— Outline —
Math 4317

I. Review of Set Theory

A. Sets

1. Basic definitions

i. We gave a naive definition of **sets**.

ii. Discussed elements of sets, subsets, the null-set, defining sets via properties and **Russel's paradox**.

2. Operations on sets

Discussed **intersection** and **union** of sets, the **complement** of a set and relations between these operations; in particular, **DeMorgan's Law**.

3. Products

Defined the **product** of two sets.

B. Functions

1. Basic definitions

i. Gave an informal definition of **functions** involving a "rule" that takes an element of one set and specifies an element in another set. Defined **range** and **domain**.

ii. Gave a formal definition of **functions** involving the **graph** of the function in the product of the range and domain space.

2. Properties of functions

i. A function $f : A \to B$ is **injective** if $f(x) = f(y)$ implies $x = y$.

ii. a function $f : A \to B$ is **surjective** if for every $z \in B$ there is an $x \in A$ such that $f(x) = z$.

iii. a functions is **bijective** or a **one-to-one correspond** if it is injective and surjective.

3. Composition and inverses

i. Defined the **composition** of two functions: $f : A \to B$ and $g : B \to C$ then $g \circ f : A \to C$ is the function that takes $x \in A$ and sends it to $g \circ f(x) = g(f(x))$ in C.

ii. A function $f : A \to B$ is **invertible** if there is a function $g : B \to A$ such that $g \circ f(x) = x$ for all $x \in A$ and $f \circ g(y) = y$ for all $y \in B$. The function g, if it exists, is called the **inverse** of f and denoted f^{-1}.

4. Direct and indirect images

i. Let $f : A \to B$ be a function and $C \subset A$. The **direct image of** C is the set $f(C) = \{z \in B : \text{ such that } z = f(x) \text{ for some } x \in A\}$.

ii. We proved

- if $C \subset D$ then $f(C) \subset f(D)$,
- $f(C \cup D) = f(C) \cup f(D)$,
- $f(C \cap D) \subset f(C) \cap f(D)$, and
- $f(C - D) \subset f(C)$.

iii. Let $f : A \to B$ be a function and $C \subset B$. The **inverse image** of C is the set $f^{-1}(C) = \{x \in A : f(x) \in C\}$. (Despite the bad, but standard, notation the inverse image is always defined even if f is not invertible.)

iv. We proved

- if $C \subset D$ then $f^{-1}(C) \subset f^{-1}(D)$,
- $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$,
- $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$, and
- $f^{-1}(C - D) = f^{-1}(C) - f^{-1}(D)$.

C. The natural numbers and cardinality

1. Cardinality: the size of sets

- **i.** Two sets S and T have the same **cardinality**, denoted $|S| = |T|$, if there is a bijection $f : S \to T$.
- **ii.** The relation "having the same cardinality" forms and equivalence relation on sets.
- **iii.** We say the cardinality of S is less than or equal to the cardinality of T, denoted $|S| \leq |T|$, if there is an injection $f : S \to T$.
- **iv.** **Schröder-Bernstein Theorem:** If $|S| \leq |T|$ and $|T| \leq |S|$ then $|S| = |T|$.
- **v.** **Theorem:** If there is a surjection $S \to T$ then $|T| \leq |S|$. (This used the **axiom of choice**.)

2. Finite and infinite sets

- **i.** Let $S_n = \{1, 2, \ldots, n\}$ and $S_0 = \emptyset$. We say a set S has **finite cardinality** if $|S| = |S_n|$ for some natural number n.
- **ii.** A set S is **countably infinite** if there is a bijection between S and the set of natural numbers $\mathbb{N}$. That is if $|S| = |\mathbb{N}|$. A set is **countable** if it is finite of countably infinite.
- **iii.** A set S is **uncountable** if it is not countable.
- **iv.** Discussed **induction** and used induction to show **Theorem:** $|S_n| = |S_m|$ if and only if $n = m$.
- **v.** **Theorem:** The integers $\mathbb{Z}$ and the rational numbers $\mathbb{Q}$ are countably infinite.
- **vi.** **Theorem:** The product of countable sets is countable and the countable union of countable sets is countable.
- **vii.** **Theorem:** The real numbers $\mathbb{R}$ are uncountable.
- **viii.** **Theorem:** The power set $\mathcal{P}(S)$ of a set S always has strictly bigger cardinality than S.
- **ix.** Mentioned the **Continuum Hypothesis.**

II. The real numbers

A. We need the real numbers

1. The **natural numbers** $\mathbb{N}$,

- **i.** Discussed their **algebraic properties**: that is addition and multiplication and the fact that they are commutative and associative operations and that multiplication distributes over addition.
- **ii.** Discussed their **order properties**: that is $\leq$, is a total order, and respects the algebraic properties.
- **iii.** Discussed the fact that they are **well ordered** (that is any non-empty subset of them has a smallest element in the order $\leq$).

2. The **integers** $\mathbb{Z}$.

- **i.** Discussed extending $\mathbb{N}$ to the integers $\mathbb{Z}$.
- **ii.** Discussed that the algebraic properties extend to make $\mathbb{Z}$ a commutative ring,
- **iii.** Discussed that the ordering extends and that $\mathbb{Z}$ is not well ordered, but satisfies the **maximum/minimum property** (that is any non-empty set bounded above has a greatest element and any non-empty set bounded below has a smallest element).

3. The **rational numbers** $\mathbb{Q}$.

- **i.** Discussed extending $\mathbb{Z}$ to the **rational numbers** $\mathbb{Q}$.
- **ii.** Discussed that the algebraic and order properties extend to make $\mathbb{Q}$ a totally ordered field.
- **iii.** The rational numbers $\mathbb{Q}$ do not satisfy the maximum or minimum property.
- **iv.** We also discussed that there is no rational number x such that $x^2 = 2$ and this is related to the fact that there is no "least upper bound" on a set of rational numbers that is bounded above.

B. The real numbers

1. The supremum property.

i. Said the **real numbers** $\mathbb{R}$ are an extension of the rational numbers satisfying the same algebraic and order properties as $\mathbb{Q}$ (that is $\mathbb{R}$ is a totally ordered field) but that $\mathbb{R}$ satisfies the **supremum property**.

ii. The **supremum property** says that a non-empty subset S of $\mathbb{R}$ that is bounded above has a **supremum** (with is also called a **least upper bound**). That is there is some $r \in \mathbb{R}$ such that $s \leq r$ for all $s \in S$ (that is, r is an upper bound on S) and if r' is also an upper bound on S then $r \leq r'$. Such a number r is called a supremum on S.

2. Other properties of the real numbers

i. Used the supremum property to show that $\mathbb{R}$ has the **Archimedean property**. That is, given any $x \in \mathbb{R}$ there is some integer n such that $x < n$.

ii. Used the Archimedean property to show

Theorem: (1) given any positive $x \in \mathbb{R}$ there is a positive integer n such that $\frac{1}{n} < x$, (2) given any $x \in \mathbb{R}$ there is an integer n such that $n \leq x < n+1$, and (3) that given any $x, y, \in \mathbb{R}$ with $x < y$ there is a rational number r such that $x < r < y$.

iii. Used the supremum property and the Archimedean property to show that there is a real number r such that $r^2 = 2$.

iv. Used the supremum property to show that $\mathbb{R}$ satisfied the **closed interval property**. That is, if I_n is a closed interval for each $n \in \mathbb{N}$ and $I_n \supset I_{n+1}$ then $\cap_{i=0}^{\infty} I_n$ is non-empty.

C. How to construct the real numbers

Discussed how to construct $\mathbb{R}$ from $\mathbb{Q}$ in terms of subsets of $\mathbb{Q}$.

D. Are the real numbers good enough

Discussed how $\mathbb{R}$ cannot be extended further if you want a totally ordered field with the supremum property. You can extend to the complex numbers $\mathbb{C}$, but you loose the ordering.

III. The topology of $\mathbb{R}^n$

A. Norms and inner products on vector spaces

1. Review definition of vector space

i. Recalled the definition of vector spaces

ii. Gave several example of vector spaces including cartesian space $\mathbb{R}^n$, the set of polynomials of degree less than or equal to k for some fixed k, the set of all polynomials, the set of all sequences in $\mathbb{R}$, the set of all functions from a set to $\mathbb{R}$.

2. Norms on vector spaces

i. A function $\|\cdot\| : V \to \mathbb{R}$ from a vector space to $\mathbb{R}$ is called a **norm** if

- $\|v\| \geq 0$ for all $v \in V$,
- $\|v\| = 0$ if and only if $v = 0$,
- $\|av\| = |a|\|v\|$ for all $a \in \mathbb{R}$ and $v \in V$,
- $\|v + w\| \leq \|v\| + \|w\|$.

ii. Gave example of the p norm on $\mathbb{R}^n$, for $p \geq 1$ set

$$\|x\|_p = (|x_1|^p + \ldots + |x_n|^p)^{1/p}$$

where $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. We also define

$$\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}.$$

When the notation $\|x\|$ is used for $x \in \mathbb{R}$ we mean $\|x\|_2$ unless otherwise specified.

iii. We defined the p norm on the set of sequences $\mathcal{S}$ by

$$\|s\| = \left(\sum_{i=1}^{\infty} |s_i|^p\right)^{1/p}$$

where $s = (s_n) \in \mathcal{S}$ and

$$\|s\|_\infty = \max\{|s_i|\}.$$

These are not norms on $\mathcal{S}$ since they do not have to be finite on a given sequence. So we define

$$l^p = \{s \in \mathcal{S} : \|s\|_p < \infty\}.$$

These are vector spaces and $\|\cdot\|_p$ is a norm on l^p.

3. Inner products on vector spaces

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

- $\langle v, v \rangle \geq 0$, for all $v \in V$,
- $\langle v, v \rangle = 0$ if and only if $v = 0$,
- $\langle v, w \rangle = \langle w, v \rangle$ for all $v, w \in V$, and
- $\langle v, aw \rangle = a\langle v, w \rangle$ and $\langle v, u + w \rangle = \langle v, u \rangle + \langle v, w \rangle$ for all $v, u, w \in V$ and $a \in \mathbb{R}$.

ii. On $\mathbb{R}^n$ we have the standard "dot product" which gives an inner product

$$\langle x, y \rangle = x \cdot y = x_1 y_1 + \cdots + x_n y_n.$$

iii. Theorem: Given an inner product $\langle \cdot, \cdot \rangle$ on V we get a norm by defining

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Moreover, this norm satisfies the **Cauchy-Schwartz inequality**

$$|\langle v, w \rangle| \leq \|v\| \|w\|,$$

with equality if and only if v and w are co-linear.

B. Open sets

1. We defined the **open ball of radius** r **about** $x \in \mathbb{R}^n$ to be

$$B_r(x) = \{y \in \mathbb{R}^n : \|x - y\| < r\},$$

and the **closed ball** to be

$$\overline{B}_r(x) = \{y \in \mathbb{R}^n : \|x - y\| \leq r\},$$

2. An **open set** in $\mathbb{R}^n$ is a set U such that for each point $x \in U$ there is some $r > 0$ for which $B_r(x) \subset U$.

3. Theorem:

- $\emptyset$ and $\mathbb{R}^n$ are open sets in $\mathbb{R}^n$.
- The intersection of two open sets if open.
- The union of any collection of open sets is open.

A collections of sets satisfying these properties is said to give a **topology** on $\mathbb{R}^n$.

4. Given a point $x \in \mathbb{R}^n$ a **neighborhood of** x is an open set N containing x. (The book says that a neighborhood is any set N containing an open set U that contains x. This is needlessly complicated, but you are welcome to use this definition if you prefer.

5. A point x in a set $A \subset \mathbb{R}^n$ is called an **interior point of** A if there is a neighborhood of x contained in A.

6. The set of all points interior to A is called the **interior of** A and is denoted

$$\text{int}A = \{x \in A : x \text{ is an interior point of } A\}.$$

7. Theorem:

- int$A$ is an open set.
- int$A$ is the largest open set contained in A.
- int$A$ is the union of all open sets contained in A.

8. Theorem: For a set B in $\mathbb{R}^n$ the following statements are equivalent
- B is open.
- int$B = B$.
- B is a neighborhood of each of its points.

C. Closed sets

1. A set C in $\mathbb{R}^n$ is **closed** if its complement, $\mathbb{R}^n - C$, is open.

2. Theorem:
- $\emptyset$ and $\mathbb{R}^n$ are closed sets in $\mathbb{R}^n$.
- The union of two closed sets if closed.
- The intersection of any collection of closed sets is closed.

3. A point $x \in \mathbb{R}^n$ is an **accumulation point**, also called a **cluster point**, of a set $A \subset \mathbb{R}^n$ if every open set containing x also contains a point in A other than x. That is, if U is an open set containing x then

$$(U - \{x\}) \cap A \neq \emptyset.$$

4. Theorem: A set $A \subset \mathbb{R}^n$ is closed if and only if every cluster point of A is contained in A.

5. The **closure** of a set $A \subset \mathbb{R}^n$, denoted $\overline{A}$, is the intersection of all closed sets containing A. (Note the closure of a set is closed.)

6. Theorem: $\overline{A}$ is A together with all its cluster points.

7. The **boundary** of a set $A \subset \mathbb{R}^n$ is defined as

$$\partial A = \overline{A} \cap \overline{\mathbb{R}^n - A}.$$

8. Theorem: A point x is in ∂A if and only if for every $\epsilon > 0$ we have $B_\epsilon(x) \cap A \neq \emptyset$ and $B_\epsilon(x) \cap (\mathbb{R}^n - A) \neq \emptyset$.

D. Sequences

1. Basic definitions and examples

i. A **sequence** in a set $A \subset \mathbb{R}^n$ is a function $s : \mathbb{N} \to A$ from the natural numbers to A. We usually denote the sequence by its image. That is let $s_k = s(k)$ for $k \in \mathbb{N}$, then denote s by $\{s_k\}$.

ii. A sequence $\{s_k\}$ **converges** to a point $x \in \mathbb{R}^n$ (we also say x is a **limit** of the sequence), if for every neighborhood U of x there is some number N such that $s_k \in U$ for all $k \geq N$. If such an x exists then we say the sequence $\{s_k\}$ is **convergent** and write $s_k \to x$ or $\lim s_k = x$. If no such x exists then we say the sequence $\{s_k\}$ is **divergent**

iii. Theorem: A sequence $\{s_k\}$ converges to a point x if and only if for all $\epsilon > 0$ there is a number N such that $\|s_k - x\| < \epsilon$ for all $k \geq N$.

iv. Theorem: $\lim s_k = x$ if and only if $\lim \|s_k - x\| = 0$.

v. Theorem: If $\{s_k\}$ is a convergent sequence then the set of point $\{s_1, s_2, \ldots\}$ that make up the sequence is bounded. (That is there is some r such that $\|s_k\| < r$ for all k.)

vi. Theorem: A sequence $\{s_k\}$ in $\mathbb{R}^n$ converges to a point y if and only if it converges point-wise. (That is if $s_k = (x_{k,1}, \ldots, x_{k,n})$ and $y = (y_1, \ldots y_n)$ then $s_k \to y$ if and only if for each $i, x_{i,k} \to y_i$.)

vii. Theorem: Let $x_k \to x$ and $y_k \to y$ in $\mathbb{R}^n$ and $z_k \to z$ in $\mathbb{R}$, then
- $(x_k + y_k) \to x + y$ this can be written $\lim(x_k + y_k) = \lim(x_k) + \lim(y_k)$.

- $(x_k y_k) \to xy$ this can be written $\lim(x_k y_k) = \lim(x_k)\lim(y_k)$.
- $(z_k y_k) \to zy$ this can be written $\lim(z_k y_k) = \lim(z_k)\lim(y_k)$.
- If $z_k \neq 0$ and $z \neq 0$ then $(y_k/z_k) \to y/z$ this can be written $\lim(y_k/z_k) = \lim(y_k)/\lim(z_k)$.

viii. Given a sequence $\{s_k\}$, a **subsequence** is $\{s_{k_i}\}$ where the k_i are a choice of increasing natural numbers $0 \leq k_1 < k_2 < \ldots < k_i < k_{i+1} < \ldots$.

ix. Theorem: $s_k \to x$ if and only if every subsequence $\{s_{k_l}\}$ of $\{s_k\}$ converges to x.

x. Thoerem (The monotone convergence theorem): Suppose $\{s_k\}$ is a sequence that is monotonically increasing (that is $x_i \leq x_{i+1}$ for all i). Then $\{s_k\}$ converges if and only if it is bounded above, in which case $\lim s_k = \sup\{x_k\}$.

2. Properties of sequences and cluster points

i. Theorem: A sequence can converge to at most one point.

ii. Theorem: Suppose $x \notin A$. Then x is a cluster point of A if and only if there is a sequence of points $\{s_k\}$ in A such that $s_k \to x$.

iii. Theorem: A set A is closed if and only if every sequence $\{s_k\}$ in A which converges has its limit in A.

E. Compact sets

1. Let A be a subset of $\mathbb{R}^n$. A collection of open sets $\{U_\alpha\}_{\alpha \in J}$ is called an **open cover** of A if $A \subset \cup_{\alpha \in J} U_\alpha$. It is called a finite open cover if J is a finite set.

2. A set A is called **compact** if ever open cover of A has a finite subcover, that is if $\{U_\alpha\}_{\alpha \in J}$ is an open cover of A then there is a finite subset J' of J such that $\{U_\alpha\}_{\alpha \in J'}$ is also an open cover of A.

3. Theorem: for a set A in $\mathbb{R}^n$ the following are equivalent:

i. A is compact.

ii. A is closed and bounded.

iii. Any sequence in A has a subsequence that converges to a point in A.

iv. Any infinite set in A has a cluster point in A.

The equivalence **i.** $\Leftrightarrow$ **ii.** is called the **Heine-Borel Theorem**. The equivalence **ii.** $\Leftrightarrow$ **iii.** and **ii.** $\Leftrightarrow$ **iv.** are both called the **Bolzano-Weierstrass Theorem**.

4. Cauchy Sequences: a sequence $\{s_k\}$ in $\mathbb{R}^n$ is said to be a **Cauchy sequence** if for any $\epsilon > 0$ there is an N such that for any $k, l \geq N$ we have $\|s_k - s_l\| < \epsilon$.

5. Theorem: A sequence in $\mathbb{R}^n$ is Cauchy if and only if it converges.

F. Connected Sets

1. A set $D \subset \mathbb{R}^n$ is **disconnected** if there exists open sets U, V in $\mathbb{R}^n$ such that

- $D \subset U \cup V$,
- $D \cap U$ and $D \cap V$ are both non-empty and
- $(U \cap D) \cap (V \cap D) = \emptyset$.

The sets U and V are called a disconnection of D. The set D is **connected** if it is not disconnected.

2. Theorem: the set $(0, 1)$ is a connected subset of $\mathbb{R}$.

3. Theorem: If C is connected and x is a cluster point of C then $C \cup \{x\}$ is connected.

4. Theorem: A subset of $\mathbb{R}$ is connected if and only if it is an interval (that is equal to $(a, b), (a, b], [a, b)$ or $[a, b]$ where for an open end point a could be $-\infty$ and b could be ∞).

5. Theorem: $\mathbb{R}^n$ is connected for all $n \geq 1$.

6. Theorem: The only subsets of $\mathbb{R}^n$ that are both open and closed are $\emptyset$ and $\mathbb{R}^n$.

IV. Continuous Functions

A. Definitions and Examples

1. A function $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ is **continuous at a point** $a \in D$ if for every open set U in $\mathbb{R}^q$ containing $f(a)$ there is an open set V in $\mathbb{R}^p$ containing a such that $V \cap D \subset f^{-1}(U)$.
2. **Theorem:** For a function $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ and point $a \in D$ the following are equivalent:
   - **i.** f is continuous at a.
   - **ii.** For all $\epsilon > 0$ there is a $\delta > 0$ such that for each $x \in D$ with $\|x - a\| < \delta$ we have $\|f(x) - f(a)\| < \epsilon$.
   - **iii.** For all sequences $\{x_n\}$ in D that converge to a we have $f(x_n) \to f(a)$.
3. We say examples of functions continuous at all point of their domain and at no points of their domain. We also saw a function $f; [0,1] \to [0,1]$ that were continuous at the irrational numbers and discontinuous at the rational numbers.

B. **Theorem:** $f : D \subset \mathbb{R}^p \to \mathbb{R}^q, g : D' \subset \mathbb{R}^p \to \mathbb{R}^q$ and $h : D'' \subset \mathbb{R}^p \to \mathbb{R}$. Then

1. If f and g are continuous at $a \in D \cap D'$ then the functions $(f+g)(x) = f(x)+g(x)$, $(f-g)(x) = f(x) - g(x)$ and $(f \cdot g)(x) = f(x) \cdot g(x)$ are all continuous at a.
2. If f and h are continuous at $a \in D \cap D''$ then the function $(hf)(x) = h(x)f(x)$ is continuous at a and if moreover $h(a) \neq 1$ then $(f/h)(x) = f(x)/h(x)$ is continuous at a.
3. **Theorem:** If f is continuous at a and g is continuous at $f(a)$ then $g \circ f$is continuous at a.
4. A function $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ is called **Lipschitz** if there is a constant $K > 0$ such that for all $x, y \in D$ we have $\|f(x) - f(y)\| \leq K\|x - y\|$.
5. **Theorem:** Lipschitz functions are continuous at all points in their domain.
6. **Theorem:** Linear functions are Lipschitz and hence continuous at all points of their domain.

C. Global Continuity

1. A function $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ is **continuous** if it is continuous at all points of its domain D.
2. Given a set $D \subset \mathbb{R}^p$ then a subset $A \subset D$ is called **relatively open** (or open relative to D) if there is some open set V in $\mathbb{R}^p$ such that $A = D \cap V$. Similarly a set $A \subset D$ is called **relatively closed** (or closed relative to D) if there is some closed set C such that $A = C \cap D$.
3. **Theorem:** For a function $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ the following are equivalent:
   - **i.** f is continuous.
   - **ii.** For every open set U in $\mathbb{R}^q$, the set $f^{-1}(U)$ is relatively open in D.
   - **iii.** For every closed set C in $\mathbb{R}^q$, the set $f^{-1}(C)$ is relatively closed in D.
   - **iv.** If $\{x_n\}$ is any sequence in D that converges to a point $a \in D$, then $f(x_n) \to f(a)$.
   - **v.** For each $x \in D$ and $\epsilon > 0$ there is a $\delta > 0$ such that for all $y \in D$ with $\|x - y\| < \delta$ we have $\|f(x) - f(y)\| < \epsilon$.

D. Properties of Continuous Functions

1. **Theorem:** Let $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$. If $H \subset D$ is a connected set and f is continuous on H then $f(H)$ is connected.
2. **Theorem (The Intermediate Value Theorem):** Let $f : D \subset \mathbb{R}^p \to \mathbb{R}$ be continuous. If $H \subset D$ is connected and $x, y \in H$, then for all $c \in \mathbb{R}$ with $f(x) \leq c \leq f(y)$ we have some $z \in C$ such that $f(z) = c$.
3. **Theorem:** Let $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$. If $H \subset D$ is a compact set and f is continuous on H then $f(H)$ is compact.
4. **Theorem (Maximum/Minimum Value Theorem):** Let $f : D \subset \mathbb{R}^p \to \mathbb{R}$. If $K \subset D$ is a compact set and f is continuous on K then there are points x_m and x_M such that for all $z \in K$ we have $f(x_m) \leq f(z) \leq f(x_M)$. That is $f(x_m) = \inf f(K)$ and $f(x_M) = \sup f(K)$.

E. Uniform Continuity

1. A function $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ is **uniformly continuous** on a set $A \subset D$ if for all $\epsilon > 0$ there is a $\delta > 0$ such that for every $x, y \in A$ with $\|x - y\| < \delta$ we have $\|f(x) - f(y)\| < \epsilon$.
2. **Theorem:** If $f : D \subset \mathbb{R}^p \to \mathbb{R}^q$ is a function and f is continuous on a compact set $K \subset D$, then f is uniformly continuous on K.

V. Sequences of Functions

A. Spaces of functions

1. If $D \subset \mathbb{R}^p$ the we denote by $\mathcal{F}(D, \mathbb{R}^q)$ the set of all functions from D to $\mathbb{R}^q$.
2. If $\{f_n\}$ is a sequence in $\mathcal{F}(D, \mathbb{R}^q)$ then we say it **converges point-wise to f on** $D_0 \subset D$ if for each point $x \in D_0$ the sequence of points $\{f_n(x)\}$ in $\mathbb{R}^q$ converges to $f(x)$. We denote this by $f_n \underset{pw}{\longrightarrow} f$ on D_0.
3. We say that a sequence $\{f_n\}$ in $\mathcal{F}(D, \mathbb{R}^q)$ **converges uniformly to f on** $D_0 \subset D$ if for every $\epsilon > 0$ there is some N such that for all $n \geq N$ and $x \in D_0$ we have $\|f(x) - f_n(x)\| < \epsilon$. We denote this by $f_n \underset{u}{\longrightarrow} f$ on D_0.
4. **Theorem:** If a sequence of continuous functions $\{f_n\}$ converges uniformly to f on D_0 then f is continuous on D_0.
5. Uniform convergence implies point-wise convergence, but point-wise convergence does not imply uniform convergence.

B. Norms on function spaces

1. We denote by $\mathcal{B}(D, \mathbb{R}^q)$ the set of bounded function on $D \subset \mathbb{R}^p$.
2. For $f \in \mathcal{B}(D, \mathbb{R}^q)$ we define the **uniform norm** (also known as the **sup norm**) of f to be
$$\|f\|_u = \sup\{\|f(x)\| : x \in D\}.$$
3. **Lemma:** The set $\mathcal{B}(D, \mathbb{R}^q)$ is a vector space (under point wise addition and scalar multiplication) and $\|\cdot\|_u$ is a norm on this vector space.
4. We say a sequence $\{f_n\}$ in $\mathcal{B}(D, \mathbb{R}^q)$ **converges to f in the uniform norm** if
$$\|f - f_n\|_u \to 0.$$
5. **Theorem:** A sequece $\{f_n\}$ in $\mathcal{B}(D, \mathbb{R}^q)$ converges to f uniformly on D if and only if $\{f_n\}$ converges to f in the uniform norm.
6. **Theorem:** Let $\{f_n\}$ be a Cauchy sequence in $\mathcal{B}(D, \mathbb{R}^q)$ (that is for all $\epsilon > 0$ there is some N such that $\|f_n - f_m\| < \epsilon$ for all $n, m \geq N$), then there is some $f \in \mathcal{B}(D, \mathbb{R}^q)$ such that $\|f_n - f\| \to 0$.
7. We call a normed vector space $(V, \|\cdot\|)$ **complete** if every Cauchy sequence in V converges in norm to some point in V. (That is if $\{v_n\}$ is a sequence such for all $\epsilon > 0$ there is some N such that $\|v_n - v_m\| < \epsilon$ for all $n, m \geq N$, then there is some $v \in V$ such that $\|v_n - v\| \to 0$). A complete normed vector space is called a **Banach space**.
8. **Theorem:** The vector spaces $\mathcal{B}(D, \mathbb{R}^q)$ and
$$\mathcal{C}_b(D, \mathbb{R}^q) = \{f \in \mathcal{B}(D, \mathbb{R}^q) : f \text{ is continuous}\}$$
are Banach spaces in the uniform norm.
9. **Theorem:** There is a continuous surjection $f : [0, 1] \to [0, 1] \times [0, 1]$. Such an f is called a **space filling curve** or a **Peano curve**.

C. Approximations of functions

1. **Theorem (Weirstrass-Bernstein Approximation):** If $f : [0, 1] \to \mathbb{R}$ is a continuous function and $\epsilon > 0$ then there is a polynomial $p(x)$ such that $\|f - p\|_u < \epsilon$.

2. In the above theorem can use the **Bernstein polynomial** of f:

$$p_n(x) = \sum_{k=0}^{n} \binom{n}{k} f(\frac{k}{n}) x^k (1-x)^{n-k}.$$

3. **Theorem (Stone-Weierstrass Approximation):** Let $A \subset \mathbb{R}^p$ be compact and $\mathcal{B} \subset \mathcal{C}(A, \mathbb{R})$ (here $\mathcal{C}(A, \mathbb{R})$ is the set of continuous functions with domain A and range $\mathbb{R}$) satisfy

- **i.** $\mathcal{B}$ is an algebra (that is, $f, g \in \mathcal{B}$ and $a \in \mathbb{R}$, implies $fg \in \mathcal{B}$ and $af \in \mathcal{B}$)
- **ii.** the constant function $\mathbf{1} : A \to \mathbb{R} : x \to 1$ is in $\mathcal{B}$ and
- **iii.** $\mathcal{B}$ separates points (that is, for each $x, y \in A$ with $x \neq y$ we have some $f \in \mathcal{B}$ such that $f(x) \neq f(y)$)

Then given any $f \in \mathcal{C}(A, \mathbb{R})$ and ϵ there is some $g \in \mathcal{B}$ such that $\|f - g\|_u < \epsilon$.

D. Interlude: series of numbers

1. If $\{x_n\}$ is a sequence in $\mathbb{R}^p$ then the **(infinite) series** generated by $\{x_n\}$ is the sequence $\{s_k\}$ where $s_k = \sum_{n=1}^{k} x_n$ is the k^{th} **partial sum** of the terms in the sequence $\{x_n\}$. We say the series **converges** or is **summable** if the sequence of partial sums converge, and denote the limit by

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

We abuse notation and also use the symbol to denote the series even if it does not converge.

2. **Lemma:** If the series $\sum_{n=1}^{\infty} x_n$ in $\mathbb{R}^p$ converges then $\lim x_n = 0$.

3. **Theorem:** The series $\sum_{n=1}^{\infty} x_n$ in $\mathbb{R}^p$ converges if and only if for all $\epsilon > 0$ there is some $M > 0$ such that for all $m \geq n \geq M$ we have

$$\|x_{n+1} + \ldots + x_m\| < \epsilon.$$

4. We say $\sum_{n=1}^{\infty} x_n$ in $\mathbb{R}^p$ is **absolutely convergent** (or **converges absolutely**) if the series $\sum_{n=1}^{\infty} \|x_n\|$ converges.

5. **Theorem:** If a series converges absolutely then it converges.

6. **Theorem:** We have the following convergence "tests" or "results"

- **i.** (Geometric series) If $|r| < 1$ is a real number then $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$. If $|r| \geq 1$ then $\sum_{n=0}^{\infty} r^n$ diverges.
- **ii.** (p-series) The series $\sum_{n=1}^{\infty} n^{-p}$ converges if $p > 1$ and diverges (to infinity) if $p \leq 1$.
- **iii.** (comparison test) If the series $\sum_{n=1}^{\infty} a_n$ converges and $0 \leq b_n \leq a_n$ then the series $\sum_{n=1}^{\infty} b_n$ converges. If the series $\sum_{n=1}^{\infty} c_n$ diverges and $0 \leq c_n \leq d_n$ then the series $\sum_{n=1}^{\infty} d_n$ diverges.
- **iv.** (ratio test) Let $\{a_n\}$ be a sequence in $\mathbb{R}$ and let $r = \lim \frac{|a_{n+1}|}{|a_n|}$ the the series $\sum_{n=1}^{\infty} a_n$ converges if $r < 1$, diverges if $r > 1$ and if $r = 1$ then the test is inconclusive.
- **v.** (root test) Let $\{a_n\}$ be a sequence in $\mathbb{R}$ and let $r = \lim |a_n|^{\frac{1}{n}}$ the the series $\sum_{n=1}^{\infty} a_n$ converges if $r < 1$, diverges if $r > 1$ and if $r = 1$ then the test is inconclusive.

7. **Theorem:** If $\sum_{n=1}^{\infty} a_n = S$ and $\sum_{n=1}^{\infty} b_n = T$ then $\sum_{n=1}^{\infty}(a_n + b_n) = S + T$ and $\sum_{n=1}^{\infty} cx_n = cS$ for any $c \in \mathbb{R}$.

8. **Theorem:** Suppose $\{a_n\}$ is a decreasing sequence and $a_n \to 0$. Then the series

$$\sum_{n=1}^{\infty} (-1)^n a_n$$

converges. Also if $s_k = \sum_{n=1}^{k} (-1)^n a_n$ and $S = \sum_{n=1}^{\infty} (-1)^n a_n$ then

- **i.** $s_{2k} > S > s_{2k+1}$ for all k.

ii. $|S - s_k| < a_{k+1}$ for all k.

9. A **rearrangement** of the series $\sum_{n=1}^{\infty} x_n$ is $\sum_{n=1}^{\infty} x_{\sigma(n)}$ where $\sigma : \mathbb{N} \to \mathbb{N}$ is a bijection.
10. **Theorem:** If $\sum_{n=1}^{\infty} x_n$ is absolutely convergent the so is any rearrangement and both series converge to the same thing.
11. **Theorem (Riemann rearrangement):** If $\sum_{n=1}^{\infty} x_n$ is convergent but not absolutely convergent and s is any real number then there is a rearrangement of $\sum_{n=1}^{\infty} x_n$ that converges to s.

E. Series of functions

1. Let $\{g_k\}$ be a sequence of functions in $\mathcal{C}(D, \mathbb{R}^q)$ for some domain $D \subset R^p$. We say the series $\sum_{k=1}^{\infty} g_k$ **converges point-wise to** $g : D \to \mathbb{R}^q$ if for every $x \in D$ the series $\sum_{k=1}^{\infty} g_k(x)$ in $\mathbb{R}^q$ converges to $g(x)$. We denote this

$$\sum_{k=1}^{\infty} g_k = g \quad \text{(point-wise)}.$$

We say the series **converges absolutely to** g if the series $\sum_{k=1}^{\infty} \|g_k(x)\|$ in $\mathbb{R}$ converges to $\|g(x)\|$ for all $x \in D$. Finally, the series **converges uniformly to** g if the sequence of partial sums $s_n = \sum_{k=1}^{n} g_k$ converges uniformly to g. We denote this

$$\sum_{k=1}^{\infty} g_k = g \quad \text{(uniformly)}.$$

2. **Theorem:** If $\{g_k\}$ is a sequence of continuous functions and $\sum_{k=1}^{\infty} g_k = g$ (uniformly), then g is continuous.
3. **Theorem (Weierstrass M-test):** Suppose
   a) $\{g_k : D \subset \mathbb{R}^p \to \mathbb{R}^q\}$ is a sequence of functions and
   b) for each k there are constants M_k satisfying $\|g_k(x)\| \leq M_k$ for all $x \in D$.

   If $\sum_{k=1}^{\infty} M_k$ converges then, $\sum_{k=1}^{\infty} g_k$ converges uniformly (and absolutely) on D.
4. **Theorem (Abel test:):** Let $\phi_n : D \subset \mathbb{R}^p \to \mathbb{R}$ be a sequence of functions satisfying
   a) (the ϕ_n are point-wise decreasing) $\phi_{n+1}(x) \leq \phi_n(x)$ for all $x \in D$ and n, and
   b) (the ϕ_n are bounded) there is some M such that $|\phi_n(x)| \leq M$ for all $x \in D$ and n.

   If $\sum_{n=1}^{\infty} g_n$ is a uniformly convergent series on D then so is $\sum_{n=1}^{\infty} \phi_n g_n$.
5. **Theorem (Dirchlet test):** Suppose
   a) $\{f_k : D \subset \mathbb{R}^p \to \mathbb{R}^q\}$ is a sequence of functions with uniformly bounded partial sums, that is, for which there is a constant M such that

   $$\|s_n(x)\| \leq M \quad \text{for all } x \in D \text{ and } n,$$

   where $s_n(x) = \sum_{k=1}^{n} f_k(x)$; and
   b) $\{g_k : D \to \mathbb{R}\}$ is a sequence of decreasing positive functions (that is $g_n(x) \geq g_{n+1}(x) \geq 0$) that converges uniformly to 0.

   Then $\sum_{k=1}^{\infty} g_k f_k$ converges uniformly on D.

F. Power series

1. Let $c \in \mathbb{R}$. A series of function $\sum_{n=0}^{\infty} f_n$ is called a **power series about** $x = c$ if each of the f_n is of the form

$$f_n(x) = a_n(x - c)^n,$$

for some constant a_n; that is a series of the form $\sum_{n=0}^{\infty} a_n(x - c)^n$.

2. Given a sequence $\{b_n\}$ of non-negative numbers that is bounded above we define the **limit superior** of $\{b_n\}$ to be

$$\limsup b_n = \inf\{v : v \text{ is larger than all but finitely many } b_n\}$$

If $\{b_n\}$ is not bounded then we set $\limsup b - n = \infty$.

3. **Properties of the limit superior:** (1) It's is always well-defined. (2) If $v > \limsup b_n$ then there is some N such that for $n \geq N$ we have $b_n \leq v$. (3) If $v < \limsup b_n$ then for any N there are $n > N$ such that $b_n > v$. (4) If $\lim b_n$ exists then $\limsup b_n = \lim b_n$. (5) If $c \geq 0$ then $\limsup cb_n = c \limsup b_n$. (6) $\limsup(a_n + b_n) \leq \limsup a_n + \limsup b_n$.

4. Given a power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ let $\rho = \limsup |a_n|^{1/n}$. Then the **radius of convergence** of the power series is

$$R = \begin{cases} \infty & \text{if } \rho = 0 \\ \frac{1}{\rho} & \text{if } 0 < \rho < \infty \\ 0 & \text{if } \rho = \infty. \end{cases}$$

The **interval of convergence** is $(c - R, c + R)$.

5. **Theorem:** If R is the radius of convergence for the power series $\sum_{n=0}^{\infty} a_n(x-c)^n$, then the series converges absolutely for $|x - c| < R$ and diverges for $|x - c| > R$.

6. **Theorem:** Given a power series $\sum_{n=0}^{\infty} a_n(x-c)^n$, the radius of convergence is given by $\lim \frac{|a_n|}{|a_{n+1}|}$ if the limit exists.

7. **Theorem:** If R is the radius of convergence for the power series $\sum_{n=0}^{\infty} a_n(x-c)^n$, then the series converges uniformly on any compact subset of $(c-R, c+R)$. In particular, the series defines a continuous function on $(c - R, c + R)$.

8. We can define the following functions using power series:

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n, \quad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \text{ and } \quad \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}.$$

Their radius of convergence is infinite for all these functions and so they all are continuous functions on $\mathbb{R}$. Once can use the power series to prove that $e^{x+y} = e^x e^y$ and $e^{iz} = \cos z + i \sin z$. These two formulas can be used to derive the angle sum formulas for sin and cos as well as all the other trigonometric formulas.

G. Fourier Series

1. The L^2 inner product

i. **Theorem:** Let $I \subset \mathbb{R}$ be an interval and $f, g : I \to \mathbb{R}$ be functions. Set

$$\langle f, g \rangle = \int_I f(x)g(x)\, dx.$$

Let $\mathcal{L}_c^2(I) = \{f : I \to \mathbb{R} \text{ continuous with } \langle f, f \rangle < \infty\}$. Then $\mathcal{L}_c^2(I)$ is a vector space and $\langle \cdot, \cdot \rangle$ is an inner product on it.

ii. We call the above inner product the L^2 **inner product** on $\mathcal{L}_c^2(I)$.

iii. Like all inner products this one induces a norm $\|\cdot\|_2$ on $\mathcal{L}_c^2(I)$ called the L^2 **norm**.

iv. We say a sequence $\{f_n\}$ in $\mathcal{L}_c^2(I)$ **converges in** L^2 to f if such that $\|f_n - f\|_2 \to 0$ as $n \to \infty$. We say the sequence is Cauchy in L^2 if for all $\epsilon > 0$ there is an N such that $\|f_n - f_m\|_2 < \epsilon$ for all $n, m \geq N$.

v. $\mathcal{L}_c^2(I)$ with the L^2 norm is not complete (that is a Cauchy sequence does not have to converge).

2. Orthonormal sets

i. A set $\{v_\alpha\}_{\alpha \in A}$ in an inner product space $(V, \langle \cdot, \cdot \rangle)$ is called **orthonormal** if
- $\langle v_\alpha, v_\beta \rangle = 0$ for all $\alpha \neq \beta$ and
- $\langle v_\alpha, v_\alpha \rangle = 1$ for all α.

ii. **Theorem:** Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. (1) if $\{v_\alpha\}_{\alpha \in A}$ is an orthonormal set of vectors then they are also linearly independent. (2) If $\{v_i\}_{i=1}^n$ (where n could be ∞) is a linearly independent set of vectors then there is an orthonormal set $\{w_i\}_{i=1}^n$ such that

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(w_1, \dots, w_k)$$

for all $k = 1, \dots, n$. (The algorithm in the proof to construct the w_i is called the **Gram-Schmidt** process.)

iii. Theorem: Suppose that $\{v_i\}_{i=1}^\infty$ is an orthonormal set in the inner product space $(V, \langle \cdot, \cdot \rangle)$. If $\sum_{i=1}^\infty a_i v_i$ converges in norm to v then $a_i = \langle v, v_i \rangle$.

iv. Theorem (Bessel's inequality): Suppose that $\{v_i\}_{i=1}^\infty$ is an orthonormal set in the inner product space $(V, \langle \cdot, \cdot \rangle)$. For any $v \in V$ the series

$$\sum_{i=1}^\infty |\langle v, v_i \rangle|^2$$

converges and

$$\sum_{i=1}^\infty |\langle v, v_i \rangle|^2 \leq \|v\|^2,$$

where $\| \cdot \|$ is the norm associated to the inner product.

v. We say an orthonormal set $\{v_i\}_{i=1}^\infty$ in an inner product space $(V, \langle \cdot, \cdot \rangle)$ is **complete** if for every $v \in V$ there are constants a_i such that $v = \sum_{i=1}^\infty a_i v_i$. (Note from above each a_i must equal $\langle v, v_i \rangle$.) The series $\sum_{i=1}^\infty \langle v, v_i \rangle v_i$ is called the **Fourier series of v with respect to $\{v_i\}_{i=1}^\infty$** and the constants $\langle v, v_i \rangle v_i$ are called the **Fourier coefficients** of v.

vi. Theorem: Suppose that $\{v_i\}_{i=1}^\infty$ is an orthonormal set in the inner product space $(V, \langle \cdot, \cdot \rangle)$. The set $\{v_i\}_{i=1}^\infty$ is complete if and only if

$$\|v\|^2 = \sum_{i=1}^\infty |\langle v, v_i \rangle|^2,$$

for all $v \in V$. (The equality is called **Parseval's equality**.)

3. The Fourier series

i. Lemma: The set $\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \sin nx\}_{n=1}^\infty$ is an orthonormal set in $\mathcal{L}_c^2([-\pi, \pi])$.

ii. Given a function $f \in \mathcal{L}_c^2([-\pi, \pi])$ (or more generally a piece-wise continuous function ...) then the **Fourier coefficients of f** are

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx, \quad n = 0, 1, \dots$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx, \quad n = 1, 2, \dots$$

and the **Fourier series of f** is

$$f(x) \sim \frac{1}{2} a_0 + \sum_{n=1}^\infty (a_n \cos nx + b_n \sin nx),$$

here you should read "$\sim$" as "has the Fourier series" (there is no guarantee that the term on the left and right are equal).

iii. Corollary (Bessel Inequality): For f as above

$$\frac{a_0^2}{2} + \sum_{i=1}^\infty (a_n^2 + b_n^2) \leq \frac{1}{\pi} \|f\|_2^2,$$

where a_n and b_n are the Fourier coefficients.

4. Convergence of Fourier series

i. Theorem: The set $\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos nx, \frac{1}{\sqrt{\pi}}\sin nx\}_{n=1}^{\infty}$ is "complete" in $\mathcal{L}_c^2([-\pi,\pi])$. In particular, the Fourier series of f approaches f in the L^2 norm.

N.B. The set is not really complete, largely because $\mathcal{L}_c^2$ is not complete as a normed vector space, so the series might converge, but not to something actually in $\mathcal{L}_c^2$. This issue will not be a problem for the following corollary though and we will see exactly how it fails below.

ii. Corollary (Parseval's equality): For f as above

$$\frac{a_0^2}{2} + \sum_{i=1}^{\infty}(a_n^2 + b_n^2) = \frac{1}{\pi}\|f\|_2^2,$$

where a_n and b_n are the Fourier coefficients.

iii. A function is called **piece-wise Lipschitz** if it is piece-wise continuous and there is some constant L such that L is a Lipschitz constant for the function on each interval of convergence. (Here when discussing piece-wise continuous we mean that the domain can be broken into intervals such that the function is continuous on the interiors of each interval and the function has a limit as you approach each end point of an interval of continuity.)

iv. Theorem (Dirichlet-Jordan): Suppose that f is a piecewise Lipschitz function on $[-\pi,\pi]$ (that is extended to be 2π periodic). Then if f is continuous at x the Fourier series converges to $f(x)$. If f has a jump discontinuity at x then the Fourier series converges to

$$\frac{f(x^+) - f(x^-)}{2},$$

where $f(x^+)$ is the limit of $f(x_n)$ where x_n is a sequence approaching x from above and $f(x^-)$ is the limit of $f(x_n)$ where x_n is a sequence approaching x from below.