorite metric”) defined on \( \mathbb{R}_+ \times \mathbb{R}_+ \) by \( \rho(x, y) = \begin{cases} 
 0 & \text{if } x = y = 0 \\
 1/x & \text{if } x \neq y = 0 \\
 1/y & \text{if } y \neq x = 0 \\
 |1/x - 1/y| & \text{otherwise} 
\end{cases} \). Now consider the sequence \( \{x_n\} \) defined by \( x_n = n \). This sequence is clearly Cauchy w.r.t. \( \rho \)—and indeed converges to a point \( x \in \mathbb{R}_+ \) (which one?)—but it isn’t bounded, at least not in the sense that I defined boundedness above.

Why do we care? One reason is that in many cases in economics, we want to prove that some scalar, vector, or function exists. E.g., proving the existence of equilibrium prices for an economy: a price could be a scalar, a vector or in some cases a function. Often, the easiest way to proceed is to construct a Cauchy sequence of, say, equilibrium prices of some sequence of nearby economies, and then argue that that the limit of this sequence is the equilibrium we need. But if the limit isn’t the same kind of object as the members of the Cauchy sequence, then we can’t pursue this route. E.g., our equilibrium price function might be required to be a continuous function of, say, some signal; but the limit of a Cauchy sequence of continuous functions need not be a continuous function, so in this case, the Cauchy approach won’t work.

The above theorem is another way of saying that the real number system doesn’t have any holes in it. Obvious as it is, the proof involves some work. Essentially the proof formalizes the following image: draw lines just above and just below the graph of the sequence; we’ll argue that these lines have to get closer and closer together as you go out to infinity. Trick is to make this idea precise.

Proof: When defined on \( \mathbb{R} \times \mathbb{R} \), the Pythagorian metric reduces to \( d(x, y) = |x - y| \).

We first need to show that every Cauchy sequence is bounded both above and below. Set \( \epsilon = 1 \); \( \exists M \in \mathbb{N} \) such that for \( n, m > M \), \( d(x_n, x_m) < 1 \). The set \( \{x_1, \ldots, x_M, x_{M+1}\} \) is a finite set and has
a maximum. Call it \( \bar{x} \). \( \text{(Question: why wouldn’t the proof below be correct if we kept everything the same, except we defined } \bar{x} \text{ to be the maximum of the set } \{ x_1, \ldots, x_M \}? \) We’ll show that \( \bar{x} + 1 \) is an upper bound for the entire sequence: let \( m = M + 1 \), and note that for all \( n > M \), since \( d(x_m, x_n) = |x_n - x_m| < 1 \), we have that for all \( n > M \), \( x_n < x_{M+1} + 1 \leq \bar{x} + 1 \). \( \text{(Note that this is where we use the specific functional form for the metric.) Similarly, the set is bounded below.} \)

Now for each \( m \in \mathbb{N} \), let \( A_m = \{ x_m, x_{m+1}, \ldots \} \) and let \( a_m = \sup(A_m) \). \( \text{From Theorem A, the sequence } \{ a_m \}_{m=1}^{\infty} \text{ is a nonincreasing sequence. Moreover the sequence is bounded below (previous paragraph). Hence the sequence } \{ a_m \}_{m=1}^{\infty} \text{ converges to a point } a \in \mathbb{R}. \) \( \text{(From the Axiom of Completeness last time.)} \)\(^2\) We’ll show \( \{ x_n \} \) also converges to \( a \). Fix \( \epsilon > 0 \).

- pick \( N_1 \) sufficiently large that \( \forall n, m \geq N_1, d(x_n, x_m) < \epsilon/3 \) (using the fact that the sequence is Cauchy.)
- pick \( N_2 \) sufficiently large that \( \forall n \geq N_2, d(a_n, a) < \epsilon/3. \) (using the fact that the sequence of sup’s converges to \( a \)).
- let \( N = \max\{N_1, N_2\} \)
- Note that for \( n, m > N \) both of the following are true:
  - \( \forall n, m \geq N, d(x_n, x_m) < \epsilon/3 \)
  - \( \forall n \geq N, d(a_n, a) < \epsilon/3. \)

That is, all of the points beyond the \( N \)’th in the sequence are bunched together and the supremum of all of these points is close to \( a \). To tie the proof together we just have to relate all of these points in the tail of the sequence to the supremum of these points and we are done.

\(^2\) Recall from last lecture the The Axiom of Completeness: Every non-decreasing sequence \( \{ x_n \} \) in \( \mathbb{R} \) that is bounded above converges to a point \( x \in \mathbb{R} \) in the Pythagorian metric.
pick $K \geq N$ such that $d(a_N, x_K) < \epsilon/3$: we can do this for the following reason: $a_N$ is the sup of $A_N = \{x_N, x_{N+1}, \ldots\}$; by Theorem B, there’s some point in $A_N$ which is within $\epsilon/3$ of $a_N$; Let $K$ denote the index of this point.

Now apply the triangle inequality which just formalizes the ideas that: the points in the sequence are bunched; the suprema are close to $a$; the top end the sequence is close to the supremum. Formally, for arbitrary $n > N$, we have

$$d(x_n, a) \leq d(x_n, x_K) + d(x_K, a_N) + d(a_N, a) \leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon \quad \square$$

**Question:** What part of the proof invokes the special property of the real line. That is, the theorem isn’t true for Cauchy Sequences in general.

**Answer:** We utilize the fact that every nonincreasing, bounded sequence in $\mathbb{R}$ converges to a point in $\mathbb{R}$. In this case the nonincreasing sequence is the sequence of sups.

To summarize this presentation on Cauchy sequences, for many students, it’s really hard to appreciate the difference between a Cauchy sequence and a convergent sequence. At first sight, it seems that the definition of Cauchy is just another way of writing the definition of convergence. But as we’ve seen, while in some special situations, the two concepts are equivalent they aren’t equivalent in general: convergentness is a strictly stronger property than Cauchyness.

To determine when the concepts are equivalent and when they are not you **have to consider the properties of the space that the sequence lives in.** If a sequence lives in a space that has no holes in it, then Cauchy sequences will be convergent; such spaces are called **complete** spaces. But as
we've seen there are many incomplete spaces—the space of continuous functions, the rationals, for example. In such spaces, Cauchy sequences *may* be convergent, but they need not necessarily be.