1. 01/18/11: Building Up to the Reals, Part One

1.1. Symbols.

We will start with a list of mathematical symbols that we will use throughout the course.

$\forall$: for all
$\exists$: there exists
$\Rightarrow$ or $\Leftarrow$: implies
$\iff$: if and only if
$\in$: belongs to
$\{x \,|\, \text{condition}\}$: set notation
$\cap$: intersection
$\cup$: union
$\amalg$: disjoint union, or if $A = B \amalg C$, then $A = B \cup C$ and $B \cap C = \emptyset$
$S \setminus A$: set difference, or $\{x \in S \,|\, x \notin A\}$

An example of a formula written (almost) completely in symbols:

$$(\forall \epsilon > 0)(\exists \delta > 0) \text{ s.t. } (|x-y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon)$$

1.2. The Natural Numbers, or $\mathbb{N}$.

The main idea for the first part of the course will be $\mathbb{R}$, or the real numbers. There are three main ways to prove their existence: decimal expansions, Dedekind cuts, or sequences of rationals. We will eventually use a hybrid of the second and third methods, but that is for later. For now...

$\mathbb{N} = \{1, 2, 3, 4, 5, 6, \dots\}^*$. This is a commutative semiring with an identity (built off 1 with operators $+$ and $\cdot$). There is a notion of order, or $x > y$, contained in the algebraic structure of $\mathbb{N}$.

Axiom (Induction)**.** *If $S \subseteq \mathbb{N}$ and:*

(1) $1 \in S$
(2) $x \in S \Rightarrow (x+1) \in S$

hold, then $S = \mathbb{N}$.

In other words, if the two conditions above are true, then nothing can fail to be in S. Note that we will say $0 \notin \mathbb{N}$.

1.3. The Integers, or $\mathbb{Z}$.

$\mathbb{Z} = \mathbb{N} \cup \{0\} \cup (-\mathbb{N})$. This is a commutative ring with multiplicative identity, meaning that it is an abelian group under $(+, 0)$ and is associative, commutative, and distributive (i.e. $x(y+z) = xy + xz$) under multiplication. We also define $0x = 0 \; \forall x$.

Note that a group G is a set with an operation $\star$ such that $g \star h \in G$ for $g, h \in G$, along with these properties:

(1) G has an identity e, for which the following holds: $e \star g = g \star e = g \; \forall g$.
(2) $(g \star h) \star k = g \star (h \star k)$
(3) $(\forall g \in G)(\exists h \in G)$ such that $g \star h = e = h \star g$.

---

*Read the book for notes on their structure.

An abelian group is a group with the property that $g \star h = h \star g$. From the properties of a group, it follows that the h in (3) is unique, namely $h = g^{-1}$ or $-g$, and that $(a \star b = a \star c) \Rightarrow b = c$.

There is an order on $\mathbb{Z}$, or some notion such that $x > y$. Namely,

$$x \geq y \iff x - y \in \mathbb{N} \cup \{0\}.$$

We then obtain three properties about this ordering:

(1) $x \leq x$
(2) $(x \leq y$ and $y \leq x) \Rightarrow x = y$
(3) $(x \leq y$ and $y \leq z) \Rightarrow x \leq z$

These are the axioms of a partially ordered set. Another example is the ordering on the subsets of a set with $\subseteq$ instead of $\leq$.

1.4. **The Rationals, or $\mathbb{Q}$, an Ordered Field.**

$\mathbb{Q} = \{\frac{m}{n} \,|\, (m,n) = 1$ and $m, n \in \mathbb{Z}$ and $n \neq 0\}$. Note that (x, y) is equivalent to the expression $\gcd(x, y)$.

We will define an order on $\mathbb{Q}$: Let $\mathcal{P} = \{\frac{m}{n} \,|\, m \in \mathbb{N}, n \in \mathbb{N}\}$. Then,

$$x \leq y \iff y - x \in \mathcal{P} \text{ or } y = x.$$

Definition. *A field F is a set with two operations $+$, $\cdot$ and two distinguished elements $0, 1$ such that, for $0 \neq 1$,*

- *$(F, +0)$ is an abelian group*
- *$(F \setminus \{0\}, \cdot, 1)$ is an abelian group*
- *$x(y + z) = xy + xz$*

As a precaution, we can extend $\cdot$ to all of F by $0x = 0 \ \forall x$.

$\mathbb{Q}$ is not only a field, however, but also an ordered field.

Definition. *An ordered field is a field F with a subset $\mathcal{P} \subseteq F$ with the property*

$$F = \mathcal{P} \amalg \{0\} \amalg -\mathcal{P}$$

where $-\mathcal{P} = \{-p \,|\, p \in \mathcal{P}\}$, and axioms

(1) *$\mathcal{P} + \mathcal{P} \subseteq P$, where $\mathcal{P} + \mathcal{P} = \{p_1 + p_2 \,|\, p_1, p_2 \in \mathcal{P}\}$*
(2) *$\mathcal{P} \cdot \mathcal{P} \subseteq \mathcal{P}$, where $\mathcal{P} \cdot \mathcal{P} = \{p_1 \cdot p_2 \,|\, p_1, p_2 \in \mathcal{P}\}$*

Proposition. *$1 \in \mathcal{P}$.*

Proof. We will proceed by contradiction. Note that if $x \in \mathcal{P}$, then $(-1)x \in -\mathcal{P}$. Then, if $1 \in -\mathcal{P}$, then $(-1) \in \mathcal{P}$, which is a contradiction because $(-1) \cdot (-1) = 1$, but $1 \notin \mathcal{P}$. □

Proposition. *$x^{-1} \in \mathcal{P}$.*

Proof. We will proceed by contradiction, and assume that $x^{-1} \in -\mathcal{P}$. If that is the case, then $x \cdot (-x^{-1}) = -1$. Then, we obtain a contradiction because $x \in \mathcal{P}$ and $(-x^{-1}) \in \mathcal{P}$, but $-1 \notin \mathcal{P}$. □

Thus, given an ordered field $(F, +, \cdot, 0, 1, \mathcal{P})$, we can define:

$$x \leq y \ [x < y] \iff y - x \in \mathcal{P} \cup \{0\} \ [y - x \in \mathcal{P}]\,.$$

This ordering gives us the following consequences, taken as axioms in other sources such as the book:

(1) $a \leq b \iff -b \leq -a$
(2) $c \geq 0 \Rightarrow ac \leq bc$
(3) $a^2 \geq 0$
(4) $(0 \leq a, 0 \leq b) \Rightarrow ab \geq 0$
(5) $a > 0 \Rightarrow a^{-1} > 0$
(6) $0 < a < b \Rightarrow 0 < b^{-1} < a^{-1}$

2. 01/20/11: Building Up to the Reals, Part Two

2.1. More on $\mathbb{Q}$.

Last time, we discussed ordered fields $(F, +, \cdot, 0, 1, \mathcal{P})$ where $F = \mathcal{P} \amalg \{0\} \amalg -\mathcal{P}$, and defined an order $x \leq y \iff y - x \in \{0\} \cup \mathcal{P}$. Note that if $x, y \in F$, then either $x \leq y$ or $y \leq x$. This property is known as the Total Order Property.

Now, it's time for our first proof.

Theorem. *The equation $x^2 = 2$ has no solution in $\mathbb{Q}$.*

Proof. By contradiction; suppose $x \in \mathbb{Q}$ satisfies $x^2 = 2$. Since $(-x)^2 = x^2 = 2$, we may suppose x is positive. Write $x = \frac{m}{n}$ with $m, n \in \mathbb{N}$ and $(m, n) = 1$. Since $x^2 = 2$, $(\frac{m}{n})^2 = 2$, which means that:

$$m^2 = 2n^2 \,. \tag{*}$$

Therefore, m^2 is even. However, because the square of an odd number is odd, it follows that m is even. Say $m = 2k$ with $k \in \mathbb{N}$. Substituting $m = 2k$ in (*), we see:

$$2k^2 = n^{2*}$$

So n^2 is even, and by a previous argument, n is even. However, this means that 2 divides both m and n, contradicting $(m, n) = 1$. $\square$

Note: When writing proofs, make sure that you are thinking about whether what you're writing makes sense and whether you're writing too much.

2.2. Terminology.

Let $(X, \leq)$ be a poset, and let $S \neq \emptyset$ be a subset of X.

Definition. *S is **bounded above** if there is an $x \in X$ with $x \geq s$ for all $s \in S$. Such an x is called an **upper bound** for S.*

Definition. *Same as above, but with **below**, $x \leq s$, **lower bound**.*

Definition. *$x \in X$ is a **least upper bound** for S if*

(1) *x is an upper bound.*
(2) *If y is an upper bound for S, then $y \geq x$.*

The least upper bound of S is called the supremum and is denoted $\sup S$ or $\vee S$. If $|S| < \infty$, we can write $S = \{x_1, x_2, \ldots, x_n\}$. In this case, $\sup S = \max\{x_1, x_2, \ldots, x_n\}$.

Definition. *Same as above, but with **greatest lower**, lower bound, $y \leq x$. The greatest lower bound of S is known as the infimum and is denoted $\inf S$ or $\wedge S$. If $|S| < \infty$, we can write $S = \{x_1, x_2, \ldots, x_n\}$. In this case, $\inf S = \min\{x_1, x_2, \ldots, x_n\}$.*

Note: If $\sup S$ or $\inf S$ exists, then it is unique.

Examples: $\{q \in \mathbb{Q} \,|\, 0 < q < 1\}$ has a supremum, 1. $\{q \in \mathbb{Q} \,|\, q^2 < 2\}$ has an upper bound, but no supremum.

Definition. *An ordered field is called **complete** if every subset $S \subseteq F$ which is bounded above has a least upper bound.*

*Note that we removed the substitution step.

2.3. The Reals, or $\mathbb{R}$, the Complete Ordered Field.

$\mathbb{R}$ is a complete ordered field. Note that all complete ordered fields are isomorphisms of one another, so there is only *one* complete ordered field.

Proposition (ϵ-characterization of the supremum)**.** *Let $S \subseteq \mathbb{R}$, $r \in \mathbb{R}$. Then $r = \sup S$ if and only if*

(1) *r is an upper bound of S, and*
(2) *$(\forall \epsilon > 0)(\exists s \in S)$ with $s > r - \epsilon$.*

Proof. We will go in both directions.

($\Rightarrow$) (1) is trivial. For (2), let $\epsilon > 0$ be given. By contradiction, suppose no such s exists, i.e. $\forall s \in S, s \leq r - \epsilon \Rightarrow r - \epsilon$ is an upper bound. However, $r - \epsilon < r$, so r is not the *least* upper bound, and we have a contradiction.

($\Leftarrow$) By (1) we only have to show that r is the *least* upper bound. Suppose p is an upper bound for S (namely, $p \geq S$), and $p < r$. Apply (2) with $\epsilon = r - p$ to obtain an $s \in S$ such that $s > r - (r - p) = p$, which contradicts that p is an upper bound.

$\square$

Theorem. *$\mathbb{R}$ is Archimedean, i.e. $\forall a \in \mathbb{R} \exists n \in \mathbb{N}$ such that $n > a$. Equivalently, $\forall a \in \mathbb{R}$ $(a > 0)$ $\exists n \in \mathbb{N}$ such that $\frac{1}{n} < a$.*

Proof. By contradiction, suppose there is an $a \geq \mathbb{N}$. Then, $\mathbb{N}$ is bounded above. By the completeness of $\mathbb{R}$, we can let $s = \sup \mathbb{N}$. By the ϵ-characterization of the supremum with $\epsilon = 1$, we obtain that $\exists n \in \mathbb{N}$ with $n > s - 1$. However, this implies that $n + 1 > s$, which is a contradiction. $\square$

Theorem. *"$\mathbb{Q}$ is dense in $\mathbb{R}$", i.e. $\forall a, b \in \mathbb{R}$ where $a \geq 0$ and $b > a, \exists q \in \mathbb{Q}$ such that $a < q < b$.*

Proof. We know that $b - a > 0$, so $\frac{1}{b-a} > 0$. Therefore, by the Archimedean property, we can choose an m such that $m > \frac{1}{b-a}$, which is equivalent to saying $mb - ma > 1$. Consider $\{n \in \mathbb{N} \mid n > ma\}$. This set is nonempty by the Archimedean property. Let p be its least element (as a subset of $\mathbb{N}$). Then, $p < mb$, because if $p \geq mb$, then $p - 1$ would be greater than ma. Therefore, $ma < p < mb$, which implies $a < \frac{p}{m} < b$. $\square$

Note: In this proof, we have implicitly assumed that $a \geq 0$. However, this does not cause any problems for us.

3. 01/25/11: Completeness and Sequences

3.1. Completeness.

Theorem. *The equation $x^2 = 2$ has a solution in $\mathbb{R}$.*

Proof. Let $S = \{x \in \mathbb{R} \mid x^2 = 2, x > 0\}$. Then, S is nonempty, because $1 \in S$. S is also bounded above, because if $x, y \geq 0$, then

$$x^2 < y^2 \Rightarrow x^2 - y^2 < 0 \Rightarrow (x-y)(x+y) < 0 \Rightarrow x - y < 0$$

so $x < y$. Thus, if $y = 2$, then $y^2 = 4 > 2$, so 2 is an upper bound for S. Therefore, by the completeness of $\mathbb{R}$, $\sup S$ exists.

Let $x = \sup S$. We want to show $x^2 = 2$.

To show x^2 cannot be less than 2, we will assume that $x^2 < 2$ and try to show that we can add $\epsilon > 0$ to x such that $(x+\epsilon)^2 < 2$, so that x is not an upper bound for S.

Scratchwork: We have $x^2 + 2\epsilon x + \epsilon^2 < 2 \Rightarrow 2\epsilon x + \epsilon^2 < 2 - x^2 = \delta$. So, we will try to make $2\epsilon x$ and ϵ^2 both less than $\frac{\delta}{2}$. Choose $\epsilon < \frac{\delta}{4x}$ for the first term and $\epsilon < \min\{1, \frac{\delta}{2}\}$ for the second term.

So, back to the proof, choose ϵ to be less than $\min\{1, \frac{\delta}{4x}, \frac{\delta}{2}\}$, where $\delta = 2 - x^2 > 0$. Then, we obtain:

$$(x+\epsilon)^2 = x^2 + 2\epsilon x + \epsilon^2 < x^2 + \delta = x^2 + 2 - x^2 = 2\,.$$

Now, we must show that x^2 cannot be greater than 2. We will use the ϵ-characterization of the supremum, and try to find an $\epsilon > 0$ such that $(x-\epsilon)^2 > 2$, so that x is not the supremum of S.

Scratchwork: We have $x^2 - 2\epsilon x + \epsilon^2 > 2 \Rightarrow x^2 - 2 > -\epsilon^2 + 2\epsilon x$. Let $\delta = x^2 - 2 > 0$. We want to make $2\epsilon x$ less than $\frac{\delta}{2}$. We can choose $\epsilon < \frac{\delta}{2x}$.

So, back to the proof, choose ϵ to be less than $\frac{\delta}{2x}$. Then, we get:

$$(x-\epsilon)^2 = x^2 - 2\epsilon x + \epsilon^2 > x^2 - 2\epsilon x > x^2 - \delta = 2\,.$$

We know that $x - \epsilon < x$, but this implies that $x - \epsilon > s$ for all $s \in S$, which is a contradiction because x is supposed to be the least upper bound of S.

Therefore, x^2 must equal 2, and thus $x = \sup S$ is a solution to the above equation in $\mathbb{R}$. □

This leads to the fact that every positive element in $\mathbb{R}$ is a square, a fact which again uses the completeness of $\mathbb{R}$.

3.2. Infinity.

∞ is a very nice symbol, with the properties that $\infty > r$ and $-\infty < r$ for all $r \in \mathbb{R}$.

$\sup S = \infty$ means that S is not bounded above.

Recall interval notation: $[a, b) = \{x \mid a \leq x < b\}$, $(a, b] = \{x \mid a < x \leq b\}$, and so on. Likewise, we define:

$$[a, \infty) = \{x \mid a \leq x\}$$
$$(-\infty, y) = \{x \mid x < y\}$$

and so on.

3.3. **Sequences.**

Definition. *A* ***sequence*** *beginning at* m *is a function from* $\{n \in \mathbb{Z} \,|\, n \geq m\}$ *to* $\mathbb{R}$*, or some other set, and is denoted* (s_n)*.*

Note: We need for a sequence to be well-defined, such as $(s_n) = n^2$ or $(s_n) = (-1)^n$, not just something like $(0, \frac{1}{2}, \frac{1}{4}, 0, 2, \dots)$.

Definition. $\lim_{x\to\infty}(s_n) = s$*, or* $s_n \to s$*, is equivalent to:*

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) \textit{ s.t. } (n > N) \Rightarrow |s_n - s| < \epsilon \,.$$

Note: There are two irrelevancies in this definition: $n \geq N$ and $|s_n - s| \leq \epsilon$ lead to the same definition. Also, think of $|a - b|$ as the *distance* between a and b. If you add the condition that $\epsilon \in \mathbb{Q}$, it is still the same definition by the density of $\mathbb{Q}$.

Example: $\lim_{x\to\infty} \frac{1+n}{1+3n} = \frac{1}{3}$.
Scratchwork: Notice that for any n, it is the case that:

$$s_n - s = \frac{1+n}{1+3n} - \frac{1}{3} = \frac{3+3n-1-3n}{(1+3n)3} = \frac{2}{3+9n}$$

and this is the same regardless of the absolute value sign. Therefore, we want $\frac{2}{3+9n} < \epsilon$.

$$\frac{2}{3+9n} < \epsilon \Rightarrow \frac{2}{\epsilon} < 3 + 9n \Rightarrow \frac{2}{\epsilon} - 3 < 9n \Rightarrow n > \frac{1}{9}(\frac{2}{\epsilon} - 3)$$

Proof. Let $\epsilon > 0$ be given. Choose any $N \in \mathbb{N}$ such that $N > \frac{1}{9}(\frac{2}{\epsilon} - 3)$. Then, for $n > N$:

$$|s_n - s| = \left|\frac{1+n}{1+3n} - \frac{1}{3}\right| = \frac{2}{3+9n} < \frac{2}{3+9N} < \epsilon$$

$\square$

4. 01/27/11: Convergence, Monotone Sequences

4.1. **Some Basic Notes.**

If $A \subseteq B \subseteq \mathbb{R}$ and B is bounded above, then $\sup A \leq \sup B$.

Proposition. *If (s_n) is a sequence where $\lim s_n = s$ and $\lim s_n = t$, then $s = t$.*

Proof. Suppose $s \neq t$ and put $\epsilon = \frac{|s-t|}{50}$ into the definition of the convergence of a sequence. Then, choose N such that $n \geq N \Rightarrow |s_n - s| < \frac{|s-t|}{50}$. Also, since $\lim s_n = t$, choose M such that $n \geq M \Rightarrow |s_n - t| < \frac{|s-t|}{50}$. Choose k such that $k \geq M$ and $k \geq N$.

Then, $|s-t| \leq |s-s_k| + |s_k - t|$ by the triangle inequality. However, then $|s-s_k| < \frac{|s-t|}{50}$ and $|s_k - t| < \frac{|s-t|}{50}$, which implies $|s-t| < 2\left(\frac{|s-t|}{50}\right)$, a contradiction. □

As an aside: The Triangle Inequality states that for $a, b, c \in \mathbb{R}$, $|a-b| \leq |a-c| + |b-c|$.

This follows from the slightly easier assertion that $|a+b| \leq |a| + |b|$, in which you can define $a = a - c$ and $b = b - c$.

4.2. **Bounded Sequences.**

Definition. *(s_n) is **bounded above** if $\{s_n \mid n \in \mathbb{N}\}$ is bounded above, and similarly for **bounded below**. (s_n) is **bounded** if $\{s_n \mid n \in \mathbb{N}\}$ is bounded above* and *below.*

Proposition. *A convergent sequence is bounded.*

Proof. Suppose $s_n \to s$, or $\lim_{x\to\infty} s_n = s$. Then by definition there exists an N such that $s - 1 < s_n < s + 1$ for all $n \geq N$.

Let $M = \max\{s_1, s_2, \ldots, s_N, (s+1)\}$. Clearly $s_n \leq M$ for all n. Then, let $m = \min\{s_1, s_2, \ldots, s_N, (s-1)\}$. Clearly $s_n \geq m$ for all n. □

Theorem. *Given $s_n \to s$ and $t_n \to t$, we have:*

(1) $s_n + t_n \to s + t$.
(2) $s_n t_n \to st$.
(3) *If $t_n \neq 0$ for all n and $t \neq 0$, then $\lim \frac{s_n}{t_n} = \frac{s}{t}$.*

Proof. We will only prove (2), as the others are similar.

Scratchwork: We are working with $|s_n t_n - st|$, so we will write it as

$$|(s_n t_n - s_n t) + (s_n t - st)| \leq |s_n t_n - s_n t| + |s_n t - st| = |s_n||t_n - t| + |s_n - s||t|$$

However, we know that all four expressions are bounded, so we are ready, and the formal proof follows.

By the previous result, choose $B > 0$ such that $|s_n| < B$ for all n. Now since $t_n \to t$, choose N_1 such that $(n \geq N_1) \Rightarrow |t_n - t| < \frac{\epsilon}{2B}$ where $\epsilon > 0$ is given. Also, since $s_n \to s$, choose N_2 such that $(n \geq N_2) \Rightarrow |t||s_n - s| < \frac{\epsilon}{2}$. Now, choose $N = \max\{N_1, N_2\}$. Then, if $n \geq N$:

$$|s_n t_n - st| \leq |s_n t_n - s_n t| + |s_n t - st| = |s_n||t_n - t| + |t||s_n - s|.$$

Therefore, since $|s_n||t_n - t| \leq B \cdot \frac{\epsilon}{2B} = \frac{\epsilon}{2}$ because $n \geq N_1$ and $|t||s_n - s| \leq \frac{\epsilon}{2}$ because $n \geq N_2$, their sum is less than or equal to $\frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. □

Note: This theorem is very convenient! For instance:

$$\frac{1+n}{1+3n} = \frac{1+\frac{1}{n}}{3+\frac{1}{n}} \to \frac{1}{3}\,.$$

So what about when a sequence doesn't converge? Well, first of all, an unbounded sequence obviously doesn't converge. But what about $s_n = (-1)^n$?

If this sequence converges, then for all $\epsilon > 0$, then for all sufficiently large n, $|s_n - a| < \epsilon$. However, if n is odd, then $|1+a| < \epsilon$ and if n is even, $|1-a| < \epsilon$. This means:

$$2 = |1-(-1)| \leq |a-1| + |a+1| < 2\epsilon$$

Set $\epsilon = \frac{1}{2}$, and we have a contradiction.

4.3. **Monotonic Sequences.**

Definition. *A sequence (s_n) is* ***nondecreasing*** *if $s_{n+1} \geq s_n$ for all n. (s_n) is* ***nonincreasing*** *if $a_n \geq a_{n+1}$ for all n. (s_n) is* ***monotonic*** *if it is either nondecreasing or nonincreasing.*

Theorem. *In an ordered field F, F is complete if and only if every bounded monotone sequence converges.*

Proof. In next lecture. □

5. 02/01/11: Monotone, limsup, Cauchy

5.1. **Monotone Sequences, continued.**

Theorem. *In an ordered field F, F is complete if and only if every bounded monotone sequence converges.*

Proof. We will go in both directions.

($\Rightarrow$) Let (s_n) be monotonic and bounded. Without loss of generality, suppose (s_n) is nondecreasing. Then, $\{s_n \mid n \in \mathbb{N}\}$ is bounded above and nonempty. By completeness, there exists an $s = \sup\{s_N \mid n \in \mathbb{N}\}$.

We will claim that $s_n \to s$. Let $\epsilon > 0$ be given. We want to find an N such that if $n \geq N$, then $|s_n - s| < \epsilon$. By the ϵ-characterization of the supremum, there exists $N \in \mathbb{N}$ such that $s_N > s - \epsilon$. Then, for all $n \geq N$, we have:

$$s_N \leq s_n \leq s \Rightarrow s - \epsilon < s_N \leq s_n$$

which implies $|s_n - s| < \epsilon$.

($\Leftarrow$) We will again assume, without loss of generality, that (s_n) is nondecreasing. Let $S \neq \emptyset$ be a subset of F that is bounded above by B. We must show the existence of $\sup S$. We will construct inductively two sequences, (x_n) and (y_n), with the following properties:

(1) $x_n \in S$ and y_n is an upper bound for S.
(2) $x_{n+1} \geq x_n$ and $y_{n+1} \leq y_n$.
(3) $|x_{n+1} - y_{n+1}| \leq \frac{1}{2}|x_n - y_n|$.

To do this, let $x_0 \in S$ and y_0 be any upper bound for S. Suppose $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ are already constructed, and satisfy the three properties. Then, either $\frac{x_n+y_n}{2}$ is an upper bound for S or it is not. If it is, then set $x_{n+1} = x_n$ and $y_{n+1} = \frac{x_n+y_n}{2}$. If it is not, then there is an $s \in S$ such that $s \geq \frac{x_n+y_n}{2}$. Set $x_{n+1} = s$ and $y_{n+1} = y_n$. Note then that every x_n stays in S and every y_n is an upper bound for S, so (1) is satisfied. By construction, (2) is true, and so is (3). Therefore, the sequence we have constructed satisfies all of the given properties.

Note that (x_n) is a nondecreasing sequence bounded above by y_0 and that (y_n) is a nonincreasing sequence bounded below by x_0. Therefore, by hypothesis, there are $x, y \in F$ such that $x_n \to x$ and $y_n \to y$. We will claim that $x = y$. It is obvious from (3) that $|x_n - y_n| \leq \frac{1}{2^n}|x_0 - y_0|$. Suppose for the moment we can show that $\frac{1}{2^n} \to 0$. Then, by choosing sufficiently large n, it will be the case that for any given $\epsilon > 0$, $|x - x_n| < \epsilon$, $|y_n - y| < \epsilon$, and $|x_n - y_n| < \epsilon^*$. Then, we can obtain that:

$$|x - y| = |x - x_n + x_n - y_n + y_n - y| \leq |x - x_n| + |x_n - y_n| + |y_n - y|.$$

Therefore, $|x - y| < 3\epsilon$, so $x = y$.

However, we still need to show that $\frac{1}{2^n} \to 0$. We know $\frac{1}{2^n}$ is a nonincreasing bounded sequence, so it does have a limit $z \in F$. Also, by the elementary properties of limits which are valid in any ordered field, the sequence $(2\frac{1}{2^n}) \to 2z$. However, the limit of $(2\frac{1}{2^n})$ is the same as that of $\frac{1}{2^n}$. Therefore, $2z = z$, which means that $z = 0$. Therefore, we now know that $x = y$.

*This condition is what we need the convergence of $\frac{1}{2^n}$ for.

Finally, we need to show that $x = y = \sup S$. y is an upper bound for S because $y_n \geq s$ for all $s \in S$ and for all $n \in \mathbb{N}$, so $y \geq s$ for all $s \in S$ as well. Also, if there existed an upper bound $t < y$, then there would be an x_n such that $x_n > t$, because $(x_n) \to x$. Therefore, t cannot be an upper bound, and thus y is the least upper bound.

□

Note: The way we proved that $x = y$ is equivalent to proving $\lim(x_n - y_n) = 0$, which, for convergent sequences, implies that $\lim x_n = \lim y_n$.

Definition.

$$\lim_{n\to\infty} s_n = +\infty \iff \forall M \in \mathbb{R}, (\exists N \in \mathbb{N})\ s.t.\ (n \geq N) \Rightarrow (s_n \geq M).$$
$$\lim_{n\to\infty} s_n = -\infty \iff \forall M \in \mathbb{R}, (\exists N \in \mathbb{N})\ s.t.\ (n \geq N) \Rightarrow (s_n \leq M).$$

Note: ∞ is *not* a number.

Proposition. *Any unbounded monotone sequence tends to ∞ or $-\infty$.*

Proof. The proof is supplied in the textbook (10.4). □

5.2. $\limsup$**, or** $\overline{\lim}$**, and** $\liminf$**, or** $\underline{\lim}$**.**

Suppose (s_n) is a sequence in $\mathbb{R}$. Let $T_n = \{s_k \,|\, k \geq n\}$*. It is obvious that $T_n \supseteq T_{n+1} \supseteq T_{n+2} \supseteq \ldots$, and if $A \subseteq B$ for any sets A and B, then it is the case that $\sup A \leq \sup B$ and $\inf A \geq \inf B$.

Therefore, the sequence $(\sup T_n)$ is nonincreasing and the sequence $(\inf T_n)$ is nondecreasing. By the completeness of $\mathbb{R}$ as well as the "suitable consideration" of $\pm\infty$ in the unbounded cases, $\lim_{n\to\infty} \sup T_n$ and $\lim_{n\to\infty} \inf T_n$ always exist!

Examples: $\limsup(-1)^n = 1$ and $\liminf(-1)^n = -1$.
Let s_n be defined such that $s_n = 1 - \frac{1}{n}$ if n is even and $s_n = -n$ if n is odd. Then, $\limsup(s_n) = 1$ and $\liminf(s_n) = -\infty$.
Let s_n be defined such that $s_n = 1 + \frac{1}{n}$ if n is even and $s_n = -n$ if n is odd. Then, $\limsup(s_n)$ is still 1!

Remember: $\limsup, \liminf$ *always exist.*

Theorem. *Let (s_n) be a sequence. Then:*

(1) $s_n \to s \Rightarrow \liminf s_n = \limsup s_n = s$
(2) *If* $\liminf s_n = \limsup s_n$, *then* $\lim s_n = \liminf s_n = \limsup s_n$.

Proof. We will prove each part.

(1) We know that $(\forall\epsilon > 0)(\exists N)$ s.t. $\{T_N\} \subseteq (s-\epsilon, s+\epsilon)$. Therefore, $\sup T_N \subseteq [s-\epsilon, s+\epsilon]$, and $\sup T_n \subseteq [s-\epsilon, s+\epsilon]$ for $n > N$. Because this is true for all ϵ, $\limsup s_n = s$. The proof of $\liminf$ is similar.

*The T stands for "tail."

(2) Suppose $\limsup s_n = r = \liminf s_n$ for some $r \in \mathbb{R}$. Let $\epsilon > 0$ be given. We want to show that $s_n \to r$. We know that:

$$\exists N_1 \text{ s.t. } \sup T_n < r + \epsilon \ \forall n \geq N_1 \text{ and}$$
$$\exists N_2 \text{ s.t. } \inf T_n > r - \epsilon \ \forall n \geq N_2 \,.$$

Let $N = \max\{N_1, N_2\}$. Then, if $n \geq N$, $s_n \in T_n$, so $s_n \leq \sup T_n < r + \epsilon$ and $s_n \geq \inf T_n > r - \epsilon$, which implies $|s_n - r| < \epsilon$.

□

5.3. **Cauchy Sequences.**

Definition. *Let (s_n) be a sequence of real numbers. We say (s_n) is **Cauchy** if:*

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) \ s.t. \ (m, n \geq N) \Rightarrow (|s_n - s_m| < \epsilon) \,.$$

Note: This makes sense in spaces with an arbitrary notion of distance.

Proposition. *A convergent sequence is Cauchy.*

Proof. Assume $s_n \to s$ for a sequence (s_n). Let $\epsilon > 0$ be given and choose N so that $(n \geq N) \Rightarrow |s_n - s| < \frac{\epsilon}{2}$. Then, if $m, n \geq N$, we obtain:

$$|s_m - s_n| = |s_m - s + s - s_n| \leq |s_m - s| + |s - s_m| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \,.$$

□

6. 02/03/11: Cauchy Sequences, Subsequences

6.1. **A Simple Proof.**

Say we want to show that for $0 < r < 1$, $r^n \to 0$.

One way to show this is to assume $r^n \to x$. Then, we know that $kr^n \to kx$. Let $k = r$. Then, $rr^n \to rx$, but r^{n+1} has the same limit as r^n. Therefore, $rx = x$ and $x = 0$.

Another way is to show that $(\frac{1}{r})^n \to \infty$, as that implies $r^n \to 0$. Write $\frac{1}{r}$ as $1 + b > 0$. Then $(1 + b)^n \geq 1 + nb$, which tends to ∞.

6.2. **Cauchy Sequences, again.**

Last time, we learned about lim sup, lim inf, and Cauchy sequences. Recall that (s_n) is Cauchy if and only if:

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) \text{ s.t. } (n, m \geq N) \Rightarrow |s_n - s_m| < \epsilon\,.$$

Note: This definition makes sense in any ordered field.

Proposition. *A Cauchy sequence is bounded.*

Proof. Choose $\epsilon = 17$ (arbitrarily, of course). Then, we know that

$$\exists N \in \mathbb{N} \text{ s.t. } \forall n, m \geq N, |s_n - s_m| < 17\,.$$

This means that for $n \geq N$, $|s_n - s_N| < 17$, or $s_N - 17 < s_n < s_N + 17$. So for all n, it is the case that

$$s_n \leq \max\{s_1, s_2, \ldots, s_{N-1}, s_N + 17\}$$

and likewise,

$$s_n \geq \min\{s_1, s_2, \ldots, s_{N-1}, s_N - 17\}\,.$$

□

Theorem. *If (s_n) is a sequence in $\mathbb{R}$, then (s_n) is Cauchy if and only if (s_n) converges.*

Proof. We will proceed in both directions.

($\Leftarrow$) This was proved in the previous lecture.

($\Rightarrow$) Let (s_n) be a Cauchy sequence in $\mathbb{R}$. Note that because (s_n) is bounded, $\limsup s_n$ and $\liminf s_n$ exist and are real numbers. Therefore, it suffices to show that $\liminf s_n = \limsup s_n$. Because $\liminf \leq \limsup$ in all cases, we only need to show that $\liminf s_n \geq \limsup s_n$. To do this, we will show that $\liminf s_n + \epsilon \geq \limsup s_n$ for any $\epsilon > 0$.

Let $\limsup s_n = x$ and $\liminf s_n = y$. Let ϵ be given, and choose an $N \in \mathbb{N}$ such that $|s_n - s_m| < \epsilon$ for $n, m \geq N$. Note that this implies $s_n - s_m < \epsilon$. Then, because $\limsup s_n = x$, there exists an $n \geq N$ such that $s_n > x - \epsilon$, and similarly there exists an $m \geq N$ such that $s_m < y + \epsilon$. Therefore:

$$x - y = x - s_n + s_n - s_m + s_m - y < 3\epsilon\,.$$

□

Proposition. *Let (s_n) be a sequence in an (Archimedean) ordered field. Then, if there exist k, r (where $0 < r < 1$) such that $|s_n - s_{n+1}| < r^n k$ for all n, then (s_n) is Cauchy.*

Proof. Let $\epsilon > 0$ be given. Choose N such that $n \geq N \Rightarrow r^n < \frac{1-r}{k}\epsilon$. This is possible because $r^n \to 0$, which we may need Archimedean for. Without loss of generality, we will assume that $m > n$. Then, for $n, m \geq N$, we have:

$$|s_n - s_m| = |s_n - s_{n+1} + s_{n+1} - s_{n+2} + \cdots + s_{m-1} - s_m| \leq$$
$$|s_n - s_{n+1}| + |s_{n+1} - s_{n+2}| + \cdots + |s_{m-1} + s_m| < kr^n(1 + r + \cdots + r^{m-n}) =$$
$$kr^n\left(\frac{1 - r^{m-n+1}}{1-r}\right) < \frac{kr^n}{1-r} < \epsilon$$

□

Theorem. *If F is an ordered field, the following are equivalent:*

(1) *F is complete, i.e. every nonempty subset bounded above has a supremum.*
(2) *Every nondecreasing bounded sequence converges.*
(3) *Every Cauchy sequence converges (We may need F is Archimedean).*

Proof. By our previous results, we only need to prove (3) $\Rightarrow$ (1). To do this, we will refer back to our proof of (2) $\Rightarrow$ (1).

We will show that $(x_n), (y_n)$ from the aforementioned proof are Cauchy. This would show that by assuming Cauchy sequences converge, we can get to the completeness of F.

We know that $|x_n - x_{n+1}| < \frac{1}{2^n}|x_0 - y_0|$ because $x_n \leq x_{n+1} \leq y_n$, and similarly for y_{n+1}. Therefore, (x_n) and (y_n) are Cauchy, and $\lim x_n = x$, $\lim y_n = y$ for some $x, y \in \mathbb{R}$. The rest of the proof follows in exactly the same way as the proof for nondecreasing bounded sequences. □

6.3. Subsequences.

Definition. *For $(s_n), n \in \mathbb{N}$ given, a* ***subsequence*** *of s_n is a sequence of the form $(t_k \,|\, t_k = s_{n_k})$ with $n_k < n_{k+1} < \ldots$ for all k.*

Example: Take the sequence (n). $(2n)$ is a subsequence of (n), but (1) is not. However, if our sequence is $\{1, 2, 3, 1, 2, 3, \ldots\}$, then (1) is a subsequence.

Proposition. *If $s_n \to s$, then any subsequence of s_n tends to s.*

Proof. Let (s_{n_k}) be a subsequence. Let $\epsilon > 0$ and $\lim_{n\to\infty} s_n = s$ be given. Choose N such that $n \geq N \Rightarrow |s_n - s| < \epsilon$. Then, if $k \geq N$, then $|s_{n_k} - s| < \epsilon$ because $n_k \geq N$. □

Theorem (Bolzano-Weierstrass)**.** *In $\mathbb{R}$, any bounded sequence has a convergent subsequence.*

Proof Sketch. Take the interval in which the sequence lies, and divide it in half. Then, choose a half with infinitely many terms (there must be at least one) and pick a term in that interval. Then, divide that half in half and choose a half with infinitely many terms again. Pick another term from that half, such that that term appears later in the sequence than the first. Repeat this sequence indefinitely. This will lead to a limit for the subsequence that you will have picked. □

7. 02/08/11: Defining the Reals, Part One

7.1. **Equivalence Relations.**

Definition. *Given a set S, an equivalence relation on S is a relation $x \sim y$ with 3 properties:*

(1) $x \sim x \ \forall x \in S$ *(Reflexivity)*
(2) $x \sim y \Rightarrow y \sim x$ *(Symmetry)*
(3) $x \sim y$ *and* $y \sim z \Rightarrow x \sim z$ *(Transitivity)*

Given $x \in S$, define $[x] = \{y \,|\, y \sim x\}$.

Here are some facts about $[x]$:

(1) $x \sim y \iff [x] = [y]$
(2) $[x] \cap [y] = \emptyset$ if $x \nsim y$ and $[x] \cap [y] = [x]$ if $x \sim y$.

Therefore, S is the disjoint union of equivalence classes.

However, if $S = \coprod_{i \in I} S_i$, then define $\sim$ by:

$$x \sim y \iff \exists i \in I \text{ s.t. } \{x, y\} \subseteq S_i$$

It is easy to check that this relation is an equivalence relation, meaning that a partition on a set gives you an equivalence relation, not just the other way around.

7.2. **Defining the Reals.**

The Equivalence Relation. First, we will define a relation $\sim$ on sequences (s_n) of rationals, such that $(s_n) \sim (t_n)$ if and only if:

$$\lim_{n\to\infty} s_n - t_n = 0, \text{ i.e. } (\forall \epsilon > 0, \epsilon \in \mathbb{Q})\ \exists N \text{ s.t. } n \geq N \Rightarrow |s_n - t_n| < \epsilon\,.$$

This is an equivalence relation. Note that in an Archimedean field, $(\forall \epsilon > 0)$ can be replaced by $(\forall \epsilon > 0 \,|\, \epsilon \in \mathbb{Q})$, by the denseness of the rationals.

We will show that if (s_n) is Cauchy and $(s_n) \sim (t_n)$, then (t_n) is Cauchy. We know that $|s_n - s_m| < \epsilon$ and want to show that this implies $|t_n - t_m| < \epsilon$.

$$|t_n - t_m| = |t_n - s_n + s_n - s_m + s_m - t_m| \leq |t_n - s_n| + |s_n - s_m| + |s_m - t_m| < 3\epsilon$$

Therefore, (t_n) is Cauchy.

So, we will define F to be the set of equivalence classes of Cauchy sequences. We will give it the structure of an ordered field and show it is complete.

Checking the Field Properties. First, we will define $+$ such that $[(s_n)] + [(t_n)] = [(s_n + t_n)]$. We must check to make sure that this is well-defined. Suppose $(s_n) \sim (s'_n)$ and $(t_n) \sim (t'_n)$. Then, we know that $\lim s_n - s'_n = 0$ and $\lim t_n - t'_n = 0$. Therefore:

$$\lim(s_n - s'_n + t_n - t'_n) = \lim((s_n + t_n) - (s'_n + t'_n)) = 0 + 0 = 0$$

and $+$ is well-defined. Let $[(0_n)]$ be the additive identity and note $-[(s_n)] = [(-s_n)]$, so inverses exist. $+$ is also commutative, so this is an abelian group.

Next, we will define $\cdot$ such that $[(s_n)] \cdot [(t_n)] = [(s_n t_n)]$. Again, we want this to be well-defined. First, we will check that $(s_n t_n)$ is also Cauchy:

$$|s_n t_n - s_m t_m| = |s_n t_n - s_n t_m + s_n t_m - s_m t_m| \leq |s_n||t_n - t_m| + |t_m||s_n - s_m|\,.$$

Obviously, the rightmost side goes to zero, so $(s_n t_n)$ is Cauchy. Now, suppose $\lim s_n - s'_n = 0$ and $\lim t_n - t'_n = 0$. We want to show that $\lim(s_n t_n - s'_n t'_n) = 0$.

$$|s_n t_n - s'_n t'_n| = |s_n t_n - s'_n t_n + s'_n t_n - s'_n t'_n| \leq |t_n||s_n - s'_n| + |s'_n||t_n - t'_n|$$

Again, obviously, the rightmost side goes to zero, so multiplication is also well-defined. It is obviously commutative, associative, distributive over addition, and has an identity, namely $[(1_n)]$. However, the existence of inverses is not trivial. For this, we will make an important observation.

Proposition. *Let (s_n) be a Cauchy sequence of rationals. Then, either:*

(1) $s_n \to 0$,
(2) $\exists N$ *s.t.* $s_n > \epsilon\ \forall n \geq N$, *or*
(3) $\exists N$ *s.t.* $s_n < -\epsilon\ \forall n \geq N$.

where $\epsilon \in \mathbb{Q}$ *and* $\epsilon > 0$.

Proof. Suppose s_n does not converge to 0. Then, the negation of the definition of convergence gives us that:

$$\exists \epsilon > 0 \text{ s.t. } \forall N \in \mathbb{N}, \exists r \geq N \text{ s.t. } |s_r| \geq \epsilon\,.$$

By the Cauchiness of (s_n), there exists an N s.t. $|s_n - s_m| < \frac{\epsilon}{2}\ \forall n, m \geq N$, where ϵ is the same as the ϵ in the above expression. Choose the r guaranteed for this ϵ such that $r \geq N$ and $|s_r| \geq \epsilon$.

Suppose $s_r \geq \epsilon$. Then, for $n \geq N$, we have $|s_r - s_n| < \frac{\epsilon}{2}$, so $s_n > \frac{\epsilon}{2}$, which is case (2).

Now, suppose $s_r \leq -\frac{\epsilon}{2}$. Then, for $n \geq N$, we have $|s_r - s_n| < \frac{\epsilon}{2}$, so $s_n < -\frac{\epsilon}{2}$, which is case (3). □

Note that it is obvious to see that (1), (2), and (3) depend only on the class of (s_n).

If $[(s_n)] \neq [(0_n)]$, then s_n does not converge to 0, and we are in case (2) or (3). Then, we can change finitely many terms, by the above proposition, in order to have $s_n \neq 0$ for all n. Then, $[(s_n)] \cdot [(\frac{1}{s_n})] = [(1_n)]$, but we must also make sure $(\frac{1}{s_n})$ is Cauchy. We obtain:

$$\left|\frac{1}{s_n} - \frac{1}{s_m}\right| = \left|\frac{s_n - s_m}{s_n s_m}\right| < \frac{\epsilon}{B^2}$$

where $|s_n| < B$, because Cauchy sequences are bounded.

Therefore, finally, we have established that this is an abelian group under multiplication, and thus that F is a field which contains the rationals as constant sequences, i.e. $q \in \mathbb{Q}$ is defined as (q_n).

Ordering the Field. We will define $\mathcal{P}$ as such:

$$\mathcal{P} = \{[(s_n)]\,|\,\exists N, \epsilon > 0 \text{ with } s_n > \epsilon\ \forall n \geq N \text{ for some, hence any } (s_n) \in [(s_n)]\}\,.$$

Essentially, $\mathcal{P}$ is the set of equivalence classes of Cauchy sequences that fall into category (2) of the above proposition. This trivially satisfies all of the conditions necessary for an order.

Observe immediately that F is Archimedean. For some $[(s_n)]$, $s_n < k$ for some large integer k, because Cauchy sequences are bounded. Thus, $2k - s_n > 0$ for all n, and $(2k_n)$ therefore bounds (s_n) for all n. A similar argument can be used to establish the nonexistence of infinitesimals.

8. 02/10/11: Defining the Reals, Part Two

8.1. Recap.

Recall that the intuition of this construction is that we are approximating reals by sequences of truncations of rational approximations of each real number; this is where the notion of equivalence classes comes from.

So far, we have that the additive group with identity $[(0_n)]$ and multiplicative group with identity $[(1_n)]$ exist and are abelian. Therefore, the previously defined F is a field. Also, we have an order with $\mathcal{P} = \{[(s_n)] \mid \text{case two of the proposition from last time}\}$, so F is an ordered field. Furthermore, F is Archimedean because given a Cauchy, and thus bounded, $[(s_n)]$, there exists a $q \geq s_n + \epsilon$ for all n where $q \in \mathbb{N}$, so in the field, "q" $> [(s_n)]$. Thus, we will use the fact that in an Archimedean field, if all Cauchy sequences converge then it is complete in order to prove the completeness of F.

8.2. Defining a Concept of "Less Than".

Note that $|[(s_n)]| = [(|s_n|)]$. We could define a notion of "less than" as such:

$|[(s_n)]| < \epsilon \Rightarrow \epsilon - [(|s_n|)] > 0 \Rightarrow \exists \eta > 0$ s.t. for large n, $\epsilon - |s_n| > \eta$, or $|s_n| < \epsilon - \eta$

In this definition, ϵ, η, s_n are all rationals. This is a perfectly serviceable definition, but in order to get away from η, we will use an alternate definition.

Definition. *"$[(s_n)]$ is* ***less than*** *ϵ" is defined through these two implications:*

1. *$[(s_n)] < \epsilon \Rightarrow \exists N \in \mathbb{N}$ s.t. $s_n < \epsilon \ \forall n \geq N$, and*
2. *$\exists N \in \mathbb{N}$ s.t. $s_n < \epsilon \ \forall n \geq N \Rightarrow [(s_n)] \leq \epsilon$.*

Now, we are finally ready.

8.3. Convergence of Cauchy Sequences in F.

Theorem. *Every Cauchy sequence in F converges.*

Proof. What we have to do is given $([(s_n)]_m)$ Cauchy, come up with a $[(t_n)]$ where $\lim_{m\to\infty}[(s_n)]_m = [(t_n)]$. Our notation will be:

$$\begin{aligned}
[(s_n)]_1 &= s_{1,1}, s_{2,1}, s_{3,1}, s_{4,1}, \ldots \\
[(s_n)]_2 &= s_{1,2}, s_{2,2}, s_{3,2}, s_{4,2}, \ldots \\
[(s_n)]_3 &= s_{1,3}, s_{2,3}, s_{3,3}, s_{4,3}, \ldots \\
&\vdots \\
[(s_n)]_k &= s_{1,k}, s_{2,k}, s_{3,k}, s_{4,k}, \ldots, s_{k,k}, \ldots
\end{aligned}$$

We are going to shoot for the diagonal, $(s_{n,n})$. However, we are dealing with equivalence classes of Cauchy sequences, so we could have made a bad choice of representatives and $(s_{n,n})$ may go to infinity. Therefore, we will change $[(s_n)]_k$ so that it only varies by $\frac{1}{k}$ in each term. We will do this by changing all elements before N (Note, a finite number) so they all live within ϵ of the next. or, more formally...

Proposition. *Let (s_n) be a Cauchy sequence. For all $\epsilon > 0$, there exists a (t_n) such that $(s_n) \sim (t_n)$ and $|t_n - t_m| < \epsilon$ for all n, m.*

Proof. $\exists N$ s.t. $|s_n - s_m| < \epsilon \; \forall n, m \geq N$. Then, fix:

$$t_1 = s_N, t_2 = s_N, \ldots, t_{n-1} = s_N, \text{ and } t_n = s_n \; \forall n \geq N\,.$$

It is obvious that $[(t_n)] = [(s_n)]$. □

So, because $[(s_n)]_m$ is Cauchy, we can suppose:

$$(**) \qquad |s_{n,m} - s_{k,m}| < \frac{1}{m} \; \forall k, m$$

Now, set the aforementioned $t_n = s_{n,n}$. Now, all that remains for us to do is to show that $[(s_n)]_m \to [(t_n)]$ as $m \to \infty$.

First, we will deal with the fact that $[(t_n)]$ must be Cauchy. We want to compare $|t_n - t_m|$, or $|s_{n,n} - s_{m,m}|$. Let $\epsilon > 0$ be given, where in this case $\epsilon \in \mathbb{Q}$. Choose an $N > \frac{1}{\epsilon}$ s.t. $|[(s_n)]_m - [(s_n)]_k| < \epsilon \; \forall m, k \geq N$. Also, choose a p s.t. $|s_{p,m} - s_{p,k}| < \epsilon$. Now, suppose $m, k \geq N$.

$$|t_m - t_k| = |s_{m,m} - s_{k,k}| \leq |s_{m,m} - s_{p,m}| + |s_{p,m} - s_{p,k}| + |s_{p,k} - s_{k,k}| < 3\epsilon$$

where the first term is bounded by (**), the second term is bounded by the choice of N and the Cauchiness of $([(s_n)]_m)$, and the third term is also bounded by (**). Therefore, we are done, and $[(t_n)] \in F$.

Now, we only need to show that $\lim_{m\to\infty}[(s_n)]_m = [(t_n)]$, or in other words that $\forall \epsilon > 0 \; \exists N$ s.t. $m \geq N \Rightarrow |[(s_n)]_m - [(t_n)]| \leq \epsilon$. Note that it is less than or equal to ϵ because of the second implication in the definition of "less than" above. Let $\epsilon > 0$ be given. By the Cauchiness of (t_n), we have $\exists N > \frac{2}{\epsilon}$ s.t. $\forall m, n \geq N, |t_n - t_m| < \frac{\epsilon}{2}$. Then, we obtain:

$$|t_n - s_{n,m}| = |s_{n,n} - s_{n,m}| = |s_{n,n} - s_{m,m} + s_{m,m} - s_{n,m}| \leq$$
$$|t_n - t_m| + |s_{m,m} - s_{n,m}| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Where the first term is bounded by the Cauchiness of (t_n) and the second is bounded by (**). Therefore, $|(t_n) - [(s_n)]_m| \leq \epsilon$.

Thus, since every Cauchy sequence in F converges and F is Archimedean, F is complete! □

8.4. **Metric Spaces.**

Definition. *A **metric space** is a set X together with a function $d\colon X \times X \to \mathbb{R}$, known as a **metric**, where $d(x,y)$ satisfies:*

1. *$d(x,y) \geq 0$ and $d(x,y) = 0 \iff x = y$,*
2. *$d(x,y) = d(y,x)$, and*
3. *$d(x,y) \leq d(x,z) + d(z,y) \; \forall x, y, z \in X$.*

Basically, all we are given when talking about metric spaces is a notion of *distance*.

Example: In $\mathbb{R}$, a possible metric is $d(x,y) = |x - y|$.

Example: Take any X, and define d such that:

$$d(x,y) = \begin{cases} 0, & \text{if } x = y; \\ 1, & \text{if } x \neq y. \end{cases}$$

This is known as the discrete metric.

9. 02/15/11: Metric Spaces, Part One

9.1. Metric Spaces on $\mathbb{R}^n$.

Definition. *A **metric space** is a set X and a function $d\colon X \times X \to \mathbb{R}$ such that:*

(1) $d(x,y) \geq 0$ *and* $d(x,y) = 0 \iff x = y$,
(2) $d(x,y) = d(y,x)$, *and*
(3) $d(x,y) \leq d(x,z) + d(z,y)$

Remember that some examples are $|x-y|$ on $\mathbb{R}$ and the discrete metric from last lecture.

Now consider $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R}\}$. We will give this set several different metrics.

9.1.1. *The Euclidian Metric.* This metric is defined by:

$$d((x_1,\ldots,x_n),(y_1,\ldots,y_n)) = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_n-y_n)^2}$$

For this metric, (1) and (2) are obvious. We will come back to (3).

We will now explore the concept of norms, with the intention of obtaining some new metrics for $\mathbb{R}^n$ during this exploration.

Definition. *A **Norm** on a (real) vector space V is a function $||\cdot|| : V \to \mathbb{R}$ such that:*

(i) $||v|| \geq 0$ *and* $||v|| = 0 \iff v = 0$,
(ii) $||\lambda v|| = |\lambda|\,||v||$ *for* $\lambda \in \mathbb{R}$, *and*
(iii) $||v+w|| \leq ||v|| + ||w||$.

9.1.2. *The "Maximum" Norm.* We will define this norm such that for $V = \mathbb{R}^n$:

$$||(x_1,\ldots,x_n)||_\infty = \max\{|x_1|,|x_2|,\ldots,|x_n|\}$$

(i) and (ii) are obvious. For (iii), use the triangle inequality on $\mathbb{R}$.

9.1.3. *The "Sum" Norm.* We will define this norm such that for $V = \mathbb{R}^n$:

$$||(x_1,\ldots,x_n)||_1 = \sum_{i=1}^{n} |x_i|$$

(i), (ii), and (iii) are all trivial for this norm.

Norms can also come from inner products (sometimes known as scalar products), which we will further investigate.

Definition. *An **inner product** is a function $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ such that:*

(a) $\langle v,v\rangle \geq 0$ *and* $\langle v,v\rangle = 0 \iff v = 0$,
(b) $\langle v,w\rangle = \langle w,v\rangle$, *and*
(c) $\langle v+w,x\rangle = \langle v,x\rangle + \langle w,x\rangle$ *and* $\langle x,v+w\rangle = \langle x,v\rangle + \langle x,w\rangle$.

9.1.4. *The "Squaring" Norm.* We will define this norm in terms of an inner product, such that for $V = \mathbb{R}^n$:

$$||v||_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

This norm has been defined from an inner product, where $\sum_{i=1}^{n} x_i^2 = \langle v, v \rangle$ for $v = (x_1, x_2, \ldots, x_n)$. While properties (i) and (ii) are immediate from the properties of an inner product, we must still verify (iii) for this norm. In other words, we must show that:

$$(*) \qquad ||v + w||_2 \leq ||v||_2 + ||w||_2 \, .$$

However, because both terms are positive, showing that the inequality holds when both sides are squared is equivalent to showing that it holds as presented in (*). Therefore, we have, by (c) (also known as bilinearity),:

$$\begin{gathered} ||v + w||_2^2 \leq ||v||_2^2 + 2\,||v||_2\,||w||_2 + ||w||_2^2 \Rightarrow \\ \langle v + w, v + w \rangle \leq ||v||_2^2 + 2\,||v||_2\,||w||_2 + ||w||_2^2 \Rightarrow \\ \langle v, v \rangle + 2\,\langle v, w \rangle + \langle w, w \rangle \leq ||v||_2^2 + 2\,||v||_2\,||w||_2 + ||w||_2^2 \Rightarrow \\ |\,\langle v, w \rangle\,| \leq ||v||_2\,||w||_2 \end{gathered}$$

Note that the absolute value signs in the last inequality are valid because if it is true without the absolute value signs, then it is true with them. This last inequality is known as the **Cauchy-Schwarz inequality**, which is equivalent to:

$$\left|\sum x_i y_i\right| \leq \sqrt{\sum x_i^2}\sqrt{\sum y_i^2}$$

So now, to show that $||\cdot||_2$ is a norm, we will prove Cauchy-Schwarz.

Proof. Take $v, w \in V$. Consider $\langle v + \lambda w, v + \lambda w \rangle \geq 0, \lambda \in \mathbb{R}$.

$$\begin{gathered} \langle v + \lambda w, v + \lambda w \rangle \geq 0 \Rightarrow \\ \langle v, v \rangle + 2\lambda\,\langle v, w \rangle + \lambda^2\,\langle w, w \rangle \Rightarrow \\ ||v||_2^2 + 2\lambda\,\langle v, w \rangle + \lambda^2\,||w||_2^2 \geq 0 \Rightarrow \\ \left(\lambda\,||w||_2 + \frac{\langle v, w \rangle}{||w||_2}\right)^2 + ||v||_2^2 - \frac{\langle v, w \rangle^2}{||w||_2^2} \geq 0 \end{gathered}$$

Because this is true for all λ, we can choose λ so that the first term is 0. Then, we obtain:

$$||v||_2^2 - \frac{\langle v, w \rangle^2}{||w||_2^2} \geq 0 \Rightarrow ||v||_2^2\,||w||_2^2 \geq \langle v, w \rangle^2 \Rightarrow |\,\langle v, w \rangle\,| \leq ||v||_2\,||w||_2 \, .$$

$\square$

Therefore, since we have established the Cauchy-Schwarz inequality, $||\cdot||_2$ is a norm.

The point of all of this norm-taking was that given a norm, we can make it a metric on $\mathbb{R}^n$ by $d(v, w) = ||v - w||$. For instance, $||\cdot||_2$ is exactly the same as the aforementioned Euclidian metric on $\mathbb{R}^n$!

9.2. **Metric Space Concepts.** Let (X, d) be a metric space.

Definition. *The **open ball centered at a of radius r** for $r > 0, a \in X$ is:*

$$B(a,r) = \{x \in X \mid d(a,x) < r\}$$

Example: In $\mathbb{R}$, the open ball is denoted $(a - r, a + r)$. In the discrete metric, the balls are either the single point or the whole space. In $\mathbb{R}^2$, $||\cdot||_2$, $B(a,r)$ is the open circle of radius r around a. With $||\cdot||_\infty$, $B(0,1)$ is the square length 2 around O. With $||\cdot||_1$, $B(0,1)$ is the interior of the diamond created by taking the line $x + y = 1$ and reflecting it across all of the axes. $B(0,1)$ is known as the unit ball.

Definition. *Given a sequence (x_n) and an $x \in X$, (x_n) **converges to** x if:*

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) \ s.t. \ d(x_n, x) < \epsilon \ \forall n \geq N$$

Definition. *Given a sequence (x_n) and an $x \in X$, (x_n) is **Cauchy** if:*

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) \ s.t. \ n, m > N \Rightarrow d(x_n, x_m) < \epsilon$$

Definition. *(X, d) is **complete** if and only if every Cauchy sequence converges.*

Note: Convergence in $\mathbb{R}^n$ (with any of the norms defined above) is componentwise convergence. So, is $\mathbb{R}^n$ complete in its three metrics? Yes! The discrete metric is also complete.

Definition. *A subset $U \subseteq X$ is **open** if:*

$$\forall x \in U \ \exists r > 0 \ s.t. \ B(x,r) \subseteq U$$

Proposition. *$B(a,r)$ is open.*

Proof. Let $x \in B(a,r)$ be given. Then, $r > d(a,x)$, so $r - d(a,x) > 0$.

We will claim that $B(x, r - d(a,x)) \subseteq B(a,r)$. Suppose $y \in B(x, r - d(a,x))$. Then, we know that $d(x,y) < r - d(a,x) \Rightarrow d(a,x) + d(x,y) < r$. So, by the triangle inequality, $d(a,y) < r$. Therefore, for any $y \in B(x, r - d(a,x))$, y is in $B(a,r)$. Therefore, $B(x, r - d(a,x)) \subseteq B(a,r)$, and $B(a,r)$ is open. □

As another example, every subset of the discrete metric is open.

Definition. *A subset $E \subseteq X$ is **closed** if and only if $X \setminus E$ is open.*

Note that closed does *not* mean not open!
Example: In the discrete metric, every subset is both open and closed.

Proposition. *$E \subseteq X$ is closed if and only if $(x_n \in E)(x_n \to x) \Rightarrow x \in E$.*

Proof. We will go in both directions.

($\Rightarrow$) Assume $X \setminus E$ is open. Also, suppose $(x_n) \in E$ and $x_n \to x$, and toward contradiction, that $x \notin E$, or $x \in X \setminus E$. This implies that $\exists r > 0$ s.t. $B(x,r) \subseteq X \setminus E$. But since $x_n \to x$, $d(x_n, x) < r$ for some n. Then, $x_n \in B(x,r) \subseteq X \setminus E$, so $x_n \in (X \setminus E) \cap (E)$, which is a contradiction.

($\Leftarrow$) Toward contradiction, suppose $X \setminus E$ is not open. Then, $\exists x \in X \setminus E$ such that $B(x, \frac{1}{n}) \cap E \neq \emptyset \ \forall n$. This means that for every $n \in \mathbb{N}$, there exists some $y_n \in E$ such that $d(y_n, x) < \frac{1}{n}$. Thus, consider the sequence of y_ns, which we will call (y_n). (y_n) then obviously converges to x, but then by assumption and hypothesis, $x \in (X \setminus E) \cap (E)$, which is a contradiction.

□

10. 02/17/11: Metric Spaces, Part Two

10.1. Review.

Given (X, d), $U \subseteq X$ is open if $\forall x \in U\ \exists r > 0$ s.t. $B(x, r) \subseteq U^*$. If U_i where $i \in I$ is open for all i, then $\bigcup_{i \in I} U_i$ is open.

$E \subseteq X$ is closed if $X \setminus E$ is open†. If E_i is closed for $i \in I$, then $\bigcap_{i \in I} E_i$ is closed.

10.2. More Metric Space Concepts.

Definition. *For $x \in X$, a **neighborhood** of x is a set $S \subseteq X$ such that there exists an open set $U \subseteq X$ with $x \in U \subseteq S$.*

Definition. *The **closure** of E is:*

$$\overline{E} = \bigcap_{F\, closed, E \subseteq F} F$$

Observe that $\overline{E}$ is closed. This is the smallest closed set that contains E. Note that $\overline{B(a, r)}$ is not necessarily $C(a, r) = \{x \in X \mid d(a, x) \le r\}$; for instance, this is not true in the discrete metric space.

Note that subsets can be both open *and* closed! For example, define a metric space where $X = [-1, 1] \cup (5, 17)$. Then, each interval is an open and closed subset of the space. Another example is if $X = \mathbb{Q}$, then $(-\sqrt{2}, \sqrt{2})$ is open and closed.

Definition. *The **interior** of $S \subseteq X$ is $\mathring{S} = \{s \in S \mid \exists r > 0$ s.t. $B(s, r) \subseteq S\}$, or $\{s \in S \mid S$ is a neighborhood of $S\}$.*

Definition. *The **boundary** of S is $\partial S = \overline{S} \setminus \mathring{S}$.*

10.3. The p-adic Metric.

We will define a metric on $\mathbb{Z}$ based on any p for p prime. We will call this a p-adic metric.

Define a metric by a norm, $|j|_p$. Write $j = p^r \cdot m$ such that $(m, p) = 1$. Let $|0|_p = 0$ and $|j|_p = 2^{-r}$ where 2 can be any prime.

Then it is the case that

$$|j|_p \ge 0 \text{ and } j = 0 \iff |j|_p = 0$$

as well as that

$$|i + j|_p \le \max\{|i|_p, |j|_p\} \le |i|_p + |j|_p\,.$$

Furthermore, $d_p(i, j) = |i - j|_p$, and this metric makes $\mathbb{Z}$ into a metric space. Note that, for instance, for $p = 3$, $3^n \to 0$.

*Note that the empty set is open.

†Note that the empty set is closed.

Example:

$$(1-3)(1+3+3^2+\cdots+3^n) = 1-3^{n+1} \to 1$$

Therefore, $2+2\cdot 3+2\cdot 3^2+\cdots+2\cdot 3^n+\cdots = -1$.

We can also define $|\cdot|_p$ on $\mathbb{Q}$, where $q = p^r \cdot \frac{m}{n}$ for $(m,p)=1, (n,p)=1$. Then, $|q|_p = 2^{-r}$ and $d_p(q_1,q_2) = |q_1-q_2|_p$.

Note that we can write numbers in "decimal" form by writing them as such:

$$x = \sum_{i=1}^{\infty} a_i \cdot p^i + \sum_{j=1}^{\infty} d_i \cdot p^{-i} := \ldots a_3a_2a_1.d_1d_2d_3\ldots$$

Then, we get a "decimal" where as you add terms to the left it converges and as you add to the right it diverges!

We can define a field with this metric by defining it like we did $\mathbb{R}$, but with this new metric instead of the familiar one for $\mathbb{R}$. This field is called the p-adic numbers. Note that this technique of completing a metric space (as we did when constructing $\mathbb{R}$ from $\mathbb{Q}$) can be applied to any metric.

10.4. **Even More Metric Space Concepts.**

Definition. *Let (X,d) a metric space and $S \subseteq X$. Then, $x \in X$ is called an **accumulation point** (or limit point or cluster point) of S if every neighborhood of x contains a point in S other than x.*

Note that $E \subseteq X$ is closed if and only if it contains all of its accumulation points.

Definition. *Let (X,d) be given, and consider $\mathcal{T} := \{U \,|\, U \text{ open in } X\}$. This collection of subsets is called a **topology** on X defined by d. A topology satisfies three axioms:*

1. *$\emptyset \in \mathcal{T}, X \in \mathcal{T}$.*
2. *If $U_i \in \mathcal{T}$, then $i \in I \Rightarrow \bigcup_{i\in I} U_i \in \mathcal{T}$.*
3. *For $U_1, U_2, \ldots, U_n \in \mathcal{T}$, it is the case that $\bigcap_{i=1}^n U_i \in \mathcal{T}$.*

Any property that can be expressed only using $\mathcal{T}$, or open sets, is called a **topological property**. For instance, we can say...

Proposition. *Convergence is a topological property, or in other words:*

$$\lim_{n\to\infty} x_n = x \iff \forall U \text{ open in } X \text{ and } x \in U,\ \exists N \text{ s.t. } n \geq N \Rightarrow x_n \in U.$$

Proof. We will go in both directions.

($\Rightarrow$) Let U be open with $x \in U$ given. Since U is open, $\exists r > 0$ s.t. $B(x,r) \subseteq U$. Therefore, $\exists N$ s.t. $n \geq N \Rightarrow d(x_n,x) < r \Rightarrow x_n \in B(x,r) \subseteq U$.

($\Leftarrow$) Let $\epsilon > 0$ be given. We know that $B(x,\epsilon)$ is open. Thus, $\exists N$ s.t. $n \geq N \Rightarrow x_n \in B(x,\epsilon) \Rightarrow d(x_n,x) < \epsilon$.

$\square$

Therefore, changing the metric does not change the notion of convergence on a given topology. Note that Cauchiness and completeness are not topological properties, but closures and interiors are.

Definition. *Let U, V be open subsets of X. A topological space X is **Hausdorff** if given that $x \in U$, $y \in V$, and $x \neq y$, then $U \cap V = \emptyset$.*

11. 02/24/11: Continuity and Cardinality, Part One

11.1. More About Metric Spaces.

Proposition. $x \in \overline{E} \iff (\forall \epsilon > 0)\ B(x, \epsilon) \cap E \neq \emptyset$.

Proof. We will go in both directions.

($\Rightarrow$) We will proceed by contradiction. Suppose there exists an $\epsilon > 0$ such that $B(x, \epsilon) \cap E = \emptyset$. Then, $X \setminus B(x, \epsilon)$ is a closed set containing E. By the definition of a closure, this implies $x \in \overline{E} \subseteq X \setminus B(x, \epsilon)$. This is a contradiction because $x \in B(x, \epsilon)$.

($\Leftarrow$) Let F be closed and $E \subseteq F$. Suppose $x \notin F$. Then there exists an $\epsilon > 0$ such that $B(x, \epsilon) \subseteq X \setminus F$. Because $E \subseteq F$, $X \setminus F \subseteq X \setminus E$. This is a contradiction, so for $x \in F$ and $\forall F$ where F is closed, $E \subseteq F \Rightarrow x \in \overline{E}$.

□

Definition. *Given (X, d) and $A, B \subseteq X$, we say that A is **dense** in B if $B \subseteq \overline{A}$.*

Note that $\overline{A^X} \cap B = \overline{A^B} \cap B$ (where $\overline{A^X}$ is the closure of A under X and $\overline{A^B}$ is the closure of A under B), by the ϵ-criteria for $x \in \overline{E}$.

11.2. Continuity.

Definition (ϵ-δ). *Given metric spaces (X, d) and (Y, D), a function $f\colon X \to Y$ is said to be **continuous at** x if*

$$(\forall \epsilon > 0)(\exists \delta > 0)\ s.t.\ d(x, y) < \delta \Rightarrow D(f(x), f(y)) < \epsilon\,.$$

Note that this is continuity at a *single point*!
Example: $f\colon \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = \begin{cases} 0, & \text{if } x \in \mathbb{Q}; \\ x, & \text{otherwise.} \end{cases}$$

This is a function which is continuous at 0 and discontinuous anywhere else.

Proposition. *f is continuous at $x \in X \iff (x_n \to x) \Rightarrow (f(x_n) \to f(x))$.*
Note: *This is known as the limit definition of continuity at a point.*

Proof. We will go in both directions.

($\Rightarrow$) Suppose $x_n \to x$ and let $\epsilon > 0$ be given. Choose a δ such that

$$d(x, t) < \delta \Rightarrow d(f(x), f(t)) < \epsilon\,.$$

Using δ in the definition of $x_n \to x$, get an N s.t. $\forall n \geq N$, $d(x_n, x) < \delta$. This implies that $d(f(x_n), f(x)) < \epsilon$.

($\Leftarrow$) Suppose f is not continuous at $x \in X$. This means that:

$$\exists \epsilon > 0 \text{ s.t. } \forall \delta > 0, \exists t \in B(x, \delta) \text{ s.t. } d(f(t), f(x)) \geq \epsilon\,.$$

Choose a sequence $x_n \in B(x, \frac{1}{n})$ with $d(f(x_n), f(x)) \geq \epsilon$. Then, we have a contradiction, as $x_n \to x$ but $f(x_n) \nrightarrow f(x)$.

□

Note that this has become the commonly accepted definition of continuity. However, this is because of the shift in focus from topological spaces to metric spaces (a reversal of a previous trend). In a general topological space, the limit definition of continuity does *not* hold!

Definition. *Given metric spaces X, Y and $f\colon X \to Y$, f is* ***continuous*** *if f is continuous at every point of X.**

Proposition. *$f\colon X \to Y$ is continuous $\iff$ $f^{-1}(U)$ is open for all open $U \subseteq Y$.*†

Proof. We will go in both directions.

($\Rightarrow$) Let U be open in Y. Take $x \in f^{-1}(U)$. This implies that $f(x) \in U$. By the definition of continuity, $\exists B(x,\delta)$ s.t. $f(B(x,\delta)) \subseteq U$, so $B(x,\delta) \subseteq f^{-1}(U)$.

($\Leftarrow$) Take $x \in X$ and let $\epsilon > 0$ be given. We must show that

$$\exists \delta > 0 \text{ s.t. } f(B(x,\delta)) \subseteq B(f(x),\epsilon)\,.$$

However, $B(f(x),\epsilon)$ is open, so $f^{-1}(B(f(x),\epsilon))$ is open by hypothesis. Also, $x \in f^{-1}(B(f(x),\epsilon))$, so $\exists \delta > 0$ such that $B(x,\delta) \subseteq f^{-1}(B(f(x),\epsilon))$, which implies that $f(B(x,\delta)) \subseteq B(f(x),\epsilon)$.

$\square$

Definition. *(X,d) and (Y,D) are* ***isometric*** *if there exists a bijection $f\colon X \to Y$ with $D(f(x),f(y)) = d(x,y)$. This implies that f^{-1} is also an isometry.*

Definition. *$f\colon X \to Y$ is called a* ***homeomorphism*** *if f is a continuous bijection whose inverse is continuous.*

Note: The last condition is necessary! For instance, the identity function from $(\mathbb{R}, \mathit{discrete})$ to $(\mathbb{R}, \mathit{usual})$ is continuous but f^{-1} is not.

11.3. **Countability and Cardinality.**

We will take a little detour from metric spaces in order to talk about Cardinality, so as to prepare for the guest lecture in two lectures from Professor Hugh Woodin.

First, we will explore the Cantor Set, defined as such:

$$S_0 = [0,1]$$

$$S_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

$$S_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$$

$$S_3 = \left[0, \frac{1}{27}\right] \cup \left[\frac{2}{27}, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{7}{27}\right] \cup \left[\frac{8}{27}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{19}{27}\right] \cup \left[\frac{20}{27}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, \frac{25}{27}\right] \cup \left[\frac{26}{27}, 1\right]$$

The Cantor set is defined as $C := \bigcap_{n \in \mathbb{N}} S_n$. Essentially, we start with $[0,1]$ and remove the middle third each time, for every interval. The endpoints of the intervals in the set are definitely countable, but the set itself as actually uncountable!

---

*The related concept of uniform continuity is defined in a footnote in Lecture 16.

†Reminder: Given sets A, B and $f\colon A \to B$, if:

$S \subseteq A$, define $f(S) = \{f(x) \,|\, x \in S\}$, and

$T \subseteq B$, define $f^{-1}(T) = \{x \in S \,|\, f(x) \in T\}$.

In this case, f^{-1} is *not* the inverse of f! f^{-1} always exists when applied to a set.

12. 02/24/11: Cardinality, Part Two

12.1. The Cantor Set, Continued.

Recall the Cantor set from last time. Here is an inductive way to define it.

Take an interval $I = [a, b]$. We will define $\mathcal{C}(I) = \{[a, a + \frac{b-a}{3}], [a + \frac{2(b-a)}{3}, b]\}$. So, given $J_0 = \{[0, 1]\}$, we have $J_{n+1} = \bigcup_{I \in J_n} \mathcal{C}(I)$. Then,

$$C = \bigcap_{n\in\mathbb{N}} \Big(\bigcup_{I \in J_n} I \Big) .$$

Proposition. *The interior of the Cantor set is empty, or* $\mathring{C} = \emptyset$.

Proof Sketch. Choose n, ϵ with $3^{-n} < \frac{\epsilon}{10}$. Then, for some n, x is in an interval of width 3^{-n}, so there is no ball around x. □

12.2. Random Facts.

There exists a continuous nondecreasing function $f : [0, 1] \to [0, 1]$ with $f(0) = 0$ and $f(1) = 1$ such that f is constant on $[0, 1] \setminus C$. The graph of f is known as the Devil's Staircase.

Russell's Paradox. Given a set S, $S \in S$ or $S \notin S$. Define $A = \{S \mid S \notin S\}$. Does A belong to A? There is no way for A to belong to or not belong to itself, as both result in contradictions. Thus, Russell's Paradox will act as a cautionary tale for the discussions to come; we must be careful when we're talking about sets of all sets.

12.3. Cardinality.

We will denote the cardinality of a set A as $|A|$ (or $\#(A)$).

Definition. *Given sets A and B, $|A| = |B|$ if there exists a bijection $f : A \to B$.*

Definition. *A is **finite** if there exists an $n \in \mathbb{N}$ such that $|A| = |\{1, 2, \ldots, n\}|$. In other words, for finite sets, $|A| = |B| \iff$ they have the same number of elements. A is **infinite** if it is not finite.*

Definition. *A is **countable** if A is finite or $|A| = |\mathbb{N}|$, or equivalently, you can enumerate the elements of A as a sequence.*

Proposition. *$S \subseteq \mathbb{N}$ is countable.*

Proof. Define (c_n) as follows. Define c_1 to be the smallest element of S, and define $S_1 = S \setminus \{c_1\}$. Then, define c_{n+1} to be the smallest element of S_n, and define $S_{n+1} = S_n \setminus \{c_n\}$. Then, it is easy to check that $\{c_k \mid k \in \mathbb{N}\} = S$. □

Definition. *$|A| \leq |B|$ (also written $A \precsim B$) if there exists an injection $f : A \to B$.*

Example: If $S \subseteq T$, then $S \precsim T$. Therefore, this order on sets says that up to a change of name, $A \subseteq B$.

Proposition. *$A \precsim B \iff \exists$ a surjection $g\colon B \to A$.*

Proof. We will go in both directions.

($\Rightarrow$) Given an injection $f\colon A \to B$, define f such that for $a \in A$, $g(f(a)) = a$ and for everything else, $g(x) = a$ for some a.

($\Leftarrow$) We want to find an injection $f\colon A \to B$. For each $a \in A$, choose an $x \in B$ with $g(x) = a$, and define $f(a) = x$.

$\square$

Proposition. *$\mathbb{Z}$, $\mathbb{Z} \times \mathbb{Z}$, and $\mathbb{Q}$ are countable.*

Proof. We will prove them one at a time.

(1) $\mathbb{Z}$: We will arrange the elements of $\mathbb{Z}$ in a sequence, namely (c_n) such that $c_0 = 0, c_{2n-1} = n, c_{2n} = -n$ for all $n \in \mathbb{N}$.

(2) $\mathbb{Z} \times \mathbb{Z}$: By the previous argument, $|\mathbb{Z}| = |\mathbb{N}|$, so it suffices to show that $\mathbb{N} \times \mathbb{N}$ is countable. We will arrange the elements of $\mathbb{N} \times \mathbb{N}$ as such:

$$\begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & \dots \\ (2,1) & (2,2) & (2,3) & (2,3) & \dots \\ (3,1) & (3,2) & (3,3) & (3,4) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{matrix}$$

Now, proceed down each diagonal, starting at $(1,1)$, then moving to the next diagonal containing $(1,2)$ and $(2,1)$, then the next, containing $(1,3)$, $(2,2)$, $(3,1)$, and so on. Alternately, define a function $f\colon \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ such that $f((m,n)) = 2^m 3^n$. By the unique prime factorization of integers, this is certainly an injection.

(3) $\mathbb{Q}$: Take the function $f\colon \mathbb{Z} \times \mathbb{N} \to \mathbb{Q}$ such that $f((m,n)) = \frac{m}{n}$. This is clearly a surjection.

$\square$

12.4. **Counting $\mathbb{R}$.**

We will define $2^{\mathbb{N}} = \{\text{sequences } (a_n) \text{ such that } a_n \in \{0,1\}\}$, such as $01101001\ldots$ Note that this is the same as the set of all subsets of $\mathbb{N}$, and in general, 2^A is the same as subsets of A, also known as the power set of A.

Theorem (Cantor)**.** *$2^{\mathbb{N}}$ is uncountable.*

Proof. Suppose it is countable. Then, we can enumerate them as such:

$$\begin{matrix} (a_n)_1 = & a_{1,1} & a_{2,1} & a_{3,1} & \dots \\ (a_n)_2 = & a_{1,2} & a_{2,2} & a_{3,2} & \dots \\ (a_n)_3 = & a_{1,3} & a_{2,3} & a_{3,3} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{matrix}$$

Construct $b_n = a_{n,m} + 1 \pmod 2$. Obviously, b_n is in $2^{\mathbb{N}}$, but it cannot be the nth term in the sequence of sequences because it differs from the nth term in the nth place. Therefore, it is not in the list, and we have a contradiction. $\square$

Corollary. $\mathbb{R}$ *is uncountable.*

Proof. It suffices to define an injection $f\colon 2^{\mathbb{N}} \to \mathbb{R}$. Let $f((a_n)) = \sum_{n=1}^{\infty} a_n 10^{-n}$*. We will claim that f is an injection. The proof is in the lecture notes on Cardinality, page 2. $\square$

There is another proof in the notes, page 3, which shows that $\mathbb{R}$ has cardinality no less than $|2^{\mathbb{N}}|$.

Theorem (Schroeder-Berinstein)**.** *If* $|A| \leq |B|$ *and* $|B| \leq |A|$*, then* $|A| = |B|$*.*

Proof. Next lecture. $\square$

By the remark after the Corollary and this Theorem, we can see that $|\mathbb{R}| = |2^{\mathbb{N}}|$.

The Continuum Hypothesis asks whether there is a subset of $\mathbb{R}$ that is uncountable but that does not have the same cardinality as $2^{\mathbb{N}}$. Or in other words, if there exists a set A such that $|\mathbb{N}| \leq |A| \leq |2^{\mathbb{N}}|$. The answer is still not known.

*Note that series are not defined yet. However, partial sums are Cauchy sequences, so they are reals.

13. 03/01/11: Set Theory: Guest Lecture by Hugh Woodin

13.1. Schroeder-Bernstein.

Two prominent set theorists were Cantor and Gödel, both of whom went insane. Anyway...

Definition. *Two sets X, Y have the same cardinality if there exists a bijection $\Pi\colon X \to Y$. We denote this as $|X| = |Y|$.*

Caution: This definition is not as intuitive as it seems. If X is infinite and $a \in X$, then it can happen that X and $X \setminus \{a\}$ have the same cardinality.

Definition. *Suppose X, Y are sets. Then, the cardinality of X is less than or equal to the cardinality of Y if there exists an injection $\Pi\colon X \to Y$. We denote this as $|X| \leq |Y|$.*

Now, we will turn our attention to the Schroeder-Bernstein Theorem.

Theorem (Schroeder-Bernstein)**.** *Let X, Y be sets and suppose $|X| \leq |Y|$ and $|Y| \leq |X|$. Then, $|X| = |Y|$. In other words, suppose X, Y are sets and there exists injections $f\colon X \to Y$ and $g\colon Y \to X$. Then, there exists a bijection $\Pi\colon X \to Y$.*

Before proving the theorem, we will first develop some concepts and properties that will help us along the way.

Consider a set Z and an injection $h\colon Z \to Z$. Start with some point $a \in Z$, and keep applying h to it. Then, one of three things will happen:

(1) We will get into a loop, so $h^n(a) = a$ for some $n \in \mathbb{N}$.
(2) We will reach a point where $h^{-n}(a)$ is not in the range of h, so we cannot "back up" any farther, and effectively, the chain starts at a certain point.
(3) We will never cycle, but we will be able to "back up" endlessly, so for all $n \in \mathbb{N}$, $h^{-n}(a)$ is in the range of h.

Notice that case 3 is just a degenerate case of case 1, as it is simply just an endless loop backwards. Now, we will define a concept using this intuition.

Definition. *Suppose $h\colon Z \to Z$ is injective and suppose $a \in Z$. Then the* ***orbit*** *of a given by h is the subset of Z generated by a, h, h^{-1} where $h^{-1}\colon \{\text{range of } h\} \to Z$. More formally, define $X_0 = \{a\}$. Then, define inductively:*

$$x_{n+1} = \{h(b) \mid b \in X_n\} \cup \{c \in Z \mid h(c) \in X_n\} \cup X_n$$

Then, the orbit of a given by h, denoted Z_a, is defined to be $\bigcup_{n=0}^{\infty} X_n$ where $n \in \mathbb{N}$.

Example: $X_1 = \{a, h(a)\} \cup \{c\}$ where $h(c) = a$ if such a c exists. Furthermore, $X_2 = X_1 \cup \{h(h(a))\} \cup \{d\}$ where $h(d) = c$ if c existed and if d exists.

Example: Let $Z = \{0, 1, 2, \dots\}$, and let $h(n) = n^2$ for $n \in Z$. Then, $Z_0 = \{0\}$, $Z_1 = \{1\}$, and $Z_2 = \{2, 4, 16, 256, \dots\}$, or in other words the even powers of 2. Z_3 can also be seen as the even powers of 3.

Note that we can also denote the orbit of a given by h as Z_a^h in order to explicitly state the injection that we are dealing with.

Lemma 1. *Suppose $h\colon Z \to Z$ is injective and suppose $a, b \in Z$. Let Z_a^h and Z_b^h be the orbits of a and b relative to h, as defined above. Also, suppose $Z_a^h \cap Z_b^h \neq \emptyset$. Then, $Z_a^h = Z_b^h$.*

Proof. Exercise. □

Another intuitive definition follows from our examination of orbits above.

Definition. *Suppose $h\colon Z \to Z$ is injective and suppose $a \in Z$. Let Z_a^h be the orbit of a given by h. Z_a^h has an* ***eve-point*** *if there exists a b in Z_a^h such that b is not in the range of h.*

Lemma 2. *If an orbit has an eve-point, then the eve-point of that orbit is unique.*

Proof. Exercise. □

Lemma 3. *Let $h\colon Z \to Z$ be injective and $a \in Z$. Suppose Z_a^h has no eve-point. Then, $h|_{Z_a^h}\colon Z_a^h \to Z_a^h$ is a surjection, and hence a bijection.*

Proof. Exercise. □

Now, we are ready for the actual proof.

Proof (Schroeder-Bernstein). We are given injections $f\colon X \to Y$ and $g\colon Y \to X$. We need to construct a bijection $\Pi\colon X \to Y$. Consider the functions $g \circ f\colon X \to X$ and $f \circ g\colon Y \to Y$. These are both injections because the composition of two injections is also always injective. So now, we can consider the orbits of points in X given by $g \circ f$ and the orbits of points in Y given by $f \circ g$. We will define $\Pi\colon X \to Y$ by defining $\Pi|_{X_a^{g\circ f}}$ for each orbit of X given by $g \circ f$. There are two cases for each orbit:

(1) $X_a^{g\circ f}$ has no eve-point. Then, for each $b \in X_a^{g\circ f}$, $\Pi(b) = f(b)$.
(2) $X_a^{g\circ f}$ has an eve-point. Let a_0 be the eve-point. Then, it follows that $Y_{f(a)}^{f\circ g}$ must also have an eve-point. Let b_0 be the eve-point of $Y_{f(a)}^{f\circ g}$, and define $\Pi(a_0) = b_0$, $\Pi(a_1) = b_1$, and so on.

We can see that every point of every orbit is hit exactly once by Π, and thus Π is a bijection from X to Y. □

This theorem proves the rather trivial-looking statement that if $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|X| = |Y|$. However, while it may look trivial, it really is not.

Proposition. $|\mathbb{R}| = |P(\mathbb{N})|$*, where* $P(\mathbb{N}) = \{A \mid A \subseteq \mathbb{N}\}$.

Proof. It suffices to show that $|\mathbb{R}| \leq |P(\mathbb{N})|$ and $|P(\mathbb{N})| \leq |\mathbb{R}|$.

(1) Clearly, there is a bijection $\Pi\colon \mathbb{N} \to \mathbb{Q}$, so we just need an injection $f\colon \mathbb{R} \to P(\mathbb{Q})$. Define $f(r) := \{q \mid q \in \mathbb{Q}, q < r\}$ for all $r \in \mathbb{R}$. This is clearly injective, so $|\mathbb{R}| \leq |P(\mathbb{Q})| = |P(\mathbb{N})|$.
(2) We need an injection $g\colon P(\mathbb{N}) \to \mathbb{R}$. Define $g(a)$, where $a \in P(\mathbb{N})$, as such:

$$g(a) = \begin{cases} \sum_{i\in a} (\frac{1}{10})^i & \text{if } a \neq \emptyset \\ 0 & \text{if } a = \emptyset \end{cases}$$

This is clearly injective, so $|P(\mathbb{N})| \leq |\mathbb{R}|$.

□

Theorem (Generalized Cantor)**.** *Suppose X is a set. Then, $|X| \neq |P(X)|$, where $P(X) = \{A \mid A \subseteq X\}$**.

*Note that this implies $|X| < |P(X)|$ and $|\mathbb{R}| = |P(\mathbb{N})| \neq |\mathbb{N}|$.

Proof. Suppose $f: X \to P(X)$. We need to produce an $A \subseteq X$ such that A is not in the range of f. Define $A = \{a \in X \mid a \notin f(a)\}$. Then, A cannot be in the range of f. To see this, suppose $A = f(b)$. Then, we have:

$$b \in A \iff b \notin f(b) \iff b \notin A$$

which is clearly a contradiction. $\square$

This theorem implies that there are an infinite number of infinities, each one bigger than some others.

Therefore, we will say $\aleph_0 = |\mathbb{N}|$. Where does $|P(\mathbb{N})| = |\mathbb{R}|$ fit in? Is it $\aleph_1$?

The Continuum Hypothesis states that $|\mathbb{R}| = \aleph_1$. We have gotten to the point where we know that given our standard axioms, we can build universes where this is true and universes where this is false. For another illustration of a conjecture like this, consider the Parallel Postulate. It is true in $\mathbb{R}^2$, but false in a universe that is a circle without the edge, despite that the usual axioms of geometry hold in this new space.

In recent times, we've gotten closer to an answer. Behind this question and a lot of other set theory lie many philosophical issues about the meaning of things that don't exist...

14. 03/03/11: Compactness, Part One

14.1. Basic Concepts.

Let $A \subseteq X$, and let (X, d) be a metric space.

Definition. *An* ***open cover*** *of A is a family $\{U_\alpha \,|\, \alpha \in I\}$ of open subsets of X such that $A \subseteq \bigcup_{\alpha \in I} U_\alpha$.*

Example: $[0,1] \subseteq \mathbb{R}$. Then, $\{(-1,2)\}$ is an open cover. $\{(x-\epsilon, x+\epsilon) \,|\, x \in [0,1]\}$ is also an open cover. Consider $\epsilon = \frac{1}{3}$. Then, the cover is very redundant, as it is an uncountably large collection of intervals, and we could come up with something more economical.

Definition. *If $\{U_\alpha \,|\, \alpha \in I\}$ is an open cover of A, a* ***subcover*** *is a subfamily $\{U_\alpha \,|\, \alpha \in J\}$, where $J \subseteq I$, that is also an open cover of A.*

Definition. *$A \subseteq X$ is* ***compact*** *if for every open cover, there is a finite subcover.*

It is the case that $A \subseteq X$ is compact if and only if $A \subseteq A$ with the inherited metric is compact. Therefore, the notion of compactness is an inherited notion, like closedness. Note that this is a very difficult definition to work with, if we want to show a set is compact. However, this does make it very easy to show that a set is not compact.

Example: Finite metric spaces are always compact for any $A \subseteq X$.

Example: Take $\mathbb{R}$. $\{(x-\epsilon, x+\epsilon) \,|\, x \in \mathbb{R}\}$ is an open cover of $\mathbb{R}$. Furthermore, you cannot cover $\mathbb{R}$ with finitely many intervals, so $\mathbb{R}$ is not compact. Another example is, for $\mathbb{Z}$, the open cover $\{(n - \frac{3}{2}, n + \frac{3}{2})\}$.

14.2. Basic Properties.

Proposition. *In any metric space, a compact subset is closed and bounded*[*].

Proof. We will prove the two points separately.

1. First, we will prove boundedness. Take $x \in A$. Consider $\{B(x,n) \,|\, n \in \mathbb{N}\}$. This family of sets is trivially an open cover, even of all of X. Take a finite subcover $\{B(x, n_1, \ldots, B(x, n_{\max})\}$. Then, $\bigcup_{i=1}^{\max} B(x, n_i) = B(x, n_{max})$. $A \subseteq \bigcup_{i=1}^{\max} B(x, n_i)$ because it's an open subcover, so $A \subseteq B(x, n_{\max})$ and A is bounded.
2. Now, we will prove closedness. Let $x \in X \setminus A$, and define:
$$U_n = X \setminus \{y \,|\, d(x,y) \leq \tfrac{1}{n} \text{ where } n \in \mathbb{N}\}.$$
We will claim that the U_ns cover A, or that every $a \in A$ is in the complement of some $\{y \,|\, d(x,y) \leq \frac{1}{n}\}$. In other words, for some n, $d(a,x) > \frac{1}{n}$, or $n > \frac{1}{d(a,x)}$. This is clearly true by the Archimedean property of $\mathbb{R}$, so our claim is true. Now, extract a finite subcover $\{U_1, U_2, \ldots, U_M\}$. Notice that $\bigcup_{i=1}^{M} U_i = U_M$, so $A \subseteq U_M$. Then, we know that:
$$\{y \,|\, d(x,y) \leq \tfrac{1}{M}\} \subseteq X \setminus A \subseteq U_M$$
so it follows that $B(x, \frac{1}{M}) \subseteq X \setminus A$. Thus, because x was arbitrary, there is an open ball around every point in $X \setminus A$, and A is closed.

[*] A subset $S \subseteq X$ is **bounded** if there exists an $a \in X$ and an $R > 0$ such that $S \subseteq B(a, R)$.

□

Proposition. *If (X, d) is a compact metric space, then any closed subset of X is compact.*

Proof. Let A be closed, and let $\bigcup_{\alpha \in I} U_\alpha$ be an open cover of A. It is clear that $\{U_\alpha \mid \alpha \in I\} \cup \{X \setminus A\}$ is an open cover of X. By the compactness of X, we know that X must then also be covered by $\bigcup_{n\in\mathbb{N}}^{N} U_{\alpha_n} \cup X \setminus A$, for some $N \in \mathbb{N}$. Then, it is clear that $\bigcup_{n\in\mathbb{N}}^{N} U_{\alpha_n}$ is a finite subcover of A, because $X \setminus A$ cannot cover anything in A. □

14.3. **Sequential Compactness.**

Definition. *Given (X, d) and $A \subseteq X$, A is* ***sequentially compact*** *if every sequence in A has a convergent subsequence in A.*

Example: By the Bolzano-Weierstrass Theorem, every closed interval $[a, b]$ is sequentially compact. Furthermore, by extension, $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ is sequentially compact in $\mathbb{R}^n$. To see this in $\mathbb{R}^2$, first get a convergent subsequence in the first coordinate, which we can do because if we consider just the sequence (a_n), it is a sequence in $\mathbb{R}$. Then, fix that subsequence and get a convergent subsubsequence in the second coordinate. Any subsequence of a convergent sequence converges, so both coordinates converge and we have a convergent subsequence in $\mathbb{R}^2$. This process can easily be generalized to any n.

Recall that when we defined compactness, we said that the definition was very hard to work with. This definition is much easier, so we would like for them to be equivalent. Lucky for us, they are!

Theorem. *Let (X, d) be a metric space. Then, X is compact if and only if X is sequentially compact.*

Proof. In the next lecture. □

15. 03/08/11: Compactness, Part Two

15.1. **(Sequential) Compactness.**

Remember: $E \subseteq X$ is compact if every open cover of E has a finite subcover, and $E \subseteq X$ is sequentially compact if every sequence in E has a convergent subsequence in E.

Theorem. *Given a metric space (X, d), X is compact if and only if X is sequentially compact.*[*]

Proof. We will go in both directions.

($\Rightarrow$) Let (x_n) be a sequence in X. If $\{x_n \,|\, n \in \mathbb{N}\}$ is finite, then some x_n must be hit infinitely many times, so we would be done. Therefore, assume $\{x_n\}$ is infinite. Let $x \in X$. Either x is the limit point of a subsequence of (x_n), or there exists an $r > 0$ such that $|B(x,r) \cap \{x_n \,|\, n \in \mathbb{N}\}| \leq 1$. To see this, assume that the latter is not true for some $x \in X$. Then, for all $r > 0$, $|B(x,r) \cap \{x_n \,|\, n \in \mathbb{N}\}| > 1$. In this case, we could take ever-decreasing values of r, and there would be at least one x_n such that $d(x, x_n) < r$ for every $r > 0$. Therefore, if you picked all of those terms, we would have a convergent subsequence of (x_n). So, in other words, there is a convergent subsequence of (x_n) or $\bigcup_{x \in X} B(x,r)$ forms an open cover for X. By compactness, we can obtain a finite subcover of this open cover, so only finitely many $B(x,r)$ cover X, and hence $\{x_n\}$. Then, because for every x, $|B(x,r) \cap \{x_n \,|\, n \in \mathbb{N}\}| \leq 1$, the finiteness of this subcover would imply that $\{x_n\}$ was also finite, which is a contradiction.

($\Leftarrow$) Let $\{U_\alpha\}$ be an open cover of X.

First, we will prove this statement:

$$\exists R > 0 \text{ s.t. } \forall x \in X, \exists U_\alpha \text{ s.t. } B(x,R) \subseteq U_\alpha\,.$$

Proceeding by contradiction, we will assume that no such R exists. Because no such R exists, in particular, for each $n \in \mathbb{N}$, there exists an $x_n \in X$ such that $B(x_n, \frac{1}{n}) \nsubseteq U_\alpha$ for any U_α. Now, because X is sequentially compact, we can get a convergent subsequence of this sequence which converges to some $x \in X$, which we will denote $(x_{n_k}) \to x$. We know that this $x \in U_\alpha$ for some U_α, and because U_α is open, there exists an $r > 0$ such that $B(x,r) \subseteq U_\alpha$. Then, by the convergence of (x_{n_k}), there is some $N \in \mathbb{N}$ such that for all $n \geq N$, the following are true:

(1) $\frac{1}{n} < \frac{r}{2}$, and
(2) $x_n \in B(x,r)$.

However, by (2), we have that $B(x_n, \frac{r}{2}) \subseteq B(x,r) \subseteq U_\alpha$, so in turn, by (1), we have that $B(x_n, \frac{1}{n}) \subseteq U_\alpha$, which contradicts that we can find $x_n \in X$ such that $B(x_n, \frac{1}{n}) \nsubseteq U_\alpha$, which means that our initial assumption that such R does not exist was false. Therefore, our initial statement is true.[†]

Now, we will show that finitely many balls of radius R (as defined above) cover X. It is sufficient to show this because each ball is then contained in some U_α, so then we know that finitely many U_α cover X, and we will

[*] Note that both of these are topological properties, but this theorem is not true in a general topological space.

[†] This is a result that is of independent interest, because it proves the existence of Lebesgue numbers.

have extracted a finite subcover of a general open cover of X. We will proceed by contradiction, and assume that it is not the case that finitely many balls of radius R cover X, and that it is impossible to generate a finite subcover. Then, we will construct a sequence (x_n) as such. First, simply pick an $x_1 \in X$. $B(x_1, R)$ does not cover X, so we can pick a point x_2 outside $B(x_1, R)$ such that $d(x_1, x_2) > \frac{R}{2}$. Then, pick another point x_3 outside of $B(x_1, R) \cup B(x_2, R)$ such that $d(x_i, x_3) > \frac{R}{2}$ for $i = 1, 2$. Continue this process, such that x_{n+1} is not in $\bigcup_{i=1}^{n} B(x_i, R)$ and $d(x_i, x_{n+1}) > \frac{R}{2}$ for $i = 1, 2, \ldots, n$, for all $n \in \mathbb{N}$. Then, the latter statement is obviously true for all subsequences of (x_n). Therefore, this sequence cannot have any convergent subsequences, a contradiction of our initial assumption. Thus, it must be the case that finitely many balls of radius R cover X, so X must be compact.

□

Theorem (Heine-Borel)**.** *A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof. We will proceed in both directions.

($\Rightarrow$) We already proved this statement for an arbitrary metric in the last lecture.

($\Leftarrow$) Let $A \subseteq \mathbb{R}^n$ be closed and bounded. By the boundedness of A, we know that $A \subseteq B(0, R+1)$ for some $R > 0$. Then, it is clearly the case that $A \subseteq [-R, R] \times [-R, R] \times \cdots \times [-R, R]$. The right hand side of that expression is sequentially compact (from last lecture) and hence compact (from this lecture), and we also proved last time that a closed subset of a compact set is also compact. Therefore, A is compact.

□

From this theorem, we obtain some very useful results, such as that the Cantor set is compact, the unit sphere is compact, and so on.

Now, we will build up to another major, but seemingly unrelated, result.

15.2. The Equivalence of Norms on $\mathbb{R}^n$.

Proposition. *Let X, Y be metric spaces such that X is compact. Then, given a continuous function $f \colon X \to Y$, $f(X)$ is compact.*

Proof. Let U_α be an open cover of $f(X)$. Then, it is clear that $f^{-1}(U_\alpha)$ covers X. Because f is continuous, we know that for any open set $U \subseteq Y$, $f^{-1}(U)$ is also an open set. Thus, $f^{-1}(U_\alpha)$ is an open cover of X, and by the compactness of X, for some $n \in \mathbb{N}$, we have that:

$$X \subseteq \bigcup_{i=1}^{n} f^{-1}(U_{\alpha_i})$$

This, in turn, implies that $f(X) \subseteq \bigcup_{i=1}^{n} U_{\alpha_i}$, so $f(X)$ is compact. □

Proposition. *Let X be a metric space. If X is compact, then any continuous function $f \colon X \to \mathbb{R}$ attains its minimum and maximum values.*

Proof. From the proposition above, we know that $f(X) \subseteq \mathbb{R}$ must be compact, so it is therefore closed and bounded. Thus, $\sup f(X)$ and $\inf f(X)$ exist, and by the closedness of $f(X)$, must lie inside $f(X)$ itself. $\square$

Corollary. *Let X be a compact metric space. Then, given a continuous function $f \colon X \to \mathbb{R}$ such that $f(x) > 0$ for all $x \in X$, it is the case that there exists an $\epsilon > 0$ such that $f(x) \geq \epsilon$ for all $x \in X$.*

Proof. This follows immediately from the above proposition. $\square$

Theorem. *Any two norms on $\mathbb{R}^n$ are equivalent, or in other words, given two norms $|\cdot|$ and $||\cdot||$ on $\mathbb{R}^n$, there exist $A, B > 0$ such that for all $x \in \mathbb{R}^n$,*

$$A|x| \leq ||x|| \leq B|x| \, .$$

Proof. In next lecture. $\square$

16. 03/10/11: Norms, Separability, Connectedness

16.1. Equivalence of Norms on $\mathbb{R}^n$.

Theorem. *Any two norms on $\mathbb{R}^n$ are equivalent, or in other words, given two norms $||\cdot||$ and $|||\cdot|||$ on $\mathbb{R}^n$, there exist $A, B > 0$ such that for all $x \in \mathbb{R}^n$,*

$$A||x|| \leq |||x||| \leq B||x||\,.$$

Proof. We will make two trivial observations to start. First, note that if we replace $|||\cdot|||$ with $||\cdot||_2$, we will obtain the same result. Second, the statement is obviously true for any A, B if $x = 0$, so we will assume that $x \neq 0$*.

First, we will prove the first inequality. Let e_i be a unit vector with a 1 in the ith place. Then, $x = \sum_{i=1}^n x_i e_i$. Then, we can obtain that

$$\left|\left|\sum_{i=1}^n x_i e_i\right|\right| \leq \sum_{i=1}^n (|x_i| \cdot ||e_i||) \leq \sqrt{\sum_{i=1}^n |x_i|^2} \cdot \sqrt{\sum_{i=1}^n ||e_i||^2}$$

by the triangle inequality, followed by the Cauchy-Schwarz inequality. Furthermore, $\sqrt{\sum_{i=1}^n |x_i|^2} = ||x||_2$, so this implies

$$\frac{1}{\sqrt{\sum_{i=1}^n ||e_i||^2}}||x|| \leq ||x||_2$$

which proves the first inequality. Observe that this result implies that $||\cdot||$ is uniformly continuous† for $||\cdot||_2$, or, in other words, that the identity function from $(\mathbb{R}^n, ||\cdot||)$ to $(\mathbb{R}^n, ||\cdot||_2)$ is uniformly continuous (and vice-versa).

Now, we will prove the second inequality. Let $\epsilon > 0$ be given. Then, choose $\delta = A\epsilon$, where A is the term we found above. Then, if $||x-y||_2 < \delta$, then

$$||x-y|| \leq \frac{1}{A}||x-y||_2 < \frac{1}{A}\delta = \frac{A\epsilon}{A} = \epsilon\,.$$

As noted above, this means that the identity function from $(R^n, ||\cdot||_2)$ to $(R^n, ||\cdot||)$ is (uniformly) continuous. In particular, let $S^n = \{x \in \mathbb{R}^n \,|\, ||x||_2 = 1\}$. Then, clearly, the identity function from $(S^n, ||\cdot||_2)$ to $(R^n, ||\cdot||)$ is also continuous. We know that for any $x \in S^n$, $||x||_2 = \sum_{i=1}^n x_i^2 = 1$, so S^n is obviously closed and bounded, and hence compact, by the Heine-Borel Theorem. Furthermore, $||\cdot||$ is a norm, so $||x|| > 0$ for all $x \in S^n$. By the theorem from last lecture, this means that there exists an $\epsilon > 0$ such that $||x|| > \epsilon$ for all $x \in S^n$. So, if we take an arbitrary $v \neq 0$ in $\mathbb{R}^n$, then we have:

$$\epsilon = \epsilon\left|\left|\frac{v}{||v||_2}\right|\right|_2 < \left|\left|\frac{v}{||v||_2}\right|\right| \Rightarrow \epsilon||v||_2 < ||v|| \Rightarrow ||v||_2 < \frac{1}{\epsilon}||v||$$

completing the proof. ☐

*Note that 0, in this case, is the zero vector in $\mathbb{R}^n$, not the real number 0.

†Given metric spaces (X, d) and (Y, D), a function $f\colon X \to Y$ is **uniformly continuous** if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x, y \in X$,

$$d(x,y) < \delta \Rightarrow D(f(x), f(y)) < \epsilon\,.$$

The difference between continuity and uniform continuity is that the δ in uniform continuity must work for *all* x, y.

Now, we will discuss two final metric space topics before diving into the subject of Calculus.

16.2. **Separability.**

Definition. *Let (X, d) be a metric space. X is **separable**[*] if there exists a countable dense subset of X.*

Example: $\mathbb{R}$ is separable because $\mathbb{R} \subseteq \overline{\mathbb{Q}}$ and $|\mathbb{Q}| = |\mathbb{N}|$.

Proposition. *Every subset of a separable metric space is separable with the inherited metric.*

Proof. Let (X, d) be a metric space, Q be the countable dense subset of X, and $A \subseteq X$. Consider $\{B(q, \frac{1}{n}) \mid q \in \mathbb{Q}, n \in \mathbb{N}\}$. This is a countable set, as $\mathbb{Q} \times \mathbb{N}$ is countable. For each q, n, if $B(q, \frac{1}{n}) \cap A \neq \emptyset$, choose a point in it, which we will denote $x_{q,n}$. Then, let $D = \{x_{q,n} \mid q \in \mathbb{Q}, n \in \mathbb{N}\}$. It is clear that D is countable and a subset of A. Therefore, we only need to show that it is dense, or in other words, given $a \in A$ and $\epsilon > 0$, that there exists an $x \in D$ such that $x \in B(a, \epsilon)$. Choose an $n \in \mathbb{N}$ such that $\frac{1}{n} < \frac{\epsilon}{2}$. Because Q is dense in X, there exists a $q \in Q$ such that $d(q, a) < \frac{1}{n} < \frac{\epsilon}{2}$. This implies that $B(q, \frac{1}{n}) \cap A \neq \emptyset$, so for some $x \in D$, $x \in B(q, \frac{1}{n})$. Therefore, we have:

$$d(x, a) \leq d(x, q) + d(q, a) < \frac{1}{n} + \frac{1}{n} < \epsilon$$

completing the proof. $\square$

The notion of separability might seem a bit random and unrelated, but in the words of Professor Jones, "Paradise is a complete separable metric space."

16.3. **Connectedness.**

Definition. *Let (X, d) be a metric space. $C \subseteq X$ is **connected**[†] if for open subsets U, V of X:*

$$C \subseteq U \amalg V \Rightarrow C \subseteq U \text{ or } C \subseteq V.$$

Note: We define the empty set to be connected.
Example: All sets which are made up of just one point are obviously connected.

Proposition. *The continuous image of a connected set is connected.*

Proof. Take a continuous function $f \colon X \to Y$, where X, Y are metric spaces. Then, we want to show that for a connected subset C of X and open subsets U, V of Y,

$$f(C) \subseteq U \amalg V \Rightarrow f(C) \subseteq U \text{ or } f(C) \subseteq V.$$

Notice that $C \subseteq f^{-1}(U) \amalg f^{-1}(V)$. Then, by the connectedness of C, $C \subseteq f^{-1}(U)$ or $C \subseteq f^{-1}(V)$. Thus, $f(C) \subseteq U$ or $f(C) \subseteq V$, so $f(C)$ is connected. $\square$

[*] This is a topological notion.
[†] This is also a topological notion.

Remember that for a continuous function, the inverse image of an open set is also open.

Theorem. *A subset of $\mathbb{R}$ is connected if and only if it is an interval.*[*]

Proof. We will go in both directions.

($\Leftarrow$) Let $I \subseteq \mathbb{R}$ be an interval. Let U and V be open sets in $\mathbb{R}$, and suppose $I \subseteq U \amalg V$. Now, we will proceed by contradiction and assume $I \cap U \neq \emptyset$ and $I \cap V \neq \emptyset$. Choose $s \in I \cap U$ and $t \in I \cap V$. Without loss of generality, assume $s < t$. Consider the set $\{x \,|\, [s,x] \subseteq U\}$. This set is clearly bounded above by t. Also, s is in the set, because $s \in U$. Therefore, the supremum exists, so we will say $a = \sup\{x \,|\, [s,x] \subseteq U\}$. Because U is open, if a was in the set, it would be the case that $(a-\epsilon, a+\epsilon) \subseteq U$, and hence $[s, a+\frac{\epsilon}{2}] \subseteq U$. However, a was the supremum of U, so this could not be the case, and $a \notin U$. Similarly, $a \notin V$, because if it was, then $(a-\epsilon, a+\epsilon) \subseteq V$, which would imply $a - \frac{\epsilon}{2} \in V$, a contradiction because anything less than a must be in U and we assumed that $U \cap V = \emptyset$. Therefore, $a \notin U$ and $a \notin V$. However, this creates a contradiction with our initial assumption, because $[s,t] \subseteq I$ and $s < a < t$. Thus, any interval in $\mathbb{R}$ is connected.

($\Rightarrow$) Let $A \subset \mathbb{R}$ be connected. We will claim that then, it has the property that if $s,t \in A$ where $s < t$, then $[s,t] \subseteq A$. To see this, assume that this is not true, and that for some $s < x < t$, $x \notin A$. Then, it is clear that $A \subseteq (-\infty, x) \cup (x, \infty)$, a disjoint union of two open sets. However, we also know that $s \in (-\infty, x)$ and $t \in (x, \infty)$, which is a contradiction because A is connected. Therefore, our claim is true, and A must be an interval.

$\square$

Corollary. *Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function. Then, $f([a,b])$ is also a closed interval, $[c,d]$.*

Proof. This follows immediately from the two previous results. $\square$

This is a very powerful corollary, as it shows that a continuous function on an interval simply stretches, shrinks, and/or shifts the interval!

[*]Interval, here, means **any** sort of interval, namely any one from this list: (a,b), $(a,b]$, $[a,b)$, $[a,b]$, $(-\infty,b)$, $(-\infty,b]$, (a,∞), $[a,\infty)$.

17. 03/15/11: More (Path) Connectedness, Series and Convergence

17.1. More on Connectedness.

Recall that $E \subseteq (X, d)$ is connected if for open subsets U, V of X:

$$E \subseteq U \amalg V \Rightarrow E \subseteq U \text{ or } E \subseteq V\,.$$

We have already shown that in $\mathbb{R}$, the only connected sets are intervals, as well as that the continuous image of a connected subset is connected. Furthermore, note that connectedness is an inherited property, like compactness.

Proposition. *Let A, B be connected subsets of X. If $A \cap B \neq \emptyset$, then $A \cup B$ is connected.*

Proof. Let U, V be open sets in X, and suppose $A \cup B \subseteq U \amalg V$. Then, $A \subseteq U \amalg V$ and $B \subseteq U \amalg V$. Without loss of generality, assume $A \subseteq U$. Then, $A \cap B \subseteq U$. We know that $A \cap B \neq \emptyset$, so there exists a $b \in B$ such that $b \in U$. By the connectedness of B it follows that $B \subseteq U$. Therefore, $A \cup B \subseteq U$, completing the proof. $\square$

Now, given a metric space X, define a relation $\sim$ on X by:

$$x \sim y \iff \{x, y\} \subseteq A$$

for some connected subset A of X. The properties of reflexivity and symmetry are trivial, so we only need to check transitivity. Assume $x \sim y$ and $y \sim z$. Then, $\{x, y\} \subseteq A$ and $\{y, z\} \subseteq B$, for connected subsets A, B of X. Note that $y \in A \cap B$, so $A \cup B$ is connected by the previous proposition, and $\{x, z\} \subseteq A \cup B$. Therefore, $x \sim z$, and this is an equivalence relation.

Definition. *The equivalence classes as defined above are called the* ***connected components*** *of X.*

Note: The connected components are themselves connected and are maximal elements for the property of being connected.
Example: In $\mathbb{Q}$, the connected components are just individual points.

17.2. Path-connectedness.

Definition. *Let (X, d) be a metric space, and $x, y \in X$. A* ***path**[*] *from x to y is a continuous function $c : [0, 1] \to X$ such that $c(0) = x$ and $c(1) = y$.*

Definition. (X, d) *is* ***path-connected*** *if for all $x, y \in X$, there exists a path from x to y. Equivalently, there exists a continuous function $c\colon [a, b] \to X$ such that $c(a) = x$ and $c(b) = y$.*

Proposition. *If a metric space is path-connected, then it is connected.*

Proof. Let (X, d) be a metric space, and let U, V be open sets in X. Suppose $X = U \amalg V$. Notice that if either U or V is the empty set, then the proposition is trivially true. Thus, assume that U and V are nonempty. Proceeding toward contradiction, let $x, y \in X$ and choose $x \in U$ and $y \in V$. Now, let $c\colon [0, 1] \to X$ be a continuous function such that $c(0) = x$ and $c(1) = y$. Then, $c^{-1}(U)$ is a nonempty open subset of $[0, 1]$ as $0 \in c^{-1}(U)$, and similarly, V, is also a nonempty open

[*] This is a topological notion.

subset of $[0,1]$ because $1 \in c^{-1}(V)$. Also, we have that $c^{-1}(U) \cup c^{-1}(V) = [0,1]$ and $c^{-1}(U) \cap c^{-1}(V) = \emptyset$ because c is a well-defined function. However, because $0 \in U$ and $1 \in V$, this contradicts that $[0,1]$ is a connected set, so we have a contradiction and X must be connected. $\square$

However, the converse is *not* true! For an example of this, see the Topologist's sine curve (or the discrete analog, with sawtooth forms instead of smooth curves).

Like with connected sets, we can define components on path-connected sets.

Definition. *Let (X, d) be a metric space. Define a relation $\sim$ on X such that for any $x, y \in X$, $x \sim y$ if and only if there exists a path from x to y.*

We will check that this is an equivalence relation.

(1) Reflexive: It is obvious that $x \sim x$, because $c(t) = x$ where $0 \leq t \leq 1$ is clearly a path from x to x.
(2) Symmetric: Assume $x \sim y$. Then, given a path $c\colon [0,1] \to X$ such that $c(0) = x$ and $c(1) = y$, define another function $d\colon [0,1] \to X$ such that $d(t) = c(1-t)$ where $0 \leq t \leq 1$. This is clearly a path from y to x.
(3) Transitive: Assume $x \sim y$ and $y \sim z$. Then, there exist two paths, $c\colon [0,1] \to X$ such that $c(0) = x$ and $c(1) = y$, and $d\colon [0,1] \to X$ such that $d(0) = y$ and $d(1) = z$. Define $cd\colon [0,2] \to X$ such that:

$$cd(t) = \begin{cases} c(t) & \text{if } 0 \leq t \leq 1, \\ d(t-1) & \text{if } 1 \leq t \leq 2 \end{cases}$$

Now, we just need to show that this new function is continuous. This is obvious for all t where $0 \leq t < 1$ and $1 < t \leq 2$, because c and d were themselves continuous. Thus, we only need to verify continuity at $t = 1$, which can be done using ϵ-δ. Therefore, $x \sim z$.

Definition. *The equivalence classes as defined above are called the **path components** of X.*

That concludes our time with metric spaces. We will now move on to Calculus! Everything in this part of the class will be done in $\mathbb{R}$.

17.3. Convergence of Series.

Definition. *Let (a_n) be a sequence. Then,*

$$\sum_{n=0}^{\infty} a_n = L \iff \lim_{n \to \infty} s_n = L$$

*where $s_n = \sum\limits_{k=0}^{n} a_n$.[*] We will say that $\sum\limits_{n=0}^{\infty} a_n$ **converges** if $\sum\limits_{n=0}^{\infty} a_n = L$ and $L \neq \pm\infty$.*

Example: $\sum\limits_{n=0}^{\infty} r^n = \frac{1}{1-r}$ for $|r| < 1$.

[*] L can be ∞ or $-\infty$.

Example: $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$, because $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, so writing out s_n, we get $1 - \frac{1}{n+1}$, which tends to 1.

Here are a couple of remarks about series:

(1) If $\sum a_n$ converges, then $a_n \to 0$ because $a_n = s_n - s_{n-1}$.
(2) If $\sum a_n$ and $\sum b_n$ converge, then $\sum(a_n + b_n)$ converges to $\sum a_n + \sum b_n$ by the limit theorems.

Theorem (Comparison Test)**.** *For $0 \le a_n \le b_n$, if $\sum b_n$ converges, then so does $\sum a_n$.*

Proof. $s_n = \sum_{k=0}^{n} a_k$ is nondecreasing if all of the terms a_n are positive. Thus,

$$s_k \le \sum_{k=0}^{n} b_k \le \sum_{k=0}^{\infty} b_k\,.$$

So, $\{s_n \,|\, n \in \mathbb{N}\}$ is bounded above, and a bounded monotone sequence always converges. Therefore, $\sum a_n$ converges. $\square$

The textbook gives an alternate proof for this test, involving the fact that any series whose partial sums form a Cauchy sequence converges (known as the Cauchy criterion). This is the case because any Cauchy sequence in $\mathbb{R}$ converges. This method is also used to prove the theorem regarding absolute convergence as well as the Alternating Series Test below.

Example: $\sum \frac{1}{n^2}$ converges because $\frac{1}{n^2} < \frac{1}{n(n-1)}$ for large n.

Definition. $\sum_{n=0}^{\infty} a_n$ ***converges absolutely*** *if* $\sum_{n=1}^{\infty} |a_n|$ *converges.*

Proposition. *If a series converges absolutely, then it converges.*

Proof. Notice that $2a_n = (|a_n| + a_n) - (|a_n| - a_n)$. Both of these two terms are greater than or equal to zero, and by the triangle inequality, we have:

$$|a_n| + a_n \le 2|a_n| \text{ and } |a_n| - a_n \le 2|a_n|\,.$$

Furthermore, because $\sum a_n$ converges absolutely, $\sum 2|a_n|$ also converges. Then, by the theorem, because the two equations above hold,

$$\sum(|a_n| + a_n) - (|a_n| - a_n) = \sum 2a_n = 2\sum a_n$$

also converges. $\square$

Theorem (Alternating Series Test)**.** *The series $\sum_{n=0}^{\infty} (-1)^n a_n$, where $a_n \ge 0$, converges if $a_n \ge a_{n+1}$ and $\lim_{n\to\infty} a_n = 0$.*

Proof. Supplied in the Textbook, Theorem 15.3. $\square$

Notice that if $\sum(-1)^n a_n = L$ where $L \ne \pm\infty$, then $\left|\sum_{n=1}^{k} a_n - L\right| \le |a_{k+1}|$. Furthermore, it is the case that the sequence of even partial sums is decreasing and that the sequence of odd partial sums is increasing.

18. 03/17/11: More Convergence, Power Series, Differentiability

18.1. More Convergence.

Theorem (Ratio Test)**.** *Let $\sum a_n$ be a series such that $a_n > 0$ for all $n \in \mathbb{N}$. Then, if $\limsup \frac{a_{n+1}}{a_n} < 1$, then $\sum a_n$ converges.*

Proof. Let $r = \limsup \frac{a_{n+1}}{a_n}$. Choose an R such that $r < R < 1$. Then, there exists an $N \in \mathbb{N}$ such that $\sup\{\frac{a_{N+1}}{a_N}, \frac{a_{N+2}}{a_{N+1}}, \dots\} < R$. This implies that $\frac{a_{N+k+1}}{a_{N+k}} < R$ for all $k \geq 0$. Thus,

$$a_{N+1} < Ra_N,\ a_{N+2} < Ra_{N+1} < R^2 a_N,\ \dots\ ,\ a_{N+k} < R^k a_N\,.$$

So, consider $\sum_{j=N}^{\infty} a_j$. This converges by comparison with $\sum_{j=N}^{\infty} R^j a_N$ because $a_j < R^j a_N$. Therefore, $\sum a_n$ converges. □

Example: $\sum \frac{n}{2^n}$ is a series that we can easily use the Ratio Test on.

Theorem (nth Root Test)**.** *Let $\sum a_n$ be a series such that $a_n \geq 0$ for all $n \in \mathbb{N}$. If $\limsup \sqrt[n]{a_n} < 1$, then $\sum a_n$ converges.*

Proof. Let $r = \limsup \sqrt[n]{a_n}$. Choose an R such that $r < R < 1$. Using $\epsilon = R - r$ in the definition of lim sup, we can see that there exists an $N \in \mathbb{N}$ such that $\sup\{(a_N)^{\frac{1}{N}}, (a_{N+1})^{\frac{1}{N+1}}, \dots\} < R$. For $n \geq N$, we have $(a_n)^{\frac{1}{n}} < R$, so $a_n < R^n$. Then, we can conclude that $\sum a_n$ converges by following the same procedure as with the Ratio Test above. □

18.2. Power Series.

In general, a power series is a particular kind of series, of the form $\sum\limits_{n=1}^{\infty} a_n(x - x_0)^n$. However, we usually just say $x_0 = 0$ and write the power series as $\sum\limits_{n=1}^{\infty} a_n x^n$.

Theorem. *Let $\sum a_n x^n$ be a power series. Also, let*

$$\beta = \limsup \sqrt[n]{|a_n|} \text{ and } R = \frac{1}{\beta}\,.$$

*Then**,

1. *If $|x| < R$, $\sum a_n x^n$ converges absolutely.*
2. *If $|x| > R$, $\sum a_n x^n$ does not converge.*

Proof. We will prove the two cases.

1. We want to show that $\sum |a_n x^n|$ converges. We can write because $|a_n x^n| = |a_n||x^n|$, we must show that $\sum |a_n||x^n|$ converges. We will use the nth root test.

   $\limsup \sqrt[n]{|a_n||x^n|} = \limsup(\sqrt[n]{|a_n|}\,|x|) = |x| \limsup \sqrt[n]{|a_n|} = |x|\beta\,.$

   So, the power series converges if $|x|\beta < 1$, or $|x| < \frac{1}{\beta} = R$†.

---
*Note that β can be infinite.

†R is known as the radius of convergence.

(2) Assume $x\beta > 1$. Then, working backwards from the chain of equalities above, we have that $\limsup \sqrt[n]{|a_n||x^n|} > 1$. If $\sqrt[n]{|a_n||x^n|} > 1$, then it is clearly the case that $|a_n||x|^n > 1$. However, this would mean that the terms of the power series do not tend to zero, so the series as a whole does not converge.

$\square$

Here are some very useful power series to know.

- "e^x" $= \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$
- $\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$
- $\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$
- $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ for $|x| < 1$.
- $\log(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}$ for $|x| < 1$.

18.3. **Differentiability.**

Let be a function $f\colon (x,y) \to \mathbb{R}$ where $x, y \in \mathbb{R}$, and let $a \in (x,y)$.

Definition. *f is **differentiable at** a if $\lim_{x\to a} \frac{f(x)-f(a)}{x-a} = L$ for some $L \in \mathbb{R}$. Then, we will write $L = f'(a)$.*

However, we have not really defined what $\lim_{x\to a} f(x)$, or in other words, the limit of a function at a, means, so we will do that now, for general metric spaces.

Definition. *Let $f\colon X \to Y$ where (X,d) and (Y,D) are metric spaces, and let $a \in X$ and $y \in Y$. Then:*

$$\lim_{x\to a} f(x) = y \iff (\forall \epsilon > 0)(\exists \delta > 0) \text{ s.t. } 0 < d(x,a) < \delta \Rightarrow D(f(x),y) < \epsilon\,.$$

An alternate definition of this concept is that if $\lim_{x\to a} f(x) = y$, then:

$$F(x) = \begin{cases} f(x), & \text{if } x \neq a \\ y, & \text{if } x = a \end{cases}$$

is continuous.

Theorem. *If f is differentiable at a, then f is continuous at a.*[*]

[*]Note that this does *not* apply for general metric spaces! We are returning to the definitions of f and a given in the very beginning of this section.

Proof. We can prove this using the ϵ-δ definition of the limit of a function. However, we will try using the alternate $F(x)$ method instead.

Because we know that $\lim_{x\to a} \frac{f(x)-f(a)}{x-a} = L$ for some $L \in \mathbb{R}$, we can define

$$F(x) = \begin{cases} \frac{f(x)-f(a)}{x-a}, & \text{if } x \neq a \\ L, & \text{if } x = a. \end{cases}$$

where $F(x)$ is continuous. We want to show that $f(x)$ is continuous at a. In particular, we know that $F(x)$ is continuous at $x = a$. Thus, for $x \neq a$, we have:

$$F(x) = \frac{f(x) - f(a)}{x - a} \Rightarrow f(x) = f(a) + F(x)(x - a)\,.$$

Then, because $F(x)$ is continuous at a, for any $\epsilon > 0$, there exists a $\delta > 0$ such that $|a - x| < \delta \Rightarrow |L - F(x)| < \epsilon$. This means:

$$|L - F(x)| = \left|L - \tfrac{f(x)-f(a)}{x-a}\right| = |L(x-a) + f(a) - f(x)| \leq \tag{1}$$

$$|L\delta + f(a) - f(x)| < \epsilon\,. \tag{2}$$

In order to prove the continuity of $f(x)$ at a, we must find some γ such that $|a - x| < \gamma \Rightarrow |f(a) - f(x)| < \epsilon$. We will proceed casewise, for when L is positive and negative.

Let L be positive and $\epsilon > 0$ be given. Then, set $\gamma = \delta$. Assuming $|a-x| < \gamma = \delta$, we obtain:

$$|f(a) - f(x)| < |L\delta + f(a) - f(x)| < \epsilon\,.$$

Therefore, $f(x)$ is continuous in this case.

Now, let L be negative and $\epsilon > 0$ be given. Using a simple change of variables, we will simply write $L = -L$, and keep $L > 0$. Notice that, in the series of inequalities (1) above, $|-L - F(x)| = |F(x) + L|$. Therefore, we can write (2) as $|f(x) - f(a) + L\delta| < \epsilon$. Then, letting $\gamma = \delta$, and assuming $|a - x| < \gamma$, we obtain:

$$|f(a) - f(x)| = |f(x) - f(a)| < |f(x) - f(a) + L\delta| < \epsilon\,.$$

Therefore, $f(x)$ is also continuous in this case, completing our proof. ☐

Now, we will simply list some basic properties of derivatives. We will not prove these because the proofs are fairly elementary, and involve only simple limit computations. However, the proofs are in the textbook and any other elementary calculus textbook. We will just assume that this theorem is true, and use its results liberally.

Theorem. *Let f, g be functions on an open interval in $\mathbb{R}$ and let a be in that interval. Assume that f, g are differentiable at a. Then, the following are true:*

(1) $(cf)'(a) = cf'(a)$, *where c is a constant.*
(2) $(f + g)'(a) = f'(a) + g'(a)$.
(3) $(fg)'(a) = f(a)g'(a) + f'(a)g(a)$.
(4) *If $g(a) \neq 0$, then* $\left(\frac{f}{g}\right)'(a) = \dfrac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}$.

For the quotient rule (part 4 above), notice that because $g(a) \neq 0$ and g is continuous at a, there exists an open interval containing a such that $g(x) \neq 0$ for all x in that interval, making the theorem in fact valid as stated.

19. 03/29/11: Theorems About the Derivative

19.1. **Review.**

Suppose we have a function $f\colon (a,b) \to \mathbb{R}$ where (a,b) is an open interval. If $\lim_{x\to a} \frac{f(x)-f(a)}{x-a} = L$, then $L = f'(a)$, and f is differentiable on a. Equivalently, we can define a function $F(x)$ such that $F(x) = \frac{f(x)-f(a)}{x-a}$ if $x \neq a$ and $F(x) = L$ if $x = a$. Then, F is differentiable at a if and only if F is continuous at a.

Example: $f(x) = x^2\chi$ where $\chi(x)$ is 1 if $x \in \mathbb{Q}$ and 0 if $x \notin \mathbb{Q}$. f is differentiable at $x = 0$. Because all sorts of functions are differentiable at a single point, this concept isn't really useful. What we really want is differentiability in the whole interval.

19.2. **The Main Results.**

Theorem (Chain Rule)**.** *Let I and J be open intervals, and let $f\colon I \to \mathbb{R}$ and $g\colon J \to \mathbb{R}$ be functions such that if $a \in I$, then $f(a) \in J$. If f is differentiable at a and g is differentiable at $f(a)$, then $g \circ f$ is differentiable at a, and*

$$(g \circ f)'(a) = g'(f(a))f'(a)\,.$$

Note: This is just the familiar chain rule, or in other words, $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$.

Proof. Define a function $h\colon I \to \mathbb{R}$ such that:

$$h(y) = \begin{cases} \frac{g(y)-g(f(a))}{y-f(a)}, & \text{if } y \neq f(a), \\ g'(f(a)), & \text{if } y = f(a). \end{cases}$$

Then, $h(y)$ is clearly continuous because g is differentiable at $f(a)$. Also, define a function $GF\colon I \to \mathbb{R}$ such that:

$$GF(x) = \begin{cases} \frac{g(f(x))-g(f(a))}{x-a}, & \text{if } x \neq a, \\ g'(f(a))f'(a), & \text{if } x = a. \end{cases}$$

Now, it suffices to show that GF is continuous at $x = a$.

Notice that $(h \circ f)(x) = \frac{g(f(x))-g(f(a))}{f(x)-f(a)}$. Then, we have $GF(x) = h \circ f(x) \cdot \frac{f(x)-f(a)}{x-a}$ when $x \neq a$ and $GF(x) = g'(f(a))f'(a)$ when $x = a$. We know that $(h \circ f)(x)$ goes to $g'(f(a))$ as x approaches a by the definition of h and $\frac{f(x)-f(a)}{x-a}$ approaches $f'(a)$ as x approaches a by the differentiability of f, so we are done. $\square$

Theorem. *Let I be an open interval in $\mathbb{R}$, and let $f\colon I \to \mathbb{R}$. Suppose $x_0 \in I$ and that $f(x_0)$ is a maximum [or minimum] on I. If f is differentiable at x_0, then $f'(x_0) = 0$.*

Proof. Suppose without loss of generality that $f'(x_0) > 0$. Then, we will claim that it is the case that for all $\epsilon > 0$, there exists an x such that $x_0 < x < x + \epsilon$ and $f(x) > f(x_0)$. Assume toward contradiction that this is not the case. Then, we have, for every ϵ, some x such that $f(x) \leq f(x_0)$. This implies that $\frac{f(x)-f(x_0)}{x-x_0} \leq 0$, which furthermore implies that $\lim\limits_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} \leq 0$, which is a contradiction because $f'(x_0) > 0$. Therefore, because there exists an x such that $f(x) > f(x_0)$, $f(x_0)$ cannot be a maximum. A similar argument can be used to see that $f(x_0)$ cannot be a minimum. $\square$

Theorem (Rolle)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be continuous and differentiable on (a,b), and suppose $f(a) = f(b)$. Then, there exists an $x \in (a,b)$ such that $f'(x) = 0$.*

Proof. By the compactness of $[a,b]$, f attains its maximum (and minimum) on $[a,b]$. If the maximum is attained at an endpoint, then it must attain its minimum inside (a,b) or the function must be constant (as the minimum would then be equal to the maximum). In the former case, the previous theorem proves our claim, and in the latter, the derivative is 0 everywhere. □

Theorem (Mean Value)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be continuous and differentiable on (a,b). Then, there exists an $x \in (a,b)$ such that*

$$f'(x) = \frac{f(b) - f(a)}{b - a}$$

Proof. Let $t(x) = f(x) - \left(f(a) + (x-a)\frac{f(b)-f(a)}{b-a}\right)$. Notice that $t(a) = 0$ and $t(b) = 0$. Also, $t'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$. So, by Rolle's Theorem, there exists an x such that $0 = f'(x) - \frac{f(b)-f(a)}{b-a}$, or $f'(x) = \frac{f(b)-f(a)}{b-a}$. □

Corollary. *Let $f\colon (a,b) \to \mathbb{R}$ and suppose $f'(x) = 0$ for all $x \in (a,b)$. Then, f is a constant function.*

Proof. If $p > q$ in (a,b), then $0 = \frac{f(p)-f(q)}{p-q}$, which implies that $f(p) = f(q)$ by the Mean Value Theorem. □

Corollary. *If $f\colon (a,b) \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for all $x \in (a,b)$, then f is strictly increasing.*

Proof. Notice that $f'(x) = \frac{f(p)-f(q)}{p-q}$ for some $p > q$ in (a,b), so $f(p) > f(q)$. □

Note that this corollary can be generalized to all of the variations of $>$, namely $\geq, <$, and $\leq$, for which the corollary would state increasing, strictly decreasing, and decreasing, respectively.

19.3. **Taylor's Theorem.**

Let $f\colon (a,b) \to \mathbb{R}$ be differentiable. Then, we can talk about f' as a function, namely $f'\colon (a,b) \to \mathbb{R}$. So, we could potentially differentiate this function and get $(f')'(x)$, or $f''(x)$, and so on, where the nth derivative of f at x is denoted $f^{(n)}(x)$, and is defined to be $(f^{(n-1)})'(x)$.

Definition. $C^n(a,b) := \{f\colon (a,b) \to \mathbb{R} \mid f', f'', \ldots, f^{(n)}$ *are continuous on* (a,b).$\}$

Notice that $C^\infty(a,b) = \bigcap_{n=1}^{\infty} C^n(a,b)$ are the smooth functions on (a,b).

As a side note, when dealing with Taylor series, we will always expand the series around 0. Refer to the textbook (p. 174) for a brief explanation of shifting series in order to expand them around arbitrary points.

Definition. *Given a function $f \in C^\infty(a,b)$ such that $a < 0 < b$, the Taylor series* of f is defined as*

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)\, x^n}{n!}$$

This is clearly a power series, but we would like to know if it converges, what it converges to, and if it actually has anything to do with the function. Luckily, we have a Theorem that tells us just that.

Theorem (Taylor)**.** *Suppose f has n derivatives on (a,b), where $a < 0 < b$. Then, for $x \in (a,b)$, there exists a y between 0 and x† such that*

$$f(x) = f(0) + f'(0)\,x + \frac{f''(0)\,x^2}{2!} + \cdots + \frac{f^{(n-1)}(0)\,x^{n-1}}{(n-1)!} + \frac{f^n(y)\,x^n}{n!}$$

Proof. Let M be the solution to the equation

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)\,x^k}{k!} + \frac{Mx^n}{n!}$$

for some fixed nonzero x. Furthermore, define the function

$$g(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)\,t^k}{k!} + \frac{Mt^n}{n!} - f(t)$$

Notice that $g(x) = 0$. Furthermore, $g(0) = 0$ and $g^{(k-1)}(0) = 0$ for all k such that $2 \le k \le n$. So, by Rolle's Theorem, there exists an x_1 between 0 and x such that $g'(x_1) = 0$. Then, again by Rolle's Theorem, because $g'(0) = 0$ and $g'(x_1) = 0$, there exists an x_2 between 0 and x_1 such that $g''(x_2) = 0$. Continuing, we get that there exists an x_{n-1} between 0 and x_{n-2} such that $g^{(n-1)}(x_{n-1}) = 0$. Finally, there exists a y between 0 and x_{n-1} such that $g^{(n)}(y) = 0$. However, $g^{(n)}(y) = M - f^{(n)}(y)$, so $M = f^{(n)}(y)$ for some y between 0 and x, proving the theorem. □

*Technically, the Maclaurin series

†This means that if $0 < x$, then $0 < y < x$, and if $x < 0$, then $x < y < 0$.

20. 03/31/11: Differentiation of Power Series

20.1. **Consequences of Taylor's Theorem.**

Recall that Taylor's Theorem states that for any n-times differentiable function $f(x)$, there exists a y between 0 and x such that:

$$f(x) = f(0) + \frac{f'(0)\,x}{1!} + \cdots + \frac{f^{(n-1)}(0)\,x^{n-1}}{(n-1)!} + \frac{f^{(n)}(y)\,x^n}{n!}$$

A simple corollary of this theorem follows.

Corollary. *If there exists a bound on the set $\{f^{(n)}(y) \,|\, y$ varies as n varies$\}$, then the Taylor series of f converges to f.*

To see this, notice that if there is a bound on the set, as n increases, $\frac{x^n}{n!}$ converges to 0.

Taylor's Theorem is very powerful, as it is what allows us to compute exact values of differentiable functions!

20.2. **Term-by-Term Differentiation of Power Series.**

Theorem. *Let $\sum a_n x^n$ be a power series with radius of convergence R. Define $f(x) = \sum a_n x^n$ on $(-R, R)$. Then, the following are true:*

1. *$\sum_{n=0}^{\infty} n a_n x^{n-1}$ converges.*
2. *f is differentiable on $(-R, R)$.*
3. *$f'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}$ on $(-R, R)$.*

Before we get to the bulk of the Theorem, we will make a few easy observations.

Recall that $R^{-1} = \overline{\lim}\, |a_n|^{\frac{1}{n}}$ for any power series. So, for the series in part (1) above, we have $R_1^{-1} = \overline{\lim}(n|a_n|)^{\frac{1}{n}}$. However, we can show that $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$, so by the limit theorems, $R = R_1$, so if $\sum a_n x^n$ converges, then $\sum n a_n x^{n-1}$ also converges.

Also, the following identity is true for any $a, b \in \mathbb{R}$ such that $a - b \neq 0$:

$$\frac{a^n - b^n}{a - b} = a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1} \tag{3}$$

Notice that the right hand side is an expression with n terms.

The following is the infinite analogue to the Triangle Inequality:

$$\left|\sum_{n=0}^{\infty} c_n\right| \leq \sum_{n=0}^{\infty} |c_n| \tag{4}$$

Finally, recall the very simple identity

$$1 + 2 + \cdots + n = \frac{n(n-1)}{2} \tag{5}$$

Now, we are ready to start.

Proof. We have already proved (1) in the first paragraph above, and proving (2) and (3) are basically equivalent.

In order to prove (2) and (3), it suffices to show that

$$\lim_{y\to x} \left|\sum_{n=0}^{\infty} \left(\frac{a_n(y^n - x^n)}{y - x} - n a_n x^{n-1}\right)\right| = 0 \tag{6}$$

If R is finite, then choose an r such that $0 < r < R$, $|x| < r < R$, and $|y| < r < R$. If R is infinite, then this step is not necessary, and we can simply pick any r.

Now, by (4), we have that (6) is equal to

$$\lim_{y\to x}\left|\sum_{n=0}^{\infty} a_n(y^{n-1}+y^{n-2}x+y^{n-3}x^2+\cdots+yx^{n-2}+x^{n-1}-nx^{n-1})\right|$$

Because (3) is n terms long, we can split up the last nx^{n-1} and distribute them to every other term to obtain:

$$\lim_{y\to x}\left|\sum_{n=0}^{\infty} a_n((y^{n-1}-x^{n-1})+x(y^{n-2}-x^{n-2})\right.$$
$$\left.+x^2(y^{n-3}-x^{n-3})+\cdots+x^{n-2}(y-x))\right|$$

Then, if we factor out a $y-x$ from each term and apply (3) to each one individually, we obtain

$$\lim_{y\to x}|y-x|\left|\sum_{n=0}^{\infty} a_n((y^{n-2}+y^{n-3}x+\cdots+x^{n-2})\right.$$
$$\left.+x(y^{n-3}+y^{n-4}x+\cdots+x^{n-3})+\cdots+x^{n-2})\right|$$

Notice that every term in the above expression is less than r^{n-2} because $|x| < r < R$ and $|y| < r < R$ and there are $n-1$ terms in the first parenthesized expression, $n-2$ in the second, and so on until there is only 1 term in the last expression. Therefore, by (5), the whole expression above is less than or equal to:

$$\lim_{y\to x}|y-x|\left|\sum_{n=0}^{\infty}|a_n|\frac{n(n-1)}{2}\,r^{n-2}\right|$$

Now, remember that the radius of convergence of $\sum_{n=0}^{\infty} b_n x^n$ is equal to B^{-1}, where $B = \sqrt[n]{|b_n|}$. It can be shown that $\sqrt[n]{\frac{n(n-1)}{2}}$ converges to 1, so the coefficients of this power series converge simply to $|a_n|$. Therefore, because $r < R$, the whole second term is a convergent power series, so the entire term can converges as y goes to x, completing the proof. $\square$

20.3. **Elementary Functions.**

Now, we can define all of these nice functions we've previously seen, and prove their familiar properties.

- $\exp(x) = \displaystyle\sum_{n=0}^{\infty}\frac{x^n}{n!}$
- $\cos(x) = \displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n\,x^{2n}}{(2n)!}$
- $\sin(x) = \displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n\,x^{2n+1}}{(2n+1)!}$

20.3.1. *The exponential function is never negative.* Suppose there exists an a such that $\exp(a) = 0$. We can assume $a < 0$ because if $a > 0$, then we can reach a contradiction immediately through direct calculation. Consider the real number A, defined to be $\sup\{a \mid \exp(a) = 0\}$. Clearly, $\exp(A) = 0$. Consider $\exp(A + x)$ for some $x \in \mathbb{R}$. It is easy to see by direct calculation that the derivative of $\exp(x)$ is $\exp(x)$. Therefore, by the Chain Rule, the derivative of $\exp(A + x) = \exp(A + x)$.

Now, consider the Taylor series expansion of $\exp(A + x)$, or

$$F(x) = \exp(A + x) = F(0) + F'(0)x + \cdots + \frac{F^{(n)}(y)\, x^n}{n!}$$

for some y between 0 and x. Notice that $\exp(A+x)$ is increasing by direct calculation, and every derivative is exactly $\exp(A + x)$. Therefore, $f^{(n)}(y) = \exp(A + x)$, so we have a bound on y, and the Taylor series converges to F. However, it is the case that $F(0) = \exp(A + 0) = \exp(A) = 0$, so plugging that in to the Taylor series, we get that $\exp(A + x) = 0$ for all x. This is clearly false, because $\exp(A + (-A)) = \exp(0) = 1$. Thus, we have a contradiction, and $\exp(x) \neq 0$.

20.3.2. *Adding real numbers in the* exp *function.* Consider the following equation.

$$\frac{d}{dx}\left(\frac{\exp(x + a)}{\exp(x)}\right) = \frac{\exp(x + a)\exp(x) - \exp(x + a)\exp(x)}{(\exp(x))^2} = 0$$

By a corollary to the Mean Value Theorem, this implies that $\frac{\exp(x+a)}{\exp(x)}$ is constant. Therefore, we have $\exp(x + a) = K \exp(x)$ for all $x \in \mathbb{R}$ and some $K \in \mathbb{R}$. Let $x = 0$. Then, $\exp(a) = K$, so $\exp(x + a) = \exp(a)\exp(x)$, a familiar identity.

20.3.3. *The limit as* exp *approaches zero.* $\exp(\mathbb{R})$ is connected because exp is a continuous function, and the range does not contain 0, so the function is always positive.

Now, choose $\epsilon > 0$ such that $\exp(-\epsilon) < 1$. Then, it is the case that, for any $n \in \mathbb{N}$,

$$\exp(-n\epsilon) = (\exp(-\epsilon))^n$$

which is arbitrarily small as n goes to ∞. Therefore, $\lim_{x\to-\infty} \exp(x) = 0$.

20.3.4. exp *is a bijection from* $\mathbb{R}$ *to* $(0, \infty)$. This follows immediately from the observation that exp is strictly increasing.

20.3.5. *The* log *function.* We can define $\log(x)$ as the inverse function of $\exp(x)$, where log has domain $(0, \infty)$. Therefore, $\log\exp(x) = x$, $\exp\log(x) = x$, and $\log(xy) = log(x) + log(y)$. Also, it is the case that the derivative of log is $\frac{1}{x}$ by the Chain Rule on $\log\exp(x)$.

20.3.6. *Arbitrary exponential functions and* e. We will define $a^r = \exp(r\log(a))$ for $a, r \in \mathbb{R}$ such that $a > 0$. Also, we will define the real number $e = \exp(1)$. Therefore, we can see that $e^x = \exp(x)$.

20.3.7. *The Pythagorean trigonometric identity.* Consider the following equation.

$$\frac{d}{dx}(cos^2(x) + sin^2(x)) = -2\cos(x)\sin(x) + 2\sin(x)\cos(x) = 0$$

Thus, as before, $\cos^2(x) + \sin^2(x)$ is a constant function. Plugging in 0, we get that $\cos^2(x) + \sin^2(x) = 1$.

20.3.8. *Pi.* Because $\cos(1) > 0$ and $\cos(2) < 0$ by the error of the alternating series function, it is the case that there exists an x between 1 and 2 such that $cos(x) = 0$. We will define $\frac{\pi}{2} = \inf\{x \mid x > 0, \cos(x) = 0\}$.

There are many more properties that we can prove at this stage, but they are left as exercises.

20.4. **Some Extra Tidbits.** [*]

This section includes some counterexamples and odd functions related to differentiability, as well as another theorem.

(1) $\lim_{x\to 0} \sin \frac{1}{x}$ does not exist.

(2) The function
$$f(x) = \begin{cases} x \sin \frac{1}{x}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$
is continuous at 0. However, it is not differentiable at 0, as
$$\lim_{\delta\to 0} \frac{f(\delta) - f(0)}{\delta} = \lim_{\delta\to 0} \frac{\delta \sin \frac{1}{\delta}}{\delta} = \lim_{\delta\to 0} \sin \frac{1}{\delta}$$
which does not exist.

(3) The function
$$g(x) = \begin{cases} x^2 \sin \frac{1}{x^2}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$
is everywhere differentiable, as if $x = 0$, then $\lim_{\delta\to 0} \frac{\delta^2 \sin \frac{1}{\delta}}{\delta} = 0$. However, consider its derivative. If $x \neq 0$, then $\frac{dg}{dx} = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$, and if $x = 0$, then $\frac{dg}{dx} = 0$, from the above calculation. Therefore, $g'(x)$ is not continuous at 0.

Theorem. *Let f be differentiable on (a, b), where $a < x_1 < x_2 < b$, and let $c \in \mathbb{R}$. If $f'(x_1) < c < f'(x_2)$, then there exists a y such that $x_1 < y < x_2$ and $f'(y) = c$.*

Proof. Consider the function $g(x) = f(x) - cx$. Clearly, g is differentiable on (a, b), and its derivative is $f'(x) - c$. $[x_1, x_2]$ is compact, so the restriction of g to $[x_1, x_2]$ attains its minimum and maximum. Therefore, for some $y \in [x_1, x_2]$, $g'(y) = 0$, or $f'(y) - c = 0$. If y was equal to either x_1 or x_2, then $f'(x_1) = c$ or $f'(x_2) = c$, which contradict our assumptions. Thus, y must not equal either of those points, and we are done. $\square$

[*]This section was originally included in lecture 24 by Professor Jones. However, I have moved it here so that the flow of the notes is not interrupted.

21. 04/07/2011: Uniform Convergence: Guest Lecture by Scott Morrison

21.1. **Introduction.**

The next two lectures will be about approximating continuous functions. We will try to answer two questions:

(1) When f_n, a sequence of continuous functions, converges to another function f, how can we tell if f is continuous?
(2) When can we approximate a continuous function by so-called "nice functions"? For instance, can we find a sequence of polynomials (f_n) such that $f_n \to f$?

We will answer the first question in this lecture, and the second question in the next.

21.2. **The Supremum Norm.**

We will start with a non-example.

Non-Example: Define $f_n\colon [0,1] \to [0,1]$ by

$$f_n(x) = \begin{cases} nx, & \text{if } x \in [0, \frac{1}{n}] \\ 1, & \text{if } x \in [\frac{1}{n}, 1] \end{cases}$$

We can see that $f_n(x)$ converges pointwise to the function

$$f(x) = \begin{cases} 0, & \text{if } x = 0 \\ 1, & \text{if } x \in (0, 1] \end{cases}$$

which is clearly not continuous. So what went wrong?

The answer is that as x approaches 0, $f_n(x)$ takes longer and longer to converge to $f(x)$.

To see this more clearly, take $\epsilon = \frac{1}{2}$. What N ensures that $|f_n(x) - f(x)| < \epsilon$ for all $n \geq N$? Well...

$$|f_n(x) - f(x)| < \frac{1}{2} \Rightarrow \frac{1}{n} < 2x$$

So, we can take $N = \lceil \frac{1}{2x} \rceil$, but that's the best we can do. As $x \to 0$, $N \to \infty$.

To avoid this problem, we will define uniform convergence. However, first, let's define the supremum norm.

Definition. *Let $f\colon X \to \mathbb{R}$ be a function. We will define a norm on the set of all functions from X to $\mathbb{R}$, namely*

$$||f||_\infty = \sup_{x \in X} |f(x)|$$

This notion can be extended to a metric on the functions from X to $\mathbb{R}$, namely

$$d(f, g) = ||f - g||_\infty = \sup_{x \in X} |f(x) - g(x)|$$

Of course, we can generalize this even further by simply replacing the $\mathbb{R}$ with any metric space C, and the familiar metric on $\mathbb{R}$ with a metric on C.

For functions from $\mathbb{R}$ to $\mathbb{R}$, $d(f, g) < \epsilon$ means that g never gets more than ϵ away from f.

Definition. *A sequence of functions f_n* ***converges uniformly*** *if it converges with respect to the metric defined above, i.e.*

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) \text{ s.t. } (\forall n \geq N),\ d(f_n, f) < \epsilon$$

Notice that the statement

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) \text{ s.t. } (\forall n \geq N) \text{ and } (\forall x \in X),\ D(f_n(x), f(x)) < \epsilon$$

is equivalent to the definition above, where D is the metric on the codomain of f and f_n. However, if we find such an ϵ for the latter statement and apply it to the former, we must change the condition so that $d(f_n, f) \leq \epsilon$, because the d in the first statement takes the supremum of values of $f_n(x)$ and $f(x)$.

So what went wrong in our first example? We had an N for each x, but we could not find an N which satisfied the condition for *all* x.

Theorem. *If (f_n) is a sequence of continuous functions converging uniformly to a function f, then f is continuous.*

Proof. Let $x_0 \in X$ and $\epsilon > 0$ be given. We want to find some $\delta > 0$ such that for all x with $D_1(x, x_0) < \delta$, $D_2(f(x), f(x_0)) < \epsilon$, where D_1 is the metric on the domain of f_n and D_2 is the metric on the range of f_n.

Notice that

$$D_2(f(x), f(x_0)) \leq D_2(f(x), f_n(x)) + D_2(f_n(x), f_n(x_0)) + D_2(f_n(x_0), f(x_0))$$

for all n. Because f_n converges to f uniformly, we can find an N such that $D_2(f_N(y), f(y)) < \frac{\epsilon}{3}$ for all y. Also, because f_N is continuous, we can choose a $\delta > 0$ such that $D_2(f_N(x), f_N(x_0)) < \frac{\epsilon}{3}$ as long as $D_1(x, x_0) < \delta$. Using this δ, we have:

$$D_2(f(x), f(x_0)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

completing the proof. □

Exercise: Can you find (f_n) which are uniformly continuous and converge uniformly to f, where f is *not* uniformly continuous? If not, can you improve this proof to show that if every f_n is continuous, then f must be uniformly continuous?

21.3. **An Explicit Construction.**

Let's now use these concepts to construct a continuous surjective function from the Cantor set to $[0, 1]$. Note that by varying this construction, you can actually map the Cantor set to any compact metric space, like $[0, 1] \times [0, 1]$.

The basic idea will be to construct a sequence of functions $f_n \colon [0, 1] \to [0, 1]$ which are "approximately surjective" when restricted to C, the Cantor set, and show that they converge uniformly to some f which must be continuous. We will then use this continuity to show that f is actually surjective when restricted to C.

Recall that $C = M_0 \cap M_1 \cap M_2 \cap \dots$ where $M_0 = [0, 1]$, $M_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, and so on, such that

$$M_k = \frac{1}{3} M_{k-1} \cup \left(\frac{2}{3} + \frac{1}{3} M_{k-1} \right)$$

We will define f_n such that $f_n|_{M_n}: M_n \to [0,1]$ is surjective, or in other words, such that f_n when restricted to M_n is surjective for each n. Let $f_0: [0,1] \to [0,1]$ be any continuous function with $f_0(0) = 0$ and $f_0(1) = 1$. Then, define

$$f_1(x) = \begin{cases} \frac{1}{2} f_0(3x), & \text{if } x \in [0, \frac{1}{3}] \\ \frac{1}{2}, & \text{if } x \in [\frac{1}{3}, \frac{2}{3}] \\ \frac{1}{2} + \frac{1}{2} f_0(3x-2), & \text{if } x \in [\frac{2}{3}, 1] \end{cases}$$

$f_1|_{M_1}$ is surjective, as $f_1|_{[0,\frac{1}{3}]} = \frac{1}{2} f_0$ is surjective onto $[0, \frac{1}{2}]$ and $f_1|_{[\frac{2}{3},1]} = \frac{1}{2} + \frac{1}{2} f_0$ is surjective on $[\frac{1}{2}, 1]$. Therefore, extending the construction, we get, for any $k > 0$:

$$f_k(x) = \begin{cases} \frac{1}{2} f_{k-1}(3x), & \text{if } x \in [0, \frac{1}{3}] \\ \frac{1}{2}, & \text{if } x \in [\frac{1}{3}, \frac{2}{3}] \\ \frac{1}{2} + \frac{1}{2} f_{k-1}(3x-2), & \text{if } x \in [\frac{2}{3}, 1] \end{cases}$$

Now, we must check that (f_n) converges pointwise. To do this, we will need two Propositions.

Proposition. *$|f_n(x) - f_{n-1}(x)| \leq 2^{-n+1}$ for all $x \in [0,1]$.*

Proof. This is clearly true for $n = 1$, as $f_1(x) - f_0(x)| \leq 1$. Now, we will induct on n. Then, we have

$$|f_n(x) - f_{n-1}(x)| = \begin{cases} \frac{1}{2}|f_{n-1}(3x) - f_{n-2}(3x)|, & \text{if } x \in [0, \frac{1}{3}] \\ 0, & \text{if } x \in [\frac{1}{3}, \frac{2}{3}] \\ \frac{1}{2}|f_{n-1}(3x-2) - f_{n-2}(3x-2)|, & \text{if } x \in [\frac{2}{3}, 1] \end{cases}$$

Then, $|f_n(x) - f_{n-1}(x)| \leq \frac{1}{2} \cdot 2^{-n+2} = 2^{-n+1}$, so we are done. □

From this Proposition, we can see that (f_n) is Cauchy for each x, so there exists a pointwise limit, which we will call f.

Proposition. *$|f_n(x) - f(x)| \leq 2^{-n+2}$ for all $x \in [0,1]$.*

Proof. First, we can see that, for $n, m \in \mathbb{N}$ and $n < m$:

$$\begin{aligned} |f_n(x) - f_m(x)| \leq |f_n(x) - f_{n+1}(x)| + |f_{n+1}(x) - f_{n+2}(x)| \\ + \cdots + |f_{m-1}(x) + f_m(x)| \leq 2^{-n+2} \end{aligned}$$

by the previous Proposition. Therefore, by the geometric series formula, we have

$$|f_n(x) - f(x)| = \lim_{m \to \infty} |f_n(x) - f_m(x)| \leq 2^{-n+2}$$

which completes the proof. □

Therefore, because 2^{-n+2} can get arbitrarily small as n increases, (f_n) converges uniformly to f, and f is continuous by the previous Theorem. Because f is a continuous function from $[0,1]$ to $[0,1]$ and clearly $f(0) = 0$ and $f(1) = 1$, by the Mean Value Theorem, f is surjective on $[0,1]$.

So, all that is left is to show that the restriction of f to C is surjective. Take $x \in [0,1]$ and pick a $y \in [0,1]$ so that $f(y) = x$. If $y \in C$, then we're done. Otherwise, $y \in [0,1] \setminus M_k$ for some k. However, $[0,1] \setminus M_k$ is simply a collection of open intervals. Therefore, we will make a claim.

Proposition. *For all $n \geq k$, f_n is constant and independent of n on the closure of each interval of $[0,1] \setminus M_k$.*

Proof. f_1 is constant on $\overline{[0,1] \setminus M_1} = [\frac{1}{3}, \frac{2}{3}]$. If f_{n-1} is constant on $[a, b]$, then f_n is constant on $[\frac{1}{3}a, \frac{1}{3}b]$ and $[\frac{2}{3} + \frac{1}{3}a, \frac{2}{3} + \frac{1}{3}b]$, because $M_k = \frac{1}{3}M_{k-1} \cup (\frac{2}{3} + \frac{1}{3}M_{k-1})$, so

$$[0,1] \setminus M_k = \frac{1}{3}([0,1] \setminus M_{k-1}) \cup \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{2}{3} + \frac{1}{3}([0,1] \setminus M_{k-1})\right)$$

□

Therefore, by this proposition, we can pick y' such that y' is an endpoint of the interval which y belongs to, so $f(y') = x$. However, the endpoint of any such interval is in the Cantor set, so f is surjective on C, and we are done.

22. 04/12/2011: Weierstrass Approximation: Guest Lecture by Scott Morrison

22.1. Banach Spaces.

Recall the supremum norm on the functions from $[0,1]$ to $\mathbb{R}$, namely

$$||f|| = \sup_{x\in[0,1]} |f(x)|$$

and the associated metric defined by

$$d(f,g) = ||f-g|| = \sup_{x\in[0,1]} |f(x)-g(x)|$$

as well as the generalization of this norm into $C(X)$, the set of continuous functions on X:

$$d(f,g) = \sup_{x\in X} d(f(x),g(x))$$

Definition. *A vector space with a norm such that the vector space is complete is called a **Banach space**.*

Theorem. *$C([0,1])$, the set of continuous functions from $[0,1]$ to $\mathbb{R}$, is a Banach space; in particular, the supremum norm makes this space complete.*

Proof. Suppose (f_n) is a Cauchy sequence of continuous functions. Certainly, f_n has a pointwise limit f because of the completeness of $\mathbb{R}$. In other words, we know that for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $n,m \geq N$ implies

$$\sup_{x\in[0,1]} |f_n(x) - f_m(x)| < \epsilon$$

Then, for any given x, $|f_n(x) - f_m(x)| < \epsilon$, so $(f_n(x))$, a sequence in $\mathbb{R}$, is Cauchy and it converges to some number which we will call $f(x)$. So, let's check that (f_n) actually converges to this f uniformly. Because this sequence is Cauchy, for any $\epsilon > 0$, we can choose N such that $||f_n - f_m|| < \frac{\epsilon}{2}$ for all $n,m \geq N$. This means that $|f_n(x) - f_m(x)| < \frac{\epsilon}{2}$ for all $x \in [0,1]$. Taking the limit as m goes to infinity, namely $\lim_{m\to\infty}(f_m(x))$, we get $|f_n(x) - f(x)| \leq \frac{\epsilon}{2} < \epsilon$ for all $x \in [0,1]$. Therefore, $||f_n - f|| < \epsilon$, so (f_n) converges uniformly to f. By the theorem in the last lecture, this f is continuous, so $f \in C([0,1])$, and we are done. □

22.2. Subalgebras.

Definition. *$A \subseteq C([0,1])$ is a **subalgebra** of $C([0,1])$ if it is closed under addition, multiplication, and scalar multiplication.*

Example: The polynomials are a subalgebra of $C([0,1])$.

Example: Linear combinations of the trigonometric functions, namely

$$\{\cos(\pi nx) \,|\, 0 \leq n < \infty\} \cup \{\sin(\pi nx) \,|\, 0 \leq n < \infty\}$$

are also a subalgebra of $C([0,1])$.

The main exploratory question for this lecture is "When is a subalgebra of $C([0,1])$ dense?" There is a Theorem that tells us exactly that, as well as an easier corollary that can be proved independently. We will start, however, with some definitions.

Definition. *If X is a metric space, $Y \subseteq X$ is **dense** in X if for all $x \in X$, there exists a sequence $(y_n) \subseteq Y$ such that (y_n) converges to x. Equivalently, $Y \subseteq X$ is dense in X if the union of Y and the set of all of its limit points is all of X. Given $Y \subseteq X$, the **closure** of Y, denoted $\overline{Y}$, is the union of Y and the set of all of its limit points in X.*

Definition. *Let X be a metric space. A subalgebra $A \subseteq C(X)$ **separates points** if for all distinct $x, y \in X$, there exists a function $f \in A$ such that $f(x) \neq f(y)$.*

The following two theorems are our main results.

Theorem (Weierstrass Approximation)**.** *The polynomials are dense in $C([0,1])$, i.e. every continuous function on $[0,1]$ is the uniform limit of a sequence of polynomials.*

Theorem (Stone-Weierstrass)**.** *If X is a compact metric space and $A \subseteq C(X)$ separates points and contains the function $\mathbb{1}$ where $\mathbb{1}(x) = 1$ for all x, then A is dense in $C(X)$.*

Notice that the Stone-Weierstrass Theorem implies the Weierstrass Approximation Theorem, because $[0,1]$ is compact and the function $f(x) = x$ separates points.

Proposition. *If $A \subseteq C(X)$ is a subalgebra which does not separate points, then A is not dense in $C(X)$.*

Proof. It is certainly the case that $C(X)$ separates points, as given distinct $x, y \in X$, we can take $f(z) = d(x, z)$ where d is any metric on X. Then, f is continuous, and $f(x) = 0$ while $f(y) > 0$.

Assume toward contradiction that $A \subset C(X)$ does not separate points. Then, in particular, for some $x, y \in X$, $g(x) = g(y)$ for all $g \in A$. We showed above that there exists a function $f \in C(X)$ such that $d(f(x), f(y)) = L > 0$. We will show that this f is not a uniform limit of functions in A if (f_n) is a sequence in A.

$$||f_n - f|| = \sup_{z \in X} |f_n(z) - f(z)| \geq \max\{|f_n(x) - f(x)|, |f_n(y) - f(y)| \tag{7}$$

$$\geq \frac{1}{2}(|f_n(x) - f(x)| + |f_n(y) - f(y)|) \tag{8}$$

$$\geq \frac{1}{2}|f(x) + f(y)| \tag{9}$$

$$\geq \frac{1}{2}|f(x) - f(y)| = \frac{L}{2} \tag{10}$$

In the above, (9) follows from the equality of $f_n(x)$ and $f_n(y)$ as well as the triangle inequality, and (10) follows from the positive definiteness of $f = d$. Therefore, it is clear that (f_n) cannot converge to f, and A is not dense in $C(X)$. □

Therefore, separating points is a necessary and sufficient condition on A in order for it to be dense in $C(X)$. Now, we will begin the task of proving the Weierstrass Approximation Theorem.

22.3. **Proof of the Weierstrass Approximation Theorem.**

This proof will consist of four propositions.

Proposition. *$f(x) = 1 - \sqrt{1-x}$ can be uniformly approximated by polynomials.*

Proof. Define a sequence of polynomials as such: let $P_0(x) = 0$ and define P_{n+1} recursively as $P_{n+1}(x) = \frac{1}{2}(P_n(x)^2 + x)$ for all $n \in \mathbb{N}$. These are all clearly polynomials. Also, observe that $0 \leq P_n(x) \leq 1$ for $x \in [0,1]$. To see this, we can simply proceed by induction on n. Also, it is the case that

$$P_{n+2}(x) - P_{n+1}(x) = \frac{1}{2}(P_{n+1}(x) + P_n(x))(P_{n+1}(x) - P_n(x))$$

From this, we can see that $P_{n+1}(x) \geq P_n(x)$, which again can be proved by induction. From these observations, it is clear that $(P_n(x))$ is bounded above and monotonic, and therefore converges to some number in $\mathbb{R}$ for each x, which we will denote as $g(x)$.

Consider that

$$\lim_{n\to\infty}(P_{n+1}(x) - \frac{1}{2}(P_n(x)^2 + x)) = 0$$

by construction. Therefore, $g(x) - \frac{1}{2}(g(x)^2 + x) = 0$, and solving this equation for $g(x)$ allows us to see that $g(x) = f(x) = 1 - \sqrt{1-x}$ for all $x \in [0,1]$.

Therefore, we now have a sequence of polynomials converging pointwise to $f(x)$, so it remains to show that it, in fact, converges uniformly.

As mentioned above, $(P_n(x))$ is a nondecreasing sequence in n. However, it can also be seen that $(P_n(x))$ is a nondecreasing sequence in x. Therefore, because the function x is also nondecreasing, $\frac{1}{2}(P_n(x)^2 + x) = P_{n+1}(x)$ is nondecreasing. Finally, from this, we can see that $P_{n+1}(x) - P_n(x)$ is also nondecreasing in x, a fact that can be proved by induction (again). Thus, we have that

$$P_{n+1}(x) - P_n(x) \leq P_{n+1}(1) + P_n(1)$$

and because this is a telescoping series, this implies that

$$P_m(x) - P_n(x) \leq P_m(1) - P_n(1)$$

for $m > n$. Now, taking the limit as m goes to infinity, we get

$$f(x) - P_n(x) \leq f(1) - P_n(1)$$

which is equivalent to the statement

$$||f - P_n|| \leq f(1) - P_n(1)$$

because $P_{n+1}(x) - P_n(x)$ is nondecreasing in x. Because $(P_n(1))$ converges to $f(1)$, we can make the right hand side arbitrarily small. Therefore, P_n converges uniformly in the supremum norm to f. □

Proposition. *The function $f(x) = |x|$ can be uniformly approximated by polynomials.*

Proof. Notice that

$$\epsilon = x^2 + \epsilon - |x|^2 = (\sqrt{x^2+\epsilon} - |x|)(\sqrt{x^2+\epsilon} + |x|)$$

Then, this implies that

$$\sqrt{x^2+\epsilon} - |x| = \frac{\epsilon}{\sqrt{x^2+\epsilon} + |x|} \leq \frac{\epsilon}{\sqrt{\epsilon}} = \epsilon^{\frac{1}{2}}$$

Now, if $1-\sqrt{1-x}$ can be uniformly approximated by polynomials, then $\sqrt{1-x}$, $\sqrt{x}$, and $\sqrt{x^2+\epsilon}$ can also be uniformly approximated by polynomials by simple manipulations. So, for each $n \in \mathbb{N}$, choose polynomials $(f_{n,m})$ such that $(f_{n,m})$ converges to $\sqrt{x^2+\frac{1}{n}}$ as m goes to ∞.*

Define $g_n = f_{n,N}$ where N is big enough so that $||f_{n,N} - \sqrt{x^2+\frac{1}{n}}|| < \frac{1}{n}$. We will claim that g_n converges uniformly to $|x|$, as:

$$||g_n - |x||| \leq ||g_n - \sqrt{x^2+\frac{1}{n}}|| + ||\sqrt{x^2+\frac{1}{n}} - |x||| \leq \frac{1}{n} + \frac{1}{\sqrt{n}}$$

where the last inequality follows from the assumption above as well as the very first series of inequalities. Therefore, because $\frac{1}{n} + \frac{1}{\sqrt{n}}$ obviously goes to 0 as n goes to infinity, g_n is a sequence of polynomials that converges uniformly to $|x|$. □

Proposition. *Every piecewise linear function is a uniform limit of polynomials.*

Proof Sketch. You can write any piecewise linear function in the form

$$f(x) = c + \sum_{n=1}^{k} a_n|x-b_n|$$

where $c \in \mathbb{R}$, $k \in \mathbb{N}$, $a_n, b_n \in \mathbb{R}$. Then, we can apply the previous Proposition to obtain the desired result. □

Proposition. *Every continuous function is a uniform limit of piecewise linear functions.*

Proof Sketch. Every continuous function on $[0,1]$ is uniformly continuous, because $[0,1]$ is compact. Thus, choose $N \in \mathbb{N}$ such that $|f(x)-f(y)| < \epsilon$ if $|x-y| \leq \frac{1}{N}$. Now, define a function g to be the piecewise linear function which agrees with f at the points $\frac{k}{N}$ where $0 \leq k \leq N$, and is linear in between. Then, as N goes to infinity, g converges to f. □

This proposition proves the Theorem.

*Note that what we're proving here is that if (x_n) is a sequence of limit points of X that converges, then the limit point of that sequence is also a limit point of X. It may actually be better to prove this fact independently.

23. 04/14/2011: The Stone-Weierstrass Theorem

23.1. **Review.**

Theorem (Stone-Weierstrass). *Let (X, d) be a compact metric space and A be a subalgebra of $C(X, \mathbb{R})$. If*

(1) *A contains the constant function $\mathbb{1}$, where $\mathbb{1}(x) = 1$ for all $x \in X$, and*
(2) *A separates the points of X, namely for all $x \neq y$ in X, there exists a function $f \in A$ such that $f(x) \neq f(y)$,*

then $\overline{A} = C(X, \mathbb{R})$, i.e. A is dense in $C(X, \mathbb{R})$.

A review of the relevant concepts follows.

Remember that for any continuous functions $f, g\colon X \to \mathbb{R}$, we define:

- $f + g = (f + g)(x) = f(x) + g(x)$
- $fg = (fg)(x) = f(x)g(x)$
- $\lambda f = (\lambda f)(x) = \lambda(f(x))$

An algebra has various properties with the above operations. A **subalgebra** is a subset of this algebra that is closed under these operations.

Example: Let $X = [0, 1]$. Then, $\{f \mid f(0) = f(1),\ f \text{ continuous}\}$ is a subalgebra of the continuous functions from $[0, 1]$ to $\mathbb{R}$, denoted $C([0, 1], \mathbb{R})$.

Example: Likewise,

$$\sum_{n=0}^{k} a_n x^n$$

where $x \in [0, 1]$ and $a_n \in \mathbb{R}$ is a subalgebra of $C([0, 1], \mathbb{R})$.

Also, remember that on $C(X, \mathbb{R})$, there is a norm called the **supremum norm** defined by

$$||f|| = \sup\{|f(x)| \mid x \in X\}$$

which makes it a complete metric space. A complete normed metric space is called a **Banach space**.

23.2. **Proof of the Stone-Weierstrass Theorem.**

Remember from last lecture that in $C([0, 1])^*$, $\sqrt{x}$ is a uniform limit of polynomials. From this, we can get the following result.

Proposition. *If $f \in A$, where A is a subalgebra of $C([0, 1])$, then $|f| \in \overline{A}$.*

Proof. Notice that $|f| = \sqrt{f^2}$, and that because A is a subalgebra of $C([0, 1])$ (and hence closed under multiplication), f^2 is in A. Therefore, it suffices to show that given any function $g \in A$ such that $g(x) \geq 0$ for all $x \in X$, $\sqrt{g}$ is in $\overline{A}$.

Rephrasing the last statement, we want to show that for all $\epsilon > 0$, there exists an $h \in A$ such that $|h(x) - \sqrt{g(x)}| < \epsilon$ for all $x \in X$. Without loss of generality, because A is closed under scalar multiplication, we can assume that $||g|| \leq 1$. Now, from the statement proved last time, namely that $\sqrt{x}$ is a uniform limit of

[*] This is the same as $C([0, 1], \mathbb{R})$.

polynomials, there exists a polynomial p on $[0,1]$ such that $|p(r)-\sqrt{r}|<\epsilon$ for all $r\in[0,1]$. Therefore, we can simply replace r with $g(x)$ in order to obtain

$$|p(g(x))-\sqrt{g(x)}|<\epsilon$$

for all $x\in X$. However, $p(f(x))=(p(f))(x)$, so we can simply set h to be $p(f)$, completing the proof. □

Corollary. *If f and g are functions in A, then $M_{f,g}(x)$ and $m_{f,g}(x)$ are in $\overline{A}$, where*

$$M_{f,g}(x):=\max\{f(x),g(x)\}$$
$$m_{f,g}(x):=\min\{f(x),g(x)\}$$

Proof. This is immediate from the previous proposition, as

$$M_{f,g}(x)=\frac{1}{2}(f(x)+g(x)+|f(x)-g(x)|)$$
$$m_{f,g}(x)=\frac{1}{2}(f(x)+g(x)-|f(x)-g(x)|)$$

and A is closed under addition. □

Notice that if A is a subalgebra of $C(X,\mathbb{R})$, then $\overline{A}$ is also a subalgebra of $C(X,\mathbb{R})$. Also, the above Corollary can be generalized to any number of finite functions by induction.

Proposition. *Suppose A is a subalgebra of $C(X,\mathbb{R})$ that separates points and contains the constant functions. Then, for every $f\in C(X,\mathbb{R})$, $x\in X$, and $\epsilon>0$, there exists a function $g_x\in\overline{A}$ such that $g_x(z)<f(z)+\epsilon$ for all $z\in X$ and $g_x(x)>f(x)$.*

Proof. By hypothesis, because A separates points, for every $y\neq x$, there is some function $\overline{h}_{x,y}$ such that $\overline{h}_{x,y}(x)\neq\overline{h}_{x,y}(y)$. Now, because the constant function is in A and it is a subalgebra, we can choose constants a and b such that $h_{x,y}(z)=a\overline{h}(z)+b$ and the following equations hold:

$$h_{x,y}(x)=f(x)+\frac{\epsilon}{2}$$
$$h_{x,y}(y)=f(y)-\frac{\epsilon}{2}$$

Now, for every $y\neq x$, define the set

$$U_{x,y}=\{z\in X\,|\,h_{x,y}(z)<f(z)+\epsilon\}$$

The continuity of $h_{x,y}$ and f make $U_{x,y}$ an open set, as given any $z\in U_{x,y}$, one only needs to adjust z by some δ in order to ensure that $z+\delta$ is also in $U_{x,y}$. By construction of $h_{x,y}$, x and y are both in $U_{x,y}$. Therefore, the set of all $U_{x,y}$ where y varies in $X\setminus\{x\}$, or $\{U_{x,y}\,|\,y\in X\setminus\{x\}\}$, is an open cover of X. Because X is compact, there must be a finite subcover $\{y_1,y_2,\ldots,y_n\}$.

Define the function $g_x=\min\{h_{x,y_1},h_{x,y_2},\ldots,h_{x,y_n}\}$. By the above corollary, $g_x\in\overline{A}$. Now, we have that $g_x(x)=f(x)+\frac{\epsilon}{2}$ which is greater than $f(x)$. Also, given any $z\in X$, it must belong to some U_{x,y_i}, so $g_x(z)\leq h_{x,y_i}(z)<f(z)+\epsilon$, completing the proof. □

Proof of Stone-Weierstrass. For this proof, we will show that for every $f \in C(X, \mathbb{R})$ and $\epsilon > 0$, there exists a function $g \in \overline{A}$ such that $f(z) < g(z) < f(z) + \epsilon$. This is sufficient because this condition is equivalent to A being dense in $C(X, \mathbb{R})$; more informally, we are showing that for any function in $C(X, \mathbb{R})$, there is a function in $\overline{A}$ that is arbitrarily close to it.

First, for every $x \in X$, define the set

$$V_x = \{z \in X \mid g_x(z) > f(z)\}$$

where g_x is defined as in the previous proof. Again, because g_x and f are both continuous, V_x is an open set. Furthermore, $g_x(x) = f(x) + \frac{\epsilon}{2}$ which is greater than $f(x)$, so $x \in V_x$. Therefore, $\{V_x \mid x \in X\}$ is an open cover of X. Now, because X is compact, we can extract a finite subcover $\{x_1, x_2, \ldots, x_n\}$.

Define the function $g = \max\{g_{x_1}, g_{x_2}, \ldots, g_{x_n}\}$. From the corollary above, we know that $g \in \overline{A}$. Now, given any $z \in X$, it is clear from the construction of each g_x that $g(z) < f(z) + \epsilon$. Furthermore, every z is in some V_{x_i}, so $g(z) > f(z)$. Therefore, we have that $f(z) < g(z) < f(z) + \epsilon$, and we are done! $\square$

24. 04/19/2011: Integration Part One

24.1. A Review of Stone-Weierstrass.

The theorem states that given $C(X, \mathbb{R})$, the supremum norm, and a subalgebra A (namely $1 \in A$ and A separates the points of X), A is dense in $C(X, \mathbb{R})$, or $\overline{A} = C(X, \mathbb{R})$.
Example: The polynomials are dense in $C([a, b], \mathbb{R})$, as the function x separates points. Likewise, the polynomials in two variables are dense in $C([a, b] \times [c, d], \mathbb{R})$, because given (p_1, q_1) and (p_2, q_2), the function $f(x, y) = (x - p_1)^2 + (y - p_2)^2$ separates the two points. It is easy to see that this can be extrapolated to $\mathbb{R}^n$.

24.2. Darboux Integration.

Definition. *Let $[a, b]$ be an interval in $\mathbb{R}$. A **partition** of $[a, b]$ is a subset of $[a, b]$, $\{t_0, t_1, \ldots, t_n\}$, such that $t_1 = a$, $t_n = b$, and for all i where $0 \le i \le n-1$, $t_i < t_{i+1}$.*

Definition. *Let $f \colon [a, b] \to \mathbb{R}$ be a bounded function. Given $S \subseteq [a, b]$, we will define the following values:*

$$M(f, S) = \sup_{s \in S} f(s)$$
$$m(f, S) = \inf_{s \in S} f(s)$$

Definition. *Let P be a partition. The **upper Darboux sum** is defined as*

$$U(f, P) = \sum_{i=1}^{n} M(f, [t_{i-1}, t_i])(t_i - t_{i-1})$$

*The **lower Darboux sum** is defined as*

$$L(f, P) = \sum_{i=1}^{n} m(f, [t_{i-1}, t_i])(t_i - t_{i-1})$$

Clearly, it is the case that

$$m(f, [a, b])(b - a) \le L(f, P) \le U(f, P) \le M(f, [a, b])(b - a)$$

for any P. This motivates our next definition.

Definition.

$$U(f) = \inf_{P} U(f, P)$$
$$L(f) = \sup_{P} L(f, P)$$

In essence, $U(f)$ is the infimum of $U(f, P)$ as you vary the partition you are taking. Now, it is clear what our definition of the integral will be.

Definition. *$f \colon [a, b] \to \mathbb{R}$ is **Darboux integrable** if $L(f) = U(f)$, and we denote the integral as $\int_a^b f(x)dx = U(f)$.*

Example: Consider a function $f\colon [0,1] \to \mathbb{R}$ such that $f(x) = 0$ if x is irrational and $f(x) = 1$ if x is rational. Then, clearly $M(f, [t_{i-1}, t_i]) = 1$ and $m(f, [t_{i-1}, t_i]) = 0$ for any partition P. Therefore, we can compute the following:

$$\begin{aligned} U(f,P) &= \sum_{i=1}^{n} 1(t_i - t_{i-1}) = 1 \\ L(f,P) &= 0 \end{aligned}$$

This shows that f is not Darboux integrable.

Example: Now, consider the more familiar function $f(x) = x$ on $[0, b]$. For ease of working with the partitions, we will assume that our partition divides the interval into segments of equal width, or in other words, let $t_i = \frac{ib}{n}$. Then, clearly $M(f, [t_{i-1}, t_i]) = \frac{ib}{n}$ and $m(f, [t_{i-1}, t_i]) = \frac{(i-1)b}{n}$. Therefore,

$$\begin{aligned} U(f,P) &= \sum_{i=1}^{n} \frac{ib}{n} \cdot \frac{b}{n} = \frac{b^2}{n^2} \cdot \frac{n(n+1)}{2} \\ L(f,P) &= \sum_{i=1}^{n} \frac{(i-1)b}{n} \cdot \frac{b}{n} = \frac{b^2}{n^2} \cdot \frac{n(n-1)}{2} \end{aligned}$$

Both $\frac{b^2 n(n+1)}{2n^2}$ and $\frac{b^2 n(n-1)}{2n^2}$ tend to $\frac{b^2}{2}$, so $\inf U(f,P) \leq \frac{b^2}{2}$ and $\sup L(f,P) \geq \frac{b^2}{2}$. However, we cannot conclude that the function is Darboux integrable until we know for sure that $L(f)$ is always less than or equal to $U(f)$. In fact, this is our next theorem.

Theorem. *Let $f\colon [a,b] \to \mathbb{R}$ be bounded. $L(f) \leq U(f)$.*

Proof. Let P and Q be partitions of $[a,b]$ such that $P \subseteq Q$. First, we will show that

$$L(f,P) \leq L(f,Q) \leq U(f,Q) \leq U(f,P) \tag{11}$$

We will assume that $|Q| = |P| + 1$, as the full argument can be shown through induction. Therefore, we have

$$\begin{aligned} P &= \{t_0, t_1, \ldots, t_{i-1}, t_i, \ldots, t_n\} \\ Q &= \{t_0, t_1, \ldots, t_{i-1}, u, t_i, \ldots, t_n\} \end{aligned}$$

The middle inequality is trivial, and the first and third inequalities can be proved similarly, so we will only prove the first inequality, namely $L(f,P) \leq L(f,Q)$. Notice that

$$\begin{aligned} L(f,Q) - L(f,P) &= m(f,[t_{i-1},u])(u - t_i) + m(f,[u,t_i])(t_i - u) \\ &\quad -m(f,[t_{i-1},t_i])(t_i - t_{i-1}) \\ &\geq m(f,[t_{i-1},t_i])(u - t_{i-1}) + m(f,[t_{i-1},t_i])(t_i - u) \\ &\quad -m(f,[t_{i-1},t_i])(t_i - t_{i-1}) = 0 \end{aligned}$$

Therefore, for any P, Q such that $P \subseteq Q$, inequality (1) holds. Now, consider $P \cup Q$, where P and Q are arbitrary partitions. Applying (1) twice, we obtain

$$L(f,P) \leq L(f,P \cup Q) \leq U(f,P \cup Q) \leq U(f,Q)$$

Thus, we now know that $L(f,P) \leq U(f,Q)$ for any P and Q. If we fix any partition

P, we know that $L(f,P) \leq U(f)$, because $L(f,P)$ is less than or equal to $U(f,Q)$ for any Q. However, this also implies that $U(f)$ is greater than or equal to $L(f,P)$ for any P, because our fixed P was arbitrary. Therefore, $L(f) \leq U(f)$, and we are done. $\square$

25. 04/21/11: Integration Part Two

25.1. **More Darboux Integration.**

Recall that last time, we discussed the Darboux integral. Given a partition P and a bounded function f on $[a,b]$, we have:

$$U(f,P) = \sum M(f,[t_{i-1},t_i])(t_i - t_{i-1})$$
$$L(f,P) = \sum m(f,[t_{i-1},t_i])(t_i - t_{i-1})$$
$$U(f) = \inf_P U(f,P)$$
$$L(f) = \sup_P L(f,P)$$
$$\int_a^b f(x)dx = U(f) = L(f)$$

where the final equation only makes sense when $U(f) = L(f)$.

Also, last time, we proved that $L(f,P) \leq U(f,P)$ and $L(f) \leq U(f)$, where the crucial lemma was that if we have two partitions P and Q where $P \subseteq Q$, then

$$L(f,P) \leq L(f,Q) \leq U(f,Q) \leq U(f,P)$$

Theorem. *Let f be bounded on $[a,b]$. f is Darboux integrable if and only if for all $\epsilon > 0$, there exists a partition P such that $L(f,P) > U(f,P) - \epsilon$.*

Proof. We will go in both directions.

($\Leftarrow$) Due to the theorem in the previous lecture, in order to show that f is Darboux integrable, we only need to show that $L(f) \geq U(f)$. Let $\epsilon > 0$ be given, and let P be a partition such that the inequality in the theorem holds. Then, notice that

$$L(f) \geq L(f,P) > U(f,P) - \epsilon \geq U(f) - \epsilon$$

Because ϵ was arbitrary, $L(f) \geq U(f)$, so f is Darboux integrable.

($\Rightarrow$) Let $\epsilon > 0$ be given. Then, because f is Darboux integrable, there exists a partition P such that $U(f,P) < U(f) + \frac{\epsilon}{2}$ and a partition Q such that $L(f,Q) > L(f) - \frac{\epsilon}{2}$. To see this, assume that this was not the case. Then, there would be some $\frac{\epsilon}{2}$ such that for all partitions P, $U(f,P) \geq U(f) + \frac{\epsilon}{2}$. However, this would imply that there is some $U_2(f)$, namely $U(f) + \frac{\epsilon}{2}$, that is greater than $U(f)$ which is still a lower bound for the set of all $U(f,P)$. This is a contradiction, because $U(f)$ is the greatest lower bound of $\{U(f,P)\}$.

Now, consider $P \cup Q$. By a result proved in the last lecture, we know that $U(f,P) \geq U(f,P\cup Q)$ and $L(f,Q) \leq L(f,P\cup Q)$. Therefore, we have that

$$\begin{aligned} U(f,P\cup Q) - L(f,P\cup Q) &\leq U(f,P) - L(f,Q) \\ &< \left(U(f) + \frac{\epsilon}{2}\right) - \left(L(f) - \frac{\epsilon}{2}\right) = U(f) - L(f) + \epsilon \end{aligned}$$

However, f is Darboux integrable, so $U(f) = L(f)$, and rearranging the inequality, we obtain $L(f,P\cup Q) > U(f,P\cup Q) - \epsilon$, which is our desired result.

□

25.2. **Riemann Integration.**

Now we will go over Riemann's definition of an integral, which we will see is actually equivalent to Darboux integration.

Definition. *We will define the* ***mesh*** *of a partition P as such:*

$$\text{mesh}(P) = \max\{t_k - t_{k-1} \mid k \geq 1\}$$

In other words, the mesh P is the length of the largest subinterval of P.

The next theorem states another equivalent condition for integrability, but is *not* the Riemann integral.

Theorem. *A bounded function $f\colon [a,b] \to \mathbb{R}$ is Darboux integrable if and only if for all $\epsilon > 0$, there exists a $\delta > 0$ such that for all partitions P of $[a,b]$, if mesh$(P) < \delta$, then $U(f,P) - L(f,P) < \epsilon$.*

Proof. We will go in both directions.

($\Leftarrow$) If we assume that $\text{mesh}(P) < \delta$ implies $U(f,P) - L(f,P) < \epsilon$, then we can simply take the partition P that divides $[a,b]$ into subintervals of equal length less than δ, and using the previous theorem, we have our result.

($\Rightarrow$) Suppose f is integrable on $[a,b]$, and let $\epsilon > 0$ be given. By the previous theorem, we can choose a partition P such that $U(f,P) - L(f,P) < \frac{\epsilon}{2}$. Now, because f is bounded, there exists a $B > 0$ such that $|f(x)| \leq B$ for all $x \in [a,b]$.

Let Q be any partition such that $\text{mesh}(Q) < \delta$, and let $R = P \cup Q$. If we assume that $|R| = |Q| + 1$, then, repeating an argument from the proof of $L(f) \leq U(f)$, we have

$$L(f,R) - L(f,Q) \leq B \cdot \text{mesh}(Q) - (-B) \cdot \text{mesh}(Q) = 2B \cdot \text{mesh}(Q)$$

Let $|P| = m$. Thus, R can only have, at most, m elements that are not in Q. By induction, we can see that, if this is the case, then

$$L(f,Q) - L(f,P) \leq 2mB \cdot \text{mesh}(P)$$

Thus, we can set $\delta = \frac{\epsilon}{8mB}$, so that $2mB \cdot \text{mesh}(P) < 2mB\delta = \frac{\epsilon}{4}$. Now, because $|P| < |R|$, we have $L(f,P) - L(f,Q) < \frac{\epsilon}{4}$, and repeating the same steps for the Upper Darboux sums, we get $U(f,Q) - U(f,P) < \frac{\epsilon}{4}$. Combining these two facts, we obtain

$$U(f,Q) - L(f,Q) < U(f,P) - L(f,P) + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

completing the proof.

□

Definition. *Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function, and let P be a partition of $[a,b]$. f is Riemann integrable if there exists an $r \in \mathbb{R}$ such that for all $\epsilon > 0$, there is a $\delta > 0$ such that if* mesh$(P) < \delta$, *then*

$$\left| r - \sum_{k=1}^{n} f(x_k)(t_k - t_{k-1}) \right| < \epsilon$$

r is known as the Riemann integral of f on $[a,b]$, and is denoted $\int_a^b f$.

Note that the symbol for the Riemann integral is the same as the one for the Darboux integral. This will be explained by the next theorem. Note, however, that for the sake of establishing these two integrals' equivalence, we will denote the Riemann integral as $\mathcal{R}\int_a^b f$.

Theorem. *A bounded function $f\colon [a,b] \to \mathbb{R}$ is Darboux integrable if and only if it is Riemann integrable, and the values for each integral are equal.*

Proof. We will go in both directions.

($\Rightarrow$) Suppose f is Darboux integrable on $[a,b]$, and let $\epsilon > 0$ be given. From the previous theorem, we can choose $\delta > 0$ such that $\text{mesh}(P) < \delta$ implies $U(f,P) - L(f,P) < \epsilon$. We want to show that the above condition for Riemann integrability is satisfied in this case. So, let P be a partition such that $\text{mesh}(P) < \delta$. We will also denote $\sum_{k=1}^n f(x_k)(t_k - t_{k-1})$ as S.

It is obvious that $L(f,P) \leq S \leq U(f,P)$. Now, notice that by the previous theorem, we have

$$U(f,P) < L(f,P) + \epsilon \leq L(f) + \epsilon = \int_a^b f + \epsilon$$

$$L(f,P) > U(f,P) - \epsilon \geq U(f) - \epsilon = \int_a^b f - \epsilon$$

Therefore, $\int_a^b f - \epsilon \leq S \leq \int_a^b f + \epsilon$, so because ϵ was arbitrary, f is Riemann integrable, and $\mathcal{R}\int_a^b = \int_a^b$.

($\Leftarrow$) Suppose f is Riemann integrable, and let $\epsilon > 0$ be given. Choose δ and r as given in the definition of Riemann integrability. Now, let P be a partition such that $\text{mesh}(P) < \delta$ such that $|P| = n$. For each $t_k \in P$, there is an $x_k \in [t_{k-1}, t_k]$ such that $f(x_k) < m(f, [t_{k-1}, t_k]) + \epsilon$. Again, such x_ks exist by the definition of infimum. Now, if we consider these x_k, it is the case that

$$\sum_{k=1}^n f(x_k)(t_k - t_{k-1}) \leq L(f,P) + \epsilon(b-a)$$

Furthermore, by assumption, $|r - \sum_{k=1}^n f(x_k)(t_k - t_{k-1})| < \epsilon$. Therefore, substituting S for the summation, we have

$$L(f) \geq L(f,P) \geq S - \epsilon(b-a) > r - \epsilon - \epsilon(b-a)$$

Thus, $L(f) \geq r$ by the arbitrary nature of ϵ, and similarly, $U(f) \leq r$. Therefore, because we know $L(f) \leq U(f)$, $L(f) = U(f) = r$, f is Darboux integrable and $\int_a^b f = \mathcal{R}\int_a^b f$.

□

Now that we have proved this theorem, we can remove the $\mathcal{R}$ from the Riemann integral and use one symbol for both, namely $\int_a^b f$.

26. 4/28/11: The Fundamental Theorem of Calculus

26.1. **Integration Facts.**

Recall the symbol $\int_a^b f(x)dx$ where f is a bounded function from $[a, b]$ to $\mathbb{R}$. We have learned two equivalent ways of calculating the integral:

- The Darboux integral: $\int_a^b f(x)dx = U(f)$ where
$$U(f) = \inf U(f, P) = \sum M(f, [t_{i-1}, t_i])(t_i - t_{i-1})$$
$$L(f) = \sup L(f, P) = \sum m(f, [t_{i-1}, t_i])(t_i - t_{i-1})$$
and the integral is defined if $U(f) = L(f)$.
- The Riemann integral: Take $x_i \in [t_{i-1}, t_i]$ for all i. Then,
$$\int_a^b f(x)dx = \sum f(x_i)(t_i - t_{i-1})$$
The integral is defined if for all ϵ there exists a δ such that there exists an r that satisfies $|\sum f(x_i)(t_i - t_{i-1}) - r| < \epsilon$ if mesh$(P) < \delta$.

Our goal today is to prove the Fundamental Theorem of Calculus.

Theorem. *Any monotonic function is integrable.*

Proof. Exercise. □

Theorem. *Any continuous function is integrable.*

Proof. We will use the previous theorem that states that a function f is integrable if and only if for all $\epsilon > 0$, there is a partition P such that $U(f, P) \leq L(f, P) + \epsilon$. Let $\epsilon > 0$ be given. f is uniformly continuous because $[a, b]$ is compact. Thus, there exists a $d > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \frac{\epsilon}{b-a}$ for all $x, y \in [a, b]$. So, choose a partition P such that $\max\{t_i - t_{i-1} \mid i = 1, \ldots, n\} < \delta$. We can see that
$$M(f, [t_{i-1}, t_i]) < m(f, [t_{i-1}, t_i]) + \frac{\epsilon}{b - a}$$
because each subinterval has width less than δ, so the function does not differ by more than $\frac{\epsilon}{b-a}$ in each subinterval. Therefore, we now have
$$\sum_{i=1}^{n} M(f, [t_{i-1}, t_i])(t_i - t_{i-1}) < \sum_{i=1}^{n} m(f, [t_{i-1}, t_i])(t_i - t_{i-1}) + \left(\sum_{i=1}^{n}(t_i - t_{i-1})\right)(\frac{\epsilon}{b - a})$$
which implies $U(f, P) < L(f, P) + \epsilon$. □

26.2. **Properties of the Integral.**

Proposition (Linearity of $\int$)**.** *Let $f, g \colon [a, b] \to \mathbb{R}$. If $\int_a^b f dx$ and $\int_a^b g dx$ exist, then $\int_a^b (f + g)dx$ exists, and $\int_a^b (f + g)dx = \int_a^b f dx + \int_a^b g dx$.*

Proof. We will use Riemann's method of integration. Let $\epsilon > 0$ be given. Because f and g are integrable, there exists a $\delta > 0$ such that mesh$(P) < \delta$ implies

$|\sum f(x_i)(t_i - t_{i-1}) - \int_a^b f(x)dx| < \epsilon$, and likewise for g, for any choice of x_is in each subinterval $[t_{i-1}, t_i]$. Therefore, we have

$$\left|\sum (f+g)(x_i)(t_i - t_{i-1}) - \int_a^b f(x)dx - \int_a^b g(x)dx\right| =$$
$$\left|\left(\sum f(x_i)(t_i - t_{i-1}) - \int_a^b f(x)dx\right) + \left(\sum g(x_i)(t_i - t_{i-1}) - \int_a^b g(x)dx\right)\right|$$

The latter expression is less than $\epsilon + \epsilon$, so because ϵ was arbitrary, we are done. $\square$

Another part of this theorem is that for any $c \in \mathbb{R}$, $\int_a^b cf(x)dx = c\int_a^b f(x)dx$. The proof of this statement is left as an exercise.

Theorem. *If f is integrable, then $|f|$ is also integrable, and*

$$\left|\int_a^b f dx\right| \leq \int_a^b |f| dx^*$$

Proof Sketch. The crucial lemmas are that

$$M(|f|, [t_{i-1}, t_i]) - m(|f|, [t_{i-1}, t_i]) \leq |M(f, [t_{i-1}, t_i]) - m(f, [t_{i-1}, t_i])|$$

which follows from $|a| - |b| \leq |a - b|$ for any $a, b \in \mathbb{R}$, and the fact that if f, g are integrable and $f(x) \leq g(x)$ for all $x \in [a, b)$, then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$. To see this fact, consider $g - f$. $g - f \geq 0$, so $\int_a^b (g - f)dx \geq 0$, which implies, by the previous theorem, that $\int_a^b g dx \geq \int_a^b f dx$. $\square$

Theorem. *If $a < b < c$ for $a, b, c \in \mathbb{R}$ and f is integrable on $[a, b]$ and $[b, c]$, then f is integrable on $[a, c]$ and*

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$$

Proof. By Riemann's method; left as an exercise. $\square$

26.3. **The Fundamental Theorem of Calculus.**

Theorem (The Fundamental Theorem of Calculus, Part 1)**.** *Let f be differentiable on (a, b) and continuous on $[a, b]$. If f' is integrable on $[a, b]$, then*

$$\int_a^b f'(x)dx = f(b) - f(a)$$

Proof. We will proceed by Riemann's method of integration. Let $\epsilon > 0$ be given. Choose a partition P such that

$$\left|\int_a^b f'(x)dx - \sum f'(x_i)(t_i - t_{i-1})\right| < \epsilon$$

for all $x_i \in [t_{i-1}, t_i]$. By the Mean Value Theorem, for each interval $[t_{i-1}, t_i]$, there exists an $x_i \in (t_{i-1}, t_i)$ such that $f'(x) = \frac{f(t_i) - f(t_{i-1})}{t_i - t_{i-1}}$. Then, we have

$$\sum_{i=1}^{n} (t_i - t_{i-1}) f'(x_i) = \sum_{i=1}^{n} (f(t_i) - f(t_{i-1})) = f(b) - f(a)$$

---

*Note that this theorem is a version of the triangle inequality.

Therefore, the above chain of equalities implies that $|f(b) - f(a) - \int_a^b f'(x)dx| < \epsilon$, so we are done. $\square$

Theorem 12 (The Fundamental Theorem of Calculus, Part 2)**.** *Let f be an integrable function on $[a, b]$, and define the function $F(x) = \int_a^x f(t)dt$. Then,*

(1) *$F(x)$ is continuous.*
(2) *If f is continuous at x_0, then F is differentiable at x_0, and $F'(x_0) = f(x_0)$.*

Proof. Not covered in class, due to time constraints. Professor Jones referred the class to the proof in the textbook. $\square$

Appendix: Significant Results Proved in Homework

This is a list of all of the results that have been assigned as homework that we should know.

(1) Let S and T be nonempty subsets of $\mathbb{R}$ with the property that $s \le t$ for all $s \in S$ and $t \in T$. Then:
   (a) S is bounded above and T is bounded below.
   (b) $\sup S \le \inf T$.

(2) Let $\mathbb{I}$ be the set of real numbers that are not rational numbers, and let $a, b \in \mathbb{R}$. If $a < b$, then there exists an $x \in \mathbb{I}$ such that $a < x < b$.

(3) Let A and B be nonempty bounded subsets of $\mathbb{R}$, and let S be the set of all sums $a + b$ where $a \in A$ and $b \in B$. Then:
   (a) $\sup S = \sup A + \sup B$.
   (b) $\inf S = \inf A + \inf B$.

(4) Every element of a complete ordered field is the supremum of the set of rationals that are less than it.

(5) In a complete ordered field, a nonzero element is positive if and only if it is a square.

(6) Let F and G be complete ordered fields. There exists a unique bijection $f \colon F \to G$ such that

$$f(x + y) = f(x) + f(y) \text{ and } f(xy) = f(x)f(y)\,.$$

(7) Let (s_n) be a sequence in $\mathbb{R}$ such that $s_n \ge 0$ for all $n \in \mathbb{N}$. If $\lim s_n = 0$, then $\lim \sqrt{s_n} = 0$.

(8) Let (s_n) be a convergent sequence. If $\lim s_n > a$, then there exists an $N \in \mathbb{N}$ such that $n > N$ implies $s_n > a$.

(9) Let $\mathbb{R}(x)$ be the field of rational functions (with real coefficients) in x. Then, the set of elements of $\mathbb{R}(x)$ of the form $\frac{P(x)}{Q(x)}$ where P and Q are polynomials such that the coefficient of the highest power of x is positive satisfies the properties required to make $\mathbb{R}(x)$ into an ordered field.

(10) $\mathbb{R}(x)$, as defined above, is not Archimedean, and hence not complete.

(11) Every positive element of $\mathbb{R}$ has a unique positive nth root, for all $n \in \mathbb{N}$.

(12) The sum of two Cauchy sequences is Cauchy.

(13) The product of two Cauchy sequences is Cauchy.

(14) Let (s_n) and (t_n) be bounded sequences in $\mathbb{R}$. Then,

$$\limsup(s_n + t_n) \le \limsup s_n + \limsup t_n\,.$$

(15) A sequence (s_n) is bounded if and only if $\limsup |s_n| < \infty$.

(16) Addition and multiplication of sequences (in the natural way) in any ordered field makes the set of all sequences into a commutative ring. However, this structure does not define a ring structure on equivalence classes of sequences, as defined in Lecture 7.

(17) Let (s_n) be a sequence of nonnegative numbers, and for each n, define $\sigma_n = \frac{1}{n}(s_1 + s_2 + \cdots + s_n)$. Then,
   (a) $\liminf s_n \le \liminf \sigma_n \le \limsup \sigma_n \le \limsup s_n$.
   (b) If $\lim s_n$ exists, then $\lim \sigma_n$ exists, and $\lim s_n = \lim \sigma_n$.

(18) Let (X, d) be a metric space. Then, $D(x, y) = \frac{d(x,y)}{1+d(x,y)}$ also defines a metric on X and a subset U of X is open in (X, d) if and only if it is open in (X, D).

(19) Let $d\colon X \times X \to \mathbb{R}$ be a function satisfying all of the properties of a metric except $d(x,y) = 0 \Rightarrow x = y$. Define $\sim$ on X by $x \sim y \iff d(x,y) = 0$. Then, $\sim$ is an equivalence relation, setting $D([x],[y]) = d(x,y)$ is well-defined, and D makes the set of equivalence classes of X into a metric space.
(20) Let (X,d) be a metric space. Then, $|d(x,y) - d(x,z)| \leq d(y,z)$.
(21) Let (x_n) and (y_n) be two Cauchy sequences in (X,d). Then, $(d(x_n,y_n))$ is a Cauchy sequence in $\mathbb{R}$.
(22) Define a function $d((x_n),(y_n)) = \lim d(x_n,y_n)$ on the set of all sequences in a metric space (X,d). Also, define a relation $\sim$ on the set of all sequences in (X,d) by $(x_n) \sim (y_n) \iff d(x_n,y_n) \to 0$. d then defines a metric on the set of equivalence classes of Cauchy sequences, which we will denote $\hat{X}$. Furthermore, given that X can be embedded into $\hat{X}$ by defining $x \in X$ to be the equivalence class of the constant sequence of x, X is dense in $\hat{X}$.
(23) The metrics defined by the norms $||\cdot||_1$, $||\cdot||_2$, and $||\cdot||_\infty$ are equivalent.*
(24) Equivalent metrics define the same topology.
(25) Every polynomial function $p(x) = a_0 + a_1x + \cdots + a_nx^n$ is continuous on $\mathbb{R}$.
(26) The Cantor set is uncountable.
(27) Let A be an infinite set and F be a finite set. Then, $|A| = |A \cup F|$.
(28) A uniformly continuous function maps Cauchy sequences to Cauchy sequences.
(29) Let A be a dense subset of the metric space (X,d), and let (Y,D) be a complete metric space. Then, given a uniformly continuous function $f\colon A \to Y$, f admits a unique continuous extension to X.
(30) Any open subset U of $\mathbb{R}$ is a countable union of disjoint open intervals.
(31) Let X be a compact metric space, and Y be a metric space. Then, any continuous function from X to Y is uniformly continuous.
(32) Let (X,d) be a compact metric space and let $X^{\mathbb{N}}$ be the set of all sequences in X. Then, $D((x_n),(y_n)) = \sum\limits_{n=1}^{\infty} \frac{d(x_n,y_n)}{2^n}$ defines a metric on $X^{\mathbb{N}}$ for which it is sequentially compact, and hence compact.
(33) Assume that any finite map in the plane can be colored with four colors so that no adjacent countries have the same color. Then, any countably infinite planar map can also be colored with four colors.
(34) A compact metric space is complete.
(35) The closure of a connected subset of a metric space is also connected.
(36) Let E and F be connected subsets of a metric space. If E and F have empty intersection, then their union is connected. However, their intersection (regardless of whether it is empty or not) need not be connected.
(37) Cauchiness is not a topological property, i.e. there exists an example of a set which can take two metrics which define the same topology, but for which a sequence exists which is Cauchy in one metric but not the other.
(38) Define $C(v,r) \subseteq \mathbb{R}^2$ to be the circle centered at v of radius r, namely if $v = (a,b)$, then $C(v,r) = \{(x,y)\,|\,(x-a)^2 + (y-b^2) = r^2\}$. $C(v,r)$ is connected and path-connected.

*The notion of equivalent metrics is the same as the notion of equivalent norms detailed in Lecture 16.

(39) $\bigcup_{n=1}^{\infty} C((\frac{1}{n}, 0), \frac{1}{n})$, also known as the Hawaiian earring, is path-connected.
(40) The continuous image of a path-connected metric space is path-connected.
(41) Let $c\colon \mathbb{R} \to \mathbb{R}$ and $s\colon \mathbb{R} \to \mathbb{R}$ be differentiable functions such that $c(0) = 1$, $s(0) = 0$, $c'(x) = -s(x)$ for all x, and $s'(x) = c(x)$ for all x. Then, $c(x) = \cos(x)$ and $s(x) = \sin(x)$ for all x.
(42) $\cos(A + B) = \cos(A)\cos(B) - \sin(A)\sin(B)$ and $\sin(A + B) = \sin(A)\cos(B) + \cos(A)\sin(B)$.
(43) $\cos(t) = 0$ for some unique t between 1 and 2, and $\cos(x) > 0$ for all $x \in (-t, t)$.
(44) Let π be the real number equal to $2t$ as defined in the item above. Then, the following are true:

$$\cos\left(\frac{\pi}{2} - x\right) = \sin(x)$$
$$\sin\left(\frac{\pi}{2} - x\right) = \cos(x)$$
$$\sin(x + \pi) = -\sin(x)$$
$$\cos(x + \pi) = -\cos(x)$$

(45) $\exp(x)$ is a bijection from $\mathbb{R}$ to $(0, \infty)$, and $\log(x)$ is differentiable such that $\log'(x) = \frac{1}{x}$.
(46)

$$\log(1 - x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}$$

for all $x \in (-1, 1)$.
(47) Let $f\colon [0, 2] \to \mathbb{R}$ be continuous and differentiable in $(0, 2)$, and let $f(0) = 0$. If $-3 < f'(x) < 4$ for all $x \in (0, 2)$, then $-6 < f(2) < 8$.
(48) Let $S^1 := \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1\}$. There is a unique $\theta \in [0, 2\pi)$ such that $x = \cos\theta$ and $y = \sin\theta$, known as the angle between the positive x axis and the ray going from the origin through (x, y).
(49) The trigonometric polynomials, i.e.

$$\sum_{n=-k}^{k} a_n \sin n\theta + b_n \cos n\theta$$

where a_n and b_n are real numbers and k is a positive integer, are dense in $C(S^1, \mathbb{R})$.
(50) $|\cos(x) - \cos(y)| \leq |x - y|$ for all $x, y \in \mathbb{R}$.
(51) Let f be a differentiable on $\mathbb{R}$ such that $a = \sup\{|f(x)| \,|\, x \in \mathbb{R}\} < 1$. Then, for any $s_0 \in \mathbb{R}$, the sequence defined by $s_{n+1} = f(s_n)$ converges.
(52) Let f be a bounded function on $[a, b]$, such that $|f(x)| < B$ for all $x \in [a, b]$. Then, for all partitions P of $[a, b]$, it is the case that

$$U(f^2, P) - L(f^2, P) \leq 2B(U(f, P) - L(f, P))$$

Also, if f is integrable on $[a, b]$, then f^2 is integrable on $[a, b]$.
(53) Let $f(x) = x\,\mathrm{sgn}(\sin\frac{1}{x})$ for $x \neq 0$ and $f(0) = 0$. f is not piecewise continuous nor piecewise monotonic on $[-1, 1]$, but it is integrable on $[-1, 1]$.
(54) Suppose f is a continuous function on $[a, b]$ such that $f(x) \geq 0$ for all $x \in [a, b]$. If $\int_a^b f(x)dx = 0$, then $f(x) = 0$ for all $x \in [a, b]$.

(55) Let f be a continuous function on $[a, b]$ such that $\int_a^b f(x)g(x)dx = 0$ for all continuous functions g on $[a, b]$. Then, $f(x) = 0$ for all $x \in [a, b]$.
(56) Assume you are standing at $(0, 0)$, the origin in $\mathbb{R}^2$. Around you are circles of equal nonzero radius on every point (a, b) where $a, b \in \mathbb{Z}$. Then, you can neither see infinitely, nor can you see arbitrarily far. In other words, there does not exist a ray from $(0, 0)$ that does not, at some point, hit a circle, and for every radius of the circles, there exists some $R \in \mathbb{R}$ such that the length of the ray is always less than R.

Note that there are further concepts not covered in these lecture notes in the notes provided by Professor Jones about Metric Spaces and Cardinality, as well as extra exposition and concepts in the course textbook, *Elementary Analysis: The Theory of Calculus*, 1st edition, by Kenneth A. Ross, part of the Springer UTM series.