The Real Numbers

Rational Field

Take rationals as given.

$\left\{\frac{m}{n} \ , m, n \in \mathbb{R}, n \neq 0\right\}$ is a field with addition and multiplication defined.

Addition:

- $x+y=y+x$,
- $(x+y)+z=x+(y+z)$,
- There exists $0 \in \mathbb{Q}$ such that $x+0=x$.
- For each $x \in \mathbb{Q}$ there exists $y \in \mathbb{Q}$ such that $x+y=0$ ($y=-x$).

Multiplication:

- $xy=yx$,
- $(xy)z=x(yz)$,
- There exists $1 \in \mathbb{Q}$ such that $x1=1x=x \quad \forall x \in \mathbb{Q}$.
- For each $x \in \mathbb{Q} \setminus \{0\}$ there exists $y \in \mathbb{Q}$ such that $xy=1$ ($y=\frac{1}{x}=x^{-1}$).

Distribution: $x(y+z)=xy+xz$.

Axioms of Order

If S is a set, an order on S is a relation denoted by < such that

1. If $x, y \in S$ then exactly one of the statements $x<y$, $y<x$, $x=y$ is true.
2. If $x<y$ and $y<z$, then $x<z$.

We also write $x>y$ if $y<x$; $x \leq y$ if $x<y$ or $x=y$; $x \geq y$ if $x>y$ or $x=y$.

Definition: Ordered Field

An ordered field is a field F which is also an ordered set such that (in addition)

1. $x<y \Rightarrow x+z<y+z \quad \forall z$,
2. If $x>0$ and $y>0$ then $xy>0$.

Bounds

Definition: Bounded Above

Suppose S is an ordered set and W is a subset. We say W is bounded above if there exists $\beta \in S$ such that for all $x \in W$, $x \leq \beta$.

Note: β need not belong to W .

Definition: Least Upper Bound

Suppose S is an ordered set and W is a subset which is bounded above. Suppose there exists $\alpha \in S$ such that

1. α is an upper bound for W .
2. Whenever β is an upper bound for W , then $\beta \geq \alpha$.

Then α is called the least upper bound of W .

Note: α is unique.

Definition: Least Upper Bound Property

An order set S has the least upper bound property if whenever W is a subset which is bounded above, W has a least upper bound.

Note: $\mathbb{Q}$ does not have this property, but $\mathbb{R}$ does.
Note: The least upper bound of W need not belong to W. If it does, then W has a largest element or a maximum.

Remark
The order property together with the least upper bound property characterize the real numbers.

Remark
An ordered set with the least upper bound property also has the greatest lower bound property, i.e. any subset which is bounded below has a greatest lower bound.
Consider $-W=\{-x \mid x\in W\}$. $-W$ is bounded above, so it has a least upper bound α. $-\alpha$ is a greatest lower bound for W.

Dedekind Cuts

Definition: Dedekind Cuts
Consider all subsets S of rationals such that
1. S is not empty.
2. S is bounded above but does not contain its least upper bound.
3. If $s\in S$ and $r<s$, then $r\in S$.

Note: These sets are in 1-1 correspondence with the real numbers
Note: A rational number q is identified with the set $q^*=\{r\in\mathbb{Q} \mid r<q\}$.

Order
If α and β are Dedekind cuts, we define $\alpha<\beta$ if α is a proper subset of β.

Least Upper Bound Property
A set Σ of Dedekind cuts is bounded above if there is a cut β such that $\alpha<\beta$ for all $\alpha\in\Sigma$. The least upper bound of such a set Σ is just the union of all sets in Σ. Thus the least upper bound property is satisfied.

Field Operations
If α and β are Dedekind cuts, define
- $\alpha+\beta=\{r+s \mid r\in\alpha, s\in\beta\}$.
- 0 element: $0^*=\{r\in\mathbb{Q} \mid r<0\}$.
- 1 element: $1^*=\{r\in\mathbb{Q} \mid r<1\}$.
- The “negative” of a cut α: $-\alpha=\{s\in\mathbb{Q} \mid \exists r>0, r\in\mathbb{Q} \text{ such that } -s-r\notin\alpha\}$.

Decimal Expansion
A Dedekind cut α can be identified with an infinite decimal expansion.
- Choose an integer n_0 such that n_0+1 is an upper bound for α, but n_0 is not. There is a unique such integer.
- There is a unique integer $n_1\in\{0,1,\ldots,9\}$ such that $n_0+\frac{n_1+1}{10}$ is an upper bound for α, but $n_0+\frac{n_1}{10}$ is not.
- Similarly, there is a unique integer $n_2\in\{0,1,\ldots,9\}$ such that $n_0+\frac{n_1+1}{10}+\frac{n_2+1}{100}$ is an upper bound for α, but $n_0+\frac{n_1+1}{10}+\frac{n_2}{100}$ is not.

Proceeding inductively, we construct an infinite decimal expansion expansion associated to a Dedekind cut α.
Note: Decimal expansions can be taken as the basis for the construction of the real numbers; however, it must be noted that

such expansions are not unique (ex: $1.000\cdots=0.999\cdots$).

Example

In the system of Dedekind cuts, $\sqrt{2}$ is associated with the cut $\{r\in\mathbb{Q}\,|\,r^2<2\text{ or }r<0\}$, and $\sqrt{2}=1.414\cdots$.

Sequences

Definition: Sequence

A sequence of real numbers is a set of real numbers indexed by the natural numbers. It is usually written as $\{a_n\}_{n=1}^{\infty}$.

Definition: Convergence

The sequence $\{a_n\}$ converges to the limit $L\in\mathbb{R}$ if given any $\varepsilon>0$ there exists $n_0\in\mathbb{N}$ such that $|a_n-L|<\varepsilon$ whenever $n>n_0$.

Remark

A convergent sequence must be bounded, i.e. there exists $B>0$ such that $|a_n|\leq B$ for all n .

Definitions

- A sequence $\{a_n\}_{n=1}^{\infty}$ of real numbers is increasing if $a_n\leq a_{n+1}$ $\forall n$.
- A sequence $\{a_n\}_{n=1}^{\infty}$ of real numbers is strictly increasing if $a_n<a_{n+1}$ $\forall n$.
- A sequence $\{a_n\}_{n=1}^{\infty}$ of real numbers is decreasing if $a_n\geq a_{n+1}$ $\forall n$.
- A sequence $\{a_n\}_{n=1}^{\infty}$ of real numbers is strictly decreasing if $a_n>a_{n+1}$ $\forall n$.

"Monotone" means either increasing or decreasing.

Theorem

An increasing sequence which is bounded above is convergent. The limit is the least upper bound of the sequence (considered as a set of points).

Note: If a sequence $\{a_n\}_{n=1}^{\infty}$ is increasing, then either it is bounded above (and thus have a limit), or else it tends to ∞ (i.e. for all $M>0$ there exists n_0 depending on M such that $a_n>M$ for $n\geq n_0$).

Nested Intervals Lemma

Let $I_n=[a_n,b_n]$, $n=1,2,3,\ldots$ be a sequence of closed intervals such that $I_{n+1}\subset I_n$ $\forall n$. Then $\bigcap_{i=1}^{\infty} I_n$ is not empty.

Note: This is not necessarily true for open intervals. For example, if $I_n=\left(0,\frac{1}{n}\right)$, then the intersection is empty.

Note: This is not necessarily true for semi-infinite intervals. For example, if $I_n=[n,\infty)$, then the intersection is empty..

Bolzano-Weierstrass Theorem

Every bounded sequence $\{a_n\}_{n=1}^{\infty}$ of real numbers has a convergent subsequence $\{a_{n_k}\}_{k=1}^{\infty}$.

Continued Fractions

Consider $\cfrac{1}{3+\cfrac{1}{3+\cfrac{1}{3+\cdots}}}$.

- Terms are positive.
- $a_{n+1}=\dfrac{1}{3+a_n}$.
- $a_{n+2}-a_{n+1}=\dfrac{1}{3+a_{n+1}}-\dfrac{1}{3+a_n}=\dfrac{a_n-a_{n+1}}{(3+a_n)(3+a_{n+1})}$. Since the denominator > 9, $|a_{n+2}-a_{n+1}|<\dfrac{|a_n-a_{n+1}|}{9}$. The difference change signs and also approach 0 as $n\to\infty$ ..
- The subsequence of odd terms $\{a_{1,}a_{3,}a_{5,}\ldots\}$ decreases and is bounded below. The subsequence of even terms $\{a_{2,}a_{4,}a_{6,}\ldots\}$ increases and is bounded above. So the greatest lower bound of the odd terms must be equal to the least upper bound of the even terms, and thus the full sequence mush converge to this number.
- Let $n\to\infty$. Since we know the sequence has a limit L , we have $L=\dfrac{1}{3+L}\Rightarrow L^2+3L-1=0\Rightarrow L=\dfrac{-3\pm\sqrt{13}}{2}$. Since $L>0$ we must have $L=\dfrac{-3+\sqrt{13}}{2}$.

Definition: Cauchy Sequence

A sequence $\{a_n\}_{n=1}^{\infty}$ is called Cauchy if for any $\varepsilon>0$ there exists $N>0$ such that $|a_n-a_m|<\varepsilon$ for all $m,n>N$.

Theorem

A sequence $\{a_n\}_{n=1}^{\infty}$ (of real numbers) is Cauchy if and only if it is convergent.

Note: This theorem does not work with the rational numbers, i.e. a Cauchy sequence of rational numbers need not converge to a rational number.

Cardinality

Definition

Two sets have the same cardinality of there is a 1-to-1 correspondence between the points of the sets.

Definition

A has cardinality less than B if there is a 1-to-1 mapping from A into B . In this case B has cardinality greater than A .

Definition: Countable

A set S is said to be countable if it can be put into a 1-to-1 correspondence with the set of the natural numbers $\mathbb{N}$.

Topology of R^n

Definition

$\mathbb{R}^n \stackrel{\text{def}}{=} \{\vec{x}=(x_{1,}\ldots,x_n)\,|\,x_j\in\mathbb{R}\}$.

Inner Product

The dot product or inner product is given by $\langle \vec{x}, \vec{y} \rangle = \sum_{j=1}^{n} x_j y_j$.
Notice that $\langle \vec{x}, \vec{y} \rangle$ is bilinear, $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$, $\langle \vec{x}, \vec{x} \rangle \geq 0$ and $\langle \vec{x}, \vec{x} \rangle = 0 \Leftrightarrow \vec{x} = \vec{0}$.

Definition: Norm
The norm of $\vec{x} \in \mathbb{R}^n$ is $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle}$.
Properties: $\|\vec{x}\| \geq 0$ and $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$; $\|c\vec{x}\| = |c| \|\vec{x}\|$.

Schwarz Inequality
$|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\|$.

Proof: If $t \in \mathbb{R}$ then $\langle \vec{x} + t\vec{y}, \vec{x} + t\vec{y} \rangle \geq 0$. Now $\langle \vec{x} + t\vec{y}, \vec{x} + t\vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle + 2t \langle \vec{x}, \vec{y} \rangle + t^2 \langle \vec{y}, \vec{y} \rangle = \|\vec{x}\|^2 + 2t \langle \vec{x}, \vec{y} \rangle + t^2 \|\vec{y}\|^2$. For fixed $\vec{x}$ and $\vec{y}$, this is a quadratic polynomial in t . If $At^2 + Bt + C \geq 0 \quad \forall t$, then $4B^2 - 4AC \leq 0$. So $4\langle \vec{x}, \vec{y} \rangle \leq 4\|\vec{x}\|^2 \|\vec{y}\|^2 \Rightarrow \sqrt{\langle \vec{x}, \vec{y} \rangle} \leq \|\vec{x}\| \|\vec{y}\|$.

Triangle Inequality
$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$.

Proof: Note that

$$\begin{aligned}
\|\vec{x} + \vec{y}\|^2 &= \langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle \\
&= \langle \vec{x}, \vec{x} \rangle + 2\langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle \\
&= \|\vec{x}\|^2 + 2\langle \vec{x}, \vec{y} \rangle + \|\vec{y}\|^2 \\
&\leq \|\vec{x}\|^2 + 2\|\vec{x}\| \|\vec{y}\| + \|\vec{y}\|^2 \\
&= (\|\vec{x}\| + \|\vec{y}\|)^2
\end{aligned}$$

Now take square root.

Definition: Distance
We define the distance between two points $\vec{x}, \vec{y} \in \mathbb{R}^n$ by $\operatorname{dist}(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$.

Convergence and Completeness of R^N

Definition: Convergence
A sequence $\{\vec{x}_k\}_{k=0}^{\infty}$ converges to $\vec{x} \in \mathbb{R}^n$ if whenever $\varepsilon > 0$ is given, there exists $N > 0$ such that $\|\vec{x} - \vec{y}\| < \varepsilon$ whenever $k > N$.

Lemma
A sequence of points $\{\vec{x}_k\}_{k=0}^{\infty}$ in $\mathbb{R}^n$ converges to $\vec{a} \in \mathbb{R}^n$ if and only if $\lim_{k \to \infty} x_{k,j} = a_j \quad j = 1, \ldots, n$ where $\vec{x}_k = (x_{k,1}, x_{k,2}, \ldots, x_{k,n})$ and $\vec{a} = (x_1, x_2, \ldots, x_n)$.

Proof: It follows from the fact that if $\vec{y} = (y_{1,} \ldots, y_n)$ then $|y_j| \leq \|\vec{y}\| \leq \sqrt{n} \cdot \max_{j=1,\ldots,n} |y_j|$.

Definition: Cauchy
A sequence of points $\{\vec{x}_k\}_{k=0}^{\infty}$ in $\mathbb{R}^n$ is Cauchy if for any $\varepsilon > 0$ there exists $N = N(\varepsilon)$ such that $\|\vec{x}_j - \vec{x}_k\| < \varepsilon$ whenever $j, k > N$.

Lemma
A sequence of points $\{\vec{x}_k\}_{k=0}^{\infty}$ in $\mathbb{R}^n$ is Cauchy if and only if each component is Cauchy, i.e. $\{\vec{x}_{kj}\}_{k=0}^{\infty}$ for each $j=1,\dots,n$.

Theorem: Completeness Theorem for $\mathbb{R}^n$
$\mathbb{R}^n$ is complete, i.e. any Cauchy sequence converges.

Proof: If $\{\vec{x}_k\}_{k=0}^{\infty}$ is a Cauchy sequence, so is $\{\vec{x}_{kj}\}_{k=0}^{\infty}$ for each $j=1,\dots,n$. Since $\mathbb{R}$ is complete, there exists $a_j\in\mathbb{R}$ such that $x_{k,j}\to a_j$ as $k\to\infty$ for each $j=1,\dots,n$. Hence $\|\vec{x}_k-\vec{a}\|\to 0$ as $k\to\infty$.

Closed and Open Sets In *n*-Dimensional Space

Definition: Limit Point
Let A be a subset in $\mathbb{R}^n$. A point $\vec{x}\in\mathbb{R}^n$ is called a limit point of a subset A of $\mathbb{R}^m$ if there exists a sequence $\{\vec{x}_k\}_{k=0}^{\infty}$ of points in A which converges to $\vec{x}$.
In particular any point $\vec{x}\in A$ is a limit point.

Definition: Closed Set
A set in $\mathbb{R}^n$ which contains all its limit points is said to be closed.
Note: Generally, set defined by inequalities which are not strict are closed (eg: $\{\vec{x}\in\mathbb{R}^n \mid 1\le\|\vec{x}\|\le 2\}$).

Proposition
The union of a finite number of closed sets is closed.

Proof: Suppose $\{\vec{a}_k\}_{k=0}^{\infty}\subseteq A_1\cup A_2\cup\cdots\cup A_m$ where $A_1,\dots,A_m$ are closed. Suppose $\vec{a}_k\to\vec{x}\in\mathbb{R}^n$. At least one of the $A_1,\dots,A_m$, say A_{j_0} must contain infinitely many terms of the sequence. Arrange these terms in order so that the subscripts are increasing, then we get a subsequence $\{\vec{a}_{k_l}\}_{l=0}^{\infty}$ of the original sequence. This subsequence converges to $\vec{x}$. Since A_{j_0} is closed, $\vec{x}\in A_{j_0}$ and hence $\vec{x}\in A_1\cup A_2\cup\cdots\cup A_m$. So the union is closed.

Definition: Closure
Let A be an arbitrary subset of $\mathbb{R}^n$. Then a closure of A, denoted $\bar{A}$, is the set consisting of all limit points.
Note: $\bar{A}\subseteq A$.

Proposition
The closure of a set A is a closed set.

Proof: Let $\{\vec{x}_k\}_{k=0}^{\infty}$ be a sequence contained in $\bar{A}$ and suppose $\vec{x}_k\to\vec{x}\in\mathbb{R}^n$. Since $\vec{x}_k\in\bar{A}$, there exists a sequence of points in A converging to $\vec{x}_k$. Choose an element of this sequence, called $\vec{a}_k$, such that $\|\vec{a}_k-\vec{x}_k\|<\dfrac{1}{n}$. Then $\lim_{k\to\infty}\vec{a}_k=\lim_{k\to\infty}[(\vec{a}_k-\vec{x}_k)+\vec{x}_k]=0+\lim_{k\to\infty}\vec{x}_k=\vec{x}$. Hence $\vec{x}\in\bar{A}$.

Definition: Open Ball
The open ball of centre $\vec{a}$ and radius r is the set $B_r(\vec{a})=\{\vec{x}\in\mathbb{R}^n \mid \|\vec{x}-\vec{a}\|<r\}$.

Definition: Open Set

A subset A of $\mathbb{R}^n$ is open if whenever $\vec{a}\in A$ there exists $r=r(a)>0$ such that $B_r(\vec{a})\subseteq A$.
Note: Generally, set defined by strict inequalities are open (eg: $B_r(\vec{a})$).

Theorem
A set $A\subset\mathbb{R}^n$ is open if and only if A^c is closed.

Proof:
1. Suppose A is open. Let $\vec{a}\in A$. Then there exists $r>0$ such that $B_r(\vec{a})\subseteq A$. Let $\{\vec{x}_k\}_{k=0}^{\infty}$ be a sequence in A^c. Then $\|\vec{x}_k-\vec{a}\|\geq r$. So $\{\vec{x}_k\}_{k=0}^{\infty}$ cannot converge to $\vec{a}$. So $\vec{a}$ is not a limit point of A^c.
2. Suppose A is not open. Then no ball centred at $\vec{a}$ is contained in A. Hence for all $n\in\mathbb{N}$, we can find a point $\vec{x}_k\in B_{1/k}(\vec{a})$ such that $\vec{x}_k\notin A$ where $\vec{x}_k\to\vec{a}$. Hence A^c is not closed.

Proposition
The intersection of a finite number of open sets is open.

Proof: If $\vec{a}\in U_1\cup\cdots\cup U_k$ and each U_j is open, then for each $j=1,2,\dots,k$ there exists r_j such that $B_{r_j}(\vec{a})\subseteq U_k$. Let $r=\min\limits_{1\leq j\leq k}(r_j)$. Then $B_r(\vec{a})\subseteq U_j$ $\forall j$ and so $B_r(\vec{a})\subseteq\bigcap\limits_{j=1}^{k}U_j$.

Proposition
If $\{U_\alpha\}_{\alpha\in J}$ is any family of open subsets of $\mathbb{R}^n$, then $\bigcup\limits_{\alpha\in J}U_\alpha$ is open.

Proof: If $\vec{a}\in\bigcup\limits_{\alpha\in J}U_\alpha$, then $\vec{a}\in U_{\alpha_0}$ for some α_0. So there exists $r>0$ such that $B_{r_j}(\vec{a})\subset U_{\alpha_0}$. Then $B_{r_j}(\vec{a})\subset\bigcup\limits_{\alpha\in J}U_\alpha$.

Compactness

Definition: Compactness
A subset A of $\mathbb{R}^n$ is compact if every sequence $\{\vec{a}_k\}_{k=1}^{\infty}$ of points in A has a subsequence which converges to a point in A.

Proposition
The Bolzano-Weierstrass Theorem remains true in $\mathbb{R}^n$.

Proof: For the case $n=2$. Let $\vec{a}_k=(x_k,y_k)$ be a bounded sequence in $\mathbb{R}^2$. Then $\{x_k\}_{k=1}^{\infty}$ and $\{y_k\}_{k=1}^{\infty}$ are bounded sequences of real numbers. So $\{x_k\}_{k=1}^{\infty}$ has a convergent subsequence. Consider $\{y_{k_j}\}_{j=1}^{\infty}$. The Bolzano-Weierstrass Theorem implies that this has a further subsequence which is convergent, say $\{y_{k_{j_m}}\}_{m=1}^{\infty}$. Now $\{x_{k_{j_m}}\}_{m=1}^{\infty}$ is also convergent because it is a subsequence of a convergent sequence. Thus $\vec{a}_{k_{j_m}}=(x_{k_{j_m}},y_{k_{j_m}})$ is a convergent subsequence of points in $\mathbb{R}^2$.

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

Proof:
($\Rightarrow$) If the set C is unbounded, then for each $k\in\mathbb{N}$ there exists $\vec{a}_k\in C$ such that $\|\vec{a}_k\|>k$; such a sequence cannot have a convergent subsequence. If C is not closed, then there exists a sequence $\{\vec{a}_k\}_{k=1}^{\infty}$ of points in C which converges to a

point $\vec{a}\notin C$; such a sequence cannot have a subsequence which converges to a point of C .
($\Leftarrow$) Let $\{\vec{a}_k\}_{k=1}^{\infty}$ be a sequence of points in the set A . Balzano-Weierstrass Theorem implies there is a convergent subsequence, and A closed implies the limit of the subsequence is in A .

Cantor's Intersection Theorem

If $C_1\supseteq C_2\supseteq C_3\supseteq\cdots$ is a decreasing sequence of non-empty compact subsets of $\mathbb{R}^n$, then $\bigcap_{k\geq 1} C_k\neq\emptyset$.

Proof: For each k there exists a point $\vec{a}_k\in C_k$. We have $\vec{a}_k\in C_1$ $\forall k$ since $C_1\supseteq C_k$. The points $\{\vec{a}_k\}_{k=1}^{\infty}$ forms a sequence which lies in C_1 . C_1 is compact, so there is a subsequence $\{\vec{a}_{k_j}\}_{j=1}^{\infty}$ which converges to a point $\vec{a}\in C_1$. We will show $\vec{a}\in C_k$ $\forall k$ and hence $\vec{a}\in\bigcap_{k\geq 1} C_k$. Consider C_m . It is possible that $\vec{a}_{1},\ldots,\vec{a}_m$ do not lie in C_m , but all remaining terms in the subsequence $\{\vec{a}_{k_j}\}_{j=1}^{\infty}$ do lie in C_m . The remaining terms form a sequence of points in C_m which converges to $\vec{a}$. By compactness a subsequence of this sequence converges to $\vec{b}\in C_m$. We must have $\vec{b}=\vec{a}$ and $\vec{a}\in C_m$ $\forall m$. Hence $\bigcap_{k\geq 1} C_k\neq\emptyset$.

Cantor Set

The Cantor Set is constructed as follows:

- $S_0=[0,1]$.
- $S_1=\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]$.
- $S_2=\left[0,\frac{1}{9}\right]\cup\left[\frac{2}{9},\frac{1}{3}\right]\cup\left[\frac{2}{3},\frac{7}{9}\right]\cup\left[\frac{8}{9},1\right]$.
- Continue with $n=3,4,5,\ldots$.

S_n is compact for each n . Since $S_1\supseteq S_2\supseteq S_3\supseteq\cdots$, $\bigcap_{n=1}^{\infty} S_n\neq\emptyset$ by Cantor's Intersection Theorem.

In some sense, the Cantor Set is “small”. S_n contains 2^n intervals of length 3^{-n} . the middle third of each interval gets removed and has length 3^{-n-1} . The sum of the lengths of the intervals removed is $\sum_{n=1}^{\infty}\frac{2^n}{3^{n+1}}=\frac{1/3}{1-2/3}=1$.

In another sense, the Cantor Set is “large”; it is uncountable and have the same cardinality as the real numbers. Consider the ternary expansion of a number in $[0,1]$, $\sum_{k=1}^{\infty}\frac{t_k}{3^k}$ $t_k=0,1,2$ (eg. 0.02022202...). A number is in S_1 if the first digit is either 0 or 2. A number is in S_2 if the second digit is either 0 or 2. So a number in $[0,1]$ belongs to Cantor's Set precisely when the ternary expansion contains only 0's and 2's. Now take the binary expansion of any number in $[0,1]$ (eg. 0.01011101...) and map this point to the point with ternary expansion by replacing each 1 in the binary expansion by a 2. This gives a 1-1 map of the interval $[0,1]$ into the Cantor Set.

Functions

Limits and Continuity

Definition: Limit

Let S be a subset of $\mathbb{R}^n$. Suppose $F:S\to\mathbb{R}^m$ is a function. Suppose $\vec{a}$ is a cluster point (i.e. a limit point of $S\setminus\{\vec{a}\}$. We say $\lim_{\vec{x}\to\vec{a}} F(\vec{x})=\vec{v}$ where $\vec{v}\in\mathbb{R}^m$ if given any $\varepsilon>0$ there exists $\delta>0$ such that $\|F(\vec{x})-\vec{v}\|<\varepsilon$ whenever $0<\|\vec{x}-\vec{a}\|<\varepsilon$ and $\vec{x}\in S$.

Remark
If $F(\vec{x})=(f_1(\vec{x}),\dots,f_m(\vec{x}))$ where $f_j: S\to\mathbb{R}$, $j=1,\dots,m$, then $\lim_{\vec{x}\to\vec{a}} F(\vec{x})=\vec{v}\Leftrightarrow\lim_{\vec{x}\to\vec{a}} f_j(\vec{x})=v_j$ where $\vec{v}=(v_1,\dots,v_m)$.

Definition: Continuity
Let $S\subset\mathbb{R}^n$ and let $F:S\to\mathbb{R}^m$. We say that F is continuous at $\vec{a}$ if for all $\varepsilon>0$ there exists $\delta>0$ such that $\|F(\vec{x})-F(\vec{a})\|<\varepsilon$ whenever $\|\vec{x}-\vec{a}\|<\delta$ and $\vec{x}\in S$.
Note: If $\vec{a}$ is an isolated point (not a cluster point) of S, then any function on S is continuous at $\vec{a}$. Otherwise, F is continuous at $\vec{a}$ if and only if $\lim_{\vec{x}\to\vec{a}} F(\vec{x})=F(\vec{a})$.

Definition: Lipschitz
A function F in a subset $S\subset\mathbb{R}^n$ is said to be Lipschitz if there is a constant $C>0$ such that $\|F(\vec{x})-F(\vec{y})\|\le C\|\vec{x}-\vec{y}\|$ for all $\vec{x},\vec{y}\in S$.
Note: This condition implies continuity. Given $\varepsilon>0$, the condition for continuity is satisfied with $\delta=\frac{\varepsilon}{C}$.

Proposition

A linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ is Lipschitz. In column notation, a linear mapping is $A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}\sum_{j=1}^n a_{1j}x_j\\ \vdots\\ \sum_{j=1}^n a_{mj}x_j\end{pmatrix}$.

<u>Proof</u>: $A\vec{x}-A\vec{y}=A(\vec{x}-\vec{y})=\begin{pmatrix}\sum_{j=1}^n a_{1j}(x_j-y_j)\\ \vdots\\ \sum_{j=1}^n a_{mj}(x_j-y_j)\end{pmatrix}$. Note that $\left(\sum_{j=1}^n a_{ij}(x_j-y_j)\right)^2\le\left(\sum_{j=1}^n a_{ij}^2\right)\left(\sum_{j=1}^n (x_j-y_j)^2\right)$. So $\|A\vec{x}-A\vec{y}\|^2=\sum_{i=1}^n\left(\sum_{j=1}^n a_{ij}(x_j-y_j)\right)^2\le\left(\sum_{i=1}^m\sum_{j=1}^n a_{ij}^2\right)\|\vec{x}-\vec{y}\|^2$, and hence $\|A\vec{x}-A\vec{y}\|\le C\|\vec{x}-\vec{y}\|$ where $C^2=\sum_{i=1}^m\sum_{j=1}^n a_{ij}^2$.

Proposition
In one variable, a differentiable function with bounded derivative is Lipschitz.

<u>Proof</u>: $f(x)-f(y)=f'(t)(x-y)$ where t is some point between x and y (Mean Value Theorem). So $|f(x)-f(y)|=|f'(t)||x-y|\le C|x-y|$ where C is a bound for $|f'(t)|$ on the interval we are considering.

Remark
A function with a continuous derivative is locally Lipschitz.

Remark
Locally Lipschitz implies continuity.

Equivalent Conditions For Continuity

Let $F: S\subset\mathbb{R}^n\to\mathbb{R}^m$. F is continuous at $\vec{a}\in S$ if and only if for all sequences $\{\vec{x}_k\}_{k=1}^{\infty}$ of points in S such that $\vec{x}_k\to\vec{a}$ we have $F(\vec{x}_k)\to F(\vec{a})$ as $k\to\infty$.

Proof:

Let $\vec{a}\in S$ and let $\{\vec{x}_k\}_{k=1}^{\infty}$ be a sequence in S which converges to $\vec{a}$. Let $\varepsilon>0$. Since F is continuous at $\vec{a}$, there exists $\delta>0$ such that $\|F(\vec{x})-F(\vec{a})\|<\varepsilon$ when $\|\vec{x}-\vec{a}\|<\delta$ for $\vec{x}\in S$. Since $\{\vec{x}_k\}_{k=1}^{\infty}$ converges to $\vec{a}$, there exists $N>0$ such that $\|\vec{x}_k-\vec{a}\|<\delta$ when $k>N$. So $\|F(\vec{x}_k)-F(\vec{a})\|<\varepsilon$ when $k>N$. Hence $F(\vec{x}_k)\to F(\vec{a})$.

For the converse, it is equivalent to show that if F is not continuous then the sequence condition is not satisfied. If F is not continuous, then for all $\delta>0$ there exists $\varepsilon>0$ such that there is a point $\vec{x}=\vec{x}(\delta)$ such that $\|\vec{x}(\delta)-\vec{a}\|<\delta$ but $\|F(\vec{x}(\delta))-F(\vec{a})\|>\varepsilon$. Construct a sequence $\{\vec{x}_k\}_{k=1}^{\infty}$ in S converging to $\vec{a}$ as follows. Take $\delta=\frac{1}{k}$. There exists $\vec{x}_k\in S$ such that $\|\vec{x}_k-\vec{a}\|<\frac{1}{k}$ but $\|F(\vec{x}_k)-F(\vec{a})\|>\varepsilon$. Then $\vec{x}_k\to\vec{a}$ as $k\to\infty$ but $F(\vec{x}_k)$ does not converge to $F(\vec{a})$.

Definition: Relatively Open

Let S be a subset of $\mathbb{R}^n$. A subset V of S is (relatively) open if there is an open subset U of $\mathbb{R}^n$ such that $U\cap S=V$.

Theorem

Let $F: S\subset\mathbb{R}^n\to\mathbb{R}^m$. F is continuous if and only if whenever W is an open subset of $\mathbb{R}^m$, $F^{-1}(W)\overset{\text{def}}{=}\{\vec{x}\in S|F(\vec{x})=\vec{w}\in W\}$ is an open subset of S.

Proof:

Suppose F is continuous. Suppose W is an open subset of $\mathbb{R}^m$. Let $\vec{a}\in F^{-1}(W)$, i.e. $F(\vec{a})=\vec{u}\in W$. Since W is open, so there exists $\varepsilon>0$ such that $B_{\varepsilon}(\vec{u})\subseteq W$. Since F is continuous, there exists $\delta>0$ such that for all $\vec{x}\in S$ such that $\|\vec{x}-\vec{a}\|<\delta$, we have $\|F(\vec{x})-F(\vec{a})\|=\|F(\vec{x})-\vec{u}\|<\varepsilon$. This implies $\vec{x}\in F^{-1}(W)$. Hence $F^{-1}(W)$ contains $B_{\delta}(\vec{a})\cap S$, and therefore $F^{-1}(W)$ is open.

For the converse, let $\vec{a}\in S$. Let $\varepsilon>0$. Consider $B_{\varepsilon}(F(\vec{a}))=W$ an open set in $\mathbb{R}^m$. Since $F^{-1}(W)$ is an open subset of S containing $\vec{a}$, so there exists $\delta>0$ such that $B_{\delta}(\vec{a})\cap S\subseteq F^{-1}(W)$. This says if $\vec{x}\in S$ and $\|\vec{x}-\vec{a}\|<\delta$ then $\|F(\vec{x})-F(\vec{a})\|<\varepsilon$.

Properties of Continuous Functions

Properties of Limits of Continuous Functions

Let $f: S\subseteq\mathbb{R}^n\to\mathbb{R}^m$ and $g: S\subseteq\mathbb{R}^n\to\mathbb{R}^m$ be defined on $S\subseteq\mathbb{R}^n$. Let $\vec{a}\in S$. Let $\lim_{\vec{x}\to a} f(\vec{x})=\vec{u}\in\mathbb{R}^m$ and $\lim_{\vec{x}\to a} g(\vec{x})=\vec{v}\in\mathbb{R}^m$. Let $c\in\mathbb{R}$. Then:

- $\lim_{\vec{x}\to a} f(\vec{x})+g(\vec{x})=\vec{u}+\vec{v}$.
- $\lim_{\vec{x}\to a} c\,f(\vec{x})=c\,\vec{u}$.
- If f and g are real valued, then $\lim_{\vec{x}\to a} f(\vec{x})g(\vec{x})=u\,v$.
- If $v\neq 0$, then $\lim_{\vec{x}\to a}\frac{f(\vec{x})}{g(\vec{x})}=\frac{u}{v}$.

Note: Coordinating functions in $\mathbb{R}^n$ are continuous; that is if $\vec{x}=(x_1,\dots,x_n)$ then x_j is a continuous function of $\vec{x}$. Therefore any polynomial in $x_1,\dots,x_n$ is a continuous function of $\vec{x}$. And hence any rational function in $x_1,\dots,x_n$ is continuous whenever the denominator is not zero.

Composition Property

The composition of continuous functions is continuous.
Suppose $f:S\subseteq\mathbb{R}^n\to\mathbb{R}^m$. Let T be a subset of $\mathbb{R}^m$ containing the range of f. Suppose $g:T\subseteq\mathbb{R}^m\to\mathbb{R}^k$. Then $g\circ f=f(g(\vec{x}))$ is a continuous if f and g are. That is if f is continuous at $\vec{a}$ and g is continuous at $f(\vec{a})$, then $g\circ f$ is continuous at $\vec{a}$.

Proof: Let $\{\vec{x}_j\}_{j=1}^{\infty}$ be a sequence of points in S converging to $\vec{a}\in S$. Then $\{f(\vec{x}_j)\}_{j=1}^{\infty}$ is a sequence of points in T converging to $f(\vec{a})$ since f is continuous. Then $\{g(f(\vec{x}_j))\}_{j=1}^{\infty}=\{g\circ f(\vec{x}_j)\}_{j=1}^{\infty}$ is a sequence of points which converges to $g(f(\vec{a}))=g\circ f(\vec{a})$ since g is continuous..

Compactness and Extreme Values

Theorem

If K is a compact subset of $\mathbb{R}^n$ and $f:K\to\mathbb{R}^m$ is a continuous function, then $f(K)$ is compact.

Proof: Let $\{\vec{y}_k\}_{k=1}^{\infty}$ be a sequence in $f(K)$. Choose $\vec{x}_k\in K$ such that $f(\vec{x}_k)=y_k$ for $k=1,2,\ldots$. $\{\vec{x}_k\}_{k=1}^{\infty}$ is a sequence in K. There exists a subsequence $\{\vec{x}_{k_j}\}_{j=1}^{\infty}$ which converges to a point $\vec{a}\in K$. Since $\{f(\vec{x}_{k_j})\}_{j=1}^{\infty}=\{\vec{y}_{k_j}\}_{j=1}^{\infty}$ converges to $f(\vec{a})$, therefore $f(K)$ is compact.

Extreme Value Theorem

If K is a compact subset of $\mathbb{R}^n$ and $f:K\to\mathbb{R}$ is a real valued continuous function on K, then f assumes a maximum and minimum in $f(K)$, i.e. there exists points $\vec{a},\vec{b}\in K$ such that $f(\vec{a})\le f(\vec{x})\le f(\vec{b})$ $\forall\,\vec{x}\in K$.

Proof: K is a compact, so $f(K)$ is compact, or equivalently, closed and bounded. Thus $f(K)$ has a least upper bound (supremum) M and a greatest lower bound (infimum) m, i.e. $m\le f(\vec{x})\le M$. There exists a sequence of points in $f(K)$ converging to M, and since $f(K)$ is closed $M\in f(K)$. Hence there exists $\vec{b}\in K$ such that $f(\vec{b})=M$. Similar for m.

Uniform Continuity

Definition: Uniform Continuous

Let $S\subset\mathbb{R}^n$. A function $f:S\to\mathbb{R}^m$ is uniformly continuous if for all $\varepsilon>0$, there exists $\delta>0$ such that for all $\vec{a}\in S$ $\|f(\vec{x})-f(\vec{a})\|<\varepsilon$ whenever $\|\vec{x}-\vec{a}\|<\delta$ and $\vec{x}\in S$.

Remarks

1. A function which is Lipschitz, i.e. $\|f(\vec{x})-f(\vec{y})\|<C\|\vec{x}-\vec{y}\|$ $\forall\,\vec{x},\vec{y}\in S$, is uniformly continuous.
2. A linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is Lipschitz, hence uniformly continuous.

Theorem

A continuous function on a compact set is uniformly continuous.

Proof: Suppose this is not true. Then there exists $K\subset\mathbb{R}^n$ compact, $f:K\to\mathbb{R}^m$ which is not uniformly continuous. So there exists $\varepsilon>0$ such that there is no $\delta>0$ such that for all $\vec{a}\in K$, $\|f(\vec{x})-f(\vec{a})\|<\varepsilon$ whenever $\|\vec{x}-\vec{a}\|<\delta$ and $\vec{x}\in K$. With this ε, take $\delta=\dfrac{1}{k}$. There exists $\vec{a}_k,\vec{x}_k\in K$ such that $\|\vec{x}_k-\vec{a}_k\|<\dfrac{1}{k}$ but $\|f(\vec{x}_k)-f(\vec{a}_k)\|\ge\varepsilon$. Since $\{\vec{a}_k\}_{k=0}^{\infty}$ is a sequence of points in K (compact), there is a subsequence $\{\vec{a}_{k_j}\}_{j=0}^{\infty}$ which converges to a point $\vec{a}\in K$. Since

$\left\|\vec{x}_{k_j}-\vec{a}_{k_j}\right\|<\frac{1}{k_j}$, $j=1,2,3,\ldots$, so $\vec{x}_{k_j}\to\vec{a}$ also as $j\to\infty$. f is continuous at $\vec{a}$, hence $f\left(\vec{a}_{k_j}\right)\to f\left(\vec{a}\right)$. But $\left\|f\left(\vec{x}_{k_j}\right)-f\left(\vec{a}_{k_j}\right)\right\|\geq\varepsilon$. Contradiction. Hence f must be uniformly continuous.

The Intermediate Value Theorem

Intermediate Value Theorem

If f is a continuous real-valued function on $[a,b]$, then f assumes every value between $f(a)$ and $f(b)$ on $[a,b]$.

Proof: It is sufficient to show that If f is a continuous on $[a,b]$ and if $f(a)<0$ and $f(b)>0$, then there exists a point $c\in[a,b]$ such that $f(c)=0$.
Let $S=\{x\in[a,b]|f(x)<0\}$. $S\neq\emptyset$ since $a\in S$. S is bounded above by b , hence S has a least upper bound c . We'll show $f(c)=0$.

Suppose $f(c)<0$. Then by continuity, there exists $\delta>0$ such that $|f(x)-f(c)|<\frac{1}{2}|f(c)|$ if $|x-c|<\delta$ and $x\in[a,b]$. So $f(x)-f(c)<-\frac{1}{2}f(c)\Rightarrow f(x)<\frac{1}{2}f(c)<0$, in particular, this is true for $x\in(c,c+\delta)$ provided $x\in[a,b]$. Thus c is not an upper bound for S .

Suppose $f(c)>0$. Then by continuity, there exists $\delta>0$ such that $|f(x)-f(c)|<\frac{1}{2}f(c)$ if $|x-c|<\delta$ and $x\in[a,b]$. This implies $f(x)>0$ for $x\in(c-\delta,c+\delta)$ provided $x\in[a,b]$. Thus c is not the least upper bound for S .
The only possibility left is $f(c)=0$.

Monotone Functions

Definition: Monotone

$f:(a,b)\to\mathbb{R}$ is monotonic if it is increasing on (a,b) or decreasing on (a,b) .

Remark

If f is increasing, then $-f$ is decreasing, and vice versa.

Proposition

If $c\in(a,b)$ and f increasing, then $\lim_{x\to c^-}f(x)=L$ and $\lim_{x\to c^+}f(x)=M$, both exist and $L\leq f(c)\leq M$.

Proof: Let $L=\sup_{x\in(a,c)}f(x)$ and $M=\inf_{x\in(c,b)}f(x)$. Let $\varepsilon>0$. Then $L-\varepsilon$ is not an upper bound for $\{f(x):a<x<c\}$, i.e. there exists $x_1\in(a,c)$ such that $f(x_1)>L-\varepsilon$. Since f is increasing, $f(x)>L-\varepsilon$ for $x\in(x_1,c)$. Let $\delta=|x_1-c|$. Then $|f(x)-L|<\varepsilon$ when $x\in(c-\delta,c)$. Thus $\lim_{x\to c^-}f(x)=L$.
Since $L=\sup_{x\in(a,c)}f(x)$, $f(x)\leq f(c)$ $\forall x\in(a,c)$, hence $\sup_{x\in(a,c)}f(x)=L\leq f(c)$.

Remark

f is continuous at c if and only if $L=f(c)=M$.

Theorem

Suppose f is monotone on a closed interval $[a,b]$. Then the number of discontinuities of f is at most countable.

Proof: Without lost of generality, we can assume f is increasing.
For any point $x_0 \in [a, b]$, define the jump of f at x_0 to be $j(x_0) = \lim_{x \to x_0^+} f(x) - \lim_{x \to x_0^-} f(x)$ ($j(x_0) = 0$ iff f continuous at x_0). How many points can there be at which the jump is greater than 1? $N_0 \le f(b) - f(a)$. How many points are there where the jump satisfies $\frac{1}{2^k} \le j(x_0) < \frac{1}{2^{k-1}}$, $k = 1, 2, 3, \ldots$? Let that number be N_k. Then $N_k \frac{1}{2^k} \le f(b) - f(a)$, so $N_k \le 2^k (f(b) - f(a))$. Every discontinuity at x_0 has the property that there exists $k \in \mathbb{N}$ such that $\frac{1}{2^k} \le j(x_0) < \frac{1}{2^{k-1}}$.
Hence the number of discontinuity is countable.

Normed Vector Spaces

Definition: Normed Vector Space

A vector space V is said to be normed if there is a function $\|\cdot\|$ on V such that

1. $\|\vec{v}\| \ge 0$, and $\|\vec{v}\| = 0 \Leftrightarrow \vec{v} = \vec{0}$.
2. For $c \in \mathbb{R}$, $\|c\,\vec{v}\| = |c| \|\vec{v}\|$.
3. $\|\vec{u} + \vec{v}\| \le \|\vec{u}\| + \|\vec{v}\|$.

Example

Even in $\mathbb{R}^n$, there are other ways to define a norm besides the standard Euclidean norm.
The p-norm is defined by $\|\vec{x}\|_p = (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$, $1 \le p < \infty$ where $\vec{x} = (x_{1,} \ldots, x_n)$.
The case $p = 2$ gives the Euclidean norm.
$\|\vec{x}\|_1 = |x_1| + \cdots + |x_n|$.
$\|\vec{x}\|_\infty = \max |x_1|, \ldots, |x_n|$

Topology in Normed Spaces

In a normed vector space, the norm can be used to define convergence of sequence, Cauchy sequence, completeness, open sets, closed sets, continuity of functions, etc.

Definition: Convergence

Let V be a normed vector space. A sequence $\{\vec{x}_m\}_{m \in \mathbb{N}}$ is said to converge to a point $\vec{a} \in V$ if for all $\varepsilon > 0$ there exists $N > 0$ such that $\|\vec{x}_m - \vec{a}\| < \varepsilon$ whenever $m > N$.

Definition: Cauchy Sequence

A sequence $\{\vec{x}_m\}_{m \in \mathbb{N}}$ is said to be Cauchy if for all $\varepsilon > 0$ there exists $N > 0$ such that $\|\vec{x}_m - \vec{x}_n\| < \varepsilon$ whenever $m, n > N$.

Definition: Completeness

V is said to be complete if any Cauchy sequence has a limit in V.

Definition: Open Ball

The open ball with centre $\vec{a} \in V$ and radius $r > 0$ is defined by $B_r(\vec{a}) = \{\vec{v} \in V \mid \|\vec{v} - \vec{a}\| < r\}$.

Definition: Open Set

A set A is open if for all $\vec{a} \in A$ there exists $r > 0$ such that $B_r(\vec{a}) \subseteq A$.

Definition: Limit Point
$\vec{a} \in V$ i a limit point of a set $A \subseteq V$ if there is a sequence in A which converges to $\vec{a}$.
Note: $\vec{a}$ need not be in A .

Definition: Closed Set
A set is closed if it contains all its limit points.

Example
Consider sequences of $\mathbb{R}$: $(x_1, x_2, \ldots)$ such that $\sum_{j=1}^{\infty} |x_j|^p < \infty$. This gives a normed vector space with norm $\left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}}$.

Example
Certain spaces of functions give normed vector spaces.
Let K be a compact subset of a normed vector space. Consider the space V for continuous functions on K with the norm $\|f\| = \max_{x \in K} |f(x)|$ (a continuous function on a compact set has a maximum). This norm satisfies the three properties of norms.

Remark
The open unit ball with respect to any norm is open with respect to any norm.

Remark
The unit ball $B_1(\vec{0})$ with respect to a general norm on a normed vector space is convex, i.e. the line segment $(1-t)\vec{x} + t\vec{y} = \vec{x} + t(\vec{y} - \vec{x})$, $0 \leq t \leq 1$ is contained in $B_1(\vec{0})$ whenever $\vec{x}, \vec{y} \in B_1(\vec{0})$.
Proof: If $\|\vec{x}\| < 1$ and $\|\vec{y}\| < 1$, then $\|(1-t)\vec{x} + t\vec{y}\| \leq \|(1-t)\vec{x}\| + \|t\vec{y}\| \leq (1-t)\|\vec{x}\| + t\|\vec{y}\| \leq 1 - t + t = 1$.

Remark
Consider the closed Euclidean unit ball $\overline{B_1^{\mathrm{E}}(\vec{0})}$. If $\vec{x}, \vec{y} \in \overline{B_1^{\mathrm{E}}(\vec{0})}$, then the line segment joining $\vec{x}$ and $\vec{y}$ is in the open Euclidean ball except possibly for the end points. The closed Euclidean ball is strictly convex.
Note that the closed unit balls in the maximum norm and 1 norm are not strictly convex; there are line segments in the boundary.

Example: Norms on Vector Spaces of Functions
1. Let V be a vector space of continuous real valued functions on a compact set K . We can define the norm to be $\|f\|_\infty = \max_{\vec{x} \in K} |f(\vec{x})|$.
2. Let V be a vector space of continuous real valued functions on an interval $[a, b]$. Consider $\|f\|_p = \left[\int_a^b |f(\vec{x})|^p \, dx\right]^{\frac{1}{p}}$ where $1 \leq p < \infty$. This satisfies the properties of norms.
3. Let $C^k[a, b] = \{f : [a, b] \to \mathbb{R} \mid f \text{ and its derivatives up to order } k \text{ are continuous on } [a, b]\}$. This is a vector space. We can define $\|f\| = \|f\|_\infty + \|f'\|_\infty + \cdots + \|f^{(k)}\|_\infty$.

Inner Product Spaces

Definition: Inner Product Space

A vector space V is an inner product space if there is a real-valued function on $V \times V$ such that

1. Positive definiteness: $\langle \vec{x}, \vec{x} \rangle \geq 0$ with equality if and only if $\vec{x} = \vec{0}$.
2. Symmetry: $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$.
3. Bilinearity: $\langle c_1 \vec{x}_1 + c_2 \vec{x}_2, \vec{y} \rangle = c_1 \langle \vec{x}_1, \vec{y} \rangle + c_2 \langle \vec{x}_2, \vec{y} \rangle$ (similarly $\langle \vec{x}, c_1 \vec{y}_1 + c_2 \vec{y}_2 \rangle = c_1 \langle \vec{x}, \vec{y}_1 \rangle + c_2 \langle \vec{x}, \vec{y}_2 \rangle$).

Remark

In an inner product space, $\|\vec{x}\| = (\langle \vec{x}, \vec{x} \rangle)^{\frac{1}{2}}$ defines a norm.

Example

Let $C[a,b]$ be the vector space of continuous real-valued functions on $[a,b]$. Define $\langle f, g \rangle = \int_a^b f(x) g(x) dx$. Then $C[a,b]$ becomes an inner product space.

Cauchy-Schwartz Inequality

$|\langle \vec{x}, \vec{x} \rangle| \leq \|\vec{x}\| \|\vec{y}\|$, with equality if and only if $\vec{x}$ and $\vec{y}$ are collinear. By convention, $\vec{0}$ is collinear with any vector.

Proof: Let $t \in \mathbb{R}$. Then $\langle \vec{x} - t\vec{y}, \vec{x} - t\vec{y} \rangle \geq 0 \Rightarrow \langle \vec{x}, \vec{x} \rangle - 2t \langle \vec{x}, \vec{y} \rangle + t^2 \langle \vec{y}, \vec{y} \rangle \geq 0$. This is a quadratic in t , and thus its discriminate is $(2\langle \vec{x}, \vec{y} \rangle)^2 - 4\|\vec{x}\|^2 \|\vec{y}\|^2 \leq 0$.

Example

The Cauchy-Schwartz Inequality is true for infinite (square-summable) sequences.

Let $\{x_j\}_{j \in \mathbb{N}}$ and $\{y_j\}_{j \in \mathbb{N}}$ be sequences such that $\sum_{j=1}^{\infty} x_j^2 < \infty$ and $\sum_{j=1}^{\infty} y_j^2 < \infty$. Then $\sum_{j=1}^{\infty} x_j y_j$ and $\sum_{j=1}^{\infty} |x_j||y_j|$ are convergent and $\left| \sum_{j=1}^{\infty} x_j y_j \right| \leq \left(\sum_{j=1}^{\infty} x_j^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{\infty} y_j^2 \right)^{\frac{1}{2}}$.

Since a series which is absolutely convergent is convergent and $\left| \sum_{j=1}^{\infty} a_j \right| \leq \sum_{j=1}^{\infty} |a_j|$, it is sufficient to show $\sum_{j=1}^{\infty} |x_j||y_j| \leq \left(\sum_{j=1}^{\infty} x_j^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{\infty} y_j^2 \right)^{\frac{1}{2}}$. By the Cauchy Schwartz Inequality, $\sum_{j=1}^{N} |x_j||y_j| \leq \left(\sum_{j=1}^{N} x_j^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{N} y_j^2 \right)^{\frac{1}{2}} \leq \left(\sum_{j=1}^{\infty} x_j^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{\infty} y_j^2 \right)^{\frac{1}{2}}$.

Taking the limit as $N \to \infty$ yields $\sum_{j=1}^{\infty} |x_j||y_j| \leq \left(\sum_{j=1}^{\infty} x_j^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^{\infty} y_j^2 \right)^{\frac{1}{2}}$.

Hence, the sum of two square-summable sequences is square summable; $\sum_{j=1}^{\infty} (x_j + y_j)^2 = \sum_{j=1}^{\infty} x_j^2 + \sum_{j=1}^{\infty} 2x_j y_j + \sum_{j=1}^{\infty} y_j^2 < \infty$. Also $\sum_{j=1}^{\infty} (t x_j)^2 = t^2 \sum_{j=1}^{\infty} x_j^2 < \infty$. So the set of square summable sequences form a vector space, and $\langle \vec{x}, \vec{y} \rangle = \sum_{j=1}^{\infty} x_j y_j$ is an inner product of this space.

Orthonormal Sets

Definition: Orthogonal

Two vectors are orthogonal if $\langle \vec{x}, \vec{y} \rangle = 0$.

Definition: Orthonormal

A set of vectors is orthonormal if they are pairwise orthogonal, and $\|\vec{v}\| = 1$ for each $\vec{v}$ in the set.

Gram-Schmidt Process

If $\{\vec{x}_j\}$ is a finite or infinite set of vectors in an inner-product space, there is a way to construct a set of orthonormal vectors with the same span.

- If $\vec{x}_1=\vec{0}$, throw it out. If $\vec{x}_1\neq\vec{0}$, let $\vec{y}_1=\vec{x}_1$.
- Let $\vec{y}_2=\vec{x}_2-\frac{\langle\vec{x}_2,\vec{x}_1\rangle}{\langle\vec{x}_1,\vec{x}_1\rangle}\vec{x}_1=\vec{x}_2-\frac{\langle\vec{x}_2,\vec{y}_1\rangle}{\langle\vec{y}_1,\vec{y}_1\rangle}\vec{y}_1$. Then $\langle\vec{y}_1,\vec{y}_2\rangle=0$.
- Let $\vec{y}_3=\vec{x}_3-\frac{\langle\vec{x}_3,\vec{y}_1\rangle}{\langle\vec{y}_1,\vec{y}_1\rangle}\vec{y}_1-\frac{\langle\vec{x}_3,\vec{y}_2\rangle}{\langle\vec{y}_2,\vec{y}_2\rangle}\vec{y}_2$.
- Continue inductively.

Let $\vec{v}_1=\frac{y_1}{\|\vec{y}_1\|}$, $\vec{v}_2=\frac{y_2}{\|\vec{y}_2\|}$, etc. Then $\{\vec{v}_j\}$ is an orthonormal basis with the same span as $\{\vec{x}_j\}$.

Remark

If $\vec{e}_1,\dots,\vec{e}_n$ is an orthonormal basis of an inner product space, and if $\vec{v}=c_1\vec{e}_1+\cdots+c_n\vec{e}_n$, then $c_j=\langle\vec{v},\vec{e}_j\rangle$.

Trigonometric Polynomials and Fourier Series

Consider functions on $[-\pi,\pi]$. Define $\langle f,g\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(\theta)g(\theta)\,d\theta$. Then:

- $\|1\|=\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}1^2\,d\theta\right)^{\frac{1}{2}}=1$.
- $\{1,\sqrt{2}\sin(n\theta),\sqrt{2}\cos(n\theta)\}_{n\in\mathbb{N}}$ is an (infinite) orthonormal set.

Expansions With Respect to Orthonormal Bases

Let $\vec{e}_1,\dots,\vec{e}_n$ be an orthonormal set in an inner product space V. Let M be the subspace spanned by $\vec{e}_1,\dots,\vec{e}_n$. Then any $\vec{x}\in M$ can be written (uniquely) as $\vec{x}=\sum_{j=1}^{n}\alpha_j\vec{e}_j$, where $\alpha_j=\langle\vec{x},\vec{e}_j\rangle$.

Also, if $\vec{y}=\sum_{j=1}^{n}\beta_j\vec{e}_j\in M$, then $\langle\vec{x},\vec{y}\rangle=\sum_{j=1}^{n}\alpha_j\beta_j$. In particular, $\|\vec{x}\|^2=\sum_{j=1}^{n}\alpha_j^2$.

Note: These facts are completely analogous to what happens when $V=\mathbb{R}^n$ with the dot product and $\{\vec{e}_j\}_{j=1}^{n}$ is the standard basis.

Orthogonal Expansions in Inner Product Spaces

Definition: Projection

A projection on vector space V is a linear map $P:V\to V$ such that $P^2=P$.

Note: $P|_{\operatorname{Ran}P}=I$ where $\operatorname{Ran}P$ is the range of P, because $P(P(\vec{x}))=P^2(\vec{x})=P(\vec{x})$.

Note: If P is a projection then $I-P$ is also a projection, because $(I-P)^2=(I-P)\circ(I-P)=I-P-P+P^2=I-P$.

Definition: Orthogonal Projection

P is said to be an orthogonal projection if $\ker P$ is orthogonal to $\operatorname{Ran}P$.

Note: We need to be in an inner product space to talk about an orthogonal projection.

Example

In $\mathbb{R}^2$, $(x, y) \to (x, 0)$ is a projection.
The x -axis is the range, the y -axis is the kernels, so this is an orthogonal projection.
$(I-P)(x, y)=(0, y)$ is a projection.

Projection Theorem

Let $\vec{e}_1, \ldots, \vec{e}_n$ be an orthonormal set in a vector space V. Let M be the subspace spanned by $\vec{e}_1, \ldots, \vec{e}_n$. Define $P: V \to M$ by $P\vec{y}=\sum_{j=1}^{n} \langle \vec{y}, \vec{e}_j \rangle \vec{e}_j$, $\vec{y} \in V$. Then:

1. P is the orthogonal projection onto M.
2. $\sum_{j=1}^{n} \langle \vec{y}, \vec{e}_j \rangle^2 \le \|\vec{y}\|^2$.
3. For all $\vec{v} \in M$, $\|\vec{y}-\vec{v}\|^2=\|\vec{y}-P\vec{y}\|^2+\|P\vec{y}-\vec{v}\|^2$. Hence $P\vec{y}$ is the closet vector in M to $\vec{y}$ (this requires P to be an orthogonal projection).

Proof:
$P^2=P$ is clear.
P is an orthogonal projection. Let $P\vec{x}=\sum_{j=1}^{n} \alpha_j \vec{e}_j \in \operatorname{Ran} P=M$. Let $\vec{y} \in \ker P$. So
$P\vec{y}=\vec{0} \Rightarrow \sum_{j=1}^{n} \langle \vec{y}, \vec{e}_j \rangle \vec{e}_j=0 \Rightarrow \langle \vec{y}, \vec{e}_j \rangle=0 \Rightarrow \left\langle \vec{y}, \sum_{j=1}^{n} \alpha_j \vec{e}_j \right\rangle=0 \Rightarrow \langle \vec{y}, \vec{x} \rangle=0$. Hence $\ker P \perp \operatorname{Ran} P=M$.
Let $P\vec{y}=\sum_{j=1}^{n} \beta_j \vec{e}_j$, then $\|P\vec{y}\|^2=\sum_{j=1}^{n} \beta_j^{\,2}$. Now let $\vec{x}=\sum_{j=1}^{n} \alpha_j e_j \in M$ and $\vec{y} \in V$. Then

$$
\begin{aligned}
\|\vec{x}-\vec{y}\|^2 &= \langle \vec{x}-\vec{y}, \vec{x}-\vec{y} \rangle \\
&= \langle \vec{x}, \vec{x} \rangle - 2\langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle \\
&= \sum_{j=1}^{n} \alpha_j^{\,2} - 2\sum_{j=1}^{n} \alpha_j \beta_j + \|\vec{y}\|^2 \\
&= \sum_{j=1}^{n} \alpha_j^{\,2} - 2\sum_{j=1}^{n} \alpha_j \beta_j + \sum_{j=1}^{n} \beta_j^{\,2} - \sum_{j=1}^{n} \beta_j^{\,2} + \|\vec{y}\|^2 \\
&= \sum_{j=1}^{n} (\alpha_j - \beta_j)^2 - \|P\vec{y}\|^2 + \|\vec{y}\|^2 \\
&= \|\vec{x}-P\vec{y}\|^2 - \|P\vec{y}\|^2 + \|\vec{y}\|^2
\end{aligned}
$$

If we set $\vec{x}=P\vec{y}$, we get $\|P\vec{y}-\vec{y}\|^2=-\|P\vec{y}\|^2+\|\vec{y}\|^2$. This shows $\|\vec{y}\|^2 \ge \|P\vec{y}\|^2=\sum_{j=1}^{n} \langle \vec{y}, \vec{e}_j \rangle^2$.
Also, we have $\|\vec{x}-\vec{y}\|^2=\|\vec{x}-P\vec{y}\|^2-\|P\vec{y}\|^2+\|\vec{y}\|^2=\|\vec{x}-P\vec{y}\|^2+\|P\vec{y}-\vec{y}\|^2$.

Bessel's Inequality

Let $S \subseteq \mathbb{N}$ and let $\{e_n | n \in S\}$ be an orthonormal set in an inner product space V. For $\vec{x} \in V$, $\sum_{n \in S} \langle \vec{x}, \vec{e}_n \rangle^2 \le \|\vec{x}\|^2$.

Proof: If S is finite, this follows from Projection Theorem. If S is infinite, then we might as well take $S=\mathbb{N}$. Consider $\sum_{n=1}^{N} \langle \vec{x}, \vec{e}_n \rangle^2 \le \|\vec{x}\|^2$ (follows from Projection Theorem). Letting $N \to \infty$ gives the result.

Definition: Hilbert Space

A Hilbert space is a complete inner product space.
Note: Any finite dimensional inner product space over $\mathbb{R}$ is complete.

Definition: Span
The span of a set of vectors T in a Hilbert space, denoted $\operatorname{span} T$, is the set of finite linear combinations of vectors in T.

Definition: Closed Span
The closed span of a set of vectors T, denoted $\overline{\operatorname{span} E}$, is the closure of $\operatorname{span} T$.
Note: This is again a subspace.

Parseval's Theorem
Let E be a finite or countably infinite orthonormal set in a Hilbert space H.
- If E is finite, say $E=\{\vec{e}_1,\ldots,\vec{e}_n\}$, then $\overline{\operatorname{span} E}=\operatorname{span} E$.
- If E is infinite, then the subspace $M=\overline{\operatorname{span} E}$ consists of all vectors $\sum_{n=1}^{\infty}\alpha_n\vec{e}_n$ where the coefficient sequence $\{\alpha_n\}_{n=1}^{\infty}$ belongs to l^2 (i.e. $\sum_{n=1}^{\infty}{\alpha_n}^2<\infty$).

In either case, if $\vec{x}\in H$, then $\vec{x}\in\overline{\operatorname{span} E}$ if and only if $\sum\langle\vec{x},\vec{e}_n\rangle^2=\|\vec{x}\|^2$ (i.e. equality occurs in Bessel's Inequality).

<u>Proof</u>: $\sum\langle\vec{x},\vec{e}_n\rangle^2=\|\vec{x}\|^2$ follows from Projection Theorem if we have a finite orthonormal set.
Otherwise, $E=\{\vec{e}_n\}_{n=1}^{\infty}$. Let $\{\alpha_n\}_{n=1}^{\infty}\in l^2$. Let $\vec{x}_k=\sum_{n=1}^{k}\alpha_n\vec{e}_n$. $\{\vec{x}_k\}_{k=1}^{\infty}$ is a Cauchy sequence. Let $\varepsilon>0$. There exists $K>0$ such that $\sum_{n=K+1}^{\infty}{\alpha_n}^2<\varepsilon^2$. Hence if $l\geq k\geq K$, $\|\vec{x}_l-\vec{x}_k\|^2=\left\|\sum_{n=k+1}^{l}\alpha_n\vec{e}_n\right\|^2=\sum_{n=k+1}^{l}{\alpha_n}^2<\varepsilon^2$, so $\|\vec{x}_l-\vec{x}_k\|<\varepsilon$. Therefore $\{\vec{x}_k\}_{k=1}^{\infty}$ converges to $\vec{x}\in H$. Since $M=\overline{\operatorname{span} E}$, $\vec{x}\in M$.
Now suppose $\vec{x}\in H$ is arbitrary. Set $\alpha_n=\langle\vec{x},\vec{e}_n\rangle$. Bessel's Inequality implies $\{\alpha_n\}\in l^2$ since $\sum{\alpha_n}^2\leq\|\vec{x}\|^2$. Let $\vec{y}=\sum\alpha_n\vec{e}_n$. Then

$$\begin{aligned}\|\vec{x}-\vec{y}\|^2&=\|\vec{x}\|^2-2\langle\vec{x},\vec{y}\rangle+\|\vec{y}\|^2\\&=\|\vec{x}\|^2-2\langle\vec{x},\textstyle\sum\alpha_n\vec{e}_n\rangle+\|\vec{y}\|^2\\&=\|\vec{x}\|^2-2\textstyle\sum\alpha_n\langle\vec{x},\vec{e}_n\rangle+\sum{\alpha_n}^2\\&=\|\vec{x}\|^2-2\textstyle\sum{\alpha_n}^2+\sum{\alpha_n}^2\\&=\|\vec{x}\|^2-\textstyle\sum{\alpha_n}^2\end{aligned}$$

Now $\|\vec{x}-\vec{y}\|^2=0\Leftrightarrow\|\vec{x}\|^2=\sum{\alpha_n}^2\Leftrightarrow\vec{x}=\vec{y}\Leftrightarrow\vec{x}=\sum\alpha_n\vec{e}_n\in M$ ..

Theorem
Define $l^2=\left\{\vec{x}=(x_1,x_2,\ldots)\,\middle|\,\sum_{n=1}^{\infty}{x_n}^2<\infty\right\}$ with inner product $\langle\vec{x},\vec{y}\rangle=\sum_{n=1}^{\infty}x_n y_n$ and norm $\|\vec{x}\|=\sqrt{\langle\vec{x},\vec{x}\rangle}$. The space l^2 is complete, hence it is a Hilbert space.

<u>Proof</u>: Let $\{\vec{x}_k\}_{k=1}^{\infty}$ be a Cauchy sequence, with $\vec{x}_k=(x_{k,n})_{n=1}^{\infty}$. So for all $\varepsilon>0$ there exists $K>0$ such that $\|\vec{x}_k-\vec{x}_l\|<\varepsilon$ whenever $k,l>K$. Note that $|x_{k,n}-x_{l,n}|\leq\|\vec{x}_k-\vec{x}_l\|<\varepsilon$, hence the n-th component of the $\vec{x}_k$'s form a Cauchy sequence of real numbers. Since $\mathbb{R}$ is complete, there exists $y_n\in\mathbb{R}$ such that $x_{k,n}\to y_n$ as $k\to\infty$. Let $\vec{y}=(y_n)_{n=1}^{\infty}=(y_1,y_2,\ldots)$.
Show: $\vec{y}\in l^2$. Note that $\big|\|\vec{x}_k\|-\|\vec{x}_l\|\big|\leq\|\vec{x}_k-\vec{x}_l\|<\varepsilon$ for $k,l>K$. Hence the sequence $\{\|\vec{x}_k\|\}_{k=1}^{\infty}$ is Cauchy and so it converges to a limit L. Fix N. Then $\sum_{n=1}^{N}{y_n}^2=\lim_{k\to\infty}\sum_{n=1}^{N}(x_{k,n})^2\leq\lim_{k\to\infty}\|\vec{x}_k\|^2=L$. Let $N\to\infty$ and get $\sum_{n=1}^{\infty}{y_n}^2\leq L$. Thus $\vec{y}\in l^2$.
Show: $\vec{x}_k\to\vec{y}$ in l^2 norm. Fix $\varepsilon>0$ and choose $K>0$ such that $\|\vec{x}_k-\vec{x}_l\|<\varepsilon$ whenever $k,l>K$. Fix N. Then $\sum_{n=1}^{N}|y_n-x_{k,n}|^2=\lim_{l\to\infty}\sum_{n=1}^{N}|x_{l,n}-x_{k,n}|^2\leq\lim_{k\to\infty}\|\vec{x}_l-\vec{x}_k\|^2<\varepsilon^2$. Let Let $N\to\infty$ and get $\|\vec{y}-\vec{x}_k\|^2<\varepsilon^2\Leftrightarrow\|\vec{y}-\vec{x}_k\|<\varepsilon$. Since ε is

arbitrary, this says $\vec{x}_k \to \vec{y}$ in l^2 norm as $k \to \infty$.
Hence l^2 is a Hilbert space.

THE L^p NORMS

Definition: L^p Norm

The L^p norm on $C[a,b]$ is defined to be $\|f\|_p = \left(\int_a^b |f(x)|^p\,dx\right)^{\frac{1}{p}}$ where $1<p<\infty$.

Remark

$f \in L^p[a,b]$ if $\left(\int_a^b |f(x)|^p\,dx\right)^{\frac{1}{p}}$ is finite.

Lemma

If $A, B>0$ and $0<t<1$, then $A^t B^{1-t} \le tA+(1-t)B$. Equality occurs for some $t\in(0,1)$ if and only if $A=B$.

Proof: Let $a, b\in\mathbb{R}$. Let $\mathrm{e}^a = A$ and $\mathrm{e}^b = B$. The statement becomes $\mathrm{e}^{at}\mathrm{e}^{b(1-t)} = \mathrm{e}^{at+b(1-t)} \le t\,\mathrm{e}^a + (1-t)\mathrm{e}^b$. Assume $a\le b$ w.l.o.g. Since $\frac{d}{dx}(\mathrm{e}^x) = \mathrm{e}^x > 0$, the result follows.

Hölder's Inequality (for integrals)

On $C[a,b]$, consider $\|f\|_p \overset{\text{def}}{=} \left(\int_a^b |f(x)|^p\,dx\right)^{\frac{1}{p}}$ where $1<p<\infty$. If $f\in L^p[a,b]$ (i.e. $\int_a^b |f(x)|^p\,dx$ is finite) and if $g\in L^q[a,b]$ where $\frac{1}{p}+\frac{1}{q}=1$, then $\left|\int_a^b f(x)g(x)\,dx\right| \le \|f\|_p \|g\|_q$.

Proof:

It is sufficient to show $\int_a^b |f(x)g(x)|\,dx \le \|f\|_p\|g\|_q$. Note that it is true for $f\equiv 0$ or $g\equiv 0$, so we can assume $\|f\|_p\neq 0$ and $\|g\|_q \neq 0$. Let $A=\frac{|f(x)|^p}{(\|f\|_p)^p}$ and $B=\frac{|g(x)|^p}{(\|g\|_p)^p}$, x fixed. Let $t=\frac{1}{p}$, then $1-t=\frac{1}{q}$ since $\frac{1}{p}+\frac{1}{q}=1$. Then $\left(\frac{|f(x)|^p}{(\|f\|_p)^p}\right)^{\frac{1}{p}}\left(\frac{|g(x)|^q}{(\|g\|_q)^q}\right)^{\frac{1}{q}} \le \frac{1}{p}\frac{|f(x)|^p}{(\|f\|_p)^p}+\frac{1}{q}\frac{|g(x)|^q}{(\|g\|_q)^q}$. Taking $\int_a^b$ of both sides we get $\int_a^b \frac{|f(x)|}{\|f\|_p}\frac{|g(x)|}{\|g\|_q}\,dx = \frac{1}{\|f\|_p\|g\|_q}\int_a^b |f(x)||g(x)|\,dx \le \frac{1}{p}\int_a^b \frac{|f(x)|^p}{(\|f\|_p)^p}\,dx + \frac{1}{q}\int_a^b \frac{|g(x)|^q}{(\|g\|_q)^q}\,dx = \frac{1}{p}\frac{(\|f\|_p)^p}{(\|f\|_p)^p}+\frac{1}{q}\frac{(\|g\|_q)^q}{(\|g\|_q)^q} = \frac{1}{p}+\frac{1}{q}=1$. Thus $\int_a^b |f(x)||g(x)|\,dx \le \|f\|_p\|g\|_q$.

Generalizations of Hölder's Inequality

1. One can introduce a weight function ($w(x)$ positive, continuous or piecewise continuous) in the integrals, i.e. $\int_a^b |f(x)g(x)|w(x)\,dx \le \left(\int_a^b w(x)|f(x)|^p\,dx\right)^{\frac{1}{p}}\left(\int_a^b w(x)|g(x)|^q\,dx\right)^{\frac{1}{q}}$. An example is $w(x)=1-x^2$ used on the interval $[-1,1]$, which leads to a system of orthogonal polynomials called the Legendre polynomials.
2. The interval could be infinite or semi-infinite. In these cases, one needs to consider convergence of the integrals.

3. f and g could be piecewise continuous (continuous except for a finite number of discontinuities if the interval is finite, or a finite number of discontinuities on any finite subinterval if the interval is finite).
4. There is a more powerful version of integration theory, called Measure Theory or Lebesgue Integration, in which Hölder's inequality still holds.

Hölder's Inequality (for sequences)

If $(a_n)_{n=1}^{\infty}\in l^p$ (i.e. sequences such that $\sum_{n=1}^{\infty}|a_n|^p<\infty$) and $(b_n)_{n=1}^{\infty}\in l^q$ where $\frac{1}{p}+\frac{1}{q}=1$, then

$$\left|\sum_{n=1}^{\infty} a_n b_n\right|\le\left(\sum_{n=1}^{\infty}|a_n|^p\right)^{\frac{1}{p}}\left(\sum_{n=1}^{\infty}|b_n|^q\right)^{\frac{1}{q}} .$$

Note: These inequalities remain valid with weights $w_1, w_2, \dots\ , w_j>0$, i.e.

$$\left|\sum_{n=1}^{\infty} w_n a_n b_n\right|\le\sum_{n=1}^{\infty} w_n|a_n b_n|\le\left(\sum_{n=1}^{\infty} w_n|a_n|^p\right)^{\frac{1}{p}}\left(\sum_{n=1}^{\infty} w_n|b_n|^q\right)^{\frac{1}{q}} .$$

Proof: Take $f(x)=a_n\ \ , x\in[n-1, n)$ and $g(x)=b_n\ \ , x\in[n-1, n)$ for $1\le n<\infty$. f and g are piecewise continuous on $[0,\infty)$. Then $\int_0^{\infty}|f(x)|^p\,dx=\sum_{n=1}^{\infty}|a_n|^p$ and $\int_0^{\infty}|g(x)|^q\,dx=\sum_{n=1}^{\infty}|b_n|^q$.

Minkowski's Inequality (for integrals)

(This is the triangle inequality for L^p spaces)

Let $f, g\in C[a,b]$ and $1<p<\infty$. Then $\left(\int_a^b|f(x)+g(x)|^p\,dx\right)^{\frac{1}{p}}\le\left(\int_a^b|f(x)|^p\,dx\right)^{\frac{1}{p}}\left(\int_a^b|g(x)|^p\,dx\right)^{\frac{1}{p}}$.

Note: This shows the p -norms are indeed norms, i.e. they satisfy the triangle inequality.
Note: This is also true for $p=1$ and $p=\infty$.

Proof: Let $\frac{1}{q}=1-\frac{1}{p}$, then $q=\frac{p}{p-1}$.

Note that $\left\||f|^{p-1}\right\|_q=\left(\int_a^b|f(x)|^{(p-1)q}\,dx\right)^{\frac{1}{q}}=\left(\int_a^b|f(x)|^p\,dx\right)^{\frac{1}{q}}=\left(\left(\int_a^b|f(x)|^p\,dx\right)^{\frac{1}{p}}\right)^{\frac{p}{q}}=(\|f\|_p)^{\frac{p}{q}}=(\|f\|_p)^{p-1}$.

Now,

$$\begin{aligned}
(\|f+g\|_p)^p&=\int_a^b|f(x)+g(x)|^p\,dx\\
&=\int_a^b|f(x)+g(x)|^{p-1}|f(x)+g(x)|\,dx\\
&\le\int_a^b|f(x)+g(x)|^{p-1}|f(x)|\,dx+\int_a^b|f(x)+g(x)|^{p-1}|g(x)|\,dx\\
&\le\left(\int_a^b|f(x)+g(x)|^{(p-1)q}\,dx\right)^{\frac{1}{q}}\left(\int_a^b|f(x)|^p\,dx\right)^{\frac{1}{p}}+\left(\int_a^b|f(x)+g(x)|^{(p-1)q}\,dx\right)^{\frac{1}{q}}\left(\int_a^b|g(x)|^p\,dx\right)^{\frac{1}{p}}\\
&\le\left(\int_a^b|f(x)+g(x)|^{(p-1)q}\,dx\right)^{\frac{1}{q}}(\|f\|_p+\|g\|_p)\\
&=(\|f+g\|_p)^{\frac{p}{q}}(\|f\|_p+\|g\|_p)
\end{aligned}$$

Since $\frac{p}{q}=p-1$, we have $(\|f+g\|_p)^p\le(\|f+g\|_p)^{p-1}(\|f\|_p+\|g\|_p)\Rightarrow\|f+g\|_p\le\|f\|_p+\|g\|_p$.

Minkowski's Inequality (for sequences)

Let $1<p<\infty$. Then $\left(\sum_{n=1}^{\infty}|a_n+b_n|^p\right)^{\frac{1}{p}}\leq\left(\sum_{n=1}^{\infty}|a_n|^p\right)^{\frac{1}{p}}\left(\sum_{n=1}^{\infty}|b_n|^p\right)^{\frac{1}{p}}$.
Note: This is also true for $p=1$ and $p=\infty$.

Remark
- Similarly, there are versions of Minkowski's Inequalities with weights.
- There are generalizations to piecewise continuous functions, infinite intervals, etc.

Proposition
If $f\in C[0,1]$ and , $1\leq r<s<\infty$, then $\|f\|_r<\|f\|_s$. Also $\|f\|_p<\|f\|_\infty$
Note: If the interval does not have length 1, then a constant (depending on r and s) is introduced in this inequality.

Proof:

1. $\int_0^1|f(x)|^r dx=\int_0^1|f(x)|^r\cdot 1\,dx$. Using Hölder's Inequality with $p=\frac{s}{r}>1$, we get
$\int_0^1|f(x)|^r dx\leq\left(\int_0^1|f(x)|^{rp}dx\right)^{\frac{1}{p}}\left(\int_0^1 1^q dx\right)^{\frac{1}{q}}=\left(\int_0^1|f(x)|^{rp}dx\right)^{\frac{1}{p}}$, so
$\|f\|_r=\left(\int_0^1|f(x)|^r dx\right)^{\frac{1}{r}}\leq\left(\int_0^1|f(x)|^{rp}dx\right)^{\frac{1}{pr}}=\left(\int_0^1|f(x)|^s dx\right)^{\frac{1}{s}}=\|f\|_s$.
2. $\|f\|_p=\left(\int_0^1|f(x)|^p dx\right)^{\frac{1}{p}}\leq\left(\int_0^1\|f\|_\infty{}^p dx\right)^{\frac{1}{p}}=\|f\|_\infty\left(\int_0^1 dx\right)^{\frac{1}{p}}=\|f\|_\infty$.

Remarks
1. If the interval does not have length 1, then a constant (depending on r and s) is introduced in these inequalities.
2. On the finite interval, convergence in L^∞ implies convergence in L^p . More generally, convergence in L^s implies convergence in L^r on a finite interval if $r<s$.
3. This is no longer true on infinite intervals. On a infinite or semi-infinite interval, $L^r\not\subseteq L^s$ and $L^s\not\subseteq L^r$ if $1\leq r<s<\infty$.

Limits of Functions
What is the relation between convergence in L^p and convergence in L^q for different values of p and q ?

Limits of Functions

Definition: Pointwise Convergence
If $\{f_n\}_{n=1}^\infty$ is a sequence of real-valued functions on $[a,b]$, then $f_n\to f$ pointwise if for each $x\in[a,b]$, $f_n(x)\to f(x)$ as $n\to\infty$.

Definition: Uniform Convergence
If $f_n\to f$ in maximum norm, then f_n is said to converge to f uniformly.
Note: $f_n\to f$ uniformly if and only if for all $\varepsilon>0$ there exists $N>0$ such that $|f_n(x)-f(x)|<\varepsilon$ for all $x\in[a,b]$ whenever $n>N$.

Remark

If S is compact, then uniform convergence is equivalent to convergence in $L^{\infty}(S)$. Any continuous function on a compact set S is in $L^{\infty}(S)$.
If S is not compact, consider the space $C_b(S)=\{f:S\to\mathbb{R} \text{ which are continuous and bounded}\}$. For functions in $C_b(S)$, uniform convergence is equivalent to convergence in $L^{\infty}(S)$.

Theorem

For a sequence of functions $\{f_n\}_{n=1}^{\infty}\in C_b(S)$, $f_n\to f$ uniformly if and only if $\max_{x\in S}|f_n-f|\to 0$ as $n\to\infty$.

Uniform Convergence and Continuity

Example

If $f_n(x)=x^n$ on $[0,1]$, f_n converges pointwise to $f(x)=\begin{cases}0 & ,0\le x<1\\ 1 & ,x=1\end{cases}$. The limit function is not continuous. This cannot happen if the convergence is uniform.

Theorem

If $\{f_n\}_{n=1}^{\infty}$ is a sequence of continuous functions on S and $f_n\to f$ uniformly on S, then f is continuous.

Proof: Let $a\in S$ and $\varepsilon>0$. Now $f(x)-f(a)=f(x)-f_n(x)+f_n(x)-f_n(a)+f_n(a)-f(a)$, so $|f(x)-f(a)|\le|f(x)-f_n(x)|+|f_n(x)-f_n(a)|+|f_n(a)-f(a)|$. Now choose n sufficiently large such that $|f(x)-f_n(x)|<\frac{\varepsilon}{3}$ $\forall x\in S$ (so in particular $|f(a)-f_n(a)|<\frac{\varepsilon}{3}$). With this choice of n, choose $\delta>0$ such that $|f_n(x)-f_n(a)|<\frac{\varepsilon}{3}$ if $|x-a|<\delta$. With this choice of δ, $|f(x)-f(a)|\le\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon$. Hence f is continuous.

Completeness Theorem

The space $C(K)$ of continuous real-valued functions on a compact set $K\subseteq\mathbb{R}^n$ is complete.

Proof: Suppose $\{f_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\|\cdot\|_{\infty}$. Fix $\vec{x}\in K$. Then $\{f_n(\vec{x})\}_{n=1}^{\infty}$ is a Cauchy sequence of real numbers, so $|f_n(\vec{x})-f_m(\vec{x})|\le\|f_n-f_m\|$. For each $\vec{x}\in K$ there exists real number $f(\vec{x})$ such that $f_n(\vec{x})\to f(\vec{x})$. The convergence of $f_n(\vec{x})$ to $f(\vec{x})$ is uniform, hence f is a continuous function. Therefore $\{f_n(\vec{x})-f(\vec{x})\}_{\infty}$ implies $C(K)$ is complete.

Uniform Convergence and Integration

Proposition

If $\{f_n\}_{n=1}^{\infty}$ is a sequence of continuous functions on $[a,b]$, then $\int_a^b f_n(x)\,dx\to\int_a^b f(x)\,dx$.

Note: This is not true in general if the convergence is not uniform.
Note: This is not true in general if the interval is infinite.

Proof: $\left|\int_a^b f_n(x)\,dx-\int_a^b f(x)\,dx\right|=\left|\int_a^b (f_n(x)-f(x))\,dx\right|\le\int_a^b|f_n(x)-f(x)|\,dx$. Let $\varepsilon>0$. Then there exists $N>0$ such that $|f_n(x)-f(x)|<\varepsilon$ $\forall x$ when $n>N$. For such N, $\int_a^b|f_n(x)-f(x)|\,dx\le\int_a^b\varepsilon\,dx=\varepsilon(b-a)$. Hence $\int_a^b f_n(x)\,dx\to\int_a^b f(x)\,dx$.

Integral Convergence Theorem

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of continuous functions on $[a,b]$ such that $f_n \to f$ uniformly on $[a,b]$. Fix a point $c \in [a,b]$. Let $F_n(x)\int_c^x f_n(t)\,dt$ and $F(x)\int_c^x f(t)\,dt$. Then $F_n \to F$ uniformly on $[a,b]$.

Theorem

Suppose $\{h_n\}_{n=1}^{\infty}$ is a sequence of C^1 functions on $[a,b]$ such that h_n' converges uniformly to a function g on $[a,b]$, and suppose there exists a point $c \in [a,b]$ such that $\lim_{n\to\infty} h_n(c) = \gamma$ exists. Then h_n converges uniformly to a differentiable function h such that $h(c) = \gamma$ and $h' = g$.

<u>Proof</u>: Let $h_n(x) = h_n(c) + \int_c^x h_n'(t)\,dt$. Since $h_n(c) \to \gamma$ and $\int_c^x h_n'(t)\,dt \to \int_c^x g(t)\,dt$ uniformly, therefore $h_n(x) \to h(x) = \gamma + \int_c^x g(t)\,dt$. This says $h'(x) = g(x)$ and $h(c) = \gamma$.

Remark

In complex variable theory, uniform convergence of a sequence of analytic functions in a domain *does* imply uniform convergence of the derivatives on compact subsets.

Example

Consider $f_n(x) = \dfrac{\sin(nx)}{n}$. It is clear that $f_n \to 0$ uniformly for $x \in \mathbb{R}$ as $n \to \infty$. However, $f'_n = \cos(nx)$ which does not converge on $\mathbb{R}$ or on any finite subinterval.

In $\mathbb{C}$, $\dfrac{\sin(nx)}{n}$ does not converge to 0.

Theorem

Suppose $f(x,t)$ (real-valued) is continuous on $[a,b]\times[c,d]$. Let $F(x) = \int_c^d f(x,t)\,dt$. Then F is continuous on $[a,b]$.

<u>Proof</u>: f is uniformly continuous on $[a,b]\times[c,d]$. So given $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x_1,t_1) - f(x_2,t_2)| < \varepsilon$ when $\|(x_1,t_1) - (x_2,t_2)\| < \delta$. In particular $|f(x_1,t) - f(x_2,t)| < \varepsilon$ when $|x_1 - x_2| < \delta$ for all t. Now $F(x_1) - F(x_2) = \int_c^d f(x_1,t)\,dt - \int_c^d f(x_2,t)\,dt = \int_c^d (f(x_1,t) - f(x_2,t))\,dt$, so if $|x_1 - x_2| < \delta$ then $|F(x_1) - F(x_2)| \le \int_c^d |f(x_1,t) - f(x_2,t)|\,dt < \int_c^d \varepsilon\,dt = \varepsilon(d-c)$. Hence if x_2 is fixed, $F(x_1) \to F(x_2)$ as $x_1 \to x_2$.

Leibniz's Rule: Differentiation Under an Integral

Suppose that $f(x,t)$ and $\dfrac{\partial f}{\partial x}(x,t)$ are continuous on $[a,b]\times[c,d]$. Define $F(x) = \int_c^d f(x,t)\,dt$. Then F is continuous on $[a,b]$ and $F'(x) = \int_c^d \dfrac{\partial f}{\partial x}(x,t)\,dt$.

Proof: Fix $x_0 \in [a,b]$ and let $h \in \mathbb{R}$, $h \neq 0$. Want $\frac{F(x_0+h)-F(x_0)}{h} - \int_c^d \frac{\partial f}{\partial x}(x_0,t)\,dt \to 0$ as $h \to 0$.

Fix h. $\frac{F(x_0+h)-F(x_0)}{h} = \int_c^d \frac{f(x_0+h,t)-f(x_0,t)}{h}\,dt$. By the Mean Value Theorem, for each $t \in [c,d]$ there exists a point $x(t)$ lying in between x_0 and x_0+h such that $f(x_0+h,t)-f(x_0,t) = \frac{\partial f}{\partial x}(x(t),t)\cdot h$, and so $\frac{f(x_0+h,t)-f(x_0,t)}{h} = \frac{\partial f}{\partial x}(x(t),t)$. Note that MVT does not say $x(t)$ depends continuously on t, however for fixed h the LHS is a continuous function of t and so the RHS is also.

Now $\frac{\partial f}{\partial x}(x(t),t) = f_x(x,t)$ is continuous on the compact set $[a,b]\times[c,d]$, so it is uniformly continuous. So given $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x_1,t_1)-f(x_2,t_2)| < \varepsilon$ when $\|(x_1,t_1)-(x_2,t_2)\| < \delta$. In particular $|f(x_1,t)-f(x_2,t)| < \varepsilon$ when $|x_1-x_2| < \delta$ for all t. Fix h with $|h| < \delta$.

$$\left|\frac{F(x_0+h)-F(x_0)}{h} - \int_c^d f_x(x_0,t)\,dt\right| = \left|\int_c^d \frac{f(x_0+h,t)-f(x_0,t)}{h} - f_x(x_0,t)\,dt\right| \leq \int_c^d |f_x(x(t),t)-f_x(x_0,t)|\,dt \leq \int_c^d \varepsilon\,dt = \varepsilon(d-c)$$

since $|x(t)-x_0| < |h| < \delta$. This implies $\lim_{h\to 0} \frac{F(x_0+h)-F(x_0)}{h} - \int_c^d \frac{\partial f}{\partial x}(x_0,t)\,dt = 0$, hence the result.

Example

$$\frac{d}{dx}\int_0^\pi \sin(xy)\,dy = \int_0^\pi \frac{\partial}{\partial x}\sin(xy)\,dy = \int_0^\pi y\cos(xy)\,dy.$$

$$\int_0^\pi \sin(xy)\,dy = -\frac{\cos(xy)}{x}\Big|_0^\pi = -\frac{\cos(\pi x)}{x} + \frac{1}{x}, \text{ so } \frac{d}{dx}\int_0^\pi \sin(xy)\,dy = \frac{\pi\sin(\pi x)}{x} + \frac{\cos(\pi x)}{x^2} - \frac{1}{x^2}.$$

$$\int_0^\pi y\cos(xy)\,dy = y\frac{\sin(xy)}{x}\Big|_0^\pi - \int_0^\pi \frac{\sin(xy)}{x}\,dy = \frac{\pi\sin(\pi x)}{x} + \left[\frac{\cos(xy)}{x^2}\right]_0^\pi = \frac{\pi\sin(\pi x)}{x} + \frac{\cos(\pi x)}{x^2} - \frac{1}{x^2}.$$

Series of Functions

Series of Real Numbers

A series of real numbers is $\sum_{n=1}^\infty a_n$. The associated sequence of partial sums is $s_n = \sum_{k=1}^n a_k$.
The series is said to converge if the partial sums have a limit s, and s is said to be the sum of the series.

Definition: Absolute Convergence

A series is absolutely convergent if $\sum_{n=1}^\infty |a_n|$ converges.

Theorem

An absolutely convergent series is convergent.

Proof: If $\varepsilon > 0$ is given, then there exists $N > 0$ such that $\sum_{n=k}^l |a_n| < \varepsilon$ whenever $N \leq k \leq l$ (this is the Cauchy criterion of the partial sums of $\sum_{n=1}^\infty |a_n|$). Hence $\left|\sum_{n=k}^l a_n\right| \leq \sum_{n=k}^l |a_n| < \varepsilon$. This is the Cauchy criterion for the convergence of $\sum_{n=1}^\infty a_n$, or equivalently, the partial sums.

Alternating Series Test

If $|a_1| \geq |a_2| \geq \cdots$ and $\lim_{n\to\infty} a_n = 0$, and if the signs of the a_n 's are strictly alternating, then the series $\sum_{n=1}^{\infty} a_n$ converges.

Definition: Series of Functions

If $\{f_k\}_{k=1}^{\infty}$ are functions defined on a subset $S \subseteq \mathbb{R}^n$ with values in $\mathbb{R}$ (or in $\mathbb{R}^m$), then we consider the series $\sum_{k=1}^{\infty} f_n(x)$.

Note: We have the notion of pointwise convergence, uniform convergence, L^p convergence, etc.

Examples

1. Power series: $\sum_{n=0}^{\infty} a_n x^n$ or $\sum_{n=0}^{\infty} a_n (x - x_0)^n$.
2. Fourier series: $f \sim A_0 + \sum_{n=0}^{\infty} (A_n \cos(nx) + B_n \sin(nx))$.

Theorem

If $\sum_{k=1}^{\infty} f_n(x)$ is a series of continuous functions which converges uniformly, then the limit is a continuous function.

Theorem

A series of functions is uniformly convergent if and only if it is uniformly Cauchy.

Theorem

A sequence of functions is uniformly convergent if and only if it is uniformly Cauchy.

Proof:

If $\{s_n\}_{n=1}^{\infty}$ is uniformly convergent, then given $\varepsilon > 0$ there exists $N > 0$ such that $|s_n(x) - s(x)| < \frac{\varepsilon}{2}$ for all x when $n > N$.

Then for $m, n > N$ and for all x , $|s_n(x) - s_m(x)| = |s_n(x) - s(x) - (s_m(x) - s(x))| \leq |s_n(x) - s(x)| + |s_m(x) - s(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

Equivalently, $\|s_n(x) - s_m(x)\|_\infty = \|s_n(x) - s(x) - (s_m(x) - s(x))\|_\infty \leq \|s_n(x) - s(x)\|_\infty + \|s_m(x) - s(x)\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

Now suppose the sequence is uniformly Cauchy. Given $\varepsilon > 0$ there exists $N > 0$ such that for all iven $\varepsilon > 0$ there exists $N > 0$ such that for all $x \in S$, $|s_n(x) - s_m(x)| < \varepsilon$ when $m, n > N$. For each fixed x , $\{s_n(x)\}_{n=1}^{\infty}$ is a Cauchy sequence of real numbers. Hence there exists $s(x) \in \mathbb{R}$ such that $s_n(x) \to s(x)$ as $n \to \infty$. Also $|s_n(x) - s_m(x)| < \varepsilon$ when $m, n > N$ independent of x . Let $m \to \infty$, then $s_m(x) \to s(x)$. Therefore $|s_n(x) - s(x)| \leq \varepsilon$ when $n > N$. Since N is independent of x , the convergence is uniform.

Weierstrass M-Test

Suppose $\{f_n\}_{n=1}^{\infty}$ are functions on $S \subseteq \mathbb{R}^k$ with values in $\mathbb{R}^m$, and suppose $\{M_n\}_{n=1}^{\infty}$ are non-negative real numbers such that $\|f_n\|_\infty = \sup_{x \in S} \|f_n(x)\| \leq m_n$. Then if $\sum_{n=1}^{\infty} M_n < \infty$, then $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on S .

Proof: Note that $\sum_{n=1}^{\infty}\{f_n(x)\}\leq\sum_{n=1}^{\infty}\{f_n(x)\}_{\infty}<\sum_{n=1}^{\infty}M_n$, hence $\sum_{n=1}^{\infty}f_n(x)$ is absolutely convergent. $\left\|f(x)-\sum_{n=1}^{k}f_n(x)\right\|=\left\|\sum_{n=k+1}^{\infty}f_n(x)\right\|\leq\sum_{n=k+1}^{\infty}\|f_n(x)\|\leq\sum_{n=k+1}^{\infty}\|f_n(x)\|_{\infty}\leq\sum_{n=k+1}^{\infty}M_n$ independent of x. Hence $\lim_{k\to\infty}\left\|f(x)-\sum_{n=1}^{k}f_n(x)\right\|_{\infty}\leq\lim_{k\to\infty}\sum_{n=k+1}^{\infty}M_n=0$.

Example

Consider the power series $\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{2n+1!}$. On the interval $[-A,A]$ with $A>0$, $\left|\frac{(-1)^n x^{2n+1}}{2n+1!}\right|\leq\frac{A^{2n+1}}{2n+1!}$. $\sum_{n=1}^{\infty}\frac{A^{2n+1}}{2n+1!}$ converges absolutely by the ratio test for any A. Hence the power series $\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{2n+1!}$ converges uniformly on $[-A,A]$.

Power Series

Root Test

Suppose $c_n\geq 0$. If $\limsup_{n\to\infty}c_n^{\frac{1}{n}}=\gamma$ and $\gamma<1$, then the series $\sum_{n=1}^{\infty}c_n$ converges.

Note: If c_n are not non-negative, we can apply the test to $\sum_{n=1}^{\infty}|c_n|$.

Proof: Choose γ' such that $\gamma<\gamma'<1$. Then $c_n^{\frac{1}{n}}<\gamma'$ for n sufficiently large, and so $c_n<(\gamma')^n$. Now $\sum_{n=1}^{\infty}(\gamma')^n$ is a convergent geometric series. Hence $\sum_{n=1}^{\infty}c_n$ converges by the Comparison Test.

Hadamard's Theorem

If $\sum_{n=1}^{\infty}a_n x^n$ is a power series, then one of the following is true.

1. The series converges for $x=0$ only.
2. The series converges for all $x\in\mathbb{R}$.
3. There exists $R>0$ such that the series converges for $|x|<R$ and diverges for $|x|>R$ (it may or may not converge for $x=\pm R$). The series converges uniformly for $|x|\leq r<R$.

Proof: Let $\alpha=\limsup_{n\to\infty}|a_n|^{\frac{1}{n}}$. We claim the radius of convergence is $R=\begin{cases}\infty & \text{, if } \alpha=0\\ 0 & \text{, if } \alpha=\infty\\ \frac{1}{\alpha} & \text{, if } \alpha\in(0,\infty)\end{cases}$. Apply the root test to $\sum_{n=1}^{\infty}|a_n x^n|$: $\limsup_{n\to\infty}|a_n x^n|^{\frac{1}{n}}=\limsup_{n\to\infty}|x||a_n|^{\frac{1}{n}}=|x|\limsup_{n\to\infty}|a_n|^{\frac{1}{n}}=|x|\alpha$.

1. If $\alpha=0$, then $\alpha|x|=0$ for all x. So the series converges for all x.
2. If $\alpha=\infty$, then $\alpha|x|=\infty$ unless $x=0$. So the series converges only if $x=0$.
3. If $\alpha|x|<1\Leftrightarrow|x|<\frac{1}{\alpha}=R$, then the series converges. If $\alpha|x|>1\Leftrightarrow|x|>\frac{1}{\alpha}=R$, then the series diverges.

If $0<r<R$ then $\sum_{n=1}^{\infty}|a_n|r^n$ converges. Since $|a_n x^n|\leq|a_n|r^n=M_n$, the Weierstrass M-Test implies $\sum_{n=1}^{\infty}a_n x^n$ converges

uniformly on $[-r,r]$.

Theorem

Suppose $\sum_{n=1}^{\infty} H_n$ is a series of C^1 functions on $[a,b]$ such that $\sum_{n=1}^{\infty} H_n{}'$ converges uniformly to a function G and there exists $c\in[a,b]$ such that $\sum_{n=1}^{\infty} H_n(c)=\beta$ exists. Then $\sum_{n=1}^{\infty} H_n$ converges uniformly to a differentiable function H such that $H(c)=\beta$ and $H'=G$.

Application to Power Series (Differentiation)

If $f(x)=\sum_{n=0}^{\infty} a_n x^n$, then $f'(x)=\sum_{n=1}^{\infty} n a_n x^{n-1}$ with the same radius of convergence.

Proof: Let $\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$. It is clear that $\sum_{n=1}^{\infty} n a_n x^{n-1}$ converges if and only if $x\sum_{n=1}^{\infty} n a_n x^{n-1}=\sum_{n=1}^{\infty} n a_n x^n$ converges. Now $\limsup_{n\to\infty} |n a_n|^{\frac{1}{n}}=\limsup_{n\to\infty} n^{\frac{1}{n}}|a_n|^{\frac{1}{n}}=\left(\lim_{n\to\infty} n^{\frac{1}{n}}\right)\left(\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}\right)=\alpha$, hence $\sum_{n=1}^{\infty} n a_n x^{n-1}$ has the same radius of convergence as the original series $\sum_{n=0}^{\infty} a_n x^n$.

Now take $H_n(x)=a_n x^n$ on $[-r,r]$ where $r<R=\frac{1}{\alpha}$. Then $\sum_{n=1}^{\infty} H_n{}'(x)$ converges to a function G on $[-r,r]$, and $\sum_{n=1}^{\infty} H_n(0)=a_0$ converges to a limit (in fact a constant). We conclude that $\sum_{n=0}^{\infty} H_n$ converges to a differentiable function H on $[-r,r]$ such that $H(0)=a_0$ and $H'=G$ ($H(x)=\sum_{n=0}^{\infty} a_n x^n$ and $G(x)=\sum_{n=1}^{\infty} n a_n x^{n-1}$).

Remark

If $f(x)=\sum_{n=0}^{\infty} a_n x^n$ on $(-R,R)$, then $f^{(k)}(x)=\sum_{n=0}^{\infty} n(n-1)\cdots(n-k+1)a_n x^{n-k}$. Also $a_n=\frac{f^{(n)}(0)}{n!}$.

Application to Power Series (Integration)

If $f(x)=\sum_{n=0}^{\infty} a_n x^n$, then $F(x)=\int_0^x f(t)\,dt=\sum_{n=0}^{\infty} a_n\int_0^x t^n\,dt=\sum_{n=0}^{\infty}\frac{a_n}{n+1}x^{n+1}$. The radius of convergence of the integrated series is the same as that of the original series.

Example

Consider $\frac{1}{1+x}=\sum_{n=0}^{\infty}(-1)^n x^n$, $|x|<1$. The convergence is uniform on $[-r,r]$ for any fixed $r\in[0,1)$ (by the Weierstrass M-test since $|(-1)^n x^n|\le r^n$ and $\sum r^n$ converges). There is no convergence at the endpoints.
$F_k(x)=\int_0^x\sum_{n=0}^{k}(-1)^n t^n\,dt=\sum_{n=0}^{k}\frac{(-1)^n x^{n+1}}{n+1}$ converges uniformly on $[-r,r]$ to $F(x)=\int_0^x\sum_{n=0}^{\infty}(-1)^n t^n\,dt=\sum_{n=0}^{\infty}\frac{(-1)^n x^{n+1}}{n+1}$. Now $\int_0^x\frac{1}{1+t}\,dt=\log(1+x)$. Notice that $\log 0$ is not defined and the series diverges when $x=-1$.

When $x=1$, the series $\sum_{n=0}^{\infty}\frac{(-1)^n}{n+1}=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots$ does converge (very slowly) by the alternating series test. Also, $\log(1+x)$ is defined at $x=1$; its value is $\log 2$.
Is it true that $\log 2=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots$?

Abel's Theorem

Let $f(x)=\sum_{n=0}^{\infty}a_n x^n$ on $(0,1)$. Suppose also that $\sum_{n=0}^{\infty}a_n$ converges, i.e. the series converges when $x=1$. Then $\lim_{x\to 1^-}f(x)$ exists and its value is $\sum_{n=0}^{\infty}a_n$.

Approximation by Polynomials

Taylor Series

Suppose $f:[c,d]\to\mathbb{R}$ and f has derivatives of all orders. Let $a\in[c,d]$. Is $f(x)=\sum_{n=0}^{\infty}a_n(x-a)^n$ when $a_n=\frac{f^{(n)}(a)}{n!}$?
Two questions:

1. Where does the power series converge?
2. When it does converge, does it converge to f or to something else?

Note: In complex analysis, if f is analytic in a neighborhood of a point $a\in\mathbb{C}$, then f has a Taylor series expansion about a which converges to f in the disc of center a and radius equal to the distance from a to the nearest singularity of f.

Example

Consider $f(x)=e^{-\frac{1}{x^2}}$. We extend f by defining $f(x)=0=\lim_{x\to 0}e^{-\frac{1}{x^2}}$. One can show that $\left.\frac{d^n}{dx^n}f(x)\right|_{x=0}=0$. Hence the Taylor series of f about $a=0$ is $0+0x+0x^2+\cdots\equiv 0$.

Example: Differential Equation Method

This is sometimes useful in studying convergence of Taylor series to the original function.
Consider $(1+x)^{\alpha}$ where $\alpha\in\mathbb{R}$. Recall that $(1+x)^{-1}=\sum_{n=0}^{\infty}x^n$ is valid in $|x|<1$. Notice that $g(x)=(1+x)^{\alpha}$ satisfies $(1+x)g'(x)=\alpha g(x)$, $g(0)=1$. Solve this differential equation by power series method. Try $f(x)=\sum_{n=0}^{\infty}a_n x^n$ where the coefficients are to be determined. $f'(x)=\sum_{n=0}^{\infty}n a_n x^{n-1}\Rightarrow(1+x)f'(x)=\sum_{n=0}^{\infty}\left[n a_n x^{n-1}+n a_n x^n\right]$. So we have $(1+x)f'(x)=\alpha f(x)\Rightarrow\sum_{n=0}^{\infty}\left[n a_n x^{n-1}+n a_n x^n\right]=\alpha\sum_{n=0}^{\infty}a_n x^n\Rightarrow\sum_{n=0}^{\infty}\left((n+1)a_{n+1}+n a_n-\alpha a_n\right)x^n=0\ \ \forall x$. Hence for $n=0,1,\ldots$, $(n+1)a_{n+1}+n a_n-\alpha a_n=0\Rightarrow a_{n+1}=\frac{\alpha-n}{n+1}a_n$. We know $a_0=1$ since $(1+x)^{\alpha}=0$ when $x=0$, and therefore we know all the a_n's; $a_n=\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!}=\binom{\alpha}{n}$.
If α is a positive integer (say $\alpha=N>0$), then the series terminates and coincides with the expansion of $(1+x)^N$.

Also, $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|=\lim_{n\to\infty}\left|\frac{\alpha-n}{n+1}\right|=1$. Therefore the radius of convergence is 1.

$f(x)=\sum_{n=0}^{\infty} a_n x^n=\sum_{n=0}^{\infty}\binom{\alpha}{n}x^n$ solves the differential equation $(1+x)f'(x)=\alpha f(x)$ and satisfies the initial condition $f(0)=1$. To show $f(x)=(1+x)^\alpha$, we show that $\frac{f(x)}{(1+x)^\alpha}=\text{constant}$. The constant must be 1 since $f(0)=1$ and $(1+0)^\alpha=1$. We take $\frac{d}{dx}\left[\frac{f(x)}{(1+x)^\alpha}\right]=\frac{f'(x)(1+x)^\alpha-f(x)\alpha(1+x)^{\alpha-1}}{(1+x)^2\alpha}=\frac{(1+x)^{\alpha-1}\alpha f(x)-f(x)\alpha(1+x)^{\alpha-1}}{(1+x)^2\alpha}=0$.

Definition: Taylor Series

If f have derivatives of all orders in a neighborhood of a point a , the Taylor series of f at a is $\sum_{k=0}^{\infty}\frac{f^{(k)}(a)}{k!}(x-a)^k$.

Definition: Taylor Polynomial

$P_n(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k$ is the Taylor polynomial of order n .

Theorem

If $f\in C^{n+1}[A,B]$, and if M is a bound for $|f^{(n+1)}(x)|$ on $[A,B]$, and if $a\in[A,B]$, then $R_n(x)=f(x)-P_n(x)$ satisfies $|R_n(x)|\le M\frac{|x-1|^{n+1}}{(n+1)!}$.

Proof: $P_n(x)$ has the same values as f for its derivatives up to order n at the point a , therefore $R_n^{(k)}(a)=f^{(k)}(a)-P_n^{(k)}(a)$ for $k=0,1,\ldots,n$.

Also $R_n^{(n+1)}(x)=f^{(n+1)}(x)-P_n^{(n+1)}(x)=f^{(n+1)}(x)$ since $P_n(x)$ is a polynomial of degree n .

Now by FTC, $R_n^{(n)}(x)=R_n^{(n)}(a)+\int_a^x R_n^{(n+1)}(t)\,dt=0+\int_a^x f_n^{(n+1)}(t)\,dt$. So $|R_n^{(n)}(x)|\le\left|\int_a^x M\,dt\right|=M|x-a|$.

Suppose we have shown that for some k , $0\le k<n$, $|R_n^{(n-k)}(x)|\le\frac{M|x-a|^{k+1}}{(k+1)!}$. Now $R_n^{(n-k-1)}(x)=R_n^{(n-k)}(a)+\int_a^x R_n^{(n-k)}(t)\,dt$,

so $|R_n^{(n-k-1)}(x)|\le\left|\int_a^x M\frac{|t-a|^{k+1}}{(k+1)!}dt\right|=M\frac{|x-a|^{k+1}}{(k+2)!}$. By induction the estimate is true for $k=0,1,\ldots,n$.

Hence $|R_n(x)|\le M\frac{|x-1|^{n+1}}{(n+1)!}$.

How Not to Approximate a Function

The Taylor polynomial use information only at one point.

Consider a continuous function f on $[a,b]$. Pick $x_1,\ldots,x_n$ in $[a,b]$. There is an unique polynomial ϕ_n of degree n such that $\phi_n(x)=f(x_j)$. Does this procedure produce a family of polynomials which converge to f in the uniform norm? Not necessarily.

Weierstrass Approximation Theorem

Given a continuous function f on $[a,b]$, there exists a sequence polynomials $\{\psi_n\}_{n=1}^{\infty}$ such that $\|\phi_n-f\|_\infty\to 0$ as $n\to\infty$.

Proof: We'll give a proof using Bernstein's polynomials. Note that it is sufficient to prove the theorem when $[a,b]=[0,1]$.

Bernstein's Proof of the Weierstrass Theorem

Bernstein Polynomials

Recall $(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}$. Set $y=1-x$. Then $(x+1-x)^n=\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}=1$. Let $P_k^n(x)=\binom{n}{k}x^k(1-x)^{n-k}$, $k=0,1,\ldots,n$. These are the Bernstein polynomials. Notice that:

- P_k^n is a polynomial of degree n .
- $P_k^n(x)\geq 0$ on $[0,1]$; zeros can occur only at the endpoints.
- P_k^n has a maximum at $\frac{k}{n}$, $k=0,1,\ldots,n$.

If f is a continuous function on $[0,1]$, we define the polynomial $B_n f(x)=\sum_{k=0}^{n} f\left(\frac{k}{n}\right)P_k^n(x)$.

Properties

1. $B_n f$ is a polynomial of degree at most n .
2. $B_n(f+g)=B_n f+B_n g$.
3. $B_n(c f)=c B_n f$
4. $B_n f\geq 0$ if $f\geq 0$.
5. $B_n f\geq B_n g$ if $f\geq g$.
6. $|B_n f|\leq B_n g$ if $|f|\leq g$.

Lemma

1. $B_n 1=1$.
2. $B_n x=x$.
3. $B_n x^2=\frac{n-1}{n}x^2+\frac{1}{n}x=x^2+\frac{x-x^2}{n}$.

Proof:

1. $B_n 1=\sum_{k=0}^{n}P_k^n(x)=1$.
2. Note that $\frac{k}{n}\binom{n}{k}=\frac{k}{n}\frac{n!}{k!(n-k)!}=\frac{(n-1)!}{(k-1)!(n-k)!}=\binom{n-1}{k-1}$. So
$$B_n x=\sum_{k=0}^{n}\frac{k}{n}\binom{n}{k}x^k(1-x)^{n-k}=x\sum_{k=1}^{n}\binom{n-1}{k-1}x^{k-1}(1-x)^{n-k}=x\sum_{k=0}^{n-1}\binom{n-1}{k}x^k(1-x)^{n-1-k}=x(x+(1-x))^{n-1}=x\cdot$$
3. Note that $\frac{k^2}{n^2}\binom{n}{k}=\frac{k-1+1}{n}\frac{(n-1)!}{(k-1)!(n-k)!}=\frac{n-1}{n}\frac{1}{n-1}\frac{(n-1)!}{(k-12)!(n-k)!}+\frac{1}{n}\frac{(n-1)!}{(k-1)!(n-k)!}=\frac{n-1}{n}\binom{n-2}{k-2}+\frac{1}{n}\binom{n-1}{k-1}$. So
$$B_n x^2=\sum_{k=0}^{n}\frac{k^2}{n^2}\binom{n}{k}x^k(1-x)^{n-k}=\frac{n-1}{n}\sum_{k=2}^{n}\binom{n-2}{k-2}x^k(1-x)^{n-k}+\frac{1}{n}\sum_{k=1}^{n}\binom{n-1}{k-1}x^k(1-x)^{n-k}.$$
$$=\frac{n-1}{n}x^2\sum_{k=0}^{n-2}\binom{n-2}{k}x^k(1-x)^{n-2-k}+\frac{1}{n}x\sum_{k=0}^{n-1}\binom{n-1}{k}x^k(1-x)^{n-1-k}=\frac{n-1}{n}x^2+\frac{1}{n}x$$

Theorem

If $f\in C[0,1]$, then $B_n f\to f$ uniformly as $n\to\infty$.

Proof: We want to show for all $\varepsilon>0$, there exists $N>0$ such that $\|f-B_n f\|_\infty<\varepsilon$ for all $n\ge N$.
Note that f is uniformly continuous on $[0,1]$, i.e. given $\varepsilon>0$ there exists $\delta>0$ such that $|f(x)-f(y)|<\varepsilon$ whenever $|x-y|<\delta$ and $x,y\in[0,1]$.

Let $M=\|f\|_\infty<\infty$. Fix $a\in[0,1]$. Then $|f(x)-f(a)|\le\frac{\varepsilon}{2}+\frac{2M}{\delta^2}(x-a)^2$ for all $x\in[0,1]$. If $|x-a|<\delta$, this is true. If $|x-a|\ge\delta$, then $|f(x)-f(a)|\le 2M\le 2M\frac{(x-a)^2}{\delta^2}\le\frac{\varepsilon}{2}+2M\frac{(x-a)^2}{\delta^2}$.

Recall that $|B_n u|\le B_n v$ if $|u|\le v$. So

$$|B_n[f-f(a)](x)|=|B_n f(x)-f(a)|\le B_n\left(\frac{\varepsilon}{2}+\frac{2M}{\delta^2}(x-a)^2\right)=\frac{\varepsilon}{2}+\frac{2M}{\delta^2}\left(x^2+\frac{x-x^2}{n}-2ax+a^2\right)=\frac{\varepsilon}{2}+\frac{2M}{\delta^2}\left((x-a)^2+\frac{x-x^2}{n}\right).$$

Set $x=a$. Then $|B_n f(a)-f(a)|\le\frac{\varepsilon}{2}+\frac{2M}{\delta^2}\left(\frac{a-a^2}{n}\right)$. The function $\phi(a)=a-a^2$ has its maximum on $[0,1]$ when $a=\frac{1}{2}$ with a value of $\phi\left(\frac{1}{2}\right)=\frac{1}{4}$. Therefore $|B_n f(a)-f(a)|\le\frac{\varepsilon}{2}+\frac{2M}{\delta^2}\left(\frac{1}{4n}\right)=\frac{\varepsilon}{2}+\frac{M}{2n\delta^2}$. The RHS is independent of a.

Choose $N>0$ such that $\frac{M}{2N\delta^2}<\frac{\varepsilon}{2}\Leftrightarrow N>\frac{M}{\delta^2\varepsilon}$, then for all $n\ge N$, $\frac{\varepsilon}{2}+\frac{M}{2n\delta^2}<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$.

Fourier Series

Definition: Fourier Series

Let $f,g\in C[-\pi,\pi]$. Let $\langle f,g\rangle=\int_{-\pi}^{\pi} f(\theta)g(\theta)\,d\theta$, then $\{1,\sqrt{2}\cos(n\theta),\sqrt{2}\sin(n\theta)\}$ $n\in\mathbb{N}$ is an orthonormal system on $[-\pi,\pi]$. Define the Fourier series of f to be $f\sim A_0+\sum_{n=1}^{\infty}(A_n\cos n\theta+B_n\sin n\theta)$, where $A_0=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\,d\theta$, $A_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos n\theta\,d\theta$, $B_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin n\theta\,d\theta$.

Note: The formulas make sense for piece-wise continuous functions; in fact for all absolutely integrable functions (i.e. $\int_{-\pi}^{\pi}|f(\theta)|\,d\theta<\infty$).

Theorem

Any continuous function on $[-\pi,\pi]$ which is periodic with period 2π can be uniformly approximated by trigonometric polynomials $C_0+\sum_{n=1}^{N}(C_n\cos n\theta+D_n\sin n\theta)$. More precisely,

1. For all $\varepsilon>0$ there exists a trigonometric polynomial $p(\theta)$ such that $|f(\theta)-p(\theta)|<\varepsilon$ for all $\theta\in[-\pi,\pi]$, i.e. $\|f-p\|_\infty<\varepsilon$.
2. There exists a sequence of trigonometric polynomial $p_n(\theta)$ such that $\|f-p_n\|_\infty\to 0$ as $n\to\infty$.

Note: This does not say that the Fourier series of f converges uniformly (or pointwise) to f.

Corollary

The Fourier series of f converges to f in L^2.
Note: This need not imply pointwise convergence.
Note: This remains true for piecewise continuous functions.

Proof: Let $\{\phi_N\}_{N=1}^{\infty}$ be the sequence of trigonometric polynomials which converge uniformly to f. ϕ_N involves 1,

$\cos n\theta$, and $\sin n\theta$ for $1 \le n \le N$. Uniform convergence on a finite interval implies L^2 convergence, hence $\phi_N \to f$ in L^2 .

Let S_N be the partial sums of the Fourier series of f, i.e. $S_N(\theta)=A_0+\sum_{n=1}^{N}(A_n\cos n\theta+B_n\sin n\theta)$. By the Projection Theorem, $\|f-S_N\|^2 \le \|f-\phi_N\|^2$; in fact, $\|f-\phi_N\|^2=\|f-S_N\|^2+\|S_N-\phi_N\|^2$. Since $\|f-\phi_N\|_2 \to 0$ as $N\to\infty$, hence $\|f-S_N\|_2\to 0$ as $N\to\infty$.

Corollary

$\frac{1}{2\pi}\int_{-\pi}^{\pi}(f(\theta))^2\,d\theta=A_0^2+\frac{1}{2}\sum_{n=1}^{N}(A_n^2+B_n^2)$. Hence if the Fourier series of a (piecewise) continuous function is 0, then the function is 0. Therefore if f and g are (piecewise) continuous functions with the same Fourier series coefficients, then $f\equiv g$.

Proof: Follows from the Fourier series of f converging to f in L^2 .

Proposition

Suppose the Fourier series of a continuous periodic function f converges uniformly to some function. Then this function mush be f.

Proof: Let $g(\theta)=A_0+\sum_{n=1}^{\infty}(A_n\cos n\theta+B_n\sin n\theta)$. g is continuous and periodic. What are its Fourier coefficients?

$\frac{1}{\pi}\int_{-\pi}^{\pi}g(\theta)\,d\theta=\frac{1}{\pi}\int_{-\pi}^{\pi}\left[A_0+\sum_{k=1}^{\infty}(A_n\cos k\theta+B_n\sin k\theta)\right]\cos n\theta\,d\theta=A_n$ after interchanging summation and integration.

Similarly for the other Fourier coefficients of g . Therefore $f\equiv g$ since the have the same Fourier coefficients.

Lemma

The Fourier coefficients of an absolutely integrable function are bounded by $|A_0|\le\|f\|_1$, $|A_n|\le 2\|f\|_1$, $|B_n|\le 2\|f\|_1$, $n\ge 1$.

Proof: $|A_0|=\left|\frac{1}{2\pi}\int_{-\pi}^{\pi}f(\theta)\,d\theta\right|\le\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(\theta)|\,d\theta=\|f\|_1$.

Theorem

If f is a (piecewise) C^1 function which is 2π -periodic and $f\sim A_0+\sum_{n=1}^{\infty}(A_n\cos n\theta+B_n\sin n\theta)$, then $f'\sim\sum_{n=1}^{\infty}(nB_n\cos n\theta-nA_n\sin n\theta)$.

Proof: Let $f'\sim a_0+\sum_{n=1}^{\infty}(a_n\cos n\theta+b_n\sin n\theta)$.

$a_0=\frac{1}{2\pi}\int_{-\pi}^{\pi}f'(\theta)\,d\theta=\frac{1}{2\pi}[f(\theta)]_{-\pi}^{\pi}=0$ since f is periodic.

$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f'(\theta)\cos n\theta\,d\theta=\frac{1}{\pi}[f(\theta)\cos n\theta]_{-\pi}^{\pi}+\frac{1}{\pi}\int_{-\pi}^{\pi}f(\theta)\,n\sin n\theta\,d\theta=0+nB_n=nB_n$.

$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f'(\theta)\sin n\theta\, d\theta = \frac{1}{\pi}[f(\theta)\sin n\theta]_{-\pi}^{\pi} - \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta) n\cos n\theta\, d\theta = 0 - nA_n = -nA_n$.

Theorem
If f is a C^2 function which is 2π -periodic, then the Fourier series of f converges uniformly to f. In particular, it converges pointwise to f.

Proof: $f \sim A_0 + \sum_{n=1}^{\infty} (A_n \cos n\theta + B_n \sin n\theta)$, $f' \sim \sum_{n=1}^{\infty} (nB_n \cos n\theta - nA_n \sin n\theta)$, $f'' \sim -\sum_{n=1}^{\infty} (n^2 A_n \cos n\theta - n^2 B_n \sin n\theta)$.
f'' is continuous on $[-\pi, \pi]$, hence $\|f''\|_\infty = M < \infty$. Hence $|-n^2 A_n| \le \frac{1}{\pi} M \Rightarrow |A_n| \le \frac{M}{\pi n^2}$ and $|-n^2 B_n| \le \frac{1}{\pi} M \Rightarrow |B_n| \le \frac{M}{\pi n^2}$. The the Weierstrass M-test, the series $A_0 + \sum_{n=1}^{\infty} (A_n \cos n\theta + B_n \sin n\theta)$ converges uniformly.
Since it converges uniformly, it must converge to f.

Pointwise Convergence of Fourier Series
If f satisfies a Lipschitz condition on $[-\pi, \pi]$, then the Fourier series of f converges pointwise to f.

Metric Spaces

Definitions and Examples

Definition: Metric Space
A metric space is a set X together with a distance function $\rho: X \times X \to [0, \infty)$ such that

- $\rho(x, y) \ge 0$, $\rho(x, y) = 0 \Leftrightarrow x = y$;
- $\rho(x, y) = \rho(y, x)$;
- $\rho(x, z) \le \rho(x, y) + \rho(y, z)$ (triangle inequality).

Remark
Any normed vector space is a metric space.
Any subset of a normed vector space is a metric space where $\rho(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$.

Example
The surface of a sphere in $\mathbb{R}^3$ can be made into a metric space into two natural ways:

1. Define $\rho(x, y)$ to be the straight line distance (by tunneling through).
2. Define $\rho(x, y)$ to be the distance from x to y along the circle joining them.

Definition: Open Ball
Let X be a metric space. An open ball of center a and radius r is $B_r(a) = \{x \in X | \rho(x, a) < r\}$.

Definition: Open Set
A set A is open if for all $a \in A$ there exists $r > 0$ such that $B_r(a) \subseteq A$.

Definition: Continuity

A function $f: X \to \mathbb{R}$ is continuous at a if for all $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$ whenever $\rho(x, a) < \delta$.

More generally, if X and Y are metric spaces and $f: X \to Y$, then f is continuous at a if for all $\varepsilon > 0$ there exists $\delta > 0$ such that $\rho_Y(f(x), f(a)) < \varepsilon$ whenever $\rho_X(x, a) < \delta$.

f is continuous if it is continuous at all $a \in X$.

Definition: Convergence

If $\{x_n\}_{n=1}^{\infty}$ is a sequence in X, we say $x_n \to x \in X$ if $\varepsilon > 0$ there exists $N > 0$ such that $\rho_Y(x_n, x) < \varepsilon$ whenever $n > N$.

Definition: Cauchy

A sequence $\{x_n\}_{n=1}^{\infty}$ in X is Cauchy if $\varepsilon > 0$ there exists $M > 0$ such that $\rho_Y(x_n, x_m) < \varepsilon$ whenever $n, m > N$.

Definition: Completeness

A metric space X is complete if any Cauchy sequence converges to a point in X.

Definition: Limit Point

Let $A \subset X$ and $a \in A$. a is a limit point of A if there is a sequence of points $\{a_n\}_{n=1}^{\infty}$ in A such that $a_n \to a$ as $n \to \infty$.

Definition: Closed Set

A subset $A \subset X$ is closed if it contains all its limit points.

Theorem

A subset $A \subset X$ is open if and only if the complement of A is closed.

Theorem: Characterizations of Continuity

Let X and Y be metric spaces. Let $f: X \to Y$ be a mapping. The following are equivalent:

1. f is continuous.
2. Whenever $\{x_n\}_{n=1}^{\infty}$ is a sequence in X such that $x_n \to a \in X$, $f(x_n) \to f(a) \in Y$.
3. Whenever U is an open subset of Y, $f^{-1}(U)$ is an open subset of X.

Compact Metric Spaces

Definition: Open Cover

An open cover of a metric space X is a collection of open sets $\{U_\alpha\}_{\alpha \in B}$ such that $X \subseteq \bigcup_{\alpha \in B} U_\alpha$.

Definition: Compact

A metric space X (or a subspace $A \subset X$) is compact if any open cover of X (or of A) has a finite subcover.
Note: "Subspace" here do not mean vector subspace.

Example

$X = (0, 1)$ is not compact. Take $U_n = \left(\frac{1}{n}, 1 - \frac{1}{n}\right)$, then $\bigcup_{n=2}^{\infty} U_n = (0, 1)$. But if we take finitely many of these sets, say

$U_{n_1}, \ldots U_{n_k}$ and let $N = \max\{n_1, \ldots, n_k\}$, then $U_{n_j} \subseteq U_N \quad j = 1, \ldots, k$. So $\bigcup_{j=1}^{k} U_{n_j} = U_N \neq (0, 1)$.

Example
$X = [0, 1]$ is compact in this sense.

Definition: Sequentially Compact
A metric space X is sequentially compact if every sequence $\{x_n\}_{n=1}^{\infty}$ in X has a subsequence $\{x_{n_j}\}_{j=1}^{\infty}$ which converges to a point in X.
Note: In a metric space, compactness is equivalent to sequential compactness.

Theorem
If $f : X \rightarrow Y$ is a continuous mapping between metric spaces X and Y and X is compact, then $f(X)$ is compact.

Proof: Let $\{U_\alpha\}_{\alpha \in B}$ be an open covering of $f(X)$. Then $f^{-1}(U_\alpha) \quad \alpha \in B$ is open by continuity, and $\{f^{-1}(U_\alpha)\}_{\alpha \in B}$ is an open covering of X. Since X is compact, finitely many of the $f^{-1}(U_\alpha)$, say $f^{-1}(U_{\alpha_1}), \ldots, f^{-1}(U_{\alpha_n})$ cover X. Hence $U_{\alpha_1}, \ldots, U_{\alpha_n}$ cover $f(X)$, i.e. $f(X) \subset \bigcup_{j=1}^{n} U_{\alpha_j}$. This is true for all open coverings of $f(X)$, so $f(X)$ is compact.