MATH321 Class Note

Haozhen Xu

September 2024

Contents

1 Dedekind Cuts

Theorem 1.1. *The set $\mathbb{R}$, constructed by means of Dedekind cuts, is complete in the sense that it satisfies:*
Least Upper Bound Property: *If S is a nonempty subset of $\mathbb{R}$ and is bounded above then in $\mathbb{R}$ there exists a least upper bound for S.*

Proof:

1. Let $\lambda \in \mathbb{R}$ be any nonempty collection of cuts which is bounded above by the cut $X|Y$.
2. Now define
$$C = \{a \in \mathbb{Q} : \text{for some cut } A|B \in \lambda \text{ s.t. } a \in A\}$$
$$D = \{\text{the rest of } \mathbb{Q}\}$$
Comment: this step basically makes C the set of all the element in the lower half of the set of cuts λ, therefore develop a new cut.

3. It is easy to see that $C|D$ is also a cut, and it is also the upper bound for λ (because the A for every element of λ is contain in C)

   Comment: this sentence means the construction of the set is constructed by making the union of all the A, therefore every A is in C.

4. Now let $z' = C'|D'$ be any upper bound for λ

   Comment: this is what we discuss why it is the *least* upper bound.

5. By assumption, we know that $A|B \leq C'|D' \forall A|B \in \lambda$, we see that the A for very member of λ is contained in C'. Hence $C \subset C'$.

   Comment: This is basically proving that for every element in the set λ's bounded above by z', therefore, the cut z is less than any upper bound z'.

6. Thus z is the least upper bound.

1.1 Natural Arithmetic of Cuts

We do arithmetic of cuts or real numbers by do the corresponding operation to the elements comprising the two halves of the cuts.

Definition. Let $x = A|B$ and $y = C|D$, the sum of x and y is $x + y = E|F$ where

$$\begin{cases} E = \{r \in \mathbb{Q} : \text{ for some } a \in A \text{ and for some } c \in C \text{ we have } r = a + c\} \\ F = \text{the rest of } \mathbb{Q} \end{cases}$$

It is easy to see that $E|F$ is a cut in $\mathbb{Q}$ and that it does not depend on the order in which x and y appear.

Thus, it implies the cut addition has commutativity, which means $x + y = y + x$.

Definition. The additive inverse of $x = A|B$ is $-x = C|D$ where

$$\begin{cases} C = \{r \in \mathbb{Q} : \text{for some } b \in B \text{ , not the smallest element of } B, r = -b\} \\ D = \text{the rest of } \mathbb{Q} \end{cases}$$

Comment: here the fact that b can not be the smallest element be is because C cannot have a greatest element.

For the additive inverse, the following property also stands

$$(-x) + x = 0*$$

Definition. The zero cut $0*$ is given by $0 * + x = x$, the $0*$ is called the additive identity.

Definition. Correspondingly, the difference of cuts $x - y = x + (-y)$.

Definition. Multiplication of cuts:

1. The cut $A|B$ is positive if $0* < x$ or negative if $x < 0*$.

2. Since 0 lies in A or B, a cut is either positive, negative, or zero.

If $A|B$ and $y = C|D$ are positive cuts, then their product is $x \cdot y = E|F$, where

$$\begin{cases} E = \{r \in \mathbb{Q} : r \leq 0 \vee (\exists a \in A, c \in C) \text{ such that } a > 0, c > 0 \text{ and } r = ac\} \\ F = \text{ the rest of } \mathbb{Q} \end{cases}$$

Comment: This definition means that all the numbers in the set that can composed as multiply of two number in the lower half of the product side.

Similarly, when some cut is negative, we can use the following method to turn them into positive cuts:

1. If x is positive and y is negative then

$$x \cdot y = -(x \cdot (-y))$$

   Comment: This process is like: we only know the product using two positive cuts, so we turn a negative cut into a positive cut. And okay now we have a positive cut, then we determine its sign.

2. If x is negative and y is positive then

$$x \cdot y = -((-x) \cdot y)$$

3. IF x is negative and y is positive then

$$x \cdot y = (-x) \cdot (-y)$$

4. If x or y is the zero cut $0*$ then we define $x \cdot y = 0*$

we will not cover every detail of verifying all the properties of multiplication, as an example, we verify the commutativity of multiplication.

Proof:

1. If x, y are positive then
2. $\{ac : a \in A, c \in C, a > 0, c > 0\} = \{ca : c \in C, a \in A, a > 0, c > 0\}$
3. Which implies that $x \cdot y = y \cdot x$.
4. On the other hand, if x is positive and y is negative then

$$x \cdot y = -(x \cdot (-y)) = -((-y) \cdot x) = y \cdot x$$

   this second equality holds because we have already checked commutativity for positive cuts.

Definition. A field is a system consisting of a set of elements and two operations, addition and multiplication, that have the preceding algebraic properties.

1. $a + b = b + a$, $a \cdot b = b \cdot a$ commutativity.
2. $(a + b) + c$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ associativity.
3. $(a + b) \cdot c = a \cdot b + a \cdot c$ distributivity.
4. $x < y < z$ implies $x < z$ transitivity
5. Either $x < y$, $y < x$, or $x = y$ trichotomy
6. $x < y$ implies $x + z < y + z$ translation.
7. if $a > 0$, $b > 0$ then $ab > 0$.

By definition, this is what we called $\mathbb{R}$ is an ordered field.

Definition. Cut arithmetic is consistent with $\mathbb{Q}$ arithmetic in the sense that if $c, r \in \mathbb{Q}$ then

$$c * + r* = (c + r)*$$

$$c * \cdot r* = (cr)*$$

by definition, this what we called a subfield of $\mathbb{R}$

Besides, the product of positive cuts is positive and cut order is consistent with $\mathbb{Q}$ order

1.2 Magnitude, Subcollection, and Properties of $\mathbb{R}$

Theorem 1.2. *The set $\mathbb{R}$ of all cuts in $\mathbb{Q}$ is a complete ordered field that contains $\mathbb{Q}$ as an ordered subfield.*

Definition. Magnitude or absolute value of $x \in \mathbb{R}$ is given by

$$|x| = \begin{cases} x \text{ if } x \geq 0 \\ -x \text{ if } x < 0 \end{cases}$$

Thus $x \leq |x|$.

Theorem 1.3. *Triangular Inequality: For all $x, y \in \mathbb{R}$ we have $|x + y| \leq |x| + |y|$.*

Proof: The translation and transitivity properties of the order relation and from the inequality we proved imply that.

$$\begin{cases} x + y \leq |x| + y \leq |x| + |y| \\ -x - y \leq |x| - y \leq |x| + |y| \end{cases}$$

where

$$|x + y| = \begin{cases} x + y \text{ if } x + y \geq 0 \\ -x - y \text{ if } x + y \leq 0 \end{cases}$$

therefore the first inequality proves both cases.

Definition. The real number system $\mathbb{R}$ exists and it satisfies the properties of complete ordered field. A complete field basically means there is no "gap" existing in the set, or every Dedekind cut exists in the field itself.

Proof:

1. Suppose we try the same cur construction in $\mathbb{R}$ that we did in $\mathbb{Q}$.
2. The definition of cut $A|B$ is that A and B are disjoint nonempty subcollection of element, while $A \cup B = \mathbb{R}$.
3. Further more A contains no largest element. However, different from $\mathbb{Q}$, the field of real number has the *lub* property.
4. Therefore $y = lub(A)$ exists $a \leq y \leq b$ for every $a \in A$ and $b \in B$.
5. In other word, we know that there is no gap in the field $\mathbb{R}$.

Definition. $\mathbb{R}$ is unique. Any complete ordered field $\mathbb{F}$ containing $\mathbb{Q}$ as an ordered subfield corresponds to $\mathbb{R}$ in a way preserving all the ordered field structure.

To see this, take any $\phi \in \mathbb{F}$ and associate to it the cut $A|B$ where

$$A = \{r \in \mathbb{Q} : r < \phi \in \mathbb{F}\}$$

and B is the rest of $\mathbb{Q}$. And the correspondence makes $\mathbb{F} = \mathbb{R}$.

Comment: The "correspondence" refers to the way that any element ϕ in a complete ordered field $\mathbb{F}$ can be uniquely associated with a Dedekind cut in $\mathbb{Q}$.

The correspondence works because the structure of a Dedekind cut uniquely determines a real number. In a complete field, every element can be described this way, and the field will have the same structure as $\mathbb{R}$

Remark. Note that $+\infty$ a $-\infty$ are not real numbers, since $\mathbb{Q}|\emptyset$ and $|\mathbb{Q}$ are not cuts. However, it is convenient to write expressions like $x \to \infty$ to indicate that a real variable x grows larger and larger without bound.

Definition. If S is a nonempty subset of $\mathbb{R}$ then its supremum is its least upper bound with S is bounded above and ∞ otherwise.

The set S's infimum is its greatest lower bound when S is bounded below and is said to be $-\infty$ otherwise.

2 Basics

2.1 Cauchy Sequences And Cauchy Criteria

Definition. Let $a_1, a_2, a_3, ... = (a_n)$, $n \in \mathbb{N}$ be a sequence of real numbers. The sequence converges to a limit $b \in \mathbb{R}$ as $n \to \infty$ if

$$\forall \epsilon,\ \exists N \in \mathbb{N} \text{ s.t. } \forall n \geq N,\ n \geq N,\ |a_n - b| < \epsilon$$

Comment: It means that the sequence (a_n) will eventually all its terms nearly equal to b.

Theorem 2.1. *Every convergent sequence obeys a Cauchy condition which states that*

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N}\ s.t.\ (n, k > N \to |a_n - a_k| < \epsilon)$$

The converse of this fact is a fundamental property of $\mathbb{R}$.

Theorem 2.2. $\mathbb{R}$ *is complete with respect to Cauchy sequences in the sense that if* (a_n) *is a sequence of real numbers which obeys a Cauchy condition then it converges to a limit in* $\mathbb{R}$.

Proof:

1. To begin with, we show that (a_n) is bounded. Take $\epsilon = 1$ in the Cauchy condition implies there is an N such that for all $n, k \geq N \to |a_n - a_k| < 1$.

2. Now we take K large enough so that

$$-K \leq a_1, ..., a_n \leq K$$

   Now set $M = K + 1$ then

$$\forall n \text{ we have } -M < a_n < M$$

   Therefore this set is bounded.

   Comment: This set basically say that we have a set that includes some points up until one point, then because all the points are close together, now by distance 1, then they must be within distance 1 radius off the set.

3. Now define a set X as

$$X = \{x \in \mathbb{R} : \exists \text{ indefinitely many } n \text{ such that } a_n \geq x\}$$

   Obviously $-M \in X$ and $M \notin X$.

4. Thus X is a nonempty subset of $\mathbb{R}$ which is bounded above by M. The least upper bound property applies to X and we have $b = lub(X)$ with $-M \leq b \leq M$.

5. Now we claim that a_n converges to b as $n \to \infty$. Given $\epsilon > 0$ we must show that

$$\exists N \text{ such that } |a_n - b| \leq \epsilon\ \forall n \geq N$$

6. Because the sequence (a_n) is Cauchy and $\frac{\epsilon}{2}$ is positive, $\exists N$ such that

$$n, k \geq N \to |a_n - a_k| < \frac{\epsilon}{2}$$

7. Because $b - \frac{\epsilon}{2}$ is less than b, it is not an upper bound of X, so there is $x \in X$ with $b - \frac{\epsilon}{2} \leq x$. Therefore by definition of set X, there are infinitely many n with have $a_n > x$.

   Because $b + \frac{\epsilon}{2} > b$ then it does not belong to X, there are only finitely many n we have $a_n > b + \frac{\epsilon}{2}$.

8. Thus there are infinitely many n between
$$b - \frac{\epsilon}{2} \leq x \leq a_n \leq b + \frac{\epsilon}{2}$$
Now pick one of the a_n, say a_{n_0} with $n_0 \geq N$ and
$$b - \frac{\epsilon}{2} \leq a_{n_0} \leq b + \frac{\epsilon}{2}$$
then for all $n \geq N$ we have
$$|a_n - b| \leq |a_n - a_{n_0}| + |a_{n_0} - b| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$
Comment: This technique is called the two-epsilon method. Basically, the logic of this proof is that:
   (a) We first show that all the points are close together by Cauchy sequence and show that all the points.
   (b) Then show that one point is close to b because the lower bound we chose. Therefore we show all the points are close to b.

Theorem 2.3. ***Cauchy Convergence Theorem:*** *A sequence (a_n) in $\mathbb{R}$ in $\mathbb{R}$ converges iff*
$$\forall \epsilon > 0 \exists N \in \mathbb{N} \text{ such that } n, k \geq N \to |a_n - a_k| < \epsilon$$

2.2 Further Description of $\mathbb{R}$

The element of $\mathbb{R}/\mathbb{Q}$ are irrational numbers.

1. If x is irrational and r is rational then $y = x + r$ is irrational.
2. If y is rational then so is $y - r = x$, the difference of rational is rational.
3. If $r \neq 0$ then rx is irrational.
4. The reciprocal of an irrational number is irrational.

Theorem 2.4. *Every interval (a, b), no matter how small, contains both infinitely many rational and irrational numbers.*

Proof:

1. Take a, b as cuts $a = A|A'$ and $b = B|B'$. The fact that $a < b$ implies the set B/A is a nonempty set of rational numbers.
2. Let $r \in B/A$. Because B has no largest element, $\exists s \in \mathbb{Q}$ such that $a < r < s < b$.
3. Now consider the transformation
$$T : t \to r + (s - r)t$$
which basically sends element from interval $[0, 1]$ to $[r, s]$. And by the properties mentioned at the start of the section, T sends rationals to rationals, irrationals to irrationals.
4. Clear, $[0, 1]$ contains infinitely many rationals (the set of all $\frac{1}{n}$) and irrational (the set of all $\frac{1}{\sqrt{2}n}$). Therefore the set $[r, s]$ also contains infinitely amount of rationals and irrationals.

Comment: $\mathbb{R}$ is like a rubber, stretch it out and it never breaks.

2.3 Euclidean Space

Definition. Given set A and B, the Cartesian Product of A and B is the set $A \times B$ of all ordered pairs (a, b) such that $a \in A$ and $b \in B$.

Definition. The Cartesian product of $\mathbb{R}$ with itself m times is denoted $\mathbb{R}^m$. In this terminology real numbers are called scalars and $\mathbb{R}$ is called the scalar field.

Definition. The dot product or inner product of two vectors $(x_1, ..., x_n)$ and $(y_1, ..., y_n)$ are given by

$$x_1y_1 + ... + x_ny_n$$

Definition. The length or magnitude of a vector $x \in \mathbb{R}^m$ is given by

$$|x| = \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + ... + x_m^2}$$

Theorem 2.5. ***Cauchy Schwartz Inequality:*** *$\forall x, y \in \mathbb{R}^m$ we have*

$$\langle x, y\rangle \leq |x||y|$$

Proof:

1. For any vector x, y consider the new vector $w = x + ty$, where $t \in \mathbb{R}$ is a varying scalar.Then

$$Q(t) = \langle w, w\rangle = \langle x + ty, x + ty\rangle$$

therefore is a real-valued function.

2. In fact, the dot product of any vector with itself is non-negative. The bilinearity of the dot product implies that

$$Q(t) = \langle x, x\rangle + 2t\langle x, y\rangle + t^2\langle y, y\rangle = c + bt + at^2$$

is a quadratic function of t.
Comment: This dot product with itself is very useful because now we can have both the three inner products $\langle x, x\rangle$, $\langle x, y\rangle$, $\langle y, y\rangle$ and with a little algebraic technique we are able to produce something we want.

3. The fact that $Q(t)$ is non-negative, shows that $Q(t)$ may not be equal to negative number, thus cannot have two distinct roots. In other word, we know that the discriminant $b^2 - 4ac \leq 0$ thus

$$4\langle x, y\rangle^2 - 4\langle x, x\rangle\langle y, y\rangle \leq 0$$

by taking the square root of both sides gives

$$\langle x, x\rangle \leq \sqrt{\langle x, x\rangle\langle y, y\rangle} = |x||y|$$

Theorem 2.6. *Triangle Inequality for vectors: For every $x, y \in \mathbb{R}^m$, we have that*

$$|x + y| = |x| + |y|$$

which can be easily proved by Cauchy-Schwartz inequality.
Additionally, the triangular inequality of distance is given by

$$|x + y| \leq |x - y| + |y - z|$$

Definition. Vector in $\mathbb{R}^m$ is referred to the points in $\mathbb{R}^m$.

The jth coordinate of the point $(x_1, ..., x_m)$ is the number x_j appearing in the jth position. And the jth coordinate axis is the set of point $x \in \mathbb{R}^m$ whose kth coordinates are 0 for all $k \neq j$.

The first orthant of $\mathbb{R}^m$ is the set of point $x \in \mathbb{R}^m$ all of whose coordinates are nonnegative.

The integer lattice is the set $\mathbb{Z}^m \subset \mathbb{R}^m$ of ordered $m-tuples$ of integers. The integer lattice is also called the integer grid.

Definition. A box is a cartesian product of intervals

$$[a_1, b_1] \times ... \times [a_m, b_m]$$

in $\mathbb{R}^m$ also called a rectangular parallelepiped. The unit cube is the box

$$[0,1]^m = [0,1] \times ... \times [0,1]$$

Definition. The unit ball in $\mathbb{R}^m$ is defined as $B^m = \{x \in \mathbb{R}^m : |x| \leq 1\}$
The unit sphere in $\mathbb{S}^{m-1} = \{x \in \mathbb{R}^m : |x| = 1\}$.

Definition. The set $E \in \mathbb{R}^m$ is convex is for each pair of points $x, y \in E$, the straight line segment between x and y is also contained in E, express mathematically

$$sx + ty \in S$$

where $s + t = 1$ and $s, t \leq 1$ these linear combination are called convex combinations. The line segement is denotes as $[x, y]$.

Theorem 2.7. *Balls in $\mathbb{R}^m$ is convex.*

Proof:

1. Now if $x, y \in B^m$ and $sx + ty = z$ is a convex combination of x and y.

$$\langle z, z\rangle = s^2\langle x, x\rangle + 2st\langle x, t\rangle + t^2\langle y, y\rangle$$

2. Using the Cauchy-Schwartz inequality and the fact that $2st \geq 0$

$$\leq s^2|x|^2 + 2st|x||y| + t^2|y|^2 \leq s^2 + 2st + t^2 = (s+t)^2 = 1$$

3. Now take the square root of both sides gives $|z| \leq 1$, which proves convexity of the ball (by the definition of the ball).

2.4 Cardinality

Definition. Let A and B be sets, a function $f : A \to B$ is a rule or mechanism which when presented with any element $a \in A$, produces an element $b = f(a)$ of B. A function is also called a mapping or a transformation.

Definition. The set A is the domain of the function B is the target or the codomain.
The range or image of f is the subset of B such that

$$\{b \in B : \text{ there exists at least one element } a \in A \text{ such that } f(a) = b\}$$

Definition. A mapping $f : A \to B$ is an injection (one-to-one) if for each pair of distinct element $a, a' \in A$, the element $f(a), f(a')$ are distinct in B.

A mapping is a surjection if for each $b \in B$ there is at least one $a \in A$ such that $f(a) = b$. In other word, the range of f is B.

A mapping a bijection if it is both injective and surjective.

Definition. The identity map of any set is the bijection that takes each $a \in A$ and sends it to itself.

Definition. If $f : A \to B$ and $g : B \to C$ then the composite $gf : A \to C$ is the function that sends $a \in A$ to $g(f(a)) \in C$. If f and g are surjective then se is gf.

Definition. If there is a bijection from A onto B then $A$and B, then we write A B which is an equivalence relation. We say the two set have the same cardinality.

1. A A

2. A $B \to B$ A

3. $A \; B \; C \to A \; C$

Definition. A set S is

1. Finite if $S = \emptyset$ or for some $n \in \mathbb{N}$ we have $S \; \{1, ..., n\}$
2. Infinite if it is not finite.
3. Denumerable if $S \; \mathbb{N}$
4. Countable if it is finite or denumerable.
5. Uncountable if it is not countable.

Theorem 2.8. $\mathbb{R}$ *is uncountable. This is prove by the Cantor's diagonal theorem, which I will not go over in this note.*

Corollary 2.8.1. $[a, b]$ *and* (a, b) *are uncountable. Not proven in this note.*

Theorem 2.9. *Each infinite set* S *contains a countable or denumerable subset.*

Proof:

1. Since S is infinite, it is nonempty and contains and element $s_1 \in S$
2. Then pick $s_2 \in S/\{s_1\}$
3. Then pick $s_3 \in S/\{s_1, s_2\}$
4. This process is always possible because S is infinite.

Theorem 2.10. *An infinite subset* A *of a denumerable set* S *is denumerbale.*

Proof:

1. Assume that $S : \mathbb{N} \to B$ exists, each element of A appears exactly once in the list $f(1), f(2), ...,$ of B.
2. Define $g(k)$ to be the kth element of A appearing in the list. Since A is infinite, $g(k)$ is defined for all $k \in \mathbb{N}$ thus $g : \mathbb{N} \to A$ is a bijection and A is denumerable.
3. *Comment:* The proof is basically saying that since $\mathbb{N}$ and B is having a bijection, then the values of f is therefore distinct and hence we can say that the elements in $f(A)$ will also be distinct as $A \subset B$.

   Additionally, because we can see that we can also make A listed and the fact that $g(A)$ is defined, we can also say that A is also denumerable.

Corollary 2.10.1. *Even numbers are denumerable, also primes are denumerable.*

Theorem 2.11. $\mathbb{N} \times \mathbb{N}$ *is denumerable.*

Proof:

1. Proved by the snakey diagonal.

Corollary 2.11.1. *If* A *and* B *are countable, then* $A \times B$ *is countable. We can prove this theorem by defining*

$$\mathbb{N} \times \mathbb{N} \to A \times B$$

by letting $g_1(\mathbb{N}) \to A$ *and* $g_2(\mathbb{N}) \to B$ *exists. Therefore the mapping is going to be a bijection.*

Theorem 2.12. *If* $f : \mathbb{N} \to B$ *is a surjection and* B *is infinite then* B *is denumerable.*

Proof:

1. For each $b \in B$ the set $\{k \in \mathbb{N} : f(k) = b\}$ is nonempty and hence contains a smallest element.
2. Say that $h(b) = k$ is the smallest element that is sent to b by f, clearly if $b, b' \in B$ and $b \neq b' \to h(b) \neq h'(b)$.

3. Then that means $h : B \to \mathbb{N}$ is a injection which bijects B to $h(B) \subset \mathbb{N}$.

4. By theorem 5.3: an infinite subset A of a denumerable set B is denumerable, we know that hB is denumerable and there B.

5. *Comment:* The logic of this proof is that we are unable to prove they have the same cardinality straight away because the equality in cardinality should always comes with a proof of bijection.

   However, because we are able to prove the injection using the fact that the target is denumerable, we can therefore construct an injection, therefore a bijection which proves the statement about cardinality.

Corollary 2.12.1. *The denumerable union of denumerable sets is denumerable.*

Proof:

1. Suppose that $A_1, A_2, ...$ is a sequence of denumerable sets. List the elements of A_i as $a_{i1}, a_{i2}, ...$ and define
$$f : \mathbb{N} \times \mathbb{N} \to A = \bigcup A_i$$
$$(i, j) \to a_{ij}$$

2. *Comment:* This proof is basically taking the denumerable set into another dimension by making the denumerable sets unioned a new dimension.

3. Clearly f is a surjection. According to the fact that $\mathbb{N}\times\mathbb{N}$ is denumerable. The composite function fg is a surjection $\mathbb{N} \to A$. (g is defined by the snakey diagonal).

Corollary 2.12.2. *The denumerable union of denumerable sets is denumerable.*

Proof:

1. $\mathbb{Q}$ is the denumerable union of the denumerable sets $A_q = \{p/q : p \in \mathbb{Z}\}$ as q ranges over $\mathbb{N}$.

Corollary 2.12.3. *For each $m \in \mathbb{N}$, the set $\mathbb{Q}^m$ is denumerable.*

Proof: Apply the induction principle.

1. If $m = 1$ then the previous corollary states that $\mathbb{Q}^1$ is denumerable.

2. Assume inductively that $\mathbb{Q}^{m-1}$ is denumerable and $\mathbb{Q}^m = \mathbb{Q}^{m-1} \times \mathbb{Q}$. From the result of the last corollary.

2.5 Continuity (Fundamental of Calculus)

Definition. The function $f : [a, b] \to \mathbb{R}$ is continuous if for each $\epsilon > 0$ if for each $\epsilon > 0$ and each $x \in [a, b]$
$$\exists\delta > 0 \text{ s.t. } t \in [a, b] \text{ and } |t - x| < \delta \to |f(t) - f(x) \leq \epsilon|$$

Theorem 2.13. *The values of a continuous function defined on an interval $[a, b]$ form a bounded subset of $\mathbb{R}$. That is, there exists m, $M \in \mathbb{R}$ such that for all $x \in [a, b]$ we have $m \leq f(x) \leq M$.*

Proof:

1. For $x \in [a, b]$, let V_x be the set
$$V_x = \{y \in \mathbb{R} : \text{ for some } t \in [a, x]\ y = f(t)\}$$
   Comment: this set means all the $f(t)$ as t varies from a to x

2. Next we let
$$X = \{x \in [a, b] : V_x \text{ is a bounded subset of } \mathbb{R}\}$$
   therefore we need to prove that $b \in X$ so that we can show that the $f([a, b])$ is bounded.

3. Clearly that $a \in X$. b is of course and upper bound of X. Thus we want to show that $b = lub(X)$

4. Now assume that $c \leq b$ to be the least upper bound for X, for the sake of contradiction. By definition of continuity at c, there is $\delta > 0$ such that

$$|x - c| < \delta \to |f(x) - f(c)| < 1$$

5. Because that c is the $lub(X)$, $\exists x \in [c - \delta, c]$. By the definition of least upper bound.

6. Now as t varies from a to c, $f(t)$ varies first in the bounded set V_x and then in the bounded set $J = (f(c) - 1, f(c) + 1)$ because that x is within δ distance from c.

7. Because the union of two bounded sets is a bounded set and it follows that V_c is bounded, thus $c \in X$.

   Comment: Here is two bounded sets are referred to as J and V_c.

8. Besides if $c < b$ then $f(t)$ continues to vary in the bounded set J for $t > c$, as the set goes to all the way to $c + \delta$. Therefore it contradicts the fact that c is an upper bound for X. Thus $c = b$. and b is the least lower bound.

9. Thus the values of f form a bounded subset of $\mathbb{R}$.

Theorem 2.14. *The values of a continuous function defined on an interval a, b from a bounded subset of $\mathbb{R}$. That is, there exists $m, M \in \mathbb{R}$ such that $\forall x \in [a, b]$, $m \leq f(x) \leq M$.*

Proof:

1. For $x \in [a, b]$, let V_x be the set of $f(t)$ varies from a to x. Expressed formally

$$V_x = \{y \in \mathbb{R} : \text{ for some } t \in [a, x], y = f(t)\}$$

2. Now we set

$$\{x \in [a, b] : V_x \text{ is a bounded set of } \mathbb{R}\}$$

   which is the set that has all the x such that $f(x)$ defined on $[a, x]$ to be a bounded set. Thus we must prove that $b \in X$.

   Comment: This setting is because we are not sure that if some element in the set $[a, b]$ have an unbounded y value.

3. We obviously know that $a \in X$ and b is an upper bound. Since X is non-empty and bounded by above, there exists a least upper bound $c \leq b$ for X.

4. Take $\epsilon = 1$ in the definition of continuity at c. $\exists \delta > 0$ such that $|x - c| < \delta \to |f(x) - f(c)| < 1$. Because c is the least upper bound for X, there exists $x \in X$ in the interval $[c - \delta, c]$.

5. Now as t varies from a to c, the value $f(t)$ varies first in the bounded set V_x and then in the bounded set $J = (f(c) - 1, f(c) + 1)$.

   Comment: This means that the part that $[a, x]$ the value is always in the set V_x. Then now after the part $[x, c]$ the value must vary in the set J because of the definition of continuity.

6. The union of two bounded sets is a bounded set and it is followed that V_c is bounded. However, if $c < b$ then $f(t)$ continues to vary in the bounded set J for $t > c$ contrary to the fact that c is the least lower bound.

   Comment: Because in the set J there also include that part that $t \in [c, c + \delta]$. Now if we pull up the same argument then we can also find an element x in the set $c, c + s$ that is also in X.

7. Thus b has the least upper bound for the set X.

Theorem 2.15. *A continuous function f defined on the interval $[a, b]$ takes on absolute minimum and absolute maximum values: For some $x_0, x_1 \in [a, b]$ and for all $x \in [a, b]$ we have*

$$f(x_0) \leq f(x) \leq f(x_1)$$

Proof:

1. Let $M = \sup f(t)$ as t varies in $[a, b]$. By the last theorem, we know that M exists. Now again consider the set
$$X = \{x \in [a, b] : \sup V_x < M\}$$
where V_x is the set of values of $f(t)$ as t varies on $[a, x]$.

   Comment: This is step is basically saying that for every set $x \in [a, b]$ such that the supremum of the set of the value that it can reach does not exceed the maximum value.

2. We split into two cases
   - $f(a) = M$. Then f takes on a maximum value and the theorem is proved.
   - $f(a) < M$. The $X \neq \emptyset$ and we can consider the least upper bound of X, say c.

     Comment: The meaning of c is the largest x such that $\sup V_x < M$. And our goal now is to prove that $f(c) = M$

3. Now by the continuity at c, we choose $\epsilon < M - f(c)$. There $\exists \delta > 0$ such that $|t - c| < \delta$ implies $|f(t) - f(c)| < \epsilon$. Thus we can conclude that $\sup V_c < M$.

   Comment: The neighborhood was chosen is to assure that all the $f(c)$ is still below the maximum M. Now by choosing ϵ, we can guarantee that $f(t) < M$ for t near c.

   Now by continuity there exists $\delta > 0$ such that $|t - c| < \delta$ the function values $f(t)$ are close to $f(c)$ and still below M.

   The least upper bound of V_c or the supremum of all values $f(t)$ for $t \in [a, c]$ must be less than M because
   - $f(t)$ is bounded above by M for all $t \in [a, c]$ and
   - Near c, $f(t)$ stays strictly below by continuity, which covers the entire region.

4. However, if $c < b$ then this implies there exist points t to the right of c at which $\sup V_c < M$. Then this contradicts the fact that c is an upper bound of such pints. Therefore $c = b$, which implies $M < M$.

5. Thus $f(c) < M$ is wrong and since $f(c) > M$ is impossible by definition of M, we choose $f(c) = M$, so f assumes a maximum at c. The situation with minima is also similar.

Theorem 2.16. ***Intermediate Value Theorem:*** *A continuous function defined on an interval $[a, b]$ takes on all the intermediate values between $f(a)$ and $f(b)$.*

Also expressed as if $f(a) = \alpha$, $f(b) = \beta$, and γ is given and $\alpha \leq \gamma \leq \beta$ then there is some $c \in [a, b]$ such that $f(c) = \gamma$.

Proof:

1. Set
$$X = \{x \in [a, b] : \sup V_x \leq \gamma\}$$
and $c = \sup X$. Just as the last proof. And of course c exists because X is nonempty and it is bounded above by b. We claim that $f(c) = \gamma$.

   Comments: This is just saying that we want to prove c is that largest element of x such that the greatest element that the function can reach within the domain a and c is γ.

2. To prove it, we eliminate the other two possibilities which are $f(c) < \gamma$ and $f(c) > \gamma$.

(a) Suppose that $f(c) < \gamma$ and now we take $\epsilon < \gamma - f(c)$. The continuity at c gives $\delta > 0$ such that $|t - c| < \delta$ implies $|f(t) - f(c)| < \epsilon$.

Comment: So basically we are trying to prove since $f(c) < \gamma$ then we can choose and interval small enough that can fit a neighborhood between the gap of $f(c)$ and γ.

So we showed that we can find a ϵ so that

$$t \in (c - \delta, c + \delta) \quad \rightarrow \quad f(t) < \gamma$$

therefore we know $c + \frac{\delta}{2} \in X$ contrary to c being an upper bound.

(b) Suppose that $f(c) > \gamma$ and take $\epsilon = f(c) - \gamma$ continuity at c gives $\delta > 0$ such that $|t - c| < \delta$ implies $|f(t) - f(c)| < \epsilon$.

Now that $c - \frac{\epsilon}{2}$ is an upper bound for X contrary to c being the least upper bound for X.

3. Now we know that because $f(c)$ is neither $< \gamma$ and $> \gamma$ we get $f(c) = \gamma$.

   Comment: How to intuitively comprehend this proof? In my opinion is just saying that we can find γ's image in the domain because we can define it as the supremum of the set X.

3 Metric Space

3.1 Metric Space

Definition. A set X, whose elements are *points* is a metric space if with any two points p and q of x there is a number $d(p, q)$ called distance between p and q.

Properties of the distance operator (any function with these three properties are called distance function or a metric):

1. $d(p, q) > 0$ if $p \neq q$; $d(p, p) = 0$
2. $d(p, q) = d(q, p)$
3. $d(p, q) \leq d(p, r) + d(r, q)$ $\forall r \in X$

Definition. By the segment (a, b) we mean the set of all real numbers $a < x < b$.

- The interval $[a, b]$ is all x such that $a \leq x \leq b$
- Half open interval $(a, b]$, $[a, b)$

Definition. If $a_i < b_1$ for $i = 1, ..., k$, the set of points $x = (x_1, ..., x_n)$ in $\mathbb{R}^k$ whose coordinates satisfies the inequalities

$$a_i \leq x_i \leq b_i$$

is a $k - cell$

- A $1 - cell$ is an *interval*
- A $2 - cell$ is an *rectangle*

Definition. If $x \in \mathbb{R}^k$ and $r > 0$, the open (closed) ball B with center at x and radius r is defined to be

$$\{y \in \mathbb{R}^k | |y - x| < (\leq) r\}$$

Definition. We call a set $E \in \mathbb{R}^k$ convex if

$$\lambda x + (1 - \lambda) y \in E$$

if $x \in E$, $y \in E$ and $0 < \lambda < 1$. For example, a open or closed ball is convex.

Definition. For a metric space, all points and sets below are elements or subset of X.

1. A neighborhood of p is the set $N_r(p)$ consisting of all q such that $d(p,q) < r$ where r is the radius of $N_r(p)$.
2. A point p is a limit point of the set E if every neighborhood of p contains a point $p \neq q$ such that $q \in E$.
3. If $p \in E$ and p is not a limit point of E, then p is called an isolated point E.
4. E is closed if every limit point of E is a point of E.
5. A point p is an interior point of E is a neighborhood N of p such that $N \subset E$.
6. E is open if every point E is an interior point of E.
7. The complement of E, denoted by (E^c) is the set of all points $p \in X$ such that $p \notin E$.
8. E is perfect if E is closed and if every point of E is a limit point of E.
9. E is bounded if there is a real number M and point $q \in X$ such that $d(p,q) < M \ \forall p \in E$.
10. E is dense in X if every point of X is a limit point of E or a point of E (or both).

Theorem 3.1. *Every neighborhood is an open set.*

Proof:

1. Consider a neighborhood $E = N_r(p)$, let q be any point of E. Then there is a positive real number h such that
$$d(p,q) = r - h$$
Comment: Because $d(p,q) < r$, there $\exists h > 0$ such that $d(p,q) = r - h$.
2. Then
$$\forall \{s \in E | d(q,s) < h\}$$
$$d(p,s) \leq d(p,q) + d(q,s) < r - h + h$$
Thus $s \in E$ and q is there an interior point of E.

Comment: $d(q,s) < h$ we attempt to prove there exists a neighborhood $N_h(q) \subset E$.

Theorem 3.2. *If p is a limit point of a set E, then every neighborhood of p contains infinitely many points of E.*

Proof:

1. Suppose there is a neighborhood N or p which contains only a finite number of points of E.
   (a) Let $q_1, ..., q_n$ be those points of $N \cap E$ and not the same as p, and find $r = \min d(p, q_m)$ with $1 \leq m \leq$.
   (b) The neighborhood $N_r(p)$ contains no point q of E such that $q \neq p$ so that p is not a limit point of p.

Corollary 3.2.1. *A finite point set has no limit point.*

Theorem 3.3. *Let $\{E_\alpha\}$ be a (finite or infinite) collection of sets E_α. Then*
$$\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha (E_\alpha)^c$$

Proof:

1. Suppose A is the LHS and B is the RHS.

2. We first prove that

$$\left(\bigcup_{\alpha} E_{\alpha}\right)^{c} \subset \bigcap_{\alpha} (E_{\alpha})^{c}$$

   (a) If $x \in A$, then $x \notin U_{\alpha} E_{\alpha}$.
   (b) That means $x \notin E_{\alpha} \quad \forall\alpha$.
   (c) Then this means $x \in E_{\alpha}^{c} \quad \forall\alpha \rightarrow x \in E_{\alpha}^{c}$, thus then we proved that $A \subset B$.

3. Now we prove that

$$\bigcap_{\alpha} (E_{\alpha})^{c} \subset \left(\bigcup_{\alpha} E_{\alpha}\right)^{c}$$

   (a) If $x \in B$, $x \in E_{\alpha}^{c} \quad \forall\alpha$.
   (b) $x \notin E_{\alpha} \forall\alpha \rightarrow x \in (U_{\alpha} E_{\alpha})^{c}$
   (c) Thus $B \subset A$

4. Therefore we proved that $A = B$.

Theorem 3.4. *A set E is open if and only if its complement is closed.*

Proof:

1. First Direction:
   (a) Suppose E^{c} is closed, choose $x \in E$, $x \notin E^{c}$ and x is not a limit point of E^{c}.
   (b) Hence there exists a neighborhood $N(x)$ such that $E^{c} \cap N$ is empty.
   (c) Therefore $N \subset E$. Thus x is an interior point of E and therefore E is open.

2. Second Direction:
   (a) Suppose E is open. Let x be a limit point of E^{c}.
   (b) Every neighborhood of x contains a point of E^{c}.
   (c) Thus x is not a_{n} interior point of E.

   Comment: Because if every neighborhood has a point in E^{c}, then the neighborhood cannot be all in E.
   (d) Now because E is open, then x is therefore in E^{c}.

   Comment: E is open then every point is an interior point, which implies that x is not an interior point $\rightarrow x \notin E \rightarrow x \in E^{c}$.
   (e) Because all the x, which is a limit point of E^{c}, is in E^{c}, hence E^{c} is closed.

Theorem 3.5. *Union and intersection of open and closed sets*

1. *For any collection $\{G_{\alpha}\}$ of open sets, $\bigcup_{\alpha} G_{\alpha}$ is open.*
2. *For any collection $\{F_{\alpha}\}$ of closed sets, $\bigcap_{\alpha} F_{\alpha}$ is closed.*
3. *For any finite collections $G_{1}, ..., G_{n}$ of open sets $\bigcap_{i=1}^{n} G_{i}$ is open.*
4. *For any finite collections $F_{1}, ..., F_{n}$ of open sets $\bigcup_{i=1}^{n}$ is closed.*

Proof:

1. If we put $G = \cup_{\alpha} G_{\alpha}$. Now if $x \in G_{\alpha}$ for some α. Because x is an interior point of G_{α} thus x must also be an interior point of x. Thus (a) was proved.

2. By theorem 7.3, we know that

$$\left(\bigcup_{\alpha} F_{\alpha}\right)^{c} \subset \bigcap_{\alpha} (F_{\alpha})^{c}$$

3. Because the fact that F_α is open (because F_α is closed). thus

$$\bigcup_\alpha (F_\alpha^c) \text{ is open } \rightarrow \bigcap_\alpha (F_\alpha)^c \text{ is open}$$

   thus (b) was proved.

4. Now, let $H = \cap_{i=1}^n G_\alpha$. For any neighborhood, there exists N_i of x with radius r_i. $N(x_i) \subset G_i$. (By the property of open set since every neighborhood is an open set).

   Now put $r = \min(r_1, ..., r_n)$ (find the smallest radius so can fit all the sets).

5. Now let N be the neighborhood of x with radius r. Then $N \subset G_i$ for $i = 1, ..., n$ so $N \subset H$, H is therefore open. Therefore (c) was proved.

6. Because it is finite union and intersection, (d) can be proved from (c) by taking complement.

Definition. If X is a metric space and $E \subset X$, then if E' denote the set of limit point of E. The closure of E is $\bar{E} = E \cup E'$.

Theorem 3.6. *If X is a metric space and $E \in X$, then*

1. *$\bar{E}$ is closed*
2. *$E = \bar{E}$ iff E is closed*
3. *$\bar{E} \subset F$ for every closed set $F \subset X$ s.t. $E \subset F$.*

In other word, $\bar{E}$ is the smallest subset of X that contains E.

Proof:

1. (a): If $p \in X$ and $p \notin \bar{E}$ then p is neither a point nor a limit point of E.

   Hence, E has a neighborhood that doesn't interact E, $\bar{E}^c$ is therefore open, $\bar{E}$ is closed,

2. (b): If $E = \bar{E}$ then E is closed (as (a) implies $\bar{E}$ is closed).Then $E \subset E'$ hence $\bar{E} = E$.

3. (c): If F is closed and $E \subset F$. Then first $F' \subset F$ and hence $E' \subset F$.

   Comment: Since $E \subset F$, then $F' \subset E'$ because every limit point of E is also a limit point of F thus $E' \subset F' \subset F$ as $E \subset F, E' \subset F, E \subset F$.

Theorem 3.7. *Let E be a non-empty set of all real numbers which is bounded above.*

Let $y = \sup E$, then $y \in \bar{E}$ if E is closed or $y \in \bar{E}$

Proof:

1. If $y \in E$, then $y \in \bar{E}$, of course.

2. Now suppose $y \notin E$, then for every $h > 0$, there exists a point $x \in E$ such that $y - h < x < y$. (Otherwise: $y - h$ would be an upper bound of E).

3. Therefore y is a limit point of E, hence $y \in \bar{E}$.

   Comment: The main idea is to use density to prove the point y is a limit point.

Remark. Suppose $E \subset Y \subset X$, where X is a metric space and suppose E is open relative to X. (To say that E is an open subset of X means for each $p \in E$, there is an associated neighborhood in E).

Because any subset of a metric space is a metric space, then Y is a metric space, therefore E is also open relative to Y.

Theorem 3.8. *Suppose $Y \subset X$. A subset E of Y open relative to Y if and only if $E = Y \cap G$ for some open subset G of X.*

Proof:

1. *Comment:* Like the example (a,b) is open in $\mathbb{R}^1$, but not open in $\mathbb{R}^2$. This theorem states the condition how can a set open relative to a subset of a set but not open to the set itself.

2. First Direction:

   (a) Suppose E is open relative to Y, then for each $p \in E$, $\exists r_p \in \mathbb{Q}^+$ such that

   $$d(p,q) < r_p \wedge q \in Y \to q \in E$$

   (Basically just mean that exists a neighborhood contains in E).

   (b) Because let

   $$V_p = \{q \in X | d(p,q) < r_p\}$$

   and now we define $G = \bigcup_{p \in E} V_p$. (In other word, we define a set that consists of all the different neighborhoods or open set containing all the points)

   Comment: The theorem is saying "some open set" now we are find an open set for it.

   (c) Obviously G is an open subset of X by theorem

      i. Every neighborhood is an open set, and

      ii. For any collection G_α, $\bigcup_\alpha G_\alpha$ is open if G_α is open.

   (d) Now because $p \in V_p, \forall p \in E$, $E \subset G \cap Y$. (Because $p \in V_p \subset G \wedge p \in E \subset Y \to E \subset G \subset Y$)

3. Second Direction:

   (a) If G is open in X and $E = G \cap Y$, then every $p \in E$ has a neighborhood such that $V_p \subset G$.

   (b) Then obviously, $V_p \cap \subset E$.

   Comment: Because $(V_p \cap Y) \subset V_p \subset E \subset G$.

3.2 Compact Sets

Definition. An open cover of a set E in a metric space X is a collection $\{G_\alpha\}$ of open subset of X so that $E \subset \cup_\alpha G_\alpha$

Definition. A subset K of a set E in a metric space X is compact if every open cover of K contains a finite subcover.

Comment: In other word, there are finitely many indices $\alpha_1, ..., \alpha_n$ such that $K \subset G_{\alpha_1} \cup ... \cup G_{\alpha_n}$. And it is obvious that every finite set is compact.

Theorem 3.9. *Suppose $K \subset Y \subset X$. Then K is compact relative to X iff K is compact relative to Y. (We can regard compact set as metric pace on their own, without paying attention to the embedded space.)*

Proof:

1. *Comment:* Different from open subsets, E can be open in Y without being open to X. We can regard compact set as metric space on their own, without paying attention to the embedded space.

2. First Direction:

   (a) By theorem 7.8, there are sets G_α open relative to X such that $V_\alpha = Y \cap G_\alpha$. Since K is compact relative to X.

   $$K \subset G_{\alpha_1} \cup ... \cup G_{\alpha_n}$$

   for some finite set $\{\alpha_1, ..., \alpha_n\}$. Since $K \subset Y$ implies

   $$K \subset V_{\alpha_1} \cup ... \cup V_{\alpha_n}$$

   because all points in K are in Y is also in G_α therefore point in $K \in Y \cap G_\alpha = V_\alpha$

(b) Thus K is compact relative to Y as there is a finite subcover.

3. Second Direction:

    (a) If K is compact relative to Y, let $\{G_\alpha\}$ be a collection of open subsets of X that covers K.
    (b) Put $V_\alpha = Y \cap G_\alpha$ then K is contained by a finite subcover for some chosen $\alpha_1, ..., \alpha_n$.
    (c) K is also compact in X.

Theorem 3.10. *Compact subsets of metric spaces are closed.*

Proof:

1. Logic of the proof: Let K be a compact subset of metric space X. We prove the complement of K is an open subset of X.

2. Suppose that $p \in X$ and $p \notin K$ then let V_q and W_q be neighborhood of p and q be neighborhood of p and q of radius less than $\frac{1}{2}d(p,q)$.

3. Because K is compact, there are finitely many point $q_1, ..., q_n$ such that

$$K \subset W_{q_1} \cup ... \cup W_{q_n} = W$$

    Comment: We use $\frac{1}{2}d(p,q)$ so that the two neighborhoods doesn't intersect.

4. If $V = V_{q_1} \cap ... \cap V_{q_n} \to V$ is a neighborhood that doesn't intersect W. Thus we find a way to set up a neighborhood that does not intersect any open cover so that doesn't intersect the whole union by doing intersection.

5. Thus p is an interior point of K^c thus K^c is open and thus K is closed.

Theorem 3.11. *Closed subsets of compact sets are compact.*

Proof:

1. Suppose $F \subset K \subset X$, F is closed and K is compact.

2. Let $\{V_\alpha\}$ be an open cover of F. Therefore if F^c is adjoined to $\{V_\alpha\}$, we obtain an open cover Ω of K (because since F is closed, F^c is open), which is just $V_\alpha \cup F^c$, we call it Φ.

3. Because K is compact, there is a finite subcover of Ω that covers K, hence if F^c is a member of Φ, we can remove it and retain an open cover of F (since F^c by definition does not contain anything in F so that we can remove it).

4. We thus have shown that there is a finite subcover that covers F, thus F is compact.

5. "Closed" is crucial is this proof because if F is open, then F^c is closed then $\{V_\alpha\} \cup F^c$ is not an open cover anymore.

Corollary 3.11.1. *If F is closed, K is compact, then $F \cap K$ is compact.*

Proof:

1. Because of theorem 7.5 which states that any intersection of closed sets are closed and theorem 7.10 which states compact subsets of metric spaces are closed.

2. Then K is closed because K is a metric space and compact and $K \subset X$. thus $F \cap K$ is closed.

3. Then by the last theorem, which states that closed subsets of compact sets are compact. Since $F \cup K \subset K$ and since K is compact therefore $F \cap K$ is compact.

Theorem 3.12. *If $\{K_\alpha\}$ is a collection of compact subsets of a metric space X such that every subcollection is non-empty. $\cap K_\alpha$ is non-empty.*

Proof:

1. Fix a member K_1 of $\{K_\alpha\}$ and put $G_\alpha = K_\alpha^c$

2. Assume for contradiction that no member of K_1 belongs to every K_α.

   Comment: We can assume this because K_1 can be covered by the rest of K_α, therefore removing K_1 from the collection makes no difference.

3. Then G_α for an open cover of K_1 - every point was contained by a G_α, because there is a K_α not contained.

4. Because K_1 is compact, there are finitely many indices $\alpha_1, ..., \alpha_n$ that $K_1 \subset G_{\alpha_1} \cup ... \cup G_{\alpha_n}$ then by the theorem 7.5, we arrive at

$$K_1 \cap K_{\alpha_1} ... \cap K_{\alpha_n}$$

   is not empty and contradicts to our hypothesis that no member of K_1 belong to every K_α.

5. Thus $\exists p \in K_1$ belong to every subset, thus $\bigcup_\alpha K\alpha$ non-empty.

Corollary 3.12.1. *If $\{K_n\}$ is a sequence of non-empty compact set K such that $K_{n+1} \subset K_n$ then $\cap_1^\infty K_n$ is not empty.*

Comment: The same proof can be used, as K_1 can be seen as the K_1 in the last proof.

Theorem 3.13. *If E is an infinite subset of a compact set K, then E has a limit point in K.*

1. If no point of K were a limit point of E. Then for each $q \in K$, there is a neighborhood V_q that contains at most 1 point (by the negation of definition of limit point).

2. There is no finite subcollection of $\{V_q\}$ can cover E and the same is true of K since $E \subset K$.

   Comment: This is because there is no subcollection that can cover infinitely many points.

3. Thus it does not match the hypothesis that K is compact.

Theorem 3.14. *If $\{I_n\}$ is a sequence of intervals in $\mathbb{R}^1$ such that $\ldots \subset I_{n+1} \subset I_n$ then $\cap_1^\infty I_n \neq \emptyset$*

Proof:

1. If $I_n = [a_n, b_n]$, let E be the set of all a_n.

2. We know that E is nonempty and E is bounded from above. So:

   (a) Let $x = \sup E$: if $m\&n$ are positive integers, then

$$a_n \leq a_{m+n} \leq b_{m+n} \leq b_m$$

   This is because $[a_{m+n}, b_{m+n}] \subset [a_m, b_m]$

   (b) Thus $x = \sup E \leq b_m \forall m$. Obviously $a_m < x$, $x \in I_m \forall m = 1, 2, 3, ...$

Theorem 3.15. *Let k be a positive integer. If $\{I_n\}$ is a sequence of k-cells such that $I_{n+1} \subset I_n$, then $\bigcap_1^\infty I_n \neq \emptyset$.*

Proof:

1. Let I_n consist of all points $\vec{x} = (x_1, ..., x_n)$ such that

$$a_{n,j} \leq x_j \leq b_{n,j} \quad (1 \leq j \leq k)$$

   Comments: x_j is all the jth coordinates in the $k - cell$

2. Now put $I_{n,j} = [a_{n,j}, b_{n,j}]$

   (a) For every j, the sequence $I_{n,j}$ satisfies the last proof, which is in $\mathbb{R}^1$.

   (b) Hence, there are real numbers x_j^* that $(1 \leq j \leq k)$ such that

$$a_{n,j} \leq x_j^* \leq b_{n,j}$$

   (c) Now set $\vec{x}^* = (x_1^*, ..., x_k^*)$ therefore $\vec{x}^* \in I_n$ for $n = 1, 2, 3, ...$

3. The analogous proof moves from $\mathbb{R}^1$ to $\mathbb{R}^n$ smoothly.

Theorem 3.16. *Every k-cell is compact.*

Proof:

1. Let I be a k-cell and $x = (x_1, ..., x_k)$ such that $a_j \leq x \leq b_j$ and $1 \leq j \leq k$. Thus we can easily prove that
$$\delta = \left(\sum_{i=1}^{k} |b_j - a_j|^2 \right)^{\frac{1}{2}}$$
thus $|x - y| \leq \delta$ if $x, y \in I$.

2. For the sake of contradiction, assume there is an open cover $\{G_\alpha\}$ with no finite subcover and we put
$$c_j = \frac{a_j + b_j}{2}$$
which is the midpoint of (a_j, b_j) the interval $[a_j, b_j]$ and $[c_j, b_j]$ determines 2^k k-cells in $\mathbb{R}^k$, and we call then Q_i.

3. Because there is at least one Q_i that does not have a finite subcover. Then, we can repeat the subdivision, which was given by
   (a) $I_n \subset I_{n-1} \subset I_(n-2) \subset \dots$
   (b) I_n is not covered by a finite subset of G_α.
   (c) $x, y \in I_n$ then $|x - y| \leq 2^{-n}\delta$, which is going to 0.

4. By repeating the subdivision we have a sequence of nested interval and since $I_n \neq \emptyset$, there $\exists x^* \in \bigcap I_n$, by the theorem 7.12.

5. Now that G_α cover I and $x^* \in I_n$ so that
$$x^* \in G_{\alpha_0}$$
Because the fact that G_{α_0} is an open set, we an find some ball around x^* at a radius $\epsilon > 0$.

6. However, if $2^{-n}\delta < \epsilon$, then we know that I_n is entirely contained in the ϵ ball around x^*. Thus I_n was entirely covered by G_{α_0} it has a finite subcover! Because the set can be covered by a single open set, therefore forms a contradiction.

Theorem 3.17. *The following three properties are equivalent in a set $R \in \mathbb{R}^k$*

1. *E is closed and bounded.*
2. *E is compact.*
3. *Every infinite subset of R has a limit point in E.*

Proof:

1. (1) $\rightarrow$ (2) : If E is a bounded set, then it must be contained in some k-cell I. Then we know that E is closed, by theorem 7.16 I is compact.

   Additionally by theorem 7.11, we know that closed subset of compact set is also compact, thus E is also compact.

2. (2) $\rightarrow$ (3) was already proved by the theorem 7.13.

3. (3) $\rightarrow$ (1): suppose E is not bounded we can pick $x_k \in E$ and that $|x_k| > n$. But that $\{x_k\}$ is an infinite and has no limit point in $\mathbb{R}^k$. Thus none of them is in E.

   Comment: This is because each point are at least distance 1 away, there cannot infinite point around the same point.

4. *Comment:* Up until now, we have proved that if E is not bounded then it has not limit point, the contrapositive argument of the half of the statement we are trying to prove.

   Again we suppose E is not closed. We know that $\exists x_0 \in \mathbb{R}^k$ that is a limit point of E and $x_0 \notin E$. Then for $n = 1, 2, 3, \ldots$ there $\exists x_n \in E$ such that

$$|\overrightarrow{x_n} - \overrightarrow{x_0}| < \frac{1}{n}$$

5. Now we let S be the set of all x_n the S is infinite and S has x_0 as a limit point because

$$\lim_{n\to\infty} \frac{1}{n} = 0$$

   and we also know that x_0 is the only limit point in $\mathbb{R}^k$, below is a simply proof:

6. Now, for if $\overrightarrow{y} \in \mathbb{R}^k$, $\overrightarrow{y} \neq \overrightarrow{x_0}$, then

$$\begin{aligned} |\overrightarrow{x_n} - \overrightarrow{y}| &\geq |\overrightarrow{x_0} - \overrightarrow{y}| - |\overrightarrow{x_n} - \overrightarrow{x_0}| \\ &\geq |\overrightarrow{x_0} - \overrightarrow{y}| - \frac{1}{n} \geq \frac{1}{2}|x_0 - y| \end{aligned}$$

   if n is big enough

   Comment: we can assume that $\frac{1}{n} \leq \frac{1}{2}|x_0 - y|$ as n can be very big. Thus we know that every point of x_n is at least $\frac{1}{2}|x_0 - y|$. Thus every point cannot be a limit point of E.

   In a more straightforward way, every point is so close to x_0 so that not a point that is closed to y, a point in E. Therefore there cannot be a limit point that in E that fulfills the property. Thus E must be closed if every subset of E has a limit point in E.

Theorem 3.18. *Every bounded infinite subset of $\mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.*

Proof:

1. Being bounded, the set E in question is a subset of a k-cell $I \subset \mathbb{R}^k$.
2. By the theorem 7.17, since I is compact therefore E has a limit point in I.
3. By theorem 7.13, E therefore has a limit point in I.

Definition. We give a new definition of continuity: let

$$f : (x, dx) \to (y, dy)$$

a map of metric spaces is called continuous if for every open set $U \subset Y$ the set $f^{-1}(U)$ is open in X.

Lemma 3.19. *$f : (X, dx) \to (Y, dy)$ is continuous if and only if for every closed set $C \subset Y$ $f^{-1}(C) \subset X$ is closed.*

3.3 Perfect Sets and Connected Sets

Definition. E is closed and very point of E is a limit point then E is a perfect set.

Theorem 3.20. *Let P be a non-empty perfect set in $\mathbb{R}^k$. P is uncountable.*

Proof:

1. Because P has limit points. P must be infinite. Now suppose P is countable then denote the points of P, $x_1, \ldots, x_n, \ldots$. We construct a sequence of neighborhood as following
   (a) Let V_1 be any neighborhood of $\overrightarrow{x_1}$.
   (b) If V_1 consist of all $y \in \mathbb{R}^k$ such that $|\overrightarrow{y} - \overrightarrow{x_1}| < r$, then the closure $\overline{V_1}$ of V is the set of all $\overrightarrow{y} \in \mathbb{R}^k$ such that $|\overrightarrow{y} - \overrightarrow{x_1}| \leq r$.

2. Now suppose V_n has been constructed , so that $V_n \cap P$ is not empty because every point of P is a limit point of P, there $\exists$ a neighborhood V_{n+1} such that
   (a) $\overline{V_{n+1}} \subset V_n$, means that there is a point not x_n in the neighborhood with its own neighborhood in V_n
   (b) $x_n \notin \overline{V_{n+1}}$
   (c) $V_{n+1} \cap P \neq \emptyset$ because every neighborhood of x_{n+1} has a point in P distinct from x_{n+1}, by the property of perfect set.
3. Now by the third condition, V_{n+1} satisfy out induction hypothesis, the construction can proceed. Now put $K_n = \overline{V_n} \cap P$. Because $\overline{V_n}$ is closed an bounded, it is also compact by the last theorem. Then
$$x_n \notin K_{n+1} \to \cap_1^\infty K_n = \emptyset$$
but each K_n is nonempty and $K_{n+1} \subset K_n$. this contradicts the theorem 7.12 which states the intersection of compact sets are also compact.

Definition. Two subsets A and B of a metric space X are separated if both $A \cap \bar{B} = \emptyset$ and $\bar{A} \cap B = \emptyset$.

Definition. A set $E \in X$ is connected if E is not a union of 2 non-empty separated sets.

Remark. Separated sets are disjoint, but disjoint sets might not be separated.

Theorem 3.21. *A subset E of the real line $\mathbb{R}^1$ is connected iff it has the property: if $x \in E$, $y \in E$ and $x < z < y$, then $z \in E$.*

Proof:

1. First direction: Suppose that there exists $x \in E$, $y \in E$ and some $z \in (x, y)$ such that $z \notin E$ then $E = A_z \cup B_z$, where
$$\begin{cases} A_z = E \cap (-\infty, z) \\ B_z = E \cap (z, \infty) \end{cases}$$
2. Because that $x \in A_z$ and $y \in B_z$, A and B are nonempty, because $A_z \subset (-\infty, z)$ and $B_z \subset (z, \infty)$, which are both separates sets, A_z and B_z are separated.
3. Second direction: Now pick $x \in A$, $y \in B$ and assume $x < y$. Define
$$z = \sup(A \cap [x, y])$$
obviously we know that $z \in \overline{A}$ then $z \notin B$ as they are separated. Therefore $x \leq z < y$
   (a) If $z \notin A$ then $x \leq z < y$ and $z \notin E$ because $E = A \cup B$.
   (b) If $z \in A$ then $z \notin B$ since $A \cap \overline{B} = \emptyset$ so there $\exists z_1$ such that $z \leq z < y$ and $z_1 \notin E$ because $z_1 \notin B$ and $z = \sup(A \cap [x, y])$ while z is in both A and $[x, y]$.

   Comment: the point exists because there $\overline{B}$ is a closed set thus $\overline{B}^c$ is an open set, therefore has a neighborhood in itself.

4 Numerical Sequences and Series

4.1 Sequence in Metric Space

Definition. A sequence $\{p_n\}$ in a metrix space X is said to converge if there is a point $p \in X$ with the property:
$$\forall \epsilon > 0, \exists N \in \mathbb{N} \text{ such that } n \geq N \to d(p_n, p) < \epsilon$$
To denote this convergent, we usually write $p_n \to p_n$ and $\lim_{n\to\infty} p_n = p$. If a sequence does not converge, then it diverge.

Definition. The set of all points p_n is the range of $\{p_n\}$. The range of a sequence may a finite set or may be infinite.

The sequence $\{p_n\}$ is said to be bounded if its range was bounded.

Theorem 4.1. *Let $\{p_n\}$ be sequence in a metric space X.*

1. *$\{p_n\}$ converges to $p \in X$ if and only if every neighborhood of p contains p_n for all but finitely many n. (It means there are only finitely many points not contained).*
2. *If $p \in X$, $p' \in X$, and if $\{p_n\}$ converges to p and p' then $p' = p$. (Limit is unique for all sequence).*
3. *If $\{p_n\}$ converges then $\{p_n\}$ is bounded.*
4. *If $E \subset X$ and if p is a limit point of E, then there is a sequence $\{p_n\}$ in E such that $p = \lim p_n$*

Proof:

1. For (a), suppose $p_n \to p$ and let V be a neighborhood of p. For some $\epsilon > 0$, the conditions $d(q, p) < \epsilon$ and $q \in X$ implies $q \in V$.

   Comment: This just simply means that neighborhood is a set that has some boundary. This step is to lead to step step which states no matter how close we get to the limit we can always see a point in the sequence.

2. Now corresponding to this ϵ, there exists N such that $n \geq N$ implies $d(p_n, p) < \epsilon$. Thus $n \geq N$ implies $p_n \in V$.

   Comment: Therefore we can build the whole sequence by building infinite neighborhood and of course, a sequence has infinitely point. Thus we proved the first direction.

3. Conversely, suppose every neighborhood of p contains all but finitely many of the p_n. We fix $\epsilon > 0$, let V be the set of all $q \in X$ such that $d(p, q) < \epsilon$. By assumption, there exists N such that $p_n \in V$ if $n \geq N$.

   Comment: This step uses the assumption that there should be infinite amount of points of $\{p_n\}$ that are in the set V, therefore we can get to the position of the last point not in V and the rest of then are all in V.

4. Therefore $d(p_n, p) < \epsilon$ if $n \geq N$ hence $p_n \to p$. The last step's deduction fulfills the definition of convergence.

1. For (b), let $\epsilon > 0$ be given. There exist integer N, N' such that

$$n \geq N \to d(p_n, p) < \frac{\epsilon}{2}$$

$$n \geq N' \to d(p_n, p') < \frac{\epsilon}{2}$$

2. Hence if $n \geq max(N, N')$ we have

$$d(p, p') \leq d(p, p_n) + d(p_n, p') < \epsilon$$

   Because ϵ is arbitrary we can make it infinitely small, therefore we can say that $d(p, p') = 0$. *Comment:* This proof is saying that

1. For (c), suppose $p_n \to p$ there is an integer N such that $n > N$ implies $d(p_n, p) < 1$ (pick $\epsilon = 1$).
2. Now put

$$r = \max\{1, d(p_1, p), ..., d(p_N, p)\}$$

   then $d(p_n, p) \leq r$ for $n = 1, 2, 3$.

   Comment: Because we know that after the index N then every point is within distance 1 range, then there might be point before index N such that the distance might be larger than 1. Therefore taking the maximum includes all the points and it makes sense because it is just finite amount of elements.

1. For every positive integer n, there is a point $p_n \in E$ such that $d(p_n, p) < \frac{1}{n}$.

2. Given that $\epsilon > 0$, choose N so that $N\epsilon > 1$. Now if $n > N$ if follows that $d(p_n, p) < \epsilon$. Hence $p_n \to p$

   Comment: This is because when $N\epsilon > 1$, for every $n \geq N$ we have that $\epsilon > \frac{1}{n}$. Thus completes the proof.

Theorem 4.2. *Suppose $\{s_n\}$ and $\{t_n\}$ are complex sequences and $\lim_{n\to\infty} s_n = s$ and $\lim_{n\to\infty} t_n = t$. Then*

$$\lim_{n\to\infty} (s_n + t_n) = s + t \tag{1}$$

$$\lim_{n\to\infty} (cs_N = cs) \tag{2}$$

$$\lim_{n\to\infty} s_n t_n = st \tag{3}$$

$$\lim_{n\to\infty} \frac{1}{s_n} = \frac{1}{s} \tag{4}$$

We will neglect the proof for (1) and (2) becasue that is just usual 2ϵ and $N\epsilon$ method.

1. For (c), we can use the identity

$$s_n t_N - st = (s_n - s)(t_n - t) + s(t_n - t) + (s_n - s)$$

   Given that $\epsilon > 0$ there are integers N_1 and N_2 such that

$$\begin{cases} n \geq N_1 \text{ implies } |s_n - s| < \sqrt{\epsilon} \\ n \geq N_2 \text{ implies } |t_n - t| < \sqrt{\epsilon} \end{cases}$$

   If we take the $N = \max(N_1, N_2)$, $n \geq N$ implies

$$|(s_n - s)(t_n - t)| < \epsilon$$

   Therefore we know that

$$\lim_{n\to\infty} (s_n - s)(t_n - t) = 0$$

   Because of course $(t_n - t)$ and $(s_n - s)$ both converge, all three terms of the sequence converge, therefore we can get the sum also converge, thus

$$\lim_{n\to\infty} (s_n t_n - st) = 0$$

   thus

$$\lim_{n\to\infty} s_n t_n = st$$

2. For (d), choosing m such that $|s_n - s| < \frac{1}{2}|s|$ if $n \geq m$, we can see that

$$|s_n| > \frac{1}{2}|s| \quad n \geq m$$

   Comment: This is the case because of triangular identity $|s_n| - |s| < |s_n - s| < \frac{1}{2}|s|$ thus we know the upper identity.

   Given that $\epsilon > 0$ there is an integer $N > m$ such that $n \geq N$ implies

$$|s_n - s| < \frac{1}{2}|s|^2\epsilon$$

   Hence we can that for $n \geq N$

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

Theorem 4.3. *1. Suppose $\vec{x}_n \in \mathbb{R}^k \quad (n = 1, 2, 3)$ and*

$$\overrightarrow{x_n} = (\alpha_{1,n}, ..., \alpha_{k,n})$$

then $\{x_n\}$ converge to $\vec{x} = \{\alpha_1, ..., \alpha_k\}$ if and only if

$$\lim_{n\to\infty} \alpha_{j,n} = \alpha_n \quad (1 \leq j \leq k)$$

2. *Suppose $\{x_n\}$ and $\{y_n\}$ are sequences in $\mathbb{R}^k$, then $\{\beta_n\}$ is a sequence of real numbers, and $\overrightarrow{x_n} \to x$, $\overrightarrow{y_n} \to y$, $\beta_n \to \beta$ then*

$$\lim_{n\to\infty} \vec{x} + \vec{y} = \vec{x} + \vec{y} \qquad \lim_{n\to\infty} \overrightarrow{x_n}\overrightarrow{y_n} = \vec{x}\,\vec{y} \qquad \lim_{n\to\infty} \beta_n \overrightarrow{x_n} = \beta\vec{x}$$

This proof is easy and trivial.

4.2 Subsequences

Definition. Given a sequence $\{p_n\}$, consider a sequence $\{n_k\}$ of positive integer such that it is monotonically increasing, then this sequence $\{p_{n_i}\}$'s limit is a subsequential limit of $\{p_n\}$.

It is obvious that $\{p_n\}$ converges to p if and only if every subsequence of $\{p_n\}$ converge to p.

Theorem 4.4. *This theorem discusses sequence and subsequence in compact metric space.*

1. *If $\{p_n\}$ is a sequence in a compact metric space X then some subsequence of $\{p_n\}$ converges to a point of X.*
2. *Every bounded sequence in $\mathbb{R}^k$ contains a convergent subsequence.*

Proof:

1. For (a), let E be the range of $\{p_n\}$. If E is infinite then there is a $p \in E$ and a sequence $\{n_i\}$ with $n_1 < n_2 < ...$ such that
$$p_{n_1} = p_{n_2} = ... = p$$
due to the pigeon hole principle. Hence the subsequence $\{p_{n_i}\}$ so obtained converges evidently to p.
2. Now if E is infinite then theorem 7.13 suggests that if E is infinite subset of a compact set K then E has a a limit point in $p \in K$.

   Now choose n_1 so that $d(p, p_{n_1}) < 1$ having chosen $n_1, ..., n_{i-1}$, then by theorem 7.2 we know that in a neighborhood of a limit point there are infinite many points. Therefore we can choose $n_1, ..., n_{i-1}$ such that $d(p, p_{n_i}) < \frac{1}{i}$ then $\{p_{n_i}\}$ converges to p.

   Comment: This step basically builds a subsequence that obeys the identity we want using the neighborhood method and using theorem 7.2.
3. For (b), because the theorem 7.17 has states that every bounded subset of $\mathbb{R}^k$ lies in a compact subset of $\mathbb{R}^k$.

Theorem 4.5. *The subsequential limits of sequence $\{p_n\}$ in a metric space form a closed subset of X.*

Proof:

1. Let E^* be the set of all subsequential limit of $\{p_n\}$ and let q be a limit point of E^* we have to show taht $q \in E^*$.
2. Now we choose n_1 so that $p_{n_1} \neq q$ (if there is not such as point then there is only one point in E^* then there is nothing to prove). Put $\delta = (q, p_{n_1})$.
3. Suppose that $n_1, ..., n_{i-1}$ are chosen, because q is limit point of E^* there is an $x \in E^*$ such that $d(x, q) < \frac{\delta}{2^i}$ and because $x \in E^*$, there is an $n_i > n_{i-1}$ such that $d(x, p_{n_1}) < \frac{\delta}{2^i}$.

   Remark: This step there is $n_{i+1} > n_i$ is quite important because we need to ensure it is a subsequence with increasing index.
4. Therefore we know that
$$d(q, p_{n_i}) \leq d(q, x) + d(x, P_{n_i}) < 2^{-i}\delta + 2^{-i}\delta < 2^{i-1}\delta$$
for $i = 1, 2, 3, ...$ this says that $\{p_{n_i}\}$ converges to q. Hence $q \in E^*$.

   Comment: This is a very importnat proof logic similar to the two ϵ proof, because A is very close B and B is very close to C, there A is also very close C. This proof is basically saying that because there are many limit point besides me and for each limit point there are also point in a sequence that is very close to it. We built a sequence using these points.

4.3 Cauchy Sequences

Definition. A sequence $\{p_n\}$ in a metric space X is said to be a Cauchy sequence if for every $\epsilon > 0$ there is a integer N such that $d(p_n, p_m) < \epsilon$ if $n \geq N$ and $m \geq N$.

Definition. Let E be a nonempty subset of a metric space X and let S be the set of all real numbers of the form $d(p, q)$ with $p \in E$ and $q \in E$. This the supremum of S is the diameter of E denoted

$$\lim_{n\to\infty} \operatorname{diam} E_n$$

Theorem 4.6. *An obvious consequence is that $\{p_n\}$ is a sequence in X and of E_N consists of the point $p_N, p_{N+1}, p_{N+2}, \ldots$ it is clear from the two preceding definition that $\{p_n\}$ is a Cauchy sequence if and only if*

$$\lim_{n\geq\infty} \operatorname{diam} E_N = 0$$

I will not give a proof because it is obvious.

Theorem 4.7. *Below are some properties of diameter of a set*

1. *If $\overline{E}$ is the closure of a set E in a metric space X then*

$$\operatorname{diam} \overline{E} = \operatorname{diam} E$$

2. *If K_n is a sequence of compact sets in X such that $K_n \supset K_{n+1}$ (n=1,2,3,...) and if*

$$\lim_{n\to\infty} \operatorname{diam} K_n = 0$$

then $\cap_1^\infty K_n$ consists of only one point.

Proof:

1. For (a), because $E \subset \overline{E}$, it is clear that

$$\operatorname{diam} E \leq \operatorname{diam} \overline{E}$$

now if we fix $\epsilon > 0$, and choose $p \in \overline{E}$ and $q \in \overline{E}$, by the definition of $\overline{E}$ there are points p', q' in E such that $d(p, p') < \epsilon$, $d(q, q') < \epsilon$.

2. Hence

$$\begin{aligned} d(p, q) &\leq d(p, p') + d(p', q') + d(q', q) \\ &\leq 2\epsilon + d(p, p') \leq 2\epsilon + \operatorname{diam} E \end{aligned}$$

Comment: The logic for this proof can be though of as if A is close to B, C is close to D while there is some distance between A and C has a certain distance, then by triangular identity the distance between B and D ought to be larger than A and C.

3. Now suppose p, q are two points that has the distance $\operatorname{diam} \overline{E}$ then that means

$$\operatorname{diam} \overline{E} \leq 2\epsilon + \operatorname{diam} E$$

because ϵ is arbitrary, we can see that $\operatorname{diam} \overline{E} \leq \operatorname{diam} E$

4. For (b) we can put $K = \cap_1^\infty K_n$, then by theorem 7.12, we know that K is not empty.

5. Now if K contains more than one point, then $\operatorname{diam} K > 0$. But for each n, $K_n \supset K$ so that $\operatorname{diam} K_n \geq K$. This contradicts the assumption that $\operatorname{diam} K_n \to 0$

Theorem 4.8. *Below are some properties of Cauchy sequences in a metric space*

1. *In any metric space X, every convergent sequence is a Cauchy sequence.*

2. *If X is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X then $\{p_n\}$ converges to some point of X.*

3. *In $\mathbb{R}^k$, every Cauchy sequence converges.*

Proof:

1. For (a), if $p_n \to p$ and if $\epsilon > 0$ there is an integer N such that $d(p, p_n) < \epsilon \, \forall n \geq N$. Hence

$$d(p_n, p_m) \leq d(p_n, p) + d(p, p_m) < 2\epsilon$$

using triangular identity. As long as $n \geq N$ and $m \geq N$ then this $\{p_n\}$ is a Cauchy sequence.

2. For (b), let $\{p_n\}$ be a Cauchy sequence in the compact metric space X. For $N = 1, 2, 3, ...$ and let E_N be the set consisting of $p_N, p_{N+1}, p_{N+2}, ...$ then

$$\lim_{N\to\infty} \operatorname{diam} \overline{E}_N = 0$$

By the definition from definition of diameter and the natural of Cauchy sequence and by the last theorem, we know the diameter of a set is the same as the diameter of the closure of the set.

3. Being a closed subset of the compact metric space X, each $\overline{E}_N$ is also a compact set.

4. Also because $E_N \supset E_{n+1}$ so that we know that $\overline{E}_N \supset \overline{E}_{N+1}$ by the theorem 8.7 (2), this set consist of only one point, therefore there is one point, call it p that lies in every $\overline{E}_N$. Next we prove that this is the unique limit of the Cauchy sequence.

5. Let $\epsilon > 0$, because the limit of diameter of the set is equal to 0, there is an integer N_0 such that $\operatorname{diam} \overline{E}_N < \epsilon$ when $N \geq N_0$ using the property of limit. Because the fact that $p \in \overline{E}_N$, it follows that $d(p, q) < \epsilon$ for every $q \in \overline{E}_N$.

   Comment: We basically use the property of limit for diameter to find a set that the furthest point in the set doesn't have distance farther than ϵ then of course every point is at most ϵ distance from p, which fulfill the definition of a limit.

6. Therefore $p_n \to p$ as needed.

7. For (c), let $\{\vec{x}_n\}$ is a Cauchy sequence in $\mathbb{R}^k$. Define E_N as we prove (b) with $\vec{x}_i$ in the place of p_i. For some N, $\operatorname{diam} E_N < 1$.

8. The range of $\{\vec{x}_n\}$, hence the range of $\{\vec{x}_n\}$ is the union of E_N and the finite set $\{\vec{x}_1, ..., \vec{x}_{N-1}\}$, therefore $\{\vec{x}_N\}$ was bounded. Because every bounded subset of $\mathbb{R}^k$ has compact closure in $\mathbb{R}^k$ thus (c) follows from (b).

   Comment: Actually (b) follows directly from (c) because we know every k-cell is indeed a closed set. Therefore the thing is obvious.

Definition. A metric space in which every Cauchy sequence converges is said to be complete.

Remark. Thus the last theorem says that all compact metric spaces and all Euclidean spaces are complete, and every closed subset E of a complete metric space X is complete.

Definition. A sequence $\{s_n\}$ of real numbers is said to be

1. Monotonically increasing if $s_n \leq s_{n+1}$ (n = 1,2,3,...);
2. Monotonically decreasing if $s_n \geq n_{n+1}$ (n = 1,2,3,...)

we say a sequence is monotonic if it monotonically increase or decrease.

Theorem 4.9. *Suppose $\{s_n\}$ is monotonic, then $\{s_n\}$ converge if and only if it is bounded.*

Proof:

1. Suppose $s_n \leq s_{n+1}$ the let E be the range of $\{s_n\}$. If $\{s_n\}$ is bounded let s be the least upper bound of E. Then of course

$$s_n \leq s$$

for every $\epsilon \geq 0$ there is an integer N such that $s - \epsilon < s_N \leq s$.

   Comment: By the least upper bound property, because if not then it is not the least upper bound.

2. For otherwise $s - \epsilon$ would be a upper bound of E. Since $\{s_n\}$ increases, $n \geq N$ therefore implies

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

therefore $\{s_n\}$ converges to s.

Comment: We can understand this as we constructed a sequence that is approaching s because the fact that it is the least upper bound and we made sure that it is not the same element by using the fact that it monotonically increase.

4.4 Upper and Lower limits

Definition. Let $\{s_n\}$ be a sequence of real numbers with the following property: for every M there is an integer N such that $n \geq N$ implies $s_n \geq M$. We then write $s_n \geq \infty$.

Similarly, if for every real M there is an integer N such that $n \geq N$ implies $s_n \leq M$, we write $s_n \to -\infty$.

Definition. Let $\{s_n\}$ be sequence of real numbers. Let E be the set of numbers x such that $s_{n_k} \to x$ for some subsequence $\{s_{n_k}\}$. This set E contains all subsequential limits plus possibly $\pm\infty$. Now we let

$$s^* = \sup E$$

$$s_* = \inf E$$

this two numbers are called

$$\lim_{n\to\infty} \sup s_n$$

$$\lim_{n\to\infty} \inf s_n$$

Theorem 4.10. *Let $\{s_n\}$ be a sequence of real numbers. Let E and s^* have the same meaning as in the last definition. Then s^* has the following two properties*

1. *$s^* \in E$*
2. *If $x > s^*$ there is an integer N such that $n \geq N$ implies $s_n < x$.*

Moreover s^* is the only number with the properties (a) and (b).

Proof:

1. If $s^* = +\infty$ then E is not bounded above, hence $\{s_n\}$ is not bounded above then there is a subsequence $\{s_{n_k}\}$ such that $s_{n_k} \to \infty$.
2. If s^* is real, then E is bounded from above, and at least one subsequential limit exists, so that (a) follows from the theorem 7.7 which states for a bounded set of real number, then the supremum is in the closure of the set. The theorem 8.5 which shows that the set of all subsequential limits form a closed subset.
3. Thus the supremum of limit point is contained in the closure of the set of limit points, which is itself, as needed.
4. If $s^* = -\infty$ then E contains only one element , therefore there is no subsequential limit, hence for any real M, $s_n > M$ for a most a finite number of values of n, so that $s_n \to \infty$. Therefore we proved (a).
5. Suppose for the sake of contradiction, there is a number $x > s^*$ such that $s_n \geq x$ for infinitely many values of n. Then there is number $y \in E$ such that $y \geq x > s^*$ which contradicts the definition of s^*.
6. To show the uniqueness, suppose there are two numbers, p and q which satisfy (a) and (b), and suppose $p < q$. Choose x such that $p < x < q$.
7. Suppose p satisfies (b), then there is an integer N such that for every $n \geq N \to s_n < x$, but then q cannot satisfy (a), because then there cannot be infinitely many point converging to it, therefore it cannot be a subsequential limit therefore cannot be in E.

5 Continuity

Definition. Let X and Y be metric spaces, suppose $E \subset X$, f maps E into Y, and p is a limit point of E. We write $f(x) \to q$ as $x \to p$, or

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

if there is a point $q \in Y$ with the following property:
For every $\epsilon > 0$ then there exists a $\delta > 0$ such that

$$d_Y(f(x), 1) < \epsilon$$

for all points $x \in E$ for which

$$0 < d_X(x, p) < \delta$$

where d_Y and d_X means the distances in metric space X and Y.

It is important that $p \in X$ but p need not be a point of E in the above definition.

Theorem 5.1. *Let X, Y, E, f and p be as in last definition, then*

$$\lim_{n \to \infty} f(x) = q$$

if and only if

$$\lim_{n \to \infty} f(p_n) = q$$

for every sequence $\{p_n\}$ in E such that $p_n \neq p$ and $\lim_{n\to\infty} p_n = p$

Proof:

1. Suppose that $\lim_{n\to\infty} f(x) = q$ holds, then choose $\{p_n\}$ in E that satisfies $\lim_{n\to\infty} p_n = p$.
2. Then there exists $\delta > 0$ such that $d_Y(f(x), q) < \epsilon$ if $x \in E$ and $0 < d_X(x, p) < \delta$. (Comes from the definition of continuity).
3. Also there exists some N such that $n > N$ implies $0 < d_X(p_n, p) < \delta$. Therefore for $n > N$, we have
$$d_Y(f(p_n), q) < \epsilon$$
which shows $\lim_{n\to\infty} f(p_n) = q$ by the definition of limit.
4. Conversely, suppose $\lim_{n\to\infty} f(x) = q$ is false, then there is some $\epsilon > 0$ such that for every $\delta > 0$ for which $d_Y(f(x), q) \geq \epsilon$. (This is by the contrapositive argument for the definition of continuity).
5. But $0 < d_X(x, p) < \delta$ then taking $\delta_n = \frac{1}{n}$ we thus find a sequence in E satisfying $\{p_n\}$ in E such that $p_n \neq p$ therefore $\lim_{n\to\infty} f(p_n) = q$ is false.

Corollary 5.1.1. *If f has a limit at p, then this limit is unique. Which could be proved using the theorem such that if a sequence converge to two limits, then the two limits must be the same and then using the last theorem.*

Definition. Suppose we have two complex functions, f and g, both defined on E. Then by $f + g$ we mean the function which assigns to each point x of E the number $f(x) + g(x)$. And similarly, we define the difference $f - g$, the product $f(x) \cdot g(x)$ and the quotient of the two functions $\frac{f(x)}{g(x)}$, where assuming $g(x) \neq 0$.

Theorem 5.2. *Suppose $E \subset X$ a metric space, p is a limit point of E, f and g are complex functions on E and*

$$\lim_{x \to p} f(x) = A \quad \lim_{x \to p} g(x) = B$$

1. $\lim_{x\to p}(f + g)(x) = A + B$
2. $\lim_{x\to p}(fg)(x) = AB$
3. $\lim_{x\to p} \frac{f}{g}(x) = \frac{A}{B}$

we will skip the proof because the theorem is exact the same as the addition and multiplication of the sequence proof.

Definition. Suppose X and Y are metric spaces, $E \subset X$, $p \in E$ and f maps E into Y. Then f is said to be continuous at p if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$d_Y(f(x), f(p)) < \epsilon$$

for every points $x \in E$ for which $d_X(x, p) < \delta$. If f is continuous at every point of E, the f is said to be continuous on E. It should be noted that f has to be defined at the point p in order to be continuous at p. This different from the definition of the limit of the function at the point p.

Comment: If p is an isolated point of E, then our definition implies that every functions f which has E as its domain of definition is continuous at p. For no matter which $\epsilon > 0$ we choose, we can pick $\delta > 0$ so that the only point $x \in E$ for which $d_X(x, p) < \delta$ is $x = p$ then

$$d_Y(f(x), f(p)) = 0 < \epsilon$$

Theorem 5.3. *In the situation given in the last definition, assume also that p is a limit point of E. Then f is continuous at p if and only if $\lim_{x \to p} f(x) = f(p)$.*

Proof:

1. This proof is obvious, when we compare the definition of the fact $\lim_{x \to p} f(x) = p$ when p is a limit point of E and the definition of continuity on a metric space.

2. Because the fact that p is a limit point, there is a always a point in E since the continuity states the $\epsilon - \delta$ definition of the continuity, the p is of course continuous on the domain.

Theorem 5.4. *Suppose that X, Y, Z are metric space and $E \subset X$, f maps E into Y and g maps the range of f, $f(E)$ into Z and h is the mapping of E onto Z defined by*

$$h(x) = g(f(x)) \quad x \in E$$

If f is continuous at a point $p \in E$ and if g is continuous at the point $f(p)$ then h is continuous at p.

Proof:

1. Let $\epsilon > 0$ be given, since g is continuous at $f(p)$ there exists $\eta > 0$ such that

$$d_Z(g(y), g(f(p))) < \epsilon \text{ if } d_Y(y, f(p)) < \eta \text{ and } x \in E$$

2. Because f is continuous at p there is a $\delta > 0$ such that $d_Y(f(x), f(p)) < \eta$ if $d_X(x, p) < \delta$ and $x \in E$. It follows that if $d_X(x, p) < \delta$

$$d_Z(h(x), h(p)) = d_Z(g(f(x)), g(f(p))) < \epsilon$$

therefore h is continuous at p.

Comment: This proof is basically saying that all the points in Y is close to the limit point in X and all the points in Z is close to the image of the point in Y then all the points in X is then close to the image point in Z.

Theorem 5.5. *A mapping f of a metric space X into a metric space Y is continuous on X if and only if $f^{-1}(V)$ is open in X for every open set V in Y.*

Proof:

1. Suppose f is continuous on X and V are open set in Y. We have to show that every point of $f^{-1}(V)$ is an interior point of $f^{-1}(V)$.

2. Suppose that $p \in X$ and $f(p) \in V$. Because V is open there is $\epsilon > 0$ such that $y \in V$ if $d_Y(f(p), y) < \epsilon$ (an open ball).

3. Since f is continuous at p then there exists $\delta > 0$ such that $d_Y(f(p), y) < \epsilon$ if $d_X(x,p) < \delta$. Thus we know that $x \in f^{-1}(V)$ as soon as $d_X(x,p) < \delta$. Therefore every point in the set $f^{-1}(V)$ is an interior thus the set is an open set.

4. Conversely, suppose that $f^{-1}(V)$ is open in X for every open set V in Y. Now fix $p \in X$ and $\epsilon > 0$ let V be the set of all $y \in Y$ such that $d_Y(y, f(p)) < \epsilon$.

5. Therefore V is open, hence $f^{-1}(V)$ is open hence there exists $\delta > 0$ such that $x \in f^{-1}(V)$ then $f(x) \in V$ so that $d_Y(f(x), f(p)) < \epsilon$. Which fulfills the definition of continuity in the metric space.

   Comment: This proof uses the technique of the second direction uses the proof in the first direction, which allow more flexibility of the proof.

Corollary 5.5.1. *A mapping f of a metric space X into a metric space Y is continuous if and only if $f^{-1}(C)$ is closed in X for every closed set C in Y.*

Proof:

1. The proof is trivial, because a set is closed if and only if its complement is open and since $f^{-1}(E^c) = [f^{-1}(E)]^c$ for every $E \subset Y$.

2. Since the pre-image of f is open if the image is open, then the complement of the pre-image so then the complement of the complement of the image is closed.

Theorem 5.6. *Let f and g be complex continuous functions on a metric space X then $f+g$, fg, f/g are continuous on X for every closed set C in Y.*

Proof

1. As isolated points of X is nothing to prove and at limit points the statement follows from the theorem 9.3 and theorem 9.2 and theorem 9.4, which states that the continuous functions are continuous under elementary operations and the f is continuous at p if and only if $\lim_{x\to p} f(x) = f(p)$.

Theorem 5.7. *This theorem discusses the continuity of function on metric space $\mathbb{R}^k$.*

1. *Let $f_1, ..., f_k$ be real functions on a metric space X and let f be the mapping of X onto $\mathbb{R}^k$ defined by*
$$f(x) = (f_1(x), ..., f_k(x))$$
then f is continuous if and only if each of the unctions $f_1, ..., f_k$ continuous.

2. *If f and g are continuous mapping of X into $\mathbb{R}^k$ then $f+g$, $f \cdot g$ are continuous on X. Note that the functions $f_1, ..., f_k$ are called components of f, note that $f+g$ is a mapping into $\mathbb{R}^k$ whereas $f \cdot g$ is a real function on X.*

This proof is trivial and we will neglect it.

5.1 Continuity and Compactness

Definition. A mapping f of a set E into $\mathbb{R}^k$ is said to be bounded if there is a real number M such that $|f(x)| \leq M$ for all $x \in E$.

Theorem 5.8. *Suppose f is a continuous mapping of a compact metric space X into a metric space Y then $f(X)$ is compact.*

Proof:

1. Let $\{V_\alpha\}$ be an open cover of $f(X)$. Because f is continuous. The last theorem shows that each of the sets $f^{-1}(V_\alpha)$ is open. Because X is compact, there are finitely many indices, say $\alpha_1, ..., \alpha_n$ such that
$$X \subset f^{-1}(V_{\alpha_1} \cup ... \cup V_{\alpha_n})$$

2. Because that $f(f^{-1}(E)) \subset E$ for every $E \subset Y$, we know that

$$f(X) \subset V_{\alpha_1} \cup, ..., \cup V_{\alpha_n}$$

Therefore the set could also be covered by a finite subcover, fulfill the definition of compactness.

Theorem 5.9. *If f is continuous mapping of a compact metric space X into $\mathbb{R}^k$ then $f(X)$ is closed and bounded thus f is bounded.*

Proof:

1. This proof is trivial, when considering the last theorem and the fact that in a metric space, closed and bounded lead to compactness.

Theorem 5.10. *Suppose f is continuous real function on a compact metric space X and*

$$M = \sup_{p \in X} f(p) \quad m = \inf_{p \in X} f(p)$$

then there exists points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$.

This theorem could also be stated as f attains its maximum and its minimum in its domain.

The notation means that M is the least upper bound of the set of all numbers $f(p)$ where p ranges over X, and that m is the greatest lower bound of this set of numbers.

Proof:

1. By theorem 9.8, $f(X)$ is closed and bounded set of real numbers hence $f(x)$ contains

   $$M = \sup f(X) \text{ and } m = \inf f(X)$$

   because both the supremum and infimum are limit points of the set, therefore since the set is compact the limit points are contained in it, as needed.

Theorem 5.11. *Suppose that f is a continuous $1-1$ mapping of a compact metric space X onto a metric space Y then the inverse mapping f^{-1} defined on Y by*

$$f^{-1}f(x) = x \quad (x \in X)$$

if a continuous mapping of Y onto X.

Proof:

1. Using theorem 9.5 which states that a mapping f of a metric space X into a metric space Y is continuous on X if and only if $f^{-1}(V)$ is open in X for every open set V in Y. Now because the f^{-1} is already defined, we can use the function directly.

2. Using the theorem 9.5 using f^{-1} in replacement of f, we can prove that $f(V)$ is open in Y for every open set Y in X. Now fix such a set V.

3. The complement V^c of V is closed in X hence compact hence $f(V^c)$ is compact subset of Y and so is closed in Y. Because that f is one-to-one, $f(V)$is the complement of $f(V^c)$ hence $f(V)$ is open.

   Comment: The logic of this proof is that $f(V)$ is compact, and the function is bijective, thus we can just use the openness theorem by taking the complement of a closed(compact) set.

Definition. Let f be a mapping of a metric space X into a metric space Y. We say that f is a uniformly continuous function on X if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$d_Y(f(p), f(q)) < \epsilon$$

for every p and q in X for which $d_X(p, q) < \delta$.

Remark. The difference between continuity and uniform continuity is that continuity is defined on a point or a set while the uniform continuity is only defined on a set.

The main difference is that if a function is continuous on a set then for every ϵ and for every point we can find a possibly different δ that fulfills the condition, while if a function is uniformly continuous on a set, then for every ϵ, there is a δ that satisfy the condition at every point.

Evidently, every uniformly continuous function is continuous.

Theorem 5.12. *Let f be a continuous mapping of a compact metric space X onto a metric space Y then f is uniformly continuous on X.*

Proof:

1. Let $\epsilon > 0$ because that f is continuous, we can associate to each point $p \in X$ a positive number $\phi(p)$ such that
$$q \in X, d_X(p, q) < \phi(p) \text{ implies } d_Y(f(p), f(q)) < \frac{\epsilon}{2}$$
Comment: This $\phi(p)$ is a function that was defined to determine the δ and then we can determine a single number easier.

2. Let $J(p)$ be the set of all $q \in X$ such that
$$d_X(p, q) < \frac{1}{2}\phi(p)$$
Because the fact that $p \in J(p)$, the collection of all sets $J(p)$ is an open cover of X and since X is compact there is a finite set of points $p_1, ..., p_n$ in X such that
$$X \subset J(p_1) \cup ... \cup J(p_n)$$

3. Now that we can put
$$\delta = \frac{1}{2}\min[\phi(p_1), ..., \phi(p_n)]$$
then we know that $\delta > 0$ because all points in the set are positive.

   Comment: This is why compactness is very important because if there is not a finite subcover, the minimum might as well go to zero even though all the open covers are having a positive radius.

4. Now we can let q and p be points of X such that $d_X(p, q) < \delta$, then there is an integer m such that $1 \leq m \leq n$ such that $p \in J(p_m)$ hence
$$d_X(p, p_m) < \frac{1}{2}\phi(p_m)$$

5. Hence we also have
$$d_X(q, p_m) \leq d_X(p, q) + d_X(p, p_m) < \delta + \frac{1}{2}\phi(p_m) \leq \phi(p_m)$$
this is because the two ϵ method.

6. Finally we showed that
$$d_Y(f(p), f(q)) \leq d_Y(f(p), f(p_m)) + d_Y(f(q), f(p_m)) < \epsilon$$
as needed, thus δ is the number that makes the function uniformly converge.

Theorem 5.13. *Let E be a non-compact set in $\mathbb{R}$ then we know that*

1. *There exists a continuous function on E which is not bounded.*

2. *There exists a continuous and bounded function on E which has no maximum.*

3. *If E is bounded then there exists a continuous function on E which is not uniformly continuous.*

Proof:

1. *Case:* Suppose E is bounded but not closed. Then there exists a limit point $x_0 \notin E$. Define the function
$$f(x) = \frac{1}{|x - x_0|}$$
for $x \in E$. This function is continuous on E but becomes unbounded as $x \to x_0$. Hence, there exists a continuous and unbounded function on E, proving (1).

2. Define $g(x) = \frac{1}{1+(x-x_0)^2}$ for $x \in E$. This function is continuous on E and bounded, with $0 < g(x) < 1$. Since $x_0 \notin E$, we find that $\sup g(x) = 1$, but $g(x) < 1$ for all $x \in E$, so g has no maximum. This establishes (2).

3. *Case:* If E is bounded, define $f(x) = \frac{1}{|x-x_0|}$ as before. To show f is not uniformly continuous, let $\delta > 0$ and consider points $x, t \in E$ close to x_0 with $|x - x_0|, |t - x_0| < \delta$. Then,
$$|f(t) - f(x)| = \left| \frac{1}{|t - x_0|} - \frac{1}{|x - x_0|} \right| = \frac{|t - x|}{|x - x_0||t - x_0|},$$
which can be made arbitrarily large as x and t approach x_0. Thus, f is not uniformly continuous, proving (3).

5.2 Continuity and Compactness

Theorem 5.14. *If f is continuous mapping of a metric space X into a metric space Y, and if E is connected subset of X then $f(E)$ is connected.*

Proof:

1. Assume that $f(E) = A \cup B$ where A and B are non-empty separated subsets of Y. Put $G = E \cap f^{-1}(A)$ and $H = E \cap f^{-1}(B)$.

2. Then $E = G \cup H$, and neither G nor H are non-empty.

3. Because that $A \subset \overline{A}$, we have $G \subset f^{-1}(\overline{A})$ while $\overline{A}$ is a is closed, because f is continuous.

4. We know that $\overline{G} \subset f^{-1}$, because a closed set contains all its limit points, so are its subset's limit points. Therefore we know that $f(\overline{G}) \subset \overline{A}$.

5. Because $f(H) = B$ and $\overline{A} \cap B$ is empty, we conclude that $\overline{G} \cap H$ is empty.

   Comment: This logic of the proof is basically saying if the set is separated, then because closeness of the preimage for a continuous function and the containment of the preimage, we know that preimage must also be not connected.

6. The same argument shows that $G \cap \overline{H}$ is empty. Thus we know that G and H are separated. This is impossible if E is connected.

Theorem 5.15. *Intermediate Value Theorem: Let f be a continuous real function on the interval $[a, b]$. Now if $f(a) < f(b)$ and if c is a number such that $f(a) < c < f(b)$ then there exists a point $x \in (a, b)$ such that $f(x) = c$.*

Proof:

1. By the theorem 7.21: A subset E of the real line $\mathbb{R}^1$ is connected iff it has the property: if $x \in E$, $y \in E$ and $x < z < y$, then $z \in E$. We know that $[a, b]$ is indeed connected.

2. Thus the last theorem suggests that $f([a, b])$ is connected subset of $\mathbb{R}$, and the assertion follows if we appeal once more to theorem 7.21, therefore we know that there must be some value in between $f(a)$ and $f(b)$ that attains the value of the function.

Remark. This theorem does not have a converse that states, if the function for any two points its image takes all its intermediate value between the image of the two points, then the function is continuous.

Definition. If x is a point in the domain of definition of the function f at which f is not continuous, we say that f is discontinuous at x or that f has a discontinuity at x.

Definition. Let f be defined on (a, b) consider any point x such that $a \le x \le b$, we write $f(x+) = q$. If $f(t_n) \to q$ as $t \to \infty$ for every sequence t_n in (x, b) such that $t_n \to x$.

To obtain the definition of $f(x-)$ for $a \le x \le b$ we restrict ourselves to sequences $\{t_n\}$ in (a, x). It is obvious that any point x of (a, b), $\lim_{n\to\infty} f(t)$ exists if and only

$$f(x+) = f(x-) = \lim_{t\to\infty} f(t)$$

Definition. Let f be defined on (a, b). If f is discontinuous at a point x and if $f(x+)$ and $f(x-)$ exists then f is said to have a discontinuity of the first kind, or a simple continuity. Otherwise the discontinuity is said to be of the second kind.

5.3 Monotonic Functions

Definition. Let f be real on (a, b), then f is said to be monotonically increasing on (a, b) if $a < x < y < b$ implies $f(x) \le f(y)$. If the last inequality is reversed we obtain the definition of a monotonically decreasing function.

Theorem 5.16. *Let f be monotonically increasing on (a, b), then $f(x+)$ and $f(x-)$ exists at every point of x of (a, b). More precisely*

$$\sup_{a<t<x} f(t) = f(x-) \le f(x) \le f(x+) = \inf_{a<t<b} f(t)$$

to be more precise, if $a < x < y < b$, then

$$f(x+) \le f(y-)$$

analogous results evidently hold for monotonically decreasing functions.

Proof:

1. By hypothesis the set of numbers $f(t)$ where $a < t < x$ is bounded above by the number $f(x)$. Therefore there must be a least upper bound which we denote by A and $A \le f(x)$ we need to show that $A = f(x-)$.

2. Let $\epsilon > 0$ be given. It follows from the definition of least upper bound that there exists $\delta > 0$ such that $a < x - \delta < x$ such that
$$A - \epsilon < f(x - \delta) \le A$$

3. Because that f is monotonic, we have that $f(x - \delta) \le f(t) \le A$. Therefore we know that
$$|f(t) - A| < \epsilon$$
for $x - \delta < t < x$. Therefore we know that $f(x-) = A$, because ϵ can be arbitrarily small therefore the difference could be just zero.

   Comment: The logic for this proof is that because x sets a natural upper bound for the set, therefore the monotonic increasing sequence leads the least upper bound to it.

4. The other side of the proof $f(x+)$ is proven the exact same way. Therefore the first equation was proved.

5. Now, if $a < x < y < b$, we can see that
$$f(x+) = \inf_{x<t<b} f(t) = \inf_{x<t<y} f(t)$$
$$f(y-) = \sup_{a<t<y} f(t) = \sup_{x<t<y} f(t)$$
which could be proved by the first equation easily.

Corollary 5.16.1. *Monotonic functions have no discontinuities of the second kind.*

This corollary implies that every monotonic function is discontinuous at a countable set of points at most

Theorem 5.17. *Let f be monotonic on (a, b), then the set of points of (a, b) at which f is discontinuous is at most countable.*

Proof:

1. Suppose that f is increasing and let E be the set of points at which f is discontinuous. With every point $x \in E$ we associate a rational number $r(x)$ such that

$$f(x-) < r(x) < f(x+)$$

   (because we know that by the last theorem, if the function is monotonically increasing then $f(x-) \leq f(x) \leq f(x+)$) there must be something in between.

Definition. For any real c, the set of real numbers x such that $x > c$ is called a neighborhood of $+\infty$ and is written $(c, +\infty)$. Similarly, the set $(-\infty, c)$ is a neighborhood of $-\infty$.

Definition. Let f be a real number system. If for every neighborhood U or A there is a neighborhood V of x such that $V \cap E$ is not empty and such that $f(t) \in U$ for every $t \in V \cap E$ and $t \neq x$.

6 Differentiation

Definition. Let f be defined on $[a, b]$. For any $x \in [a, b]$ form the quotient

$$\phi(t) = \frac{f(t) - f(x)}{t - x}$$

and define

$$f'(x) = \lim_{t \to x} \phi(t)$$

f' is the derivative of f. If f' is defined at a point x, we say f is differentiable on E. If f' is defined at every point of a set $E \subset [a, b]$ then we say f is differentiable on E.

Theorem 6.1. *Let f be defined on $[a, b]$. If f is differentiable at a point $x \in [a, b]$ then f is continuous at x.*

Proof:

1. As $t \to x$ we have

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

   by theorem 9.2.

2. The first term is just $f'(x)$ and the second term is 0 therefore. Therefore $f(t) - f(x) \to 0$ therefore by the definition of being continuous, we know that this is true.

Remark. The converse for this theorem does not hold.

Theorem 6.2. *Suppose f and g are defined on $[a, b]$ and are differentiable at a point x. The $f + g$, fg, $\frac{f}{g}$ are differentiable at x. To be specific, we know that*

1. $(f + g)'(x) = f'(x) + g'(x)$
2. $(fg)'(x) = f'(x)g(x) + g(x)f'(x)$
3. $(\frac{f}{g})(x) = \frac{f'(x)g(x) - g'(x)f(x)}{g^2(x)}$

the proof is pure algebra so I will skip it.

Theorem 6.3. ***Chain rule:*** *Suppose that f is continuous on $[a, b]$, we know $f'(x)$ exists at some point $x \in [a, b]$, g is defined on an interval I which contains the range of f and g is differentiable at the point $f(x)$. If*

$$h(t) = g(f(t)) \quad (a \le t \le b)$$

then h is differentiable at x, and

$$h'(x) = g'(f(x))f'(x)$$

Proof:

1. Let $y = f(x)$. Then by definition of the derivative, we have that

$$\begin{cases} f(t) - f(x) = (t - x)[f'(x) + u(t)] \\ g(s) - g(y) = (s - y)[g'(y) + v(s)] \end{cases}$$

   where u and v are the error terms such that $u(t) \to 0$ as $t \to 0$ as $s \to y$.

2. Now let $s = f(t)$ then we know that

$$\begin{aligned} h(t) - h(x) &= g(f(t)) - g(f(x)) \\ &= [f(t) - f(x)] \cdot [g'(y) + v(s)] \\ &= (t - x) \cdot [f'(x) + u(t)] \cdot [g'(y) + v(s)] \end{aligned}$$

   where the first equality used the second equation and the second equality used the first equation in the cases.

3. Therefore the last equality could be written as

$$\frac{h(t) - h(x)}{t - x} = [f'(x) + u(t)] \cdot [g'(y) + t(s)]$$

   now take the limit on both sides, we get

$$f'(t) = f'(x)g'(y)$$

   Note step this step is suppose to remove the brackets and do the product, then use theorem of computation of limit to calculate the limit, but here we just simplify the proof.

4. Letting $t \to x$, we see that $s \to y$ now by the continuity of f, we know that the right side will tend to $g'(y)f'(x)$, which gives

$$h'(x) = g'(f(x))f'(x)$$

Definition. Let f be a real function defined on a metric space X. We say that f has a local maximum at a point $p \in X$ if there is $\delta > 0$ such that $f(q) \le f(p)$ for every $q \in X$ with $d(p, q) < \delta$, local minimum is defined likewise.

Theorem 6.4. *Let f be defined on $[a, b]$, if f has a local maximum (minimum) at a point $x \in (a, b)$ and if $f'(x)$ exists, then $f'(x) = 0$.*

Proof:

1. Choose δ accordance with the last definition so that

$$a < x - \delta < x < x + \delta < b$$

2. If $x - \delta < t < x$, then $\frac{f(t)-f(x)}{t-x} \ge 0$ by the definition (because since the two points are within the δ neighborhood, on the let of x, therefore the difference is always positive.)

   Now by taking the limit, we know that $f'(x) \ge 0$.

3. On the other hand, using the same reasoning, if $x < t < x + \delta$, then

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

   which shows that $f'(x) \le 0$.

4. Because both $0 \leq f'(x) \leq 0$, $f'(x) = 0$

Theorem 6.5. *If f and g are continuous real functions on $[a, b]$ which are differentiable in (a, b) then there is a point $x \in (a, b)$ at which*

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x)$$

note that differentiability is not required at the endpoints.

What does this Mean Value Theorem, or more specifically "Cauchy Mean Value Theorem" says?

1. **Graphical Visualization**:
   - Plot the two functions $f(x)$ and $g(x)$ on the interval $[a, b]$.
   - Draw the secant lines connecting $(a, f(a))$ to $(b, f(b))$ and $(a, g(a))$ to $(b, g(b))$.
   - The theorem states that there is a point x in (a, b) where the tangent line to f and the tangent line to g at x are "proportional," meaning the slopes of the tangents at x match the ratio of the secant slopes over the interval.
2. **Special Case: Mean Value Theorem**:
   - If $g(x) = x$, this theorem reduces to the standard Mean Value Theorem, where there exists a point $x \in (a, b)$ with $f'(x) = \frac{f(b)-f(a)}{b-a}$.

Proof:

1. Put
$$h(t) = [f(b) - f(a)]g(t) - [g(b) - g(a)]f(t) \quad a \leq t \leq b$$
   then of course h is continuous on $[a, b]$ and h is differentiable in (a, b) and $h(b) = h(a) = f(b)g(a) - f(a)g(b)$. To prove this theorem we need to show that $h'(x) = 0$ for some $x \in (a, b)$.

   Comment: We define a function that attains a value that are the same on both ends. This construction is very useful for proving this theorem because this theorem is essentially saying that the two functions are having a point with the ratio of their derivative is equal to the ratio of average increment.

   Therefore having this function, the average change at both ends are 0, and hence we are find the point of the function having its derivative being zero.
2. Now if f is constant, this is going to hold for every $x \in (a, b)$.
3. Now if $h(t) > h(a)$ for some $t \in (a, b)$ at which h attains its maximum.

   Comment: This exists because theorem 9.9, since the two endpoints are both defined and continuous on a closed interval, we are always able to take the supremum and infimum of the set and say there is point of the function that equal to this value.

   By the last equation, $x \in (a, b)$ and since it is a maximum $h'(x) = -$. Now if $h(t) < h(a)$ for some $t \in (a, b)$, the same argument applies if we choose for s a point on $[a, b]$ where h attains its maximum.

Theorem 6.6. ***The Mean Value Theorem*** *If f is a real continuous function on $[a, b]$ which is differentiable in (a, b), then there is a point $x \in (a, b)$ at which*

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

The proof is simply taking the last proof and use $g(x) = x$

Theorem 6.7. *Suppose that f is differentiable in (a, b).*

1. *If $f'(x) \geq 0$ for all $x \in (a, b)$, then f is monotonically increasing.*
2. *If $f'(x) = 0$ for all $x \in (a, b)$ then f is constant.*

3. *If $f'(x) \leq 0$ for all $x \in (a, b)$ then f is monotonically decreasing.*

Theorem 6.8. *Suppose f is a real differentiable function on $[a, b]$ and suppose that $f'(a) < \lambda < f'(b)$ then there is a point $x \in (a, b)$ such that $f'(x)\lambda$. Note that we does not guarantee that this function has to be a continuous*

Proof:

1. Put $g(t) = f(t) - \lambda t$, then
   - At $t = a$, we have $g'(a) = f'(a) - \lambda < 0$. This implies that $g(t)$ is decreasing just to the right of a, so g will drop below $g(a)$ for some nearby point $t_1 \in (a, b)$.
   - At $t = b$, we have $g'(b) = f'(b) - \lambda > 0$. This implies that $g(t)$ is increasing just to the left of b, so g will also drop below $g(b)$ for some nearby point $t_2 \in (a, b)$.
2. Since $g(t)$ decreases near a and increases near b, it must reach a minimum at some point in the interval (a, b). By the theorem 9.9, which is the intermediate value theorem.
3. Then at this minimum, we know that its derivative is going to be zero, thus $f'(x) = \lambda$.

Remark. The *Darboux Property* states that if f is differentiable on an interval, its derivative f' must satisfy the intermediate value property: for any a and b in the interval, if $f'(a) < \lambda < f'(b)$, then there exists $x \in (a, b)$ such that $f'(x) = \lambda$. This property holds even if f' is not continuous, as it arises from the limit-based definition of derivatives.

The Darboux Property is foundational to the Mean Value Theorem for derivatives, ensuring that f' takes on all values between $f'(a)$ and $f'(b)$, allowing us to conclude that certain intermediate values of the derivative must exist.

Corollary 6.8.1. *If f is differentiable on $[a, b]$ then f' cannot have any simple discontinuities on $[a, b]$.*

Theorem 6.9. *Suppose f and g are real and differentiable in (a, b) and $g'(x) \neq 0$ for every $x \in (a, b)$, where $-\infty \leq a < b < \infty$. Suppose that*

$$\frac{f'(x)}{g'(x)} \to A \quad x \to a$$

If

$$f(x) \to 0 \text{ and } g(x) \to 0 \text{ as } x \to a$$

or if

$$g(x) \to \infty \text{ as } x \to a$$

then

$$\frac{f(x)}{g(x)} \to a \text{ as } x \to a$$

Proof:

1. We first consider the case in which $-\infty \leq A < +\infty$. Choose a real number q such that $A < q$, and then choose r such that $A < r < q$. By first equation in the theorem, there is a point $c \in (a, b)$ such that $a < x < c$ implies
$$\frac{f'(x)}{g'(x)} < r.$$
2. If $a < x < y < c$, then by Cauchy IVT shows that there is a point $t \in (x, y)$ such that
$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r.$$
3. Suppose the second equation holds. Letting $x \to a$ in (18), we see that
$$\frac{f(y)}{g(y)} \leq r < q \quad (a < y < c).$$

4. Next, suppose the third limit holds. Keeping y fixed in last equation, we can choose a point $c_1 \in (a, y)$ such that $g(x) > g(y)$ and $g(x) > 0$ if $a < x < c_1$. Multiplying the last equation by $\frac{[g(x)-g(y)]}{g(x)}$, we obtain
$$\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)} \quad (a < x < c_1).$$
5. If we let $x \to a$ in the last equation, then the third equation shows that there is a point $c_2 \in (a, c_1)$ such that
$$\frac{f(x)}{g(x)} < q \quad (a < x < c_2).$$
6. Summing up, show that for any q, subject only to the condition $A < q$, there is a point c_2 such that $\frac{f(x)}{g(x)} < q$ if $a < x < c_2$.
7. In the same manner, if $-\infty < A \leq +\infty$, and p is chosen so that $p < A$, we can find a point c_3 such that
$$p < \frac{f(x)}{g(x)} \quad (a < x < c_3),$$
and the conclusion follows from these two statements.

7 Riemann-Stieltjes Integral

Definition. Let $[a, b]$ be a given interval, by a partition P of $[a, b]$ we mean a finite set of points $x_0, x_1, ..., x_n$ where
$$a = x_0 \leq x_1 \leq ... \leq x_{n-1} \leq x_n = b$$
and we write $\Delta x_i = x_i - x_{i-1}$.

Definition. Now suppose f is a bounded real function defined on $[a, b]$ then corresponding to each partition P of $[a, b]$ we give
$$M_i = \sup f(x) \quad (x_{i-1} \leq x \leq x_i)$$
$$m_i = \inf f(x) \quad (x_{i-1} \leq x \leq x_i)$$
$$U(p, f) = \sum_{i=1}^{n} M_i \Delta x_i$$
$$L(p, f) = \sum_{i=1}^{n} m_i \Delta x_i$$
$$\overline{\int_a^b} f dx = \inf U(p, f)$$
$$\underline{\int_a^b} f dx = \sup L(p, f)$$

The last two things defined are called the upper and lower Riemann integral of f over $[a, b]$ respectively.

Definition. If the upper and lower integrals are equal we say that f is Riemann integrable on $[a, b]$ we write $f \in \mathcal{R}$. Where $\mathcal{R}$ is the set of Riemann integrable function.

Comment: This definition is essentially saying that we first define the sum of rectangle using the lowest value of function and the greatest value of the function, then find the partition that generates the lowest sum of the greatest value and the greatest sum of the lowest value, and check if they are equal, if yes then they are Riemann integrable.

Remark. This is the Riemann integral of f over [a,b]. Since f is bounded there exists two numbers m and M such that
$$m \leq f(x) \leq M$$
thus for every P, we can define $m(b-a) \leq L(P, f) \leq U(P, f) \leq M(b-a)$ thus this set is bounded. The question about their equality hence the question of the integrability of f is a more delicate one.

Definition. Let α be a monotonically increasing function on $[a, b]$. Corresponding to each partition P of $[a, b]$, we write

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$$

it is clear that $\Delta\alpha_i \geq 0$ for any real function f which is bounded on $[a, b]$ we put that

$$U(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i$$

$$L(P, f, \alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i$$

where M_i, m_i have the same meaning as the former definition. And then we define

$$\overline{\int_a^b} f d\alpha = \inf U(P, f, \alpha)$$

$$\underline{\int_a^b} f d\alpha = \sup L(P, f, \alpha)$$

If the left members of the two integrals are equal, we denote their common value by

$$\int_a^b f d\alpha$$

This is called Riemann-Stieltjes integral of f with respect to α over $[a, b]$. If the integral exist, then we say that f is integrable with respect to α in the Riemann sense and write $f \in \mathcal{R}(\alpha)$.f

Remark. By taking $\alpha = x$, we obtain the Riemann integral, therefore it is a special case of Riemann-Stieltjes integral. Note that here α does not even need to be continuous.

Definition. We say that the partition P^* is a refinement of P if $P^* \supset P$. Now given two partitions, P_1 and P_2 we say that P^* is their common refinement if $P^* = P_1 \cup P_2$

Theorem 7.1. *If P^* us a refinement of P then*

$$L(P, f, \alpha) \leq L(P^*, f, \alpha)$$

$$U(P^*, f, \alpha) \leq U(P, f, \alpha)$$

Proof:

1. Suppose that P^* contains just one point more than P. Let this extra point be x^*, and suppose $x_{i-1} < x^* < x_i$. Where x_i and x_{i-1} are two consecutive points of P.

2. Let $w_1 = \inf f(x)$ for $x_{i-1} \leq x \leq x^*$ and $w_2 = \inf f(x)$ for $x^* \leq x \leq x_i$. Clearly we $w_1 \geq m_i$ and $w_2 \geq m_i$, because it is at least the infimum of the union for the two sets can could get larger.

3. Thus we know that

$$\begin{aligned}
L(P^*, f, \alpha) - L(P, f, \alpha) &= w_1[\alpha(x^*) - \alpha(x_{i-1})] + w_2[\alpha(x_i) - \alpha(x^*)] - m_i[\alpha(x_i) - \alpha(x_i)] \\
&= w_1[\alpha(x^*) - \alpha(x_{i-1})] + w_2[\alpha(x_i) - \alpha(x^*)] - m_i[\alpha(x_i) - \alpha(x_i)] - m_i[\alpha(x^*) - \alpha(x^*)] \\
&= (w_1 - m_i)[\alpha(x^*) - \alpha(x_{i-1})] + (w_2 - m_i)[\alpha(x_i) - \alpha(x^*)] \geq 0
\end{aligned}$$

Because all of the four terms in the brackets are greater than 0. If P^* contains k points more than P, we repeat this reasoning k times. The proof for the other equation is also analogous.

Comment: This is essentially saying that refined supremum and infimum will always have better approximation than the ones without refinement.

Theorem 7.2.

$$\underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha$$

Proof:

1. Let P^* be the common refinement of two partitions P_1 and P_2 then by the last theorem we know that
$$L(P_1, f, \alpha) \leq L(P^*, f, \alpha) \leq U(P^*, f, \alpha) \leq U(P_2, f, \alpha)$$
hence
$$L(P_1, f, \alpha) \leq U(P_2, f, \alpha)$$
now if P_2 is fixed and the supremum is taken over all the P_1. Therefore we know that
$$\underline{\int} f d\alpha \leq U(P_2, f, \alpha)$$
then take the infimum over all P_2 for the RHS of the inequality.

Theorem 7.3. *$f \in \mathcal{R}(\alpha)$ on $[a, b]$ if and only if for every $\epsilon > 0$ there exists a partition P so that*
$$U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$$

Proof:

1. For the first direction, we show that the condition shows the f is integrable wrt α. For every P we have
$$L(P, f, \alpha) \leq \underline{\int} f d\alpha \leq \overline{\int} f d\alpha \leq U(P, f, \alpha)$$

2. By Archimedean principle, we can always find an ϵ so that
$$0 \leq \underline{\int} f d\alpha - \overline{\int} f d\alpha \leq U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$$
thus because this equation is true for every ϵ, we can have the upper integral just equal to the lower integral thus $f \in \mathcal{R}(\alpha)$.

3. Conversely, suppose that $f \in \mathcal{R}(\alpha)$ and let $\epsilon > 0$ also be given. Then there is partitions P_1 and P_2 such that
$$U(P_2, f, \alpha) - \int f d\alpha < \frac{\epsilon}{2}$$
$$\int f d\alpha - L(P_1, f, \alpha) < \frac{\epsilon}{2}$$
because we know that the supremum and infimum are two limit points, thus there is a point no matter how close to it.

4. Now we choose P to be the common refinement of P_1 and P_2 so that together with the last theorem shows that
$$U(P, f, \alpha) \leq U(P_2, f, \alpha) < \int f d\alpha + \frac{\epsilon}{2} < L(P_2, f, \alpha) + \epsilon \leq L(P, f, \alpha) + \epsilon$$
thus
$$U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$$
thus we found a partition that satisfy the condition. This theorem furnishes a convenient criterion for integrability.

Theorem 7.4.
1. *If $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$ holds for some P and ϵ then this condition holds for every refinement of P with the same ϵ.*

2. *If $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$ holds for the partition $P = \{x_0, ..., x_n\}$ and if s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$ then*
$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta\alpha_i < \epsilon$$

3. *If $f \in \mathcal{R}(\alpha)$ and the hypothesis of (b) hold, then*

$$\left|\sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \int_a^b f d\alpha\right| < \epsilon$$

Proof:

1. (a) follows directly from theorem 11.1, which states that if P^* is a refinement of P then

$$L(P, f, \alpha) \leq L(P^*, f, \alpha)$$

$$U(P^*, f, \alpha) \leq U(P, f, \alpha)$$

that means the difference between the refinement is bounded by the difference for U and L, thus we can see that (a) of course should be less than ϵ.

2. (b) If both $f(s_i)$ and $f(t_i)$ lie in $[m_i, M_i]$ so that

$$|f(s_i) - f(t_i)| \leq M_i - m_i$$

Thus we know that

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)|\Delta\alpha_i \leq \sum (M_i - m_i)\Delta\alpha_i = U(P, f, \alpha) - L(P, f, \alpha)$$

which is exactly (b).

3. (c) The obvious inequalities

$$L(P, f, \alpha) \leq \sum f(t_i)\Delta\alpha_i \leq U(P, f, \alpha)$$

Which we can then take the infimum of maximum and supremum of minimum which are equal then we know that

$$L(P, f, \alpha) \leq \int f d\alpha \leq U(P, f, \alpha)$$

this proves (c).

Theorem 7.5. *If f is continuous on $[a, b]$ then $f \in \mathcal{R}(\alpha)$ on $[a, b]$.*

Proof:

1. Let $\epsilon > 0$ be given. Now choose $\eta > 0$ so that

$$[\alpha(b) - \alpha(a)]\eta < \epsilon$$

Since f is uniformly continuous on $[a, b]$ (because of theorem 7.12, which states that continuous function on compact subset is uniformly continuous) there exists a $\delta > 0$ such that

$$|f(x) - f(t)| < \eta$$

if $x \in [a, b], t \in [a, b]$ and $|x - t| < \delta$.

2. Now if P is any partition of $[a, b]$ such that $\Delta x_i < \delta$ for every i, then the last equation implies that $M_i - m_i \leq \eta$, thus

$$U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \leq \eta \sum_{i=1}^{n} \Delta\alpha_i = \eta[\alpha(b) - \alpha(a)] < \epsilon$$

Comment: We can also construct this proof by saying that η can be made arbitrarily small so that the difference between U and L are arbitrarily small.

Theorem 7.6. *If f is monotonic on $[a, b]$ and if α is continuous on $[a, b]$ the $f \in \mathcal{R}(\alpha)$.*

Proof:

1. Let $\epsilon > 0$ be given. For any positive integer n choose a partition such that

$$\Delta\alpha_i = \frac{\alpha(b) - \alpha(a)}{n}$$

this is possible because α is continuous. We choose this by picking the first time that α reach the next $\frac{\alpha(b)-\alpha(a)}{n}$ and the then proceed, because we know that α is monotonically increasing and by the nature of α, it has to reach every point between $\alpha(a)$ and $\alpha(b)$.

2. Now suppose that f is monotonically increasing. Then

$$M_i = f(x_i) \quad m_i = f(x_{i-1})$$

thus

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \\ &= \frac{\alpha(b) - \alpha(a)}{n} [f(b) - f(a)] < \epsilon \end{aligned}$$

as long as we pick n large enough, we then by the theorem 11.3, we can see that $f \in \mathcal{R}(\alpha)$.

Theorem 7.7. *Suppose f is bounded on $[a, b]$, and f has only finitely many points of discontinuity on $[a, b]$ and α is continuous at every point at which f is discontinuous then $f \in \mathcal{R}(\alpha)$.*

Proof:

1. Suppose $\epsilon > 0$ be given. Put $M = \sup |f(x)|$ and let E be the set of points at which f is discontinuous.

2. Since E is finite and α is continuous at every point of E, we can always cover E by finitely many disjoint intervals $[u_j, v_j] \subset [a, b]$ such that the sum of the corresponding differences $\alpha(v_j) - \alpha(u_j)$ is less than ϵ.

3. Then we can place these intervals in such a way that every point of $E \cap (a, b)$ lies in the interior of some $[u_j, v_j]$.

4. Now remove the segments (u_j, v_j) from $[a, b]$ the remaining set K is compact. (Because these are open interval, removing them, the set is still closed and bounded).

5. Hence f is continuous on K (by the theorem which states that continuous function on compact set is uniform continuous), and there exists $\delta > 0$ such that $|f(s) - f(t)| < \epsilon$ if $s \in K, t \in K$, $|s - t| < \delta$.

6. Now we define a partition $P = x_0, ..., x_n$ of $[a, b]$ as each u_j occurs in P. Each v_j occurs in P. No point of any segment (u_j, v_j) occurs in P. IF x_{i-1} is not one of the u_j, then $\Delta x_i < \delta$. (The Δx_i is either very small there close to each other due to continuity, or is close because α is continuous).

7. Note that $M_i - m_i \leq 2M$ for every i and that $M_i - m_i \leq \epsilon$ unless x_{i-1} is one of the u_j. Hence as the last proof

$$U(P, f, \alpha) - L(P, f, \alpha) \leq [\alpha(b) - \alpha(a)]\epsilon + 2M\epsilon$$

Because that ϵ is arbitrary. The theorem 11.3 shows $f \in \mathcal{R}(\alpha)$.