UM 101 (FALL 2022)

PART A: Real Analysis

What is this course about? This is “calculus for thinking practitioners". We will lay the foundation for “rules" that you are already very good at. So, we may work very hard to prove something that you already know, but it is the process which is the focus of the course. You will enjoy it if you are curious about “why" certain rules work, or “when do they fail", which is the kind of curiosity that makes a mathematician. For others, think of this as a training in rigor that will prepare you for the rest of this program.

1. SET THEORY AND THE REAL NUMBER SYSTEM

1.1. **The ZFC & Peano axioms.** To do analysis on real and complex numbers, we must talk about what those really are. And rather than take them completely for granted, we must understand where they come from. For that, we must understand rational numbers, which come from integers, which themselves come from natural numbers. Is it possible to synthesize a basic set of rules that eventually yield all the properties of natural numbers that we have been happily using? Peano (1858-1932) tried to do this:

Definition 1.1 (Peano's axioms)**.** A *set* A, is called a Peano set if it satisfies the following axioms or “rules".

- (P1) A contains a distinguished *element*, which is called 0. In particular A is *non-empty*.
- (P2) There exists a *function* S from A into itself.
- (P3) $S(a) \neq 0$ for any $a \in A$.
- (P4) S is injective, i.e., if $S(a) = S(b)$ for any $a, b \in A$, then $a = b$.
- (P5) (Principle of Mathematical Induction) if B is a *subset* of A such that $0 \in B$, and $S(a) \in B$ whenever $a \in B$, then $B = A$.

The function S is called a successor function on A.

Class discussion. Explain why the following choices of A and S do not make A a Peano set. You must explicitly mention which of the above axioms are failing.

a. $A = \{0, 1\}$, $S(0) = 1$ and $S(1) = 1$
b. $A = \{0, 0.5, 1, 1.5, 2, 2.5, \ldots\}$ and $S(x) = x + 1$.

It is “clear" to us that the “usual" natural numbers $\{0,1,2,\ldots\}$ endowed with $S(n) = n+1$ is a Peano set. One could take this for granted and proceed ahead (as we have done for many years in school). If you take only the existence of a Peano set for granted, you may argue as follows: why don't we define

$$\begin{aligned} 1 &:= S(0); \\ 2 &:= S(S(0)) \end{aligned}$$

and *so on*. Then using the Principle of Mathematical Induction, we show that the original Peano set contains nothing extra, and we are almost there (barring the issue of defining addition). However, this approach forces use to understand what “so on" (definition by *recursion*) means, and why the above procedure yields a subset and so on. In fact, we have to ascribe meaning to all of the italicized words above. In short, we must describe the fundamental rules governing the formation of sets and functions!

Definition 1.2. A set is a well-defined collection of (mathematical) objects that are called elements of the set. If A is a set and x is an object, then we use the notation

$$x \in A$$

to denote the fact that x is an element of A. Otherwise, we write

$$x \notin A.$$

Given two sets A, B, we say that A is a subset of B, or $A \subseteq B$, if every element of A is an element of B. Otherwise, we write that $A \nsubseteq B$. This is different from $A \subsetneq B$, i.e., A is a proper subset which means that every element of A is an element in B, but there is an element in B which is not in A.

Example 1. Going back to our known understanding of sets, let $A = \{1,2,\{3\}\}$. Then, $2 \in A$, $\{2\} \subseteq A$, but $\{2\} \notin A$. On the other hand, $\{3\} \in A$ but $3 \notin A$, and therefore, $\{3\} \nsubseteq A$.

END OF LECTURE 1

The ZFC Axioms - “Naive" approach. Note that not *every* collection of mathematical objects is considered a set (read about Russell's Paradox)! The following rules determine what is and isn't allowed to be called a set. (There are some redundancies in this list of axioms.)

A. **Basic.** Every mathematical object is a set. (This means that if you know that something is an element of a set, then that element itself is considered to be a set.)
B. **Extension.** Two sets, A and B, are said to be equal, i.e., $A = B$, if they have the same elements. Practically, to show $A = B$, one often shows that $A \subseteq B$ *and* $B \subseteq A$.

Example 2. Show that the following two sets are equal

$$\begin{aligned} A &= \{0,1\}, \\ B &= \text{ the set of possible remainders obtained when} \\ &\quad \text{dividing a perfect square by 4.} \end{aligned}$$

Let $a \in A$, we will show that there is a natural number c such that $c^2 \equiv a \pmod 4$. For $a = 0$, take $c = 2$. For $a = 1$, take $c = 1$. Thus, $A \subseteq B$. On the other hand, let $b \in B$. Then, $b = c^2 - 4q$ for some natural numbers c and q, and $0 \le b \le 3$. If c is even, then 4 divides b, which can only happen if $b = 0$. If c is odd, i.e., $c = 2k+1$, then $b = (2k+1)^2 - 4q = 4(k^2 + k - q) + 1$. The only number in $\{0,1,2,3\}$ of the form $4 \cdot +1$ is 1. Thus, we have shown that $B \subseteq A$.

C. **Existence.** There exists a set with no elements, called the empty set, denoted by $\emptyset$.

D. **Specification.** Let A be a set and $P(a)$ be a “property" that applies to every $a \in A$, i.e., $P(a)$ is either true or false, then

$$B = \{b \in A : P(b) \text{ is true}\}$$

exists as a set and is a subset of A.

Example 3. $B = \{a \in \mathbb{N} : 4|a\}$ is the set of natural numbers divisible by 4.

E. **Pairing.** Given two sets A, B, there exists a set containing exactly A and B as its elements, which is denoted by $\{A, B\}$. Note that this also gives the existence of $\{A\}$.

F. **Union.** For a set $\mathscr{F}$ of sets, there exists a set, called the union of the sets in $\mathscr{F}$, whose elements are precisely the elements of the elements of $\mathscr{F}$. I.e.,

$$x \in \bigcup_{A \in \mathscr{F}} A \iff x \text{ is an element of at least one } A \in \mathscr{F}.$$

Consequence. 1. For a non-empty set $\mathscr{F}$ of sets, there exists a set, called the intersection of the sets in $\mathscr{F}$ such that

$$x \in \bigcap_{A \in \mathscr{F}} A \iff x \text{ is an element of every} A \in \mathscr{F}.$$

G. **Power Set.** Given a set A, there exists a set $\mathscr{P}(A)$ whose elements are precisely the subsets of A. This set is called the power set of A.

Remark. The axioms so far give the existence of unions, intersections (why?), set differences, and (finite) Cartesian products (this is more complicated) as sets. In particular, given two sets A, B, the set $A \times B$ is the set of all *ordered pairs* (a, b), where $a \in A$ and $b \in B$.

END OF LECTURE 2

Definition 1.3. A relation from A to B is a subset R of $A \times B$. We say that aRb if and only if $(a, b) \in R$. The domain of R is

$$\operatorname{dom}(R) = \{x \in A : (x, y) \in R \text{ for some } y \in B\}.$$

The range of R is

$$\operatorname{ran}(R) = \{y \in B : (x, y) \in R \text{ for some } x \in A\}.$$

A function from A to B is a relation $f \subset A \times B$ such that $\operatorname{dom}(f) = A$ and for each $x \in A$, there is at most one $y \in B$ such that $x f y$. This y is denoted by $f(x)$.

H. **Replacement.** Skip!
I. **Regularity.** Skip!
J. **Choice.** Skip!

Definition 1.4. Given a set A, the successor of A is the set $A^+ = A \cup \{A\}$. A set which contains the empty set as an element and the successors of each of its elements is called an inductive set.

Question. Why does the A^+ exists?

K. **Infinity.** There exists an inductive set.

Lemma 1.5. *Let $\mathscr{C}$ be a non-empty set of inductive sets. Then, $B = \bigcap_{A \in \mathscr{C}} A$ is an inductive set.*

Proof. Homework! □

Theorem 1.6. *There exists a unique minimal inductive set. That is, there is an inductive set ω such that $\omega \subseteq S$ for every inductive set S, and if ω' is an inductive set with the same property, then $\omega = \omega'$.*

Proof. Let A be an inductive set, whose existence is guaranteed by the Axiom of Infinity. By the Axiom of Power Set, $\mathscr{P}(A)$ is a set. Then, by the Axiom of Specification,

$$F = \{W \in \mathscr{P}(A) : W \text{ is an inductive set}\}$$

is a set. Since F is non-empty (why?), we may consider

$$\omega = \bigcap_{W \in F} W.$$

Step 1. *ω is inductive.* This is a direct consequence of Lemma 1.5.

Step 2. *ω is minimal.* Let S be an inductive set. Then, by Lemma 1.5, $S \cap A$ is an inductive set. By the definition of F, $S \cap A \in F$. By the definition of intersections, $\omega \subseteq S \cap A \subseteq S$. Thus, $\omega \subseteq S$ (do you know what this property of $\subseteq$ is called?).
Step 3. *ω is unique.* By minimality, $\omega \subset \omega'$ and $\omega' \subset \omega$. By the Axiom of Extension, $\omega = \omega'$. $\square$

Theorem 1.7. *Let ω be* the *minimal inductive set. Set* $0 := \emptyset$ *and* $S(A) = A^+$. *Then, ω is a Peano set.*

Theorem 1.8 (The recursion principle)**.** *Let X be a non-empty set, $f : X \to X$ be a function, and $a \in X$. Then, there exists a unqiue function $F : \omega \to X$ such that*

(*a*) $F(\emptyset) = a$
(*b*) $F(A^+) = f(F(A))$ *for each* $A \in \omega$.

Remember that the above theorem is claiming the existence of a particular subset of $\omega \times X$. When we define sequences recursively, we are essentially applying the above theorem, as we will see shortly.

1.2. **The Natural numbers.** Within the ZFC regime, ω is defined as the set of natural numbers. But, how do we recover the familiar set of natural numbers from the above discussion? We identify $\mathbb{N}$ and ω as follows

$$\begin{aligned} 0 &:= \emptyset \\ 1 &:= 0^+ = \{\emptyset\} = \{0\} \\ 2 &:= 1^+ = \{\emptyset, \{\emptyset\}\} = \{0, 1\} \\ 3 &:= 2^+ = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = \{0, 1, 2\} \end{aligned}$$

and so on.

END OF LECTURE 3

Definition 1.9 (Peano Addition)**.** Given $m \in \mathbb{N}$, the recursion principle gives the existence of a unique function $\mathrm{sum}_m : \mathbb{N} \to \mathbb{N}$ such that

(1) $\mathrm{sum}_m(0) = m$;
(2) $\mathrm{sum}_m(n^+) = (\mathrm{sum}_m(n))^+$.

Define, for $m, n \in \mathbb{N}$,

$$m + n := \mathrm{sum}_m(n).$$

Let us prove an “obvious" statement!

Proposition 1.10. $2+3=5$.

Proof. We start from the left-hand side.

$$\begin{aligned} 2+3 &= \mathrm{sum}_2(3) = \mathrm{sum}_2(2^+) = (\mathrm{sum}_2(2))^+ = (\mathrm{sum}_2(1^+))^+ \\ &= ((\mathrm{sum}_2(1))^+)^+ = (((\mathrm{sum}_2(0))^+)^+)^+ = ((2^+)^+)^+ = (3^+)^+ = 4^+ = 5. \end{aligned}$$

$\square$

Trying to prove $2+8=10$ would be quite tedious, but not so much if we could say that $2+8=8+2$. Let us recover the basic properties of addition that we have taken for granted.

Remark. Note that $m^+ = (\mathrm{sum}_m(0))^+ = \mathrm{sum}_m(0^+) = m+1$ by this definition. So will now always write $m+1$ instead of m^+.

Definition 1.11 (Peano Multiplication.)**.** Given $m \in \mathbb{N}$, the recursion principle gives the existence of a unique function $\mathrm{prod}_m : \mathbb{N} \to \mathbb{N}$ such that

(1) $\mathrm{prod}_m(0) = 0$;
(2) $\mathrm{prod}_m(n^+) = \mathrm{prod}_m(n) + m$.

Define, for $m, n \in \mathbb{N}$,

$$m \cdot n := \mathrm{prod}_m(n).$$

Theorem 1.12. *The following hold.*

(1) *(Commutativity) $m+n=n+m$ and $m\cdot n = n\cdot m$ for all $m,n\in\mathbb{N}$.*
(2) *(Associativity) $m+(n+k)=(m+n)+k$ and $m\cdot(n\cdot k)=(m\cdot n)\cdot k$ for all $m,n,k\in\mathbb{N}$.*
(3) *(Distributivity) $m\cdot(n+k)=(m\cdot n)+(m\cdot k)$ for all $m,n,k\in\mathbb{N}$.*
(4) *$m+n=0$ implies that $m=n=0$ for $m,n,k\in\mathbb{N}$.*
(5) *$m\cdot n=0$ implies that either $m=0$ or $n=0$ for $m,n\in\mathbb{N}$.*
(6) *(Cancellation) $m+k=n+k$ implies that $m=n$, and $m\cdot k=n\cdot k$ with $k\neq 0$ implies that $m=n$ for $m,n,k\in\mathbb{N}$.*
(7) *SKIP or Give as exercise!(Order) Given $m,n\in\mathbb{N}$, exactly one of the following three statements hold.*
   (a) *$m=n$;*
   (b) *there is a $k\in\mathbb{N}\setminus\{0\}$ such that $m+k=n$;*
   (c) *there is a $k\in\mathbb{N}\setminus\{0\}$ such that $n+k=m$.*
(8) *SKIP or Give as exercise! (Order preservation) Given $m,n,k\in\mathbb{N}$ such that $n>m$ and $k>0$, $n\cdot k > m\cdot k$.*

Proof. We will only prove the additive part of (6). The rest are for self-study. Note that subtraction is not available to us! We will fix an arbitrary $m, n \in \mathbb{N}$ and prove the statement by induction on k.

Consider the case $k = 0$. We are given that $\mathrm{sum}_m(0) = \mathrm{sum}_n(0)$, which by definition gives that $m = n$. Next, assume the statement holds for k. We must prove it for $k+1$. For this, suppose $m + (k+1) = n + (k+1)$. I.e., $\mathrm{sum}_m(k+1) = \mathrm{sum}_n(k+1)$. By definition, this gives that

$$\mathrm{sum}_m(k) + 1 = \mathrm{sum}_n(k) + 1.$$

But the successor map is injective. So, $\mathrm{sum}_m(k) = \mathrm{sum}_n(k)$. By the induction hypothesis, $m = n$. By the principle of mathematical induction, the cancellation law holds for all $k \in \mathbb{N}$. Since $m, n \in \mathbb{N}$ were arbitrary, the claim holds. □

SKIP or Give as exercise! A natural number $n \in \mathbb{N}$ is said to be positive if $n \neq 0$. We say that n is greater than m if (b) holds, and write $n > m$. If either $n > m$ or $n = m$, we write $n \geq m$. **Question** Why does the positivity of n imply that $n > 0$?

END OF LECTURE 4

1.3. **Field axioms and ordered fields.** The set of natural numbers are missing many nice operations such as subtraction and division, for instance, we cannot solve

$$\begin{aligned} x + 3 &= 2, \\ 3x &= 5 \end{aligned}$$

within the set of natural numbers. In fact, we would like to work with number systems where we can even solve $x^2 + 2 = 0$ and $x^2 - 2 = 0$! First, we introduce some abstract concepts that capture some of these ideas.

Definition 1.13. A set F with two operations $+ : F \times F \to F$ and $\cdot : F \times F \to F$ is said to be a field if $(F, +, \cdot)$ satisfies the following "field axioms":

(F1) $+$ and $\cdot$ are **commutative** on $\mathbb{R}$.
(F2) $+$ and $\cdot$ are **associative** on $\mathbb{R}$.
(F3) $+$ and $\cdot$ satisfy the **distributive law** on $\mathbb{R}$.
(F4) There exists two distinct elements, denoted by 0 (the **additive identity**) and 1 (the **multiplicative identity**), such that for all $x \in F$,

$$\begin{aligned} x + 0 &= x \\ x \cdot 1 &= x. \end{aligned}$$

(F5) For every $x \in F$, there is a $y \in F$ such that

$$x + y = 0.$$

(F6) For every $x \in F$, $x \neq 0$, there is a $y \in F$ such that

$$xy = 1.$$

EXERCISE: Read Theorems 1.1-1.15 from Apostol.

Remark. You should be tempted to denote the y in (F5) by $-x$ and the y in (F6) by $1/x$. However, this notation would be meaningless unless we established the uniqueness of y in both the axioms. These are Theorems 1.1 and 1.7 in the book!

Remark. Define $-x$ as the unique y given by Theorem 1.1. It is called the additive inverse of x. We write $a + (-b)$ as $a - b$ for convenience.

Similarly, define $1/x$ as the unique multiplicative inverse of a nonzero x. We write $a \cdot (1/b)$ as a/b for convenience.

Theorem 1.14 (Apostol, Theorem 1.6)**.** $0 \cdot x = x \cdot 0 = 0$ *for all* $x \in F$.

Proof. By commutativity of $\cdot$ (F1), we already have the first equality. Now, by (F4), $1 + 0 = 1$ and $x \cdot 1 = x$. Thus, by (F3)

$$x = x \cdot 1 = x \cdot (1 + 0) = (x \cdot 1) + (x \cdot 0) = x + (x \cdot 0).$$

By the cancellation law (Theorem 1.1),

$$0 = x \cdot 0.$$

□

Definition 1.15. A set A with a relation $<$ is called an ordered set if the following holds:

(O1) For every $x, y \in A$, exactly one of the following three holds: $x = y$, $x < y$ or $y < x$.
(O2) Order is transitive. Given $x, y, z \in F$, if $x < y$ and $y < z$, then $x < z$.

If $x < y$, we say that x is strictly less than y. The notation $x \leq y$ denotes the statement $x < y$ or $x = y$, and we say that x is less than or equal to y. Similarly, $x > y$ and $x \geq y$ are interpreted.

Example 4. $\mathbb{N}$ is an ordered set if we define $<$ as follows: $m < n$ if there is a $k \in \mathbb{N} \setminus \{0\}$ such that $m + k = n$.

Definition 1.16. An ordered field is a set F that admits two operations $+, \cdot$ and a relation $<$ such that $(F, +, \cdot)$ is a field, $(F, <)$ is an ordered set and the following conditions hold.

(O3) For $x, y, z \in F$, if $y < z$, then $x + y < x + z$.

(O4) For $x, y \in F$, if $0 < x$ and $0 < y$, then $0 < xy$.

EXERCISE: Read Theorems 1.20-1.25 from Apostol. Our order axioms are DIFFERENT from Apostol's. His theorems 1.16-1.19 are our order axioms!!

Theorem 1.17 (Apostol, Theorem 1.21)**.** *In an ordered field,* $0 < 1$. *We may use Theorem 1.12 from Apostol:* $(-a) \cdot b = -(a \cdot b)$ *and* $(-a) \cdot (-b) = ab$.

Proof. From the field axioms, we know that $0 \neq 1$. Thus, by Axiom (O1), either $0 < 1$ or $1 < 0$. If the former holds, we are done.

If the latter holds, by (O3), $1 + (-1) < 0 + (-1)$. Thus, $0 < -1$. By (O4), $0 < (-1) \cdot (-1)$. Thus, $0 < 1$, which is a contradiction.

Thus, $0 < 1$ hold. $\square$

END OF LECTURE 5

1.4. **Upper bound and least upper bound.** Let $(F, +, \cdot, <)$ be an ordered field (you can think of the usual real numbers, with the usual $+$, $\cdot$ and $<$ as the key example).

Definition 1.18. A non-empty subset $S \subseteq F$ is said to be bounded above if there is an element $b \in F$ such that

$$x \leq b, \quad \forall x \in S.$$

We call b an upper bound of S. An upper bound of S that is contained in S is called a maximum of S.

Example 5. Let $S = \{x \in F : 0 \leq x \leq 1\}$ and $T = \{x \in F : 0 \leq x < 1\}$. Both S and T are bounded above. In both cases, 1 is an upper bound. In the cases, of S, 1 is the maximum of S.

Remark. Upper bounds clearly need not be unique. However, if a bounded set has a maximum, then the maximum must be unique (why?).

Definition 1.19. Let $S \subseteq F$ be a bounded above set. An element $b \in F$ is said to be a least upper bound of S or supremum of S if

(i) b is an upper bound of S,

(ii) for any $a \in F$ such that $a < b$, a is not an upper bound of S. In other words, given $a < b$, there exists an $s \in S$ such that $a < s$.

Theorem 1.20. *Let* $S \subset F$ *be a bounded set that admits a least upper bound. Suppose* b_1 *and* b_2 *are least upper bounds of* S*. Then,* $b_1 = b_2$

Proof. Note that there are three possiblities: either $b_1 = b_2$, $b_1 < b_2$ or $b_2 < b_1$. There is nothing to prove in the first case.

Suppose $b_2 < b_1$. Then, by item (ii) of Definition 1.19 applied to b_1, b_2 is not an upper bound of S. However, by item (i) of Definition 1.19 applied to b_2, b_2 is an upper bound of S. This is a contradiction.

Suppose $b_1 < b_2$. The same argument holds with the roles of b_1 and b_2 exchanged. □

Remarks. 1. We denote the least upper bound of a set by $\sup S$.
2. $\sup S$ need not belong to S!
3. The maximum of a set, if it exists, is its least upper bound.

Example 6. Let $S = \{x \in \mathbb{R} : 0 \le x < 1\}$. Then, $\sup S = 1$.

Proof. As noted earlier, 1 is an upper bound of S.

Now, let $a \in F$ such that $a < 1$. Then, either $a < 0$ or $0 \le a < 1$. In the former case, a is not an upper bound of S since $0 \in S$ and $a < 0$.

In the latter case, let $s = \frac{a+1}{2} = \frac{1}{2} \cdot (a+1)$. Since $0 < 1$, we have $0 \le 0 + a < a + 1$ (by O2). Thus, by (O2) $0 < a + 1$. Next, by (O4), $0 < s$. Next, $a < 1$. Thus, $a + 1 < 2$ by (O2). Similar to Problem 4 in HW 2, $s < 1$. Thus, $s \in S$. However, we also have that $2a < a + 1$. Thus $a < s$. This says that a is not an upper bound of S. We have established both (i) and (ii) of Definition 1.19. □

1.5. **The set of real numbers.** We assume the existence of a set $\mathbb{R}$ which admits operations $+$, $\cdot$ and a relation $<$ so that (F1)-(F6) and (O1)-(O4) are satisfied **and** the set admits the *least upper bound property*, i.e.,
(LUB) every bounded set $S \subseteq \mathbb{R}$ admits a least upper bound.

Some special subsets of $\mathbb{R}$ are:

(1) any $x > 0$ is called a positive number, and any $x < 0$ is called a negative number,
(2) $\mathbb{N} = \{0, 1, 2, \ldots\}$, the set of natural numbers,
(3) $\mathbb{P}$=\{1,2,...\}, the set of positive natural numbers,
(4) $\mathbb{Z} = \mathbb{N} \cup \{-n : n \in \mathbb{P}\}$, the set of integers,
(5) $\mathbb{Q} = \left\{\frac{p}{q} : q \in \mathbb{P}, p \in \mathbb{Z}\right\}$, the set of rational numbers,
(6) $\mathbb{Q}' = \mathbb{R} \setminus \mathbb{Q}$, the set of irrational numbers.

Theorem 1.21 (Archimedean propery)**.** *Let $x, y \in \mathbb{R}$ such that $x > 0$. There exists a positive integer n such that $nx > y$.*

Proof. May Skip! Fix $x > 0$. Let

$$S = \{nx : n \in \mathbb{P}\}.$$

Clearly, S is non-empty. Suppose the claim was not true. Then $nx \leq y$ for all $n \in \mathbb{P}$ and y us ab upper bound of S. Thus, by the l.u.b. property of $\mathbb{R}$. $b = \sup S$ exists. Now, since $b - x < b$, by item (ii) of Definition 1.19, there exists an $n \in \mathbb{P}$ such that $b - x < nx$. Thus, $b < (n+1)x \in S$. This contradicts the fact that b is an upper bound of S. $\square$

END OF LECTURE 6

Going forward, we will allow all the properties of real numbers that we have used so far, including exponentiation, the absolute value function, etc.

2.1. **Sequences.**

Definition 2.1. A sequence in $\mathbb{R}$ is a function $f : \mathbb{N} \to \mathbb{R}$. It is customary to write a sequence as $\{a_n\}_{n=0}^{\infty}$ or $\{a_n\}_{n\in\mathbb{N}}$, where

$$a_n := f(n), \quad n \in \mathbb{N},$$

and is called the n^{th}-term of the sequence.

Given a sequence, we would like to rigorously capture the idea that the terms of the sequence are approaching a single real number.

Definition 2.2. A sequence $\{a_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}$ is said to be convergent if there exists an $L \in \mathbb{R}$ such that for every $\varepsilon > 0$. there exists an $N_{\varepsilon,L} \in \mathbb{N}$ such that

$$|f(n) - L| < \varepsilon \quad \text{for all } n \geq N_{\varepsilon}.$$

In this case, we call L a limit of $\{a_n\}_{n\in\mathbb{N}}$, and write

$$a_n \to L \text{ as } n \to \infty.$$

A sequence $\{a_n\}_{n\in\mathbb{N}}$ in $\mathbb{R}$ is said to be divergent if it is not convergent.

Theorem 2.3. *Let $\{a_n\}_{n\in\mathbb{N}}$ be a convergent sequence in $\mathbb{R}$. Suppose ℓ_1 and ℓ_2 are limits of $\{a_n\}_{n\in\mathbb{N}}$. Then, $\ell_1 = \ell_2$.*

Proof. Let $\varepsilon > 0$. By definition, there exists an $N_1 \in \mathbb{N}$ such that $|a_n - \ell_1| < \varepsilon/2$ for all $n \geq N_1$. Similarly, there exists an $N_2 \in \mathbb{N}$ such that $|a_n - \ell_2| < \varepsilon/2$ for all $n \geq N_2$. Let $N = \max N_1, N_2$. Then,

$$0 \leq |\ell_1 - \ell_2| \leq |\ell_1 - a_N| + |a_N - \ell_2| < \varepsilon.$$

But, $\varepsilon > 0$ was arbitrary. HW. If $0 \leq x < \varepsilon$ for all $\varepsilon > 0$, then $x = 0$.
Note. This proof can also be done via contradiction. $\square$

Remark. We may now safely write

$$\lim_{n\to\infty} a_n = L.$$

This statement means that the limit exists and is equal to L.

Example 7. (a) Let $p > 0$ be fixed. Let $a_n = 1/n^p$, $n \in \mathbb{P}$. Then, $\{a_n\}_{n\in\mathbb{P}}$ is convergent and 0 is a limit.

Proof. Let $\varepsilon > 0$. By the Archimedean property of $\mathbb{R}$ applied to $x = \varepsilon^{1/p}$ and $y = 1$, there is an $N = N_\varepsilon \in \mathbb{P}$ such that

$$N\varepsilon^{1/p} > 1.$$

Thus, for $n \geq N$,

$$\left|\frac{1}{n^p} - 0\right| = \frac{1}{n^p} \leq \frac{1}{N^p} < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, our claim holds. □

END OF LECTURE 7

Example 8. (b) $a_n = (-1)^n$, $n \in \mathbb{N}$. Then, $\{a_n\}_{n\in\mathbb{N}}$ is divergent.

Proof. Suppose $\{a_n\}_{n\in\mathbb{N}}$ is convergent and admits a limit $L \in \mathbb{R}$. Let $\varepsilon = 1$. By definition, there exists an $N \in \mathbb{N}$ such that

$$|a_n - L| < 1 \quad \forall n \geq N.$$

Thus, by the triangle inequality,

$$|a_{2N} - a_{2N+1}| \leq |a_{2N} - L| + |a_{2N+1} - L| < 1 + 1 = 2.$$

But the left-hand side is 2. This is a contradiction.

□

Definition 2.4. A sequence $\{a_n\}_{n\in\mathbb{N}}$ is said to be bounded if there exists an $M > 0$ such that

$$|a_n| < M$$

for all $n \in \mathbb{N}$.

Definition 2.5. A sequence $\{a_n\}$ is said to be (montonically) increasing if $a_n \leq a_{n+1}$ for all n. A sequence $\{a_n\}$ is said to be (montonically) decreasing if $a_n \geq a_{n+1}$ for all n. A sequence is said to be monotone if it is either increasing or decreasing.

Theorem 2.6. *A monotone sequence is convergent if and only if it is bounded.*

Proof. (=>) **Case 1.** Let $\{a_n\}$ be an increasing and bounded sequence. Then, there exists an $M > 0$ such that

$$|a_n| < M \quad \forall n.$$

In other words $-M < a_n < M$ for all n. Let $S = \{a_n : n \in \mathbb{N}\}$. S is nonempty and bounded above. By LUB, $b = \sup S$ exists in $\mathbb{R}$.

Let $\varepsilon > 0$. By the definition of the supremum and monotonicity, $\exists N \in \mathbb{N}$ such that

$$b - \varepsilon < a_N \leq a_n \quad \forall n \geq N.$$

On the other hand, $a_n \leq b < b+\varepsilon$ for all $n \in \mathbb{N}$. Thus, for all $n \geq N$,

$$|a_n - b| < \varepsilon.$$

But ε was arbitrary.

Case 2. What if $\{a_n\}$ is decreasing?

(<=) In homework 03, you will show that *every* convergent sequence is bounded! □

⚠ Divergent sequences may diverge for different reasons!

Definition 2.7. We say that a sequence $\{a_n\}$ diverges to ∞ or $\lim_{n\to\infty} a_n = +\infty$ if for every $R \in \mathbb{R}$, there exists an $N_R \in \mathbb{N}$ such that $a_n > R$ for every $n \geq N_R$. One can similarly give meaning to $\lim_{n\to\infty} a_n = -\infty$.

Example 9. We show that $\lim_{n\to\infty} n = +\infty$. Let $R \in \mathbb{R}$. If $R \leq 0$, then $n \geq R$ for all $n \geq 1$. So we may choose $N_R = 1$. If $R > 0$, then by the Archimedean property, there is a positive natural number N such that $N > R$. Choose $N_R = N$ in this case, we get that $n > R$ for all $n \geq N_R$. Thus, we have shown that for any $R \in \mathbb{R}$, there is an $N_R \in \mathbb{N}$ such that $|a_n| > R$ for all $n \geq N_R$.

Theorem 2.8 (Limit Laws)**.** *Let $\{a_n\}$ and $\{b_n\}$ be convergent sequences in $\mathbb{R}$ with limits a and b, respectively.*

1. *For any $c \in \mathbb{R}$, $\{a_n + c\}$ converges to $a + c$ and $\{ca_n\}$ converges to ca.*
2. *The sequence $\{a_n + b_n\}$ converges to $a + b$.*
3. *The sequence $\{a_n b_n\}$ converges to ab.*
4. *Suppose $b \neq 0$ and $\exists M \in \mathbb{N}$ such that $b_n \neq 0\ \forall n \geq M$ then $\{1/b_n\}_{n\geq M}$ converges to $1/b$.*
5. *Suppose $b \neq 0$ and $\exists M \in \mathbb{N}$ such that $b_n \neq 0\ \forall n \geq M$ then $\{a_n/b_n\}_{n\geq M}$ converges to a/b.*

Proof. We will prove (4). Scrapwork: want:

$$\left|\frac{b-b_n}{bb_n}\right| < \varepsilon.$$

If we could show that

$$\left|\frac{b-b_n}{bb_n}\right| < M|b-b_n|$$

for some M independent of n, we would be done. For this, we need an (eventual) lower bound on $|b_n|$.

Let $\varepsilon_1 = |b|/2$. By convergence of $\{b_n\}$ to b, there exists an $N_1 \in \mathbb{N}$ such that, for all $n \geq N_1$,

$$|b_n - b| < \varepsilon_1 = |b|/2.$$

By the reverse triangle inequality,

$$\Big||b_n| - |b|\Big| < |b_n - b| < |b|/2.$$

Thus,

$$|b_n| > |b| - |b|/2 = |b|/2 \qquad \forall n \geq N_1. \tag{2.1}$$

This gives that for all $n \geq N_1$

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \left|\frac{b_n - b}{b_n b}\right| < \frac{2}{|b|^2}|b_n - b|.$$

Let $\varepsilon > 0$.

Set $\varepsilon_2 = \varepsilon/M$, where $M = \dfrac{2}{|b|^2}$. Then, there exists an $N_2 \in \mathbb{N}$ such that for all $n \geq N_2$,

$$|b_n - b| < \varepsilon/M. \tag{2.2}$$

Let $N = \max\{N_1, N_2\}$. Then, by (2.1) and (2.2), for all $n \geq N$,

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| < M|b_n - b| < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we are done.

$\square$

Remark. The converse statements generally do not hold in the above theorem, mostly because the convergence of sequences such as $\{a_n + b_n\}$, $\{a_n b_n\}$ and $\{a_n/b_n\}$ may not in general imply the convergence of the individual sequences $\{a_n\}$ and $\{b_n\}$.

END OF LECTURE 8

2.2. **Infinite Series.**

Definition 2.9. A infinite series of real numbers is a formal expression of the form

$$a_0 + a_1 + a_2 + \cdots \quad \text{or} \quad \sum_{n=0}^{\infty} a_n.$$

Given a series $\sum_{n=0}^{\infty} a_n$, its sequence of partial sums (sops) is the sequence $\{s_n\}_{n\in\mathbb{N}}$ given by

$$s_n = a_0 + \cdots + a_n, \quad n \in \mathbb{N}.$$

We say that $\sum_{n=0}^{\infty} a_n$ is convergent and has sum ℓ if $\{s_n\}_{n\in\mathbb{N}}$ is convergent with limit ℓ. Otherwise, we say that $\sum_{n=0}^{\infty} a_n$ is divergent.

Example 10. (a) (Harmonic series) Claim. $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent.

Proof. Observe that

$$\begin{aligned}
s_1 &= 1 \\
s_2 &= 1 + \frac{1}{2} \\
s_4 &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \geq 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} \\
s_8 &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{8} \geq 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} \\
&\vdots \\
s_{2^k} &\geq 1 + \frac{1}{2} + 2\frac{1}{2} + 4\frac{1}{8} + \cdots + 2^{k-1}\frac{1}{2^k} = 1 + k\frac{1}{2}.
\end{aligned}$$

For any $m \in \mathbb{R}$, there is some $k \in \mathbb{N}$ such that $s_{2^k} > m$. Thus, $\{s_n\}$ is divergent. □

(b) Claim. $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent.

Proof. Observe that

$$s_n < 1 + \sum_{k=2}^{n} \frac{1}{k(k-1)} = 1 + \sum_{k=2}^{n} \frac{1}{k-1} - \frac{1}{k} = 1 + 1 - \frac{1}{n} \to 2 \quad \text{as } n \to \infty.$$

□

Remark. We have seen an example of a telescoping series above. These are series of the form $\sum_{n=1}^{\infty} c_n - c_{n+1}$ for some sequence $\{c_n\}$. Note that, in this case, $s_n = c_1 - c_{n+1}$, so the convergence of the series entirely depends on the convergence of $\{c_n\}$.

(c) (Geometric) Claim. Let $-1 < x < 1$. Then, $\sum_{n=0}^{\infty} x^n$ converges and its sum is $\frac{1}{1-x}$. For $|x| \geq 1$, $\sum_{n=1}^{\infty} x^n = \frac{1}{1-x}$ diverges.

Proof. Observe that for $x \neq 1$,

$$s_n = 1 + x + \cdots + x^n = \frac{1 - x^{n+1}}{1-x} = \frac{1}{1-x} - \frac{x}{1-x}x^n.$$

We inspect the behavior of $\{x_n\}$ as $n \to \infty$. In the case when $|x| < 1$, we use the fact that $(1+y)^n \geq ny$ for any $y > 0$ and $n \in \mathbb{N}$, to observe that

$$(1 + (1/|x|) - 1)^n > n\left(\frac{1}{|x|} - 1\right) = nc.$$

Thus, $|x|^n < 1/(nc)$. By the “squeeze theorem" stated in HW04, we have that $\lim_{n\to\infty} x^n = 0$.

In the case, when $|x| > 1$, $(1+|x|-1)^n > n(|x|-1)$. Thus, for any $R \in \mathbb{R}$, by the Archimedean principle, there exists an $N \in \mathbb{N}$ such that $|x|^N > R$. Thus, $\{x^n\}$ is unbounded, and therefore, divergent.

Returning to the series $\sum_{n=0}^{\infty} x^n$. When $|x| < 1$, we use the limit laws of convergent sequences to say that

$$\lim_{n\to\infty} s_n = \frac{1}{1-x}.$$

When $|x| > 1$, $x \neq 1$, we use the fact that the sum of a convergent and divergent sequence is, in fact, divergent. Thus, $\{s_n\}$ is divergent.

For $x = 1$, observe that $s_n = n+1 \to \infty$ as $n \to \infty$. Thus, $\sum_{n=0}^{\infty}(1)^n$ diverges. $\square$

Theorem 2.10. *Let $\sum a_n$ be a convergent series. Then,* $\lim_{n\to\infty} a_n = 0$.

Proof. Let $\ell = \sum a_n$. Let $\varepsilon > 0$. Then, there exists $N \in \mathbb{N}$ such that $|s_n - \ell| < \varepsilon/2$ for all $n \geq N$. Observe that $a_n = S_n - S_{n-1}$, $n \geq 1$. Thus, $|a_n| \leq |S_n - \ell| + |S_{n-1} - \ell| < \varepsilon$ for all $n \geq N+1$. $\square$

END OF LECTURE 9

We will work with series with non-negative terms. Note that the sops for such series are monotonically increasing, so the task of showing convergence reduces to showing the (upper) boundedness of the sops.

Theorem 2.11 (Comparison Test)**.** *Suppose there exist constants $M \in \mathbb{N}$ and $C > 0$ such that*

$$0 \leq a_n \leq C b_n \quad \forall n \geq M.$$

Then, $\sum a_n$ if $\sum b_n$ is convergent. In other words, $\sum b_n$ diverges if $\sum a_n$ diverges.

Proof. Let $\{s_n\}$ and $\{t_n\}$ denote the sequence of partial sums of $\sum a_n$ and $\sum b_n$ respectively. Since $\sum b_n$ converges, there exists an $N > M$ and an $L > 0$ such that $t_n \leq L$ for all $n \geq N$. Now,

$$s_n \leq C t_n \quad \forall n \geq N.$$

Thus, $\{s_n\}$ — being a bounded monotone sequence — converges. $\square$

Example 11. (a) (p-series) Claim. $\sum_{n=1}^{\infty} \dfrac{1}{n^p}$ diverges for $p \leq 1$ and converges for $p > 1$.

Proof. For $p \leq 0$, $1/n^p \to \infty$, so the series must diverge.

For $0 < p \leq 1$, $n^p \leq n$. Thus, $\frac{1}{n^p} \geq \frac{1}{n}$. By the Comparison Test, the series diverges.

For $p \geq 2$, $n^p \geq n^2$. Thus, $\frac{1}{n^p} \leq \frac{1}{n^2}$. By the Comparison Test, the series converges.

For $1 < p < 2$, one must directly analyze the s.o.p.s. as we did with the case of $\sum 1/n$:

$$\begin{aligned} s_1 &= 1 \\ s_3 &= 1 + \frac{1}{2^p} + \frac{1}{3^p} \le 1 + 2\frac{1}{2^p} \\ s_7 &= 1 + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} + \cdots + \frac{1}{7^p} \le 1 + 2\frac{1}{2^p} + 4\frac{1}{4^p} \\ &\vdots \\ s_{2^k-1} &\le 1 + 2\frac{1}{2^p} + \cdots + 2^{k-1}\frac{1}{2^{p(k-1)}} < \frac{1}{1-2^{p-1}}. \end{aligned}$$

Exercise: complete the proof. □

Many other tests are derived from the comparison test. We will note two here, but prove only one.

Theorem 2.12 (Ratio Test)**.** *Let $\sum a_n$ be a series of non-negative terms. Suppose*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L.$$

(1) *If $L < 1$, the series converges.*
(2) *If $L > 1$, the series diverges.*
(3) *If $L = 1$, then the test is inconclusive.*

END OF LECTURE 10

Proof. **Case 1.** $L < 1$**.** Choose an r such that $L < r < 1$. Choosing $\varepsilon = r - L > 0$, we obtain an $N \in \mathbb{N}$ such that $\dfrac{a_{n+1}}{a_n} < L + \varepsilon = r$ for all $n \ge N$. Thus,

$$\begin{aligned} a_{N+1} &< a_N r \\ a_{N+2} &< a_{N+1} r < a_N r^2 \\ a_{N+3} &< a_{N+2} r < a_N r^3 \\ &\vdots \\ a_{N+k} &< a_N r^k. \end{aligned}$$

In other words, for $n \ge N$, $a_n \le \frac{a_N}{r^N} r^n = c r^n$. By the Comparison Test, and the convergence of the geometric series $\sum r^n$, $r < 1$, we have the convergence of $\sum a_n$.

Case 2. $L > 1$**.** Choose R such that $1 < R < L$. Then, choosing $\varepsilon = L - R$, we have that for some $N \in \mathbb{N}$, $\dfrac{a_{n+1}}{a_n} > L - \varepsilon = R > 1$ for all $n \ge N$. Thus, $a_{n+1} > a_n$ for all $n \ge N$. The sequence $\{a_n\}$ cannot converge to 0. Thus, $\sum a_n$ diverges.

Case 3. $L = 1$. $\sum \frac{1}{n^2}$ converges while $\lim_{n\to\infty} \frac{n+1}{n} = 1$. On the other hand, $\sum \frac{1}{n}$ diverges while $\lim_{n\to\infty} \frac{n^2+1}{n^2} = 1$. □

Theorem 2.13 (Root Test)**.** *Let $\sum a_n$ be a series of non-negative terms. Suppose*

$$\lim_{n\to\infty} \sqrt[n]{a_n} = R.$$

(1) *If $R < 1$, the series converges.*
(2) *If $R > 1$, the series diverges.*
(3) *If $R = 1$, then the test is inconclusive.*

Proof. Exercise!

For Case (3), we can take $\sum 1/n$ and $\sum 1/n^2$. How does one compute $\lim_{n\to} n^{1/n}$ without using Fms? □

Theorem 2.14 (Limit Laws for Series)**.** *Suppose $\sum a_n$ and $\sum b_n$ converge with sums a and b respectively. Then, for constants ℓ and m, $\sum \ell a_n + mb_n$ converges to $la + mb$. Suppose $\sum |a_n|$ and $\sum |b_n|$ converge. Then, so does $\sum |\ell a_n + mb_n|$ for any choice of ℓ and m in $\mathbb{R}$.*

Proof. Exercies! □

Corollary 2.15. *Suppose $\sum a_n$ converges and $\sum b_n$ diverges. Let $m \in \mathbb{R} \setminus \{0\}$. Then, $\sum(a_n + b_n)$ diverges, and $\sum mb_n$ diverges.*

Definition 2.16. A series $\sum a_n$ of real numbers is said to converge absolutely if $\sum |a_n|$ converges. A series $\sum a_n$ of real numbers is said to converge conditionally if $\sum |a_n|$ diverges but $\sum a_n$ converges.

Theorem 2.17. *If $\sum a_n$ converges absolutely, it must converge. Moreover, $|\sum a_n| \leq \sum a_n$*

Proof. We will construct a new series as follows:

$$b_n = a_n + |a_n|.$$

Observe that $0 \leq b_n \leq 2|a_n|$. Thus, by the comparison test, $\sum b_n$ converges. Now, by the limit laws for convergent series, $\sum a_n = \sum (b_n - |a_n|)$ converges. □

Example 12. Claim. $\sum \frac{(-1)^n}{n}$ is convergent.

Proof. Note that

$$\begin{aligned} s_1 &= -1 \\ s_3 &= -1 + \frac{1}{2} - \frac{1}{3} > s_1 \\ s_5 &= -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} > s_3 \\ &\vdots \\ s_{2k+1} &= \left(-1 + \frac{1}{2}\right) + \left(-\frac{1}{3} + \frac{1}{4}\right) + \cdots \left(-\frac{1}{2k+1}\right) < 0. \end{aligned}$$

Thus, $\{s_{2k+1}\}$ being a bounded increasing sequence, converges to some limit, say ℓ.

$$\begin{aligned} s_2 &= -\frac{1}{2} \\ s_4 &= -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} < s_2 \\ &\vdots \\ s_{2k} &= -1 + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{4} - \frac{1}{5}\right) + \cdots + \left(\frac{1}{2k}\right) \geq -1. \end{aligned}$$

Thus, $\{s_{2k}\}$ being a bounded decreasing sequence, converges to some limit, say m. Moreover, $s_{2k+1} = s_{2k} + \frac{1}{2k+1}$. So, by limit laws for sequences, $\ell = m$. Exercise: why does this suffice to claim that $\{s_n\}$ converges? $\square$

Theorem 2.18 (Alternating Series Test/Leibniz Test)**.** *Suppose $\{a_n\}$ is an decreasing sequence of positive numbers going to* 0. *Then, $\sum(-1)^n a_n$ converges. Denoting the sum by S, we have that*

$$0 < (-1)^n (S - s_n) < a_{n+1}.$$

Proof. Same principle as the example of $\sum \frac{(-1)^n}{n}$. $\square$

Remark. The estimate in AST allows us to estimate sums of alternating series within any prescribed error. For instance, to know $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ up to an error of 0.01. I need to find n so that

$$|S - s_n| < \frac{1}{100}.$$

Take $n = 99$, or the sum of the first 99 terms.

END OF LECTURE 11

3. LIMITS AND CONTINUITY

3.1. Limit of a function.

Definition 3.1. Given a real number p and an $\varepsilon > 0$, the ε-neighborhood of p is the open interval

$$N_\varepsilon(p) = (p-\varepsilon, p+\varepsilon) = \{x \in \mathbb{R} : p-\varepsilon < x < p+\varepsilon\} = \{x \in \mathbb{R} : |x-p| < \varepsilon\}.$$

Definition 3.2. (a) Given a function f that is defined on some $I = (a,p) \cup (p,b)$, $a < b$, we say that f has a limit L as x approaches p if:
for every $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $0 < |x-p| < \delta$, we have that $|f(x) - L| < \varepsilon$. OR
for every $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $x \in N_\delta(p) \setminus \{p\}$, $f(x) \in N_\varepsilon(L)$.
This is denoted by

$$\lim_{x \to p} f(x) = L.$$

——————Skipped one-sided limits!!——————

Given a function $f : (a,p) \to \mathbb{R}$ we say that f has a left-hand limit L as x approaches p if:
for every $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $p - \delta < x < p$, we have that $|f(x) - L| < \varepsilon$.
This is denoted by

$$\lim_{x \to p^-} f(x) = L.$$

Given a function $f : (p,b) \to \mathbb{R}$ we say that f has a right-hand limit L as x approaches p if:
for every $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $p < x < p + \delta$, we have that $|f(x) - L| < \varepsilon$.
This is denoted by

$$\lim_{x \to p^+} f(x) = L.$$

Remark. Check that $\lim_{x \to p} f(x) = L$ is equivalent to $\lim_{h \to 0} f(x+h) = L$.

Theorem 3.3 (Uniqueness of limits)**.** *Suppose f has limits L_1 and L_2 at p. Then, $L_1 = L_2$.*

Proof. Similar to the proof of uniqueness of sequential limits. □

Example 13. (1) (Constant functions) $f(x) = c$, for some fixed $c \in \mathbb{R}$. Let $p \in \mathbb{R}$. Let $\varepsilon > 0$. Let $\delta = 1$ (or any ohter positive number!). Then, whenever $0 < |x-p| < 1$, we have that $|f(x) - c| = 0 < \varepsilon$. Thus, for every $p \in \mathbb{R}$,

$$\lim_{x \to p} f(x) = c.$$

(2) (Identity function) $f(x) = x$. Fix $p \in \mathbb{R}$. Let $\varepsilon > 0$. Choose $\delta = \varepsilon$. Then, whenever $|x-p| < \delta$, we have that $|f(x) - f(p)| = |x = p| < \delta = \varepsilon$. Since, $\varepsilon > 0$ and $p \in \mathbb{R}$ were arbitrary, we have that $\lim_{x \to p} x = p$ for all $p \in \mathbb{R}$.

(3) $f(x) = \sqrt{x}$, $x > 0$. We will show that $\lim_{x\to p} f(x) = \sqrt{p}$, $p > 0$. Let $\varepsilon > 0$.

Scrapwork. We want to produce a $\delta > 0$ such that if $0 < |x-p| < \delta$, then $|\sqrt{x} - \sqrt{p}| < \varepsilon$. Note that

$$|\sqrt{x} - \sqrt{p}| = \frac{|x-p|}{|\sqrt{x} + \sqrt{p}|} \le \frac{1}{\sqrt{p}}|x-p|.$$

END OF LECTURE 12

Example 14. Let $\delta = \sqrt{p}\varepsilon > 0$. Then, for $0 < |x-p| < \delta$,

$$|\sqrt{x} - \sqrt{p}| = \frac{|x-p|}{\sqrt{x} + \sqrt{p}} \le \frac{1}{\sqrt{p}}|x-p| < \varepsilon. \tag{3.1}$$

Since $\varepsilon > 0$ was arbitrary, we are done.

(4) $f(x) = \frac{1}{x}$, $x \neq 0$. Then, $\lim_{x\to 0} f(x)$ does not exist. Suppose it did, and was L. Then, for any $\varepsilon > 0$, there exists a $\delta > 0$ such that whenever $0 < |x| < \delta$, $|1/x - L| < \varepsilon$. In particular, for any nonzero $x, y \in (-\delta, \delta)$, we must have that

$$|1/x - 1/y| < 2\varepsilon.$$

Let $\varepsilon = 1/2$ and δ correspond to this ε. Then, by the Archimedean property of $\mathbb{R}$, there is an $N \in \mathbb{N}$ such that $N\delta > 1$, or $\frac{1}{N} < \delta$. Choose, $x = N$ and $y = N+1$. Then. $0 < |x|, |y| < \delta$, but $|1/x - 1/y| = |N+1-N| = 1 > 1/2$.

Theorem 3.4 (Limit laws for functions)**.** *Let f and g be functions such that* $\lim_{x\to p} f(x) = A$ *and* $\lim_{x\to p} g(x) = B$. *Then,*

(1) $\lim_{x\to p}(f(x) \pm g(x)) = A \pm B$.
(2) $\lim_{x\to p}(f(x)g(x)) = AB$.
(3) $\lim_{x\to p}(f(x)/g(x)) = A/B$, *if* $B \neq 0$.

Proof. We will prove (2). Let $\varepsilon > 0$.

Scrapwork. Note that

$$|f(x)g(x) - AB| = |f(x)g(x) - f(x)B + f(x)B - AB| \le |f(x)||g(x) - B| + |B||f(x) - A|.$$

Let $\varepsilon_1 = \dfrac{\varepsilon}{(2|B|+1)}$. Let $\delta_1 > 0$ be such that whenever $0 < |x-p| < \delta_1$, then $|f(x) - A| < \varepsilon_1$. In particular,

$$|f(x)| < |A| + \varepsilon_1 = M.$$

Let $\varepsilon_2 = \dfrac{\varepsilon}{2M}$. Let $\delta_2 > 0$ be such that whenever $0 < |x-p| < \delta_2$, then $|g(x) - A| < \varepsilon_2$.

Finally, let $\delta = \min\{\delta_1, \delta_2\}$. When $0 < |x-p| < \delta$,

$$|f(x)g(x) - AB| = |f(x)g(x) - f(x)B + f(x)B - AB| \leq |f(x)||g(x)-B| + |B||f(x)-A|$$
$$\leq M\varepsilon_2 + |B|\varepsilon_1 < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

□

END OF LECTURE 13

3.2. Continuity.

Definition 3.5. Let $S \subset \mathbb{R}$, \$ $f : S \to \mathbb{R}$ and $p \in S$. We say that f is continuous at p if: for every $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $|x-p| < \delta$ and $x \in S$, we have that $|f(x) - f(p)| < \varepsilon$, in other words

$$\lim_{x \to p} f(x) = f(p).$$

We say that f is continuous on S if f is continuous at each $p \in S$.

Theorem 3.6 (Algebraic combinations of continuous functions)**.** *Let f, g be functions that are continuous at p. Then, $f \pm g$, fg and f/g (when $g(p) \neq 0$) are continuous at p.*

Example 15. (*a*) Based on our previous computations, constant functions and the identity function ($f(x) = x$) are continuous on all of $\mathbb{R}$. Using the above theorem, every polynomial, i.e., function of the form

$$p(x) = a_n x^n + \cdots a_0,$$

where $a_0, \ldots, a_n$ are constants, is continuous. Every rational function, i.e., function of the form

$$r(x) = \frac{p(x)}{q(x)},$$

where p and q are polynomials, is continuous on the set $\{x \in \mathbb{R} : p(x) \neq 0 \text{ and } q(x) \neq 0\}$. **Note.** $r(x)$ may be **extended** to a continuous function on certain points x, where $p(x) = q(x) = 0$, but a priori, the function $r(x)$ itself is considered **undefined** on such points! E.g., the function

$$r(x) = \frac{x^2-1}{x-1}$$

is only defined and continuous on $\mathbb{R} \setminus \{1\}$. The function

$$R(x) = \begin{cases} \frac{x^2-1}{x-1}, & x \neq 1, \\ 2, & x = 1. \end{cases}$$

is defined on all of $\mathbb{R}$, and is in fact also continuous everywhere (but this requires proof at $x = 1$).

(b) $f(x) = x^r$ for any $r \geq 0$ is continuous on $(0, \infty)$. In HW 5, you will show that $\lim_{x \to p} x^n = p^n$ for every $n \in \mathbb{N}$ and $p \in \mathbb{R}$. Now, let $q > 0$. There exists a natural number n such that $q \leq n$. Thus,

$$0 \leq y^q - 1 \leq y^n - 1, \quad y > 1$$
$$0 \leq 1 - y^q \leq 1 - y^n, \quad y \leq 1.$$

By the squeezing principle (HW05), $\lim_{x \to 1} x^q = 1$. Now, for any $a > 0$, note that

$$|x^q - a^q| = a^q \left| (x/a)^q - 1 \right|.$$

(c) Trigonometric functions will have the meaning (in terms of triangles) that they did in school (angles are measured in radians!). You may assume all the trigonometric formulas from school. You will establish their continuity in HW5.

END OF LECTURE 14

Theorem 3.7 (Compositions)**.** *Given $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$, assume that ran$(f) \subset B$. Then, $g \circ f : A \to \mathbb{R}$ given by $g \circ f(x) = g(f(x))$ is well-defined. If f is continuous at $p \in A$, and g is continuous at $q = f(p)$, then $g \circ f$ is continuous at p.*

Proof. Let $\varepsilon > 0$. We must produce a $\delta > 0$ so that whenever $|x - p| < \delta$ and $x \in A$, we have that

$$|g \circ f(x) - g \circ f(p)| < \varepsilon.$$

By the continuity of g at $q = f(p)$, there is a τ such that whenever $|y - q| < \tau$ and $y \in B$, then

$$|g(y) - g(q)| < \varepsilon. \tag{3.2}$$

Now, set $\varepsilon_2 = \tau$. By the continuity of f at p, there is a $\delta > 0$ such that whenever $|x - p| < \delta$ and $x \in A$, we have that

$$|f(x) - f(p)| < \tau.$$

Thus, whenever $|x - p| < \delta$ and $x \in A$, we have that $|f(x) - q| < \tau$ and $f(x) \in B$. Thus, by (3.2),

$$|g(f(x)) - g(f(x))| < \varepsilon.$$

$\square$

Before we proceed to the next theorem about continuous functions, let us prove a result about sequences.

Lemma 3.8 (Comparison for sequences)**.** *Suppose $\{a_n\}, \{b_n\} \subset \mathbb{R}$ are convergent sequences such that $a_n \leq b_n$, then $L_1 = \lim_{n \to \infty} a_n \leq \lim_{n \to \infty} b_n = L_2$.*

Proof. Suppose not. Say that $L_1 > L_2$. Let $\varepsilon = (L_1 - L_2)/2$. Then, there exists an $N_1 \in \mathbb{N}$ and $N_2 \in \mathbb{N}$ such that $L_1 - \varepsilon < a_n$ and $b_n < L_2 + \varepsilon$ for all $n \geq \max\{N_1, N_2\}$. I.e., $b_n < (L_1 + L_2)/2 < a_n$ for all $n \geq \max\{N_1, N_2\}$. This is a contradiction. $\square$

A common test used to detect zeros of a polynomial are to look for places where the polynomial “changes sign". This comes from an important property of continuous functions!

Theorem 3.9 (The Intermediate Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Suppose y is a real number lying between $f(a)$ and $f(b)$. Then, there is a $c \in [a, b]$ such that $f(c) = y$.*

We will assume that $f(a) \leq y \leq f(b)$. The case $f(b) \leq y \leq f(a)$ can be handled similarly.

Proof. f $y = f(a)$ or $y = f(b)$, then $c = a$ or $c = b$, respectively, works. So, we may assume that $f(a) < y < f(b)$. Let

$$S = \{x \in [a, b] : f(x) < y\} \subset [a, b].$$

Since $a \in S$, S is nonempty. Since $b \geq x$ for all $x \in S$, S is bounded above. Thus,

$$c := \sup S \quad \text{exists.}$$

We first show that $f(c) \leq y$. Let $n \in \mathbb{P}$. Since $c - 1/n$ is not an upper bound of S, there is an $x_n \in S$ such that

$$c - \frac{1}{n} < x_n \leq c.$$

Thus, by the above lemma, $\lim_{n\to\infty} x_n = c$. By the above lemma and sequential characterization of continuity, $y \geq \lim_{n\to\infty} f(x_n) = f(c)$.

Next, we show that $y \leq f(c)$. Since b is an upper bound of S, $c \leq b$. Moreover, $f(c) \leq y < f(b)$, so $c \neq b$. Let $N \in \mathbb{N}$ such that $c + 1/n < b$ for all $n \geq N$. Then, $f(c + 1/n) \geq y$. By the sequential char. of cont. and the above lemma, $f(c) = \lim_{n\to} f(c + 1/n) \geq y$. $\square$

Corollary 3.10 (Bolzano’s theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function such that $f(a)$ and $f(b)$ assume opposite signs (so, are non-zero). Then, there is at least one $c \in (a, b)$ such that $f(c) = 0$.*

Informal discussion. A continuous function on a circle is the same as a continuous 2π-periodic function on the real line. For such a function f, let

$$g(\theta) = f(\theta) - f(\theta + \pi).$$

Then, $g(0) = f(0) - f(\pi) = -(f(\pi) - f(2\pi)) = -g(\pi)$. Either $g(0) = 0$, or g takes opposite signs at 0 and π. In either case: f takes the same values on a pair of anti-podal points. This is the 1-dimensional case of the Borsuk–Ulam theorem. This is the basis of statements such as, *"At any given time, there are two polar opposite points on the Earth's equator where the temperature is exactly the same!"*

END OF LECTURE 15

———–Skip!———–

Theorem 3.11 (Existence of nth roots)**.** *Given a positive integer n and a positive number a, there is exactly one positive nth root of a.*

Proof. Let $c > 1$ such that $a \in [0, c]$. Let $f(x) = x^n$. Then, $f(0) = 0$ and $f(c) = c^n > c > a > 0$. By IVT, there is some $b \in (0, c)$ such that $b^n = a$. Now, since $0 < x < y \Rightarrow x^n < y^n$, we are done. □

———–Skip!———–

Definition 3.12. A function $f : S \to \mathbb{R}$ is said to be bounded above on S if there is some $U \in \mathbb{R}$ such that

$$f(x) \leq U \quad \forall x \in S.$$

A function $f : S \to \mathbb{R}$ is said to be bounded below on S if there is some $L \in \mathbb{R}$ such that

$$f(x) \geq L \quad \forall x \in S.$$

A function $f : S \to \mathbb{R}$ is said to be bounded on S if there is some $M > 0$ such that

$$|f(x)| \leq M \quad \forall x \in S.$$

Theorem 3.13. *A function $f : S \to \mathbb{R}$ is bounded if and only if it is bounded above and bounded below on S.*

Example 16. The function $1/x$, $x \neq 0$, is neither bounded above nor bounded below on its domain. However, it is bounded below on $(0, 1)$.

Theorem 3.14. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then f is bounded on $[a, b]$.*

Proof. Given any closed interval $I = [c, d]$ with midpoint e, if f is bounded on $I^- = [c, e]$ and $I^+ = [e, d]$, then f is bounded on $[c, d]$. Suppose f is not bounded on $[a, b]$. Let $I_0 = [a, b]$ and $a_0 = a$. Then, by the contra-positive of the above statement, f is either not bounded on $I_0^- = [a, (a + b)/2]$ or on $I_0^+ = [(a + b)/2, b]$. Pick the one on which it is not bounded (pick left to break a tie), and call it I_1. Call its left endpoint a_1. Continuing this way, inductively, we obtain a sequence of intervals

$I_0 = [a_0, b_0], \ldots, I_k = [a_k, b_k], \ldots$, so that f is unbounded on I_k and I_k is either I_{k-1}^+ or I_{k-1}^-. Since we are halving the intervals at each stage, $|b_j - a_j| < (b-a)/2^j$.

Let $s = \sup\{a_j : j \in \mathbb{N}\}$. Justify why s exists, and show that $s \in [a, b]$.

By continuity, there is a $\delta > 0$ such that if $x \in N_\delta(s) \cap [a, b]$, then $|f(x) - f(s)| < 1$. I.e., $|f(x)| < 1 + |f(s)|$ on $x \in N_\delta(s) \cap [a, b]$. Let $N \in \mathbb{N}$ such that for all $n \geq N$, $a_n \in N_{\delta/2}(s) \cap [a, b]$. This works because of the monotonicity of $\{a_n\}$. Choose $N' \geq N$ such that for all $n \geq N'$, $(b-a)/2^n < \delta/2$. Thus,

$$|b_n - s| < |b_n - a_n| + |a_n - s| < \delta \qquad \forall n \geq N'.$$

This contradicts the fact that f is unbounded on these intervals! □

END OF LECTURE 16

Definition 3.15. A function $f : S \to \mathbb{R}$ is said to have a global maximum (minimum) at $p \in S$ if $f(p) \geq (\leq) f(x)$ for all $x \in S$.

Theorem 3.16 (Extreme value theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then, f attains a global maximum and a global minimum in $[a, b]$.*

Proof. Let $s = \sup\{f(x) : x \in [a, b]\}$. We show that there is some $c \in [a, b]$ such that $f(c) = s$. Suppose not. Let $g(x) = s - f(x)$. Then, g is continuous and $g(x) > 0$ on $[a, b]$. Thus, $1/g(x)$ is bounded on $[a, b]$. Thus, by the above theorem, there is an $M > 0$ such that $1/g(x) \leq M$ for all $x \in [a, b]$. Thus, $s - f(x) \geq 1/M$ for all $x \in [a, b]$. Thus, $s - (1/M) \geq f(x)$ on $[a, b]$. This contradicts the fact that s is the least upper bound.

The first part of the argument applied to $-f$ on $[a, b]$ shows that there exists a $d \in [a, b]$ such that $-f(d) \geq -f(x)$ for all $x \in [a, b]$. Thus, $f(x) \leq f(d)$ on $[a, b]$. □

Corollary 3.17. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then, $f([a, b]) = [\min_{[a,b]} f(x), \max_{[a,b]} f(x)]$.*

Theorem 3.18 (Composition for limits, based on HW05)**.** *Let f and g be functions such that*

$$\lim_{x \to p} f(x) = L \quad \textit{and} \quad \lim_{y \to L} g(y) = M.$$

Moreover, suppose that for some $\delta > 0$, if $0 < |x - p| < \delta$, then $|f(x) - L| > 0$. Then,

$$\lim_{x \to p} g(f(x)) = \lim_{y \to L} g(y).$$

Fun fact: continuous functions can be *very* weird. They can have “corners" or “be jagged" at every point!

Definition 4.1. Given a function $f : (a,b) \to \mathbb{R}$ and $p \in (a,b)$ we say that f is differentiable at p if the limit

$$\lim_{h\to 0} \frac{f(p+h)-f(p)}{h} = \lim_{x\to p} \frac{f(x)-f(p)}{x-p}$$

exists, in which case, we call the above limit $f'(p)$. $f'(p)$ is also sometimes known as the rate of change of f at p. If f is differentiable at each $p \in (a,b)$, it is said to be differentiable on (a,b), and $f' : (a,b) \to \mathbb{R}$ is the derivative function of f.

Example 17. (a) (Constant functions) $f(x) = c$, $c \in \mathbb{R}$, is differentiable on $\mathbb{R}$ and $f'(x) \equiv 0$. For any $x \in \mathbb{R}$ and $h \neq 0$,

$$\frac{f(x+h)-f(x)}{h} = 0.$$

(b) (Linear functions) $f(x) = ax+b$, $x, a, b \in \mathbb{R}$, is differentiable at each $x \in \mathbb{R}$, and $f'(x) = a$. Proof: homework!
(c) (Natural powers) $f(x) = x^n$, $n \in \mathbb{P}$. We will use the formula (telescoping sum)

$$a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k}.$$

Note that, for any $x \in \mathbb{R}$ and $n \in \mathbb{P}$,

$$\frac{(x+h)^n - x^n}{h} = \sum_{k=0}^{n-1} (x+h)^k x^{n-1-k}.$$

But polynomials are continuous, so $\lim_{h\to 0}\sum_{k=0}^{n-1}(x+h)^k x^{n-1-k} = \sum_{k=0}^{n-1} xn-1 = nx^{n-1}$.
(d) (Sine and cosine) Homework! Uses addition formulas, and $\lim_{x\to 0} \frac{\sin x}{x} = 1$.

Theorem 4.2. *Let $f : (a,b) \to \mathbb{R}$ be a function that is differentiable at $p \in (a,b)$. Then, f is continuous at p.*

Proof. $x \in (a,b) \setminus \{p\}$, we may write

$$f(x) - f(p) = (x-p)\frac{f(x)-f(p)}{x-p}.$$

By the limit laws, the limit of the R.H.S., as $x \to p$, exists, and is equal to 0. Thus,

$$\lim_{x\to f(x)=f(p).}$$

$\square$

Remark. The converse need not be true. Consider $f(x) = |x|$. The function is continuous (why?), but observe that

$$\frac{f(h) - f(0)}{h} = \begin{cases} 1, & h > 0, \\ -1, & h < 0. \end{cases}$$

Thus, taking $\{(-1)^n n^{-1}\}$, we see that the derivative at 0 does not exist.

Theorem 4.3 (Algebra of derivatives)**.** *Let f and g be defined on (a,b) and differentiable at $p \in (a,b)$. Then,*

(1) *so is $f+g$, and $(f+g)'(p) = f'(p) + g'(p)$,*
(2) *(product rule) so is fg, and $(fg)'(p) = f(p)g'(p) + f'(p)g(p)$,*
(3) *(quotient rule) if $g(p) \neq 0$, so is f/g, and $(f/g)'(p) = \frac{g(p)f'(p) - f(p)g'(p)}{g^2(p)}$.*

Proof. We only prove the quotient rule in the special case of $f \equiv 1$. Note that

$$\frac{\frac{1}{g(p+h)} - \frac{1}{g(p)}}{h} = -\frac{g(p+h) - g(p)}{h} \frac{1}{g(p)} \frac{1}{g(p+h)}.$$

By the continuity of g at p and the algebra of limits of functions,

$$\lim_{h \to 0} \frac{\frac{1}{g(p+h)} - \frac{1}{g(p)}}{h} = -\lim_{h \to 0} \frac{g(p+h) - g(p)}{h} \frac{1}{g(p)} \frac{1}{g(p+h)} = -\frac{g'(p)}{g(p)^2}.$$

□

Example 18. We now have the differentiability of polynomials and rational functions in their domains of definition.

END OF LECTURE 18

Definition 4.4. Let $f : A \to B$ be a one-to-one and onto function. f is said to be invertible on A, and given, any $y \in B$, there is a unique $x_y \in A$ such that $f(x_y) = y$. Define the inverse of f function $f^{-1} : B \to A$ as

$$f^{-1}(y) = x_y.$$

Then, $f^{-1}(f(x)) = x$ for all $x \in A$, and $f^{-1}(f(y)) = y$ for all $y \in B$.

Theorem 4.5 (Monotone/Continuous/Differentiable Inverse Function Theorem)**.** *Let $f : [a,b] \to \mathbb{R}$ be an invertible function with range J. Then*

(i) *If f is strictly increasing (decreasing), so is $f^{-1} : J \to \mathbb{R}$.*
(ii) *If f is continuous on $[a,b]$, then f is strictly monotone on $[a,b]$, and f^{-1} is continuous on J.*

(iii) *If f is continuous on $[a,b]$, f is differentiable at $p \in (a,b)$ and $f'(p) \neq 0$, then f^{-1} is differentiable at $q = f(p)$, and*

$$\left(f^{-1}\right)'(f(p)) = \frac{1}{f'(p)}.$$

Proof. (i) Almost follows from definition, so left as an exercise.

(ii) First, we show the strict monotonicity of f.

Case 1. $f(a) < f(b)$. Let $g : [a,b] \to \mathbb{R}$ be given by $g(y) = f(y) - f(a)$. Note that g is cont. on $[a,b]$ and $g(b) > 0$. Suppose $g(y) \leq 0$ for some $y \in (a,b)$. Then g vanishes at some $c \in [y,b)$ by IVT, and $f(c) = f(b)$. This contradicts the one-t-one-ness of f. So, $f(a) < f(y)$ for all $y \in (a,b]$. Similarly, by considering $h(y) = f(y) - f(b)$ on $[a,b]$, we have that $f(y) < f(b)$ for all $y \in [a,b)$.

Thus,

$$f(a) < f(y) < f(b), \text{ if } a < y < b. \tag{4.1}$$

Now, let $x, y \in [a,b]$ such that $x < y$. If $x = a$, we already have that $f(x) < f(y)$ by (4.1). Now, suppose $a < x < y$. Then, applying the above argument to f restricted to the domain $[a,y]$, we have that $f(a) < f(x) < f(y)$. Thus, f is strictly increasing on $[a,b]$

Case 2. $f(a) > f(b)$. Apply Case 1 to $-f$ to obtain that f is strictly decreasing. □

END OF LECTURE 19

Proof. (ii) We have already shown that f is strictly monotone. Assume, WLOG (why?), that f is strictly increasing. Then, by IVT, we have that $J = [f(a), f(b)]$.

Case 1. $p \in (a,b)$. Let $q = f(p)$. Let $d = \frac{1}{2}\min\{p-a, b-p\} > 0$. Let $\varepsilon > 0$. Suppose $\varepsilon \leq d$. Then

$$a < p - \varepsilon < p < p + \varepsilon < b.$$

Thus, since f is strictly increasing, we have that

$$f(a) < f(p-\varepsilon) < q < f(p+\varepsilon) < f(b).$$

Let $\delta_\varepsilon = \min\{f(p+\varepsilon) - q, q - f(p-\varepsilon)\}$. Now, whenever $|y-q| < \delta$, we have that

$$f(p-\varepsilon) < q - \varepsilon < y < q + \varepsilon < f(p+\varepsilon).$$

Thus,

$$p - \varepsilon < f^{-1}(y) < p + \varepsilon.$$

If $\varepsilon > d$, choose δ_ε as δ_d. Then, whenever $|y-q| < \delta_d$, we have that $|f^{-1}(y) - f^{-1}(q)| < d < \varepsilon$.

Case 1. $p = a$ or $p = b$. Homework!

(iii) We prove the differentiability of f^{-1} at $q = f(p)$. Given k such that $q + k \in J$, Let $h(k) = f^{-1}(q+k) - f^{-1}(q)$. Then, $h(k) \neq 0$ whenever $k \neq 0$, (since f^{-1} is is one-to-one). Then $h(k) + p = f^{-1}(q+k)$, or $k = f(h(k)+p) - f(p)$. Thus,

$$\frac{f^{-1}(q+k) - f^{-1}(q)}{k} = \frac{h(k)}{f(p+h(k)) - f(p)} = \frac{1}{\frac{f(p+h(k))-f(p)}{h(k)}}.$$

By continuity of f^{-1}, $\lim_{k\to 0} h(k) = 0$. Since $h(k) \neq 0$, whenever $k \neq 0$, we have by the composition rule and the limit laws that

$$\lim_{k\to 0} \frac{f^{-1}(y+k) - f^{-1}(y)}{k} = \lim_{k\to 0} \frac{1}{\frac{f(x+h(k))-f(x)}{h(k)}} = \frac{1}{\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}} = \frac{1}{f'(x)}.$$

□

Example 19. The above theorem immediately gives the continuity and differentiability of $f(x) = x^{1/n}$ for $x > 0$ and $n \in \mathbb{P}$. Moreover, $f'(x) = \frac{1}{nx^{(n-1)/n}}$.

By the algebra of derivatives and induction (or the chain rule — which we haven't done yet!), one obtains the continuity and differentiability of $f(x) = \left(x^{1/q}\right)^p$, $x > 0$, for each $p \in \mathbb{Z}$ and $q \in \mathbb{P}$. Moreover, $f'(x) = \frac{p}{q}x^{(p-q)/q}$.

Example 20. We may now define inverse trigonometric functions. As convention, we invert the sine function on $[-\pi/2, \pi/2]$, the cosine function over $[0,\pi]$ and the tangent function over $(-\pi/2, \pi/2)$.

END OF LECTURE 20

Theorem 4.6. *Let $f : (a,b) \to \mathbb{R}$ and $g : (c,d) \to \mathbb{R}$ such that $f((a,b)) \subset (c,d)$. Say f is differentiable at p and g is differentiable at $q = f(p)$. Then, $g \circ f$ is differentiable at p, and*

$$(g\circ f)'(p) = g'(f(p)) \cdot f'(p).$$

Proof. Let $h = g \circ f$. Then, the relevant difference quotient is

$$\frac{h(p+k) - h(p)}{k} = \frac{g(f(p+k)) - g(f(p))}{k}.$$

Let $\ell(k) = f(p+k) - f(p)$. Note that ℓ is a continuous function of k. We have that

$$\frac{g(f(p+k)) - g(f(p))}{k} = \frac{g(q+\ell(k)) - g(q)}{f(p+k) - f(p)} \frac{f(p+k) - f(p)}{k},$$

where the last equality holds only if $f(p+k) \neq f(p)$. Idea: what if we could directly write $g(q + \ell(k)) - g(q)$ directly as $\ell(k)G(k)$, where $G(k)$ admits a limit as $k \to 0$?

Define

$$G(\ell) = \begin{cases} \frac{g(q+\ell)-g(q)}{\ell}, & \ell \neq 0, q+\ell \in (c,d) \\ g'(q), & \ell = 0. \end{cases}$$

Since $\lim_{\ell \to 0} G(\ell) = G(0)$, G is continuous at $\ell = 0$. Note that

$$g(q+\ell) - g(q) = \ell G(\ell).$$

Thus,

$$\frac{h(p+k)-h(p)}{k} = \frac{g(q+\ell)-g(q)}{k} = \frac{\ell(k)G(\ell(k))}{k}.$$

Now, $\lim_{k \to 0} \frac{\ell(k)}{k} = f'(p)$, by assumption, and $\lim_{k\to 0} G(\ell(k)) = G(0) = g'(q)$ by the continuity of ℓ and G at 0. Thus, we are done. $\square$

We will now see some applications of derivatives.

Definition 4.7. Let $f : A \to \mathbb{R}$. We say that f attains a local maximum (minimum) at $a \in A$ if there exists a $\delta_a > 0$ such that $f(a) \geq f(x)$ ($f(a) \leq f(x)$) for all $x \in A \cap N_{\delta_a}(a)$.

Theorem 4.8 (Extreme=>Critical)**.** *Let $f : (a,b) \to \mathbb{R}$. Let $c \in (a,b)$ such that f is differentiable at c and f attains a local extremum at c. Then, $f'(c) = 0$.*

Remarks. All the conditions are very important!

(1) If the extremum occurs at one of the endpoints of an interval, the function may have a nonzero derivative there. E.g. $f(x) = x$ when restricted to $[0,1]$ has local max and min on $[0,1]$ at 1 and 0 resp.

(2) f need not even be differentiable at points where local extrema are defined. E.g. $f(x) = |x|$ at $x = 0$.

(3) The converse of the above theorem is not true. Take $f(x) = x^3$. f does not attain a local extremum at $x = 0$, and yet the derivative vanishes there.

END OF LECTURE 21

Proof. We assume (WLOG) that f attains a local maximum at c. Then, there is a $\delta > 0$ such that $f(c) \geq f(x)$ for all $x \in N_\delta(c) \cap (a,b)$. We may further assume that $N_\delta(c) \subset (a,b)$.

By the sequential characterization of limits, we have that for any sequence $\{c_n\} \subset (a,b)$ such that $\lim_{n\to\infty} c_n = c$, we have that

$$\lim_{n\to\infty} \frac{f(c_n) - f(c)}{c_n - c} = f'(c).$$

Now, for $c - \delta < x < c$, we have that

$$\frac{f(x) - f(c)}{x - c} \geq 0.$$

Let $N \in \mathbb{N}$ be such that $c - \delta < c - 1/n$ for all $n \geq N$. Then, by the comparison lemma for sequences, we have that

$$\lim_{n \to \infty} \frac{f(c - 1/n) - f(c)}{1/n} = f'(c) \geq 0.$$

On the other hand, for $c < x < c + \delta$, we have that

$$\frac{f(x) - f(c)}{x - c} \leq 0.$$

Thus, using the sequence $\{c + 1/n\}$, we have that $f'(c) \leq 0$. Combining the two bounds, we have our result. $\square$

The theorem is most helpful in the following way: to identify **potential** points of local extrema within open intervals where f is differentiable. Thereafter, one needs to do some local analysis at the potential points. For the local analysis, one inspects the sign of the derivative 'just a bit before' and 'just a bit after' the point.

Example 21. Let $f : \mathbb{R} \to \mathbb{R}$ be given by

$$f(x) = (x - 1)^2 + |x| + x^3.$$

The potential local extrema are at $x = 0$ (point of non-differentiability), $x = -3$ and $x = \sqrt{2} - 1$. We need some more machinery to say which of these are points of extrema, and of what type.

Theorem 4.9 (Mean Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then, there is some $c \in (a, b)$ such that $f'(c) = \frac{f(b) - f(a)}{b - a}$.*

Proof. **Special case.** $f(a) = f(b)$**.** This is referred to as Rolle's theorem. By the extreme value theorem, there exist $c_1, c_2 \in [a, b]$ where f attains its global extrema on $[a, b]$. If at least one of c_1 or c_2 are in (a, b), we are done. If $f(a) = f(b) = c_1 = c_2$. Then, f is a constant function, in which case, $f' \equiv 0 =$ on $[a, b]$.

END OF LECTURE 22

General case. We want to apply Rolle's to a function G such that $G'(c) = 0 \iff f'(c) - \frac{f(b) - f(a)}{b - a} = 0$. Let $g(x) = f(x) - x\frac{f(b) - f(a)}{b - a}$. Then, g is cont. on $[a, b]$ and differentiable on (a, b). Moreover, $g(a) = bf(a) - af(b) = g(b)$. Thus, by Rolle's theorem, there is some $c \in (a, b)$ such that $g'(c) = 0$, i.e.

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

□

Applications of MVT

Theorem 4.10. *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Then*

(1) *if $f'(x) > 0$ for all $x \in (a,b)$, f is strictly increasing on $[a,b]$,*
(2) *if $f'(x) < 0$ for all $x \in (a,b)$, f is strictly decreasing on $[a,b]$,*
(3) *if $f'(x) = 0$ for all $x \in (a,b)$, f is a constant function on $[a,b]$.*

Proof. We prove only (1) and leave the rest as an exercise. Let $x, y \in [a,b]$ such that $x < y$. Apply the MVT on f restricted to $[x,y]$, we have the existence of some $z \in (x,y)$ such that

$$0 < f'(z) = \frac{f(y) - f(x)}{y - x}.$$

Thus, $f(y) > f(x)$. □

Remarks. (1) What if $f'(c) > 0$ or $f'(c) < 0$ for some $c \in (a,b)$?
(2) What about the converse of each part above? Remember that x^3 is a strictly increasing function.
(3) You are allowed to use the correct variations of the above result.

Example 22. The above theorem gives the **first derivative test** whose formal statement is given as Theorem 4.8 in Apostol. Let us return to $f(x) = (x-1)^2 + |x| + x^3/3$. Note that

$$f'(x) = \begin{cases} x^2 + 2x - 3 = (x-1)(x+3), & x < 0, \\ \left(x - (\sqrt{2} - 1)\right)\left(x + \sqrt{2} + 1\right), & x > 0. \end{cases}$$

Test at $x = -3$: $f'(x) > 0$ on $(-4,-3)$ and $f'(x) < 0$ on $(-3,0)$, so $f(x) \le f(-3)$ for all $x(-4,0)$. So, -3 is a point of local max.
Test at $x = 0$: $f'(x) < 0$ on $(-1,0)$, and $f'(x) < 0$ on $(0, \sqrt{2} - 1)$. So, $f(x) > f(0) > f(y)$ for every $x \in (-1,0)$ and $y \in (0, \sqrt{2} - 1)$. Thus, $x = 0$ is not a point of local max/min.
Test at $x = \sqrt{2} - 1$ Homework!

Definition 4.11. Let $f : (a,b) \to \mathbb{R}$ be a differentiable function such that $f' : (a,b) \to \mathbb{R}$ is also differentiable. Then, the $(f')'$ is denoted by f'' and is called the second derivative of f. One may defined the nth derivative of f inductively, and denote it by $f^{(n)}$. If $f^{(n)}$ exists for all $n \in \mathbb{N}$, we say that f is infinitely differentiable on (a,b)

Self-study: second derivative test

Many functions we encounter such as polynomials and trigonometric functions are infinitely differentiable.

END OF LECTURE 23

Definition 4.12. If a function f has an nth derivative at x_0, its nth Taylor polynomial at x_0 is the polynomial

$$P_n^{x_0}(x) = P_n(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)(x - c)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n.$$

P_n^c can be used to approximate the function f. It would be useful to under stand what the error term looks like. There are many known forms of the error term. We mention one.

Theorem 4.13 (Taylor's theorem)**.** *Let $f : (a,b) \to \mathbb{R}$ be $(n+1)$-times differentiable on (a,b). Let $x_0 \in (a,b)$. Then, for any $x \in (a,b)$, there exists a $c(x)$ between x and x_0 such that*

$$f(x) = P_n^{x_0}(x) + \frac{f^{(n+1)}(c(x))}{(n+1)!}(x - x_0)^{n+1}.$$

Idea: Want to apply Rolle's theorem to a function G such that

$$G'(c) = 0 \iff f^{(n+1)}(c) = (n+1)! \frac{f(x) - P_n^{x_0}(x)}{(x - x_0)^{n+1}}.$$

Could try

$$f^{(n)}(t) - t(n+1)! \frac{f(x) - P_n^{x_0}(x)}{(x - x_0)^{n+1}}.$$

But $G(y) \neq G(c)$. We will skip the proof, but work out an example.

Question from class. What if we wish to approximate f by P_{n+1} instead of P_n?

Answer. The best we can say is that there is a function $h = h_{x_0} : (a,b) \to \mathbb{R}$ such that $\lim_{x \to x_0} h(x) = 0$ and

$$f(x) = P_{n+1}^{x_0}(x) + h(x)(x - x_0)^{n+1} \tag{4.2}$$

for $x \in (a,b)$. Thus, we don't get as much information on the form of the error term as we do in Taylor's theorem. Now, if f was $(n+2)$-times differentiable, then by Taylor's theorem, h in (4.2) would become

$$h(x) = \frac{f^{(n+2)}(c(x))}{(n+2)!}(x - x_0)$$

Example 23. Let $f(x) = \cos(x)$. We compute the 3rd Taylor polynomials of f at $c = 0$. Note that $f(0) = 1$, $f'(0) = 0$ and $f''(0) = -1$. Then,

$$P_3(x) = 1 - \frac{1}{2!}x^2.$$

By Taylor's theorem, there is some y between 0 and x such that

$$\cos(x) = 1 - \frac{1}{2!}x^2 + \frac{\sin(y)}{3!}x^3.$$

But the error term is nonnegative on $[-\pi, \pi]$. Thus,

$$\cos x \geq 1 - \frac{x^2}{2}$$

on $[-\pi, \pi]$.

On trigonometric functions: read Section 2.5 in Apostol.

END OF LECTURE 24

5. INTEGRATION

Our goal will be to construct a “reasonable" theory of “area under a graph". There are two equivalent theories of integration that appear in introductory calculus books — that of Riemann integration and Darboux integration. We will follow Apostol and discuss Darboux integration. The word “reasonable" above refers to the fact that we want area, which is a function whose inputs are sets and outputs are real numbers, to satisfy certain rules (axioms) such as positivity, finite additivity, monotonicity under containments, etc. It is possible that for any such theory, the areas of certain sets just cannot be measured, i.e., the sets are not ‘measurable'! You will encoutner all of this more formally in a measure theory course, but we will see some indication of this in this course as well.

we will perform integration on a closed interval, i.e., a set of the form $[a, b]$, $a < b$.

Definition 5.1. (i) Let $a < b$. A partition of $[a, b]$ is a finite set

$$\mathscr{P} = \{x_0, \ldots, x_n\}$$

such that $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$. The partition $\mathscr{P}$ determines n subintervals, $[x_{j-1}, x_j]$, $1 \le j \le n$, of $[a, b]$. We refer to $[x_{j-1}, x_j]$ as the j^{th} interval of $\mathscr{P}$.

(ii) Given a partition $\mathscr{P}$ of $[a, b]$, a refinement of $\mathscr{P}$ is a partition $\mathscr{P}'$ such that, as sets, $\mathscr{P} \subseteq \mathscr{P}'$, i.e., every subinterval determined by $\mathscr{P}'$ is a subset of some subinterval determined by $\mathscr{P}$.

(iii) Given two partitions $\mathscr{P} = \{x_0 < x_1 < \cdots < x_n\}$ and $\mathscr{Q} = \{y_0 < y_1 < \cdots < y_m\}$ of $[a, b]$, their common refinement is the partition

$$\mathscr{R} = \mathscr{P} \cup \mathscr{Q} = \{z_0 < z_1 < \cdots < z_N\}.$$

Remark. It is a (tedious to prove) fact that for any $\ell \in \{1, \ldots, N\}$ there is a unique $j(\ell) \in \{1, \ldots, n\}$ and unique $k(\ell) \in \{1, \ldots, m\}$ such that

$$(z_{\ell-1}, z_\ell) = (x_{j(\ell)-1}, x_{j(\ell)}) \cap (y_{k(\ell)-1}, y_{k(\ell)}).$$

Definition 5.2. A function $f : [a, b] \to \mathbb{R}$ is called a step function if there is a partition $\mathscr{P} = \{x_0 < \cdots < x_n\}$ such that s is constant on each open subinterval of $\mathscr{P}$, i.e., for each $1 \le k \le n$, there is a real number s_k such that

$$s(x) = s_k \text{ for } x \in (x_{k-1}, x_k).$$

At the end points of these intervals, the function may be defined as anything.

Example 24. The floor and ceiling functions restricted to closed and bounded intervals. Inifnitely many steps are not allowed in the above definition.

Remark. The sum, difference, product and quotient (when defined) of two step functions, with partitions $\mathscr{P}$ and $\mathscr{Q}$ is also a step function with partition $\mathscr{P} \cup \mathscr{Q}$. We demonstrate this with a few pictures, since the proof is long and boring.

Definition 5.3. Let $s : [a,b] \to \mathbb{R}$ be a step function with partition $P = \{x_0 < ... < x_n\}$. Its integral is denoted by $\int_a^b s(x)dx$ and is given by

$$\sum_{k=1}^{n} s_k(x_k - x_{k-1}).$$

We also define $\int_b^a s(x)dx$ to be $-\int_a^b s(x)dx$.

END OF LECTURE 26

Remark. The partition of a step funciton is not uniquely determined. A refinement of $\mathscr{P}$ would also work, but it would leave the integral unaffected.

Theorems 1.2-1.8 state certain key properties of the above definitions such as

$$\int_a^b (c_1 s(x) + c_2 t(x))dx = c_1 \int_a^b s(x)dx + c_2 \int_a^b t(x)dx,$$

$$s \le t \Rightarrow \int_a^b s(x)dx \le \int_a^b t(x)dx,$$

and

$$\int_{ka}^{kb} s(x/k)dx = k \int_a^b s(x)dx.$$

These can formally proved using the hints listed in 1.15. See also Theorem 11.2.16 in Tao.

To integrate a general function, $f : [a,b] \to \mathbb{R}$, we want to approximate it with step functions (from above and below), i.e., we want to consider two classes of functions

Definition 5.4.

$$S_f = \{s : [a,b] \to \mathbb{R} : s \text{ is a step function and } s \le f\},$$
$$T_f = \{t : [a,b] \to \mathbb{R} : t \text{ is a step function and } t \ge f\}.$$

For an unbounded function, such as $1/x$ on $[0,1]$, the latter set is empty! So, we stick to bounded functions.

Lemma 5.5. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function, i.e., there is an $M > 0$ such that*

$$-M \le f(x) \le M \ \forall x \in [a,b].$$

Then, $\sup S_f$ *and* $\inf T_f$ *exist, and*

$$-M(b-a) \le \sup\left\{\int_a^b s(x)\,dx : s \in S_f\right\} \le \inf\left\{\int_a^b t(x)\,dx : t \in T_f\right\} \le M(b-a).$$

Proof. Let $g(x) = -M$, $x \in [a,b]$, and $h(x) = M$, $x \in [a,b]$. Then, $g \in S_f$ and $h \in T_f$. Moreover, $g \le t$ for all $t \in T_f$ and $h \ge s$ for all $s \in S_f$. Thus,

$$-M(b-a) = \int_a^b g(x)\,dx \le \int_a^b t(x)\,dx,\ \forall t \in T_f$$
$$M(b-a) = \int_a^b h(x)\,dx \ge \int_a^b s(x)\,dx,\ \forall s \in S_f.$$

□

END OF LECTURE 27

Definition 5.6. Let $f : [a,b] \to \mathbb{R}$ be a bounded function. The lower integral of f is the quantity

$$\underline{I}(f) = \sup\left\{\int_a^b s(x)\,dx : s \in S_f\right\},$$

and the upper integral of f is the quantity

$$\overline{I}(f) = \inf\left\{\int_a^b t(x)\,dx : t \in T_f\right\}.$$

We say that f is Riemann integrable if $\underline{I}(f) = \overline{I}(f)$, in which case this quantity is called the integral of f over $[a,b]$ and denoted by $\int_a^b f(x)\,dx$. We set

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx,$$
$$\int_a^a f(x)\,dx = 0.$$

Example 25. Show that every step function is integrable and there is no ambiguity in using the symbol $\int_a^b s(x)\,dx$ for a step function.

Theorem 5.7. *Every bounded monotone function on* $[a,b]$ *is Riemann integrable.*

Proof. (WLOG?) Assume f is increasing. Let $n \in \mathbb{N}$. Let

$$x_j = a + j\frac{b-a}{n},\ j = 1,\dots,n.$$

Let s,t be step functions on $[a,b]$ given by

$$\begin{aligned} s_n(x) &= f(x_{j-1}),\ x \in [x_{j-1}, x_j), \\ t_n(x) &= f(x_j),\ x \in [x_{j-1}, x_j). \end{aligned}$$

Let $s_n(b) = f(b) = t_n(b)$. Note that

$$\int_a^b t_n(x)dx - \int_a^b s_n(x)dx = (f(b) - f(a))\frac{b-a}{n}.$$

Note

$$\int_a^b s_n(x)dx \le \underline{I}(f) \le \overline{I}(f) \le \int_a^b t_n(x)dx.$$

Letting $C = (f(b) - f(a))(b-a)$, we have that

$$\overline{I}(f) - \underline{I}(f) \le \frac{C}{n}.$$

Taking $n \to \infty$, we obtain the integrability of f. □

Example 26. We claim that for $b > 0$,

$$\int_0^b x^p dx = \frac{b^{p+1}}{p+1}.$$

Since x^n is monotone on $[0, b]$, it is Riemann integrable. Motivated from the previous proof, let $x_\ell = \ell b/n$, $\ell = 0, 1, ..., n$. Then,

$$\frac{b}{n}\sum_{k=0}^{n-1}\left(\frac{kb}{n}\right)^p \le I(x^p) \le \frac{b}{n}\sum_{k=1}^{n}\left(\frac{kb}{n}\right)^p.$$

I.e.,

$$\left(\frac{b}{n}\right)^{p+1}\left(\sum_{k=0}^{n-1} k^p\right) \le I(x^p) \le \left(\frac{b}{n}\right)^{p+1}\left(\sum_{k=1}^{n} k^p\right).$$

Now, we use that

$$\lim_{n\to\infty}\frac{1^p + \cdots + (n-1)^p}{n^{p+1}} = \frac{1}{p+1}.$$

This follows from the inequalities

$$1^p + \cdots + (n-1)^p < \frac{n^{p+1}}{p+1} < 1^p + \cdots + n^p,$$

which can be proved by induction (on n).

END OF LECTURE 28

We will take the following properties of Riemann integration for granted: linearity with respect to integrand (Thm. 1.16); additivity with respect to interval of integration (Thm 1.17); invariance under translation (Thm 1.18); epxansion/contraction of interval of integration (Thm 1.19); comparison (Thm 1.20). Try to prove these for extra practice, or see Section 1.27 for the proofs.

Definition 5.8. A function $f : A \to \mathbb{R}$ is s.t.b. uniformly continuous on A if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $|x - y| < \delta$, $x, y \in A$, we have that $|f(x) - f(y)| < \varepsilon$.

Remarks. (1) Uniform cont $\Rightarrow$ cont.
(2) If f is u.c. on A, then it must be u.c. on every $B \subseteq A$.

Example 27. (1) $f(x) = x$ is uniformly continuous on $\mathbb{R}$. Note that $\delta = \varepsilon$ works for each $\varepsilon > 0$.

(2) $f(x) = x^2$ is **not** uniformly continuous on $\mathbb{R}$. We need to produce an $\varepsilon > 0$ such that for every δ, there exist x_δ and y_δ such that $|x_\delta - y_\delta| < \delta$, but $|f(x_\delta) - f(y_\delta)| \geq \varepsilon$.

Choose $\varepsilon = 1$. Let $\delta > 0$. Let $x_\delta = \frac{1}{\delta}$ and $y_\delta = \frac{1}{\delta} + \frac{\delta}{2}$. Then, $|x_\delta - y_\delta| < \delta$, but

$$|f(x_\delta) - f(y_\delta)| = 1 + \delta^2/4 > 1.$$

FACT. Every continuous function on a closed and bounded interval is, in fact, uniformly continuous! We will not prove this in class, however, this can be proved using a similar technique as was used for showing that continuous functions on closed and bounded intervals are bounded.

Theorem 5.9. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then, f is Riemann integrable.*

Proof. Since f is continuous on $[a, b]$, it is bounded. Thus, $\underline{I}(f)$ and $\overline{I}(f)$ exist.

Let $\varepsilon > 0$. By the uniform continuity of f on $[a, b]$, there exists a $\delta > 0$ such that whenever $|x - y| < \delta$ for $x, y \in [a, b]$, we have that $|f(x) - f(y)| < \varepsilon$. Let $n \in \mathbb{N}$ such that $\frac{b-a}{n} < \delta$. As done in previous proofs, choose

$$x_j = a + j\frac{b-a}{n}, \quad j = 0, \ldots, n.$$

Now, $f|_{[x_{j-1}, x_j]}$ is a continuous, therefore bounded, function that attains its max. value, M_j, at some $u_j \in [x_{j-1}, x_j]$ and min. value, m_j, at some $l_j \in [x_{j-1}, x_j]$. Thus, on $[x_{j-1}, x_j]$,

$$m_j \leq f \leq M_j.$$

Moreover, since $|u_j - l_j| < \delta$, we have that $M_j - m_j < \varepsilon$.

Now, set

$$\begin{aligned} s_n(x) &= m_j, \quad x \in [x_{j-1}, x_j) \\ s_n(b) &= f(b). \end{aligned}$$

and

$$\begin{aligned} t_n(x) &= M_j, \quad x \in [x_{j-1}, x_j) \\ t_n(b) &= f(b). \end{aligned}$$

Then,

$$\int_a^b s(x)dx \le \underline{I}(f) \le \overline{I}(f) \le \int_a^b t(x)dx.$$

Thus,

$$0 \le \overline{I}(f) - \underline{I}(f) \le \frac{b-a}{n} \sum_{j=1}^{n} M_j - m_j < \varepsilon(b-a).$$

Since $\varepsilon > 0$ was arbitrary, we are done! □

END OF LECTURE 29

Theorem 5.10 (M integrals)**.** *Let $f : [a,b] \to \mathbb{R}$ be continuous. Then, there is some $c \in [a,b]$ such that*

$$f(c)(b-a) = \int_a^b f(x)dx.$$

Proof. Let $M = \max f$ and $m = \min f$. Then,

$$m \le \frac{1}{b-a} \int_a^b f(x)dx \le M.$$

Thus, by IVT, there is some $c \in [a,b]$ such that $f(c)$ attains the claimed value. □

Theorem 5.11. *Let $a < b$. Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable. Then, so is $|f|$, and*

$$\left|\int_a^b f(x)dx\right| = \left|\int_b^a f(x)dx\right| \le \int_a^b |f(x)|dx.$$

We will now relate the notions of integrability and differentiability.

Theorem 5.12 (The first Fundamental Theorem of Calculus)**.** *Let $f : [a,b] \to \mathbb{R}$ be a Riemann integrable function. Let $F : [a,b] \to \mathbb{R}$ be the function given by*

$$F(x) = \int_a^x f(t)dt.$$

Then, F is a continuous function on $[a,b]$. Moreover, if f is continuous at $p \in (a,b)$, then F is differentiable at p, and $F'(p) = f(p)$.

Proof. Since f is bounded, there exists an $M > 0$ such that $-M \le f(x) \le M$ for all $x \in [a,b]$. Now, let $x, y \in [a,b]$ such that $x < y$. Then,

$$F(y) - F(x) = \int_x^y f(t)dt \le M(y-x).$$

Similarly, $-M(y-x) \le F(y) - F(x)$. Thus,

$$|F(y) - F(x)| \le M|y-x|.$$

Interchanging the role of x and y, we get the same conclusion if $y < x$. You will show in the next assignment that the above condition implies continuity of F.

Now, we assume that f is continuous at some $p \in (a, b)$. Let $h \in \mathbb{R}$ such that $p + h \in (a, b)$. Then,

$$\begin{aligned}\frac{F(p+h)-F(p)}{h} &= \frac{1}{h}\int_{p}^{p+h} f(x)dx \\ &= \frac{1}{h}\int_{p}^{p+h} f(p)+f(x)-f(p)dx \\ &= f(p)+\frac{1}{h}\int_{p}^{p+h}(f(x)-f(p))dx \\ &= f(p)+G(h).\end{aligned}$$

We wish to show that $\lim_{h\to 0} G(h) = 0$, in other words, given $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $0 < |h| < \delta$, we have that $|G(h)| < \varepsilon$.

Let $\varepsilon > 0$. By the continuity of f at p, there exists a $\delta > 0$ such that whenever $x \in (a, b)$ and $|x - p| < \delta$, we have that $|f(x) - f(p)| < \varepsilon$. Thus, using the previous theorem

$$\begin{aligned}\left|\frac{1}{h}\int_{p}^{p+h}(f(x)-f(p))dx\right| &= \begin{cases}\frac{1}{h}\left|\int_{p}^{p+h} f(x)-f(p)dx\right|, \ h>0, \\ -\frac{1}{h}\left|\int_{p}^{p+h} f(x)-f(p)dx\right|, \ h<0,\end{cases} \\ &\leq \begin{cases}\frac{1}{h}\int_{p}^{p+h}\left|f(x)-f(p)\right|dx, \ h>0, \\ -\frac{1}{h}\int_{p+h}^{p}\left|f(x)-f(p)\right|dx, \ h<0,\end{cases} \quad < \varepsilon.\end{aligned}$$

□

Remarks. (1) If f is not continuous at p, that does not necessarily imply that F is not differentiable at p. For instance, in the next assignment, you will show that if

$$f(x) = \begin{cases}0, \ x \in [-1,0) \cup (0,1], \\ 1, \ x = 0,\end{cases}$$

then f is integrable,r and $F(x) = 0$ for all $x \in [-1, 1]$. However, $F'(0) \neq f(1)$. On the other hand, consider

$$f(x) = \begin{cases}0, \ x \in [-1,0), \\ 1, \ x \in [0,1],\end{cases}$$

Then,

$$F(x) = \begin{cases}and 0, \ x \in [-1,0), \\ x, \ x \in [0,1],\end{cases}$$

which is simply not differentiable at 0.

(2) The function F above is sometimes called an indefinite integral of f. If f is nonnegative, then F is increasing, since

$$F(y) - F(x) = \int_x^y f(x)dx \geq 0.$$

If f is increasing, then F inherits the property of "convexity". We will not discuss this in this course, but you may read Section 2.18 from Apostol.

(3) One could have also defined

$$G(x) = \int_b^x f(t)dt.$$

Then, $F(x) - G(x) = \int_a^b f(t)dt$, which is a constant. Thus, G has the same continuity and differentiability properties of F!

Definition 5.13. Let $f : (a,b) \to \mathbb{R}$ be a function. A differentiable function $F : (a,b) \to \mathbb{R}$ is called a primitive/anti-derivative of f on (a,b) if $F' = f$ on (a,bW).

Remark. If F and G are two primitives of f on (a,b), then $(F-G)' = 0$. Thus, $F = G + C$ for some constant C.

Example 28. $\sin(x)$ and $\cos(x)$ are primitives of $\cos(x)$ and $-\sin(x)$, respectively. $\frac{x^{r+1}}{r+1}$ is a primitive of x^r on $(0,\infty)$ for all $r \in \mathbb{R} \setminus \{1\}$.

By the First FTOC, we also know that

$$F(x) = \int_1^x \frac{1}{t}dt, \quad x > 0,$$

is a primitive of $f(x) = \frac{1}{x}$ on $(0,\infty)$.

Theorem 5.14 (The second Fundamental Theorem of Calculus)**.** *Let I be an open interval, and $f : I \to \mathbb{R}$ be a function that admits an anti-derivative F on I. Let $[a,b] \subseteq I$, and suppose $f\Big|_{[a,b]}$ is Riemann integrable. Then,*

$$\int_a^b f(x)dx = F(b) - F(a).$$

END OF LECTURE 31

Proof. We will show that

$$I - \varepsilon \leq F(b) - F(a) \leq I + \varepsilon$$

for all $\varepsilon > 0$.

Let $\varepsilon > 0$. By the integrability of f, there exist $s_\varepsilon \in S_f$ and $t_\varepsilon \in T_f$ such that

$$\int_a^b s_\varepsilon(x)dx > I - \varepsilon,$$
$$\int_a^b t_\varepsilon(x)dx < I + \varepsilon.$$

By taking common refinements, we may assume that s_ε and t_ε are step functions with respect to the same partition $P = \{x_0 < \cdots < x_n\}$, i.e., there exist $s_1, ..., s_n \in \mathbb{R}$ and $t_1, ..., t_n \in \mathbb{R}$ such that $s_\varepsilon(x) = s_j$ and $t_\varepsilon(x) = t_j$ for all $x \in (x_{j-1}, x_j)$.

Now, by MVT, there exist $c_j \in (x_{j-1}, x_j)$ such that

$$F(b) - F(a) = \sum_{j=1}^n F(x_j) - F(x_{j-1}) = \sum_{j=1}^n F'(c_j)(x_j - x_{j-1}) = \sum_{j=1}^n f(c_j)(x_j - x_{j-1}).$$

But, $s_j = s_\varepsilon(c_j) \le f(c_j) \le t_\varepsilon(c_j) = t_j$ for all $j = 1, ..., n$. Thus,

$$\begin{aligned} I - \varepsilon \le \int_a^b s_\varepsilon(x)dx = \sum_{j=1}^n s(c_j)(x_j - x_{j-1}) &\le \sum_{j=1}^n f(c_j)(x_j - x_{j-1}) \\ &\le \sum_{j=1}^n t(c_j)(x_j - x_{j-1}) = \int_a^b t_\varepsilon(x)dx < I + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we are done. □

Remark. We can now write formulas such as

$$\int_a^b \frac{1}{x^{1/3}}dx = \int_a^b x^{-1/3}dx = \frac{3b^{2/3}}{2} - \frac{3a^{2/3}}{2}.$$

Remark (Leibniz' notation)**.** Leibniz used the notation $\int f(x)dx$ to denote a general primitive of f, e.g., the equation

$$\int \cos(x)dx = \sin(x) + C$$

merely means that the derivative of $\sin(x)$ is $\cos(x)$. This symbol (without limits) has nothing to do with integration a priori. You should simply read "$\int f(x)dx$" as "the general primitive of f" The two fundamental theorems tell us that:

1. For a continuous f, $\int f(x)dx = \int_a^x f(t)dt + C$.
2. $\int_a^b f(x)dx = \int f(x)dx\Big|_a^b$

The FTOC's allow us to convert theorems about differentiabilty to theorems about integrability. For instance, suppose $F' = f$. We know that $\left[F \circ g\right]'(x) = F'(g(x))g'(x) = f(g(x))g'(x)$. Not worrying about integrability for now, this should tell us that

$$\int_a^b f(g(x))g'(x)\,dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} F'(x)\,dx = \int_{g(a)}^{g(b)} f(x)\,dx.$$

This is precisely integration by substitution!

Theorem 5.15 (Integration by substitution)**.** *Let I be an open interval, and $g : I \to \mathbb{R}$ be a function that admits a continuous derivative on I. Let J be an open interval containing $g(I)$, and $f : J \to \mathbb{R}$ be a continuous function. Then, for an $[a,b] \subseteq I$,*

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(x)\,dx.$$

Proof. We will find a primitive of $f(g(x))g'(x)$ on $[a,b]$. Let

$$F(x) = \int_{g(a)}^{x} f(t)\,dt, \quad x \in J.$$

The above integral exist due the continuity of f of J. By the FFTOC, we have that F is differentiable on J, and

$$F'(x) = f(x), \quad x \in J.$$

Since $g(I) \subset J$, we have that $F \circ g$ is well-defined and differentiable on I. By the chain rule,

$$(F \circ g)'(x) = F'(g(x))g'(x) = f(g(x))g'(x), \quad x \in I.$$

for all $x \in [a,b]$. Thus, $F \circ g$ is an anti-derivative of $(f \circ g)g'$ on I. By the SFTOC, applied twice,

$$\int_a^b f(g(x))g'(x)\,dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(x)\,dx.$$

□

One can similarly obtain the formula for Integration by Parts as a consequence of the Fundamental Theorems of Calculus and the Product Rule for derivatives, but we will skip that.

END OF LECTURE 32

5.1. **Logarithm and exponentiation.** For any $b > 0$, we wish to define $\log_b(x)$ as that real number z such that

$$b^z = x. \tag{5.1}$$

However, in this course, we are yet to properly define what b^z means if z is an irrational number. Moreover, why should such a z exist? Why should it be unique. Let us extract the key features of the desired log function from (5.1).

Say, $z =$ "$\log_b(x)$" and $w =$ "$\log_b(y)$", i.e., $b^z = x$ and $b^w = y$. Then, if powers are to behave as expected, $b^{z+w} = xy$, i.e., $\log_b(x) + \log_b(y) = z + w =$ "$\log_b(xy)$". Thus, we seek a function ℓ such that

$$\ell(xy) = \ell(x) + \ell(y)$$

for all x, y in its domain. Some observations based on this *functional equation*:

- If $0 \in \text{dom}(\ell)$, then taking $x = y = 0$, we have that $\ell(0) = 2\ell(0) \Rightarrow \ell(0) = 0$. Taking only $y = 0$, we have that $\ell(0) = \ell(x)$ for all $x \in \text{dom}(\ell)$. Thus, we only get the constant zero function.
- If $1, -1 \in \text{dom}(\ell)$, then taking $x = y = 1$ gives that

  $$\ell(1) = 0,$$

  and taking $x = y = -1$ yields that $\ell(-1) = 0$.
- Now, if $x, -x, 1, -1 \in \text{dom}(\ell)$, taking $y = -1$ yields that $f(x) = f(-x)$. Thus, f is an even function on its domain.
- Assuming ℓ is differentiable, taking y as a constant and differentiating in x, we obtain that

  $$y\ell'(yx) = \ell'(x).$$

  Evaluating at $x = 1$ yields that

  $$\ell'(y) = \frac{\ell'(1)}{y}.$$

  By FTOC, we would obtain that

  $$\ell(y) = \begin{cases} \ell(y) - \ell(1) = \ell'(1)\int_1^y \frac{dt}{t}, & y > 0, \\ \ell(y) - \ell(-1) = \ell'(1)\int_1^{-y} \frac{dt}{t}, & y < 0. \end{cases}$$

  The choice $\ell'(1)$ gives the trivial solution. For any non-zero choice of $\ell'(1)$, $\ell(x)/\ell'(1)$ gives yet another function that satisfies the functional equation (f.e.)!

Definition 5.16. Let $x > 0$. The *natural logarithm of* x is defined as

$$\ln(x) = \int_1^x \frac{dt}{t}.$$

Theorem 5.17. *The function* $\ln : \mathbb{R} \to$ *has the following properties:*

- $\ln(1) = 0$,
- $\ln(xy) = \ln(x) + \ln(y)$, $x, y > 0$,
- $\ln$ *is increasing and continuous on* $(0, \infty)$
- $\ln(x)$ *is differentiable on* $(0, \infty)$, *and* $\ln'(x) = \frac{1}{x}$, $x > 0$,
- *(in Leibniz notation)* $\int \frac{dt}{t} = \ln|t| + C$, $t \neq 0$,

- *(Leibniz)* $\int \ln(x)\,dx = x\ln(x) - x + C$,
- $\ln$ *is one-to-one and onto* $\mathbb{R}$.

Proof. □

Definition 5.18. Let e denote the unique positive number such that $\ln(e) = 1$. For any $x \in \mathbb{R}$, let $\exp(x)$ denote the unique positive number y such that $\ln(y) = x$, i.e., the function $\exp(x)$ is the inverse of $\ln(x)$.

The constant e was defined earlier using an infinite series. This turns out to be the same e!

Theorem 5.19. *The function* $\exp : \mathbb{R} \to \mathbb{R}$ *has the following properties:*

- $\exp(0) = 1$,
- $\exp(x + y) = \exp(x)\exp(y)$, $x, y \in \mathbb{R}$,
- $\exp$ *is increasing and continuous on* $\mathbb{R}$,
- $\exp$ *is differentiable on* $\mathbb{R}$ *and* $\exp'(x) = \exp(x)$, $x \in \mathbb{R}$,
- *(Leibniz)* $\int \exp(x)\,dx = \exp(x) + C$,
- $\exp$ *is one-to-one and onto* $(0, \infty)$
- $\exp(r) = e^r$ *for every rational* r.

Proof. □

Definition 5.20. Let $a > 0$ and $x \in \mathbb{R}$. Define

$$a^x = \exp(x \ln a).$$

END OF LECTURE 33