MATH600: Analysis I

Jonathan Ma
johnma@udel.edu

March 10, 2025

Contents

1 Week 1 **4**
1.1 Lecture 1. Wed, Aug 28 4
1.1.1 Orderings, Upper/Lower Bounds 4
1.2 Lecture 2. Fri, Aug 30 4
1.2.1 Ring/Field Axioms 5

2 Week 2 **5**
2.1 Lecture 3. Wed, Sep 4 5
2.1.1 Basic Point-set Topology 5
2.2 Lecture 4. Fri, Sep 6 6
2.2.1 Countability 6

3 Week 3 **7**
3.1 Lecture 5. Mon, Sep 9 7
3.1.1 Metric Spaces 7
3.2 Lecture 6. Wed, Sep 11 9
3.2.1 Open and Closed Sets 9
3.3 Lecture 7. Fri, Sep 13 10
3.3.1 Compactness 10

4 Week 4 **11**
4.1 Lecture 8. Mon, Sep 16 11
4.1.1 Sequential Compactness 11
4.2 Lecture 9. Wed, Sep 18 12
4.2.1 Heine-Borel 12
4.2.2 Separability, Connectedness 13
4.3 Lecture 10. Fri, Sep 20 14
4.3.1 Alternative Proof for Uncountability of Reals 15

5 Week 5 **15**
5.1 Lecture 11. 15
5.1.1 Sequential Convergence 16
5.2 Lecture 12. 17
5.2.1 More on Sequential Convergence 17
5.3 Lecture 13. 19

5.3.1 Bolzano-Weierstrauss . . . 19
5.3.2 Cauchy Sequences, Completeness . . . 20

6 Week 6 **21**
6.1 Lecture 14. . . . 21
6.1.1 Monotone Convergence . . . 21
6.2 Lecture 15. . . . 23
6.2.1 Subsequential Limits, Limit Inferiors / Superiors . . . 23
6.2.2 Series . . . 25
6.3 Lecture 16. . . . 26
6.3.1 Convergence of Series . . . 26

7 Week 7 **29**
7.1 Lecture 17. Wed Oct 9 . . . 29
7.1.1 Series Convergence Tests . . . 29
7.2 Lecture 18. Fri Oct 11 . . . 31
7.2.1 Absolute Convergence of Series . . . 33

8 Week 8 **35**
8.1 Lecture 19. . . . 35
8.1.1 Products of Absolutely Convergent Series . . . 35
8.1.2 Power Series . . . 36
8.2 Lecture 20. Wed Oct 16 . . . 38
8.2.1 Continuity . . . 38
8.3 Lecture 21. Fri Oct 18 . . . 40
8.3.1 Continuity and Compactness . . . 40

9 Week 9 **42**
9.1 Lecture 22. Tue Oct 22 . . . 42
9.1.1 Uniform Continuity . . . 42
9.1.2 Intermediate Value Theorem . . . 43
9.2 Lecture 23. Wed Oct 23 . . . 45
9.2.1 Monotone Convergence Theorem . . . 45
9.2.2 Differentiation . . . 47
9.3 Lecture 24. Thu Oct 24 . . . 47
9.3.1 More on Differentiation . . . 47
9.3.2 Local Extrema, Stationary Points, Fermat's Theorem . . . 49

10 Week 10 **50**
10.1 Lecture 25. Mon Oct 28 . . . 50
10.1.1 Mean Value Theorem . . . 50
10.1.2 L'Hôpital's Rule . . . 51
10.2 Lecture 26. Wed Oct 30 . . . 52
10.2.1 Taylor's Theorem . . . 52
10.3 Lecture 27 . . . 55
10.3.1 Riemann Integrals . . . 55
10.4 Lecture 28. Wed Nov 6 . . . 58
10.4.1 Conditions for Riemann Integrability . . . 58

10.5 Lecture 29. Fri Nov 8 . . . 61
10.5.1 Properties of Riemann Integration . . . 61

11 Week 11 **65**
11.1 Lecture 30. Mon Nov 11 . . . 65
11.1.1 Fundamental Theorem of Calculus . . . 65
11.2 Lecture 31. Wed Nov 13 . . . 68
11.2.1 Integration By Parts . . . 68
11.2.2 Sequences of Functions . . . 71
11.3 Lecture 32. . . . 72
11.3.1 Uniform Convergence . . . 72
11.3.2 Weierstrauss M-test . . . 75

12 Week 12 **75**
12.1 Lecture 33. Mon Nov 18 . . . 75
12.1.1 Sequences of Functions Applied to Power Series . . . 77
12.2 Lecture 34. Wed Nov 20 . . . 79
12.2.1 Equicontinuity, Pointwise Boundedness . . . 79
12.2.2 Arzela-Ascoli Theorem . . . 81
12.3 Lecture 35. Fri Nov 22 . . . 82
12.3.1 Functions of Several Variables . . . 82

13 Week 13 **86**
13.1 Lecture 36. Mon Dec 2 . . . 86
13.1.1 Operator Norms . . . 86
13.1.2 Chain Rule for Functions of Several Variables . . . 88
13.2 Lecture 37. Wed Dec 4 . . . 89
13.2.1 Contractive Mappings . . . 89
13.2.2 Implicit Function Theorem . . . 90
13.2.3 Partial Derivatives . . . 90
13.3 Lecture 38. . . . 91

14 Week 14 **94**
14.1 Lecture 39. . . . 94
14.1.1 Inverse Function Theorem . . . 95

15 Preliminary Review **97**
15.1 Vector Calculus . . . 97
15.1.1 Line Integrals . . . 98
15.1.2 Line Integrals of Vector Fields . . . 100
15.1.3 Fundamental Theorem of Line Integrals . . . 101
15.1.4 Green's Theorem . . . 103
15.1.5 Divergence, Curl . . . 104
15.1.6 Vector Form of Green's Theorem . . . 106
15.2 Surface Integrals, Flux . . . 107
15.2.1 Parametric Surfaces . . . 107
15.2.2 Surface Area . . . 107
15.2.3 Surface Integrals . . . 108

15.2.4 Surface Integrals of Vector Fields 110
15.2.5 Stoke's Theorem 111
15.2.6 Divergence Theorem 112

Preface

Analysis is a mathematical discipline that seeks to understand certain collections of mathematical objects in a rigorous and precise fashion.

1 Week 1

1.1 Lecture 1. Wed, Aug 28

1.1.1 Orderings, Upper/Lower Bounds

Definition 1.1 (Ordered set)**.** Let S be a set. An order on S is a binary relation $<$ satisfying

1. $\forall\, x, y \in S$, exactly one of $x < y$ or $x = y$ or $y < x$ holds (trichotomy),
2. if $x, y, x \in S$, then $((x < y) \wedge (y < z) \Rightarrow (x < z))$ (transitivity).

Definition 1.2. Let S be an ordered set. If $E \subset S$ and there exists $\beta \in S$ such that $x \leq \beta$, for all $x \in E$, then we say E is bounded above, and that β is an upper bound for E. Also, if $\beta \leq \alpha$ for every upper bound α of E, then we call β a (or the) supremum of E.

Definition 1.3. See also bounded below, lower bound, and infimum.

Definition 1.4 (Least-upper-bound property)**.** Let S be an ordered set. We say S has the least-upper-bound property if and only if every non empty subset of S that is bounded above has a least upper bound.

Theorem 1.5. *Let S be an ordered set that satisfies the LUB property, and let $E \subset S$ be a nonempty set that is bounded below. If L is the set of all lower bounds of E, $\alpha = \sup L$ exists, and $\alpha = \inf E$.*

Proof. Since E is bounded below, $L \neq \emptyset$. By definition of L, for any $y \in L$, and any $x \in E$, $y \leq x$. Equivalently, for any $x \in E$, and any $y \in L$, $y \leq x$, which implies that x shall be an upper bound of L. Since $E \neq \emptyset$, we can conclude at least one such x exists. Hence, L is nonempty and bounded above, and it is a subset of S, so L has a supremum, call it α. We claim $\alpha = \inf E$. Suppose α is not the infimum of E. Then there exists a lower bound β, with $\alpha < \beta$. But then β is also an element of L, and this implies α is not the supremum of L. □

Corollary 1.5.1. *An ordered set with the LUB property also has the GLB property.*

1.2 Lecture 2. Fri, Aug 30

Definition 1.6. Let S be an ordered set. We say $E \subset S$ is well ordered iff every nonempty subset A of E has a least element. $(\forall A \subset E, A \neq \emptyset \Rightarrow (\exists x \in A, \forall y \in A, x \leq y))$.

1.2.1 Ring/Field Axioms

Theorem 1.7. *Let R be an ordered ring. Then*

1. $\forall \alpha \in R, \alpha > 0 \Leftrightarrow -\alpha < 0$,
2. $\forall \alpha, \beta, \gamma \in R, (\alpha > 0 \wedge \beta < \gamma) \Rightarrow \alpha\beta < \alpha\gamma$,
3. $\forall \alpha, \beta, \gamma \in R, (\alpha < 0 \wedge \beta < \gamma) \Rightarrow \alpha\gamma < \alpha\beta$,
4. $\forall \alpha, \beta, \gamma, \delta \in R, (\alpha < \beta \wedge \gamma < \delta \wedge \alpha, \beta, \gamma, \delta > 0) \Rightarrow \alpha\gamma < \beta\delta$,
5. $\alpha \neq 0 \Rightarrow \alpha^2 > 0$,
6. *if R is a ring with unity,* $1 > 0$,
7. *if R is a field,* $\forall \alpha, \beta \in R$, $(\alpha > 0 \wedge \beta > 0) \Rightarrow (\alpha^{-1} > 0 \wedge \beta^{-1} > 0)$. *Also,* $\alpha < \beta \Rightarrow \beta^{-1} < \alpha^{-1}$.

Proposition 1.8. *The following are the foundational assumptions of the course:*

- *$\mathbb{Z}$ is an ordered integral domain in which $\mathbb{Z}^+$ is well ordered.(Also, if D is an ordered integral domain in which D^+ is well ordered, then D is isomorphic to $\mathbb{Z}$.)*
- *$\mathbb{R}$ is an ordered field that satisfies the least upper bound property. (Also, if F is an ordered field satisfying the LUB property, then F is isomorphic to $\mathbb{R}$.)*
- *It's not as easy to characterize $\mathbb{Q}$, but in algebra, we construct $\mathbb{Q}$ to be the field of fractions of $\mathbb{Z}$. Also, it is the smallest field containing $\mathbb{Z}$.*
- *We have that $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.*

Theorem 1.9 (Archimedean property of $\mathbb{R}$). *If $x \in \mathbb{R}^+$, and $y \in \mathbb{R}$, then there exists $n \in \mathbb{Z}^+$ such that $nx > y$.*

Proof. Assume that $x \in \mathbb{R}^+$, and $y \in \mathbb{R}$, and assume that $nx < y$ for all $n \in \mathbb{Z}^+$. It follows that $E = \{nx \mid n \in \mathbb{Z}^+\}$ is bounded above, and obviously, $E \neq \emptyset$. Thus, $\alpha = \sup E$ exists, and $\alpha \in \mathbb{R}$. Then $\alpha - x$ is not an upper bound for E, since $x > 0$. Hence, there exists $n \in \mathbb{Z}^+$ such that $nx > \alpha - x$. But then $nx > \alpha - x \Rightarrow (n+1)x > \alpha$. Therefore, α is not an upper bound for E, since $(n+1)x \in E$. □

Theorem 1.10. *Let $x, y \in \mathbb{R}$ with $x < y$. Then there exists $r \in \mathbb{Q}$ such that $x < r < y$.*

Proof. We have $x < y \Rightarrow y - x > 0$. Therefore, there must be some $n \in \mathbb{Z}^+$ such that $n(y-x) > 1$. Hence, $\frac{1}{n} < y - x$. Let m be the largest integer such that $\frac{m}{n} \leq x$. Consider $r = \frac{m+1}{n}$. Then $r > x$ by definition of m. □

2 Week 2

2.1 Lecture 3. Wed, Sep 4

2.1.1 Basic Point-set Topology

Definition 2.1. Let A, B be sets. A function $f\colon A \to B$ is (informally) a rule that associates to each $a \in A$ a unique $b = f(a) \in B$. More formally, a function $f\colon A \to B$ is a subset R of $A \times B = \{(a,b) \mid a \in A, b \in B\}$, with the property that for each $a \in A$, there exists a unique pair $(a,b) \in R$ that we write $b = f(a)$. We call A the domain, and B the codomain of f. The range of f is $R(f) = \{f(a) \mid a \in A\}$.

Definition 2.2. Let A, B be sets and let $f\colon A \to B$ be a function. We say that f is injective iff $(a_1, a_2 \in A \wedge f(a_1) = f(a_2)) \Rightarrow a_1 = a_2$. We say that f is surjective iff $\forall b \in B, \exists a \in A, f(a) = b$. We say that f is bijective iff it is injective and surjective.

Definition 2.3. Given a subset S of A. The image of $f\colon A \to B$ under f is denoted $f(S) = \{f(s) \mid s \in S\}$. Note that $R(f) = f(A)$.

Definition 2.4. Given $T \subset B$, the inverse image of T under f is $f^{-1}(T) = \{a \in A \mid f(a) \in T\}$.

Definition 2.5. Let A, B be sets and $f\colon A \to B$ a function. We say is invertible iff there exists $g : B \to A$ such that $f(g(b)) = b, \forall b \in B$, and $g(f(a)) = a, \forall a \in A$.

Theorem 2.6. *A function f is invertible iff it is bijective. If f is invertible, its inverse is unique. We write $g = f^{-1}$.*

Proof. Suppose f is bijective. Define $g\colon B \to A$ by $g(b) = a \Leftrightarrow f(a) = b$. Then g is well defined because f is injective and surjective, guaranteeing that a exists and it is unique. Note that for all $a \in A$, $g(f(a)) = a$ because if $b = f(a)$, then $g(f(a)) = g(b) = a$. Similarly, for all $b \in B$, $f(g(b)) = b$ because if $g(b) = a$, then $f(g(b)) = f(a) = b$. Then $g = f^{-1}$, and f is invertible.
Conversely, assume that f is invertible, so that there exists $g : B \to A$ such that $f(g(b)) = b, \forall b \in B$, and $g(f(a)) = a, \forall a \in A$. For any $b \in B$, define $a = g(b)$; then $f(a) = f(g(b)) = b$. Thus, f is surjective. Also, if $a_1, a_2 \in A$, then if $f(a_1) = f(a_2) \Rightarrow g(f(a_1)) = g(f(a_2)) \Rightarrow a_1 = a_2$.
To show that the inverse of f is unique, assume that $g : B \to A$ satisfies $f(g(b)) = b, \forall b \in B$, and $g(f(a)) = a, \forall a \in A$. Also, assume $h : B \to A$ satisfies $f(h(b)) = b, \forall b \in B$, and $h(f(a)) = a, \forall a \in A$. Then let $b \in B$. We have $f(g(b)) = b = f(h(b))$. This implies that $g(b) = h(b)$. Thus, $g(b) = h(b)$, for all $b \in B$, i.e. $g = h$. □

Definition 2.7. Let A, B be sets. We say that A, B have the same cardinality, and write $A \sim B$ iff there exists a bijection between A and B. (One can prove that $\sim$ is in fact an equivalence relation.)

Definition 2.8. Let A be a set. We say A is finite iff $A \sim [n]$, $n \in \mathbb{Z}^+$, or $A = \emptyset$. We say A is infinite if it is not finite. We say A is countably infinite iff $A \sim \mathbb{Z}^+$. We say that A is countable iff A is finite or countably infinite. We say A is uncountable if A is not countable.

Theorem 2.9. *If S is a countable set, then every subset of S is also countable.*

2.2 Lecture 4. Fri, Sep 6

2.2.1 Countability

Theorem 2.10. *Let A, B, C be sets. Then $A \cup B = B \cup A$, and $A \cap B = B \cap A$, and $(A \cup B) \cup C = A \cup (B \cup C)$, and $(A \cap B) \cap C = A \cap (B \cap C)$. Also, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.*

Theorem 2.11. *Let A be a set and suppose there exists a surjection $f : \mathbb{Z}^+ \to A$. Then A is countable.*

Proof. For each $a \in A$, define $S_a = f^{-1}(a)$. Since f is surjective, $S_a \neq \emptyset$, for all $a \in A$. Also $S_a \subset \mathbb{Z}^+$. Define $g : A \to \mathbb{Z}^+$ by the condition that $g(a)$ is the least element of S_a. Note that $R(g) \subset \mathbb{Z}^+$, hence, $R(G)$ is countable. Also, g is injective. If $g(a) = g(b)$, then $g(a) \in S_a$ and $g(b) \in S_b$, but if $a \neq b$, then $S_a \cap S_b = \emptyset$. (If $a \neq b$, and $n \in S_a \cap S_b$, then $f(n) = a$, and $f(n) = b$, but f is well defined.) Thus, g defines a bijection from A to $R(g)$, therefore, A is countable. □

Theorem 2.12. *A countable union of countable sets is countable.*

Proof. Rudin draws a proof in the book. □

Theorem 2.13. *Let A be a countable set and let $n \in \mathbb{Z}^+$. Then $A^n = A \times A \times \ldots \times A$ is countable.*

Proof. We argue by induction on n. For $n = 1$, $A = A$ is countable. Suppose $n \geq 2$i and A^{n-1} is countable. We have

$$\begin{aligned} A^n &= \{(a_1, a_2, \ldots, a_n \mid a_1, \ldots, a_n \in A)\} \\ &= \{(a_1, a_2, \ldots, a_{n-1}, a_n) \mid (a_1, \ldots, a_{n-1}) \in A^{n-1}, a_n \in A)\} \\ &= \bigcup_{a_n \in A} \{(a_1, a_2, \ldots, a_n) \mid (a_1, \ldots, a_{n-1}) \in A^{n-1}\} \end{aligned}$$

Note that for fixed $a_n \in A$, $S := \{(a_1, \ldots, a_{n-1}, a_n) \mid (a_1, \ldots, a_{n-1}) \in A^{n-1}\} \sim A^{n-1}$. (Construct a bijection.) Since A^{n-1} is countable by our inductive hypothesis, A^n is a countable union of countable sets, so A^n is countable. □

Corollary 2.13.1. *The set of rationals $\mathbb{Q}$ is countable.*

Theorem 2.14. *Define $S := \{(p, q) \mid p, q \in \mathbb{Z}, q \neq 0\}$. Then $S \subset \mathbb{Z} \times \mathbb{Z}$, which is countable. But $f : S \to \mathbb{Q}$, $f(p, q) = \frac{p}{q}$ is clearly a surjection. Thus, $\mathbb{Q}$ is countable.*

Theorem 2.15. *The set of reals $\mathbb{R}$ is uncountable.*

Proof. Suppose that $f : \mathbb{Z}^+ \to (0, 1)$ is a surjection. Recall that every $x \in (0, 1)$ can be written uniquely as $x = 0d_1d_2d_3, \ldots,\ d_j \in \{0, 1, \ldots, 9\}$, and $\not\exists n, d_j = 9, \forall j \geq n$. Define $x \in (0, 1)$ by $x = 0.e_1e_2e_3\ldots$, where $e_j = 2$ if the jth digit of $f(j) \neq 2$, and 3 if the jth digit of $f(j) = 2$. Then $x \neq f(n)$, for any $n \in \mathbb{Z}^+$, since the nth digit of x is different from the nth digit of $f(n)$. Thus, f cannot be surjective. Therefore, $(0, 1)$ is not countable, and therefore, $\mathbb{R}$ is not countable. □

3 Week 3

3.1 Lecture 5. Mon, Sep 9

3.1.1 Metric Spaces

Definition 3.1. Let X be a set and suppose $d : X \times X \to \mathbb{R}$. Suppose

1. $d(x, y) \geq 0$, for all $x, y \in X$; $d(x, x) = 0$,
2. $d(x, y) = d(y, x)$, for all $x, y \in X$,
3. $d(x, z) \leq d(x, y) + d(y, z)$, for all $x, y, z \in X$.

Then d is called a metric on X, and (X, d) is called a metric space.

Example 3.2.
1. Let $X = \mathbb{R}$, and $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $d(x, y) = |x - y|$.

2. Let $X = \mathbb{R}^n$, and let $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $d(x, y) = ||x - y||_2$, where $\|x\|_2 = \left(\sum_{j=1}^{n} x_j^2\right)^{\frac{1}{2}}$ for all

$x \in \mathbb{R}^n$.
3. Let X be any set, and define

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

This metric is called the discrete metric.
Consider the second proposed metric. Proving that $d(x,y) = \|x - y\|_2$ defines a metric on $\mathbb{R}^n$ is tricky. Note that $\|x\|_2 = (x \cdot x)^{\frac{1}{2}}$ where $x \cdot y = \sum_{j=1}^{n} x_j y_j$, for all $x, y \in \mathbb{R}^n$. This is called the dot product of two terms. The dot product is symmetric and bilinear: $x \cdot y = y \cdot x$, and $(\alpha x + \beta y) \cdot z = \alpha(x \cdot z) + \beta(y \cdot z)$, for $\alpha, \beta \in \mathbb{R}$. The dot product satisfies the Cauchy-Schwarz inequality: $|x \cdot y| \leq \|x\|_2 \|y\|_2$, with equality iff $y = \alpha x$ or $x = \alpha y$ for $\alpha \in \mathbb{R}$. Then for $x, y \in \mathbb{R}^n$,

$$\begin{aligned} ||x \cdot y||_2^2 &= (x+y) \cdot (x+y) = x \cdot x + 2x \cdot y + y \cdot y \\ &\leq x \cdot x + 2\|x\|_2 \|y\|_2 + y \cdot y \\ &= \|x\|_2^2 + 2\|x\|_2 \|y\|_2 + \|y\|_2 \\ &= (\|x\|_2^2 + \|y\|_2^2). \end{aligned}$$

Then for $sx, y, x \in \mathbb{R}^n$,

$$\begin{aligned} d(x,z) = \|x - z\|_2 &= \|(x-y) + (y-z)\|_2 \\ &\leq \|x - y\|_2 + \|y - z\|_2 \\ &= d(x,y) + d(y,z). \end{aligned}$$

The remaining properties of the metric are straightforward to verify.

Definition 3.3.

1. Let (X, d) be a metric space. Given $x \in X$, and $r > 0$, the open ball of radius r centered at x is the set $B_r(x) = \{y \in X \mid d(x,y) < r\}$.

2. Suppose $E \subset X$. We say that $x \in X$ is an interior point of E iff there exists $r > 0$ such that $B_r(x) \subset E$.

3. Let $E \subset X$. We say that E is open iff every point of E is an interior point of E.

Theorem 3.4. *Let (X, d) be a metric space. Let $x \in X$, and let $r > 0$. Then $B_r(x)$ is an open set.*

Theorem 3.5. *We must show that for all $y \in B_r(x)$, there exists $r' > 0$, such that $B_{r'}(y) \subset B_r(x)$. Let $r_1 = d(y,x)$, and define $r' = r - r_1$. Note that $r_1 < r$ so that $r' > 0$. Suppose $z \in B_{r'}(y)$. Then $d(z,x) \leq d(z,y) + d(y,x) < r' + r_1 = r$. This holds for all $z \in B_{r'}(y)$, so $B_{r'}(y) < B_r(x)$.*

Definition 3.6. Let (X, d) be a metric space.

1. Let $E \subset X$. We say that $x \in X$ is a limit point of E iff for all $r > 0$, $B_r(x)$ contains at least one point of E not equal to x. Note that x need not be in E.

2. We write E' for the set of limit points of E. We define the closure of E to be the set $\overline{E} = E' \cup E$.

3. We say that E is closed iff $E' \subset E$ (iff $\overline{E} = E \cup E$)

4. We say that x is an isolated point of E iff $x \in E$ and x is not a limit point of E (i.e. iff $x \in E$ and there exists $r > 0$ such that $B_r(x) \cap E = \{x\}$).

Theorem 3.7. *Let (X, d) be a metric space, let $E \subset X$, and let $x \in E'$. Then for all $r > 0$, $B_r(x)$ contains infinitely many points of E.*

Corollary 3.7.1. *A finite set has no limit points.*

Definition 3.8. Let x be a set and $E \subset X$. We define the complement of E in X as $X \setminus E = \{x \in X \mid x \notin E\}$. If X is understood, then we write $E^C = X \setminus E$.

Theorem 3.9. *Let (X, d) be a metric space, and let E be a subset of X. Then E is open if and only if E^C is closed.*

3.2 Lecture 6. Wed, Sep 11

3.2.1 Open and Closed Sets

Proof. Suppose first that E is open. We wish to prove that E^C is closed, i.e., that E^C contains all of its limit points. Equivalently, we need to prove that if $x \in E$, then x is not a limit point of E^C. Suppose that $x \in E$, then, since E is open, there exists $r > 0$ such that $B_r(x) \subset E$. But then $B_r(x) \cap E^C = \emptyset$, which proves that x is not a limit point of E^C.
Conversely, assume that E^C is closed. Then E^C contains all of its limit points, and so, if $x \in E$ then x is not a limit point of E^C. However, there exists $r > 0$ such that $B_r(x) \cap E^C = \emptyset$, which implies that there is some $r > 0$ with $B_r(x) \subset E$. This shows that E is open. □

Theorem 3.10. *Let X be a set. Let A be another set, not necessarily a subset of X. Assume that $E_\alpha \in X$, for all $\alpha \in A$. Then $\left(\bigcup_{\alpha \in A} E_\alpha\right)^C = \bigcap_{\alpha \in A} E_\alpha^C$, and $\left(\bigcap_{\alpha \in A} E_\alpha\right)^C = \bigcup_{\alpha \in A} E_\alpha^C$.*

Proof. First, $x \in \left(\bigcup_{\alpha \in A} E_\alpha\right)^C \Leftrightarrow x \notin \bigcup_{\alpha \in A} E_\alpha \Leftrightarrow \forall \alpha \in A, x \notin E_\alpha \Leftrightarrow \forall a \in A, x \in E_\alpha^C \Leftrightarrow x \in \bigcap_{\alpha \in A} E_\alpha^C$. □

Theorem 3.11. *Let (X, d) be a metric space. If $E_\alpha \subset X$ is open for all $\alpha \in A$, then $\bigcup_{\alpha \in A} E_\alpha$ is open, and if $E_1, \ldots, E_n$ are open subsets of X, then $\bigcap_{i=1}^n E_i$ is open.*

Proof. Let $E = \bigcup_{\alpha \in A} E_\alpha$. We must show that for all $x \in E$, there exists $r > 0$ such that $B_r(x) \subset E$. If $x \in E$, then there exists $\alpha' \in A$ such that $x \in E_{\alpha'}$. This implies that there exists some > 0, $B_r(x) \subset E_{\alpha'}$, since $E_{\alpha'}$ is open. This implies that $B_r(x) \subset E$, and therefore, E is open.
Now, let $E = \bigcap_{j=1}^n E_j$, and let $x \in E$. Then for all $j \in [n]$, $x \in E_j$. This means that for all $j \in [n]$, there exists an $r_j > 0$, such that $B_{r_j}(x) \subset E_j$. Set $r = \min\{r_1, \ldots, r_n\}$, then, for all $j \in [n]$, $B_r(x) \subset B_{r_j}(x) \subset E_j \Rightarrow B_r(x) \subset \bigcap_{j=1}^n E_j = E$. □

Corollary 3.11.1. *Let (X, d) be a metric space. If E_α is closed for all $\alpha \in A$, then $\bigcap_{\alpha \in A} E_\alpha$ is closed. If $E_1, \ldots, E_n$ are closed subsets of X, then $\bigcup_{j=1}^n E_j$ is closed.*

Proof. Note that it suffices to prove that $\left(\bigcap_{\alpha \in A} E_\alpha\right)^C$ is open. Note that $\left(\bigcup_{\alpha \in A} E_\alpha\right)^C = \bigcup_{\alpha \in A} E_\alpha^C$, and we know that E_α^C is open, therefore, $\bigcup_{\alpha \in A} E_\alpha^C$ is open. Thus $\bigcap_{\alpha \in A} E_\alpha$ is closed. □

Theorem 3.12 (Rudin 2.27)**.** *Let (X, d) be a metric space and let $E \subset X$. First, $\overline{E}$ is closed. Then, E is closed iff $E = \overline{E}$. Also, if F is a closed subset of X and $E \subset F$, then $\overline{E} \subset F$. (I.e. $\overline{E}$ is the smallest closed subset of X containing E.)*

Definition 3.13. Let (X, d) be a metric space, and suppose $Y \subset X$. Note that Y can be considered a metric space under d. We say that $E \subset Y$ is open relative to Y iff for all $y \in Y$, there exists $r > 0$ such that $Y \cap B_r(y) \subset E$.

Theorem 3.14. *Let (X, d) be a metric space, and let $Y \subset X$. Then $S \subset Y$ is open relative to Y iff there exists an open set $E \subset X$ such that $S = Y \cap E$.*

Proof. First suppose $S = Y \cap E$, where $E \subset X$ is open. We wish to show that S is open relative to Y. Let $y \in S$. Since $y \in S \subset E$, and E is open (in X), then there exists $r > 0$ such that $B_r(Y) \subset E$. But then $Y \cap B_r(y) \subset Y \cap E = S$. Since $y \in S$ was chosen arbitrarily, this shows S is open relative to Y. Conversely, suppose $S \subset Y$ is open relative to Y. Then for all $y \in S$, there exists some $r_y > 0$, such that $Y \cap B_{r_y}(y) \subset S$. Define $E = \bigcup_{y \in S} B_{r_y}(y)$, and since each $B_{r_y}(y)$ is open in X, E is open in X. Also,

$$Y \cap E = Y \cap \left(\bigcap_{y \in Y} B_{r_y}(y) \right) = \bigcup_{y \in Y} (Y \cap B_{r_y}(y)) \subset S,$$

since $Y \cap B_{r_y}(y) \subset S$, for all $y \in S$. Verifying $S \subset Y \cap E$ is simple, therefore, $S = Y \cap E$, as desired. ☐

3.3 Lecture 7. Fri, Sep 13

3.3.1 Compactness

Definition 3.15. Let (X, d) be a metric space, and let $E \subset X$. We say that E is compact iff for any collection $\{G_\alpha \mid \alpha \in A\}$ of open sets with $E \subset \bigcup_{\alpha \in A} G_\alpha$, there is a finite subcollection $\{G_{\alpha_1}, \ldots, G_{\alpha_n}\} \subset \{G_\alpha \mid \alpha \in A\}$, such that $E \subset \bigcup_{j=1}^{n} G_{\alpha_j}$.

Definition 3.16. If the sets G_α are open for all $\alpha \in A$, and $E \subset \bigcup_{\alpha \in A}$, then we call $\{G_\alpha \mid \alpha \in A\}$ an open cover of E. If $\{G_{\alpha_1}, \ldots, G_{\alpha_n}\} \subset \{G_\alpha \mid \alpha \in A\}$ and $E \subset \bigcup_{j=1}^{n} G_{\alpha_j}$, then we call $\{G_{\alpha_1}, \ldots, G_{\alpha_n}\}$ is a finite subcover of E.

Example 3.17. The open interval $\big(0, 1\big) \subset \mathbb{R}$ is not compact. Let $G_n = (\frac{1}{n}, 1)$, for all $n \in \mathbb{Z}^+$. Then $(0, 1) \subset \bigcup_{i=1}^{\infty} G_n$. Then $\{G_n \mid n \in \mathbb{Z}^+\}$ is an open cover of $(0, 1)$. But no $\{G_{n_1}, \ldots, G_{n_k}\} \subset \{G_n \mid n \in \mathbb{Z}^+\}$ can cover $(0, 1)$, since

$$\bigcup_{j=1}^{k} G_{n_j} = \bigcup_{i=1}^{k} \left(\frac{1}{n}, 1 \right) = \left(\frac{1}{\ell}, 1 \right),$$

where $\ell = \max\{n_1, \ldots, n_k\}$.

Definition 3.18. Let (X, d) be a metric space. A set $E \subset X$ is bounded iff $\exists\, x \in X$, $r > 0$, such that $\forall y \in E$, $d(y, x) \leq r$.

Theorem 3.19. *Let (X, d) be a metric space, and let $E \subset X$ be compact.*

1. *Then E is closed.*
2. *If $F \subset E$, and F is closed, then F is compact.*

Proof. Let x be any point in E^C. For each $y \in E$, define $r_y = \frac{d(y,x)}{2}$. Note that $B_{r_y}(y) \cap B_{r_y}(x) = \varnothing$. Also note that the collection $\{B_{r_y}(y) \mid y \in E\}$ is an open cover of E. Since E is compact, let $\{B_{r_y}(y_1), \ldots, B_{r_y}(y_n)\}$ be a finite subcover of E. Define $r := \min\{r_{y_1}, \ldots, r_{y_n}\}$. Then $E \subset$

$\bigcup_{j=1}^n B_{r_{y_j}}(y_j)$, and $B_r(x) \cap B_{r_{y_j}}(y_{r_j}) = \varnothing$, for all $j = 1, \dots, n$. Thus, $B_r(x) \cap E = \varnothing$. Hence $B_r(x) \subset E^C \Rightarrow E^c$ is open $\Rightarrow E$ is closed.
For the second assertion, suppose $F \subset E$ is closed. Let $\{G_\alpha \mid \alpha \in A\}$ be an open cover of F. Since F is closed, F^C is open, and so $\{G_\alpha \mid \alpha \in F\} \cup \{F\}$ is an open cover of E (since $E = (E \cap F) \cup (E \cap F^C)$). Since E is compact, there exists a finite subcover of either the form $\{G_{\alpha_1}, \dots, G_{\alpha_n}\}$ or $\{G_{\alpha_1}, \dots, G_{\alpha_n}, F^C\}$. In either case, $F \subset \bigcup_{j=1}^n G_{\alpha_j}$ (since $F \cap F^C = \varnothing$). Thus, F is compact. □

Corollary 3.19.1. *Let (X, d) be a metric space, and let $E \subset X$ be compact, and let $F \subset X$ be closed. Then $E \cap F$ is compact.*

Theorem 3.20. *Let (X, d) be a metric space, and let $E \subset Y \subset X$. Then E is compact relative to Y iff E is compact relative to X.*

Proof. Suppose E is compact relative to X. Let $\{U_\alpha \mid x \in A\}$ be a collection of subsets of Y open relative to Y such that $E \subset \bigcup_{\alpha \in A} U_\alpha$. By an earlier theorem, for each $\alpha \in A$, there exists an open subset G_α of X such that $U_\alpha = Y \cap G_\alpha$. But then $E \subset \bigcup_{\alpha \in A} U_\alpha \subset \bigcup_{\alpha \in A} G_\alpha$. Since E is compact in X, there exists $\alpha_1, \dots, \alpha_n$ such that $E \subset \bigcup_{j=1}^n G_{\alpha_j}$. Then $E \subset Y \Rightarrow E \subset Y \cap \left(\bigcup\limits_{j=1}^n G_{\alpha_j}\right) = E \subset \bigcup\limits_{j=1}^n \left(Y \cap G_{\alpha_j}\right) \Rightarrow E \subset \bigcup\limits_{j=1}^n U_{\alpha_j}$. Therefore, E is compact in Y. Conversely, suppose E is compact relative to Y. Let $\{G_\alpha \mid \alpha \in A\}$ be an open cover in X ($G_\alpha \subset E$ is open, $E \subset \bigcup_{\alpha \in A} G_\alpha$). Then

$$E \subset Y, E \subset \bigcup_{\alpha \in A} G_\alpha \Rightarrow E \subset U \cap (\bigcup_{\alpha \in A} G_\alpha) \Rightarrow E \subset \bigcup_{\alpha \in A} (Y \cap G_\alpha),$$

and each $Y \cap G_\alpha$ is open relative to Y. Since E is compact relative to Y, there exists $\alpha_1, \dots, \alpha_n$ such that $E \subset \bigcup\limits_{j=1}^n (Y \cap G_{\alpha_j}) \subset \bigcup\limits_{j=1}^n G_{\alpha_j}$. Thus, there is a finite subcover of E and hence E is compact in X. □

4 Week 4

4.1 Lecture 8. Mon, Sep 16

4.1.1 Sequential Compactness

Theorem 4.1. *Let (X, d) be a metric space and let $\{E_\alpha \mid \alpha \in A\}$ be a collection of compact subsets of X with the property that any finite subcollection has a nonempty intersection. Then $\bigcap\limits_{\alpha \in A} E_\alpha \neq \varnothing$.*

Proof. By contrapositive, assume $\bigcap\limits_{\alpha \in A} E_\alpha = \varnothing$. Choose any $\alpha_0 \in A$, and define $A' = A \setminus \{\alpha_0\}$. Then $E_{\alpha_0} \bigcap \left(\bigcap\limits_{\alpha \in A'} E_\alpha\right) = \varnothing$. Therefore, $E_{\alpha_0} \subset \left(\bigcap\limits_{\alpha \in A'} E_\alpha\right)^C \Rightarrow E_{\alpha_0} \subset \bigcup\limits_{\alpha \in A'} E_\alpha^C$. Since E_α is compact, E_α is closed, so E_α^C is open for all $\alpha \in A'$. Thus, $\{E_\alpha^C \mid \alpha \in A'\}$ is an open cover of the compact set E_{α_0}, hence there exists $\alpha_1, \dots, \alpha_n \in A'$ such that $E_{\alpha_0} \subset \bigcup\limits_{j=1}^n E_{\alpha_j}^C$. This implies

$E_{\alpha_0} \subset \left(\bigcap_{j=1}^{n} E_\alpha\right)^C \Rightarrow E_{\alpha_0} \bigcap \left(\bigcap_{j=1}^{n} E_\alpha\right) = \varnothing \Rightarrow \bigcap_{j=0}^{n} E_{\alpha_j} = \varnothing$. Hence there is a finite subcollection with an empty intersection. ☐

Corollary 4.1.1. *Let (X, d) be a metric space and let E_n be a nonempty compact subset of X, for all $n \in \mathbb{Z}^+$. If $E_{n+1} \subset E_n$, for all $n \in \mathbb{Z}^+$, then $\bigcap_{j=1}^{\infty} E_n \neq \varnothing$.*

Theorem 4.2. *Let (X, d) be a metric space, and let $K \subset X$ be compact. If $E \subset K$ is infinite, then E has a limit point in K.*

Proof. By contrapositive, assume that $E \subset K$ is infinite, and E has no limit points in K. We want to show that K is not compact. Then for all $x \in K$, there exists $r_x > 0$ such that $B_{r_x}(x)$ contains at most one point of E (specifically, $B_{r_x}(x) \cap E = \{x\}$ if $x \in E$, and $B_{r_x}(x) \cap E = \varnothing$). Then $\{B_{r_x}(x) \mid x \in K\}$ is an open cover of K, and it has no finite subcover (any finite subcollection contains only finitely many points of E, hence we cannot cover E or K). Thus, K is not compact. ☐

Corollary 4.2.1. *If $\{X_n\} \subset K$ and K is compact, then there exist a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ converging to some x in K.*

Theorem 4.3. *Let $\{[a_n, b_n]\}$ be a sequence of nonempty closed intervals in $\mathbb{R}$. Suppose that $[a_n + 1, b_n + 1] \subset [a_n, b_n]$, for all $n \in \mathbb{Z}^+$. Then $\bigcap_{n=1}^{\infty} [a_n, b_n] \neq \varnothing$.*

Proof. Note that if $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$, for all $n \in \mathbb{Z}^+$, then $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$. This implies that $a_n \leq b_m$, for all $n, m \in \mathbb{Z}^+$. Thus, $\{a_n \mid n \in \mathbb{Z}^+\}$ is bounded above by any $b_m, m \in \mathbb{Z}^+$, so $a = \sup\{a_n\}$ exists in $\mathbb{R}$. But then $a_n \leq a \leq b_n$ for all $n \in \mathbb{Z}^+$. This implies that $a \in \bigcap_{n=1}^{\infty} [a_n, b_n]$. ☐

Definition 4.4. Let $k \in \mathbb{Z}^+$. A k-cell is a subset of $\mathbb{R}^k$ of the form

$$C = \{x \in \mathbb{R}^k \mid a_j \leq x_j \leq b_j, \forall j = 1, \ldots, k\},$$

where $a_1, \ldots, a_k, b_1, \ldots, b_k$ are fixed constants, $a_i \leq b_j, \forall j$.

Theorem 4.5. *Let $k \in \mathbb{Z}^+$ and let $\{C_n\}$ be a sequence of k-cells. If $C_{n+1} \subset C_n$, for all $n \in \mathbb{Z}^+$, then $\bigcap_{n=1}^{\infty} C_n \neq \varnothing$.*

Theorem 4.6. *Let $k \in \mathbb{Z}^+$, and let C be a k-cell. Then C is compact. (For $k = 1$, we have C is a closed interval.)*

4.2 Lecture 9. Wed, Sep 18

4.2.1 Heine-Borel

Lemma 4.7. *Let (X, d) be a metric space and let $E \subset X$. If E is compact, then E is bounded.*

Proof. Choose any point $x \in E$ and note that $\{B_n(x) \mid n \in \mathbb{Z}^+\}$ is an open cover of E (every point $y \in E$ satisfies $d(y, x) < n$ for some $n \in \mathbb{Z}^+$). Thus, if E is compact, a finite subcollection $\{B_{n_1}(x), B_{n_2}(x), \ldots, B_{n_k}(x)\}$ covers E. But then $E \subset B_r(x)$, where $r = \max\{n_1, \ldots, n_k\}$. Thus, E is bounded. ☐

Theorem 4.8 (Heine-Borel). *Let $k \in \mathbb{R}^k$, and let $E \subset \mathbb{R}^k$ be closed and bounded. Then E is compact.*

Proof. Since E is bounded, it is a subset of some k-cell C. We have seen that a k-cell is compact, and we also have seen that a closed subset of a compact set is compact. Thus, E is compact. □

Definition 4.9. Let (X, d) be a metric space. We say that $E \subset X$ is *dense* in X iff for all $x \in X$ and all $r > 0$, there exists $e \in E$ such that $d(e, x) < r$.

4.2.2 Separability, Connectedness

Definition 4.10. Let (X, d) be a metric space. We say that X is separable iff it contains a countable dense subspace.

Example 4.11. Let $k \in \mathbb{Z}^+$. Then $\mathbb{R}^k$ is separable.

Proof. Define $S = \{(r_1, \ldots, r_k) \mid r_1, \ldots, r_k \in \mathbb{Q}\}$. By an earlier theorem, $\mathbb{Q}^k$ is countable. Let $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$, and $r > 0$ be given. For each $j \in [k]$, there exists $r_j \in (x_j - \frac{r}{\sqrt{k}}, x_j + \frac{r}{\sqrt{k}}) \cap \mathbb{Q}$. It follows that

$$d(r, x) = \|r - x\|_2 = \left[\sum_{j=1}^{k}(r_j \cdot x_j)\right]^{\frac{1}{2}} < \left[\sum_{j=1}^{k}\frac{r^2}{k}\right]^{\frac{1}{2}} = r.$$

Thus, $\mathbb{Q}^k$ is dense in $\mathbb{R}^k$. □

Definition 4.12. Let (X, d) be metric space. A collection $\{U_\alpha \mid \alpha \in A\}$ of open sets of X is called a base for X (or for the topology on X) iff for every open subset G of X and for every $x \in G$, there exists $\alpha' \in A$ such that $x \in U_{\alpha'} \subset G$.

Theorem 4.13. *Let (X, d) a separable metric space. Then there exists a countable base for X.*

Proof. Let $S = \{X_n\}$ be a countable dense subset of X and define

$$\begin{aligned}\aleph &= \{B_r(x_n) \mid n \in \mathbb{Z}^+, r \in \mathbb{Q}^+\} \\ &= \bigcup_{n=1}^{\infty}\{B_r(x_n) \mid r \in \mathbb{Q}^+\}.\end{aligned}$$

Then $\aleph$ is countable (since it is a countable union of of countable sets). We claim that $\aleph$ is a base for X. Let G be an open subset of X, and let $x \in G$. Then there exists $r \in \mathbb{R}^+$ such that $B_r(x) \subset G$. Choose $r' \in (0, \frac{r}{2}) \cap \mathbb{Q}$ and let $x_n \in S$ such that $d(x_n, x) < r!$. Then $x \in B_{r'}(x_n) \subset B_r(x) \subset G$. (To see this, note that $y \in B_{r'}(x_n) \implies d(y, x) \leq d(y, x_n) + d(x_n, x) < r' + r' = 2r' < r$). Thus $B_{r'}(x_n) \subset B_r(x)$.) This completes the proof. □

Theorem 4.14. *Let (X, d) be a metric space with a countable base. If $E \subset X$ is any set and $\{G_\alpha \mid \alpha \in A\}$ is an open cover for E. Then there is a countable subcover A' of A such that $E \subset \bigcup_{\alpha \in A} G_\alpha$.*

Proof. Let $\{U_n \mid x \in \mathbb{Z}^+\}$ be a countable base for X, and let $E \subset X$, and let $\{G_\alpha \mid \alpha \in A\}$ be an open cover for E. For each $x \in E$, there exists $\alpha_x \in A$ such that $x \in G_{\alpha_x}$. Since $\{U_\alpha\}$ is a base for the topology of X, for each $x \in E$, there exists $n_x \in \mathbb{Z}^+$ such that $x \in U_{n_x} \subset G_{\alpha_x}$. Define $B = \{U_{n_x} \mid x \in E\} \subset \{U_n\}$ and note that B is countable. By construction for each $U \in B$, there exists G_{α_u} such that $U \subset G_{\alpha_u}$ (there may be many such G_{α_u}, but we need only one). But then $\{G_{\alpha_u} \mid U \in B\}$ is a countable open cover for E. □

4.3 Lecture 10. Fri, Sep 20

Theorem 4.15. *Let (X, d) be a metric space with the property that every infinite set in X has a limit point in X. Then X is separable.*

Proof. Let $\delta > 0$ be arbitrary. We will construct $x_1, \dots, x_n$ such that $d(x_i, x_j) \geq \delta$, for all $i, j, i \neq j$. Moreover, we will show that any such set is finite. So choose any $x_1 \in X$. Given $x_1, \dots, x_k \in X$ such that $d(x_i, x_j) \geq \delta$, for all $i, j = 1, \dots, k, i \neq j$, choose x_{k+1} so that $d(x_i, x_{k+1}) \geq \delta$, for all $i = 1, \dots, k$. We claim that this process must end after finitely many x_j's are chosen. If not, then we obtain a sequence $x_1, x_2, \dots$, such that $d(x_i, x_j) \geq \delta$, for all $i, j \in \mathbb{Z}^+, i \neq j$. But then $\{x_n\}$ contains no limit point in X. (Any ball of radius $\frac{\delta}{2}$ contains at most one x_n.) This contradicts the hypothesis, so only finitely many such $x_1, \dots, x_n$ can exist.
Given $\delta > 0$, write $S_\delta = \{x_1, \dots, x_{n_\delta}\}$. Note that $x_1, \dots, x_{n_\delta}$ are not uniquely determined. We choose one such set for each $\delta > 0$. Now define

$$S = \bigcup_{n=1}^{\infty} S_{\frac{1}{n}}.$$

Then S is countable. Moreover, for all $x \in X$, there exists $y \in S_{\frac{1}{n}}$ such that $d(y, x) < \frac{1}{n}$. (In constructing S_δ, we keep adding more points until we can't find another $x \in X$ at distance at least δ from the points in S_δ. Thus, for all $x \in X$, there exists $y \in S_\delta$, $d(y, x) < \delta$.) Therefore, S is dense in $X : \forall x \in X, \forall \varepsilon > 0, (\exists n \in \mathbb{Z}^+, \frac{1}{n} < \varepsilon$, and $\exists y \in S_{\frac{1}{n}}, d(y, x) < \frac{1}{n} < \varepsilon)$. This shows that X is separable. □

Theorem 4.16. *Let (X, d) be a metric space with the property that every infinite subset of X has a limit point in X. Then X is compact.*

Proof. By the last theorem, X is separable, then X has a countable base. [Let $\mathcal{U}$ be a collection of open sets in X. Then $\mathcal{U}$ is a base for X iff for all open sets $G \subset X$, and all $x \in G$, there exists $U \in \mathcal{U}$ such that $x \in U \in G$.] Hence, it suffices to prove that every countable cover $\{G_n \mid n \in \mathbb{Z}^+\}$ of X has a finite subcover. So assume that $\{G_n\}$ is a countable open cover of X. For all $n \in \mathbb{Z}^+$, define

$$F_n = \left(\bigcup_{i=1}^{n} G_j \right)^C = \bigcap_{j=1}^{n} G_j^C.$$

Let us argue by contradiction and assume that $\{G_n\}$ contains no finite subcover. Then $\{G_1, \dots, G_n\}$ does not cover X for all $n \in \mathbb{Z}^+$. Hence $F_n \neq \varnothing$ for all $F_n \neq \varnothing$ for all $n \in \mathbb{Z}^+$. However,

$$\bigcap_{j=1}^{\infty} F_n = \bigcap_{j=1}^{\infty} G_j^C = \left(\bigcup_{j=1^\infty G_j} \right)^C = X^C = \varnothing.$$

For each $n \in \mathbb{Z}^+$, let $x_n \in F_n$. By assumption, $\{x_n\}$ has a limit point $x \in X$. Note that each F_n is closed and

$$F_{n+1} \subset F_n \forall n \in \mathbb{Z}^+ \Rightarrow \forall n \in \mathbb{Z}^+, \{x_k \mid k \geq n\} \subset F_n.$$

Thus x is a limit point for each F_n and hence, since each F_n is closed, $x \in F_n, \forall n \in \mathbb{Z}^+$. But then $x \in \bigcap_{n=1}^{\infty}$, a contradiction. Thus $\{G_n\}$ must contain a finite subcover, and we have proven that X is compact. □

Corollary 4.16.1. *Let (X, d) be a metric space and let $E \subset X$ have the property that every infinite subset of E has a limit point in E. Then E is compact.*

Proof. E is compact in (E, d) iff E is compact in (X, d). $\square$

Theorem 4.17. *In $\mathbb{R}^k$, E is compact iff E is closed and bounded.*

Theorem 4.18. *In any metric space, if E is compact, then E is closed and bounded.*

4.3.1 Alternative Proof for Uncountability of Reals

Definition 4.19. Let (X, d) be a metric space and let $E \subset X$. We say that E is *perfect* iff every point of E is a limit point of E.

Definition 4.20. Let (X, d) be a metric space. If $A, B \subset X$, we say that A, B are separated iff $(A \cap \overline{B} = \varnothing$, and $\overline{A} \cap B = \varnothing)$.

5 Week 5

5.1 Lecture 11.

Theorem 5.1. *A subset E of $\mathbb{R}$ is connected iff it is an interval, that is, iff*

$$(x, y \in E \wedge x < z < y) \Rightarrow z \in E. \tag{*}$$

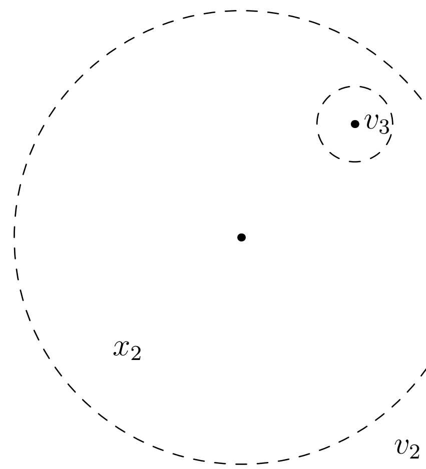
Proof. Suppose first that E is not connected, that is, there exists nonempty separated sets $A, B \subset \mathbb{R}$, such that $E = A \cup B$. We wish to prove that $(*)$ fails. Choose $x \in A, y \in B$, and assume WLOG $x < y$. Define $z = \sup\{A \cap [x, y]\}$. Note that $A \cap [x, y]$ is bounded above y, so z is well defined, and $z \in \overline{A}$. (If $z \notin A$, then z must be a limit point of A, otherwise, there would be a smaller upper bound). Since A, B are separated, $z \notin B$. In particular, $z < y$. If $z \notin A$, then $z \notin E = A \cup B$, and $x < z < y$, so $(*)$ fails. If $z \in A$, then $z \notin \overline{B}$, so there exists $z_1 \in (z, y)$ such that $z_1 \notin B$, but then $z_1 > z \Rightarrow z_1 \notin A \cap [x, y] \Rightarrow z_1 \notin A$. We see that $x < z_1 < y$ and $z \notin E = A \cup B$, so $(*)$ fails.
Conversely, suppose $(*)$ fails. Then there exists $x, y, z \in \mathbb{R}$ such that $x, y \in E$, and $z \notin E$, and $x < z < y$. Define $A = E \cap (-\infty, z), B = E \cap (z, \infty)$. Then A, B are nonempty ($x \in A, y \in B$), and A, B are separated, since $A \subset (-\infty, z), B \subset (z, \infty)$ and $(-\infty, z), (z, \infty)$ are separated. Also, $E = A \cup B$, so E is separated. $\square$

Theorem 5.2. *If $k \in \mathbb{Z}^+$, and $E \subset \mathbb{R}^k$ is nonempty and perfect, then E is uncountable.*

Proof. Since E is nonempty and perfect, E is infinite (a finite set has no limit points). Let us assume by contradiction that E is countable. Then write E as a sequence: $E = \{x_n\}$. Choose $r_1 > 0$ arbitrarily and define $V_1 = B_{r_1}(x_1)$. Choose V_2 to be an open ball with the following properties:

1. V_2 is centered at a point lying in E,
2. $x_1 \notin \overline{V_2}$,
3. $\overline{V_2} \subset V_1$.

This is possible since x_1 is a limit point and hence V_1 contains infinitely many points of E. (Note that x_2 may or may not belong to V_2.) Next, choose V_3 to be an open ball centered at a point of E, and such that $x_2 \notin \overline{V_3}$, and $\overline{V_3} \subset V_2$. Continue in this fashion to construct a sequence $\{v_n\}$ of open balls, each centered at a point of E such that $x_n \notin \overline{V_{n+1}}$ and $\overline{V_{n+1}} \subset V_n$, for all $n \in \mathbb{Z}^+$. Define



$C_n = \overline{V_n} \cap E$. Then each C_n is compact, since $\overline{V_n}$ is compact by the Heine-Borel theorem and E is closed. Also, each C_n is nonempty (since the center of V_n lies in E for each $n \in \mathbb{Z}^+$). Hence, by a previous theorem,

$$C = \bigcap_{n=1}^{\infty} \neq \varnothing,$$

and obviously $C \subset E$. But by construction, $x_n \notin C_{n+1}$, for all $n \in \mathbb{Z}^+$, and hence $x_n \notin C$ for all $n \in \mathbb{Z}^+$. This is a contradiction, since $E = \{x_n\}$, and $C \subset E$. □

Corollary 5.2.1. *The sets $\mathbb{R}, \mathbb{R}^k$ are uncountable. Any open interval in $\mathbb{R}$ is uncountable, and any nonempty open set in $\mathbb{R}^k$ is uncountable.*

5.1.1 Sequential Convergence

Definition 5.3. Let (X, d) be a metric space, and let $\{x_n\}$ be a sequence in X. We say that $\{x_n\}$ converges iff there exists $x \in X$ such that for all $\varepsilon > 0$, there exists $N \in \mathbb{Z}^+$ such that $n \geq N \Rightarrow d(x_n, x) < \varepsilon$. In this case, we write $\{x_n\}$ converges to x, and write $x_n \to x$, or $x = \lim_{n\to\infty} x_n$.

Note. A sequence $\{x_n\} \subset X$ is a function $x : \mathbb{Z}^+ \to X$, where we write x_n instead of $x(n)$. We also the symbol $\{x_n\} = \{x_n \mid n \in \mathbb{Z}^+\}$ to denote x; that is, we denote x by its range. To say $\{x_n\}$ is bounded is to say that the set $\{x_n\}$ is bounded.

Theorem 5.4. *Let (X, d) be a metric space and let $\{x_n\}$ be a sequence in X:*

1. *$\{x_n\}$ converges to $x \in X$ iff for all $r > 0$, $B_r(x)$ contains all but finitely many terms in $\{x_n\}$.*
2. *If $\{x_n\}$ converges, its limit is unique.*
3. *If $\{x_n\}$ converges, it is bounded.*
4. *If $E \subset X$ and $x \in X$ is a limit point of E, then there exists a sequence $\{x_n\} \subset E$ such that $x_n \to x$.*

Proof.
1. Suppose $x_n \to x$. Then for all $r > 0$, there exists $N \in \mathbb{Z}^+$ such that $n \geq N \Rightarrow d(x_n, x) < r$. That is, $\{x_n \mid n \geq N\} \subset B_r(x)$. Thus, all but finitely many terms of $\{x_n\}$ belong to $B_r(x)$. Conversely, suppose $\{x_n\} \subset X$, $x \in X$, and for all $r > 0$, all but finitely many terms in $\{x_n\}$ belong to $B_r(x)$.

Let ε be given. Since all but finitely many terms of $\{x_n\}$ belong to $B_\varepsilon(x)$, there exists $N \in \mathbb{Z}^+$ such that $\{x_n \mid n \in N\} \subset B_\varepsilon(x)$, that is, $n \geq N \Rightarrow d(x_n, x) < \varepsilon$, thus, $x_n \to x$. □

5.2 Lecture 12.

5.2.1 More on Sequential Convergence

Theorem 5.5. *Let $\{s_n\}, \{t_n\}$ be sequences in $\mathbb{R}$ and suppose that $s_n \to s$, $t_n \to t$. Then*

1. $s_n + t_n \to s + t$,
2. *for all* $c \in \mathbb{R}$, $cs_n \to cs$, *and* $c + s_n \to c + s$,
3. $s_n t_n \to st$,
4. $\frac{s_n}{t_n} \to \frac{s}{t}$, *if* $t \neq 0$.

Proof of 3. Let ε be given. We must show that there exists $N \in \mathbb{Z}^+$ such that $n \geq N \Rightarrow |s_n t_n - st| < \varepsilon$. Note that

$$\begin{aligned} s_n t_n - st &= s_n t_n - st_n + st_n - st \\ &= (s_n - s)t_n + s(t_n - t). \end{aligned}$$

Now, there exists $N_1 \in \mathbb{Z}^+$ such that

$$n \geq N_1 \Rightarrow |t_n - t| < 1 \Rightarrow |t_n| = |t_n - t + t| \leq |t_n - t| + |t| < |t| + 1.$$

Since $s_n \to s$, there exists $N_2 \in \mathbb{Z}^+$ such that

$$n \geq N_2 \Rightarrow |s_n - s| < \frac{\varepsilon}{2(|t| + 1)},$$

and since $t_n \to t$, there exists $N_3 \in \mathbb{Z}^+$ such that

$$n \geq N_3 |t_n - t| < \frac{\varepsilon}{2|S|}$$

(take $N_3 = 1$ if $s = 0$). Then, with $N = \max\{N_1, N_2, N_3\}$, we have

$$\begin{aligned} n \geq N \Rightarrow |s_n t_n - st| &\leq |s_n - s||t_n| + |s||t_n - t| \\ &< \frac{\varepsilon}{2(|t| + 1)}(|t| + 1) + |s|\frac{\varepsilon}{2|s|} \\ &= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Thus $s_n t_n \to st$. □

Note. In Theorem 5.5.4, it is understood that

$$\begin{aligned} t_n \to t, s \neq 0 \Rightarrow t_n \neq 0, \forall n \in \mathbb{Z}^+ &\text{ sufficiently large} \\ \Rightarrow \frac{s_n}{t_n} &\text{ is well defined for all } n \in \mathbb{Z}^+ \text{ sufficiently large.} \end{aligned}$$

But we may have to ignore a finite number of terms in $\left\{\frac{s_n}{t_n}\right\}$ that are undefined (t may equal 0 for a finite number of values of n).

Theorem 5.6. *Let $k \in \mathbb{Z}^+$ and let $\{x_n\}$ be sequence in $\mathbb{R}^k$, where*

$$x_n = (x_{1,n}, x_{2,n}, \dots, x_{k,n}).$$

Then $x_n \to x = (x_1, x_2, \dots, x_k)$ iff

$$\forall j \in [k], x_{j,n} \to x_j.$$

Proof. First, suppose $x_n \to x$, which means that

$$\|x_n - x\|_2 \to 0.$$

We have

$$|x_{j,n} - x_j| \le \|x_n - x\|_2, \forall j \in [k], \forall n \in \mathbb{Z}^+,$$

and hence

$$\begin{aligned} \|x_n - x\|_2 \to 0 \Rightarrow |x_{j,n} - x_j| \to 0 \\ \Rightarrow x_{j,n} \to x_j. \end{aligned}$$

(Given $\varepsilon > 0$, there exists $N \in \mathbb{Z}^+$ such that

$$n \ge N \Rightarrow \|x_n - x\|_2 < \varepsilon.$$

But then, for each $j = 1, \dots, k$,

$$n \ge N \Rightarrow |x_{j,n} - x_j| < \varepsilon.$$

Thus, for each $j = 1, \dots, k$, $x_{j,k} \to x_j$.) Conversely, suppose $\forall j = 1, \dots, k$, that $x_{j,n} \to x_j$. Let $\varepsilon > 0$, be given. Then, for each $j = 1, \dots, k$, there exists $N_j \in \mathbb{Z}^+$ such that

$$n \ge N_j \Rightarrow |x_{j,n} - x_j| < \frac{\varepsilon}{\sqrt{k}}.$$

Define $N = \max\{N_1, \dots, N_k\}$. Then

$$\begin{aligned} n \ge N &\Rightarrow |x_{j,n} - x_j| < \frac{\varepsilon}{\sqrt{k}}, \forall j = 1, \dots, k \\ &\Rightarrow \|x_n - x\|_2 = \left[\sum_{j=1}^{k} |x_{j,n} - x_j|^2\right]^{\frac{1}{2}} < \left[\sum_{j=1}^{k} \frac{\varepsilon^2}{k}\right]^{\frac{1}{2}} \\ &= \varepsilon. \end{aligned}$$

Thus shows that $x_n \to x$. □

Corollary 5.6.1. *Let $k \in \mathbb{Z}^+$, let $\{x_n\}, \{y_n\}$ be sequences in $\mathbb{R}^k$, and suppose $x_n \to x$, and $y_n \to y$. Then*

1. $x_n \pm y_n \to x \pm y$,
2. $x_n y_n \to x \cdot y$.

If $\{\beta_n\}$ is a sequence in $\mathbb{R}$ such that $\beta_n \to \beta$, then $\beta_n x_n \to \beta x$.

Definition 5.7. Let (X, d) be a metric space, let $\{x_n\}$ be a sequence in X, and let $\{n_k\}$ be a strictly increasing sequence in $\mathbb{Z}^+$. The sequence $\{x_{n_k}\}$ is called a subsequence of $\{x_n\}$. If $x_{n_k} \to x$, then x is called a subsequential limit of $\{x_n\}$.

Theorem 5.8. *Let (X, d) be a metric space and let $\{x_n\}$ be a sequence in X. Then $x_n \to x \in X$ iff every subsequence of $\{x_n\}$ converges to x.*

Proof. Homework. ☐

Lemma 5.9. *Let (X, d) be a metric space and let $\{x_n\}$ be a sequence in X. If $x \in X$ is a limit point of $\{x_n\}$, then there is a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ that converges to x.*

Proof. Suppose x is a limit point of $\{x_n\}$. We construct a subsequence of $\{x_n\}$ converging to x as follows.

1. Choose any $x_{n_1} \in B_1(x)$.
2. Then $B_{\frac{1}{2}}(x)$ contains infinitely many elements of $\{x_n\}$, so choose x_{n_2} such that $x_{n_2} \in B_{\frac{1}{2}}(x)$ and $n_2 > n_1$.
3. Continuing by induction, given $x_{n_1}, \ldots, x_{n_k}$, such that $x_{n_j} \in B_{\frac{1}{j}}(x)$, $\forall j = 1, \ldots, k$, and $n_1 < n_2 < \ldots < n_k$, choose $x_{n_{k+1}}$ such that $x_{n_{k+1}} \in B_{\frac{1}{k+1}}(x)$, and $n_{k+1} > n_k$. It then follows immediately that $x_{n_k} \to x$. If $\varepsilon > 0$, choose $\ell \in \mathbb{Z}^+$, such that $\frac{1}{\ell} < \varepsilon$. Then

$$k \geq \ell \Rightarrow d(x_{n_k}, x) < \frac{1}{\ell} < \varepsilon.$$

Note that the converse is not true.

☐

5.3 Lecture 13.

5.3.1 Bolzano-Weierstrauss

Theorem 5.10. *Let (X, d) be a metric space and let $E \subset X$. Then E is compact iff every sequence contained in E has a subsequence that converges to a point in E.*

Proof. Suppose first that E is compact. Then by an earlier theorem, each infinite subset of E has a limit point in E (Theorem 4.16). If the set $\{x_n \mid n \in \mathbb{Z}^+\}$ is infinite, then it has a limit point $x \in E$, and by an earlier lemma (Lemma 5.9), there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \to x$. If the set $\{x_n \mid n \in \mathbb{Z}\}$ is finite, then there exists some $x \in E$, and an increasing sequence $\{n_k\}$ in $\mathbb{Z}^+$ such that $x_{n_k} = x$ for all $k \in \mathbb{Z}^+$. But then $x_{n_k} \to x$.
Conversely, suppose that every sequence in E has a subsequence that converges to a point of E. Let S be any infinite subset of E. Then there exists a sequence $\{x_n\}$ in S with distinct terms. Then there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ and $x \in E$ such that $x_{n_k} \to x$. It follows immediately that x is a limit point of S. Since every infinite subset of E has a limit point in E, it follows that E is compact. ☐

Theorem 5.11. *Let (X, d) be a metric space and let $\{x_n\}$ be a sequence in X. Then the set of all subsequential limits of $\{x_n\}$ is closed.*

Proof. Let S be the set of all subsequential limits of $\{x_n\}$, and suppose $x \in S^C$. Then, since x is not a subsequential limit of $\{x_n\}$, there exists $r > 0$ such that $B_r(x)$ contains no term of $\{x_n\}$ other than possibly x itself. But then, since $B_r(x)$ is open, for each $y \in B_r(x)$, there exists $r_y > 0$ such that $B_{r_y}(y) \subset B_r(x)$, and hence, $B_{r_y}(y)$ contains no term of $\{x_n\}$. Thus, every $y \in B_r(x)$ is not a subsequential limit of $\{x_n\}$, and it follows that $B_r(x) \subset S^C$. Therefore, S^C is open, and hence S is closed. ☐

5.3.2 Cauchy Sequences, Completeness

Definition 5.12. Let (X,d) be a metric space, and let $\{x_n\}$ be a sequence in X. We say that $\{x_n\}$ is Cauchy iff

$$\forall \varepsilon > 0, \exists N \in \mathbb{Z}^+ (m, n \geq N \Rightarrow d(x_m, x_n) < \varepsilon).$$

Note. Cauchy sequences describe sequences which look like they should converge. The question posed by Cauchy sequences is "whether that limit exists in X."

Example 5.13. Let x_n equal the rational number defined by the first n digits of π:

$$x_1 = 3.1, x_2 = 3.14, x_3 = 3.141, x_4 = 3.1415, \ldots$$

Note that $m, n \geq N \Rightarrow |x_m - x_n| < 10^{-N}$. It follows that $\{x_n\}$ is Cauchy. If we regard $\{x_n\}$ as a sequence in $\mathbb{Q}$, it is Cauchy, but not convergent. Clearly though, $x_n \to \pi$, as a sequence in $\mathbb{R}$, and $\{x_n\}$, as a Cauchy sequence, "behaves" as a convergent sequence.

Lemma 5.14. *Let (X,d) be a metric space, and let $\{x_n\}$ be a Cauchy sequence in X. Then $\{x_n\}$ is bounded.*

Proof. Since $\{x_n\}$ is Cauchy, there exists $N \in \mathbb{Z}^+$ such that

$$m, n \geq N \Rightarrow d(x_m, x_n) < 1.$$

In particular,

$$n \geq N \Rightarrow d(x_n, x_N) < 1.$$

Define

$$R = \max\{d(x_1, x_N), \ldots, d(x_{N-1}, x_N), 1\}.$$

Then $d(x_n, x_N) \leq R$, $\forall n \in \mathbb{Z}^+$, and hence $\{x_n\}$ is bounded. □

Theorem 5.15.
1. If (X,d) is a metric space, and $\{x_n\}$ is a convergent sequence in X, then $\{x_n\}$ is Cauchy.
2. If (X,d) is a compact metric space, and $\{x_n\}$ is a Cauchy sequence in X, then $\{x_n\}$ converges to a point of X.
3. If $k \in \mathbb{Z}^+$, and $\{x_n\}$ is a Cauchy sequence in $\mathbb{R}^k$, then $\{x_n\}$ converges to a point of $\mathbb{R}^k$.

Proof.
1. Suppose $x_n \to x$ in X. Let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{Z}^+$, such that

$$n \geq N \Rightarrow d(x_n, x) < \frac{\varepsilon}{2}.$$

But then

$$\begin{aligned} m, n \geq N \Rightarrow (x_m, x_n) &\leq d(x_m, x) + d(x_n, x) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Then $\{x_n\}$ is Cauchy.
2. Suppose that X is compact and $\{x_n\}$ is Cauchy. Then there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ that converges to a point $x \in X$. (Either $\{x_n \mid n \in \mathbb{Z}^+\}$ is infinite, in which case it has a limit point

in X, or it is finite, in which case $x_{n_k} = x$ for some $x \in X$, and infinitely many $k \in \mathbb{Z}^+$.) Let $\varepsilon > 0$. Since $\{x_n\}$ is Cauchy, there exists $N \in \mathbb{Z}^+$ such that

$$m, n \geq N \Rightarrow d(x_m, x_n) < \frac{\varepsilon}{2}.$$

Since $x_{n_k} \to x$, and $n_k \to \infty$, there exists $K \in \mathbb{Z}^+$ such that

$$k > K \Rightarrow (d(x_{n_k}, x) < \frac{\varepsilon}{2} \text{ and } n_k \geq N).$$

For any $n \geq N$, we have

$$d(x_n, x) \leq d(x_n, x_{n_K}) + d(x_{n_K}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus, $x_n \to x$.
3. If $\{x_n\} \subset \mathbb{R}^k$ is Cauchy, then it is also bounded, and hence it belongs to a k-cell C. Since every k-cell is compact, it follows from part 2 that $\{x_n\}$ converges. □

Definition 5.16. Let (X, d) be a metric space. We say that X is complete iff every Cauchy sequence in X converges to a point of X.

Example 5.17. By the previous theorem, $\mathbb{R}^k$ is complete for all $k \in \mathbb{Z}^+$, in particular, $\mathbb{R}$ is complete. Obviously $\mathbb{Q}, \mathbb{Q}^k$ are not complete.

6 Week 6

6.1 Lecture 14.

6.1.1 Monotone Convergence

Definition 6.1. Let $\{x_n\}$ be a sequence of real numbers. We say that $\{x_n\}$ is increasing iff $x_n \leq x_{n+1}, \forall n \in \mathbb{Z}^+$, and strictly increasing iff $x_n < x_{n+1}, \forall n \in \mathbb{Z}^+$. Analog definitions exist for decreasing, and strictly decreasing sequences. A sequence is monotonic if it is either increasing or decreasing.

Theorem 6.2. *Let $\{x_n\}$ be a sequence of real numbers.*

1. *If $\{x_n\}$ is increasing and bounded above, then it converges to some $x \in \mathbb{R}$.*
2. *If $\{x_n\}$ is decreasing and bounded below, it converges to some $x \in \mathbb{R}$.*

Proof. We will prove the first result. The proof of the second is analogous. Since $\{x_n\}$ is bounded below, $x = \sup\{x_n\}$ exists. We have

$$x_n \leq x, \forall n \in \mathbb{Z}^+,$$

and, for all $\varepsilon > 0$, then exists $N \in \mathbb{Z}^+$ such that

$$x_N > x - \varepsilon$$

(otherwise $x - \varepsilon$ would be an upper bound of $\{x_n\}$ less than the least upper bound). Since $\{x_n\}$ is increasing, it follows that

$$n \geq N \Rightarrow x_n \geq X_N \Rightarrow x_n \in (x - \varepsilon, x] \subset B_\varepsilon(x).$$

This shows that $x_n \to x$. □

Definition 6.3. Let $\{x_n\}$ be a sequence of real numbers. We say that $\{x_n\}$ diverges to ∞, and write $x_n \to \infty$ iff

$$\forall M \in \mathbb{R}, \exists N \in \mathbb{Z}^+ (n \geq N \Rightarrow x_n \geq M).$$

Analogously, we say that $\{x_n\}$ diverges to $-\infty$, and write $x_n \to -\infty$ iff

$$\forall M \in \mathbb{R}, \exists N \in \mathbb{Z}^+ (n \geq N \Rightarrow x_n \leq M).$$

Lemma 6.4. *Let $\{x_n\}$ be a sequence of positive real numbers. Then $x_n \to \infty$ iff $\frac{1}{x_n} \to 0$.*

Proof. Exercise. □

Lemma 6.5. *Let $a \in \mathbb{R}^+$. Then $a^{\frac{1}{n}} \to 1$.*

Proof. Suppose that $a > 1$. Then $a^{\frac{1}{n}} > 1$, since it is easy to prove by induction that $y \leq 1 \Rightarrow y^n \leq 1$. Recall that

$$x > 0 \Rightarrow (1+x)^n = 1 + nx + \ldots + x^m > 1 + nx,$$

and thus,

$$\begin{aligned} 1 + n(a^{\frac{1}{n}} - 1) &< (1 + a^{\frac{1}{n}} - 1)^n = a \\ \Rightarrow a^{\frac{1}{n}} - 1 &< \frac{a-1}{n} \\ \Rightarrow a^{\frac{1}{n}} - 1 &\to 0 \qquad \text{(since we know that } a^{\frac{1}{n}} - 1 > 0) \\ \Rightarrow a^{\frac{1}{n}} &\to 1. \end{aligned}$$

Now suppose that $0 < a < 1$. Then $b = \frac{1}{a} > 1$, and we have

$$\begin{aligned} b^{\frac{1}{n}} &\to 1 \\ \Rightarrow (\frac{1}{a})^{\frac{1}{n}} &\to 1 \\ \Rightarrow \frac{1}{\left(\frac{1}{a}\right)^{\frac{1}{n}}} &\to 1 \\ \Rightarrow a^{\frac{1}{n}} &\to 1. \end{aligned}$$

□

Example 6.6.
1. If $p > 0$, then $n^p \to \infty$.
2. If $a > 1$, then $a^n \to \infty$.
3. If $p > 0$, and $a > 0$, then $\frac{n^p}{a^n} \to 0$.

Proof.
1. Let $M > 0$ be given. Then

$$n^P \geq < M \Leftrightarrow n \geq M^{\frac{1}{p}}.$$

Choose $N \in \mathbb{Z}^+$ such that $N \geq M^{\frac{1}{p}}$. Then

$$n \geq N \Rightarrow n \geq n \geq m^{\frac{1}{p}} \Rightarrow n^p \geq M.$$

Thus $n^p \to \infty$.

Note. We are using standard properties of powers, such as the fact that x^p is well defined for all $x \in \mathbb{R}^+, p \in \mathbb{R}$, and that $x \to x^p$ is an increasing function of x if $p > 0$. See exercise 6, Rudin, Ch 1 for the proofs of the most important properties. 2. Let $M > 0$ be given. Then

$$a^n > M \Leftrightarrow a > M^{\frac{1}{n}}.$$

By the previous lemma, $M^{\frac{1}{n}} \to 1$, and $a > 1$, by assumption. Hence, there exists $N \in \mathbb{Z}^+$ such that

$$n \geq N \Rightarrow m^{\frac{1}{n}} < a \Rightarrow a^n > M.$$

Thus, $a^n \to \infty$.
3.Suppose $p > 0$, and $a > 1$. Choose any $k \in \mathbb{Z}^+$ such that $k > p$. Write $a = (1 + \alpha)$, for some $\alpha > 0$. Note that

$$a^n = (1+\alpha)^n = \sum_{n=0}^{n} \binom{n}{j} 1^{n-j} \alpha^j = \sum_{j=0}^{n} \binom{n}{j} \alpha^j > \binom{n}{k} \alpha^k \qquad \text{(if } n \geq k))$$

Assume that $n \geq 2k$. Then

$$\begin{aligned} a^n &> \binom{n}{k} \alpha^k \\ &= \frac{n!}{(n-k)!k!} \alpha^k \\ &= \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} \alpha^k \\ &> \frac{\left(\frac{n}{2}\right)^k}{k!} \alpha^k \\ &= \frac{n^k \alpha^k}{2^k k!}. \end{aligned}$$

Thus, if $n > 2k$, we have

$$0 < \frac{n^p}{a^n} < \frac{n^p}{\frac{n^k \alpha^k}{2^k k!}} = \frac{2^k k!}{\alpha^k} n^{p-k} = C n^{p-k},$$

Since $k > p$, $n^{p-k} \to 0$ as $n \to \infty$. Thus $\frac{n^p}{a^n} \to 0$. □

Note. The last result is surprising and important. It shows that any increasing exponential grows faster than any power function. For example,

$$\frac{n^{10^6}}{(1+10^{-6})^n} \to 0.$$

6.2 Lecture 15.

6.2.1 Subsequential Limits, Limit Inferiors / Superiors

Definition 6.7. Let $\{x_n\}$ be a sequence in $\mathbb{R}$. The limit superior and limit inferior of $\{x_n\}$ are defined by

$$\limsup_{n\to\infty} = \lim_{N\to\infty} (\sup\{x_n \mid n \geq N\})$$

and

$$\liminf_{n\to\infty} = \lim_{N\to\infty} (\inf\{x_n \mid n \geq N\}).$$

Note. The sequence $\{s_N\}$ defined by $s_N = \sup\{x_n \mid n \geq N\}$ is decreasing, since $\{x_n \mid n \geq N+1\} \subset \{x_n \mid n \geq N\}$. Similarly, $\{\inf\{x_n \mid n \geq N\}\}$ is decreasing. Hence, $\limsup_{n\to\infty} x_n$ is well defined, though it can equal either a real number, of $\pm\infty$. Analogously, $\liminf_{n\to\infty} x_n$ is well defined as a real number or $\pm\infty$.

Theorem 6.8. *Let $\{x_n\}$ be a sequence of real numbers.*

1. *If $\limsup\limits_{n\to\infty} x_n = \infty$, then there is a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \to \infty$.*
2. *If $\limsup\limits_{n\to\infty} x_n = U \in \mathbb{R}$, then there is a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \to U$, and U is the largest subsequential limit of $\{x_n\}$.*
3. *If $\liminf\limits_{n\to\infty} x_n = -\infty$, then there is a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \to -\infty$.*
4. *If $\liminf\limits_{n\to\infty} x_n = L \in \mathbb{R}$, then there is a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \to L$, and L is the smallest subsequential limit of $\{x_n\}$.*

Proof.
1. Suppose $\limsup_{n\to\infty} x_n = \infty$. This implies that

$$\sup\{x_n \mid n \geq N\} = \infty, \forall N \in \mathbb{Z}^+,$$

since $\{\sup\{x_n \mid n \geq N\}\}$ is a decreasing sequence. We construct a sequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \to \infty$ as follows. Choose $n_1 \in \mathbb{Z}^+$ such that $x_{n_1} > 1$. Clearly such an n_1 exists, otherwise, $\{x_n\}$ is bounded above. Next, choose $n_2 > n_1$ such that $x_{n_1} > 2$. Continuing, suppose $n_1 < n_2 < \ldots < n_k$ so that $x_{n_j} > j$, $\forall\, 1, 2, \ldots, k$. There must exist $n_{k+1} > n_k$ such that $x_{n_{k+1}} > k+1$, otherwise, $\{x_n \mid n \geq n_{k+1}\}$ is bounded above. In this way, we've constructed a subsequence $\{x_{n_k}\}$ such that $x_{n_k} > k, \forall k$, so that $x_{n_k} \to \infty$ as desired.
2. Suppose $\limsup_{n\to\infty} x_n = x \in \mathbb{R}$. We will construct a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that

$$x_{n_k} \in \left(x - \frac{1}{k}, x + \frac{1}{k}\right), \forall k \in \mathbb{Z}^+.$$

Note that $\{\sup\{x_n \mid n \geq N\}\}$ is decreasing, hence,

$$\sup\{x_n \mid n \geq N\} \geq x, \forall N \in \mathbb{Z}^+.$$

Since

$$\sup\{x_n \mid n \geq N\} \to x \quad \text{as } N \to \infty,$$

there exists $N' \in \mathbb{Z}^+$ such that

$$N \geq N' \Rightarrow \sup\{x_n \mid n \geq N\} \in [x, x+1).$$

So there exists $n_1 \geq N'$ such that $x_{n_1} \in (x-1, x+1)$. Now suppose we have $n_1 < n_2 < \ldots < n_k$ such that

$$x_{n_j} \in \left(x - \frac{1}{j}, x + \frac{1}{j}\right), \quad \forall j = 1, \ldots, k.$$

Since

$$\sup\{x_n \mid n \geq N\} \to x \quad \text{as } N \to \infty,$$

there exists $N' \in \mathbb{Z}^+$ such that

$$N \geq N' \Rightarrow \sup\{x_n \mid n \geq N\} \in \left[x, x + \frac{1}{k+1}\right).$$

Hence, there exists $n_{k+1} > \max\{N', n_k\}$ such that

$$x_{n_{k+1}} \in \left(x - \frac{1}{k+1}, x + \frac{1}{k+1}\right),$$

$\big($if not, $\sup\{x_n \mid n > n_k\} \leq x - \frac{1}{k+1}\big)$. Thus, there exists subsequence $\{x_{n_k}\}$ such that

$$x_{n_k} \in \left(x - \frac{1}{k}, x + \frac{1}{k}\right), \forall k \in \mathbb{Z}^+,$$

and clearly $x_{n_k} \to x$.
(Parts 3, 4 are proven analogously.) □

6.2.2 Series

Looking forward, recall from Calculus that

$$\forall x \in \mathbb{R}, e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$
$$e = e^1 = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots.$$

Proposition 6.9. *The number e is irrational.*

Proof. Suppose e is rational, say, $e = \frac{p}{q}$, where $p, q \in \mathbb{Z}^+$. Then

$$\frac{p}{q} = e = \sum_{n=0}^{q} \frac{1}{n!} + \sum_{n=q+1}^{\infty} \frac{1}{n!}$$
$$\Rightarrow \frac{q!p}{q} = \sum_{n=0}^{q} \frac{q!}{n!} + \sum_{n=q+1}^{\infty} \frac{q!}{n!}.$$

Note that

$$\frac{q!p}{q} = p(q-1)! \in \mathbb{Z}^+$$

and

$$\sum_{n=0}^{q} \frac{q!}{n!} \in \mathbb{Z}^+ \quad \text{(since } \frac{q!}{n!} = q(q-1)\cdots(n+1), \forall n \leq q).$$

Also, for $n > q$,

$$\frac{q!}{n!} = \frac{1}{(q+1)\cdots n} \leq \frac{1}{(q+1)^{n-q}}.$$

Thus,

$$\begin{aligned}
\sum_{n=q+1}^{\infty} \frac{q!}{n!} &\leq \sum_{n=q+1}^{\infty} \left(\frac{1}{q+1}\right)^{n-q} \\
&= \sum_{k=1}^{\infty} \left(\frac{1}{q+1}\right)^{k} \\
&= \frac{1}{1-\frac{1}{q+1}} - 1 \\
&= \frac{q+1}{q} - \frac{q}{q} = \frac{1}{q} < 1.
\end{aligned}$$

But then

$$\frac{q!p}{q} = \sum_{n=0}^{q} \frac{q!}{n!} + \sum_{n=q+1}^{\infty} \frac{q!}{n!},$$

where the LHS is an integer, and the RHS is an integer plus a number in $(0,1)$, but there are no integers in $(0,1)$, so e cannot be rational. □

6.3 Lecture 16.

6.3.1 Convergence of Series

Definition 6.10. Let $\{x_n\}$ be a sequence of real numbers. We call

$$\sum_{n=1}^{\infty} x_n = x_1 + x_2 + \dots$$

the infinite series with terms $x_1, x_2, \dots$. For each $N \in \mathbb{Z}^+$,

$$\sum_{n=1}^{N} x_n = x_1 + x_2 + \dots + x_N$$

is called a partial sum of the infinite series. We say that an infinite series converges, and write

$$\sum_{n=1}^{\infty} x_n = S$$

iff the sequence of partial sums converges to S, that is

$$\lim_{n\to\infty} \sum_{n=1}^{N} x_n = S.$$

The series diverges iff the sequence of partial sums diverges.

Proposition 6.11. *Recall the following algebraic identity:*

$$(1-r)(1+r+r^2+\dots+r^N) = 1 - r^{N+1}.$$

Proof. We have

$$(1-r)(1+r+r^2+\ldots+r^N) = 1+r+r^2+\ldots+r^N \\ -(r+r^2+r^3+\ldots+r^{N+1}) \\ = 1-r^{N+1}.$$

□

Example 6.12. Consider

$$\sum_{n=0}^{N} r^n = 1+r+\ldots+r^N = \frac{1-r^{N+1}}{1-r}$$

$$\Rightarrow \lim_{N\to\infty}\sum_{n=0}^{N} r^n = \frac{1}{1-r}, \text{ provided } |r|<1$$

$$\Rightarrow \sum_{n=0}^{\infty} r^n = \frac{1}{1-r}, \text{ if } |r|<1.$$

If $|r| > 1$, then $\sum_{n=0}^{\infty} r^n$ diverges. We call such a series a geometric series.

Note. Note that if $M > N$, then

$$\sum_{n=1}^{M} x_n - \sum_{n=1}^{N} x_n = \sum_{n=N+1}^{M} x_n.$$

We call the following theorem the Cauchy criterion for convergence of series.

Theorem 6.13. *Let $\{x_n\}$ be a sequence of real numbers. Then*

$$\sum_{k=1}^{\infty} x_k$$

converges iff for all $\varepsilon > 0$ there exists $N \in \mathbb{Z}^+$ such that

$$m \geq n \geq N \Rightarrow \left|\sum_{k=n}^{m} x_k\right| < \varepsilon.$$

Proof. Just apply the Cauchy criterion for sequences to the sequence of partial sums. □

Theorem 6.14. *If $\{x_n\}$ is a sequence of real numbers and*

$$\sum_{n=1}^{\infty} x_n$$

converges, then $x_n \to 0$.

Proof. Let $\varepsilon > 0$. Since the series converges, the previous theorem implies that there exists $N \in \mathbb{Z}^+$ such that

$$m \geq n \geq N \Rightarrow \left|\sum_{k=n}^{m} x_k\right| < \varepsilon \\ \Rightarrow |x_n| < \varepsilon.$$

Thus $x_n \to 0$. □

Note. The converse of the previous theorem is not true.

Example 6.15. Consider the harmonic series

$$\sum_{n=1}^{\infty} \frac{1}{n}.$$

Write

$$S_N = \sum_{n=1}^{N} \frac{1}{n},$$

and note that

$$\begin{aligned}
S_1 &= 1 \\
S_2 &= 1 + \frac{1}{2} \\
S_4 &= S_2 + \frac{1}{3} + \frac{1}{4} > S_2 + \frac{2}{4} \geq 1 + 2\frac{1}{2} \\
S_8 &= S_4 + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > S_4 + \frac{4}{8} \geq 1 + 3\frac{1}{2} \\
S_{16} &= S_8 + \frac{1}{9} + \ldots + \frac{1}{16} > S_8 + \frac{8}{16} \geq 1 + 4\frac{1}{2}.
\end{aligned}$$

Continuing in this fashion, it is easy to show that

$$S_{2^n} > 1 + \frac{n}{2}.$$

Since $\{S_n\}$ is increasing, it follows that $S_n \to \infty$.

Theorem 6.16. *Let $\{x_n\}$ be a sequence of nonnegative real numbers. Then $\sum_{n=1}^{\infty} x_n$ converges iff the sequence of partial sums is bounded.*

Proof. This follows immediately from the fact

$$x_n \geq 0,\ \forall n \Rightarrow \left\{ \sum_{n=1}^{N} x_n \right\} \text{ is increasing.}$$

□

Theorem 6.17 (Cauchy condensation test)**.** *Let $\{x_n\}$ be a decreasing sequence of nonnegative real numbers. Then $\sum_{n=1}^{\infty} x_n$ converges iff $\sum_{k=0}^{\infty} 2^k x_{2^k}$ converges.*

Proof. Since the partial sums of both sequences are increasing, it suffices to prove that

$$\sum_{k=1}^{2^n-1} x_k \leq \sum_{k=0}^{n-1} 2^k x_{2^k} \leq 2 \sum_{k=1}^{2^{n-1}} x_k. \tag{*}$$

We have

$$\sum_{k=1}^{2^1-1} x_k = x_1,$$
$$\sum_{k=0}^{0} 2^k x_{2^k} = x_1,$$
$$2\sum_{k=1}^{2^{1-1}} x_k = 2x_1.$$

So (*) holds for $n = 1$. Assume that $(*)$ holds for some $n \geq 1$. Then

$$\begin{aligned}
\sum_{k=1}^{2^{n+1}-1} x_k &= \left(\sum_{k=1}^{2^n-1} x_k\right) + \overbrace{x_{2^n} + \ldots + x_{2^{n+1}-1}}^{2^n \text{ terms}} \\
&\leq \left(\sum_{k=0}^{n-1} 2^k x_{2^k}\right) + 2^n x_{2^n} && (\text{since } x_{2^n+j} \leq x_{2^n}, \forall j \geq 0) \\
&= \sum_{k=0}^{n} 2^k x_{2^k} \\
&= \left(\sum_{k=0}^{n-1} 2^k x_{2^k}\right) + 2^n x_{2^n} \\
&\leq 2\left(\sum_{k=1}^{2^{n-1}} x_k\right) + 2\left(2^{n-1} x_{2n}\right) \\
&\leq 2\left(\sum_{k=1}^{2^{n-1}} x_k\right) + 2\left(x_{2^{n-1}+1} + \ldots + x_{2^n}\right) \\
&= 2\sum_{k=1}^{2^n} x_k.
\end{aligned}$$

Then by the principle of mathematical induction, our result follows. □

7 Week 7

7.1 Lecture 17. Wed Oct 9

7.1.1 Series Convergence Tests

Theorem 7.1 (Comparison Test)**.** *Suppose* $\{x_n\}, \{y_n\}$ *are sequences of real numbers and there exists* $n_0 \in \mathbb{Z}^+$ *such that* $|x_n| \leq y_n$, $\forall n \geq n_0$*. If* $\sum_{n=1}^{\infty} y_n$ *converges, then* $\sum_{n=1}^{\infty} x_n$ $\left(\textit{or } \sum_{n=1}^{\infty} |x_n|\right)$ *converges.*

Proof. Let $\varepsilon > 0$ be given. Since $\sum_{n=1}^{\infty} y_n$ converges, there exists $N \in \mathbb{Z}^+$, $N \geq n_0$, such that

$m \geq n \geq N$ implies $\sum_{k=n}^{m} y_k < \varepsilon$. Then

$$
\begin{aligned}
m \geq n \geq N \Rightarrow \left| \sum_{k=n}^{m} x_k \right| &\leq \sum_{k=n}^{m} |x_k| && \text{(triangle inequality)} \\
&\leq \sum_{k=n}^{m} y_k && \text{(since } N \geq n_0\text{)} \\
&< \varepsilon.
\end{aligned}
$$

But this means that $\sum_{n=1}^{\infty} x_n$ satisfies the Cauchy criterion, and hence, $\sum_{n=1}^{\infty} x_n$ converges. □

Corollary 7.1.1. *If $\{x_n\}, \{y_n\}$ are sequences of nonnegative real numbers, if there exists $n_0 \in \mathbb{Z}^+$ such that $x_n \leq y_n$ for all $n \geq n_0$, and if $\sum_{n=1}^{\infty} x_n$ diverges, then $\sum_{n=1}^{\infty} y_n$ diverges.*

Example 7.2.

1. Does $\sum_{n=1}^{\infty} \frac{1}{n+100}$ converge or diverge?

Solution. Note that

$$\sum_{n=1}^{\infty} \frac{1}{2n}$$

diverges to ∞, and

$$\frac{1}{n+100} \geq \frac{1}{2n}, \forall n \geq 100.$$

Thus the given series diverges. □

2. Does $\sum_{n=1}^{\infty} \frac{1}{n2^n}$ converge or diverge?

Solution. Since

$$\frac{1}{n2^n} \leq \frac{1}{2^n}, \forall n \geq 1,$$

and $\sum_{n=1}^{\infty} 2^{-n}$ is a convergent geometric series, we conclude that the given series converges. □

Theorem 7.3 (Root Test)**.** *Let $\{x_n\}$ be a sequence of real numbers and let $L = \limsup_{n\to\infty} |x_n|^{\frac{1}{n}}$. If $L < 1$, then $\sum_{n=1}^{\infty} x_n$ converges. If $L > 1$, then $\sum_{n=1}^{\infty} x_n$ diverges.*

Proof. Suppose $L < 1$, and choose $r \in (L, 1)$. Since $\limsup_{n\to\infty} |x_n|^{\frac{1}{n}} = L < r$, then there exists some $N \in \mathbb{Z}^+$ such that $n \geq N$ implies that $\sup\{|x_n|^{\frac{1}{n}} \mid n \geq N\} < r$. Thus $n \geq N$ implies that $|x_n| \leq r^n$. Since $\sum_{n=1}^{\infty} r^n$ is convergent, $\sum_{n=1}^{\infty} x_n$ converges by the comparison test.

Now, suppose $L > 1$. Then there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $|x_{n_k}|^{\frac{1}{n_k}} \to L$. Write $r = |L-1|$. This implies that there exists $N \in \mathbb{Z}^+$ such that $k \geq N \Rightarrow ||x_{n_k}|^{\frac{1}{n_k}} - L| < \frac{r}{2} \implies |x_{n_k}|^{\frac{1}{n_k}} > 1$. But then $n \geq N \Rightarrow |x_{n_k}| \geq 1$. Hence, $x_n \not\to 0$, and thus $\sum_{n=1}^{\infty} x_n$ diverges. □

Note. If $\limsup_{n\to\infty} |x_n|^{\frac{1}{n}} = 1$, then the series might converge or it might not diverge.

Theorem 7.4 (Ratio Test). *Let $\{x_n\}$ be a sequence of nonzero real numbers (or at least $x_n \neq 0$ for all n sufficiently large).*

1. *If $\limsup_{n\to\infty} \left|\frac{x_{n+1}}{x_n}\right| = L < 1$, then the series converges.*
2. *If there exists $N \in \mathbb{Z}^+$ such that*
$$\left|\frac{x_{n+1}}{x_n}\right| \geq 1, \forall\, n \geq N,$$
then $\sum_{n=1}^{\infty} x_n$ diverges.

Proof.
1. Since $\limsup_{n\to\infty} \left|\frac{x_{n+1}}{x_n}\right| = L < 1$, we can choose $r \in (L, 1)$, such that there exists N such that
$$n \geq N \Rightarrow \left|\frac{x_{n+1}}{x_n}\right| \leq r.$$
That is, $n \geq N \Rightarrow |x_{n+1}| \leq r|x_n|$. Hence,
$$\begin{aligned} |x_{N+1}| &\leq r|x_N| \\ |x_{N+2}| &\leq r|x_{N+1}| \leq r^2|x_N| \\ |x_{N+3}| &\leq r|x_{N+2}| \leq r^3|x_N| \\ &\vdots \\ |x_{N+k}| &\leq r^k|x_N|, \forall k \geq 0. \end{aligned}$$
Since $\sum_{k=0}^{\infty} r^k|x_N| = |x_N|\sum_{k=0}^{\infty} r^k$ is a convergent geometric series, it follows from the comparison test that $\sum_{k=0}^{\infty} x_{N+k}$ and hence $\sum_{n=1}^{\infty} x_n$ converges.
2. Suppose there exists $N \in \mathbb{Z}^+$ such that $n \geq N \Rightarrow \left|\frac{x_{n+1}}{x_n}\right| \geq 1$. Then $n \geq N \Rightarrow |x_{n+1}| \geq |x_n|$, which implies that for all $n \geq N$, $|x_n| \geq |x_N| > 0$. Thus, $x_n \not\to 0$, and the series must diverge. □

7.2 Lecture 18. Fri Oct 11

Theorem 7.5 (Alternating series test). *Suppose $\{x_n\}$ is a sequence of real positive numbers, and that $\{x_n\}$ is decreasing, and $x_n \to 0$, Then*
$$\sum_{n=1}^{n}(-1)^{n+1}x_n \text{ and } \sum_{n=1}^{\infty}(-1)^n x_n$$
both converge.

Proof. Write $S_N = \sum_{n=1}^{N}(-1)^{n+1}x_n$. Then

$$\begin{aligned} S_1 &= x_1 > 0 \\ S_2 &= S_1 - x_2 \in [0, S_1) && (x_1 > 0, x_2 \le x_1) \\ S_3 &= S_2 + x_3 \in (S_2, S_1] && (x_3 > 0, x_3 \le x_2) \\ S_4 &= S_3 - x_4 \in [S_2, S_3). \end{aligned}$$

Note that in general

$$\begin{aligned} S_{2k+2} &= S_{2k} + x_{2k+1} - x_{2k+2} \\ &\ge S_{2k} && (\text{since } x_{2k+2} \le s_{2k+1}) \\ &\Rightarrow \{S_{2k}\} \text{ is increasing.} \end{aligned}$$

Similarly,

$$\begin{aligned} S_{2k+3} &= S_{2k+1} - x_{2k+2} + x_{2k+3} \\ &\le S_{2k+1}. && (\text{since } x_{2k+2} \ge x_{2k+3}) \\ &\Rightarrow \{S_{2k+1}\} \text{ is decreasing.} \end{aligned}$$

Also, every even term is less than every odd term. Consider $S_{2k+1}, S_{2\ell}$. If $k \ge 1$, then

$$S_{2\ell} \le S_{2k} < S_{2k} + S_{2k+1} = S_{2k+1}.$$

If $k < \ell$, then

$$S_{2k+1} \ge S_{2\ell+1} = S_{2\ell} + x_{2\ell+1} > S_{2\ell}.$$

Thus $\{S_{2k}\}$ and $\{S_{2k+1}\}$ are bounded and monotone, and have $S_{2k} \to S_*$, and $S_{2k+1} \to S^*$, with $S_* \le S^*$. Finally, for all k,

$$S_{2k} \le S_* \le S^* \le S_{2k} + x_{2k+1}, \text{ for all } k.$$

Since $x_{2k+1} \to 0$, we see that $S_* = S^*$. Thus

$$\sum_{n=1}^{\infty}(-1)^{n+1}x_n = S_* = S^*.$$

□

Theorem 7.6.

1. If $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ converge, then $\sum_{n=1}^{\infty}(x_n + y_n)$ converges.

2. If $\sum_{n=1}^{\infty} x_n$ converges, and $c \in \mathbb{R}$, then $\sum_{n=1}^{\infty} cx_n$ converges, and

$$\sum_{n=1}^{\infty} cx_n = c\sum_{n=1}^{\infty} x_n.$$

Theorem 7.7. *Let $\{x_n\}$ be a sequence of real numbers, and suppose $\sum_{n=1}^{\infty} |x_n|$ converges, then $\sum_{n=1}^{\infty} x_n$ converges.*

Proof. By the Cauchy criterion, for all $\varepsilon > 0$, there exists an $N \in \mathbb{Z}^+$, such that

$$
\begin{aligned}
& m \geq n \geq N \Rightarrow \left(\Big| \sum_{n=1}^{\infty} |x_k| \Big| < \varepsilon \right) \\
\Longrightarrow\ & m \geq n \geq N \Rightarrow \left(\sum_{n=1}^{\infty} |x_k| < \varepsilon \right) \\
\Longrightarrow\ & m \geq n \geq N \Rightarrow \left(\Big| \sum_{n=1}^{\infty} x_k \Big| \leq \sum_{n=1}^{\infty} |x_k| < \varepsilon \right)
\end{aligned}
$$

Thus $\sum_{n=1}^{\infty} x_n$ converges. □

7.2.1 Absolute Convergence of Series

Definition 7.8. Let $\{x_n\}$ be a sequence of real numbers. We say that $\sum_{n=1}^{\infty} x_n$ converges absolutely iff $\sum_{n=1}^{\infty} |x_n|$ converges. We say that $\sum_{n=1}^{\infty} x_n$ converges conditionally iff $\sum_{n=1}^{\infty}$ converges but $\sum_{n=1}^{\infty} |x_n|$ diverges.

Question. How do you write a product of two series?

$$
\begin{aligned}
\left(\sum_{m=0}^{\infty} x_m \right) \left(\sum_{n=0}^{\infty} y_n \right) &= \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} x_m \right) y_n \\
&= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} x_m y_n ?
\end{aligned}
$$

Or,

$$
\begin{aligned}
\left(\sum_{m=0}^{\infty} x_m \right) \left(\sum_{n=0}^{\infty} y_n \right) &= \sum_{m=0}^{\infty} \left(\sum_{n=0}^{\infty} y_n \right) x_m \\
&= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} x_m y_n ?
\end{aligned}
$$

Or,

$$
\left(\sum_{m=0}^{\infty} x_m \right) \left(\sum_{n=0}^{\infty} y_n \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\infty} x_k y_{n-k} \right) ?
$$

Theorem 7.9 (Riemann)**.** *Let $\{x_n\}$ be a sequence of real numbers such that $\sum_{n=1}^{\infty} x_n$ converges conditionally. Let L be any (extended) real number. Then there exists a rearrangement of the series, that is,* *there exists a bijection $m : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that*

$$\sum_{k=1}^{\infty} x_{m(k)} = L.$$

Definition 7.10. Let $\{x_n\}, \{y_n\}$ be sequences of real numbers ($n = 0, 1, 2, \ldots$). Then the product of $\sum_{n=0}^{\infty} x_n$ and $\sum_{n=0}^{\infty} y_n$ is

$$\Big(\sum_{n=0}^{\infty} x_n\Big)\Big(\sum_{m=0}^{\infty} y_m\Big) = \sum_{n=0}^{\infty}\Big(\sum_{k=0}^{\infty} x_k y_{n-k}\Big)$$

Theorem 7.11. *Let $\{x_n\}$ be a sequence of real numbers. If $\sum_{n=1}^{\infty} x_n$ converges absolutely, then every rearrangement of $\sum_{n=1}^{\infty} x_n$ converges to the same limit.*

Proof. Let $m : \mathbb{Z}^+ \to \mathbb{Z}^+$ be a bijection. We want to prove that $\sum_{k=1}^{\infty} x_{m(k)}$ converges to $\sum_{n=1}^{\infty} x_n$. Let ε be given. Then there exists $N \in \mathbb{Z}^+$ such that

$$p \geq n \geq N \Rightarrow \sum_{k=n}^{p} |x_k| < \varepsilon.$$

Choose $M \in \mathbb{Z}^+$ such that $\{1, 2, \ldots, N-1\} \subset \{m(1), m(2), \ldots, m(M-1)\}$. Then, for $\ell \geq M$, write

$$\begin{aligned} J_1 &= \{1, 2, \ldots, \ell\} \\ J_2 &= \{m(1), m(2), \ldots, m(\ell)\} \\ J &= (J_1 \cup J_2) \setminus (J_1 \cap J_2). \end{aligned}$$

Then

$$\begin{aligned} |S'_\ell - S_\ell| &= \Big|\sum_{k=1}^{\ell} x_{m(k)} - \sum_{k=1}^{\ell} x_k\Big| \\ &= \Big|\sum_{j \in J} x_j\Big| \\ &\leq \sum_{j \in J} |x_j| < \varepsilon \qquad \text{(since } \{1, 2, \ldots, N\} \cap J = \varnothing) \end{aligned}$$

Thus,

$$\lim_{p\to\infty} s'_p = \lim_{p\to\infty} s_p = s.$$

□

8 Week 8

8.1 Lecture 19.

8.1.1 Products of Absolutely Convergent Series

Theorem 8.1. *Let $\{x_n\}, \{y_n\}$ be sequences of real numbers. If at least one of $\sum_{n=0}^{\infty} x_n, \sum_{n=0}^{\infty} y_n$ converges absolutely, then their product converges, and*

$$\sum_{n=0}^{\infty}\left(\sum_{n=0}^{\infty} x_k y_{n-k}\right) = \left(\sum_{n=0}^{\infty} x_n\right)\left(\sum_{n=0}^{\infty} y_n\right).$$

Proof. WLOG, assume $\sum_{n=0}^{\infty} x_n$ converges absolutely, and $\sum_{n=0}^{\infty} y_n$ converges. Define

$$s = \sum_{n=0}^{\infty} x_n, \quad t = \sum_{n=0}^{\infty} y_n,$$

and for all $n \in \mathbb{Z}^+$,

$$S_n = \sum_{k=0}^{n} x_k, \quad t_n = \sum_{k=0}^{n} y_k, \quad u_n = \sum_{k=0}^{\infty} z_k,$$

where

$$z_k = \sum_{j=0}^{k} x_j y_{k-j}.$$

Also, write

$$d_n = t_n - t, \ \forall n \in \mathbb{Z}^+.$$

Note that

$$\begin{aligned}
u_n &= x_0y_0 + (x_0y_1 + x_1y_0) + (x_0y_2 + x_1y_1 + x_2y_0) + \ldots + (x_0y_n + x_1y_{n-1} + \ldots + x_ny_0) \\
&= x_0(y_0 + y_1 + \ldots + y_n) + x_1(y_0 + y_1 + \ldots + y_{n-1}) + x_2(y_0 + y_1 + \ldots + y_{n-2}) + \ldots + x_ny_0 \\
&= x_0t_n + x_1t_{n-1} + x_2t_{n-2} + \ldots + x_nt_0 \\
&= x_0(t + d_n) + x_1(t + x_{n-1}) + x_2(t + d_{n-2}) + \ldots + x_n(t + d_0) \\
&= (x_0 + x_1 + \ldots + x_n)t + x_0d_n + x_1d_{n-1} + \ldots + x_nd_0 \\
&= s_nt + \sum_{k=0}^{n} d_kx_{n-k}.
\end{aligned}$$

We know that $s_nt \to st$, and we want to prove that $u_n \to st$. Thus, it suffices to prove that $\sum_{k=0}^{n} d_kx_{n-k} \to 0$. Let $\varepsilon > 0$ be given. Since $t = \sum_{n=0}^{\infty} y_n$,

$$d_n = t_n - t \to 0.$$

Hence, there exists $N \in \mathbb{Z}^+$ such that

$$n \geq N \Rightarrow |d_n| < \frac{\varepsilon}{2S'},$$

where $S' = \sum_{n=0}^{\infty} |x_n|$. But then

$$\begin{aligned} n \geq N \Rightarrow \left| \sum_{k=0}^{n} d_k x_{n-k} \right| &\leq \sum_{k=0}^{n} |d_k||x_{n-k}| \\ &= \sum_{k=0}^{N-1} |d_k||x_{n-k}| + \sum_{k=N}^{n} |d_k||x_{n-k}| \\ &< \sum_{k=0}^{N-1} |d_k||x_{n-k}| + \frac{\varepsilon}{2S'} \sum_{k=N}^{n} |x_{n-k}|. \end{aligned}$$

Now,

$$\sum_{k=N}^{n} |x_{n-k}| \leq \sum_{k=0}^{\infty} |x_k| = S'.$$

Also, if $M = \sum_{k=0}^{N-1} |d_k|$, then there exists $N' \in \mathbb{Z}^+$ such that

$$n \geq N' \Rightarrow |x_n| < \frac{\varepsilon}{2M}.$$

But then

$$\begin{aligned} n \geq N + N' \Rightarrow n - (N-1) \geq N' &\Rightarrow |x_{n-k}| < \frac{\varepsilon}{2M}, \ \forall k = 0, 1, \ldots, N-1 \\ &\Rightarrow \sum_{k=0}^{N-1} |d_k||x_{n-k}| < \frac{\varepsilon}{2M} \sum_{k=0}^{N-1} |d_k| = \frac{\varepsilon}{2M} \cdot M = \frac{\varepsilon}{2}. \end{aligned}$$

We thus obtain

$$n \geq N + N' \Rightarrow \left| \sum_{k=0}^{n} d_k x_{n-k} \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence,

$$\sum_{k=0}^{n} d_k x_{n-k} \to 0$$

as desired. □

8.1.2 Power Series

Definition 8.2. Let $\{c_n\}$ be a sequence of real numbers, and $a \in \mathbb{R}$. We can define a real-valued function f by

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n,$$

where the domain is the set of all x for which the series converges. Such a series is called a power series.

Proposition 8.3. *We can almost determine the domain of f by the root test. We compute*

$$\limsup_{n\to\infty} |c_n(x-a)^n|^{\frac{1}{n}} = \limsup_{n\to\infty} |c_n|^{\frac{1}{n}}|x-a|$$
$$= |x-a|(\limsup_{n\to\infty} |c_n|^{\frac{1}{n}}).$$

We see that the series converges iff

$$|x-a| < R = \frac{1}{\limsup_{n\to\infty} |c_n|^{\frac{1}{n}}},$$

and diverges iff

$$|x-a| > R.$$

Note. Consider two special cases:

1. $\limsup\limits_{n\to\infty} |c_n|^{\frac{1}{n}} = 0$. Then the series converges for all $x \in \mathbb{R}$. We write $R = \infty$, and the domain of f is $\mathbb{R} = (-\infty, \infty)$.

2. If $\limsup\limits_{n\to\infty} |c_n|^{\frac{1}{n}} = \infty$, then the series converges only for $x = a$. We write $R = 0$, and the domain of f is the degenerate interval $[a, a]$.

Note. If

$$0 < \limsup_{n\to\infty} |c_n|^{\frac{1}{n}} < \infty,$$

then $0 < R < \infty$, and the domain of f contains

$$(a-R, a+R).$$

It may or may not contain $a-R$ and $a+R$ (recall that the ratio test is inconclusive for $L = 1$). Thus, the domain in this case is one of the following:

$$(a-R, a+R), [a-R, a+R), (a-R, a+R], [a-R, a+R].$$

We call R the radius of convergence of the power series, and the interval of convergence is the domain of f (which is always an interval if we are willing to call $[a, a]$ an interval).

Note. If we start with $a \in \mathbb{R}$, and $f : \mathbb{R} \to \mathbb{R}$ (or $f : I \to \mathbb{R}$, where I is an open interval containing a), then we can define the Taylor series of f at a (or in powers of $x - a$) as follows:

- $p_0(x) = f(x)$ is the unique constant polynomial that agrees with f at $x = 0$.
- Set $p_1(x) = f(a) + f'(a)(x-a)$ as the unique linear polynomial that agrees with f and f' at $x = a$. (That is, $p(a) = f(a), p'(a) = f'(a)$.)
- Set $p_2(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2$ as the unique quadratic polynomial that agrees with f, f', and f'' at $x = a$.
- In general, let

  $$p_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

  be the unique polynomial of degree n that agrees with $f, f', \ldots, f^{(n)}$ at $x = a$.

Thus, it is natural to define the Taylor series of f at $x = a$ (or in powers of $x - a$) by

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

(assuming f is infinitely differentiable). We can then ask:

- does the series converge?
- Does the series converge to $f(x)$?

8.2 Lecture 20. Wed Oct 16

8.2.1 Continuity

Definition 8.4 (Functional convergence)**.** Let $(X, d_x), (Y, d_y)$ be metric spaces. Let $E \subset X$, and let p be a limit point of E, and let $f : E \to Y$. We say $f(x)$ converges to $q \in Y$ as $x \to p$, written

$$\lim_{x \to p} f(x) = q \text{ or } f(x) \to q \text{ as } x \to p.$$

iff

$$\forall \varepsilon > 0 \exists \delta > 0 (x \in E \text{ and } 0 < d_x(x,p) < \delta) \Rightarrow d_y(f(x), q) < \varepsilon.$$

Note that the condition $0 < d_x(x,p) < \delta$ implies that the value of $f(p)$ (if it exists, p may not belong to E) is irrelevant.

Theorem 8.5. *Let $(X, d_x), (Y, d_y)$ be metric spaces, let $E \subset X$, let p be a limit point of E, and let $f : E \to Y$. Then*

$$\lim_{x \to p} f(x) = q \tag{*}$$

iff

$$\forall \{p_n\} \subset E(p_n \neq p, \forall n \wedge p_n \to p) \Rightarrow f(p_n) \to q. \tag{**}$$

Proof. Suppose first that (*) holds. Let $\{p_n\} \subset E$ satisfy $p_n \neq p, \forall n \in \mathbb{Z}^+$, and $p_n \to p$. We must prove that $f(p_n) \to q$. Let $\varepsilon > 0$ be given. Since $f(x) \to q$, as $x \to p$, there exists $\delta > 0$ such that

$$(x \in E \wedge 0 < d_x(x,p) < \delta) \Rightarrow d_y(f(x), q) < \varepsilon.$$

Since $p_n \to p$ and $p_n \neq p, \forall n \in \mathbb{Z}^+$, there exists $N \in \mathbb{Z}^+$ such that

$$n \geq N \Rightarrow 0 < d_x(p_n, p) < \delta.$$

But then

$$\begin{aligned} n \geq N &\Rightarrow 0 < d_x(p_n, p) < \delta \\ &\Rightarrow d_y(f(p_n), q) < \varepsilon. \end{aligned}$$

Thus $f(p_n) \to q$. Conversely, suppose (*) fails. Therefore, there exists $\varepsilon > 0$ such that

$$\forall \delta > 0, \exists x \in E(B_\delta(p) \setminus \{p\}), d_y(f(x), q) \geq \varepsilon.$$

In particular, $\forall n \in \mathbb{Z}^+$, there exists

$$p_n \in E \cap (B_{y_n}(p) \setminus \{p\}) \text{ such that } d_y(f(p_n), q) \geq \varepsilon.$$

Then $\{p_n\} \subset E$, $p_n \neq p \forall n$, and $p_n \to p$, but $f(p_n) \not\to q$ (since $d_Y(f(p_n), q) \geq \varepsilon, \forall n$). Thus, (**) fails. ☐

Corollary 8.5.1. *Given the same hypotheses as above, if* $\lim_{x\to p} f(x)$ *exists, it is unique.*

Proof. This follows from the fact that sequential limits are unique. □

Corollary 8.5.2. *Let* (X, d_x) *be metric spaces, let* $E \subset X$*, let* p *be a limit point of* E*, and let* $f : E \to \mathbb{R}, g : E \to \mathbb{R}$*. If* $\lim_{x\to p} f(x), \lim_{x\to p} g(x)$ *both exist, then*

1. $$\lim_{x\to p}(f+g)(x) = \lim_{x\to p} f(x) + \lim_{x\to p} g(x),$$

2. $$\lim_{x\to p}(f-g)(x) = \lim_{x\to p} f(x) - \lim_{x\to p} g(x),$$

3. $$\lim_{x\to p}(fg)(x) = \Big(\lim_{x\to p} f(x)\Big)\Big(\lim_{x\to p} g(x)\Big),$$

4. $$\lim_{x\to p}(f/g)(x) = \frac{\lim_{x\to p} f(x)}{\lim_{x\to p} g(x)} \text{ if } \lim_{x\to p} g(x) \neq 0,$$

   Also, if $c \in \mathbb{R}$*,*

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

Definition 8.6. Let $(X, d_x), (Y, d_y)$ be metric spaces. Let $E \subset X$ and let $f : E \to Y$, and suppose $p \in E$. We say that f is continuous at p iff

$$\forall \varepsilon > 0 \exists \delta > 0 (x \in E \wedge d_x(x,p) < \delta) \Rightarrow d_y(f(x), f(p)) < \varepsilon.$$

Note. If $p \in E$ is an isolated point of E, ($\exists \delta > 0, B_\delta(p) \cap E = \{p\}$), then f is automatically continuous at P. (The same δ works for every $\varepsilon > 0$.)

Theorem 8.7. *Let* $(X, d_y), (Y, d_y)$ *be metric spaces. Let* $E \subset X$*, let* $f : E \to Y$*, and let* $p \in E$ *be a limit point of* E*. Then* f *is continuous at* p *iff*

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

Proof. Just compare the definitions:

$$f \text{ is continuous at } p \Leftrightarrow \forall \varepsilon > 0, \exists \delta > 0 (x \in E \wedge d_x(x,p) < \delta) \Rightarrow d_y(f(x), f(p)) < \varepsilon,$$

$$\lim_{x\to p} f(x) = f(p) \Leftrightarrow \forall \varepsilon > 0, \exists \delta > 0 (x \in E \wedge 0 < d_x(x,p) < \delta) \Rightarrow d_y(f(x), f(p)) < \varepsilon.$$

□

Theorem 8.8. *Let* $(X, d_x), (Y, d_x), (Z, d_z)$ *be metric spaces. Let* $E \subset X, F \subset Y$*, and let* $f : E \to Y, g : F \to Z$*, and assume* $\operatorname{range} F \subset F$*. Define* $h = g \circ f$ *(i.e.* $h(x) = g(f(x)), \forall x \in E$*). If* $p \in E$*,* f *is continuous at* p*,* g *is continuous at* $f(p)$*, then* h *is continuous at* p*.*

Proof. Let $\varepsilon > 0$ be given. Since g is continuous at $f(p)$, there exists $\delta > 0$ such that

$$(y \in F \wedge d_y(y, f(p)) < \delta) \Rightarrow d_z(g(y), g(f(p))) < \varepsilon.$$

Since f is continuous at p, there exists $\gamma > 0$ such that

$$\begin{aligned}(X \in E \wedge d_x(x,p) < \gamma) &\Rightarrow d_y(f(x), f(p)) < \delta \\ &\Rightarrow d_z(g(f(x)), g(f(p))) < \varepsilon \\ &\Rightarrow d_x(h(x), h(p)) < \varepsilon.\end{aligned}$$

Thus, h is continuous at p. □

Theorem 8.9. *Let (X, d) be a metric space, let $E \subset X$, and let $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}$. If f, g are continuous at $p \in E$, then so are $f \pm g$, fg, and f/g (provided $g(p) \neq 0$ for that last statement). Also cf is continuous at p for every $c \in \mathbb{R}$.*

Proof. Because f, g continuous at p, we have

$$\begin{aligned}&\lim_{x\to p} f(x) = f(p), \lim_{x\to p} g(x) = g(p) \\ &\Rightarrow \lim_{x\to p}(f+g)(x) = \lim_{x\to p} f(x) + \lim_{x\to p} g(x) = f(p) + g(p) = (f+g)(p) \\ &\Rightarrow f + g \text{ is continuous at } p.\end{aligned}$$

□

8.3 Lecture 21. Fri Oct 18

8.3.1 Continuity and Compactness

Theorem 8.10. *Let $(X, d_x), (Y, d_y)$ be metric spaces, and let $f : X \to Y$. Then f is continuous iff $f^{-1}(V)$ is open in X for every open subset V of Y.*

Proof. Suppose first that f is continuous, and let $V \subset Y$ be open. Define $U = f^{-1}(V)$, and let $x \in U$. Since V is open and $f(x) \in V$, there exists $\varepsilon > 0$ such that $B_\varepsilon(f(x)) \subset V$. Since f is continuous at x, there exists $\delta > 0$ such that if $u \in B_\delta(x)$, then

$$d_x(u, x) < \delta \Rightarrow d_y(f(u), f(x)) < \varepsilon.$$

That is,

$$\begin{aligned}&f(B_\delta(x)) \subset B_\varepsilon(f(x)) \subset V \\ &\Rightarrow B_\delta(x) \subset f^{-1}(V) = U.\end{aligned}$$

Thus, U is open. Conversely, suppose $f^{-1}(V)$ is open in X for every open subset V of Y. Let $x \in X$ and let $\varepsilon > 0$ be given. Since $B_\varepsilon(f(x))$ is open in Y, we know that $f^{-1}(B_\varepsilon(f(x)))$ is open in X. Moreover, $x \in f^{-1}(B_\varepsilon(f(x)))$. Thus there exists $\delta > 0$ such that $B_\delta(x) \subset f^{-1}(B_\varepsilon(f(x)))$. This tells us that $f(B_\delta(x)) \subset B_\varepsilon(f(x))$. This shows f is continuous at x, but since x is chosen arbitrarily, f is continuous. □

Lemma 8.11. *Let X, Y be sets and suppose $f : X \to Y$. Then the following hold true.*

1. *For all $V \subset Y$, $f^{-1}(V^C) = f^{-1}(V)^C$ (i.e. $f^{-1}(Y \setminus V)) = X \setminus f^{-1}(V)$).*

2. *If f is bijective, then $\forall U \subset X$, $f(U^C) = f(U)^C$. That is $f(X \setminus U) = Y \setminus f(U)$.*

3. *If $\{S_a \mid a \in A\}$ is a collection of subsets of Y, then*

$$f^{-1}\left(\bigcup_{a\in A} S_a\right) = \bigcup_{a\in A} f^{-1}(S_a),$$

and

$$f^{-1}\left(\bigcap_{a\in A} S_a\right) = \bigcap_{a\in A} f^{-1}(S_a).$$

Proof. Homework. □

Corollary 8.11.1. *Let $(X, d_x), (Y, d_y)$ be metric spaces, and let $f : X \to Y$. Then f is continuous iff the inverse image of every closed subset in Y is closed in X.*

Theorem 8.12. *Let $(X, d_x), (Y, d_y)$ be metric spaces, and let $f : X \to Y$ be continuous. If $E \subset X$ is compact, the $f(E)$ is compact, and hence, closed and bounded.*

Proof. Let $E \subset X$ be compact, and let $\{G_a \mid a \in A\}$ be an open cover of $f(E)$. Then

$$\begin{aligned} f(E) &\subset \bigcup_{a\in A} G_a \\ \Rightarrow E &\subset f^{-1}\left(\bigcup_{a\in A} G_a\right) = \bigcup_{a\in A} f^{-1}(G_a) \end{aligned}$$

By the previous theorem (Theorem 8.10), $f^{-1}(G_a)$ is open, for all $a \in A$ (since G_a is open and f is continuous). But then $\{f^{-1}(G_a) \mid a \in A\}$ is an open cover of E, so there exists $a_1, \ldots, a_n \in A$ such that

$$\begin{aligned} E &\subset \bigcup_{j=1}^{n} f^{-1}(G_{a_j}) = f^{-1}\left(\bigcup_{j=1}^{n} g_{a_j}\right) \\ \Rightarrow f(E) &\subset \bigcup_{j=1}^{n} G_{a_j}. \end{aligned}$$

Thus, $f(E)$ is compact. □

Corollary 8.12.1 (Weierstrauss). *Let (X, d) be a metric space, and let $f : X \to \mathbb{R}$ be continuous, and let $E \subset X$ be compact. Then f attains its maximum and minimum over E, that is, there exists $x_1, x_2 \in E$ such that*

$$f(x_1) \le f(x), \forall x \in E \text{ and } f(x_2) \ge f(x), \forall x \in E.$$

Example 8.13 (Counterexamples, why compactness is necesssary). Consider $f : (0, 1) \to (0, 1)$, where $f(x) = x^2$. Note that $(0, 1)$ is not compact, since it isn't closed. We can easily see that $0 = \inf\{f(x) \mid 0 < x < 1\}$, and $1 = \sup\{f(x) \mid 0 < x < 1\}$. Neither values are ever attained by $f(x)$.

Proof (Corollary 8.12.1). Since E is compact and f is continuous, $f(E) \subset \mathbb{R}$ is compact, and hence, closed and bounded. Thus, $\sup\{f(x) \mid x \in E\}$ exists and $\sup\{f(x) \mid x \in E\} \in f(E)$. Thus, $\sup f(E) = f(x_2)$ for some $x_2 \in E$. Same is done for the infimum. □

Theorem 8.14. *Let $(X, d_x), (Y, d_y)$ be metric spaces, and assume X is compact. Let $f : X \to Y$ be continuous and invertible. Then f^{-1} is continuous.*

Proof. It suffices to prove that

$$U \subset X \text{ is open} \Rightarrow (f^{-1})^{-1}(U) \text{ is open in } Y \Rightarrow f(U) \text{ is open in } Y \text{ since } ((f^{1})^{-1}(U) = f(U)).$$

(This is because

$$\begin{aligned} y \in (f^{-1})^{-1}(U) &\Leftrightarrow f^{-1}(y) \in U && (f(f^{-1}(u)) = y) \\ &\Leftrightarrow y \in f(U).) \end{aligned}$$

So let $U \subset X$ be open. Then

$$\begin{aligned} U^C \text{ is closed} &\Rightarrow U^C \text{ is compact} \\ &\Rightarrow f(U^C) \text{ is closed} \\ &\Rightarrow f(U)^C \text{ is closed} \\ &\Rightarrow f(U) \text{ is open.} \end{aligned}$$

Note that $f(U^C) = f(U)^C$ is valid only because f is bijective (whereas $f^{-1}(V^C) = f^{-1}(V)^C$ is always valid). Also, $(f^{-1})^{-1}(U) = f(Y)$ is only true (in fact, it's only meaningful) because f is bijective. □

9 Week 9

9.1 Lecture 22. Tue Oct 22

9.1.1 Uniform Continuity

Definition 9.1. Let $(X, d_x), (Y, d_y)$ be metric spaces, and let $f : X \to Y$. We say f is uniformly continuous on X iff

$$\forall \varepsilon > 0, \exists \delta > 0 (u, x \in X \wedge d_x(x, u) < \delta \Rightarrow d_y(f(u), f(x)) < \varepsilon).$$

Note. The difference between uniform continuity and continuity is that if f is only continuous on X, then the given δ corresponding to any ε value depends both on which $x \in X$ chosen as well as which ε chosen. If f is uniformly continuous on X, then the given δ only relies on ε.

Example 9.2.
1. Let $f : [0, 1] \to \mathbb{R}, f(x) = x^2$. Then

$$\begin{aligned} f(u) - f(x) &= u^2 - x^2 \\ &= (u - x)(u + x) \\ \Rightarrow |f(u) - f(x)| &= |u + x||u - x| \\ &\leq 2|u - x|. \end{aligned}$$

So given $\varepsilon > 0$, choose $\delta = \frac{\varepsilon}{2}$. Then

$$u, x \in [0,1], |u - x| < \delta \Rightarrow |f(u) - f(x)| < 2\delta = \varepsilon.$$

Then f is uniformly continuous.
2. Let $f : (0,1) \to \mathbb{R}, f(x) = \frac{1}{x}$. Note that

$$\begin{aligned} |f(u) - f(x)| &= \left| \frac{1}{u} - \frac{1}{x} \right| \\ &= \frac{|u - x|}{ux}. \end{aligned}$$

So $|u - x| < \delta \Rightarrow |f(u) - f(x)| < \frac{\delta}{ux}$, but $\frac{1}{x} \to \infty$ as $x \to 0$ (or $\frac{\delta}{ux} < \varepsilon \Leftrightarrow \delta < \varepsilon ux$, but there is not fixed δ which satisfies this for all $x > 0$). Thus, f is continuous but not uniformly continuous.

Theorem 9.3. *Let $(X, d_x), (Y, d_y)$ be metric spaces. Let X be compact, and let $f : X \to Y$ be continuous. Then f is uniformly continuous.*

Proof. Let $\varepsilon > 0$ be given. Since f is continuous, $\forall x \in X, \exists \delta_x > 0$ such that if $u \in X$, and $d_x(u, x) < \delta_x$, then $d_y(f(u), f(x)) < \frac{\varepsilon}{2}$. Define for all $x \in X$, $U_x = B_{\frac{1}{2}\delta_x}(x)$. Then $\{U_x \mid x \in X\}$ is an open cover of X, hence there exists $x_1, \ldots, x_n \in X$ such that

$$X \subset \bigcup_{j=1}^{n} U_{x_j}. \tag{*}$$

Define $\delta = \frac{1}{2} \min\{\delta_{x_1}, \delta_{x_2}, \ldots, \delta_{x_n}\}$. Then let $u, x \in X$, $d_x(u, x) < \delta$. By (*), there exists $j \in \{1, 2, \ldots, n\}$ such that

$$x \in U_{x_j} \Rightarrow d_x(x, x_j) < \frac{1}{2}\delta_{x_j}.$$

Since $d_x(u, x) < \delta \le \frac{1}{2}\delta_{x_j}$,

$$\begin{aligned} d_x(u, x_j) &\le d_x(u, x) + d_x(x, x_j) \\ &< \delta + \frac{1}{2}\delta_{x_j} \\ &< \frac{1}{2}\delta_{x_j} + \frac{1}{2}\delta_{x_j} = \delta_{x_j}. \end{aligned}$$

This implies that $u, x \in B_{\delta_{x_j}}(x_j)$. Therefore,

$$\begin{aligned} d_y(f(u), f(x)) &\le d_y(f(u), f(x_j)) + d_y((x_j), f(x)) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Thus, f is uniformly continuous. □

9.1.2 Intermediate Value Theorem

Theorem 9.4. *Let $(X, d_x), (Y, d_y)$ be metric spaces, and let $f : X \to Y$ be continuous. If $E \subset X$ is connected, then $f(E)$ is connected.*

Proof. Suppose $E \subset X$, and $f(E)$ is not connected, i.e. $\exists C, D \subset Y, C, D \neq \varnothing, C \cup D = f(E), \overline{C} \cap D = C \cap \overline{D} = \varnothing$. Consider that

$$\begin{aligned}
&E \subset f^{-1}(f(E)) = f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D) \\
&\Rightarrow E = (E \cap f^{-1}(C)) \cup (E \cap f^{-1}(D)) \\
&\Rightarrow E = A \cup B, A = E \cap f^{-1}(C), B = E \cap f^{-1}(D).
\end{aligned}$$

Now, $C \neq \varnothing$, and $C \subset f(E)$, so there exists $x \in E$ such that $f(x) \in C$, i.e., $x \in E \cap f^{-1}(C) = A$. Thus, A is nonempty. Similarly, $D \neq \varnothing \Rightarrow B \neq \varnothing$. Also,

$$\begin{aligned}
A \subset f^{-1}(C) &\Rightarrow A \subset f^{-1}(\overline{C}) \\
&\Rightarrow \overline{A} \subset f^{-1}(\overline{C}) && (\overline{C} \text{ closed} \Rightarrow f^{-1}(C) \text{ closed b.c } f \text{ is cont.}) \\
&\Rightarrow f(\overline{A}) \subset \overline{C}.
\end{aligned}$$

Similarly, $B \subset f^{-1}(D) \Rightarrow f(B) \subset D$. But then,

$$\begin{aligned}
x \in \overline{A} \cap B \Rightarrow f(x) \in f(\overline{A} \cap B) &\subset f(\overline{A}) \cap f(B) \\
&\subset \overline{C} \cap D = \varnothing.
\end{aligned}$$

This shows $\overline{A} \cap B$ is empty, and similarly, $A \cap \overline{B}$ is empty. Thus, E is not connected. □

Corollary 9.4.1 (Intermediate value theorem). *Suppose $f : I \to \mathbb{R}$, where $I \subset \mathbb{R}, [a, b] \subset I$, and $f(a) \neq f(b)$. If f is continuous and v lies between $f(a), f(b)$, (either $f(a) < v < f(b)$ or $f(b) < v < f(a)$) then $\exists c \in (a, b)$ such that $f(c) = v$.*

Proof. Since $[a, b]$ is connected and f is continuous, then $f([a, b])$ is connected, and by a previous theorem, $f([a, b])$ is an interval. □

Definition 9.5. Let $f : I \to \mathbb{R}$ where $I \subset \mathbb{R}$. If $(c, a) \subset I$, for $c < a, c, a \in \mathbb{R}$, then we say

$$\lim_{x \to a^-} f(a)$$

exists iff there exists $L \in \mathbb{R}$ such that

$$\forall \varepsilon > 0, \exists \delta > 0 (x \in (a - \delta, a) \Rightarrow |f(x) - L| < \varepsilon).$$

In this case, we write

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

Similarly, $\lim_{x \to a^+} f(x) = L$ iff

$$\forall \varepsilon > 0 \exists \delta > 0 (x \in (a, a + \delta) \Rightarrow |f(x) - L| < \varepsilon).$$

This requires $(a, b) \subset I$ for some $b > a$.

Definition 9.6.
1. We say $f : X \to Y$ has a removable discontinuity at $x = a$ iff $\lim_{x \to a} f(x)$ exists, but does not equal $f(a)$ (or $f(a)$ is not defined).
2. We say that $f : I \to \mathbb{R}, I \subset \mathbb{R}$ has a jump discontinuity at $x = a$, for a in the interior of I iff

$$\lim_{x \to a^-} f(x) \neq \lim_{x \to a^+} f(x),$$

assuming both limits exist.
3. We say that $f : I \to \mathbb{R}$ has an infinite discontinuity at $x = a$ if

$$\lim_{x \to a^-} f(x) = \pm\infty \text{ or } \lim_{x \to a^+} f(x) = \pm\infty.$$

Note. The discontinuities above are not the only discontinuities which exist. Consider

$$f(x) = \begin{cases} \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

This function "oscillates" infinitely as $x \to 0$. This kind of discontinuity doesn't have a name.

Example 9.7.
1. Consider $f(x) = \frac{\sin(x)}{x}$, whose domain if $(-\infty, 0) \cup (0, \infty)$. We claim that

$$\lim_{x \to 0} \frac{\sin(x)}{x} = 1.$$

We can define

$$f(x) = \begin{cases} \frac{\sin(x)}{x} & x \neq 0 \\ 1 & x = 0. \end{cases}$$

The resulting function f is continuous.

9.2 Lecture 23. Wed Oct 23

9.2.1 Monotone Convergence Theorem

Definition 9.8. Let $f : I \to \mathbb{R}$ where $I \subset \mathbb{R}$, and let $(a, b) \subset I$. We say that f is increasing on (a, b) iff

$$x_1, x_2 \in (a, b) \wedge x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2),$$

and decreasing on (a, b) iff

$$x_1, x_2 \in (a, b) \wedge x_1 < x_2 \Rightarrow f(x_1) \geq f(x_2).$$

Strictly increasing and strictly decreasing have analog meanings, and we say that f is moniotonic on (a, b) iff it is increasing or decreasing on (a, b).

Theorem 9.9. *Let $f : I \to \mathbb{R}$, where $I \subset \mathbb{R}$ and suppose $(a, b) \subset I$. If f is increasing on (a, b), then for all $c \in (a, b)$,*

$$\lim_{x \to c^-} f(x) \text{ and } \lim_{x \to c^+} f(x)$$

exist and

$$\lim_{x \to c^-} f(x) = \sup\{f(x) \mid a < x < c\} \leq \inf\{f(x) \mid c < x < b\} = \lim_{x \to c^+} f(x).$$

Analogous results hold for decreasing functions.

Proof. Note that $S = \{f(x) \mid a < x < c\}$ is nonempty and bounded above by $f(t)$ for any $t \in (c, b)$. Thus, $L = \sup S$ exists in $\mathbb{R}$. Let $\varepsilon > 0$ be given. Then there exists $s \in (a, c)$ such that

$$L - \varepsilon < f(s) \leq L,$$

(otherwise $L - \varepsilon < L$ is an upper bound for S). Define $\delta = c - s$. Then

$$\begin{aligned} x \in (a, c) \text{ and } |x - c| < \delta &\Rightarrow s < x < c \\ &\Rightarrow f(s) \leq f(x) \leq L && \text{(since } f \text{ is increasing)} \\ &\Rightarrow |f(x) - L| \leq |f(s) - L| < \varepsilon. \end{aligned}$$

Thus, $L = \lim_{x\to c^-} f(x)$. The proof that

$$\lim_{x\to c^+} f(x) = \inf\{f(x) \mid c < x < b\}$$

is similar, and

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

follows immediately from

$$f(x) \le f(t), \forall x \in (a, c), \forall t \in (c, b).$$

□

Corollary 9.9.1. *Let $f : I \to \mathbb{R}$ be monotonic on $(a, b) \subset I$. Then the only discontinuities of f in (a, b) are jump discontinuities.*

Theorem 9.10. *Let $f : I \to \mathbb{R}$ be montonic on $(a, b) \subset I$. Then the set of discontinuities of f in (a, b) is countable.*

Proof. WLOG assume that f is increasing. Let E be the set of discontinuities of f in (a, b). For each $x \in E$, choose $r_x \in \mathbb{Q}$ satisfying

$$\lim_{t\to x^-} f(t) < r_x < \lim_{t\to x^+} f(t).$$

Then define $\varphi : E \to \mathbb{Q}$ by $\varphi(x) = r_x, \forall x \in E$. Clearly φ is injective, hence E is equivalent to a subset of $\mathbb{Q}$. Thus, E is countable. □

Note. We previously define

$$\lim_{x\to\infty} f(x), \lim_{x\to-\infty} f(x).$$

The usual properties of limits apply:

- these limits, if they exist, are unique.
- If $\lim_{x\to\infty} f(x), \lim_{x\to\infty} g(x)$ exist (in $\mathbb{R}$), then

$$\lim_{x\to\infty}(f(x) \pm g(x)) = \lim_{x\to\infty} f(x) \pm \lim_{x\to\infty} g(x),$$
$$\lim_{x\to\infty} f(x)g(x) = \Big(\lim_{x\to\infty} f(x)\Big)\Big(\lim_{x\to\infty} g(x)\Big),$$
$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \frac{\lim_{x\to\infty} f(x)}{\lim_{x\to\infty} g(x)}. \qquad \text{(if } \lim_{x\to\infty} g(x) \neq 0\text{)}$$

These rules remain true if

$$\lim_{x\to\infty} f(x) = \pm\infty \text{ or } \lim_{x\to\infty} g(x) = \pm\infty,$$

provided we avoid the indeterminate forms

$$\infty - \infty, 0 \cdot \infty, \frac{\infty}{\infty}.$$

The proofs for all of these properties are straightforward, but tedious.

9.2.2 Differentiation

Definition 9.11. Let $f : I \to \mathbb{R}$ where $I \subset \mathbb{R}$. Suppose t lies in the interior of I. We say that f is differentiable at t iff

$$\lim_{x \to t} \frac{f(x) - f(t)}{x - t}$$

exists, in which case this limit is called the derivative of f at t, and denoted $f'(t)$. That is

$$f'(t) = \lim_{x \to t} \frac{f(x) - f(t)}{x - t},$$

if this limit exists. If f is differentiable at every $t \in I$, we say that f is differentiable on I. If $f : [a, b] \to \mathbb{R}$, we can define the (one-sided) derivatives at the endpoints of

$$f'(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a},$$
$$f'(b) = \lim_{x \to b^-} \frac{f(x) - f(b)}{x - b}.$$

We don't have a special notation for the one-sided derivative. We don't use them much.

Theorem 9.12 (Differentiation is linear)**.** *Suppose $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ are both differentiable at t, and let $\alpha, \beta \in \mathbb{R}$. Then*

$$\begin{aligned}
(\alpha f + \beta g)'(t) &= \lim_{x \to t} \frac{(\alpha f + \beta g)(x) - (\alpha f + \beta g)(t)}{x - t} \\
&= \lim_{x \to t} \left(\alpha \frac{f(x) - f(t)}{x - t} + \beta \frac{g(x) - g(t)}{x - t} \right) \\
&= \alpha \lim_{x \to t} \frac{f(x) - f(t)}{x - t} + \beta \lim_{x \to t} \frac{f(x) - f(t)}{x - t} \\
&= \alpha f'(t) + \beta f'(t).
\end{aligned}$$

Proposition 9.13. *If we define vector spaces*

$$C[a, b] = \{f : [a, b] \to \mathbb{R} \mid f \text{ is continuous}\},$$

and

$$C'[a, b] = \{f \in C[a, b] \mid f \text{ is differentiable on } [a, b] \text{ and } f' \in C[a, b]\}.$$

Then

$$D : C'[a, b] \to C[a, b], \quad Df = f'$$

is a linear map.

9.3 Lecture 24. Thu Oct 24

9.3.1 More on Differentiation

Theorem 9.14. *Let $f : I \to \mathbb{R}$, where $I \subset \mathbb{R}$, and assume $t \in I^0$. If f is differentiable at t, then f is continuous at t.*

Proof. We must prove that

$$f(x) \to f(t), \text{ as } x \to t$$
$$\Leftrightarrow f(x) - f(t) \to 0, \text{ as } x \to t.$$

But

$$f(x) - f(t) = \frac{f(x) - f(t)}{x - t}(x - t) \to f'(t) \cdot 0 = 0, \text{ as } x \to t.$$

(Since $\frac{f(x)-f(t)}{x-t} \to f'(t)$ and $x - t \to 0$.) □

Proposition 9.15. *Assume that f is differentiable at t. Then*

$$\begin{aligned} &\frac{f(x) - f(t)}{x - t} \to f'(t) \text{ as } x \to t \\ \Longleftrightarrow\ &\frac{f(x) - f(t)}{x - t} - f'(t) \to 0 \text{ as } x \to t \\ \Longleftrightarrow\ &\frac{f(x) - f(t) - f'(t)(x - t)}{x - t} \to 0 \text{ as } x \to t \\ \Longleftrightarrow\ &\frac{f(x) - (f(t) + f'(t)(x - t))}{x - t} \to 0 \text{ as } x \to t. \end{aligned}$$

What is $y = f(t) + f'(t)(x - t)$? This is the tangent line approximation of $y = f(t)$ near $x = t$ or the local linearization of $y = f(x)$ at $x = t$. The condition

$$\frac{f(x) - (f(t) + f'(t)(x - t))}{x - t} \to 0 \text{ as } x - t \to 0$$

says that the error $f(x) - (f(t) + f'(t)(x - t)$ is small compared to $|x - t|$ as $|x - t| \to 0$. We write

$$f(x) - (f(t) + f'(t)(x - t)) = o(|x - t|), \tag{9.15.1}$$

where $o(h)$ denotes a quantity that satisfies

$$\frac{o(h)}{h} \to 0 \text{ as } h \to 0.$$

There's a converse to Eq. (9.15.1). If $m \in \mathbb{R}$ satisfies

$$f(t + h) = f(t) + mh + o(h) \text{ as } h \to 0,$$

then f is differentiable at t, and $m = f'(t)$.

Example 9.16.
1. Let $h = fg$, where f, g are differentiable at t. Then

$$\begin{aligned} h(t + h) &= f(t + h)g(t + h) \\ &= (f(t) + f'(t)h + o(h))(g(t) + g'(t)h + o(h)) \\ &= f(t)g(t) + f(t)g'(t)h + f(t)o(h) \\ &+ f'(t)g(t) + f'(t)g'(t)h^2 + f'(t)ho(h) \\ &+ g(t)o(h) + g'(t)ho(h) + o(h)^2 \\ &= f(t)g(t) + (f'(t)g(t)) \\ &= h(t) + (f'(t)g(t) + f(t)g'(t))h + o(h). \end{aligned}$$

Therefore, $h'(t) = f'(t)g(t) + f(t)g'(t)$.
2. Suppose $h(x) = f(g(x))$. Let g be differentiable at t, and f be differentiable at $f(t)$. Then

$$\begin{aligned} h(t+h) &= f(g(t+h)) = f(g(t) + g'(t)h + o(h)) \\ &= f(g(t)) + f'(g(t))(g'(t)h + o(h) + o(g'(t)h + o(h)) \\ &= f(g(t)) + f'(g(t))g'(t)h + o(h). \end{aligned}$$

Thus $h'(t)$ must equal $f'(g(t))g'(t)$.
3. Let $\varphi(x) = x^{-1}$. Then

$$\begin{aligned} \varphi'(t) &= \lim_{x\to t} \frac{\varphi(x) - \varphi(t)}{x - t} \\ &= \lim_{h\to 0} \frac{\varphi(t+h) - \varphi(t)}{h} \\ &= \lim_{h\to 0} \frac{\frac{1}{t+h} + \frac{1}{t}}{h} \\ &= \lim_{h\to 0} \frac{t - t - h}{ht(t+h)} \\ &= \lim_{h\to 0} -\frac{h}{ht(t+h)} \\ &= \lim_{h\to 0} -\frac{1}{t(t+h)} \to -\frac{1}{t^2}. \end{aligned}$$

4. Suppose $h(x) = \frac{f(x)}{g(x)}$. Then

$$\begin{aligned} h(x) &= f(x)\varphi(g(x)) \\ \Rightarrow h'(x) &= f'(x)\varphi(g(x)) + f(x)\varphi'(g(x))g'(x) \\ &= \frac{f'(x)}{g(x)} - \frac{f(x)g'(x)}{g(x)^2} \\ &= \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \end{aligned}$$

9.3.2 Local Extrema, Stationary Points, Fermat's Theorem

Definition 9.17. Let $f : I \to \mathbb{R}$, where $I \subset \mathbb{R}$, and suppose a is an interior point of I. We say that a is a local minimizer (and $f(a)$ is a local minimum) of f iff there exists $\varepsilon > 0$ such that $(a - \varepsilon, a + \varepsilon) \subset I$, and

$$f(a) \le f(x), \forall x \in (a - \varepsilon, a + \varepsilon).$$

We say a ia a local maximizer of f and $f(a)$ is a local maximum of f iff there exists $\varepsilon > 0$ such that $(a - \varepsilon, a + \varepsilon) \subset I$, and

$$f(a) \ge f(x), \forall x \in (a - \varepsilon, a + \varepsilon).$$

Strict local minimizers, strict local maximizers are defined in the obvious way.

Theorem 9.18 (Fermat)**.** *Let $f : I \to \mathbb{R}$, where $I \subset \mathbb{R}$, has a local maximizer or local minimizer at $a \in I^0$. If f is differentiable at a, then $f'(a) = 0$.*

Definition 9.19. Let $f : I \to \mathbb{R}$ where $I \subset \mathbb{R}$, and let a be an interior point of I. If f is differentiable at a and $f'(a) = 0$, we call a a stationary point of f. (American calculus courses call a a critical point.)

Proof of Theorem 9.18. WLOG suppose a is a local minimizer of f, and assume f is differentiable at a. Then

$$\begin{aligned} f'(a) &= \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} \\ &\geq 0, && \text{(since } f(a+h) - f(a) \geq 0\text{, and } h > 0\text{)} \end{aligned}$$

and

$$\begin{aligned} f'(a) &= \lim_{x \to a^-} \frac{f(x) - f(a)}{x - a} \\ &\leq 0. && \text{(since } f(x) - f(a) \geq 0 \text{ and } h < 0\text{)} \end{aligned}$$

Conclude that $f'(a) = 0$. □

Note. Why is Fermat's theorem a big deal? Suppose we claim that $x = a$ is a local minimizer of f. How can one verify that claim? By definition, we have to check $f(x) \geq f(a)$, for all infinitely many values near a. Fermat's theorem let's us simply check that $f'(a) = 0$. Take note that Fermat's theorem is only necesssary, and not sufficient.

10 Week 10

10.1 Lecture 25. Mon Oct 28

10.1.1 Mean Value Theorem

Definition 10.1 (Local max over general metric space)**.** We say $f : X \to \mathbb{R}$ has a local maximizer at $p \in X$ iff there is some $\delta > 0$ such that $f(q) \leq f(p)$, for all $q \in X$, with $d(p, q) < \delta$.

One more proof of Theorem 9.18. Let $f : [a, b] \to \mathbb{R}$, WLOG suppose x is a local maximizer of f on $[a, b]$, that is let $\delta > 0$ be given, such that $[x - \delta, x + \delta] \subset (a, b)$, and

$$f(q) \leq f(x), \forall q \in (x - \delta, x + \delta).$$

If $t \in (x - \delta, x)$, then

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

If $t \in (x, x + \delta)$, then

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

Then we conclude $f'(x) = 0$. □

Theorem 10.2 (Generalized MVT)**.** *Let $f, g : [a, b] \to \mathbb{R}$ both be continuous and differentiable in (a, b). Then there exists $x \in (a, b)$ such that*

$$(f(b) - f(a))g'(x) = (g(b) - g(a))f'(x).$$

Theorem 10.3 (Standard MVT). *Let $f : [a, b] \to \mathbb{R}$ be continuous and differentiable in (a, b). Then there exists $x \in (a, b)$ such that*

$$f(b) - f(a) = (b - a)f'(x).$$

Note. By choosing $g(x) = x$, we can derive the standard MVT from the generalized MVT.

Proof of Theorem 10.2. Define $h : [a, b] \to \mathbb{R}$ to be continuous and differentiable in (a, b), such that $h(t) := (f(b) - f(a))t - (b - a)f(t)$. Then

$$\begin{aligned} h(a) &= f(b)a - f(a)a - bf(a) + af(a) \\ &= f(b)a - bf(a). \end{aligned}$$

Also,

$$\begin{aligned} h(b) &= f(b)b - f(a)b - bf(b) + af(b) \\ &= af(b) - f(a)b. \end{aligned}$$

Hence, $h(b) = h(a)$. We want to show that $h(x) = 0$, for some $x \in (a, b)$. If h is a constant function, then we are done. Otherwise, let $x \in (a, b)$, so that $h(x)$ is either maximal or minimal (which exists on $[a, b]$ since $[a, b]$ is compact, and f is continuous, but h is not constant, and thus $x \neq a$ and $x \neq b$). Hence, $h'(x) = 0$. □

Theorem 10.4. *Assume $f : [a, b] \to \mathbb{R}$ is differentiable. Then*

1. *$f'(x) \geq 0, \forall x \in (a, b) \implies f$ is monotonically increasing,*
2. *$f'(x) \leq 0, \forall x \in (a, b) \implies f$ is monotonically decreasing,*
3. *$f'(x) = 0, \forall x \in (a, b) \implies f$ is constant.*

Proof. If the assumption of 1 is taken, consider that $\forall x_1, x_2 \in (a, b), \exists x \in (x_1, x_2)^*$, such that

$$f'(x) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq 0.$$

□

10.1.2 L'Hôpital's Rule

Theorem 10.5 (L'Hôpital's Rule). *Suppose $f : (a, b) \to \mathbb{R}$, $g : (a, b) \to \mathbb{R}$ are differentiable, and satisfy*

$$\lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = \infty.$$

If $g'(x) \neq 0$, for all $x \in (a, b)$, ang

$$L = \lim_{x \to a^+} \frac{f'(x)}{g'(x)}$$

exists in $\mathbb{R}$, then

$$\lim_{x \to a^+} \frac{f(x)}{g(x)}$$

exists and

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = \lim_{x \to a^+} \frac{f'(x)}{g'(x)} = L.$$

Moreover the above holds if $x \to a^+$ is everywhere replaced by $x \to b^-$.

Proof. Let $y \in (a,b)$ be fixed. Then for all $x \in (a,y)$ there exists $c_x \in (x,y)$ such that

$$(f(x) - f(y))g'(c_x) = (g(x) - g(y))f'(c_x)$$
$$\implies \frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)} - \frac{f'(c_x)}{g'(c_x)} \cdot \frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}.$$

Write

$$L = \lim_{x \to a^+} \frac{f'(x)}{g'(x)}.$$

Let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that

$$c \in (a, a = \delta) \implies \left| \frac{f'(c)}{g'(c)} - L \right| < \frac{\varepsilon}{3}.$$

Let y be such that $y \in (a, a+\delta)$. Then since $c_x \in (x,y) \subset (a,y) \subset (a, a+\delta)$. Then

$$x \in (a,y) \implies c_x \in (a, a+\delta) \implies \left| \frac{f'(c_x)}{g'(c_x)} - L \right| < \frac{\varepsilon}{3}.$$

Since

$$\frac{f'(c_x)}{g'(c_x)} \cdot \frac{g(y)}{g(x)}, \frac{f(y)}{g(x)} \to 0, \text{ as } x \to a^+,$$

there exists $\gamma > 0$ making

$$x \in (a, a+\gamma) \implies \left| \frac{f'(c_x)}{g'(c_x)} \frac{g(y)}{g(x)} \right| < \frac{\varepsilon}{3}, \quad \left| \frac{f(y)}{g(x)} \right| < \frac{\varepsilon}{3}.$$

Set $\delta' = \min \delta, \gamma$. Then

$$x \in (a, a+\delta') \implies \frac{f(x)}{g(x)} = L = \frac{f'(c_x)}{g'(c_x)} - L - \frac{f'(c_x)}{g'(c_x)} \frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}$$
$$\implies \left| \frac{f(x)}{g(x)} - L \right| \le \left| \frac{f'(c_x)}{g'(c_x)} - L \right| + \left| \frac{f'(c_x)}{g'(c_x)} \frac{g(y)}{g(x)} \right| + \left| \frac{f(y)}{g(x)} \right|$$
$$\implies \left| \frac{f(x)}{g(x)} - L \right| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Thus,

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = L.$$

□

10.2 Lecture 26. Wed Oct 30

10.2.1 Taylor's Theorem

Note. Recall that f'' means $(f')'$, thus,

$$f''(t) = \lim_{x \to t} \frac{f'(x) - f(t)}{x - t},$$

and the domain of f'' is the set of all t for which f' exists on an open interval containing t (So that $\frac{f'(x)-f'(t)}{x-t}$ makes sense) and

$$\lim_{x\to t}\frac{f'(x)-f'(t)}{x-t} \text{ exists.}$$

Similarly, $f''' = (f'')', f^{(4)} = (f''')'$, etc. If we say that $f^{(n)}$ exists on (a,b), this implies that f is differentiable on (a,b), (so that f' exists), f' is differentiable on (a,b) (so that f'' exists), ..., $f^{(n-1)}$ is differentiable on (a,b).

Definition 10.6. Suppose f is $(n-1)$-times differentiable on an open interval containin α. The Taylor polynomial p_{n-1} of f at α (i.e. in powers of $(x-\alpha)$) is

$$p_{n-1}(x) = \sum_{k=0}^{n-1}\frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k.$$

Recall the defining property of p_{n-1}: it is the unique polynomial of degree at most $n-1$ such that

$$\begin{aligned} p_{n-1}(\alpha) &= f(\alpha) \\ p_{n-1}(\alpha) &= f'(\alpha) \\ &\vdots \\ p_{n-1}^{(n-1)}(\alpha) &= f^{(n-1)}(\alpha). \end{aligned}$$

Typically, p_{n-1} provides a "good" approximation of f on an interval that gets larger as $n\to\infty$. *This is not always true.*

Theorem 10.7. *Let f by n-times differentiable on (a,b), and let $\alpha\in(a,b)$. Then for all $t\in(a,b)$, there exists c_t lying between α and t such that*

$$f(t) = \sum_{k=0}^{n-1}\frac{f^{(k)}(\alpha)}{k!}(t-\alpha)^k + \frac{f^{(n)}(c_t)}{n!}(t-\alpha)^n.$$

Proof. Set $p_{n-1}(t) = \sum_{k=0}^{n-1}\frac{f^{(k)}(\alpha)}{k!}(t-\alpha)^k$. Define $M\in\mathbb{R}$ by

$$f(t) = p_{n-1}(t) + M(t-\alpha), M = \frac{f(t)-p_{n-1}(t)}{(t-\alpha)^n}$$

We want to show that $M = \frac{f^{(n)}(c_t)}{n!}$ for some c_t between α, t. Define $g:(a,b)\to\mathbb{R}$ by

$$g(x) = f(x) - p_{n-1}(x) - M(x-\alpha)^n.$$

Note that

$$\begin{aligned} g(\alpha) &= f(\alpha) - p_{n-1}(\alpha) - M\cdot 0 = 0 \\ g'(\alpha) &= f'(\alpha) - p'_{n-1}(\alpha) - nM\cdot 0 = 0 \\ &\vdots \\ g^{(n-1)}(\alpha) &= f^{n-1}(\alpha) - p_{n-1}^{(n-1)}(\alpha) - n!M\cdot 0 = 0 \end{aligned}$$

and

$$\begin{aligned} g^{(n)}(x) &= f^{(n)}(x) - p_{n-1}^{(n)}(x) - \overbrace{p_{n-1}^{(n)}(x)}^{\deg(p_{n-1})\leq n-1} - n!M \\ &= f^{(n)}(x) - n!M. \end{aligned}$$

Note that $g(\alpha) = 0$ and $g(t) = 0$, by definition of M. This implies that there exists c_1 between α, t such that $g'(c_1) = 0$, by the mean value theorem. Then $g'(\alpha) = 0$, and $g'(c_1) = 0$, implies that there exists c_2 between α and c_2 such that $g'(c_2) = 0$ and $g''(\alpha) = 0$ and $g''(c_2) = 0$. Continuing in this fashion, $g^{(n-1)}(\alpha) = 0$ and $g^{(n-1)}(c_{n-1}) = 0$, which implies that there exists c_n between α and c_{n-1} such that $g^{(n)}(c_n) = 0$. Define $c_t = c_n$, and write

$$g^{(n)}(c_t) = f^{(n)}(c_t) - n!M = 0 \Longrightarrow M = \frac{f^{(n)}(c_t)}{n!},$$

i.e.

$$f(t) = p_{n-1}(t) + M(t-\alpha)^n = p_{n-1}(t) + \frac{f^{(n)}(c_t)}{n!}(t-\alpha)^n.$$

□

Example 10.8.
1. Let $f(x) = e^x$, with $\alpha = 0$. Let R be any positive real number. Note that

$$f^{(n)}(x) = e^x, \forall x \in \mathbb{R}, \forall n \in \mathbb{Z}^+.$$

Thus, if $t \in (-R, R)$, then $c_t \in (-R, R)$, and hence

$$|f^{(n)}(c_t)| = e^{c_t} \leq e^R.$$

Hence,

$$|f(t) - p_{n-1}(t)| = \frac{|f^{(n)}(c_t)|}{n!}|t|^n \leq \frac{e^R|t|^n}{n!}.$$

Since

$$\frac{|t|^n}{n!} \to 0 \text{ as } n \to \infty,$$

we see that

$$|f(t) - p_{n-1}(t)| \to 0 \text{ as } n \to \infty, \forall t \in (-R, R),$$

that is,

$$p_{n-1}(t) \to f(t) \text{ as } n \to \infty.$$

Thus,

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}t^k = e^t, \forall t \in (-R, R).$$

Since R was arbitrary, we see that

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}t^k = e^t, \forall t \in \mathbb{R}.$$

Moreover, $f^{(k)}(0) = e^0 = 1$, for all k, so we obtain

$$e^t = \sum_{k=0}^{\infty} \frac{t^k}{k!}, \forall t \in \mathbb{R}.$$

2. Consider $f(x) = \sin(x)$, $\alpha = 0$. Then

$$\begin{aligned} f'(x) &= \cos(x), \quad f'(0) = 1, \\ f''(x) &= -\sin(x), \quad f''(0) = 0 \\ f'''(x) &= -\cos(x), \quad f'''(0) = -1. \end{aligned}$$

(The pattern 0, 1, 0, -1, 0, 1, 0, -1, ... continues.) We obtain the following Taylor series:

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}.$$

Taylor's theorem is

$$\sin(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + \frac{f^{(2n+2)(c_x)}}{(2n+2)!} x^{2n+2}.$$

We have

$$|f^{(2n+2)}(c_x)| \leq 1,$$

since the even derivatives are all $\pm$sine, and hence

$$\left| \sin(x) - \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} \right| \leq \frac{|x|^{2n+2}}{(2n+2)!} \to 0 \text{ as } n \to \infty, \forall x \in \mathbb{R}.$$

Thus

$$\sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}, \forall x \in \mathbb{R}.$$

Similarly,

$$\cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}, \forall x \in \mathbb{R}.$$

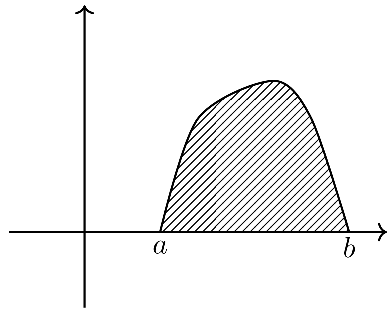
10.3 Lecture 27

10.3.1 Riemann Integrals

To this point, Mark hasn't defined what it means for a function $f : (X, d_X) \to \mathbb{R}$ to be bounded. I wrote my own definition here.

Definition 10.9. Let $f : (X, d_X) \to \mathbb{R}$. Let $E \subset X$. We say f is bounded over the set E iff there exists $M > 0$, such that $|f(x)| \leq M$, for all $x \in E$.

Consider $f : [a, b] \to \mathbb{R}$, and assume for convenience that $f(x) \geq 0, \forall x \in [a, b]$. *Throughout this section, assume f is bounded.*



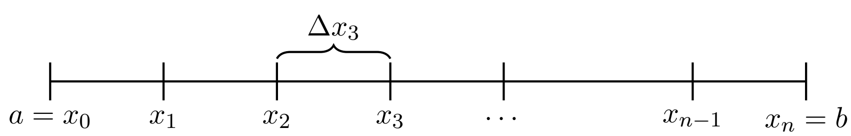
We often wish to compute the area between $y = 0$, and $y = f(x)$ on $[a, b]$. Computing areas is important in geometry, but the real reason is that if $y = f(t)$ represents the instantaneous rate of change of some quantity, w.r.t. time, at time $t \in [a, b]$, then this area is numerically the change in that quantity from $t = a$ to $t = b$. The following analysis should be familiar from Calculus.

Definition 10.10. A partition of $[a, b]$ is a set $P = \{x_0, x_1, \ldots, x_n\} \subset [a, b]$, where

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

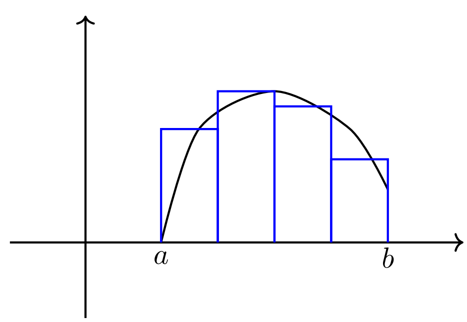
We write $\Delta x_j = x_j - x_{j-1}$, for all $j = 1, 2, \ldots, n$. The mesh size of P is

$$|P| = \max\{\Delta x_j \mid j \in [n]\}.$$



Note. "Partition" is a poor choice of word, since partition has another well-established meaning in mathematics.

Definition 10.11 (Riemann Sums). Given a partition $\{x_1, x_2, \ldots, x_n\}$ of $[a, b]$, we approximate the desired area as the sum of the areas of rectangles, one on each subinterval $[x_{j-1}, x_j]$.



The height of the rectangle on $[x_{i-1}, x_i]$ is $f(x_i^*)$, where x_i^* is any point in $[x_{i-1}, x_i]$. Then

$$A \approx \sum_{j=1}^{n} f(x_j^*)\Delta x_j, \quad (x_j^* \in [x_{j-1}, x_j], \forall j \in [n]).$$

The expression $\sum_{j=1}^{n} f(x_j^*)\Delta x_j$ is called a Riemann sum.

Definition 10.12. In a calculus class, one then takes the limit as $|P| \to 0$, if this limit exists. We call this the Riemann integral of f on $[a, b]$:

$$\int_a^b f(x)\,dx = \lim_{|P|\to 0} \sum_{j=1}^{n} f(x_j^*)\Delta x_j.$$

Indeed, this is not a very convenient definition. This is in fact a new kind of limit which requires a new definition which we are not familiar with in this course. We avoid this by taking a different approach.

Definition 10.13 (Upper and lower Darboux sums)**.** Given $f : [a, b] \to \mathbb{R}$ and a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$, we define the upper Darboux sum of f on P by

$$U(P, f) = \sum_{j=1}^{n} M_j \Delta x_j,$$

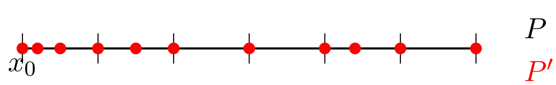
where $M_j = \sup\{f(x) \mid x_{j-1} \leq x \leq x_j\}$, for $j = 1, \ldots, n$, and the lower Darboux sum of f on P by

$$L(P, f) = \sum_{j=1}^{n} m_j \Delta x_j,$$

$m_j = \inf\{f(x) \mid x_{j-1} \leq x \leq x_j\}$, for $j = 1, \ldots, n$. Hence, every Riemann sum for f relative to P satisfies

$$L(P, f) \leq \sum_{j=1}^{n} f(x_j^*)\Delta x_j \leq U(P, f).$$

Definition 10.14. Let P, P' be partitions of $[a, b]$. We say that P' is a refinement of P iff $P \subset P'$.



Lemma 10.15. *If P, P' are partitions of $[a, b]$ and P' is a refinement of P, then*

$$L(P, f) \leq L(P', f) \leq U(P', f) \leq U(P, f).$$

Proof. The inequality $L(P', f) \leq U(P', f)$ follows by definition. Let us prove that $L(P, f) \leq L(P', f)$. It suffices to prove this inequality in the case that $P' = P \cup \{x_{\ell'}\}$, where $x_\ell \in (x_{\ell-1}, x_\ell)$. (If P' contains m more points than P, we can define $P'_1, P'_2, \ldots, P'_m = P'$ where P'_1 has one more point than P and P'_{i+1} has one more point than P'_j, for $j = 1, \ldots, m-1$. Then we would have

$$L(P, f) \leq L(P', f) \leq \ldots \leq L(P'_m, f) = L(P'.f).)$$

Assume that $P' = P \cup \{x'_\ell\}$, $x_{\ell-1} < x'_\ell < x_\ell$. Then

$$\begin{aligned} L(P', f) - L(P, f) = \inf\{f(x) \mid x_{\ell-1} \leq x \leq x'_\ell\}(x'_\ell - x_{\ell-1}) + \inf\{f(x) \mid x'_\ell \leq x \leq x_\ell\}(x_\ell - x'_\ell) \\ - \inf\{f(x) \mid x_{\ell-1 \leq x \leq x_\ell}\}(x_\ell - x_{\ell-1}) \end{aligned}$$

Since

$$\inf\{f(x \mid x_{\ell-1} \le x \le x'_\ell)\} \ge \inf\{f(x) \mid x_{\ell-1} \le x \le x_\ell\}$$

and

$$\inf\{f(x) \mid x'_\ell \le x \le x_\ell\} \ge \inf\{f(x) \mid x_{\ell-1} \le x \le x_\ell\},$$

we have

$$L(P', f) - L(P, f) \ge \inf\{f(x) \mid x_{\ell-1} \le x \le x_{\ll}\}(x'_\ell - x_{\ell-1}) + \inf\{f(x) \mid x_{\ell-1} \le x \le x_\ell\}(x_\ell - x'_\ell)$$
$$- \inf\{f(x) \mid x_{\ell-1} \le x \le x_\ell\}(x_\ell - x_{\ell-1}) = 0.$$

(Since $x'_\ell - x_{\ell-1} + x_\ell - x'_\ell = x_\ell - x_{\ell-1}$.) This completes the proof. The proof that $U(P', f) \le U(P, f)$ is similar. □

Definition 10.16. Let $f : [a, b] \to \mathbb{R}$ and let $\mathcal{P}$ be the set of all partitions on $[a, b]$. The upper and lower Riemann integrals of f on $[a, b]$ are

$$\overline{\int_a^b} f(x)\, dx = \inf\{U(P, f) \mid P \in \mathcal{P}\},$$

and

$$\underline{\int_a^b} f(x)\, dx = \sup\{L(P, f) \mid P \in \mathcal{P}\}$$

respectively. If

$$\underline{\int_a^b} f(x)\, dx = \overline{\int_a^b} f(x)\, dx,$$

then we say f is Riemann integrable on $[a, b]$ and define the Riemann integral of f on $[a, b]$ to be this common value:

$$\int_a^b f(x)\, dx = \underline{\int_a^b} f(x)\, dx = \overline{\int_a^b} f(x)\, dx.$$

10.4 Lecture 28. Wed Nov 6

10.4.1 Conditions for Riemann Integrability

Definition 10.17. We will write $\mathcal{P}$ to denote the set of partitions of $[a, b]$.

Lemma 10.18. *If $f : [a, b] \to \mathbb{R}$ is bounded, then*

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

Proof. Let P_1, P_2 be any partition of $[a, b]$, and define $P' = P_1 \cup P_2$. Then P' is a refinement of both P_1 and P_2 and so

$$L(P_1, f) \le L(P', f) \le U(P', f) \le U(P_2, f).$$

Then $L(P_1, f) \le U(P_2, f), \forall P_1, P_2 \in \mathcal{P}$. This implies

$$\underline{\int_a^b} f(x)\, dx = \sup\{L(P_1, f) \mid P_1 \inf \mathcal{P}\} \le U(P_2, f), \forall P_2 \in \mathcal{P},$$

thus

$$\underline{\int_a^b} f(x)\,dx \leq \inf\{U(P_2, f) \mid P_2 \in \mathcal{P}\} = \overline{\int_a^b} f(x)\,dx.$$

□

Lemma 10.19. *Let $f : [a, b] \to \mathbb{R}$ be bounded. Then f is Riemann integrable on $[a, b]$ iff for all $\varepsilon > 0$, there exists $P \in \mathcal{P}$ such that $U(P, f) - L(P, f) < \varepsilon$.*

Proof. Suppose first that f is Riemann integrable on $[a, b]$. Let $\varepsilon > 0$ be given. Write

$$I = \int_a^b f(x)\,dx = \inf\{U(P, f) \mid P \in \mathcal{P}\} = \sup\{L(P, f) \mid P \in \mathcal{P}\}.$$

Therefore, there exists $P_1 \in \mathcal{P}$, such that

$$I \leq U(P_1, f) < I + \frac{\varepsilon}{2}.$$

Similarly, there exists $P_2 \in \mathcal{P}$ such that

$$I - \frac{\varepsilon}{2} < L(P_2, f) \leq I.$$

Define $P' = P_1 \cup P_2$. Then

$$I - \frac{\varepsilon}{2} < L(P_2, f) \leq L(P', f) \leq U(P', f) \leq U(P_1, f) < I + \frac{\varepsilon}{2}.$$

This implies

$$U(P', f) - L(P', f) < I + \frac{\varepsilon}{2} - I + \frac{\varepsilon}{2} = \varepsilon.$$

For the converse, suppose that for all $\varepsilon > 0$, there exists $P \in \mathcal{P}$ such that

$$U(P, f) - L(P, f) < \varepsilon.$$

It suffices to prove that

$$\varepsilon \geq \overline{\int_a^b} f(x)\,dx - \underline{\int_a^b} f(x)\,dx, \forall \varepsilon > 0.$$

Let $\varepsilon > 0$ be given, and choose P_1 such that

$$U(P_1, f) - L(P_1, f) < \varepsilon.$$

Then

$$\begin{aligned}
\inf\{U(P, f) \mid P \in \mathcal{P}\} - L(P_1, f) &\leq U(P_1, f) - L(P_1, f) < \varepsilon \\
&\implies \overline{\int_a^b} f(x)\,dx - L(P_1, f) < \varepsilon \\
\implies \overline{\int_a^b} f(x)\,dx - \sup\{L(P, f) \mid P \in \mathcal{P}\} &\leq \overline{\int_a^b} f(x)\,dx - L(P, f) < \varepsilon \\
&\implies \overline{\int_a^b} f(x)\,dx - \underline{\int_a^b} f(x)\,dx < \varepsilon.
\end{aligned}$$

□

Theorem 10.20. *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is Riemann integrable on $[a, b]$.*

Proof. Note that since f is continuous, for all $P \in \mathcal{P}$,

$$U(P, f) = \sum_{j=1}^{n} M_j \Delta x_j, \text{ where } M_j = \max\{f(x) \mid x_{j-1} \leq x \leq x_j\}$$
$$L(P, f) = \sum_{j=1}^{n} m_j \Delta x_j, \text{ where } m_j = \min\{f(x) \mid x_{j-1} \leq x \leq x_j\}.$$

Also, since f is continuous, and $[a, b]$ is compact, f is uniformly continous. Let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that

$$x, y \in [a, b] \text{ and } |x - y| < \delta \Longrightarrow |f(x) - f(y)| < \frac{\varepsilon}{b - a}.$$

Choose a partition $P = \{x_0, x_1, \ldots, x_n\} \in \mathcal{P}$ such that $|P| < \delta$. Then for all $j \in [n]$,

$$\begin{aligned} M_j - m_j &= f(x_j^*) - f(x_j'), \text{ for some } x_j^*, x_j' \in [x_{j-1}, x_j] \\ &< \frac{\varepsilon}{b-a} \qquad \text{(since } |x_j^* - x_j'| \leq |P| < \delta) \\ \Longrightarrow U(P, f) - L(P, f) &= \sum_{j=1}^{n} M_j \Delta x_j - \sum_{j=1}^{n} m_j \Delta x_j \\ &= \sum_{j=1}^{n} (M_j - m_j) \Delta x_k \\ &< \sum_{j=1}^{n} \frac{\varepsilon}{b-a} \Delta x_j \\ &= \frac{\varepsilon}{b-a}(b - a) = \varepsilon. \end{aligned}$$

Thus, for all $\varepsilon > 0$, there exists $P \in \mathcal{P}$ such that $U(P, f) - L(P, f) < \varepsilon$. Therefore, f is Riemann integrable on $[a, b]$. ☐

Theorem 10.21. *Let $f : [a, b] \to \mathbb{R}$ be monotonic. Then f is Riemann integrable on $[a, b]$.*

Proof. Assume f is increasing. Let $P = \{x_0, \ldots, x_n\}$ be a uniform partition on $[a, b]$. That is,

$$x_j = a + j\Delta x, \quad j = 0, 1, \ldots, n, \Delta x = \frac{b - a}{n}.$$

Note that

$$M_j = f(x_j), m_j = f(x_{j-1}).$$

Choose n sufficiently large such that

$$\Delta x < \frac{\varepsilon}{f(b) - f(a)}.$$

(If $f(b) = f(a)$, then f is constant, and the result is trivial.) Then

$$\begin{aligned}
U(P,f) - L(P,f) &= \sum_{j=1}^{n} (M_j - m_j)\Delta x \\
&= \sum_{j=1}^{n} (f(x_j) - f(x_{j-1}))\Delta x \\
&= \Delta x \sum_{j=1}^{n} (f(x_j) - f(x_{j-1})) \\
&= \Delta(f(x_n) - f(x_0)) \\
&= \Delta(f(b) - f(a)) \\
&< \frac{\varepsilon}{f(b) - f(a)}(f(b) - f(a)).
\end{aligned}$$

Thus, for all $\varepsilon > 0$, there exists a $P \in \mathcal{P}$ such that

$$U(P,f) - L(P,f) < \varepsilon.$$

Therefore, f is Riemann integrable on $[a,b]$. □

10.5 Lecture 29. Fri Nov 8

10.5.1 Properties of Riemann Integration

Theorem 10.22. *Let $f : [a,b] \to \mathbb{R}$ be bounded. If f has only finitely many discontinuities in $[a,b]$, then f is Riemann integrable on $[a,b]$.*

Proof. Since f is bounded, assume $|f(x)| \le M$, for all $x \in [a,b]$. Let the discontinuities of f be labeled $t_0, t_1, \ldots, t_k \in [a,b]$. Assume

$$\min\{|t_i - t_j| \mid 1 \le i < j < k\} > \frac{\varepsilon}{2kM}.$$

(If this were not true, we could replace ε by a smaller number.) Define

$$u_j = \max\left\{t_j - \frac{\varepsilon}{4kM}, a\right\}, \quad v_j = \min\left\{t_j + \frac{\varepsilon}{4kM}, b\right\},$$

for $j = 1, \ldots, k$. Then $t_j \in [u_j, v_j]$, for all $j = 1, \ldots, k$, and the intervals $[u_i, v_j]$, $j = 1, \ldots, k$ are disjoint and contained in $[a,b]$. Define

$$S = [a,b] - \bigcup_{j=1}^{k} (u_j, v_j).$$

Then S is compact (because it is a closed subset of $[a,b]$) and therefore f is uniformly continuous on S. Therefore, there exists $\delta > 0$ such that

$$x, y \in S \text{ and } |x - y| < \delta \Longrightarrow |f(x) - f(y)| < \frac{\varepsilon}{2(b-a)}.$$

Now define a partition P on $[a, b]$, written $P = \{x_0, x_1, \ldots, x_n\}$, to satisfy.

$$\begin{aligned} u_j \in P, v_j \in P, \quad \forall j = 1, \ldots, k, \\ (u_j, v_j) \cap P = \varnothing, \quad \forall j = 1, \ldots, k \\ \forall i \in [k], j \in [n], x_{j-1} \in P \wedge x_{j-1} \neq u_i \implies \Delta x_j < \delta. \end{aligned}$$

Now

$$U(P, f) - L(P, f) = \sum_{j=1}^{n} (M_j - m_j) \Delta x_j.$$

Write $\{1, 2, \ldots, n\} = J_1 \cup J_2$, where

$$\begin{aligned} i \in J_1 &\iff [x_{i-1}, x_i] \neq [u_j, v_j], \forall j = 1, \ldots, k \\ i \in J_2 &\iff [x_{i-1}, x_i] = [u_j, v_j], \text{ for some } j = 1, \ldots, k. \end{aligned}$$

Then

$$U(P, f) - L(P, f) = \sum_{j \in J_1} (M_j - m_j) \Delta x_j + \sum_{j \in J_2} (M_j - m_j) \Delta x_j.$$

For $j \in J_1$,

$$\begin{aligned} M_j - m_j &= \sup\{f(x) \mid x_{j-1} \leq x \leq x_j\} - \inf\{f(x) \mid x_{j-1} \leq x \leq x_j\} \\ &< \frac{\varepsilon}{2(b-a)} && \text{(since } \Delta x_j < \delta) \\ \implies \sum_{j \in J_1} (M_j - m_j) \Delta x_j &< \sum_{j \in J_1} \frac{\varepsilon}{2(b-a)} \Delta x_j \\ &= \frac{\varepsilon}{2(b-a)} \sum_{j \in J_2} \Delta x_j \\ &\leq \frac{\varepsilon}{2(b-a)} \cdot (b-a) \\ &= \frac{\varepsilon}{2}. \end{aligned}$$

For $j \in J_2$,

$$\begin{aligned} M_j - m_j \leq 2M, \quad \Delta x_j \leq \frac{\varepsilon}{4kM} + \frac{\varepsilon}{4kM} = \frac{\varepsilon}{2kM} \\ \implies \sum_{j \in J_2} (M_j - m_j) \Delta x_j \leq \sum_{j \in J_2} 2M \cdot \frac{\varepsilon}{2kM} = \sum_{j \in J_2} \frac{\varepsilon}{k} = \varepsilon. \end{aligned}$$

(Mark lost a factor of 2! Can you help me find it?) Thus

$$U(P, f) - L(P, f) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore, for all $\varepsilon > 0$, there exists a partition P on $[a, b]$ such that $U(P, f) - L(P, f) < \varepsilon$ so f is Riemann integrable on $[a, b]$. □

Theorem 10.23. *Let $[a, b] \subset \mathbb{R}$ be given, $(b > a)$, and let V be the set of all $f : [a, b] \to \mathbb{R}$ such that f is Riemann integrable. Then f is closed under addition and scalar multiplication. That is*

$$\begin{aligned} f, g \in V \implies f + g \in V, \\ f \in V, c \in \mathbb{R} \implies cf \in V. \end{aligned}$$

Moreover,

$$f, g \in V \implies \int_a^b (f+g) = \int_a^b f + \int_a^b g,$$
$$f \in V, c \in \mathbb{R} \implies \int_a^b cf = c\int_a^b f.$$

Proof. Let $f, g \in V$. Let $\varepsilon > 0$ be given. We must show that there exists $P \in \mathcal{P}$ such that

$$U(P, f+g) - L(P, f+g) < \varepsilon.$$

We know there exists $P_1 \in \mathcal{P}$ such that

$$U(P_1, f) - L(P_1, f) < \frac{\varepsilon}{2},$$

and $P_2 \in \mathcal{P}$ such that

$$U(P_2, f) - L(P_2, f) < \frac{\varepsilon}{2}.$$

Define $P' = P_1 \cup P_2$. Then

$$U(P', f) - L(P', f) < \frac{\varepsilon}{2},$$
$$U(P', g) - L(P', g) < \frac{\varepsilon}{2}.$$

Now,

$$\begin{aligned}
&\sup\{f(x) + g(x) \mid x_{j-1} \le x \le x_j\} \\
&\le \sup\{f(x) \mid x_{j-1} \le x \le x_j\} + \sup\{g(x) \mid x_{j-1} \le x \le x_j\} \\
&\implies U(P', f+g) \le U(P', f) + U(P', g).
\end{aligned}$$

Similarly,

$$L(P', f+g) \ge L(P', f) + L(P', g).$$

Hence

$$\begin{aligned}
U(P', f+g) - L(P', f+g) &\le U(P', f) + U(P', g) - (L(P', f) + L(P', g)) \\
&= U(P', f) - L(P', f) + U(P', g) - L(P', g) \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Hence, $f+g$ is Riemann integrable on $[a, b]$. Also,

$$\begin{aligned}
L(P', f) + L(P', g) \le L(P', f+g) &\le U(P', f+g) \\
&\le U(P', f) + U(P', g)
\end{aligned}$$
$$\implies \int_a^b f + \int_a^b g \le \int_a^b f + g \le \int_a^b f + \int_a^b g$$
$$\implies \int_a^b (f+g) = \int_a^b f + \int_a^b g.$$

□

Theorem 10.24.
1. If $f : [a,b] \to \mathbb{R}, g : [a,b] \to \mathbb{R}$ are Riemann integrable, and $f(x) \leq g(x)$, for all $x \in [a,b]$, then

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

2. If $f : [a,b] \to \mathbb{R}$ is Riemann integrable, then so is $|f|$, and

$$\left|\int_a^b f\right| \leq \int_a^b |f|.$$

Proof.
1. Let $P = \{x_0, x_1, \dots, x_n\} \in \mathcal{P}$ be given. Then

$$\begin{aligned} &f(x) \leq g(x) \forall x \in [a,b] \\ \implies \sup\{f(x) \mid x_{j-1} \leq x \leq x_j\} &\leq \sup\{g(x) \mid x_{j-1} \leq x \leq x_j\}, \forall j \\ &\implies U(P,f) \leq U(P,g). \end{aligned}$$

This holds for all $P \in \mathcal{P}$, so

$$\int_a^b f = \inf\{U(P,f) \mid P \in \mathcal{P}\} \leq \inf\{U(P,g) \mid P \in \mathcal{P}\} = \int_a^b g.$$

2. Let $P = \{x_0, x_1, \dots, x_n\}$ be given. Then

$$U(P,f) - L(P,f) = \sum_{j=1}^n (M_j - m_j)\Delta x_j,$$

and

$$\begin{aligned} M_j - m_j &= \sup\{f(x) \mid x_{j-1} \leq x \leq x_j\} - \inf\{f(x) \mid x_{j-1} \leq x \leq x_j\} \\ &= \sup\{f(x) - f(y) \mid x, y \in [x_{j-1}, x_j]\} \\ &= \sup\{|f(x) - f(y)| \mid x, y \in [x_{j-1}, x_j]\} \\ &\geq \sup\{|f(x)| - |f(y)| \mid x, y \in [x_{j-1}, x_j]\} \\ &= \sup\{|f(x)| \mid x_{j-1} \leq x \leq x_j\} - \inf\{|f(x)| \mid x_{j-1} \leq x \leq x_j\} \\ &= M_j' - m_j', \end{aligned}$$

where

$$U(P,|f|) - L(P,|f|) = \sum_{j=1}^n (M_j - m_j)\Delta x_j.$$

Therefore $|f|$ is Riemann integrable if f is. Then

$$\begin{aligned} &-|f(x)| \leq f(x) \leq |f(x)|, \forall x \in [a,b] \\ \implies &-\int_a^b |f| \leq \int_a^b f \leq \int_a^b |f| && \text{(by 1)} \\ \implies &\left|\int_a^B f\right| \leq \int_a^b |f|. \end{aligned}$$

□

11 Week 11

11.1 Lecture 30. Mon Nov 11

11.1.1 Fundamental Theorem of Calculus

Theorem 11.1. *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable on $[a, b]$, and let $c \in (a, b)$. Then f is also Riemann integrable on $[a, c]$ and on $[c, b]$, and*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Proof. Let P be any partition on $[a, b]$, and define $P' = P \cup \{c\}$. Then $P_1 = P' \cap [a, c]$ is a partition on $[a, c]$, and $P_2 = P' \cap [c, b]$ is a partition on $[c, b]$. We have

$$\begin{aligned} U(P', f) &= U(P_1, f) + U(P_2, f) \\ L(P', f) &= L(P_1, f) + L(P_2, f) \\ \implies U(P', f) - L(P', f) &= U(P_1, f) + U(P_2, f) - (L(P_1, f) + L(P_2, f)) \\ &= (U(P_1, f) - L(P_1, f)) + (U(P_2, f) - L(P_2, f)). \end{aligned}$$

Now, recall that

$$\begin{aligned} U(P', f) &\leq U(P, f), \\ L(P', f) &\geq L(P, f). \end{aligned}$$

Therefore,

$$\begin{aligned} U(P', f) - L(P', f) &\leq U(P, f) - L(P, f) \\ \implies (U(P_1, f) - L(P_1, f)) + (U(P_2, f) - L(P_2, f)) &\leq U(P, f) - L(P, f). \end{aligned}$$

Since

$$U(P_1, f) - L(P_1, f) \geq 0, \quad U(P_2, f) - L(P_2, f) \geq 0,$$

we have

$$U(P_1, f) - L(P_1, f) \leq U(P, f) - L(P, f), \quad U(P_2, f) - L(P_2, f) \leq U(P, f) - L(P, f).$$

But this holds for all $P \in \mathcal{P}$, and hence we can make $U(P_1, f) - L(P_1, f)$ and $U(P_2, f) - L(P_2, f)$ as small as desired, because f is Riemann integrable on $[a, b]$. Hence, f is Riemann integrable on $[a, c]$ and on $[c, b]$. Also, for all $P \in \mathcal{P}$,

$$\begin{aligned} &L(P, f) \leq L(P_1, f) + L(P_2, f) \leq U(P_1, f) + U(P_2, f) \leq U(P, f), \\ &\implies \int_a^b f = \int_a^c f + \int_c^b f. \end{aligned}$$

□

Theorem 11.2 (Fundamental Theorem of Calculus, first version)**.** *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable on $[a, b]$, and define $F[a, b] \to \mathbb{R}$ by*

$$F(x) = \int_a^x f$$

(Note that F is well defined by the last theorem, and by convention, $F(a) = 0$.) Then F is uniformly continuous on $[a, b]$, and if f is continuous at $x \in [a, b]$, then F is differentiable at x, and $F'(x) = f(x)$.

Proof. Since f is Riemann integrable, it must be bounded. Suppose that $|f(x)| \leq M$, for all $x \in [a, b]$. Then, for all $x, y \in [a, b]$, with $x > y$,

$$\begin{aligned} F(x) - F(y) &= \int_a^x f - \int_a^y f \\ &= \int_a^y f + \int_y^x f - \int_a^y f \\ &= \int_y^x f \\ \implies |F(x) - F(y)| &= \left| \int_x^y f \right| \leq \int_y^x |f| \leq \int_y^x M = M(x - y). \end{aligned}$$

So for all $x, y \in [a, b]$,

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

Thus, given any $\varepsilon > 0$, if $\delta = \frac{\varepsilon}{M}$, then

$$\begin{aligned} x, y \in [a, b], |x - y| < \delta \implies |F(x) - F(y)| &\leq M|x - y| \\ &< M \cdot \frac{\varepsilon}{M} \\ &= \varepsilon. \end{aligned}$$

Therefore, F is uniformly continuous on $[a, b]$. Now suppose that f is continuous at $x \in [a, b)$. Then for $h > 0$,

$$\begin{aligned} \frac{F(x + h) - F(x)}{h} &= \frac{1}{h} \left[\int_a^{x+h} f - \int_a^x f \right] \\ &= \frac{1}{h} \int_x^{x+h} f \\ \implies \frac{F(x + h) - F(x)}{h} - f(x) &= \left(\frac{1}{h} \int_x^{x+h} f \right) - f(x) \\ &= \frac{1}{h} \int_x^{x+h} f(t)\, dt - \frac{1}{h} \int_x^{x+h} f(x)\, dt \\ &= \frac{1}{h} \int_x^{x+h} (f(t) - f(x))\, dt. \end{aligned}$$

Let $\varepsilon > 0$, then since f is continuous at x, there exists $\delta > 0$ such that

$$y \in [a, b] \text{ and } |y - x| < \delta \implies |f(y) - f(x)| < \varepsilon.$$

But then

$$
\begin{aligned}
0 < h < \delta &\Longrightarrow |f(t) - f(x)| < \varepsilon, \quad \forall t \in [x, x+h] \\
\Longrightarrow \left| \frac{F(x+h) - F(x)}{h} - f(x) \right| &\leq \frac{1}{h} \int_x^{x+h} |f(t) - f(x)|\, dt \\
&< \frac{1}{h} \int_x^{x+h} \varepsilon\, dt = \varepsilon.
\end{aligned}
$$

Thus

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

and similarly

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

Thus F is differentiable at x, and

$$F'(x) = f(x).$$

□

Theorem 11.3 (Fundamental Theorem of Calculus, second version)**.** *Suppose $f : [a, b] \to \mathbb{R}$ is Riemann integrable on $[a, b]$, and $F : [a, b] \to \mathbb{R}$, such that F is continuous on $[a, b]$, and differentiable on (a, b), and satisfies $F' = f$ on (a, b). Then*

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

Proof. Let $P = \{x_0, x_1, \ldots, x_n\}$ be any partition on $[a, b]$. Then by the mean value theorem,

$$
\begin{aligned}
\exists c_j \in (x_{j-1}, x_j), \quad F'(c_j)\Delta x_j &= F(x_j) - F(x_{j-1}), \forall j = 1, \ldots, n \\
\Longrightarrow \sum_{j=1}^{n} F'(c_j)\Delta x_j &= \sum_{j=1}^{n} (F(x_j) - F(_{x_{j-1}})) \\
\Longrightarrow \sum_{j=1}^{n} f(c_j)\Delta x_j &= F(b) - F(a).
\end{aligned}
$$

But we have that

$$L(P, f) \leq \sum_{j=1}^{n} f(c_j)\Delta x_j \leq U(P, f),$$

and thus

$$
\begin{aligned}
L(P, f) \leq F(b) - F(a) \leq U(P, f), \forall P \in \mathcal{P} \\
\Longrightarrow \sup\{L(P, f) \mid P \in \mathcal{P}\} \leq F(b) - F(a) \leq \inf\{U(P, f) \mid P \in \mathcal{P}\} \\
\Longrightarrow \int_a^b f = F(b) - F(a).
\end{aligned}
$$

□

Theorem 11.4 (Change of variables, or U-substitution). *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, and suppose that $\varphi : [A, B] \to [a, b]$ is surjective, and $\varphi(A) = a$, and $\varphi(B) = b$. Suppose that φ is differentiable on $[A, B]$ and φ' is continuous on $[A, B]$. Then*

$$\int_a^b f(x)\,dx = \int_A^B f(\varphi(t))\varphi'(t)\,dt.$$

Proof. Define $F[a, b] \to \mathbb{R}$ by

$$F(x) = \int_a^x f,$$

and define $G : [A, B] \to \mathbb{R}$ by $G(t) = F(\varphi(t))$. By the fundamental theorems of calculus,

$$F' = f \implies \int_a^b f = F(b) - F(a)$$

and by the chain rule,

$$\begin{aligned} G'(t) &= F'(\varphi(t))\varphi'(t) \\ &= f(\varphi(t))\varphi'(t) \\ \implies \int_A^B f(\varphi(t))\varphi'(t) &= G(B) - G(A) \\ &= F(\varphi(B)) - F(\varphi(A)) \\ &= F(b) - F(a) \\ \implies \int_a^b f(x)\,dx &= \int_A^B f(\varphi(t))\varphi'(t)\,dt. \end{aligned}$$

□

11.2 Lecture 31. Wed Nov 13

11.2.1 Integration By Parts

Theorem 11.5. *Suppose $f : [a, b] \to \mathbb{R}$ is Riemann integrable, with $\mathcal{R}(f) \subset [a, b]$, and $\varphi : [A, B] \to \mathbb{R}$ is continuous. Then $h = \varphi \circ f$ is Riemann integrable in $[a, b]$.*

Proof. Note that h is bounded on $[a, b]$. Suppose $|h(x)| \leq k, \forall x \in [a, b]$. Let $\varepsilon > 0$ be given. We must show that there exists a partition P on $[a, b]$ such that

$$U(P, f) - L(P, f) < \varepsilon.$$

Since φ is uniformly continuous on $[A, B]$, there exists a $\delta \in (0, \varepsilon)$ such that

$$s, t \in [A, B] \text{ and } |s - t| < \delta \implies |\varphi(s) - \varphi(t)| < \varepsilon.$$

Since f is Riemann integrable on $[a, b]$, there exists $P = \{x_0, x_1, \dots x_n\}$, $P \in \mathcal{P}$ such that

$$U(P, f) - L(P, f) < \delta^2.$$

Write

$$\begin{aligned} M_j &= \sup\{f(x) \mid x_{j-1} \le x \le x_j\}, \\ m_j &= \inf\{f(x) \mid x_{j-1} \le x \le x_j\}, \end{aligned}$$

for $j = 1, \dots, n$ and write M_j', m_j' for the same quantities relative to h. So

$$U(P,h) - L(P,h) = \sum_{j=1}^{n} (M_j' - m_j')\Delta x_j.$$

Write $\{1, 2, \dots, n\} = J_1 \cup J_2$, where

$$\begin{aligned} j \in J_1 &\Longrightarrow M_j - m_j < \delta, \\ j \in J_2 &\Longrightarrow M_j - m_J \ge \delta. \end{aligned}$$

Then

$$U(P,h) - L(P,h) = \sum_{j \in J_1} (M_j' - m_j')\Delta x_j + \sum_{j \in J_2} (M_j' - m_j')\Delta x_j.$$

Now,

$$j \in J_1 \Longrightarrow M_j - m_j < \delta \Longrightarrow M_j' - m_j' < \varepsilon.$$

since

$$\begin{aligned} M_j' - m_j' &= \sup\{\varphi(f(x)) \mid x_{j-1} \le x \le x_j\} - \inf\{\varphi(f(x)) \mid x_{j-1} \le x \le x_j\} \\ &= \sup\{\varphi(f(x)) - \varphi(f(y)) \mid x_{j-1} \le x, y \le x_j\} \\ &= \sup\{\varphi(s) - \varphi(t) \mid m_j \le s, t \le M_j\} \\ &< \varepsilon. \qquad (m_j \le s, t \le M_j \Longrightarrow |s - t| < \delta) \end{aligned}$$

Thus,

$$\sum_{j \in J_1} (M_j' - m_j')\Delta x_j \le \sum_{j \in J_1} \varepsilon \Delta x_j \le \varepsilon(b - a).$$

For $j \in J_2$, we only have

$$M_j - m_j \le 2k, \qquad \text{(recall we bounded } h \text{ by } k)$$

and so

$$\sum_{j \in J_2} (M_j' - m_j')\Delta x_j \le 2k \sum_{j \in J_2} \Delta x_j.$$

Also,

$$\begin{aligned} & U(P,f) - L(P,f) < \delta^2 \\ \Longrightarrow & \sum_{j=1}^{n} (M_j - m_j)\Delta x_j < \delta^2 \\ \Longrightarrow & \sum_{j \in J_2} \delta x_j \le \sum_{j \in J_2} (M_j - m_j)\Delta x_j < \delta^2 \qquad (M_j - m_j \ge \delta, \forall j \in J_2) \\ \Longrightarrow & \ \delta \sum_{j \in J_2} \Delta x_j < \delta^2 \\ \Longrightarrow & \sum_{j \in J_2} \Delta x_j < \delta. \end{aligned}$$

Thus,

$$\sum_{j \in J_2} (M_j' - m_j') < 2k\delta < 2k\varepsilon. \qquad (\delta \in (0, \varepsilon))$$

So finally,

$$\begin{aligned} U(P,h) - L(P,h) &= \sum_{j \in J_1} (M_j' - m_j')\Delta x_j + \sum_{j \in J_2} (M_j' - m_j')\Delta x_j \\ &< \varepsilon(b-a) + \varepsilon 2k \\ &= \varepsilon(b - a + 2k). \end{aligned}$$

Thus we can make $U(P,h) - (P,h)$ arbitrarily small (replace the original ε by $\frac{\varepsilon}{b-a+2k}$) and so h is Riemann integrable on $[a, b.]$ □

Corollary 11.5.1. *Suppose $f : [a,b] \to \mathbb{R}$ is Riemann integrable. Then so is $|f|$.*

Proof. Let $\varphi : \mathbb{R} \to \mathbb{R}, \varphi(t) = |t|$. Note that φ is continuous and $|f| = \varphi \circ f$. □

Corollary 11.5.2. *If $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ are Riemann integrable, then so if fg.*

Proof. We have

$$\begin{aligned} (f+g)^2 - (f-g)^2 &= f^2 + 2fg + g^2 - (f^2 - 2fg + g^2) \\ &= 4fg \\ \implies fg &= \frac{1}{4}((f+g)^2 - (f-g)^2). \end{aligned}$$

Now, $(f+g)^2, (f-g)^2$ are Riemann integrable. (Consider $\varphi(t) = t^2$.) Thus, fg is Riemann integrable. □

Theorem 11.6 (Integration by parts)**.** *Let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be differentiable on $[a,b]$, and assume that f', g' are Riemann integrable on $[a,b]$. Then*

$$\int_a^b fg' = f(b)f(b) - f(a)g(a) - \int_a^b f'g.$$

Note. Modern PDE theory is based on integration by parts – Mark.

Proof. Note that $fg', f'g$ are Riemann integrable by the previous corollary. If $h = fg$, then $h' = fg' + f'g$. Therefore,

$$\begin{aligned} \int_a^b h' &= h(b) - h(a) \\ \implies \int_a^b (fg' + f'g) &= f(b)g(b) - f(a)g(a) \\ \implies \int_a^b &= f(b)g(b) - f(a)g(a) - \int_a^b f'g. \end{aligned}$$

□

11.2.2 Sequences of Functions

Let $(X, d_X), (Y, d_Y)$ be metric spaces. Let $E \subset X$, and for each $n \in \mathbb{Z}^+$, let $f_n : E \to Y$ be a function. Let $f : E \to Y$ be another function.

- We say that $f_n \to f$ pointwise on E iff

  $$\forall x \in E, \lim_{n\to\infty} f_n(x) = f(x).$$

  That is,

  $$\forall x \in E, \forall \varepsilon > 0, \exists N \in \mathbb{Z}^+ (n \geq N \implies d_Y(f_n(x), f(x)) < \varepsilon)).$$

- We say that $f_n \to f$ uniformly on E iff

  $$\forall \varepsilon > 0, \exists N \in \mathbb{Z}^+ (x \in E \wedge n \geq N) \implies d_Y(f_n(x), f(x)) < \varepsilon.$$

Example 11.7. Suppose $f_n \to f$. If every f_n is continuous, is f necessarily continuous? Note that

$$\begin{aligned} f \text{ is continuous at } x = a &\iff \lim_{x\to a} f(x) = f(a) \\ &\iff \lim_{x\to a}(\lim_{n\to\infty} f_n(x)) = f(a) \\ &\overset{?}{\iff} \lim_{n\to\infty}(\lim_{x\to a} f_n(x)) = f(a) \\ &\implies \lim_{x\to a} f_n(a) = f(a). \end{aligned}$$

The interchange of limit operations indicated by the $?$ is not valid if the convergence is only pointwise. For example, consider

$$f_n : [0,1] \to \mathbb{R}, f_n(x) = x^n.$$

Then $f_n \to f$ pointwise, where

$$f(x) = \begin{cases} 0, & 0 \leq x < 1 \\ 1, & x = 1. \end{cases}$$

Example 11.8. If $f_n : [a, b] \to \mathbb{R}$ is Riemann integrable, and $f_n \to f$, is f Riemann integrable and does

$$\int_a^b f_n \to \int_a^b f?$$

We are asking if

$$\int_a^b f = \int_a^b \left(\lim_{n\to\infty} f_n\right) \overset{?}{=} \lim_{n\to\infty} \int_a^b f_n?$$

If the convergence is only pointwise, the interchange of limits and integrals indicated by the $?$ is not necessarily valid. Take as an example, consider for all $n \in \mathbb{Z}^+$, $f_n : [0,1] \to \mathbb{R}$ as the piecewise linear functions satisfying

$$\begin{aligned} f_n(0) &= 0, \\ f_n\left(\frac{1}{2n}\right) &= 2n, \\ f_n\left(\frac{1}{n}\right) &= 0, \\ f_n(1) &= 0. \end{aligned}$$

Then $f_n \to f$ where $f = 0$. (Since $f_n(0) = 0, \forall n \implies f_n(0) \to 0$. For $x \in (0,1]$, $\frac{1}{n} < x, \forall n$ sufficiently large, which implies $f_n(x) = 0, \forall x$ sufficiently large, which implies $f_n(x) = 0$.) But

$$\int_0^1 f_n = 1, \quad \forall n \implies \int_0^1 f_n \not\to 0 = \int_0^1 f.$$

Example 11.9. If $f_n : [a,b] \to \mathbb{R}$ is differentiable, for all $n \in \mathbb{Z}^+$, and $f_n \to f$, is f necessarily differentiable? If so, does $f_n' \to f'$? Actually, neither poinwise nor uniform convergence guarantees this. We would need an even stronger form of convergence to guarantee this. Take as an example $f_n = \frac{\sin(nx)}{\sqrt{n}}, \forall n \in \mathbb{Z}^+$. Then $f_n \to f, f = 0$ uniformly, since

$$|f_n(x) - f(x)| \leq \frac{1}{\sqrt{n}}, \forall x \in [a,b].$$

But

$$f_n'(x) = \sqrt{x}\cos nx,$$

and $f_n' \not\to 0$, e.g. $f_n'(0) = \sqrt{n} \to \infty$.

11.3 Lecture 32.

11.3.1 Uniform Convergence

Suppose $(X, d_X), (Y, d_Y)$ are metric spaces. Let $E \subset X$. Recall that if $f_n : E \to Y, \forall n \in \mathbb{Z}^+, f : E \to Y$, then $f_n \to f$ uniformly on E iff

$$\forall \varepsilon > 0, \exists n \in \mathbb{Z}^+ (n \geq N \wedge x \in E) \implies d_Y(f_n(x), f(x)) < \varepsilon.$$

Theorem 11.10. *Let $(X, d_X), (Y, d_Y)$ be metric spaces, and let $E \subset X$. Assume that for all $n \in \mathbb{Z}^+$, $f_n : E \to Y$ is a continuous function. If $f_n \to f$ uniformly on E, where $f : E \to Y$, then f is continuus on E.*

Proof. Let $x \in E$ and let $\varepsilon > 0$ be given. Since $f_n \to f$ uniformly, there exists $N \in \mathbb{Z}^+$ such that

$$n \geq N \wedge u \in E \implies d_Y(f_n(u), f(u)) < \frac{\varepsilon}{3}.$$

In particular,

$$\forall u \in E, d_Y(f_N(u), f(u)) < \frac{\varepsilon}{3}.$$

By assumption, f_N is continuous. Hence, there exists $\delta > 0$ such that

$$(u \in E \wedge d_X(u,x) < \delta) \implies d_Y(f_N(u), f_N(x)) < \frac{\varepsilon}{3}.$$

But then

$$\begin{aligned}
&(u \in E \wedge d_X(u,x) < \delta) \\
&\implies d_Y(f(u), f(x)) \leq d_Y(f(u), f_N(u)) + d_Y(f_N(u), f_N(x)) + d_Y(f_N(x), f(x)) \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

Thus f is continuous on x. Since $x \in E$ was arbitrarily, it follows that f is continuous on E. □

Theorem 11.11. *For each $n \in \mathbb{Z}^+$, let $f_n : [a, b] \to \mathbb{R}$ be Riemann integrable and suppose $f_n \to f$ uniformly on $[a, b]$ where $f : [a, b] \to \mathbb{R}$. Then f is Riemann integrable on $[a, b]$ and*

$$\int_a^b f_n \to \int_a^b f.$$

Proof. First, we prove that f is Riemann integrable on $[a, b]$. Let $\varepsilon > 0$ be given. Since $f_n \to f$ uniformly on $[a, b]$, there exists $N \in \mathbb{Z}^+$ such that

$$n \geq N \Longrightarrow |f_n(x) - f(x)| < \frac{\varepsilon}{3(b-a)}, \quad \forall x \in [a, b].$$

In particular,

$$|f_N(x) - f(x)| < \frac{\varepsilon}{3(b-a)}, \quad x \in [a, b].$$

Since f_N is Riemann integrable on $[a, b]$, there exists a partition $P = \{x_0, x_1, \ldots, x_n\}$ such that

$$U(P, f_N) - L(P, f_N) < \frac{\varepsilon}{3}.$$

But

$$\begin{aligned} U(P, f) &= \sum_{j=1}^{n} \sup\{f(x) \mid x_{j-1} \leq x \leq x_j\} \Delta x_j \\ &< \sum_{j=1}^{n} \left(\sup\{f_N(x) \mid x_{j-1} \leq x \leq x_j\} + \frac{\varepsilon}{3(b-a)} \right) \Delta x_j \\ &= U(P, f_N) + \frac{\varepsilon}{3(b-a)} \sum_{j=1}^{n} \Delta x_j \\ &= U(P, f_N) + \frac{\varepsilon}{3}. \end{aligned}$$

Similarly,

$$L(P, f) > L(P, f_N) - \frac{\varepsilon}{3}.$$

Therefore,

$$\begin{aligned} U(P, f) - L(P, f) &< \left(U(P, f_N) + \frac{\varepsilon}{3} \right) - \left(L(P, f_N) - \frac{\varepsilon}{3} \right) \\ &= U(P, f_N) - L(P, f_N) + \frac{2\varepsilon}{3} \\ &< \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} = \varepsilon. \end{aligned}$$

Thus f is Riemann integrable on $[a, b]$. Now we show $\int_a^b f_n \to \int_a^b f$. Let $\varepsilon > 0$ be given. Then there exists an $N \in \mathbb{Z}^+$ such that

$$n \geq N \Longrightarrow \left(|f_n(x) - f(x)| < \frac{\varepsilon}{b-a}, \quad \forall x \in [a, b] \right).$$

But then

$$n \geq N \Longrightarrow \left| \int_a^b f_n - \int_a^b f \right| \leq \int_a^b |f_n - f| < \int_a^b \frac{\varepsilon}{b-a} = \varepsilon.$$

This completes the proof. □

The above theorem shows why uniform convergence is so powerful. We now give some technical results that are useful for verifying uniform convergence.

Theorem 11.12. *Let $(X, d_X), (Y, d_Y)$ be metric spaces, and let $E \subset X$ and let $f_n : E \to Y$, $n = 1, 2, \ldots,$ and let $f : E \to Y$ be given functions. Define*

$$M_n = \sup\{d_Y(f_n(x), f(x)) \mid x \in E\}.$$

Then $f_n \to f$ uniformly on E iff $M_n \to 0$.

Proof. Immediately from the definition of uniform convergence. □

Theorem 11.13 (Cauchy criterion for uniform convergence)**.** *Let $(X, d_X), (Y, d_Y)$ be metric spaces, let $E \subset X$, and let $f_n : E \to Y$, for all $n \in \mathbb{Z}^+$, be given functions. Then if $\{f_n\}$ converges uniformly on E, then*

$$\forall \varepsilon > 0, \exists N \in \mathbb{Z}^+ (m, n \geq N \implies d_Y(f_m(x), f_n(x)) < \varepsilon, \quad \forall x \in E). \tag{11.13}$$

If Y is complete, then the converse holds.

Proof. Suppose first that $\{f_n\}$ converges uniformly on E, say, $f_n \to f$ uniformly on E. Let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{Z}^+$ such that

$$n \geq N \implies \left(d_Y(f_n(x), f(x)) < \frac{\varepsilon}{2}, \quad \forall x \in E\right).$$

But then

$$m, n \geq N \implies \left(d_Y(f_m(x), f_n(x)) \leq d_Y(f_m(x), f(x)) + d_Y(f(x), f_n(x)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \forall x \in E\right).$$

Thus Eq. (11.13) holds. Conversely, suppose Eq. (11.13) holds and Y is complete. Then Eq. (11.13) implies that $\{f_n(x)\}$ is a Cauchy sequence for each $x \in E$ and converges since Y is complete. Define $f : E \to Y$ by

$$f(x) = \lim_{n \to \infty} f_n(x), \quad \forall x \in E.$$

Let $\varepsilon > 0$ be given. By Eq. (11.13), there exists $N \in \mathbb{Z}^+$ such that

$$m, n \geq N \implies \left(d_Y(f_m(x), f_n(x)) < \frac{\varepsilon}{2}, \quad \forall x \in E\right).$$

We claim that

$$n \geq N \implies (d_Y(f_n(x), f(x)) < \varepsilon, \quad \forall x \in E).$$

To see this, let $n \geq N$ be fixed. For any $x \in E$, there exists $N_x \geq N$ such that

$$m \geq N_x \implies d_Y(f_m(x), f(x)) < \frac{\varepsilon}{2}.$$

But then

$$\begin{aligned} d_Y(f_n(x), f(x)) &\leq d_Y(f_n(x), f_{N_x}(x)) + d_Y(f_{N_x}(x), f(x)) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Thus, we have shown $f_n \to f$ uniformly. □

11.3.2 Weierstrauss M-test

We will now show that a series $\sum_{n=1}^{\infty} f_n$ of real-valued functions converges uniformly iff the sequence $\left\{\sum_{n=1}^{N} f_n\right\}$ converges uniformly.

Theorem 11.14 (Weierstrauss M-test)**.** *Let $E \subset \mathbb{R}$, and suppose $f_n : E \to \mathbb{R}$ is a given function for each $n \in \mathbb{Z}^+$. If there exists a sequence $\{M_n\}$ of nonnegative real numbers such that*

$$\sum_{n=1}^{\infty} M_n \quad \textit{converges,}$$

and

$$|f_n(x)| \leq M_n, \quad \forall x \in E, \forall n \in \mathbb{Z}^+,$$

then $\sum_{n=1}^{\infty} f_n$ converges uniformly on E.

Proof. Let $\{S_n\}$ denote the sequence of partial sums:

$$S_n(x) = \sum_{k=1}^{n} f_k(x), \quad \forall x \in E.$$

Let $\varepsilon > 0$ be given. Then since $\sum_{k=1}^{\infty} M_k$ converges, there exists some $N \in \mathbb{Z}^+$ such that

$$m \geq n \geq N \implies \left|\sum_{k=n}^{m} M_k\right| = \sum_{k=n}^{n} M_k < \varepsilon,$$

by the Cauchy criterion for series. But then

$$m, n \geq N \implies |S_m(x) - S_n(x)| = \left|\sum_{k=n+1}^{m} f_k(x)\right| \leq \sum_{k=n+1}^{m} |f_k(x)| \leq \sum_{k=n+1}^{m} M_k < \varepsilon, \quad \forall x \in E.$$

Hence, $\{S_n\}$ is uniformly Cauchy on E. It follows that $\{S_n\}$ converges uniformly on E, that is, $\sum_{n=1}^{\infty} f_n$ converges uniformly on E. □

12 Week 12

12.1 Lecture 33. Mon Nov 18

Up until now we have defined

$$\int_a^b f,$$

for $b > a$ (or if $b = a$, $\int_a^b f = 0$.) We now define

$$\int_b^a f = -\int_a^b f \quad \text{if } a < b.$$

Now suppose $a < b < c$. Then

$$\begin{aligned} \int_a^c f &= \int_a^b f + \int_b^c f \\ \implies \int_a^b f = \int_a^c f - \int_b^c f &= \int_a^c f + \int_c^b f && \left(\text{since } \int_c^b f = -\int_b^c f\right) \\ \implies \int_a^b f &= \int_a^c f + \int_c^b f. \end{aligned}$$

In fact, this holds for any ordering of a, b, c. Also, suppose f is continuous on $[a, b]$ and $F : [a, b] \to \mathbb{R}$ is defined by

$$F(x) = \int_{x_0}^x f,$$

where $x_0 \in [a, b]$. We have

$$\begin{aligned} F'(x) &= \lim_{h\to\infty} \frac{F(x+h) - F(x)}{h} \\ &= \lim_{h\to\infty} \frac{1}{h}\left[\int_{x_0}^{x+h} f - \int_{x_0}^x f\right] && (f(x+h) \approx f(x) \text{ for } h \text{ arb. small}) \\ &= \lim_{h\to\infty} \frac{1}{h}\int_x^{x+h} f = f(x). \end{aligned}$$

Theorem 12.1. *Let $f_n : [a, b] \to \mathbb{R}$ be continuously differentiable on $[a, b]$. (i.e. f_n' exists and is continuous on $[a, b]$) for each $n \in \mathbb{Z}^+$. Suppose $f_n' \to g$ uniformly on $[a, b]$, where $g : [a, b] \to \mathbb{R}$, and suppose there exists $x_0 \in [a, b]$ such that $\{f_n(x_0)\}$ converges. Then there exists $f : [a, b] \to \mathbb{R}$ such that $f_n \to f$ uniformly on $[a, b]$ and $f' = g$.*

Proof. Suppose $f_n(x_0) \to c \in \mathbb{R}$. Then define $f : [a, b] \to \mathbb{R}$ by

$$f(x) = c + \int_{x_0}^x g.$$

Since each f_n' is continuous and $f_n' \to g$ uniformly, we know that g is continuous and hence Riemann integrable. Thus f is well defined. By the fundamental theorem of calculus,

$$\int_{x_0}^x f_n' = f_n(x) - f_n(x_0).$$

Since $f_n' \to g$ uniformly, we have

$$\int_{x_0}^x f_n' \to \int_{x_0}^x g.$$

Thus,

$$f_n(x) = f_n(x_0) + \int_{x_0}^x f_n' \to c + \int_{x_0}^x g = f(x).$$

Also by the fundamental theorem, $f' = g$. So far we have proved $f_n \to f$ pointwise. We have

$$\begin{aligned}
|f_n(x) - f(x)| &= \left| f_n(x_0) + \int_{x_0}^{x} f_n' - c - \int_{x_0}^{x} g \right| \\
&= \left| (f_n(x_0) - c) + \int_{x_0}^{x} (f_n' - g) \right| \\
&\leq |f_n(x_0) - c| + \left| \int_{x_0}^{x} |f_n' - g| \right|.
\end{aligned}$$

Let $\varepsilon > 0$ be given. Since $f_n(x_0) \to c$, there exists $N_1 \in \mathbb{Z}^+$ such that

$$n \geq N_1 \Longrightarrow |f_n(x_0) - c| < \frac{\varepsilon}{2}.$$

Since $f_n' \to g$ uniformly on $[a, b]$, there exists $N_2 \in \mathbb{Z}^+$ such that for all $x \in [a, b]$,

$$\begin{aligned}
n \geq N_2 &\Longrightarrow |f_n'(x) - g(x)| < \frac{\varepsilon}{2(b-a)}, \forall x \in [a, b] \\
&\Longrightarrow \left| \int_{x_0}^{x} |f_n' - g| \right| < \left| \int_{x_0}^{x} \frac{\varepsilon}{2(b-a)} \right| \\
&\qquad = \frac{\varepsilon}{2(b-a)} |x - x_0| \\
&\qquad \leq \frac{\varepsilon}{2}. \qquad \left(\tfrac{x - x_0}{b - a} \leq 1\right)
\end{aligned}$$

Thus, if $N = \max\{N_1, N_1\}$, then for all $x \in [a, b]$,

$$n \geq N \Longrightarrow |f_n(x) - f(x)| \leq |f_n(x_0) - c| + \left| \int_{x_0}^{x} |f_n - g| \right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus $f_n \to f$ uniformly on $[a, b]$. □

12.1.1 Sequences of Functions Applied to Power Series

Let $\{c_n\}$ be a sequence of real numbers, let $x_0 \in \mathbb{R}$ be given, and let

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n$$

have radius of convergence $R > 0$ (possibly $R = \infty$). Recall that

$$R = \frac{1}{\limsup_{n \to \infty} |c_n|^{\frac{1}{n}}}.$$

We exclude the possibility that $c_n \to \infty$ or $c_{n_k} \to \infty$ so quickly that

$$\limsup_{n \to \infty} |c_n|^{\frac{1}{n}} = \infty.$$

Define $f : (x_0 - R, x_0 + R) \to \mathbb{R}$ by

$$f(x) = \sum_{n=0}^{\infty} c_n (x - x_0)^n. \tag{12.1.1}$$

Also define $f_n : (x_0 - R, x_0 + R) \to \mathbb{R}$ by

$$f_n(x) = \sum_{k=0}^{n} c_k (x - x_0)^k. \qquad \text{(the } n\text{th partial sum)}$$

Then we know $f_n \to f$ pointwise on $(x_0 - R, x_0 + R)$.

Proposition 12.2. *For each $r \in (0, R)$, $f_n \to f$ uniformly on $[x_0 - r, x_0 + r]$. Thus f is continuous on $(x_0 - R, x_0 + R)$.*

Proof. Let $r \in (0, R)$. Then for all $x \in [x_0 - r, x_0 + r]$,

$$|c_n(x - x_0)^n| \le |c_n| r^n$$

and

$$\sum_{n=0}^{\infty} |c_n| r^n$$

converges since

$$\limsup_{n\to\infty} ||c_n| r^n|^{\frac{1}{n}} = r \limsup_{n\to\infty} |c_n|^{\frac{1}{n}} = \frac{r}{R} < 1.$$

Thus, by the Weierstrauss M-test, $f_n \to f$ uniformly on $[x_0 - r, x_0 + r]$, and hence f is continuous on $[x_0 - r, x_0 + r]$. Since every $x \in (x_0 - R, x_0 + R)$ lies in $[x_0 - r, x_0 + r]$ for some $r \in (0, R)$, it follows that f is continuous at every $x \in (x_0 - R, x_0 + R)$. ☐

Theorem 12.3. *The given f (Eq. (12.1.1)) is differentiable on $(x_0 - R, x_0 + R)$, and*

$$f'(x) = \sum_{n=1}^{\infty} n c_n (x - x_0)^{n-1}. \tag{12.1.2}$$

In other words, "term-by-term" differentiation is valid.

Proof. We have

$$\begin{aligned}
\limsup_{n\to\infty} |n c_n (x - x_0)^{n-1}|^{\frac{1}{n}} &= \limsup_{n\to\infty} n^{\frac{1}{n}} |c_n|^{\frac{1}{n}} |x - x_0|^{1-\frac{1}{n}} \\
&= |x - x_0| \limsup_{n\to\infty} \frac{n^{\frac{1}{n}} |c_n|^{\frac{1}{n}}}{|x - x_0|^{\frac{1}{n}}} \\
&= |x - x_0| \limsup_{n\to\infty} |c_n|^{\frac{1}{n}} = \frac{|x - x_0|}{R},
\end{aligned}$$

since

$$n^{\frac{1}{n}} \to 1 \text{ and } |x - x_0|^{\frac{1}{n}} \to 1, \quad \forall x \neq x_0.$$

Thus by the root test, the series Eq. (12.1.2) converges for all x such that $|x - x_0| < R$, i.e. $\forall x \in (x_0 - R, x_0 + R)$. By the previous theorem, the convergence is uniform on $[x_0 - r, x_0 + r]$ for any $r \in (0, R)$. Write

$$\begin{aligned}
f_n(x) &= \sum_{k=0}^{n} c_k (x - x_0)^k \\
\implies f_n'(x) &= \sum_{k=0}^{n} k c_k (x - x_0)^{k-1},
\end{aligned}$$

and we have that $f_n' \to g$ uniformly on $[x_0 - r, x_0 + r]$, where

$$g(x) = \sum_{k=0}^{\infty} kc_k(x - x_0)^{k-1}.$$

Since $f_n(x_0) = 0$, for all $n \in \mathbb{Z}^+$, $\{f_n(x_0)\}$ converges, and hence there exists $f : [x_0 - r, x_0 + r] \to \mathbb{R}$ such that $f_n \to f$ uniformly on $[x_0 - r, x_0 + r]$ and $f' = g$. We already know f :

$$f(x) = \sum_{k=0}^{\infty} c_k(x - x_0)^k,$$

thus

$$\frac{d}{dx}\left[\sum_{k=0}^{\infty} c_k(x - x_0)^k\right] = \sum_{k=0}^{\infty} kc_k(x - x_0)^{k-1}.$$

□

"Power series are great!"

— Mark Gockenbach

They look like polynomials of infinite degree, and act like them. We can differentiate them like polynomials, add them like polynomials, multiply them like polymomials, and in our next homework, we will prove that you can integrate them like polynomials.

12.2 Lecture 34. Wed Nov 20

12.2.1 Equicontinuity, Pointwise Boundedness

Definition 12.4. Let (X, d) be a metric space, and let $E \subset X$ be compact. We define

$$C(E) = \{f : E \to \mathbb{R} \mid f \text{ is continuous on } E\}.$$

For all $f \in C(E)$, we define

$$\|f\|_\infty = \max\{|f(x)| \mid x \in E\}.$$

It can be verified that $\|\cdot\|_\infty$ defines a norm on $C(E)$. The corresponding metric on $C(E)$ is

$$d_\infty(f, g) = \|f - g\|, \forall f, g \in C(E).$$

Theorem 12.5. *Let (X, d) be a metric space, and let $E \subset X$ be compact. Then $(C(E), d_\infty)$ is a complete metric space.*

Proof. Suppose $\{f_n\} \subset C(E)$ is Cauchy. Let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{Z}^+$ such that

$$m, n \geq N \implies \|f_m - f_n\| < \varepsilon,$$

that is

$$m, n \geq N \implies (|f_m(x) - f_n(x)| < \varepsilon, \quad \forall x \in E).$$

This shows that $\{f_n(x)\}$ is Cauchy in $\mathbb{R}$ for all $x \in E$, and hence, since $\mathbb{R}$ is complete,

$$\lim_{n\to\infty} f_n(x)$$

exists. Define $f : E \to \mathbb{R}$ by

$$f(x) = \lim_{n\to\infty} f_n(x), \quad \forall x \in E.$$

Again, let $\varepsilon > 0$ be given and let $N \in \mathbb{Z}^+$ satisfy

$$m, n \geq N \Longrightarrow \left(|f_m(x) - f_n(x)| < \frac{\varepsilon}{2}, \quad \forall x \in E \right).$$

For each $x \in E$, there exists $N_x \in \mathbb{Z}^+$, $N_x \geq N$ such that

$$n \geq N_x \Longrightarrow |f_n(x) - f(x)| < \frac{\varepsilon}{2}.$$

But then

$$\begin{aligned} n \geq N &\Longrightarrow \left(\forall x \in E, |f_n(x) - f(x)| \leq |f_n(x) - f_{N_x}(x)| + |f_{N_x}(x) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \right) \\ &\Longrightarrow \|f_n - f\| < \varepsilon. \end{aligned}$$

Thus $f_n \to f$ in $C(E)$ (i.e. $f_n \to f$ uniformly on E), and we have shown that $C(E)$ is complete. □

We now wish to show that certain subsets of $C(E)$ are compact. More specifically, we wish to derive conditions on $\{f_n\} \subset C(E)$ guaranteeing that $\{f_n\}$ has a subsequence converging in $C(E)$ (i.e. converging uniformly on E).

Definition 12.6. Let (X, d) be a metric space, and let $E \subset X$, and let $f_n : E \to \mathbb{R}$ for all $n \in \mathbb{Z}^+$. We say that $\{f_n\}$ is pointwise bounded on E iff

$$\forall x \in E, \exists M > 0 \left(|f_n(x)| \leq M, \forall n \in \mathbb{Z}^+ \right).$$

We say $\{f_n\}$ is uniformly bounded on E iff

$$\exists M > 0 (|f_n(x)| \leq M, \forall x \in E, \forall n \in \mathbb{Z}^+).$$

Definition 12.7. Let (X, d) be a metric space, and let $E \subset X$, and let $\mathscr{F}$ be any set of functions of the type $f : E \to \mathbb{R}$. We say that $\mathscr{F}$ is equicontinuous ("uniformly uniformly continuous") iff

$$\forall \varepsilon > 0, \exists \delta > 0 (x, y \in E \wedge d(x, y) < \delta) \Longrightarrow (|f(x) - f(y)| < \varepsilon, \forall f \in \mathscr{F}).$$

Note that if $\mathscr{F}$ is equicontinuous, then every $f \in \mathscr{F}$ is uniformly continuous.

Theorem 12.8. *Let (X, d) be a metric space, and let $E \subset X$ be compact. Let $\{f_n\} \subset C(E)$. If $\{f_n\}$ converges in $C(E)$ (i.e. if $\{f_n\}$ converges uniformly on E) then $\{f_n\}$ is equicontinuous.*

Proof. Suppose $f_n \to f$ uniformly on E, and let $\varepsilon > 0$ be given. We know that f is continuous (the uniform limit of continuous functions is continuous) and hence uniformly continuous (since E is compact). hence there exists $\delta_0 > 0$ such that

$$(x, y \in E \wedge d(x, y) < \delta_0) \Longrightarrow |f(x) - f(y)| < \frac{\varepsilon}{3}.$$

Since $f_n \to f$ uniformly on E, there exists $N \in \mathbb{Z}^+$ such that

$$n \geq N \Longrightarrow |f_n(x) - f(x)| < \frac{\varepsilon}{3}, \quad \forall x \in E.$$

Suppose $n \geq N$. Then

$$\begin{aligned}(x, y \in E \wedge d(x,y) < \delta_0) \Longrightarrow |f_n(x) - f_n(y)| &\leq |f_n(x) - f(x)| + |f(x) - f(y)| - |f(y) - f_n(y)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon.\end{aligned}$$

Since each f_n is continuous on E and hence uniformly continuous on E, for each $n = 1, 2, \ldots, N-1$, there exists $\delta_n > 0$ such that

$$(x, y \in E \wedge d(x,y) < \delta_n) \Longrightarrow |f_n(x) - f_n(y)| < \varepsilon.$$

Therefore, if $\delta = \min\{d_0, d_1, \ldots, d_{N-1}\}$, then

$$(n \in \mathbb{Z}^+ \wedge x, y \in E \wedge d(x,y) < \delta) \Longrightarrow |f_n(x) - f_n(y)| < \varepsilon.$$

It follows that $\{f_n\}$ is equicontinuous. □

Lemma 12.9. *Let E be a countable set and suppose $f_n : E \to \mathbb{R}$ for all $n \in \mathbb{Z}^+$. If $\{f_n\}$ is pointwise bounded on E, then $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E$.*

Proof. Let $E = \{x_n\}$. Since $\{f_n(x_1)\}$ is bounded, by the Heine-Borel theorem, there exists a subsequence of $\{f_n(x_1)\}$ that converges, let us call it $\{f_{1,k}(x)\}$. Now suppose we have identified subsequences $\{f_{1,k}\}, \{f_{2,k}\}, \ldots, \{f_{\ell,k}\}$ such that $\{f_{j,k}(x_\ell)\}$ converges for all $j = 1, 2, \ldots, \ell$, and $\{f_{j+1,k}\}$ is a subsequence of $\{f_{j,k}\}$, for all $j = 1, 2, \ldots, \ell - 1$. Consider $\{f_{\ell,k}(x_{\ell+1})\}$. This sequence is bounded (since it's a subsequence of $\{f_n(x_{\ell+1})\}$), and hence it has a subsequence $\{f_{\ell+1,k}(x_\ell + 1)\}$ that converges in $\mathbb{R}$.
In this way, we have constructed a sequence of subsequences:

$$\{f_{\ell,k}\}, \ell = 1, 2, 3, \ldots.$$

Now define $\{f_{n_k}\}$ by $f_{n_k} = f_{k,k}, \forall k \in \mathbb{Z}^+$. For each $j \in \mathbb{Z}^+$, $\{f_{n_k}(x_k)\}$ is a subsequence of $\{f_{j,k}(x_j)\}$, and hence $\{f_{n_k}(x_j)\}$ converges. This completes the proof. □

12.2.2 Arzela-Ascoli Theorem

Theorem 12.10 (Arzela-Ascoli theorem). *Let (X, d) be a metric space, and let $E \subset X$ be compact. Let $f_n \in C(E)$, for all $n \in \mathbb{Z}^+$. Suppose $\{f_n\}$ is pointwise bounded and equicontinuous on E. Then $\{f_n\}$ is uniformly bounded on E and $\{f_n\}$ contains a uniformly convergent subsequence.*

Proof. Let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that

$$\forall n \in \mathbb{Z}^+, \forall x, y \in E(d(x,y) < \delta \Longrightarrow |f_n(x) - f_n(y)| < \varepsilon).$$

Note that $\{B_\delta(x) \mid x \in E\}$ is an open cover for E, and since E is compact, there exists $p_1, \ldots, p_k \in E$ such that

$$E \subset \bigcup_{j=1}^{k} B_\delta(p_j).$$

For each $j = 1, \ldots, k$, there exists $M_j > 0$ such that

$$\forall n \in \mathbb{Z}^+, |f_n(p_j)| \leq M_j. \quad \Longrightarrow \forall n \in \mathbb{Z}^+, \forall j = 1, 2, \ldots, k, |f_n(p_j)| \leq M := \max\{M_1, \ldots, M_k\}.$$

But then, for all $x \in E$, there exists $j \in \{1, 2, \ldots, k\}$ such that $x \in B_\delta(p_j)$, and hence, $d(x, p_j) < \delta$, which implies that

$$\left(|f_n(x) - f_n(p_j)| < \varepsilon, \forall n \in \mathbb{Z}^+\right) \Longrightarrow \left(|f_n(x)| < |f(p_j)| + \varepsilon \leq M + \varepsilon, \forall n \in \mathbb{Z}^+\right).$$

That is,

$$\forall n \in \mathbb{Z}^+, \forall x \in E, |f_n(x)| \leq M + \varepsilon.$$

Thus, $\{f_n\}$ is uniformly bounded on E. Now, since E is compact, it contains a countable dense subset S. By the previous lemma, there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in S$. We will prove that $\{f_{n_k}\}$ converges uniformly on E. Let $\varepsilon > 0$ be given and choose $\delta > 0$ such that

$$\forall n \in \mathbb{Z}^+, \forall x, y \in E \left(d(x, y) < \delta \Longrightarrow |f_n(x) - f_n(y)| < \frac{\varepsilon}{3}\right).$$

Note that

$$E \subset \bigcup_{x \in S} B_\delta(x),$$

since S is dense in E. (For all $y \in E$, there exists $x \in S$ such that $d(y, x) < \delta$.) Therefore, since E is compact, there exists $x_1, \ldots, x_n \in S$ such that

$$E \subset \bigcup_{j=1}^{n} B_\delta(x_j).$$

We know that $\{f_{n_k}(x_j)\}$ converges for each $j = 1, 2, \ldots, n$, so there exists $N \in \mathbb{Z}^+$ such that

$$i, j \geq N \Longrightarrow \left(|f_{n_i}(x_\ell) - f_{n_j}(x_\ell)| < \frac{\varepsilon}{3}, \quad \forall \ell = 1, \ldots, n\right).$$

Now let $x \in E$ be arbitrary. Then $x \in B_\delta(x_\ell)$ for some $\ell \in \{1, 2, \ldots, n\}$, and hence

$$|f_{n_k}(x) - f_{n_k}(x_\ell)| < \frac{\varepsilon}{3}, \quad \forall k \in \mathbb{Z}^+.$$

But then $(i, j \geq N \wedge x \in E)$ imply

$$\begin{aligned} |f_{n_i}(x) - f_{n_j}(x)| &\leq |f_{n_i}(x) - f_{n_i}(x_\ell)| + |f_{n_i}(x_\ell) - f_{n_j}(x_\ell)| + |f_{n_j}(x_\ell) - f_{n_j}(x)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Thus, by the Cauchy criterion, $\{f_{n_k}\}$ converges uniformly. □

12.3 Lecture 35. Fri Nov 22

12.3.1 Functions of Several Variables

Consider functions of the form $f : \mathbb{R}^n \to \mathbb{R}^m$. We've intensely studied the case when $n = m = 1$.

Definition 12.11. Let $L : \mathbb{R}^n \to \mathbb{R}^m$. We say that L is linear if

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y), \quad \forall x, y \in \mathbb{R}^n, \forall \alpha, \beta \in \mathbb{R}.$$

Definition 12.12 (Grown-up definition of derivative)**.** Let $E \subset \mathbb{R}^n$ be an open subset, and let $f : E \to \mathbb{R}^n$. We say that f is differentiable at $x \in E$ iff there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h\to 0} \frac{\|f(x+h) - f(x) - L(h)\|}{\|h\|} = 0, \quad (h \in \mathbb{R}^n).$$

(Using the Euclidean norm, although all norms are topologically equivalent on $\mathbb{R}^n$.) In the case that f is differentiable at x, we write $Df(x)$ for L, and call it the derivative of f at x.

What does this say about $f'(x)$ when $m = n = 1$? Then we have

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

But in a linear algebra class, we'd learn that every linear map from $\mathbb{R}$ to $\mathbb{R}$ has the form $L(x) = ax$ for some constant $a \in \mathbb{R}$. Thus,

$$\begin{aligned} \lim_{h\to 0} \frac{|f(x+h) - f(x) - ax|}{h} = 0 &\Longrightarrow \lim_{h\to 0} \left| \frac{f(x+h) - f(x)}{h} - a \right| = 0 \\ &\Longrightarrow \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = a. \end{aligned}$$

So every linear map $L : \mathbb{R}^n \to \mathbb{R}^n$ is represented by a scalar $a \in \mathbb{R}$, and $f'(x) \in \mathbb{R}$ is the representative of $Df(x)$.

Theorem 12.13. *If $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then there exists $M \in \mathbb{R}^{m\times n}$ such that*

$$L(x) = Mx, \quad \forall x \in \mathbb{R}^n.$$

We've deduced that when $m = n = 1$, $f'(x) \in \mathbb{R}$ is the representative of $Df(x)$. Let us shift our focus to the case when $m > n$, and $n = 1$. Let us refer to $x(t)$ rather than $f(x)$. Then $Dx(t)$ is a linear map from $\mathbb{R}$ to $\mathbb{R}^m$, and it's represented by an m-by-1 matrix, or a vector in $\mathbb{R}^m$. We write $x'(t)$, or $\dot{x}(t)$. So,

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_m(t) \end{bmatrix} \Longrightarrow \dot{x}(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_m'(t) \end{bmatrix}.$$

How did we get here? Consider

$$
\begin{aligned}
&\lim_{h\to 0}\frac{\|x(t+h)-x(t)-L(h)\|}{|h|}=0, \quad L(h)=hv, v\in\mathbb{R}^m\\
\implies &\lim_{h\to 0}\frac{\|x(t+h)-x(t)-hv\|}{|h|}=0\\
\implies &\lim_{h\to 0}\left\|\frac{x(t+h)-x(t)}{h}-v\right\|=0\\
\implies &\lim_{h\to 0}\left\|\begin{bmatrix}\frac{x_1(t+h)-x_1(t)}{h}\\ \frac{x_2(t+h)-x_2(t)}{h}\\ \vdots\\ \frac{x_m(t+h)-x_m(t)}{h}\end{bmatrix}-\begin{bmatrix}v_1\\ v_2\\ \vdots\\ v_m\end{bmatrix}\right\|=0\\
\implies &\lim_{h\to 0}\left|\frac{x_j(t+h)-x_j(t)}{h}-v_j\right|=0, \quad j=1,2,\ldots,m\\
\implies &v_j=\lim_{h\to 0}\frac{x_j(t+h)-x_j(t)}{h}=x_j'(t).
\end{aligned}
$$

Now, let us study the case when $n, m > 1$. Then $Df(x)$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ that is represented by a matrix $f'(x) \in \mathbb{R}^{m\times n}$, called the Jacobian matrix of f at x:

$$
f'(x)=\begin{bmatrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n}\\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n}\end{bmatrix}, \quad f(x)=\begin{bmatrix}f_1(x)\\ f_2(x)\\ \vdots\\ f_n(x)\end{bmatrix}.
$$

Why is this formula correct? We know that the jth column of $f'(x)$ is $Df(x)e_j$ where e_j is the jth standard basis vector. Then

$$\begin{aligned}
&\lim_{h\to 0^+}\frac{\|f(x+he_j)-f(x)-Df(x)(he_j)\|}{\|he_j\|}=0\\
\implies &\lim_{h\to 0^+}\frac{\|f(x+he_j)-f(x)-hDf(x)e_j\|}{h}=0\\
\implies &\lim_{h\to 0^+}\left\|\frac{f(x+he_j)-f(x)}{h}-Df(x)\right\|=0\\
\implies &\lim_{h\to 0^+}Df(x)=\lim_{h\to 0^+}\frac{f(x+he_j)-f(x)}{h}\\
&=\lim_{h\to 0}\begin{bmatrix}\frac{f_1(x+he_j-f_1(x))}{h}\\ \frac{f_2(x+he_j-f_2(x))}{h}\\ \vdots\\ \frac{f_m(x+he_j-f_m(x))}{h}\end{bmatrix}\\
&=\begin{bmatrix}\lim_{h\to 0^+}\frac{f_1(x+he_j-f_1(x))}{h}\\ \lim_{h\to 0^+}\frac{f_2(x+he_j-f_2(x))}{h}\\ \vdots\\ \lim_{h\to 0^+}\frac{f_m(x+he_j-f_m(x))}{h}\end{bmatrix}\\
&=\begin{bmatrix}\frac{\partial f_1}{\partial x_j}(x)\\ \frac{\partial f_2}{\partial x_j}(x)\\ \vdots\\ \frac{\partial f_m}{\partial x_j}(x)\end{bmatrix},
\end{aligned}$$

which is the jth column of $f'(x)$. Lastly, let us study the case when $m=1$, $n>1$. Then $Df(x)$ is represented by the $1\times n$ matrix,

$$f'(x)=\begin{bmatrix}\frac{\partial f}{\partial x_1}(x) & \cdots & \frac{\partial f}{\partial x_n}(x)\end{bmatrix},$$

and

$$\begin{aligned}
f(x+h)\approx f(x)+f'(x)h&=f(x)+\begin{bmatrix}\frac{\partial f}{\partial x_1}(x) & \cdots & \frac{\partial f}{\partial x_n}(x)\end{bmatrix}\begin{bmatrix}h_1\\ \vdots\\ h_m\end{bmatrix}\\
&=f(x)+\overbrace{\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(x)}^{\text{a dot product!}}\\
&=f(x)+\nabla f(x)\cdot h.
\end{aligned}$$

Generally $\nabla f(x)$ is the preferred representative of $Df(x)$, not $f'(x)$ This is an example of the Riesz representation theorem.

13 Week 13

13.1 Lecture 36. Mon Dec 2

13.1.1 Operator Norms

Recall that $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x \in \mathbb{R}^n$ iff there exists $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ such that

$$f(x+h) = f(x) + Lh + o(\|h\|), \quad \text{as } \|h\| \to 0.$$

Definition 13.1. If X, Y are vector spaces over $\mathbb{R}$, then $\mathcal{L}(X, Y)$ is the space of all linear maps from X to Y. Also, $\mathcal{L}(X, Y)$ is a vector space over $\mathbb{R}$.

Definition 13.2. The operator norm on $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is defined by

$$\|A\| = \max\{\|Ax\| \mid x \in \mathbb{R}^n, \|x\| = 1\}, \quad \forall A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m).$$

Note that $\{x \in \mathbb{R}^n \mid \|x\| = 1\}$ is compact, and $x \mapsto \|Ax\|$ is continuous, so $\|A\|$ is well defined. It is straightforward to prove that the operator norm is actually a norm.

Theorem 13.3.
1. For all $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$,

$$\|Ax\| \le \|A\|\|x\|, \quad \forall x \in \mathbb{R}^n.$$

2. For all $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$,

$$\|A\| = \inf\{\beta \ge 0 \mid \|Ax\| \le \beta\|x\|, \forall x \in \mathbb{R}^n\}.$$

3. If $B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, $A \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)$, then

$$\|AB\| \le \|A\|\|B\|.$$

Proof. For 1, we have

$$\begin{aligned} \|Ax\| &\le \|A\|, \quad \forall x \in \mathbb{R}^n, \|x\| = 1 \\ \implies \left\| A\frac{x}{\|x\|} \right\| &\le \|A\|, \quad \forall x \in \mathbb{R}^n, x \ne 0 \\ \implies \frac{\|Ax\|}{\|x\|} &\le \|A\|, \quad \forall x \in \mathbb{R}^n, x \ne 0 \\ \implies \|Ax\| &\le \|A\|\|x\|, \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

For 2, write $S = \{\beta \ge 0 \mid \|Ax\| \le \beta\|x\|, \forall x \in \mathbb{R}^n\}$. By 1, $\|A\| \in S$, and so $\inf S \le \|A\|$. Also,

$$\begin{aligned} \beta \in S &\implies \|Ax\| \le \beta, \quad \forall x \in \mathbb{R}^n, \|x\| = 1 \\ &\implies \max\{\|Ax\| \mid x \in \mathbb{R}^n, \|x\| = 1\} \le \beta \\ &\implies \|A\| \le \beta. \end{aligned}$$

So $\|A\|$ is a lower bound for S, and $\|A\| \in S$, so $\|A\| = \inf S$.
For 3, let $x \in \mathbb{R}^n$. Then

$$\begin{aligned} \|ABx\| = \|A(Bx)\| &\le \|A\|\|Bx\| && (1) \\ &\le \|A\|\|B\|\|x\| && (1) \\ \implies \|AB\| &\le \|A\|\|B\|. && (2) \end{aligned}$$

□

Theorem 13.4. *Define $\Omega = \{A \in \mathcal{L}(R^n) \mid A \text{ is invertible}\}$. Then Ω is open, in fact, if $A \in \Omega$, then $B_r(A) \subset \Omega$ for $r = \frac{1}{\|A^{-1}\|}$.*

Proof. Recall that an operator in $\mathcal{L}(\mathbb{R}^n)$ is invertible iff it is nonsingular ($Ax = 0 \iff x = 0$). Let $B \in \mathcal{L}(R^n)$, such that

$$\|B - A\| < \frac{1}{\|A^{-1}\|}.$$

Suppose that $x \in \mathbb{R}^n, x \neq 0$. Then

$$\begin{aligned}
Bx &= Ax + (Bx - Ax) = Ax + (B - A)x \\
\implies \|Bx\| &\geq \|Ax\| - \|(B - A)x\| \\
&\geq \|Ax\| - \|B - A\|\|x\| \\
&> \|Ax\| - \frac{\|x\|}{\|A^{-1}\|} \\
&= \|Ax\| - \frac{\|A^{-1}Ax\|}{\|A^{-1}\|} \\
&> \|Ax\| - \frac{\|A^{-1}\|\|Ax\|}{\|A^{-1}\|} = 0.
\end{aligned}$$

□

Corollary 13.4.1. *The map $f : \Omega \to \Omega$ defined by $f(A) = A^{-1}$ is continuous and invertible, with $f^{-1} = f$.*

Proof. Since $(A^{-1})^{-1} = A$, f is invertible, and $f^{-1} = f$. Let $A \in \Omega$. We will show that f is continuous at A. If $B \in \mathcal{L}(\mathbb{R}^n)$ satisfies $\|B - A\| \leq \frac{1}{2\|A^{-1}\|}$, then $B \in \Omega$. (If A is invertible, $A \neq 0$.) Then

$$\begin{aligned}
B^{-1} - A^{-1} &= B^{-1}AA^{-1} - B^{-1}BA^{-1} \\
&= B^{-1}(A - B)A^{-1} \\
\implies \|B^{-1} - A^{-1}\| &= \|B^{-1}(A - B)A^{-1}\| \\
&\leq \|B^{-1}\|\|A - B\|\|A^{-1}\| \\
&\leq \frac{\|B^{-1}\|\|A^{-1}\|}{2\|A^{-1}\|} \\
&= \frac{\|B^{-1}\|}{2} \\
\implies \|B^{-1}\| &\leq \|A^{-1}\| + \frac{\|B^{-1}\|}{2} \\
\implies \frac{\|B^{-1}\|}{2} &\leq \|A^{-1}\| \\
\implies \|B^{-1}\| &\leq 2\|A^{-1}\|.
\end{aligned}$$

Now let $\varepsilon > 0$ be given, and define

$$\delta = \max\left\{\frac{1}{2\|A^{-1}\|}, \frac{\varepsilon}{2\|A^{-1}\|^2}\right\}.$$

Then, if $\|B-A\| < \delta$, we have

$$\begin{aligned} \|B^{-1} - A^{-1}\| &\le \|B^{-1}\|\|A-B\|\|A^{-1}\| \\ &< 2\|A^{-1}\|\frac{\varepsilon}{2\|A^{-1}\|^2}\|A^{-1}\| = \varepsilon. \end{aligned}$$

Then f is continuous at A. □

13.1.2 Chain Rule for Functions of Several Variables

Theorem 13.5 (Chain Rule). *Suppose $E \subset \mathbb{R}^n$ is open, and $F \subset \mathbb{R}^m$ is open. Then $f : E \to \mathbb{R}^m$, and $\mathcal{R}(f) \subset F$, and $g : F \to \mathbb{R}^k$. If f is differentiable at $x \in E$, and g is differentiable at $f(x) \in F$, then $h = g \circ f$ is differentiable at x, and*

$$Dh(x) = Dg(f(x))Df(x).$$

Note. Be wary that composing linear maps is not generally commutative, so the order of the derivative operators matters.

Proof. We have

$$\begin{aligned} h(x+p) - h(x) &= g(f(x+p)) - g(f(x)) \\ &= g(f(x) + Df(x)p + o(\|p\|)) - g(f(x)) \\ &= g(f(x)) + Dg(f(x))(Df(x)p + o(\|p\|)) + o(\|Df(x)p + o(\|p\|)\|) - gf(x) \\ &= Dg(f(x))Df(x)p + Dg(f(x))o(\|p\|) + o(\|Df\|(x)p + o(\|p\|)). \end{aligned}$$

If we can show that

$$Dg(f(x))o(\|p\|) + o(\|Df(x)p + o(\|o\|)\|) = o(\|p\|),$$

then h is differentiable at x, and $Dh(x) = Dg(f(x))Df(x)$. First,

$$\|Dg(f(x))o(\|p\|)\| \le \|Dg(f(x))\|\|o(\|p\|)\| = o(\|p\|),$$

where the last equality holds since $\|Dg(f(x))\|$ is a constant. Secondly,

$$\begin{aligned} \|Df(x)p + o(\|p\|)\| &\le \|Df(x)p\| + \|o(\|p\|)\| \\ &\le \|Df(x)\|\|p\| + \|o(\|p\|)\| \\ &= \|Df(x)\|\|p\| + \frac{\|o(\|p\|)\|}{\|p\|}\|p\| \\ &= \left(\|Df(x)\| + \frac{\|o(\|p\|)\|}{\|p\|}\right)\|p\| \\ &\le (\|Df(x)\| + 1)\|p\|, \quad \forall p \text{ sufficiently small.} \end{aligned}$$

Thus,

$$\begin{aligned} \frac{o(\|Df(x)p + o(\|p\|)\|)}{\|Df(x)p + o(\|o\|)\|} \to 0 &\implies \frac{o(\|Df(x)p + o(\|p\|)\|)}{(\|Df(x)\| + 1)\|p\|} \to 0 \\ &\implies \frac{o(\|Df(x)p + o(\|p\|)\|)}{\|p\|} \to 0 \\ &\implies o(\|Df(x)p + o(\|p\|)\|) = o(\|p\|). \end{aligned}$$

Therefore,

$$h(x+p) - h(x) = Dg(f(x))Df(x)p + o(\|p\|).$$

□

13.2 Lecture 37. Wed Dec 4

13.2.1 Contractive Mappings

Definition 13.6. Let (X, d) be a metric space, and let $\varphi : X \to X$. We say that φ is a contractive mapping (contraction) iff there exists $\lambda \in (0, 1)$ such that

$$\forall x, y \in X, d(\varphi(x), \varphi(y)) \leq \lambda d(x, y).$$

Clearly φ is uniformly continuous.

Theorem 13.7 (Contractive Mapping Theorem)**.** *Let (X, d) be a complete metric space. Let $\varphi : X \to X$ be a contractive mapping. Then there exists a unique $x \in X$ such that $\varphi(x) = x$. Such an x is called a fixed point of x.*

Proof. Suppose $d(\varphi(x), \varphi(y)) \leq \lambda d(x, y), \forall x, y \in X$, where $\lambda \in (0, 1)$. Let $x_0 \in X$ be arbitrary, and define $\{x_n\} \subset X$ by

$$x_{n+1} = \varphi(x_n), n = 0, 1, \ldots.$$

(This process is called fixed point iteration.) Note that

$$\begin{aligned}
d(x_2, x_1) &= d(\varphi(x_2), \varphi(x_1)) \leq \lambda d(x_1, x_0), \\
d(x_3, x_2) &= d(\varphi(x_3), \varphi(x_2)) \leq \lambda d(x_2, x_1) \leq \lambda^2 d(x_1, x_0), \\
d(x_4, x_3) &= d(\varphi(x_4), \varphi(x_3)) \leq \lambda d(x_3, x_2) \leq \lambda^2 d(x_2, x_1) \leq \lambda^3 d(x_1, x_0), \\
&\vdots \\
d(x_{n+1}, x_n) &\leq \lambda^n d(x_1, x_0).
\end{aligned}$$

Now let $m \geq n$. Then

$$\begin{aligned}
d(x_m, x_n) &\leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \ldots + d(x_{n+1}, x_n) \\
&\leq \lambda^{m-1} d(x_1, x_0) + \lambda^{m-2} d(x_1, x_0) + \ldots + \lambda^n d(x_1, x_0) \\
&= (\lambda^{m-1} + \lambda^{m-2} + \ldots + \lambda^n) d(x_1, x_0) \\
&< (\lambda^n + \lambda^{n+1} + \ldots) d(x_1, x_0) = \lambda^n \sum_{j=0}^{\infty} \lambda^j d(x_1, x_0) \\
&= \frac{\lambda^n}{1 - \lambda} d(x_1, x_0) \to 0, \quad \text{as } n \to \infty.
\end{aligned}$$

Thus $\{x_n\}$ is Cauchy, and hence $x_n \to x$, for some $x \in X$, since X is complete. Since φ is continuous,

$$\begin{aligned}
\varphi(x) &= \lim_{n\to\infty} \varphi(x_n) = \lim_{n\to\infty} x_{n+1} = x, \\
\implies \varphi(x) &= x,
\end{aligned}$$

hence, x is a fixed point. Suppose also that $\varphi(x') = x'$. Then

$$d(x', x) = d(\varphi(x'), \varphi(x)) \leq \lambda d(x', x),$$

which is possible only if $d(x', x) = 0$, i.e. $x' = x$. $\square$

13.2.2 Implicit Function Theorem

Suppose $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$. Then suppose

$$f(x, y) = 0 \iff \begin{cases} f_1(x, y_1, \ldots, y_n) = 0 \\ f_2(x, y_1, \ldots, y_n) = 0 \\ \quad\quad\vdots \\ f_n(x, y_1, \ldots, y_n) = 0. \end{cases}$$

For each fixed x, this represents n nonlinear equations in n unknowns. Hopefully, for each x, we can solve for $y = \psi(x)$ such that

$$f(x, \psi(x)) = 0.$$

I.e. for each x, solve for y and define $\psi : \mathbb{R}^n \to \mathbb{R}^n$ by $\psi(x) = y$. Then we are implicitly defining y as a function of x, under certain assumptions.

Theorem 13.8 (Implicit Function Theorem)**.** *Suppose that $E \subset \mathbb{R}^m \times \mathbb{R}^n$ is open, and $f : E \to \mathbb{R}^n$ is C^1, (i.e. f is differentiable on E, and Df is continuous on E). Suppose a point $(x_0, y_0) \in E$ satisfies*

$$f(x_0, y_0) = 0.$$

Suppose also that $D_y(x_0, y_0)$ is nonsingular (invertible). Then there exists open subsets U of $\mathbb{R}^m$ and V of $\mathbb{R}^n$ such that $(x_0, y_0) \in U \times V$ and $\psi : U \to V$ such that

$$f(x, \psi(x)) = 0, \quad \forall x \in U.$$

Moreover, for all $x \in U$, $y = \psi(x)$ is the only solution of $f(x, y) = 0$ that lies in V. Finally, ψ is C^1.

We'll prove this in our final lectures. You might have noticed that we used D_y, a partial derivative without defining partial derivatives.

13.2.3 Partial Derivatives

Consider that $Df(x, y) \in \mathcal{L}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^n)$. Consider

$$f(x + u, y + v) = f(x, y) + Df(x, y)(u, v) + o(\|u, v\|).$$

Write $L = Df(x, y)$. Then $(u, v) = (u, 0) + (0, v)$, and

$$\begin{aligned} f(x + u, y) &= f(x, y) + L(u, , 0) + o(\|u, v\|) \\ \implies f(x + u, y) &= f(x, y) + L(u, 0) + o(\|u\|). \end{aligned}$$

We can define $L_x \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ by

$$L_x u = L(u, 0), \quad \forall x \in \mathbb{R}^m.$$

Then

$$f(x + u, y) = f(x, y) + L_x(u) + o(\|u\|).$$

Thus, $f(\cdot, y)$ is differentiable at x, and the derivative is L_x. So if f is differentiable at (x, y), then $f(\cdot, y)$ is differentiable at x, and we call the derivative the partial derivative of f at x, written

$$D_x f(x, y), \text{ where } D_x f(x, y)u = Df(x, y)(u, 0), \quad \forall u \in \mathbb{R}^m.$$

Similarly, if f is differentiable at (x, y), then $f(x, \cdot)$ is differentiable at y, and its derivative is the partial derivative of f w.r.t. y:

$$D_y f(x, y), \text{ where } D_y f(x, y)v = Df(x, y)(0, v), \quad \forall v \in \mathbb{R}^n.$$

Theorem 13.9. *If $Df(x, y)$ exists, then $D_x f(x, y)$ and $D_y f(x, y)$ exists.*

The converse is not always true. Intuitively, if we wanted to show that f is differentiable, we would have

$$f(x + u, y + v) - f(x, y) = \underbrace{f(x + u, y + v) - f(x + u, y)}_{?} + \underbrace{f(x + u, y) - f(x, y)}_{D_x f(x,y)u + o(\|u\|)}$$

We don't know that $f(x + u, \cdot)$ is differentiable. From here, we garner some insight into how to construct a counterexample.

13.3 Lecture 38.

Let us continue with the proof of the implicit function theorem.

Theorem 13.10. *Let E be an open subset of $\mathbb{R}^m \times \mathbb{R}^n$. Let $f : E \to \mathbb{R}^n$ be differentiable on E, and assume that Df is continuous on E. Suppose that $(x_0, y_0) \in E$ satisfies $f(x_0, y_0) = 0$, and $D_y f(x_0, y_0)$ is nonsingular, or invertible. Then there exists open subsets $U \in \mathbb{R}^m$, $V \in \mathbb{R}^n$ such that $x_0 \in U, y_0 \in V$, $U \times V \subset E$, and $\psi : U \to V$ such that $f(x, \psi(x)) = 0$, for all $x \in U$. Furthermore, for all $x \in U$, $y = \psi(x)$ is the only point in V satisfying $f(x, y) = 0$. Finally, ψ is continuously differentiable, and $\psi'(x) = D_y f(x, \psi(x))^{-1} D_x f(x, \psi(x))$.*

Proof. We will prove existence and uniqueness by setting up a contractive mapping. Define $\varphi : B_\varepsilon(x_0) \times B_\delta(y_0) \to B_\delta(y_0)$ by

$$\varphi(x, y) = y - D_y f(x_0, y_0)^{-1} f(x, y),$$

where $\varepsilon > 0$, and $\delta > 0$ are to be specified. Note that

$$\begin{aligned}
\varphi(x, y_1) - \varphi(x, y_2) &= y_1 - D_y f(x_0, y_0)^{-1} f(x, y_1) - y_2 + D_y f(x_y, y_0)^{-1} f(x, y_2) \\
&= y_1 - y_2 - D_y f(x_0, y_0)^{-1}(f(x, y_1) - f(x, y_2)).
\end{aligned}$$

Now,

$$\begin{aligned}
f(x, y_1) - f(x, y_2) &= \int_0^1 D_y f(x, y_2 + t(y_1 - y_2))(y_1 - y_2)\, dt \\
&= \left(\int_0^1 D_y f(x, y_2 + t(y_1 - y_2)\, dt) \right) (y_1 - y_2)
\end{aligned}$$

$$\implies D_y f(x_0, y_0)^{-1}(f(x, y_1) - f(x, y_2)) = \left(\int_0^1 D_y(x_0, y_0)^{-1} D_y f(x, y_2 + t(y_1 - y_2))\, dt \right) (y_1 - y_2)$$

$$\begin{aligned}
&\implies y_1 - y_2 - D_y f(x_0, y_0)^{-1}(f(x, y_1) - f(x, y_2)) \\
&\quad = y_1 - y_2 - \left(\int_0^1 D_y f(x_0, y_0)^{-1} D_y f(x, y_2 + t(y_1 - y_2)\, dt \right) (y_1 - y_2).
\end{aligned}$$

Now choose $\varepsilon' > 0, \delta' > 0$ such that $B_{\varepsilon'}(x_0) \times B_{\delta'}(y_0) \subset E$, and define $A : B_{\varepsilon'}(x_0) \times B_{\delta'}(y_0) \times B_{\delta'}(y_0) \to \mathcal{L}(\mathbb{R}^n)$ by

$$A(x, y_1, y_2) = y_1 - y_2 - \int_0^1 D_y f(x_0, y_0)^{-1} D_y f(x, y_2 + t(y_1 - y_2))\, dt.$$

□

Since the mapping $L \mapsto L^{-1}$ is continuous, we can choose ε', δ' sufficiently small such that $D_y f(x, y)$ is invertible for all $x, y \in B_{\varepsilon'}(x_0) \times B_{\delta'}(y_0)$. Note that A is continuous (why?), and $A(x_0, y_0, y_0) = 0$. Therefore, there exists $\varepsilon \in (0, \varepsilon']$, and $\delta \in (0, \delta']$ such that

$$(x, y_1, y_2) \in B_\varepsilon(x_0) \times B_\delta(y_1) \times B_\delta(y_2) \implies \|A(x, y_1, y_2)\| \le \frac{1}{2}.$$

We have

$$\begin{aligned}
\varphi(x, y) - y_0 &= y - y_0 - D_y f(x_0, y_0)^{-1} f(x, y) \\
&= y - y_0 - D_y f(x_0, y_0)^{-1} f(x, y) - f(x, y_0) + f(x, y_0) - f(x_0, y_0)) \\
&= A(x, y, y_0)(y - y_0) + \left(\int_0^1 D_y f(x_0, y_0)^{-1} D_x f(x_0 + t(x - x_0), y_0)\, dt \right) (x - x_0) \\
\|\varphi(x, y) - y_0\| &\le \|A(x, y, y_0)\| \|y - y_0\| + \left\| \int_0^1 D_y f(x_0, y_0)^{-1} D_x f(x_0 + t(x - x_0), y_0)\, dt \right\| \|x - x_0\| \\
&\le \frac{1}{2}\delta + \left\| \int_0^1 D_y f(x_0, y_0)^{-1} D_x f(x_0 + t(x - x_0), y_0)\, dt \right\| \varepsilon.
\end{aligned}$$

Since $D_x f(x, y_0)$ depends continuously on x, we can reduce ε, if necessary, to ensure that the second term is less than $\frac{1}{2}\delta$ for all $x \in B_\varepsilon(x_0)$, and hence that

$$\|\varphi(x, y_0) - y_0\| < \delta, \quad \forall x \in B_\varepsilon(x_0).$$

Then for all $x \in B_\varepsilon(x_0), \varphi(x_j)$ maps into $B_\delta(y_0)$ into $B_\delta(y_0)$. Next,

$$\begin{aligned}
&(x, y_1, y_2) \in B_\varepsilon(x_0) \times B_\delta(y_1) \times B_\delta(y_2) \\
\implies \|\varphi(x, y_1) - \varphi(x, y_2)\| &= \|A(x, y_1, y_2)(y_1 - y_2)\| \\
&\le \|A(x, y_1, y_2)\| \|y_1 - y_2\| \\
&\le \lambda \|y_1 - y_2\|,
\end{aligned}$$

thus $\varphi(x, \cdot)$ is a contractive mapping. Therefore, for each $x \in B_\varphi(\varphi(x_0))$, there exists a unique $y \in B_\delta(y_0)$ such that

$$\begin{aligned}
\varphi(x, y) &= y \\
\iff y - D_y f(x_0, y_0)^{-1} f(x, y) &= y \\
\iff -D_y f(x_0, y_0)^{-1} f(x, y) &= 0 \\
\iff f(x, y) &= 0.
\end{aligned}$$

Define $U = B_\varepsilon(x_0), V = B_\delta(y_0)$, and $\psi : U \to V$ by the condition that $\psi(x) = y$, where y is the unique point in V such that $f(x, y) = 0$. This proves the existence and uniqueness of ψ. We will save the rest of the proof for the next lecture.

Example 13.11 (Analysis Preliminary, University of Delaware)**.** Let $u((x,y),v(x,y))$ be the unique simultaneous solution of the equations

$$\begin{cases} xu^3 + (y+1)uv = 6 \\ yu^2 + v^2 + xy = 9, \end{cases}$$

for (x,y) near $(0,0)$, and (u,v) near $(2,3)$. Compute u_x, u_y, v_x, and v_y at the point $(x,y) = (0,0)$. Clearly state every theorem that you use.

Solution. Define $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ by

$$f((x,y),(u,v)) = \begin{bmatrix} xu^3 + (y+1)uv - 6 \\ yu^2 + v^2 + xy - 9 \end{bmatrix}.$$

□

Note that

$$\begin{aligned} f((0,0),(2,3)) &= \begin{bmatrix} 0 + 1 \cdot 6 - 6 \\ 0 + 9 + 0 - 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \\ f'_{(u,v)}((x,y),(u,v)) &= \begin{bmatrix} 3xu^2 + (y+1)v & (y+1)u \\ 2yu + x & 2v \end{bmatrix}, \\ f'_{(u,v)}((0,0),(2,3)) &= \begin{bmatrix} 0 + 1 \cdot 3 & 1 \cdot 2 \\ 0 + 0 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 0 & 6 \end{bmatrix}. \end{aligned}$$

Since $f((0,0),(2,3)) = 0$, and $f'_{(u,v)}((0,0),(2,3))$ is nonsingular, the implicit function theorem applies. There exist open sets U, V in $\mathbb{R}^2$ such that $(0,0) \in U$, $(2,3) \in V$, and $\psi : U \to V$ such that

$$f((x,y),\psi(x,y)) = 0, \quad \forall (x,y) \in U$$

and $(u,v) = \psi(x,y)$ is the unique solution of

$$f((x,y),(u,v)) = 0,$$

that lies in V. Also,

$$\begin{aligned} &\psi'(x,y) = -f'_{(u,v)}((x,y),\psi(x,y))^{-1} f'_{(x,y)}((x,y),\psi(x,y)) \\ \implies &\psi'(0,0) = -f'_{(u,v)}((0,0),(2,3))^{-1} f'_{(x,y)}((0,0),(2,3)). \end{aligned}$$

We computed $f'_{(u,v)}((0,0),(2,3))$ above. We have

$$\begin{aligned} f((x,y),(u,v)) &= \begin{bmatrix} xu^3 + (y+1)uv - 6 \\ yu^2 + v^2 + xy - 9 \end{bmatrix} \\ \implies f_{(x,y)}((x,y),(u,v)) &= \begin{bmatrix} u^3 & uv \\ y & u^2 + x \end{bmatrix} \\ \implies f'_{(x,y)}((0,0),(2,3)) &= \begin{bmatrix} 8 & 6 \\ 0 & 4 \end{bmatrix}. \end{aligned}$$

Thus,

$$\psi'(0,0) = -\begin{bmatrix} 3 & 2 \\ 0 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 8 & 6 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} -8/3 & -14/9 \\ 0 & -2/3 \end{bmatrix}.$$

Since

$$\psi'(0,0) = \begin{bmatrix} u_x(0,0) & u_y(0,0) \\ v_x(0,0) & v_y(0,0) \end{bmatrix},$$

we see that

$$\begin{aligned} u_x(0,0) = -8/3, \quad u_y(0,0) = -14/9 \\ v_x(0,0) = 0, \quad v_y(0,0) = -2/3. \end{aligned}$$

14 Week 14

14.1 Lecture 39.

Final bits of the proof for Theorem 13.8. We have thus far shown there exists open sets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^n$ and $\psi : U \to V$ such that

$$x_0 \in U, y_0 \in V, U \times V \subset E, f(x, \psi(x)) = 0, \forall x \in U,$$

and that $y = \psi(x)$ is the only solution of $f(x,y) = 0$ that lies in V. We must show that ψ is continuously differentiable and that $\psi'(x) = -D_y f(x, \psi(x))^{-1} D_x f(x, \psi(x))$ for all $x \in U$. Note that we can assume taht $\overline{U} \times \overline{V} \subset E$ (by reducing ε' and δ' if necessary earlier in the proof) and hence $\|D_x f(x,y)\|$ is uniformly bounded for $(x,y) \in U \times V$. This will be needed below. We begin by showing that ψ is Lipschitz continuous on U. That is, there exists $C > 0$ such that

$$\|\psi(x_1) - \psi(x_2)\| \leq C\|x_1 - x_2\|, \forall x_1, x_2 \in U.$$

Let $x_1, x_2 \in U$. Then

$$\begin{aligned} \psi(x_1) - \psi(x_2) &= \varphi(x_1, \psi(x_1)) - \varphi(x_2, \psi(x_2)) \\ &= \varphi(x_1, \psi(x_1)) - \varphi(x_2, \psi(x_1)) + \varphi(x_2, \psi(x_1)) - \varphi(x_2, \psi(x_2)) \\ \implies \|\psi(x_1) - \psi(x_2)\| &\leq \|\varphi(x_1, \psi(x_1)) - \varphi(x_2, \psi(x_1))\| + \|\varphi(x_2, \psi(x_1)) - \varphi(x_2, \psi(x_2))\|. \end{aligned}$$

Now, for any $y \in V$, (including $y = \psi(x_1)$),

$$\begin{aligned} \varphi(x_1, y) - \varphi(x_2, y) &= y - D_y f(x_0, y_0)^{-1} f(x_1, y) - y + D_y f(x_0, y_0)^{-1} f(x_2, y) \\ &= D_y f(x_0, y_0)^{-1}(f(x_2, y) - f(x_1, y)) \\ &= D_y f(x_0, y_0)^{-1} \int_0^1 D_x f(x_1 + t(x_2 - x_1), y)(x_2 - x_1)\, dt \\ &= \left(\int_0^1 D_y f(x_0, y_0)^{-1} D_x f(x_1, t(x_2 - x_1), y)\, dt \right)(x_2 - x_1) \\ \implies \|\varphi(x_1, y) - \varphi(x_2, y)\| &\leq \left\| \int_0^1 D_y f(x_0, y_0)^{-1} D_x f(x_1 + t(x_2 - x_1), y)\, dt \right\| \|x_2 - x_1\| \\ &\leq \int_0^1 \|D_y f(x_0, y_0)^{-1}\| \|D_x f(x_1 + t(x_2 - x_1), y)\|\, dt \|x_2 - x_1\| \\ &\leq C'\|x_2 - x_1\|. \end{aligned}$$

(Note that we have applied the triangle inequality for integrals, which we haven't formally proven before. Also, as noted before, $\|D_x f(x,y)\|$ is uniformly bounded for $(x,y) \in U \times V$.) We previously proved that $\varphi(x, \cdot)$ is a contraction and hence

$$\|\varphi(x_2, \psi(x_1)) - \varphi(x_1, \psi(x_2))\| \leq \lambda\|\psi(x_1) - \psi(x_2)\|, \quad 0 < \lambda < 1.$$

Therefore,

$$
\begin{aligned}
\|\psi(x_1)-\psi(x_2)\| &\le \|\varphi(x_1,\psi(x_1))-\varphi(x_2,\psi(x_1))\| + \|\varphi(x_2,\psi(x_1))-\varphi(x_2,\psi(x_2))\| \\
&\le C'\|x_1-x_2\| + \lambda\|\psi(x_1)-\psi(x_2)\| \\
\implies (1-\lambda)\|\psi(x_1)-\psi(x_2)\| &\le C'\|x_1-x_2\| \\
\implies \|\psi(x_1)-\psi(x_2)\| &\le \frac{C'}{1-\lambda}\|x_1-x_2\| \\
&= C\|x_1-x_1\|, \quad C=\frac{C'}{1-\lambda}.
\end{aligned}
$$

Now, we prove that ψ is differentiable at $x \in U$:

$$
\begin{aligned}
&f(x+p,\psi(x+p))=0, \quad \forall p\in\mathbb{R}^n \text{ sufficiently small} \\
&\implies f(x,\psi(x)) + Df(x,\psi(x))(p,\psi(x+p)-\psi(x)) + o(\|(p,\psi(x+p)-\psi(x))\|)=0.
\end{aligned}
$$

Note that

$$
\begin{aligned}
\|(p,\psi(x+p)-\psi(x))\| &= \sqrt{\|p\|^2+\|\psi(x+p)-\psi(x)\|^2} \\
&\le \sqrt{\|p\|^2+C^2\|p^2\|} && \text{(since } \psi \text{ is Lipschitz)} \\
&= \sqrt{1+c^2}\|p\|.
\end{aligned}
$$

Hence

$$
o(\|p,\psi(x+p)-\psi(p)\|) = o(\|p\|).
$$

Also,

$$
\begin{aligned}
f(x,\psi(x)) &= 0, \\
D_f(x,\psi(x))(p,\psi(x+p)-\psi(x)) &= D_x f(x,\psi(x))p + D_y f(x,\psi(x))(\psi(x+p)-\psi(x)).
\end{aligned}
$$

Thus we obtain

$$
\begin{aligned}
&D_x f(x,\psi(x))pD_y f(x,\psi(x))(\psi(x+p)-\psi(x)) + o(\|p\|) = 0 \\
&\implies D_y f(x,\psi(x))(\psi(x+p)-\psi(x)) = -D_x f(x,\psi(x))p + o(\|p\|) \\
&\implies \psi(x+p) = \psi(x) - D_y f(x,\psi(x))^{-1} D_x f(x,\psi(x))p + D_y f(x,\psi(x))^{-1} o(\|p\|) \\
&\qquad\qquad\quad = \psi(x) - D_y f(x,\psi(x))^{-1} D_x(f(x),\psi(x))p + o(\|p\|).
\end{aligned}
$$

This proves ψ is differentiable at x, and that

$$
D\psi(x)p = -D_y f(x,\psi(x))^{-1} D_x f(x,\psi(x))p.
$$

14.1.1 Inverse Function Theorem

Theorem 14.1. *Suppose that $f : E \to \mathbb{R}^n$, where $E \subset \mathbb{R}^n$ is open, f is differentiable at $x_0 \in E$, and $Df(x_0)$ is invertible. Let us define $F : \mathbb{R}^n \times E \to \mathbb{R}^n$ by*

$$
F(y,x) = f(x) - y.
$$

Then if $y_0 = f(x_0)$, we have

$$
\begin{aligned}
F(y_0,x_0) &= 0, \\
D_x F(y_0,x_0) &\text{ is nonsingular}.
\end{aligned}
$$

Hence, the implicit function applies, and there exists open sets $U, V \subset \mathbb{R}^n$ and $\psi : U \to V$ such that

$$y_0 \in U, x_0 \in V, V \subset E, F(y, \psi(y)) = 0, \quad \forall y \in U,$$

and $x_0 = \psi(y)$ is the unique solution of $F(y, x) = 0$ that lies in V. Since

$$F(y, \psi(y)) = 0 \iff y - f(\psi(y)) = 0 \implies f(\psi(y)) = y,$$

we have

$$f(\psi(y)) = y, \quad \forall y \in U.$$

Note that f maps $\psi(U)$ onto U, and $f|_{\psi(U)}$ is injective. Thus $f|_{\psi(U)}$ is invertible. Moreover,

$$\psi(U) = (f|_{\psi(U)})^{-1}(U),$$

which shows that $\psi(U)$ is open. We can then redefine $V := \psi(U)$, and thus $\psi : U \to V$ is the inverse of $f|_V$. The implicit function theorem guarantees that ψ is C^1. We can compute $D\psi$ by implicit differentiation:

$$\begin{aligned} f(\psi(y)) &= y \\ \implies Df(\psi(y))D\psi(y) &= I \\ \implies D\psi(y) &= Df(\psi(y))^{-1}. \end{aligned}$$

Example 14.2. Define $f : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$f(x) = \begin{bmatrix} x_1x_2 + x_2^2 \\ x_1^2 + x_1x_2 \end{bmatrix}.$$

Then

$$f'(x) = \begin{bmatrix} x_2 & x_1 + 2x_2 \\ 2x_1 + x_2 & x_1 \end{bmatrix}.$$

If $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, then

$$f'(x_0) = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix},$$

which is nonsingular. What is $(f^{-1})'(y_0)$, where

$$y_0 = f(x_0) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}?$$

Answer. We have

$$(f^{-1})'(y_0) = f'(x_0)^{-1} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1/8 & 3/8 \\ 3/7 & -1/8 \end{bmatrix}.$$

Note that we should write $(f|_U)^{-1}$, not f^{-1}. There's no reason to think that f itself is invertible. The theorem also gives no indication of how big the set U is. □

□

15 Preliminary Review

This section contains review on topics covered in Calculus III, to be examined in the Analysis prelim. These are just some notes I ripped from `https://math.berkeley.edu/~arash/53/notes`. After writing this, I must apologize. I hadn't realized how much material is actually covered in vector calculus. The notes I have taken are a semester's work of material at Berkeley.

15.1 Vector Calculus

Definition 15.1. Let D be a subset of $\mathbb{R}^2$. A vector field on $\mathbb{R}^2$ is a function $\vec{F}$ that assigns to each point (x, y) in D a two-dimensional vector $\vec{F}(x, y)$. We may write $\vec{F}$ in terms of component functions P, Q:

$$\vec{F}(x, y) = \vec{F}(\vec{x}) = P(x, y)\hat{\imath} + Q(x, y)\hat{\jmath} = \langle P, Q\rangle.$$

Here, P, Q are scalar functions of x, y, and are called scalar fields.

Definition 15.2. Let E be a subset of $\mathbb{R}^3$. A vector field on $\mathbb{R}^3$ is a function that assigns to each point (x, y, z) in E a three-dimensional vector $\vec{F}(x, y, z)$.

Example 15.3. By Newton's Law of Gravitation, we may represent the gravitational force field by

$$\vec{F}(\vec{x}) = -\frac{mMG}{\|\vec{x}\|^3}\vec{x}.$$

The force exerted by an electric charge at Q on a charge q located at (x, y, z), with position vector $\vec{x} = \langle x, y, z\rangle$ is

$$\vec{F}(\vec{x}) = \frac{\varepsilon q Q}{\|x\|^3}\vec{x}.$$

The electric field of this charge Q on $\mathbb{R}^3$ is

$$\vec{E}(\vec{x}) = \frac{1}{q}\vec{F}(\vec{x}) = \frac{\varepsilon Q}{\|x\|^3}\vec{x}.$$

Example 15.4. The gradient of a function $f(x, y)$ is $\nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle$. Thus ∇f is a vector field on $\mathbb{R}^2$, called a *gradient vector field.*

Example 15.5. A vector field $\vec{F}$ is called a *conservative vector field* if it is the gradient of some scalar function. That is, if there exists a function f such that $\vec{F} = \nabla f$. Then f is called the potential function of $\langle F\rangle$.

Example 15.6. The vector field

$$\vec{F}(\vec{x}) = -\frac{mMG}{\|x\|^3}\vec{x}$$

is conservative, with potential function

$$f(x, y, z) = \frac{mMG}{\sqrt{x^2 + y^2 + z^2}}.$$

15.1.1 Line Integrals

Definition 15.7. Suppose C is a curve in $\mathbb{R}^2$ such that for the position (x, y), each x, y is a function of a parameter t, with $a < t < b$. In other words, suppose we had a function $\vec{r}(t) = \langle x(t), y(t)\rangle$. Suppose C is smooth: $\vec{r'}$ is continuous, and its derivative is nowhere 0. Divide C into n small segments or sub-arcs, each with length Δs_i. Define the line integral of f along C as

$$\int_C f(x,y)\,ds = \lim_{n\to\infty}\sum_{i=1}^{n} f(x_i^*, y_i^*)\Delta s_i,$$

if the limit exists. Recall that the length of the curve C is

$$L = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt$$

Then the line integral is

$$\int_C f(x,y)\,ds = \int_a^b f(x(t), y(t))\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt.$$

The value of the line integral does not depend on the parameterization of the curve, as long as we traverse the curve exactly once as t increases from a to b. Thus our original single integral over an interval becomes a special case, where we evaluate a line integral from $(a, 0)$ to $(b, 0)$, with $x = t$, $y = 0$:

$$\int_C f(x,y)\,dx = \int_a^b f(x,0)\,dx.$$

Just as we may interpret the single integral as the area under a nonnegative curve, we may interpret the line integral of a nonnegative curve as the area under the "curtain" under the curve.

Example 15.8. Evaluate the line integral

$$\int_C \frac{x}{y}\,ds, \quad C : x = t^3, y = t^4, 1 \le t \le 2.$$

Solution. First,

$$\frac{dx}{dt} = 3t^2, \quad \frac{dy}{dt} = 4t^3.$$

So

$$ds = \sqrt{9t^4 + 16t^6}\,dt = t^2\sqrt{9 + 16t^2}\,dt,$$

and

$$\int_C \frac{x}{y}\,ds = \int_1^2 t\sqrt{9+16t^2}\,dt.$$

Let $u = 9 + 16t^2 \implies dt = \frac{1}{32t}\,dt$, and

$$\begin{aligned}\int_1^2 t\sqrt{9+16t^2}\,dt &= \frac{1}{32}\int_{10}^{13}\sqrt{u}\,du \\ &= \frac{1}{32}\left[\frac{2}{3}u^{\frac{3}{2}}\right]_{10}^{13}.\end{aligned}$$

□

Definition 15.9. If C is a piecewise smooth curve, that is, if C is the union of a finite number of smooth curves $C_1, \ldots, C_n$, where each C_i is a smooth curve, and the initial point of C_{i+1} is the terminal point of C_i, then we define the line integral of f along C as the sum of the integrals of f along each of the smooth pieces of C:

$$\int_C f(x,y)\,ds = \sum_{i=1}^{n} \int_{C_i} f(x,y)\,ds.$$

Proposition 15.10. *We may express a line integral in terms of x or in terms of y:*

$$\int_C f(x,y)\,dx = \int_a^b f(x(t),y(t))x'(t)\,dt \text{ is the line integral with respect to } x,$$

$$\int_C f(x,y)\,dy = \int_a^b f(x(t),y(t))y'(t)\,dt \text{ is the line integral with respect to } y.$$

Hence, our original line integral is a line integral with respect to arc length. It is customary to write the line integrals with respect to x, y together:

$$\int_C P(x,y)\,dx + \int_C Q(x,y)\,dy = \int_C P(x,y)\,dx + Q(x,y)\,dy.$$

We may have to set up the parametric equations so that we would start at a point on C and end at another point on C. Recall the parametric equation for a line segment that starts at $P_0(x_0,y_0)$ and ends at $P_1(x_1,y_1)$:

$$\vec{r}(t) = \langle x_0, y_0\rangle + t\langle x_1 - x_0, y_1 - y_0\rangle, \quad 0 \le t \le 1.$$

In general, the value of the line integral depends on the path as well as on the endpoints of the curve. Also, for line integrals with respect to x, or y, if we switch the directions, the sign of the integral reverses:

$$\int_{-C} f(x,y)\,dx = -\int_C f(x,y)\,dx, \quad \int_{-C} f(x,y)\,dy = -\int_C f(x,y)\,dy.$$

This is because Δx_i and Δy_i change sign when we reverse the orientation of C. However, since Δs_i is always positive, the sign of the line integral with respect to arc length does not change when we reverse the traversal of the path:

$$\int_{-C} f(x,y)\,ds = \int_C f(x,y)\,ds.$$

Every concept we've discussed thus far generalizes to $\mathbb{R}^3$,

$$\begin{aligned}\int_C f(x,y,z)\,ds &= \int_a^b f(x(t),y(t),z(t))\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\,dt \\ &= \int_a^b f(\vec{r}(t))|\vec{r}(t)|\,dt,\end{aligned}$$

where $a \le t \le b$. When $f(x,y,z) = 1$, we would get the arc length:

$$\int_C ds = \int_a^b |\vec{r}(t)|\,dt = L.$$

15.1.2 Line Integrals of Vector Fields

Recall that the work done by a variable force $f(x)$ in moving an object from a to b along the x-axis is

$$W = \int_a^b f(x)\,dx.$$

The work done by a constant force $\vec{F}$ in moving a particle from a point P to another point Q in space is

$$W = \vec{F} \cdot \vec{PQ},$$

where $\vec{PQ}$ is the displacement vector. If $\vec{T}(t_i)$ is the unit tangent vector at point P_i, then the work done by the force field $\vec{F}$ is the limit of the Riemann sum

$$\sum_{i=1}^{n} \Big(\vec{F}(x_i, y_i, z_i) \cdot \vec{T}(x_i, y_i, z_i)\Big)\,\Delta s_i,$$

that is

$$W = \int_C \vec{F}(x, y, z) \cdot \vec{T}(x, y, z)\,ds = \int_C \vec{F} \cdot \vec{T}\,ds.$$

Definition 15.11. Let $\vec{F}$ be a continuous vector field defined over a smooth curve C given by vector function $\vec{r}(t)$, $a \le t \le b$. Then the line integral of $\vec{F}$ along C is

$$\int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}(t)\,dt = \int_C \vec{F} \cdot \vec{T}\,ds,$$

where $\vec{T} = \vec{r'}(t)/|\vec{r'}(t)|$ is the unit tangent vector at the point (x, y, z). Therefore, $d\vec{r} = \vec{r'}(t)\,dt$.

Suppose $\vec{F} = \langle P, Q, R\rangle$. Using the above definition, we have

$$\begin{aligned}
\int_C \vec{F} \cdot d\vec{r} &= \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r'}(t)\,dt \\
&= \int_a^b \langle P, Q, R\rangle \cdot \langle x'(t), y'(t), z'(t)\rangle\,dt \\
&= \int_a^b \left(Px' + Qy' + Rz'\right)\,dt.
\end{aligned}$$

The last integral is precisely the line integral. Therefore,

$$\int_C \vec{F} \cdot d\vec{r} = \int_C P\,dx + Q\,dy + R\,dz.$$

Example 15.12. Calculate $\int_C \vec{F} \cdot d\vec{r}$, where C is given by the vector function $\vec{r}(t)$, and

$$\vec{F}(x, y, z) = \langle x, y, xy\rangle, \quad \vec{r}(t) = \cos t\hat{i} + \sin t\hat{j} + t\hat{k}, \quad 0 \le t \le \pi.$$

Solution. We find

$$\vec{r'}(t) = \langle -\sin t, \cos t, 1\rangle.$$

Using the definition above,

$$\begin{aligned}
\int_C \vec{F}\, d\vec{r} &= \int_0^\pi \vec{F}(\vec{r}(t)) \cdot \vec{r}'(t)\, dt \\
&= \int_0^\pi \langle \cos t, \sin t, \cos t \sin t\rangle \cdot \langle -\sin t, \cos t, 1\rangle\, dt \\
&= \int_0^\pi -\cos t \sin t + \cos t \sin t + \cos t \sin t\, dt \\
&= \int_0^\pi \cos t \sin t\, dt && (u\text{-sub}) \\
&= \Big[\sin t\Big]_0^\pi = 0.
\end{aligned}$$

□

15.1.3 Fundamental Theorem of Line Integrals

The following theorem says that we can evaluate the line integral of a conservative vector field, that is, the gradient field of a potential function f, simply by knowing the value of f at the endpoints of C. Therefore, the line integral of a conservative vector field is independent of the path C. It only depends on the initial point and the terminal point of C.

Theorem 15.13. *Let C be a smooth curve given by the vector function $\vec{r}(t)$, $a \le t \le b$. Let f be a differentiable function of two or three variables whose gradient vector ∇f is continuous on C. Then*

$$\int_C \nabla f \cdot d\vec{r} = f(\vec{r}(b)) - f(\vec{r}(a)).$$

Proof. We have

$$\begin{aligned}
\int_C \nabla f \cdot d\vec{r} &= \int_a^b \nabla f(\vec{r}(t)) \cdot \vec{r}'(t)\, dt \\
&= \int_a^b \left(\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}\right) dt \\
&= \int_a^b \frac{d}{dt} f(\vec{r}(t))\, dt \\
&= f(\vec{r}(b)) - f(\vec{r}(a)).
\end{aligned}$$

□

Definition 15.14. A curve C is called closed if its terminal point and initial point are the same, that is, $\vec{r}(a) = \vec{r}(b)$.

If $\int_C \vec{F} \cdot d\vec{r}$ is independent of path, then the line integral of a closed path is zero.

Theorem 15.15. *We claim $\int_C \vec{F} \cdot d\vec{r}$ is independent of path iff $\int_C \vec{F} \cdot d\vec{r} = 0$ for every closed path C.*

Proof. Suppose that $\int_C \vec{F} \cdot d\vec{r}$ is independent of path. Let C be an arbitrary closed path containing distinct points A, B. Let C_1, C_2 be the two distinct paths from A to B along C. Then $C = C_1 + (-C_2)$. Thus

$$\int_C \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} + \int_{-C_2} \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} - \int_{C_2} \vec{F} \cdot d\vec{r} = 0,$$

where the last equality holds by assumption. For the converse implication, assume that $\int_C \vec{F} \cdot d\vec{r} = 0$ for every closed path C. Let C_1, C_2 be arbitrary distinct paths from point A to point B. Then $C = C_1 + (-C-2)$ is a closed path, and

$$0 = \int_C \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} + \int_{-C_2} \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot d\vec{r} \; - \int_{C_2} \vec{F} \cdot d\vec{r}.$$

□

Theorem 15.16. *Suppose that $\vec{F}$ is a vector field that is continuous on an open connected region D. If $\int_C \vec{F} \cdot d\vec{r}$ is independent of path in D, then $\vec{F}$ is a conservative vector field on D. That is, there exists a function f such that $\nabla f = \vec{F}$.*

Now the question is to determine whether or not a vector field $\vec{F}$ is conservative? Suppose that we know $\vec{F} = \langle P, Q \rangle$ is conservative, where P, Q have continuous first order partial derivatives. Then we know that there exists a function f such that $\vec{F} = \nabla f$, and

$$P = \frac{\partial f}{\partial x}, \quad Q = \frac{\partial f}{\partial y}.$$

Then, by Clairaut's theorem,

$$\frac{\partial P}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial Q}{\partial x}.$$

Theorem 15.17. *If $\vec{F}(x,y) = \langle P(x,y), Q(x,y) \rangle$ is a conservative vector field, where P, Q have continuous first-order partial derivatives on a domain D, then throughout D,*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

Definition 15.18. A simple curve is a curve which does not intersect itself.

Definition 15.19. A simply connected region is a connected region D such that every simple closed curve in D encloses only points that are in D. Thus, a simply connected region contains no holes and cannot consist of two separate pieces. *The converse of the above theorem is true for those regions and curves!*

Theorem 15.20. *Let $\vec{F} = P\hat{\imath} + Q\hat{\jmath}$ be a vector field on an open simply connected region D. Suppose that P and Q have continuous first-order partial derivatives and*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \textit{throughout } D.$$

Then $\vec{F}$ is conservative.

We can use "partial integration" to find a potential function f.

Example 15.21. Determine whether or not $\vec{F}$ is a conservative vector field. If it is, find a function f such that $\vec{F} = \nabla f$:

$$\vec{F}(x,y) = ye^x\hat{\imath} + (e^x + e^y)\hat{\jmath}.$$

Proof. Since $\vec{F}$ is defined for all points in $\mathbb{R}^2$, the above theorem applies. We can note that

$$\frac{\partial P}{\partial y} = e^x,$$

and

$$\frac{\partial Q}{\partial x} = e^x.$$

Thus, $\vec{F}$ is conservative. To find the potential function of $\vec{F}$, we integrate ye^x with respect to x:

$$\int ye^x \, dx = ye^x + C(y),$$

where we let $C(y)$ be a constant of integration with respect to y. Then taking the derivative of what we've found with respect to y reveals:

$$\frac{\partial}{\partial y} \left(ye^x + C(y) \right) = e^x + C'(y).$$

This should be equal to $Q(x, y)$, so

$$e^x + C'(y) = e^x + e^y \implies C'(y) = e^y.$$

This implies that $C(y) = e^y + C_0$, where C_0 is an arbitrary constant of integration. Then we've found the family of potential functions for $\vec{F}$:

$$\nabla \left(ye^x + e^y + C_0 \right) = \vec{F}.$$

□

15.1.4 Green's Theorem

To review, let D be a simply connected region, and let C be a simple curve on D. We have

$$\int_C \vec{F} \cdot d\vec{r} = 0, \forall \text{ closed paths } C \iff \int_C \vec{F} \cdot d\vec{r} \text{ independent of path } \iff \vec{F} \text{ conservative } (\vec{F} = \nabla f).$$

For Green's theorem, we use a positively oriented curve.

Definition 15.22. A positive orientation on a simple closed curve C refers to a single counter-clockwise traversal of C. Thus, the region D enclosed by C is always on the left as the point $\vec{r}(t)$ traverses C. Negative orientation is the reverse direction.

Theorem 15.23 (Green's Theorem)**.** *Let C be a positively oriented, piecewise-smooth, simple closed curve in the plane, and let D be the region bounded by C. If P, Q have continuous partial derivatives on an open region that contains D, then*

$$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \int_C P \, dx + Q \, dy.$$

In a sense, Green's theorem generalizes the fundamental theorem of calculus:

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

In both cases, we have an integral of derivatives on the left hand side, and on the right hand side we have the original functions on the boundary of the domain. These are other notations for the line integral:

$$\oint_C P\,dx + Q\,dy = \int_{\partial D} P\,dx + Q\,dy.$$

(There is one more symbol of integral, denoting the orientation of traversal, but that LaTeX package isn't working for me.) We may use Green's theorem to compute the area of D, which is $\iint_D 1\,dA$. Thus, we would like to have

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1.$$

There are many useful combinations of choices for P, Q. For example:

$$P = 0, Q = x \implies \oint_C x\,dy,$$
$$P = -y, Q = 0 \implies -\oint_C y\,dx,$$
$$P = \frac{-1}{2}y, Q = \frac{1}{2}x \implies \frac{1}{2}\oint_C x\,dy - y\,dx.$$

Example 15.24. Use Green's theorem to evaluate the line integral

$$\int_C y^4\,dx + 2xy^3\,dy,$$

along the positively oriented curve C, where C is the circle $x^2 + y^2 = 1$.

15.1.5 Divergence, Curl

We will now study two operations on vector fields: curl and divergence.

Definition 15.25. Define the "del" operator $\nabla : D^1(\mathbb{R}^3) \to \mathcal{F}(\mathbb{R}^3)$ by

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right), \quad \forall f \in D^1(\mathbb{R}^3).$$

Definition 15.26. Define the "curl" operator $\mathrm{curl} : D^1(\mathbb{R}^3) \to \mathcal{F}(\mathbb{R}^3)$ by

$$\mathrm{curl}\,\vec{F} = \nabla \times \vec{F} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ P & Q & R \end{vmatrix}.$$

(An abuse of notation sadly, but literature finds this consistent.)

Theorem 15.27. *Let f be a function of three variables that has continuous second-order partial derivatives. Then*

$$\mathrm{curl}(\nabla f) = \vec{0}.$$

Proof. Expand

$$\begin{aligned}
\operatorname{curl}(\nabla f) &= \nabla \times (\nabla f) \\
&= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial f/\partial x & \partial f/\partial y & \partial f/\partial z \end{vmatrix} \\
&= \left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y}\right)\hat{i} + \left(\frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z}\right)\hat{j} + \left(\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial z}\right)\hat{k} \\
&= \vec{0}\hat{i} + \vec{0}\hat{j} + \vec{0}\hat{k} \\
&= \vec{0},
\end{aligned}$$

where we used Clairaut's theorem to equate the second-order partial derivatives. □

Corollary 15.27.1. *If $\vec{F}$ is conservative, then $\operatorname{curl}\vec{F} = \vec{0}$.*

The converse of this holds iff the domain of $\vec{F}$ is simply-connected, that is, it has no holes.

Theorem 15.28. *If $\vec{F}$ is a vector field on a simply-connected domain, whose component functions have continuous partial derivatives, and $\operatorname{curl}\vec{F} = \vec{0}$, then $\vec{F}$ is a conservative vector field.*

The reason why this operator is called *curl*, is that if a particle has $\operatorname{curl}\vec{F} = \vec{0}$, then, in a fluid, that particle will not rotate. If the opposite is true, then the particle will rotate, like in a whirlpool, or an eddy.

Definition 15.29. If $\vec{F} = \langle P, Q, R\rangle$ is a vector field defined on $\mathbb{R}^3$, and $\partial P/\partial x, \partial Q/\partial y$, and $\partial R/\partial z$ exist, then the divergence of $\vec{F}$ is defined to be

$$\operatorname{div}\vec{F} = \nabla \cdot \vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

Theorem 15.30. *If $\vec{F} = \langle P, Q, R\rangle$ is a vector field on $\mathbb{R}^3$ and P, Q, R have continuous second-order partial derivatives. Then*

$$\operatorname{div}\operatorname{curl}\vec{F} = 0.$$

Proof. Expanding, we find

$$\begin{aligned}
\operatorname{div}\operatorname{curl}\vec{F} &= \nabla \cdot (\nabla \times \vec{F}) \\
&= \frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \\
&= 0.
\end{aligned}$$

□

Consider the following:

$$\operatorname{div}(\nabla f) = \nabla(\nabla f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

This expression occurs so often that there is a special notation for it: $\nabla^2 f$. The operator is called the Laplace operator, or Laplacian, because of its relation to Laplace's equation.

Definition 15.31. A function g is called harmonic on D if it satisfies Laplace's equation, that is, $\nabla^2 g = 0$ on D.

15.1.6 Vector Form of Green's Theorem

Suppose $\vec{F} = \langle P, Q, 0\rangle$. Then

$$\int_C \vec{F}\, d\vec{r} = \int_C P\, dx + Q\, dy,$$

and

$$\operatorname{curl} \vec{F} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ P(x,y) & Q(x,y) & 0 \end{vmatrix} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \hat{k}.$$

Thus,

$$\operatorname{curl} \vec{F} \cdot \hat{k} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

Therefore, Green's theorem in vector form is

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_D (\operatorname{curl} \vec{F}) \cdot \hat{k}\, dA.$$

If C is given by

$$\vec{r}(t) = \langle x(t), y(t)\rangle, \quad a \le t \le b,$$

then the unit tangent vector is

$$\vec{T}(t) = \frac{1}{\|\vec{r}(t)\|} \langle x'(t), y'(t)\rangle.$$

The outward unit normal vector is

$$\vec{n}(t) = \frac{1}{\|r'(t)\|} \langle y'(t), -x'(t)\rangle.$$

Then

$$\begin{aligned} \oint_C \vec{F} \cdot \vec{n}\, ds &= \int_a^b (\vec{F} \cdot \vec{n}) \|\vec{r}(t)\|\, dt \\ &= \int_a^b (Py' - Qx')\, dt \\ &= \int_C P dy - Q dx \\ &= \iint_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) dA, \end{aligned}$$

by Green's theorem. Therefore, another form of Green's theorem is

$$\oint_C \vec{F} \cdot \vec{n}\, ds = \iint_D \operatorname{div} \vec{F}(x,y)\, dA,$$

which means the line integral of the normal component of $\vec{F}$ along C is equal to the double integral of the divergence of $\vec{F}$ over the region D enclosed by C.

15.2 Surface Integrals, Flux

15.2.1 Parametric Surfaces

We have seen that we may describe a curve with parametric equations:

$$\vec{r}(t) = \langle x(t), y(t), z(t)\rangle.$$

We can do something similar with surfaces in space with parametric equations:

$$\vec{r}(t) = \langle x(u,v), y(u,v), z(u,v)\rangle.$$

A parametric surface S is the set of all points defined as above, with (u,v) varying throughout a domain D.

Example 15.32. The parametric equations of a sphere $x^2+y^2+z^2=a^2$ are

$$x = a\sin\phi\cos\theta, \quad y = a\sin\phi\sin\theta, \quad z = a\cos\phi,$$

with parameters $0 \le \phi \le \phi$, and $0 \le \theta \le 2\pi$. The grid lines are the longitudes when θ is constant, and when ϕ is constant, we would get latitudes.

15.2.2 Surface Area

By taking the limit of a Riemann sum, we can define the area as follows:

Definition 15.33. If a smooth parametric surface S is given by the equation

$$\vec{r}(u,v) = \langle x(u,v), y(u,v), z(u,v)\rangle, \quad (u,v) \in D,$$

and S is covered just once as (u,v) ranges throughout the parameter domain D, then the surface area of S is

$$A(S) = \iint_D \|\vec{r}_u \times \vec{r}_v\| dA,$$

where

$$\vec{r}_u = \left\langle \frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right\rangle,$$

and

$$\vec{r}_u = \left\langle \frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right\rangle.$$

Example 15.34. Find the surface area of a sphere with radius a.

When $z = f(x,y)$, we may select the parameters to be x, y. Thus,

$$x = x, y = y, z = f(x,y),$$

and

$$\vec{r}_x = \left\langle 1, 0, \frac{\partial f}{\partial x}\right\rangle,$$

and

$$\vec{r}_y = \left\langle 0, 1, \frac{\partial f}{\partial y}\right\rangle.$$

Therefore,

$$\vec{r}_x \times \vec{r}_y = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 0 & \partial f/\partial x \\ 0 & 1 & \partial f/\partial y \end{vmatrix} = -\frac{\partial f}{\partial x}\hat{i} - \frac{\partial f}{\partial y}\hat{j} + \hat{k}.$$

Then

$$\begin{aligned} \|\vec{r}_x \times \vec{r}_y\| &= \sqrt{(\partial f/\partial x)^2 + (\partial f/\partial y)^2 + 1} \\ &= \sqrt{(\partial z/\partial x)^2 + (\partial z/\partial y)^2 + 1} \end{aligned}$$

The surface area definition becomes

$$A(S) = \iint_D \sqrt{(\partial z/\partial x)^2 + (\partial z/\partial y)^2 + 1}\, dA$$

Definition 15.35. The tangent plane to a surface at a point where $(u, v) = (u_0, v_0)$ contains the tangent vectors

$$\vec{r}_u = \langle x_u(u_0, v_0), y_u(u_0, v_0), z_u(u_0, v_0)\rangle, \quad \vec{r}_v = \langle x_v(u_0, v_0), y_v(u_0, v_0), z_v(u_0, v_0)\rangle.$$

If $\vec{r}_u \times \vec{r}_v \neq 0$, then the surface is called smooth, and the normal to the tangent plane is this cross product.

15.2.3 Surface Integrals

Suppose we have a function $f(x, y, z)$ over a surface S given by

$$\vec{r}(u, v) = \langle x(u, v), y(u, v), z(u, v)\rangle, \quad (u, v) \in D.$$

We divide the region D into small subregions, each with dimension $\Delta u \cdot \Delta v$. Then the surface S is also divided into corresponding patches S_{ij}. We evaluate the function f at some point P_{ij}^* and multiply the function value by the area of the patch ΔS_{ij}. Adding all of these products would give us the Riemann sum

$$\sum_{i=1}^{m}\sum_{j=1}^{n} f(P_{ij}^*)\Delta S_{ij}.$$

The limit of this Riemann sum, as the number of patches increases without bound is the surface integral of f over the surface S:

$$\iint_S f(x, y, z)\, dS = \lim_{m,n\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n} f(P_{ij}^*)\Delta S_{ij}.$$

Also,

$$\iint_S f(x, y, z)\, dS = \iint_D f(\vec{r}(u, v))\|\vec{r}_u \times \vec{r}_v\|\, dA.$$

We saw that we may integrate a function over a curve, called a line integral. If the function is 1, then the line integral gives us the length of the curve. Similarly, the surface integral is the integral over a surface, and if the function is 1, the surface integral gives us the area of the surface, as we saw in the previous chapter.

Proposition 15.36. *When $z = g(x, y)$, we can regard x, y as parameters, and then the surface integral becomes*

$$\iint_S f(x,y,z)\,dS = \iint_D f(x,y,g(x,y))\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\,dA.$$

Likewise, if S is a surface with equation $y = h(x, z)$, then we would regard x, z as the parameters.

Example 15.37. Evaluate

$$\iint_S y^2 z^2\,dS,$$

where S is the part of the cone $y = \sqrt{x^2 + z^2}$, given by $0 \leq y \leq 5$.

Solution. Let us begin by parametrizing S in cylindrical coordinates. Let $x = r\cos\theta, z = r\sin\theta$. Then $y = r$, since $x^2 + z^2 = r^2$. Now,

$$\begin{aligned}
\vec{r}(\theta, r) &= \langle r\cos\theta, r, r\sin\theta\rangle, \\
\vec{r}(\theta, r)_r &= \langle \cos\theta, 1, \sin\theta\rangle, \\
\vec{r}(\theta, r)_\theta &= \langle -r\sin\theta, 0, r\cos\theta\rangle,
\end{aligned}$$

and

$$\vec{r}(\theta, r)_r \times \vec{r}(\theta, r)_\theta = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \cos\theta & 1 & \sin\theta \\ -r\sin\theta & 0 & r\cos\theta \end{vmatrix} = \langle r\cos\theta, -r, r\sin\theta\rangle,$$

so

$$\|\vec{r}(\theta, r)_r \times \vec{r}(\theta, r)_\theta\| = r\sqrt{2}.$$

Altogether,

$$\begin{aligned}
\iint_S y^2 z^2\,dS &= \sqrt{2}\int_0^5 \int_0^{2\pi} r^5 \sin^2\theta\,d\theta\,dr \\
&= \sqrt{2}\int_0^5 r^5 \int_0^{2\pi} \frac{1 - \cos 2\theta}{2}\,d\theta\,dr \\
&= -\frac{\sqrt{2}}{4}\int_0^5 r^5 \Big[\sin 2\theta - 2\theta\Big]_0^{2\pi}\,dr \\
&= \pi\sqrt{2}\int_0^5 r^5\,dr \\
&= \frac{15625\pi\sqrt{2}}{6}.
\end{aligned}$$

□

Definition 15.38. Suppose a surface S has a tangent plane at every point on S. There are two possible choices for a unit normal vector at each point, call it $\vec{n_1}$ and $\vec{n_2} = -\vec{n_1}$. If it is possible to choose a unit normal vector $\vec{n}$ at every point (x, y, z), so that $\vec{n}$ varies continuously over S, we say that S is an oriented surface, and the given choice of $\vec{n}$ provides S with an orientation. There are two orientations for any orientable surface. For a closed surface, that is, a surface that is the boundary of a solid region E, the convention is that the positive orientation is the one for which the normal vectors face outward from E. The inward facing normals give the negative orientation.

15.2.4 Surface Integrals of Vector Fields

Suppose S is an oriented surface with unit normal vector $\vec{n}$. Suppose S is porous, like a fishing net across a stream, and the stream flows through S with density $\rho(x, y, z)$ and velocity field $\vec{v}(x, y, z)$. The rate of flow, mass per unit time per unit area, is $\rho\vec{v}$. If we divide S into small patches, the mass of the stream per unit time crossing a small patch S_{ij} in the direction of $\vec{n}$ is approximately

$$(\rho\vec{v} \cdot \vec{n})A(S_{ij}),$$

where $\rho, \vec{v}, \vec{n}$ are evaluated at some point on S_{ij}. We may add all of these quantities and obtain the following integral as the result:

$$\iint_S \rho\vec{v} \cdot \vec{n}\, d\vec{S}.$$

The above integral is the rate of flow through S. If $\vec{F} = \rho\vec{v}$, then the integral becomes

$$\iint_S \vec{F} \cdot \vec{n}\, dS.$$

Definition 15.39. If $\vec{F}$ is a continuous vector field on an oriented surface S with unit normal vector $\vec{n}$, then the surface integral of $\vec{F}$ over S is

$$\iint_S \vec{F} \cdot\, d\vec{S} = \iint_D \vec{F} \cdot \vec{n}\, dS.$$

This integral is also called the flux of $\vec{F}$ across S.

The above formula means the surface integral of a vector field over S is equal to the surface integral of its normal component over S. If S is given by $\vec{r}(u, v)$, then

$$\vec{n} = \frac{\vec{r}_u \times \vec{r}_v}{\|\vec{r}_u \times \vec{r}_v\|},$$

and

$$\begin{aligned}\iint_S \vec{F} \cdot d\vec{S} &= \iint_S \vec{F} \cdot \frac{\vec{r}_u \times \vec{r}_v}{\|\vec{r}_u \times \vec{r}_v\|}\, dS \\ &= \iint_D \left(\vec{F} \cdot \frac{\vec{r}_u \times \vec{r}_v}{\|\vec{r}_u \times \vec{r}_v\|}\right) \|\vec{r}_u \times \vec{r}_v\|\, dA.\end{aligned}$$

Therefore,

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_D \vec{F} \cdot (\vec{r}_u \times \vec{r}_v)\ dA.$$

If we have $z = g(x, y)$, then we may take x, y as parameters, and

$$\vec{F} \cdot (\vec{r}_u \times \vec{r}_v) = \langle P, Q, R\rangle \cdot \left\langle -\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1\right\rangle.$$

Thus, the formula for flux becomes

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_D \left(-P\frac{\partial g}{\partial x} - Q\frac{\partial g}{\partial y} + R\right)\ dA.$$

The above formula is for when S is upward-oriented. For a downward orientation, we multiply by -1. Similar formulas apply when $y = h(x, z)$ or $x = k(y, z)$.

Example 15.40. Evaluate the surface integral $\iint_S \vec{F} \cdot d\vec{S}$ for the vector field

$$\vec{F}(x, y, z) = x\hat{i} + y\hat{j} + 5\hat{k},$$

and the oriented surface S, where S is the boundary of the region enclosed by the cylinder $x^2+z^2 = 1$, and the planes $y = 0$, $x + y = 2$.

Solution. Use cylindrical coordinates to parametrize S: the surface area of the "cylindrical wrapper" of the given surface. Let $x = \cos\theta$, $z = \sin\theta$, $y = y$. Now,

$$\begin{aligned}\vec{r}(\theta, y) &= \langle \cos\theta, y, \sin\theta \rangle,\\ \vec{r}(\theta, y)_\theta &= \langle -\sin\theta, 0, \cos\theta \rangle,\\ \vec{r}(\theta, y)_y &= \langle 0, 1, 0 \rangle.\end{aligned}$$

Then

$$\begin{aligned}\vec{r}(\theta, y)_\theta \times \vec{r}(\theta, y)_y &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ -\sin\theta & 0 & \cos\theta \\ 0 & 1 & 0 \end{vmatrix}\\ &= \langle -\cos\theta, 0, -\sin\theta \rangle.\end{aligned}$$

Note we have calculated the inward facing normal, so we negate this, and find

$$\begin{aligned}\iint_S \vec{F} \cdot d\vec{S} &= \iint_D \vec{F} \cdot (\vec{r}(\theta, y)_y \times \vec{r}(\theta, y)_\theta)\, dA\\ &= \iint_D \langle \cos\theta, y, 5 \rangle \cdot \langle \cos\theta, 0, \sin\theta \rangle\, dA\\ &= \int_{\theta=0}^{\theta=2\pi} \int_{y=0}^{y=2-\cos\theta} \cos^2\theta + 5\sin\theta\, dA\\ &= \int_{\theta=0}^{\theta=2\pi} (2 - \cos\theta)(\cos^2\theta + 5\sin\theta)\, d\theta\\ &= 2\pi.\end{aligned}$$

The last equality can be found by using half-angle identities, and u-substitutions. Honestly, it's kinda tedious to do this without Divergence theorem, so I'm going to leave it like this. □

Flux is not just for fluids. If $\vec{E}$ is an electric field, then the surface integral $\iint_S \vec{E} \cdot d\vec{S}$ is the electric flux of $\vec{E}$. Gauss's law says that the net charge enclosed by a closed surface S is

$$Q = \varepsilon_0 \iint_S \vec{E} \cdot d\vec{S},$$

where ε_0 is a constant, called the permitivity of free space.

15.2.5 Stoke's Theorem

Green's theorem relates a double integral over a plane region D to a line integral around the boundary C of D. Stoke's theorem relates a surface integral over a surface S to a line integral around the boundary curve of S. Note the difference, S is a space curve in the case of Stoke's theorem, and a plane curve in the case of Green's theorem.

Theorem 15.41. *Let S be an oriented piecewise-smooth surface that is bounded by a simple, closed, piecewise-smooth boundary curve C with positive orientation. Let $\vec{F}$ be a vector field whose components have continuous partial derivatives on an open region in $\mathbb{R}^3$ that contains S. Then*

$$\int_C \vec{F} \cdot d\vec{r} = \iint_S \operatorname{curl} \vec{F} \cdot d\vec{S}.$$

We showed earlier that Green's theorem is a special case of Stoke's theorem, when S is a flat region in the plane. In this case, $\vec{n} = \hat{k}$, and Stoke's theorem becomes

$$\int_C \vec{F} \cdot d\vec{r} = \iint_S \operatorname{curl} \vec{F} \cdot d\vec{S} = \iint \operatorname{curl} \vec{F} \cdot \hat{k}\, dA.$$

Note. Let $\vec{F} = \langle P, Q \rangle$. Furthermore, suppose we traverse the curve C, for $a \leq t \leq b$. Then

$$\begin{aligned}
\int_C \vec{F} \cdot d\vec{r} &= \int_a^b \left(\vec{F} \cdot \frac{d\vec{r}}{dt} \right) dt \\
&= \int_a^b \left(\vec{F} \cdot \frac{ds}{dt} \vec{T} \right) dt \\
&= \int_C \left(\vec{F} \cdot \vec{T} \right) ds,
\end{aligned}$$

where $\vec{T}$ is the unit tangent vector at each point. In other words, we can regard our original definition of the line integral as a special case of the line integral of a scalar field, the scalar in this case being the tangential component of the vector field. Then we may write two versions of Green's theorem:

$$\int_C \left(\vec{F} \cdot \vec{T} \right) ds = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA,$$

interpreted as the work done by a force $\vec{F}$ around C, and

$$\int_C (\vec{F} \cdot \vec{n})\, ds = \iint_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right), dA,$$

the rate of fluid flow past C if $\vec{F}$ is the flux. On the other hand, in the first version, $\vec{F}$ might be the flux of a fluid as well, and we would interpret the integral of its tangential component as the circulation around C. We would then say there is a whirlpool inside C if the integral is nonzero. This is the reason why the operator is named "curl." So the line integral of the tangential component of $\vec{F}$ along curve C is the double integral of the vertical component of $\operatorname{curl} \vec{F}$ over the region D enclosed by C. The line integral of the normal component of $\vec{F}$ is

$$\int_C (\vec{F} \cdot \vec{n})\, ds = \iint_D (\nabla \cdot \vec{F})\, dA = \iint_D \operatorname{div} \vec{F}\, dA = \iint_D (\operatorname{curl} \vec{F}) \cdot \hat{k}\, dA.$$

From a mathematical standpoint, none of this matters. Although of course, it is important in applications. Both versions of Green's theorem are correct for any vector field.

15.2.6 Divergence Theorem

We have reached the end of what is covered usually in vector calculus. Calculus I builds its way up to the Fundamental theorem of Calculus:

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

Green's theorem tells us

$$\int_C \vec{F} \cdot \vec{n}\, ds = \iint_D \operatorname{div} \vec{F}(x, y)\, dA,$$

when C is a positively oriented boundary curve of the plane region D. We are now prepared to generalize this idea of relating integrals and derivatives to the case of solids volumes and surface areas.

Theorem 15.42 (Divergence theorem)**.** *Let E be a solid region in $\mathbb{R}^3$. Let S be the boundary surface of E, given with positive (outward) orientation. Let $\vec{F}$ be a vector field whose component functions have continuous partial derivatives on an open region that contains E. Then*

$$\iint_S \vec{F} \cdot d\vec{S} = \iiint_E \operatorname{div} \vec{F}(x, y, z)\, dV.$$

In words, the Divergence theorem states that the flux of $\vec{F}$ across the boundary surface of E is the same as the triple integral of the divergence of $\vec{F}$ over E. If you have been reading thus far, congratulations, this is the end of the material which is expected to be covered on the University of Delaware Analysis preliminary. For being such a dedicated reader, here is the generalized Stokes theorem.

Theorem 15.43 (Stokes-Cartan)**.** *Let ω be a smooth $(n-1)$-form with compact support on an oriented n-dimensional manifold, with boundary Ω, where $\partial\Omega$ is the induced orientation. Then*

$$\int_\Omega d\omega = \int_{\partial\Omega} \omega.$$

Here, d is the exterior derivative.

This theorem is one of the rewards you receive upon finishing a differential geometry or differential topology course. For example, https://www.ekzhang.com/assets/pdf/Math_132_Notes.pdf.

Example 15.44. Let

$$\vec{F}(x, y, z) = \langle z \tan^{-1}(y^2), z^3 \ln(x^2 + 1), z \rangle.$$

Find the flux of $\vec{F}$ across the part of the paraboloid $x^2 + y^2 + z = 2$ that lies above the plane $z = 1$, and is oriented upward.

Solution. We know

$$\iint_S \vec{F}(x, y, z) = \iiint_E \operatorname{div} \vec{F}(x, y, z)\, dA,$$

by the Divergence theorem. Let us parametrize S. Let $x = r\cos\theta, y = \sin\theta$. Then z ranges from $z = 1$ to

$$z = 2 - x^2 - y^2 = 2 - r^2.$$

Also, $r^2 + z = 2 \implies r = 1$, when $z = 1$, so r ranges from $r = 0$ to $r = 1$. Let us calculate $\operatorname{div} \vec{F}(x, y, z)$:

$$\begin{aligned}\operatorname{div} \vec{F}(x, y, z) &= \frac{\partial}{\partial x}\left(z \tan^{-1}(y^2)\right) + \frac{\partial}{\partial y}\left(z^3 \ln(x^2 + 1)\right) + \frac{\partial}{\partial z}(z) \\ &= 0 + 0 + 1 = 1.\end{aligned}$$

So we have a nice integral to compute:

$$\begin{aligned}
\iiint_E \operatorname{div} \vec{F}(x,y,z)\, dA &= \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=1} \int_{z=1}^{z=2-r^2} r\, dz\, dr\, d\theta \\
&= \int_{\theta=0}^{\theta=2\pi} \int_{r=0}^{r=1} r - r^3\, dr\, d\theta \\
&= \int_{\theta=0}^{\theta=2\pi} \frac{1}{2} - \frac{1}{4}\, d\theta \\
&= \pi - \frac{1}{2}\pi \\
&= \frac{\pi}{2}.
\end{aligned}$$

□