Convex Function, L^p Spaces, Space of Continuous Functions, Lusin's Theorem

By Ng Tze Beng

To study the space of functions on $\mathbb{R}^n$ from the point of view of continuity and measurability, we may start by considering the space of functions on the topological space *X*, which is in some way like the Euclidean space $\mathbb{R}^n$ for integer $n \geq 1$, with the usual metric topology. As for the measure on the topological space *X*, we may consider one that has the properties that the Lebesgue measure on $\mathbb{R}^n$ has, namely it must contain the Borel subsets of *X*, is finite on compact subsets, regular and complete. The results that hold true for the space of functions on such a topological space *X* with or without the required measure are true when *X* is replaced by $\mathbb{R}^n$.

The closure of any open *n*-disk in $\mathbb{R}^n$ is compact by the Heine-Borel Theorem. Any neighbourhood *N*, of a point *x* in $\mathbb{R}^n$, contains a closed *n*-disk, which is compact and is also a neighbourhood of *x*. To take this property into the topology of the space *X*, we require that for any point $x \in X$, *x* has a compact neighbourhood. We say *X* is *locally compact* if every point in *x* has a compact neighbourhood. We know that $\mathbb{R}^n$ with the usual topology is *Hausdorff*, that is, any two distinct points in $\mathbb{R}^n$ can be separated by open neighbourhoods. So, we would want our space *X* to be Hausdorff, that is, if *x* and *y* are distinct points in *X*, then there are open neighbourhoods *U* and *V* of *x* and *y* respectively such that $U \cap V = \varnothing$. Hence, we would want *X* to be a locally compact Hausdorff space. Now a compact subset of a Hausdorff space *X* is closed in *X* and a closed subset of a compact set is compact. Hence a compact neighbourhood *K* of a point *x* in a Hausdorff space is closed. There is an open set *U* such that $x \in U \subseteq K$. As $\overline{U}$ is closed in *X* and is contained in *K*, $\overline{U}$ is compact and so $\overline{U}$ is a compact neighbourhood of *x*. We say a subset *E* of *X* is *relatively compact* if its closure $\overline{E}$ in *X* is compact. Hence a Hausdorff space is locally compact if and only if every point *x* in *X* has a relatively compact open neighbourhood. It is now clear that $\mathbb{R}^n$ with the usual topology is a locally compact Hausdorff topological space.

We shall investigate the space of functions on a locally compact Hausdorff topological space when the space of functions is endowed with a suitable norm to form a *normed linear space*.

A *norm* on a linear space, i.e., a vector space *V* over $\mathbb{R}$ or $\mathbb{C}$, is a non-negative function $\|.\|: V \to [0,\infty)$ such that

(a) $\|a+b\| \leq \|a\| + \|b\|$ for all $a, b \in V$ (triangle inequality),

(b) $\|\alpha x\| = |\alpha| \|x\|$ for any *x* in *V* and any scalar α,

(c) $\|x\| = 0 \Leftrightarrow x = 0$.

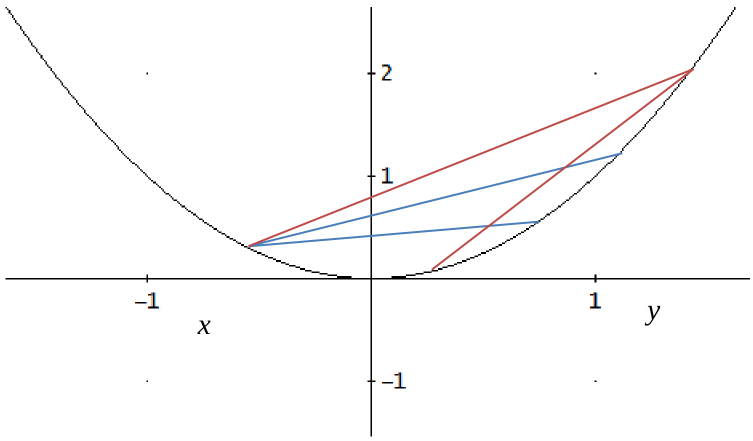
The vector space, V, with such a norm is called a *normed linear space.*

A norm on a linear space V gives rise to a metric on V. We shall investigate when a linear space can be regarded as a *complete* metric space with the metric induced by its norm, i.e., if every Cauchy sequence in V is convergent in V. A normed linear space is a *Banach* space if it is complete as a metric space with the metric induced by its norm.

We begin with the investigation into, among other things, some inequalities including the triangle inequality that the norm, which we shall define and elaborate later, must satisfy. The property of a convex function plays a very useful and effective role in proving some of these inequalities. Our first section will be devoted to convex functions.

1. Convex Functions.

In Calculus, the graph of a real valued function, f, on an open interval, is said to be *concave upward*, if f is differentiable and its derivative is an increasing function or the graph of f lies above each of its tangent line. A typical example is the function $f(x) = x^2$.



A careful examination of the above graph will reveal that it satisfies the property (3) stated in the next definition.

Definition 1. Let $-\infty \le a < b \le \infty$. A function $f:(a,b) \to \mathbb{R}$ is *convex* if

(1) $\forall x, y \in (a,b)$, for $0 \le \lambda \le 1$, $f((1-\lambda)x + \lambda y) \le (1-\lambda) f(x) + \lambda f(y)$, or equivalently,

(2) for $a < x < z < y < b$, the chord function

$$p(x,z) = \frac{f(x) - f(z)}{x - z} \le p(z,y) = p(y,z) = \frac{f(y) - f(z)}{y - z} ,$$

or equivalently,

(3) the chord slope function, $p(x,z)=\dfrac{f(x)-f(z)}{x-z}$, is monotone increasing on each argument with the other kept fixed.

We shall show later that a function whose graph is concave upward is a convex function. Hence our example, the function $f(x)=x^2$, is convex on $(-\infty, \infty)$.

We now show that the three conditions above are equivalent.

(1) $\Rightarrow$ (2)

Suppose condition (1) is satisfied for the function f.

For $a<x<z<y<b$

$p(x,z)=\dfrac{f(x)-f(z)}{x-z}\le p(z,y)=p(y,z)=\dfrac{f(y)-f(z)}{y-z}$ if and only if

$$(y-z)\big(f(x)-f(z)\big)\ge\big(f(y)-f(z)\big)(x-z)$$

$$\Leftrightarrow (x-y)f(z)\ge(x-z)f(y)+(z-y)f(x)$$

$$\Leftrightarrow f(z)\le\frac{z-x}{y-x}f(y)+\frac{y-z}{y-x}f(x) \quad \text{----------------} \quad (*)$$

If we take $\lambda=\dfrac{z-x}{y-x}$, then $0\le\lambda\le1$, $1-\lambda=\dfrac{y-z}{y-x}$ and $(1-\lambda)x+\lambda y=\dfrac{y-z}{y-x}x+\dfrac{z-x}{y-x}y=z$. Thus (1) implies $f(z)=f\big((1-\lambda)x+\lambda y\big)\le\lambda f(y)+(1-\lambda)f(x)=\dfrac{z-x}{y-x}f(y)+\dfrac{y-z}{y-x}f(x)$ and by (*) implies (2).

From (*) we have for $a<x<z<y<b$,

$$p(x,z)=\frac{f(x)-f(z)}{x-z}\le p(z,y)=p(y,z)=\frac{f(y)-f(z)}{y-z}$$

$$\Leftrightarrow f(z)-f(x)\le\frac{z-x}{y-x}f(y)+\frac{x-z}{y-x}f(x)$$

$$\Leftrightarrow \frac{f(z)-f(x)}{z-x} \le \frac{1}{y-x} f(y) + \frac{-1}{y-x} f(x) = \frac{f(y)-f(x)}{y-x}, \qquad \text{--------------- (**)}$$

and

$$p(x,z) = \frac{f(x)-f(z)}{x-z} \le p(z,y) = p(y,z) = \frac{f(y)-f(z)}{y-z}$$

$$\Leftrightarrow f(z) - f(y) \le \frac{(z-y)}{y-x} f(y) + \frac{(y-z)}{y-x} f(x)$$

$$\Leftrightarrow \frac{f(z)-f(y)}{z-y} \ge \frac{1}{y-x} f(y) + \frac{-1}{y-x} f(x) = \frac{f(y)-f(x)}{y-x} = \frac{f(x)-f(y)}{x-y}. \quad \text{----- (***)}$$

Note that (2) ⇔ (*) ⇔ (**) ⇔ (***)

(2) ⇒(3).

By keeping *x* fixed we see that for all $x < z < y$, $p(x,z) \le p(x,y)$ by (**) and by keeping *y* fixed, for all $x < z < y$, $p(y,x) = p(x,y) \le p(y,z)$ by (***) and together with $p(z,x) \le p(z,y)$, we see that the chord function $p(x, y)$ is monotone increasing in the second argument with the first kept fixed, i.e. on $(a, b) - \{x\}$. This is because if $y > z > x$, then $p(x,y) \ge p(x,z)$; if $x > y > z$, then $p(x,y) \ge p(x,z)$ and if $y > x > z$, then $p(x,y) \ge p(x,z)$. Since $p(x, y)$ is symmetric it follows that it is also monotone increasing in the first argument with the second kept fixed.

(3) ⇒ (2),

This is obvious.

(2) ⇒ (1)

Assuming (2). Then if we take $z = (1-\lambda)x + \lambda y$, $0 \le \lambda \le 1$, $\lambda = \frac{z-x}{y-x}$ and $1-\lambda = \frac{y-z}{y-x}$, assuming $x < y$. Then $x < z < y$ and by (*)

$$f\left((1-\lambda)x + \lambda y\right) = f(z) \le \frac{z-x}{y-x} f(y) + \frac{y-z}{y-x} f(x) = \lambda f(y) + (1-\lambda) f(x).$$

If $y < x$, then letting $\mu = 1-\lambda$, we have $z = \mu x + (1-\mu)y$, $\mu = \frac{z-y}{x-y}$, $1-\mu = \frac{x-z}{x-y}$ and $y < z < x$.

Assuming (2) we have $p(y,z)=\dfrac{f(y)-f(z)}{y-z}\le p(z,x)=p(x,z)=\dfrac{f(x)-f(z)}{x-z}$ and by (*) we have

$$f(z)\le\frac{z-y}{x-y}f(x)+\frac{x-z}{x-y}f(y)=\mu f(x)+(1-\mu)f(y)=\lambda f(y)+(1-\lambda)f(x)$$

and so $f\left((1-\lambda)x+\lambda y\right)\le\lambda f(y)+(1-\lambda)f(x)$.

Properties of a convex function $f:(a,b)\to\mathbb{R}$, $-\infty\le a<b\le\infty$.

(1) The left and right derivatives of f exist (finite) at each point of (*a*, *b*).

(2) The left derived function and the right derived function, $f_-'(x)$ and $f_+'(x)$, are equal except possibly for a countable number of points *x* in (*a*, *b*).

(3) The left derived function and the right derived function, $f_-'(x)$ and $f_+'(x)$, are both increasing function on (*a*, *b*) and $f_-'(x)\le f_+'(x)$ and in particular, if $x<y$, then $f_+'(x)\le f_-'(y)$.

(4) f is continuous on (*a*,*b*).

A *support line* at a point (*x*, *y*) on a curve Γ in $\mathbb{R}^2$ is a line ℓ through (*x*, *y*) such that Γ lies either on or entirely on one side of the line ℓ.

(5) If $f:(a,b)\to\mathbb{R}$ is convex, then for any *x* in (*a*, *b*), any line with a slope α such that $f_-'(x)\le\alpha\le f_+'(x)$ and passing through $\left(x,f(x)\right)$ on its graph is a support line at $\left(x,f(x)\right)$, where the graph lies either on or entirely above the line. i.e., $f(y)\ge(y-x)\alpha+f(x)$.

Proof.

Property (1)

Take *z* in (*a*, *b*), then for any *x* and *y* such that $a<x<z<y<b$,

$$p(x,z)=\frac{f(x)-f(z)}{x-z}\le p(z,y)=p(y,z)=\frac{f(y)-f(z)}{y-z}.\ \text{----------------------}\ (1)$$

So, by fixing *y*, $p(x,z)=\dfrac{f(x)-f(z)}{x-z}$ is bounded above by $p(y,z)=\dfrac{f(y)-f(z)}{y-z}$. Since $p(x,z)$ is increasing in the first argument, $\lim\limits_{x\to z^-}p(x,z)=\lim\limits_{x\to z^-}\dfrac{f(x)-f(z)}{x-z}$ exists. That is, $f_-'(z)$ exists.

Similarly, by fixing *x*, $p(y,z)=\dfrac{f(y)-f(z)}{y-z}$ is bounded below by $p(x,z)=\dfrac{f(x)-f(z)}{x-z}$.

Since $p(y,z)$ is increasing in the first argument, $\lim\limits_{y\to z^+} p(y,z)=\lim\limits_{y\to z^+}\dfrac{f(y)-f(z)}{y-z}=f_+'(z)$ exists since $\dfrac{f(y)-f(z)}{y-z}$ is decreasing as *y* decreases to *z*.

Properties (2) and (3)

Moreover, $\dfrac{f(x)-f(z)}{x-z}\le f_+'(z)$ for all *x* < *z* in (*a*, *b*) since by (1) $\dfrac{f(x)-f(z)}{x-z}$ is a lower bound for $\left\{\dfrac{f(y)-f(z)}{y-z} : y\in(z,b)\right\}$ and so

$$\frac{f(x)-f(z)}{x-z}\le \inf\left\{\frac{f(y)-f(z)}{y-z} : y\in(z,b)\right\}=f_+'(z).$$

It follows that $f_+'(z)$ is an upper bound for $\left\{\dfrac{f(x)-f(z)}{x-z} : x\in(a,z)\right\}$ and so

$$f_+'(z)\ge \sup\left\{\frac{f(x)-f(z)}{x-z} : x\in(a,z)\right\}=f_-'(z).$$

Hence, $f_-'(z)\le f_+'(z)$ for all *z* in (*a*, *b*) --- (2)

As $\dfrac{f(z)-f(x)}{z-x}$ is decreasing as *z* decreases to *x*, for any *x* < *y* in (*a*, *b*)

$$f_+'(x)=\lim_{z\to x^+}\frac{f(z)-f(x)}{z-x}\le p(x,y)=\frac{f(y)-f(x)}{y-x}$$

and as $\dfrac{f(y)-f(z)}{y-z}$ is increasing as *z* increases to *y*, so that for any *x* < *y* in (*a*, *b*) by we get

$$\frac{f(x)-f(y)}{x-y}\le f_-'(y)=\lim_{z\to y^-}\frac{f(z)-f(y)}{z-y}.$$

Therefore, for any *x* < *y* in (*a*, *b*), $f_+'(x)\le p(x,y)=\dfrac{f(y)-f(x)}{y-x}\le f_-'(y)$.

It then follows from (2) that for *x* < *y* in (*a*, *b*),

$$f_-'(x)\le f_+'(x)\le p(x,y)=\frac{f(x)-f(y)}{x-y}\le f_-'(y)\le f_+'(y)\,.\quad \text{----------- (3)}$$

This means $f'_-(x)$ and $f'_+(x)$ are both increasing function on (a, b).

Since $f'_+(x)$ is increasing on (a, b), $f'_+(x)$ is continuous except for a countable number of points in (a, b).

Let $x_0 \in (a,b)$ be a point of continuity of f'_+. Then $\lim_{x \to x_0^-} f'_+(x) = f'_+(x_0)$. But it follows from (3) that for $x < x_0$, $f_+'(x) \le f_-'(x_0) \le f_+'(x_0)$. Hence $f_+'(x_0) = \lim_{x \to x_0^-} f_+'(x) \le f_-'(x_0) \le f_+'(x_0)$. It follows that $f_-'(x_0) = f_+'(x_0)$ and so f is differentiable at x_0. We can now conclude that $f'_-(x)$ and $f'_+(x)$ are equal except possibly for a countable number of points x in (a, b). That is to say, f is differentiable except possibly for a countable number of points x in (a, b).

Property (4)

We next show that f is locally Lipschitz.

Take any closed interval $[c,d] \subseteq (a,b)$. For any $x < y$ in $[c, d]$,

$$f_+'(c) \le f_+'(x) \le p(x,y) = \frac{f(x) - f(y)}{x - y} \le f_-'(y) \le f_-'(d). \quad \text{------------- (*)}$$

Let $M = \max\left\{\left|f'_+(c)\right|, \left|f'_-(d)\right|\right\}$. Then it follows from (*) that for any $x \ne y$ in $[c, d]$, $\left|\frac{f(x) - f(y)}{x - y}\right| \le M$. Therefore, for all x, y in $[c, d]$,

$$\left|f(x) - f(y)\right| \le M\left|x - y\right|.$$

Hence f is Lipschitz on $[c, d]$ and so it is continuous on $[c, d]$ and it follows that f is continuous on (a, b).

Property (5)

We have, from (3), that for any $x < y$ in (a, b), $f_+'(x) \le p(x,y) = \frac{f(x) - f(y)}{x - y} \le f_-'(y)$. It follows then that for $x_0 \in (a,b)$, $f_+'(x_0) \le \frac{f(x) - f(x_0)}{x - x_0}$ for $x > x_0$ and so

$$f(x) \ge \left(x - x_0\right) f_+'(x_0) + f(x_0)$$

for $x \ge x_0$.

For $x < x_0$ in (a, b), $\frac{f(x_0) - f(x)}{x_0 - x} \le f_-'(x_0) \le f_+'(x_0)$ and $f(x) \ge f_+'(x_0)(x - x_0) + f(x_0)$ for $x \le x_0$. This means $f(x) \ge (x - x_0) f_+'(x_0) + f(x_0)$ for all x in (a, b).

We also have that $f(x) \ge f_-'(x_0)(x - x_0) + f(x_0)$ for $x \le x_0$. And from $f_-'(x_0) \le f_+'(x_0) \le \frac{f(x) - f(x_0)}{x - x_0}$ for $x > x_0$, we deduce that $f(x) \ge (x - x_0) f_-'(x_0) + f(x_0)$.

It follows that $f(x) \ge (x - x_0) f_-'(x_0) + f(x_0)$ for all x in (a, b).

Thus, for any $f_-'(x_0) \le \alpha \le f_+'(x_0)$, for $x \ge x_0$,

$$f(x) \ge (x - x_0) f_+'(x_0) + f(x_0) \ge (x - x_0)\alpha + f(x_0)$$

and for $x \le x_0$, $f(x) \ge (x - x_0) f_-'(x_0) + f(x_0) \ge (x - x_0)\alpha + f(x_0)$. Therefore, for all x in (a, b), $f(x) \ge (x - x_0)\alpha + f(x_0)$. Hence the line $y = (x - x_0)\alpha + f(x_0)$ is a support line for the graph of f at $(x_0, f(x_0))$ with the graph of f lying on or above the line.

For the next inequality, we shall recall some definitions and facts from measure theory.

A *measure space* $(X, \mathscr{M}, \mu)$ consists of a non-empty set X, a σ*- algebra* $\mathscr{M}$ of subsets of X and a *positive measure* $\mu : \mathscr{M} \to [0, \infty]$.

A σ*- algebra* on a set X is a collection $\mathscr{C}$ of subsets of X such that (i) $X \in \mathscr{C}$, (ii) if $A \in \mathscr{C}$, then its complement in X, $A^c \in \mathscr{C}$ and (iii) if $\{A_n : n = 1, 2, \cdots\} \subseteq \mathscr{C}$, then the countable union $\bigcup_{n=1}^{\infty} A_n \in \mathscr{C}$. It then follows from de-Morgan's law, that $\bigcap_{n=1}^{\infty} (A_n)^c \in \mathscr{C}$ and so since every subset can be written as a complement of its complement, $\bigcap_{n=1}^{\infty} A_n \in \mathscr{C}$. In particular if $A, B \in \mathscr{C}$, then $A - B = A \cap B^c \in \mathscr{C}$.

A subset X and a σ- algebra $\mathscr{C}$ of subsets of X is also called a *measure space* when the measure function is yet to be defined. If $(X, \mathscr{C})$ is a measure space and $(Y, \mathscr{T})$ a topological space, where $\mathscr{T}$ is a collection of subsets (called open sets) of Y, which is closed under finite intersection and arbitrary union, a function $f : X \to Y$ is said to be *measurable*, if for any open set U in $\mathscr{T}$, $f^{-1}(U) \in \mathscr{C}$. If $g : X \to [0, \infty]$ is a function into the extended non-negative real numbers and $(X, \mathscr{C})$ is a measure space, then g is measurable if

$g^{-1}((c,\infty])$ is measurable or $g^{-1}((c,\infty]) \in \mathscr{C}$ for any real number c. Likewise, $g : X \to [-\infty,\infty] = \overline{\mathbb{R}}$ is measurable if $g^{-1}((c,\infty]) \in \mathscr{C}$ for any real number c.

Given any non-empty collection of subsets of S of X, there is a smallest σ- algebra on X containing S. We say this is the *σ- algebra generated by S*. If (X, $\mathscr{T}$) is a topological space, then the σ- algebra $\mathscr{B}$ generated by the open sets, that is by $\mathscr{T}$, is called a *Borel measure* and members of $\mathscr{B}$ are called *Borel subsets*. Therefore, any continuous function $f : X \to Y$, where we give X the Borel measure or any measure that contains the Borel subsets, is automatically measurable and we say f is Borel measurable.

A *measure* or a p*ositive measure* on a measure space (X, $\mathscr{M}$) is a function $\mu : \mathscr{M} \to [0,\infty]$ such that $\mu(\varnothing) = 0$ and μ is countably additive, i.e., if $\{E_n\}_{n=1}^{\infty}$ is a countable disjoint collection of measurable subsets in $\mathscr{M}$, then $\mu\left(\bigcup_{n=1}^{n} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$.

The triplet, (X, $\mathscr{M}$, μ) is called a *measure space* and we sometimes refer to sets in $\mathscr{M}$ as *μ-measurable* sets. If $\mu(X) = 1$, we say (X, $\mathscr{M}$, μ) is a *probability space*.

We shall assume some familiarity with integration over a general measure space.

Jensen's Inequality

Theorem 2. Let (X, $\mathscr{M}$, μ) be a probability space (i.e., a measure space with positive measure μ and $\mu(X) = 1$). Suppose $f : X \to (a,b)$, with $-\infty \le a < b \le \infty$, is in $L^1(\mu)$, i.e., f is measurable and $\int_X |f| \, d\mu < \infty$, and φ is a convex function on (a, b).

Then $\varphi\left(\int_X f \, d\mu\right) \le \int_X \varphi \circ f \, d\mu$.

Proof.

Note that $\varphi \circ f$ is measurable since φ is continuous on (a, b) and f is measurable.

Since $a < f(x) < b$ for all x in X and $\mu(X) > 0$, $\int_X a \, d\mu < \int_X f \, d\mu < \int_X b \, d\mu$ and so

$a\mu(X) < \int_X f \, d\mu < b\mu(X)$ and we have that $a < \int_X f \, d\mu < b$. This is true if $a = -\infty$ or $b = \infty$ since $\left|\int_X f \, d\mu\right| \le \int_X |f| \, d\mu < \infty$.

Let $\beta = \int_X f \, d\mu$. Take a support line $y = \varphi(\beta) + (x - \beta)\gamma$ at $(\beta, \varphi(\beta))$ with $\varphi'_-(\beta) \le \gamma \le \varphi'_+(\beta)$. Hence $\varphi(x) \ge \varphi(\beta) + (x - \beta)\gamma$, for all x in (a, b). It follows that

$$\varphi(f(x)) \geq \varphi(\beta) + (f(x) - \beta)\gamma \text{ for all } x \text{ in } X.$$

Therefore,

$$\int_X \varphi \circ f \, d\mu \geq \varphi(\beta)\int_X d\mu + \left(\int_X f \, d\mu - \beta \int_X d\mu\right)\gamma = \varphi(\beta) = \varphi\left(\int_X f \, d\mu\right).$$

Corollary 3. Suppose f is convex on (a, b). Suppose $g : [c,d] \to (a,b)$ is Lebesgue integrable. Then

$$f\left(\frac{\int_c^d g}{d-c}\right) \leq \frac{1}{d-c}\int_c^d f(g(x))dx. \quad \text{------------------- } (*)$$

Proof. The measure on [c, d] is given by the measure subspace of the Lebesgue measure on $\mathbb{R}$ and the measure is given by η, the restriction of the Lebesgue measure to the interval [c, d]. For any Lebesgue measurable subset U in [c, d], let $\mu(U) = \frac{1}{\eta([c,d])}\eta(U) = \frac{1}{d-c}\eta(U)$.

Then μ is a measure on the σ algebra of Lebesuge measurable subsets of [c, d]. (Note that for the Lebesgue measure space ($\mathbb{R}$, $\mathcal{M}$, ℓ), where $\mathcal{M}$ is the σ algebra of Lebesgue measurable subsets of $\mathbb{R}$, $\ell : \mathcal{M} \to [0, \infty]$ is the Lebesgue measure on $\mathbb{R}$, ([c, d], $\mathcal{M}_{[c,d]}$, η), where $\mathcal{M}_{[c,d]} = \{ E \in \mathcal{M} : E \subseteq [c, d]\} = \{ E \cap [c, d]: E \in \mathcal{M} \}$ and $\eta : \mathcal{M}_{[c,d]} \to [0, \infty]$ is the restriction of ℓ to $\mathcal{M}_{[c,d]}$, is a measure space.)
Then $\mu([c,d]) = 1$ and ([c, d], $\mathcal{M}_{[c,d]}$, μ) is a probability space. Since g is Lebesgue integrable, i.e., g is η integrable and so is μ integrable. By Theorem 2,

$$f\left(\int_{[c,d]} g d\mu\right) \leq \int_{[c,d]} f(g(x))d\mu.$$

But $f\left(\int_{[c,d]} g d\mu\right) = f\left(\frac{1}{d-c}\int_{[c,d]} g d\eta\right) = f\left(\frac{1}{d-c}\int_c^d g\right)$

and $\int_{[c,d]} f(g(x))d\mu = \frac{1}{d-c}\int_c^d f(g(x))dx$ and so (*) follows.

Similarly we have

Corollary 4.
Let (X, $\mathcal{M}$, μ) be a measure space with μ a positive measure and $0 < \mu(X) < \infty$. Suppose $f : X \to (a,b)$, with $-\infty \leq a < b \leq \infty$, is in $L^1(\mu)$ and φ is a convex function on (a, b).

Then $\varphi\left(\frac{1}{\mu(X)}\int_X f \, d\mu\right) \leq \frac{1}{\mu(X)}\int_X \varphi \circ f \, d\mu$.

We can extend the idea of convex function to real vector space.

The Euclidean norm $\|\ \|$ for $\mathbb{R}^n$ is convex since for any x , y in $\mathbb{R}^n$ and $0 \leq \lambda \leq 1$,

$$\|(1-\lambda)x + \lambda y\| \leq (1-\lambda)\|x\| + \lambda\|y\|.$$

In particular, the modulus function on $\mathbb{R}$ is convex on $\mathbb{R}$.

Proposition 5. If $g:(a,b)\to\mathbb{R}$ is twice differentiable on (a, b) and $g''(x)\geq 0$ for all x in (a, b), or if g' is increasing on (a, b), then g is convex on (a, b).

Proof. $g''(x)\geq 0$ for all x in (a, b) implies that $g':(a,b)\to\mathbb{R}$ is increasing on (a, b).

For $a<x<z<y<b$, by the Mean Value Theorem, $p(x,z)=\dfrac{g(x)-g(z)}{x-z}=g'(c)$ for some c such that $x < c < z$ and also by the Mean Value Theorem, $p(y,z)=\dfrac{g(y)-g(z)}{y-z}=g'(d)$, for some d such that $z<d<y$. Thus, since $c < d$,

$p(x,z)=\dfrac{g(x)-g(z)}{x-z}=g'(c)\leq g'(d)=\dfrac{g(y)-g(z)}{y-z}=p(y,z)$. Therefore, by condition (2) of Definition 1, g is convex on (a, b).

Remark. The graph of a differentiable function on an open interval (a, b) is said to be *concave upward* if its derivative is increasing on (a, b). Proposition 5 implies that such a function is convex on (a, b).

Definition 6. $f:(a,b)\to\mathbb{R}$ is said to be *strictly convex* on (a, b) if

(1) $\forall x,y\in(a,b)$ and $x\neq y$, for $0<\lambda<1$, $f((1-\lambda)x+\lambda y)<(1-\lambda)f(x)+\lambda f(y)$, or equivalently,

(2) for $a<x<z<y<b$, the chord function

$$p(x,z)=\frac{f(x)-f(z)}{x-z}<p(z,y)=p(y,z)=\frac{f(y)-f(z)}{y-z}\ .$$

Proposition 7. If $g:(a,b)\to\mathbb{R}$ is twice differentiable on (a, b) and $g''(x)>0$ for all x in (a, b), or if g' is strictly increasing on (a, b), then g is strictly convex on (a, b).

Proof. The proof is similar to the proof of Proposition 5 with inequality replaced by strict inequality.

The Arithmetic-Geometric Inequality

Let $\varphi:\mathbb{R}\to\mathbb{R}$ be defined by $\varphi(x)=e^x$. Then $\varphi''(x)=e^x>0$ and so φ is (strictly) convex on $\mathbb{R}$.

Suppose (X, $\mathcal{M}$, μ) is a probability space (i.e., a measure space with $\mu(X) = 1$) and f is in $L^1(\mu)$. Then by Jensen's inequality (Theorem 2), $\exp\left(\int_X f\, d\mu\right) \le \int_X \exp\circ f\, d\mu$.

Take $X = \{1, 2, \cdots, n\}$, $n > 1$, $\mathcal{M}$ = power set of X and $\mu(\{i\}) = \frac{1}{n}$ for $1 \le i \le n$ and f a real valued function on X such that $f(i) = x_i$ for $1 \le i \le n$. Then (X, $\mathcal{M}$, μ) is a probability space and any function f on X is integrable with respect to μ. $\int_X f\, d\mu = \frac{1}{n}\left(x_1 + x_2 + \cdots + x_n\right)$ and $\int_X \exp\circ f\, d\mu = \frac{1}{n}\left(e^{x_1} + e^{x_2} + \cdots + e^{x_n}\right)$. Therefore,

$$e^{\frac{1}{n}(x_1+x_2+\cdots+x_n)} \le \frac{1}{n}\left(e^{x_1} + e^{x_2} + \cdots + e^{x_n}\right).$$

Letting $y_i = e^{x_i}$, we get the Arithmetic-Geometric Inequality,

$$\left(y_1 \cdot y_2 \cdots y_n\right)^{\frac{1}{n}} \le \frac{1}{n}\left(y_1 + y_2 + \cdots + y_n\right).$$

2. Inequalities in function spaces.

In preparation of presenting the L^p space as a normed linear space, where the triangle inequality is required for a norm, we discuss here the pertinent inequalities that may be used or referred to.

Definition 8. If $1 < p < \infty$, $1 < q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, then p and q are called *conjugate indices.*

Generalized Arithmetic-Geometric Inequality

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be defined by $\varphi(x) = e^x$. Then φ is convex on $\mathbb{R}$. Let $X = \{1,2\}$, $\mu(\{1\}) = \frac{1}{p}$, $\mu(\{2\}) = \frac{1}{q}$, $\mathcal{M}$ = power set of X. Then (X, $\mathcal{M}$, μ) is a probability space and any real-valued function f on X is integrable with respect to μ. Therefore, by Theorem 2,

$$\varphi\left(\int_X f\, d\mu\right) \le \int_X \varphi\circ f\, d\mu.$$

Take $f : X \to \mathbb{R}$, let $x_1 = f(1)$ and $x_2 = f(2)$. Then $\int_X f\, d\mu = \frac{1}{p}x_1 + \frac{1}{q}x_2$ and $\int_X \exp\circ f\, d\mu = \frac{1}{p}e^{x_1} + \frac{1}{q}e^{x_2}$. Therefore,

$$e^{\frac{1}{p}x_1+\frac{1}{q}x_2} \le \frac{1}{p}e^{x_1} + \frac{1}{q}e^{x_2} .$$

Thus letting $y_i = e^{x_i}$, i =1, 2, we have the generalized arithmetic-geometric inequality

$$y_1^{\frac{1}{p}} \cdot y_2^{\frac{1}{q}} \le \frac{1}{p}y_1 + \frac{1}{q}y_2 .$$

We denote the extended real numbers $[-\infty,\infty]$ by $\overline{\mathbb{R}}$.

Definition 9. Let $p \ge 1$ and $L^p(\mu) = \{ f ;\ f : X \to \mathbb{C}$ is measurable, $(X, \mathscr{M}, \mu)$ any measure space and $\int_X |f|^p d\mu < \infty \}$. For $f \in L^p(\mu)$, define $\|f\|_p = \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}}$.

Theorem 10. Let p and q be conjugate indices and $1 < p < \infty$. Suppose $(X, \mathscr{M}, \mu)$ a measure space. Let f, g :$X \to \overline{\mathbb{R}}$ be measurable. Then we have:

(a) Hölders Inequality. $\|f \cdot g\|_1 \le \|f\|_p \|g\|_q$.

(b) Minkowski Inequality. $\|f + g\|_p \le \|f\|_p + \|g\|_p$.

(c) For $1 \le k < \infty$, $L^k(\mu)$ is a vector space and $\|\cdot\|_k$ is a norm on equivalence classes of almost everywhere equal measurable functions in $L^k(\mu)$.

(d) In particular, $L^2(\mu)$ is an inner product space with $\langle f, g \rangle = \int_X f \cdot \overline{g}\, d\mu$.

Proof.

Part (a) If $\|f\|_p = 0$ or $\|g\|_q = 0$, then $|f|^p = 0$ or $|g|^q = 0$ almost everywhere with respect to μ. Therefore, $|f \cdot g| = 0$ almost everywhere with respect to μ and so (a) is trivially true if we follow the convention on multiplication in $[0, \infty]$: $0 \cdot \infty = \infty \cdot 0 = 0$, $x + \infty = \infty + x = \infty$ for $x \ge 0$. If $\|f\|_p > 0$ and $\|g\|_p = \infty$ or $\|g\|_p > \infty$ and $\|f\|_p = \infty$, then we have nothing to prove.

Assume now that $0 < \|f\|_p < \infty$ and $0 < \|g\|_q < \infty$. Let $F = \dfrac{|f|}{\|f\|_p}$ and $G = \dfrac{|g|}{\|g\|_q}$.

Apply the Generalised Arithmetic-Geometric Inequality to F^p and G^q to get

$$F \cdot G \le \frac{1}{p}F^p + \frac{1}{q}G^q .$$

Integrating we get:

$$\int_X (F\cdot G)\,d\mu \le \frac{1}{p}\int_X F^p d\mu + \frac{1}{q}\int_X G^q d\mu = \frac{1}{p}+\frac{1}{q}=1, \quad \text{--------------------- (4)}$$

since $\int_X F^p d\mu = \dfrac{\int_X |f|^p d\mu}{(\|f\|_p)^p} = \dfrac{\int_X |f|^p d\mu}{\int_X |f|^p d\mu} = 1$ and $\int_X G^q d\mu = \dfrac{\int_X |g|^q d\mu}{(\|g\|_q)^q} = \dfrac{\int_X |g|^q d\mu}{\int_X |g|^q d\mu} = 1$.

But $\int_X (F\cdot G)\,d\mu = \dfrac{\int_X |f\cdot g|\,d\mu}{\|f\|_p \|g\|_q}$ and it follows from (4) that

$$\|f\cdot g\|_1 = \int_X |f\cdot g|\,d\mu \le \|f\|_p \|g\|_q .$$

This proves part (a).

Part (b) $\|f+g\|_p \le \|f\|_p + \|g\|_p$

If $\|f+g\|_p = 0$, then we have nothing to prove. If $\|f\|_p = \infty$ or $\|g\|_p = \infty$, then we too have nothing to prove.

We now assume $\|f+g\|_p \ne 0$, $\|f\|_p < \infty$ and $\|g\|_p < \infty$.

If $\|f\|_p < \infty$ and $\|g\|_p < \infty$, then $\int_X |f|^p d\mu < \infty$ and $\int_X |g|^p d\mu < \infty$. Since $\left(\dfrac{|f|+|g|}{2}\right)^p \le \dfrac{1}{2}\left(|f|^p + |g|^p\right)$, on account of the convexity of the function t^p on (0, ∞), $\int_X (|f|+|g|)^p d\mu \le 2^{p-1}\left(\int_X |f|^p\,d\mu + \int_X |g|^p\,d\mu\right) < \infty$. Therefore, as $|f+g| \le |f|+|g|$, $|f+g|^p \le (|f|+|g|)^p$, it follows that $\int_X |f+g|^p\,d\mu \le \int_X (|f|+|g|)^p\,d\mu < \infty$. This means that if $\|f\|_p < \infty$ and $\|g\|_p < \infty$, then $\int_X |f+g|^p d\mu < \infty$ and $\|f+g\|_p < \infty$.

So now we assume that $\|f\|_p < \infty$, $\|g\|_p < \infty$ and $0 < \|f+g\|_p < \infty$.

Observe that $|f+g|^p = |f+g|\,|f+g|^{p-1} \le \big||f|+|g|\big|\,|f+g|^{p-1} = |f|\,|f+g|^{p-1} + |g|\,|f+g|^{p-1}$.

Therefore,

$$(\|f+g\|_p)^p = \int_X |f+g|^p\,d\mu \le \int_X |f|\,|f+g|^{p-1}\,d\mu + \int_X |g|\,|f+g|^{p-1}\,d\mu$$

$$\le \|f\|_p \left(\int_X |f+g|^{(p-1)q}\,d\mu\right)^{\frac{1}{q}} + \|g\|_p \left(\int_X |f+g|^{(p-1)q}\,d\mu\right)^{\frac{1}{q}} \quad \text{by Hölders Inequality (part(a)).}$$

But $\frac{1}{p}+\frac{1}{q}=1$ so that $\frac{1}{q}=1-\frac{1}{p}=\frac{p-1}{p}$ and $(p-1)q=p$. Hence

$$\|f\|_p\left(\int_X |f+g|^{(p-1)q}\,d\mu\right)^{\frac{1}{q}}+\|g\|_p\left(\int_X |f+g|^{(p-1)q}\,d\mu\right)^{\frac{1}{q}}$$

$$=\left(\|f\|_p+\|g\|_p\right)\left(\int_X |f+g|^{p}\,d\mu\right)^{\frac{p-1}{p}}.$$

Therefore,

$$\left(\|f+g\|_p\right)^p \le \left(\|f\|_p+\|g\|_p\right)\left(\int_X |f+g|^{p}\,d\mu\right)^{\frac{p-1}{p}}=\left(\|f\|_p+\|g\|_p\right)\left(\|f+g\|_p\right)^{p-1}.$$

Dividing by $\left(\|f+g\|_p\right)^{p-1}$ on both sides, we have $\|f+g\|_p \le \|f\|_p+\|g\|_p$. This proves part (b)

Part (c). $L^k(\mu)$ is a vector space for $k \ge 1$.

$f,\ g \in L^k(\mu) \Rightarrow |f|, |g|$ and $|f+g|$ are measurable and $\|f\|_k, \|g\|_k < \infty$.

Suppose $1 < k < \infty$. Then by Minkowski inequality (part b)
$\big\||f|+|g|\big\|_k \le \big\||f|\big\|_k+\big\||g|\big\|_k=\|f\|_k+\|g\|_k<\infty$. Note that $|f+g|\le|f|+|g|$ and it follows that $\int_X |f+g|^k\,d\mu \le \int_X \left(|f|+|g|\right)^k d\mu$. Hence

$$\|f+g\|_k=\left(\int_X |f+g|^k\,d\mu\right)^{\frac{1}{k}} \le \left(\int_X \left(|f|+|g|\right)^k d\mu\right)^{\frac{1}{k}}=\big\||f|+|g|\big\|_k \le \|f\|_k+\|g\|_k<\infty.$$

Therefore, $f+g \in L^k(\mu)$.

If $k = 1$, for $f, g \in L^1(\mu)$, by the triangle inequality, $|f+g|\le|f|+|g|$,

$$\|f+g\|_1=\int_X |f+g|\,d\mu \le \int_X |f|\,d\mu+\int_X |g|\,d\mu=\|f\|_1+\|g\|_1<\infty.$$

Hence, $f+g \in L^1(\mu)$.

If $\lambda \in \mathbb{C}$, then plainly $\|\lambda f\|_k=|\lambda|\|f\|_k<\infty$ and so $\lambda f \in L^k(\mu)$. Thus $L^k(\mu)$ is a vector space for $k \ge 1$.

We note that (1) $\|f\|_k \ge 0$, (2) $\|f\|_k=0 \Leftrightarrow \int_X |f|^k\,d\mu=0 \Leftrightarrow f=0$ almost everywhere on X. (3) $\|\lambda f\|_k=|\lambda|\|f\|_k$ and $\|f+g\|_k \le \big\||f|+|g|\big\|_k \le \|f\|_k+\|g\|_k$. Hence $\|\cdot\|_k$ is a norm on equivalence classes of almost everywhere equal measurable functions in $L^k(\mu)$.

Part (d). In $L^2(\mu)$, $\|f\|_2 = \left(\int_X |f|^2 \, d\mu\right)^{\frac{1}{2}} = \left(\int_X f \; \bar{f} \; d\mu\right)^{\frac{1}{2}}$. If we define the inner product on $L^2(\mu)$ by $\langle f, g\rangle = \int_X f \; \bar{g} \; d\mu$, then $\|f\|_2 = \langle f, f\rangle^{\frac{1}{2}}$. Note that $\left|\langle f, g\rangle\right| = \left|\int_X f \; \bar{g} \; d\mu\right| \leq \int_X |f \; \bar{g} \;| d\mu = \|f \cdot g\|_1 = \||f| \cdot |g|\|_1 \leq \|f\|_2 \|g\|_2 < \infty$ by Hölders Inequality. Observe that $\langle g, f\rangle = \int_X g \; \bar{f} \; d\mu = \int_X \overline{f \; \bar{g}} \, d\mu = \overline{\int_X f \; \bar{g} d\mu} = \overline{\langle f, g\rangle}$. It is then easily seen that $L^2(\mu)$ is an inner product space.

Remark.

When thought of as a normed linear space, the function space $L^p(\mu)$ consists of equivalence classes of almost everywhere equal measurable functions. In practice in most of our argument we merely proceed by taking a representative of an equivalence class without explicitly mentioning the class it represents.

3. L^p Spaces, L^∞ Space

The first result that we present here is that the equivalence classes of almost everywhere equal functions of $L^p(\mu)$, with the metric induced by its norm is a Banach space.

For a normed linear space *V* with the norm $\|\cdot\|$, the metric *d induced by the norm* on *V* is defined by $d(x, y) = \|x - y\|$. Then it is easily seen that for all *x*, *y* and *z* in *V*,

(1) $d(x, y) \geq 0$,

(2) $d(x, y) = 0 \Leftrightarrow x = y$,

(3) $d(x, y) = d(y, x)$,

(4) Triangle inequality, $d(x, y) \leq d(x, z) + d(z, y)$.

Recall that a set *M* with a metric function $d: M \times M \to \mathbb{R}$ is called a *metric space* if the function *d* satisfies the properties (1) to (4) above. An open ball of radius *r* > 0 and centred at *x* in *M* is the set $B_r(x) = \{y \in M : d(y, x) < r\}$. A sequence (x_n) in *M* is *convergent* in *M*, if there is a point *x* in *M* such that for any $\varepsilon > 0$, there is an integer *N* such that

$$n \geq N \Rightarrow x_n \in B_\varepsilon(x), \text{ i.e., } n \geq N \Rightarrow d(x_n, x) < \varepsilon .$$

A sequence (x_n) in *M* is said to be *Cauchy*, if for any $\varepsilon > 0$, there is an integer *N* such that $n, m \geq N \Rightarrow d(x_n, x_m) < \varepsilon$. Plainly any convergent sequence is a Cauchy sequence. But not all Cauchy sequence is convergent. A metric space (M, d) is *complete* if every Cauchy sequence in (M, d) converges to a point in M.

For a normed linear space V, we can consider V as a metric space with the metric induced by the norm on V. If the normed linear space with the metric induced by its norm is complete as a metric space, then we say the normed linear space is a *Banach space*. It is to be noted that this definition is dependent on the norm specified on V. As a topological space, V has the topology given by the open balls of V. In general, for a metric space (M, d), the induced metric topology T_d on (M, d) is generated by the family of open sets, i.e., open balls and arbitrary union of open balls in (M, d). Hence different norms on the linear space may give rise to different metric spaces with different topologies induced by the respective induced metric.

If the norm of a Banach space V arises from an inner product, then it is called a *Hilbert space*.

More precisely, an *inner product* on a (real or complex) linear space V is a scalar valued function on $V \times V$, whose value on (x, y) in $V \times V$ is denoted by $\langle x, y \rangle$ and the function satisfies the following properties:

(1) $\langle x, x \rangle \geq 0; \langle x, x \rangle = 0 \Leftrightarrow x = 0$;

(2) $\langle x, y \rangle = \overline{\langle y, x \rangle}$, the complex conjugate of $\langle y, x \rangle$;

(3) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$;

(4) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$.

An inner product gives rise to an associated norm $\|\cdot\|$ on V defined by $\|x\| = \sqrt{\langle x, x \rangle}$. That this is a norm is a consequence of the (Schwarz Inequality) for inner product:
$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle} = \|x\| \|y\|$. For instance,
$\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + \|y\|^2 + 2\text{Real part} \langle x, y \rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| = (\|x\| + \|y\|)^2$
and the triangle inequality follows.

Hence, we can consider a linear space with an inner product as a normed linear space with the associated norm defined above. If the inner product space, considered as a normed linear space with the associated norm, is a Banach space, then it is called a *Hilbert space.*

Now the equivalence classes of almost everywhere equal measurable functions in $L^p(\mu)$ under the norm $\| f \|_p = \left(\int_X | f |^p d\mu \right)^{\frac{1}{p}}$ for f a representative of [f], the equivalence class which f belongs, is a normed linear space and the same collection of equivalence classes with the induced metric is a metric space. By abuse of notation we simply say $L^p(\mu)$ is a

Banach space if the equivalence classes of almost everywhere equal measurable functions in $L^p(\mu)$ is a complete metric space under the metric induced by the norm.

Theorem 11. For $1 \le p < \infty$, $L^p(\mu)$ is a Banach space. In particular, $L^2(\mu)$ is a Hilbert space.

Proof.

We shall show that any Cauchy sequence (f_n) in $L^p(\mu)$ has a subsequence that converges pointwise almost everywhere to a function f in $L^p(\mu)$.

(This subsequence will be an "effective" Cauchy sequence. We shall find integers $n_1 < n_2 < \cdots$ such that $\left\| f_{n_{i+1}} - f_{n_i} \right\|_p < \frac{1}{2^i}$.)

Since (f_n) is a Cauchy sequence, there exists an integer n_1 such that for all $n \ge n_1$,

$\left\| f_n - f_{n_1} \right\|_p < \frac{1}{2^1}$. Then there exists an integer $n_2 > n_1$, such that for all $n \ge n_2$,

$\left\| f_n - f_{n_2} \right\|_p < \frac{1}{2^2}$, in particular, $\left\| f_{n_2} - f_{n_1} \right\|_p < \frac{1}{2^1}$. Next there exists an integer $n_3 > n_2$, such that for all $n \ge n_3$, $\left\| f_n - f_{n_3} \right\|_p < \frac{1}{2^3}$ and $\left\| f_{n_3} - f_{n_2} \right\|_p < \frac{1}{2^2}$. Inductively we find integers $n_1 < n_2 < \cdots$ such that for all $n \ge n_i$, $\left\| f_n - f_{n_i} \right\|_p < \frac{1}{2^i}$ and $\left\| f_{n_{i+1}} - f_{n_i} \right\|_p < \frac{1}{2^i}$.

Define for each integer $k \ge 1$, $g_k = \sum_{i=1}^{k} \left| f_{n_{i+1}} - f_{n_i} \right|$. Then $g_k = \sum_{i=1}^{k} \left| f_{n_{i+1}} - f_{n_i} \right|$ is measurable and $\left\| g_k \right\|_p = \left\| \sum_{i=1}^{k} \left| f_{n_{i+1}} - f_{n_i} \right| \right\|_p \le \sum_{i=1}^{k} \left\| f_{n_{i+1}} - f_{n_i} \right\|_p < \sum_{i=1}^{k} \frac{1}{2^i} < 1$ and so $g_k \in L^p(\mu)$.

This means $\left(\int_X |g_k|^p \, d\mu \right)^{\frac{1}{p}} = \left(\int_X g_k^p d\mu \right)^{\frac{1}{p}} < 1$, and so $\int_X g_k^p d\mu < 1$. It follows that $g_k^p \in L^1(\mu)$. Note that (g_k^p) is an increasing sequence of non-negative functions.

We also have that $g_k \to \sum_{i=1}^{\infty} \left| f_{n_{i+1}} - f_{n_i} \right|$ and so if we let $g = \sum_{i=1}^{\infty} \left| f_{n_{i+1}} - f_{n_i} \right|$ then $g_k^p \to g^p$ and g^p is measurable. (Note that we only assert that the limit exists and may be infinity.)

Therefore, by Fatou's Lemma,

$$\int_X g^p d\mu = \int_X \liminf_{k \to \infty} g_k^p d\mu \le \liminf_{k \to \infty} \int_X g_k^p d\mu \le 1.$$

It follows that g^p is finite almost everywhere with respect to μ and so g is finite almost everywhere with respect to μ. Hence the series $\sum_{i=1}^{\infty}\left(f_{n_{i+1}} - f_{n_i} \right)$ converges pointwise absolutely almost everywhere with respect to μ. That is to say there is a set M of μ measure zero, such that $\sum_{i=1}^{\infty}\left(f_{n_{i+1}}(x) - f_{n_i}(x) \right)$ converges pointwise absolutely, for all x in X but outside M.

Define $f(x) = \begin{cases} f_{n_1}(x) + \sum_{i=1}^{\infty}\left(f_{n_{i+1}}(x) - f_{n_i}(x) \right) & x \in X \setminus M \\ 0 \quad , \; x \in M \end{cases}$.

Then $f(x) = \begin{cases} \lim_{k\to\infty}\left(f_{n_1}(x) + \sum_{i=1}^{k-1}\left(f_{n_{i+1}}(x) - f_{n_i}(x) \right) \right) & x \in X \setminus M \\ 0 \quad , \; x \in M \end{cases}$

$$= \begin{cases} \lim_{k\to\infty} f_{n_k}(x) & x \in X \setminus M \\ 0 \quad , \; x \in M \end{cases}$$

Hence f_{n_k} converges pointwise to f almost everywhere with respect to μ.

We shall show next that $f \in L^p(\mu)$ and that $f_n \to f$ in $L^p(\mu)$.

Since (f_n) is a Cauchy sequence in $L^p(\mu)$, given $\varepsilon > 0$, there exists an integer N such that for all $n, m \geq N$,

$$\| f_n - f_m \|_p < \varepsilon .$$

Fixed an $n \geq N$. Then $\left| f_{n_i} - f_n \right|^p$ converges pointwise almost everywhere with respect to μ, as i tends to ∞, to $\left| f - f_n \right|^p$. Therefore, by Fatou's Lemma,

$$\int_X | f - f_n |^p \, d\mu = \int_X \liminf_{k\to\infty} | f_{n_k} - f_n |^p \, d\mu \leq \liminf_{k\to\infty} \int_X | f_{n_k} - f_n |^p \, d\mu \leq \varepsilon^p .$$

The last inequality holds since for $n_k \geq N$, $\| f_{n_k} - f_n \|_p = \left(\int_X \left| f_{n_k} - f_n \right|^p \right)^{\frac{1}{p}} < \varepsilon$.

This implies that $\| f - f_n \|_p = \left(\int_X | f - f_n |^p \, d\mu \right)^{\frac{1}{p}} \leq \varepsilon < \infty$ and we conclude that for all $n \geq N$, $\| f - f_n \|_p \leq \varepsilon$ and $f - f_n \in L^p(\mu)$. Since $f_n \in L^p(\mu)$, $f = (f - f_n) + f_n \in L^p(\mu)$. This means $f_n \to f$ in $L^p(\mu)$. Thus, any Cauchy sequence (f_n) in $L^p(\mu)$ is convergent in $L^p(\mu)$ and so $L^p(\mu)$ is a Banach space.

Now $L^2(\mu)$ is also a Banach space and the norm is given by an associated inner product (see Theorem 10 part (d)) and so it is a Hilbert space.

Definition 12. Let $(X, \mathscr{M}, \mu)$ be a measure space. Suppose $f: X \to \overline{\mathbb{R}}$ is a measurable function. Then the *essential supremum* of f is defined by

$$\operatorname{ess\,sup} f = \inf_{N \subseteq X, \mu(N)=0} \sup\{f(x) : x \notin N\} .$$

Equivalently, let $S = \left\{\alpha \in \mathbb{R} : \mu\left(f^{-1}((\alpha,\infty])\right) = 0 \text{ and } \alpha \geq \inf f(X)\right\}$,

$$\operatorname{ess\,sup} f = \begin{cases} \inf S \ , & \text{if } S \neq \varnothing \\ \infty \ , & \text{if } S = \varnothing \end{cases} .$$

Note that if $S \neq \varnothing$, then either $\inf f(X) < \inf S < \infty$ or $\inf S = \inf f(X)$.

It is not clear that the two definitions above are equivalent. We elaborate the proof of this fact below.

Lemma 13. The two definitions in Definition 12 are equivalent.

Proof.

We show the equivalence of the two definitions for the case one of the definitions gives ∞.

Suppose $S = \left\{\alpha \in \mathbb{R} : \mu\left(f^{-1}(\alpha,\infty]\right) = 0 \text{ and } \alpha \geq \inf f(X)\right\} = \varnothing$.

Now either $\inf f(X) = \infty$ or $\inf f(X) < \infty$,

If $\inf f(X) = \infty$, then $f(X) = \infty$ and for any proper measurable subset N of X with $\mu(N) = 0$, $\sup\{f(x) : x \notin N\} = \infty$ and so $\inf_{N \subseteq X, \mu(N)=0} \sup\{f(x) : x \notin N\} = \infty$.

Suppose $\inf f(X) < \infty$ and so $S = \varnothing$ implies for any real number $\alpha > \inf f(X)$, $\mu\left(f^{-1}(\alpha,\infty]\right) \neq 0$. It follows that for any measurable $N \subseteq X$ and $\mu(N) = 0$, $f^{-1}(\alpha,\infty] \not\subset N$ and so $\sup\{f(x) : x \notin N\} > \alpha$. Hence $\sup\{f(x) : x \notin N\} = \infty$. Therefore, $\inf_{N \subseteq X, \mu(N)=0} \sup\{f(x) : x \notin N\} = \infty$. Note that if $-\infty < \inf f(X) < \infty$, as $\mu\left(f^{-1}(\alpha,\infty]\right) \neq 0$ for $\alpha > \inf f(X)$, $\mu\left(f^{-1}(\inf f(X),\infty]\right) \neq 0$.

Conversely, suppose $\inf_{N\subseteq X,\mu(N)=0} \sup\{f(x): x\notin N\}=\infty$, then $\sup\{f(x): x\notin N\}=\infty$ for any proper measurable subset $N \subseteq X$ and $\mu(N)=0$. It follows that for any real $\alpha \geq \inf f(X)$, $f^{-1}(\alpha,\infty] \not\subset N$ for any proper subset $N \subseteq X$ and $\mu(N)=0$ for if $f^{-1}(\alpha,\infty] \subseteq N$, then $\sup\{f(x): x\notin N\} \leq \alpha$ contradicting $\sup\{f(x): x\notin N\}=\infty$. Hence $\mu\left(f^{-1}(\alpha,\infty]\right)\neq 0$ for any real $\alpha \geq \inf f(X)$ and so

$$S=\left\{\alpha\in\mathbb{R}: \mu\left(f^{-1}(\alpha,\infty]\right)=0 \text{ and } \alpha \geq \inf f(X)\right\}=\varnothing .$$

Case $S\neq\varnothing$**.**

Then either $\inf f(X) < \inf S < \infty$ or $\inf S = \inf f(X)$.

Case $S\neq\varnothing$ **and** $\inf S = \inf f(X)$

When $S\neq\varnothing$ and $\inf S = \inf f(X)$, we discuss the two possibilities according to (1) $\inf f(X)\in f(X)$ or (2) $\inf f(X)\notin f(X)$.

Note that if $S\neq\varnothing$ and $\inf S = \inf f(X)$, then $\mu\left(f^{-1}\left((\alpha,\infty]\right)\right)=0$ for any $\alpha > \inf f(X)$.

(1) $S\neq\varnothing$, $\inf S = \inf f(X)$ and $\inf f(X)\in f(X)$. Then either $\inf f(X)=-\infty$ or $-\infty < \inf f(X) < \infty$.

Case $S\neq\varnothing$, $\inf S=\inf f(X)$**,** $\inf f(X)\in f(X)$ **and** $\inf f(X)=-\infty$**.**

For any $\alpha > \inf f(X) = -\infty$, $f^{-1}\left((\alpha,\infty]\right)\neq X$, $\mu\left(f^{-1}\left((\alpha,\infty]\right)\right)=0$

and $\inf f(X) \leq \sup\left\{f(x): x\notin f^{-1}(\alpha,\infty]\right\} \leq \alpha$

and so $\inf_{N\subseteq X,\mu(N)=0} \sup\{f(x): x\notin N\} = \inf f(X) = \inf S$.

Case $S\neq\varnothing$, $\inf S=\inf f(X)$**,** $\inf f(X)\in f(X)$ **and** $-\infty < \inf f(X) < \infty$**.**

For any $\alpha > \inf f(X)$, $f^{-1}\left((\alpha,\infty]\right)\neq X$, $\mu\left(f^{-1}\left((\alpha,\infty]\right)\right)=0$ and so

$\inf f(X) \leq \sup\left\{f(x): x\notin f^{-1}(\alpha,\infty]\right\} \leq \alpha$. Taking $N=\mu\left(f^{-1}\left((\alpha,\infty]\right)\right)$, $\sup\{f(x): x\notin N\}\leq\alpha$. Hence $\inf_{N\subseteq X,\mu(N)=0} \sup\{f(x): x\notin N\} \leq \alpha$ for any $\alpha > \inf f(X)$. Therefore, $\inf_{N\subseteq X,\mu(N)=0} \sup\{f(x): x\notin N\} = \inf f(X) = \inf S$.

Moreover, for integer $n \geq 1$, $\mu\left(f^{-1}\left(\inf f(X)+\frac{1}{n},\infty\right]\right)=0$ so that

$$\mu\left(f^{-1}\left((\inf f(X),\infty]\right)\right)=\mu\left(\bigcup_{n=1}^{\infty} f^{-1}\left(\left(\inf f(X)+\frac{1}{n},\infty\right]\right)\right)$$

$$=\lim_{n\to\infty}\mu\left(f^{-1}\left(\left(\inf f(X)+\frac{1}{n},\infty\right]\right)\right)=0$$

and so $\inf S=\inf f \in S$.

(2) Case $S \neq \varnothing$, $\inf S=\inf f(X)$, $\inf f(X)\notin f(X)$

If $\inf f(X)>-\infty$, then $\mu\left(f^{-1}\left(\inf f(X)+\frac{1}{n},\infty\right]\right)=0$ for any integer $n \geq 1$.

Therefore, $\mu(X)=\mu\left(f^{-1}\left((\inf f(X),\infty]\right)\right)=\mu\left(\bigcup_{n=1}^{\infty} f^{-1}\left(\left(\inf f(X)+\frac{1}{n},\infty\right]\right)\right)$

$$=\lim_{n\to\infty}\mu\left(f^{-1}\left(\left(\inf f(X)+\frac{1}{n},\infty\right]\right)\right)=0$$

and so $\inf S=\inf f(X)\in S$.

Note that for any integer $n \geq 1$, $\mu\left(f^{-1}\left(\left(\inf f(X)+\frac{1}{n},\infty\right]\right)\right)=0$, so taking $N=f^{-1}\left(\left(\inf f(X)+\frac{1}{n},\infty\right]\right)$, $\sup\left\{f(x):x\notin N\right\}\leq \inf f(X)+\frac{1}{n}$. It follows that $\inf_{N\subseteq X,\mu(N)=0}\sup\left\{f(x):x\notin N\right\}\leq \inf f(X)$ and so $\inf_{N\subseteq X,\mu(N)=0}\sup\left\{f(x):x\notin N\right\}=\inf f(X)=\inf S$.

If $\inf f(X)=-\infty$, then

$$\mu(X)=\mu\left(f^{-1}\left((\inf f(X),\infty]\right)\right)=\mu\left(\bigcup_{n=1}^{\infty} f^{-1}\left((-n,\infty]\right)\right)$$

$$=\lim_{n\to\infty}\mu\left(f^{-1}\left((-n,\infty]\right)\right)=0.$$

Observe that for each integer $n \geq 1$, $f^{-1}\left(\left(-n,\infty\right]\right) \neq X$, $\mu\left(f^{-1}\left(\left(-n,\infty\right]\right)\right) = 0$ and so taking $N = f^{-1}\left(\left(-n,\infty\right]\right)$ we have $\sup\left\{f(x) : x \notin N\right\} \leq -n$. It follows that $\inf_{N\subseteq X,\mu(N)=0} \sup\left\{f(x) : x \notin N\right\} = -\infty = \inf f(X) = \inf S$.

In both cases, we get $\inf_{N\subseteq X,\mu(N)=0} \sup\left\{f(x) : x \notin N\right\} = \inf S$.

In the case when $S \neq \varnothing$, $\inf S = \inf f(X)$ and $\inf f(X) \notin f(X)$, $\mu(X) = 0$. This can happen only if the measure space (X, $\mathscr{M}$, μ) is trivial.

Case $S \neq \varnothing$ and $\inf f(X) < \inf S < \infty$.

We claim that if $\inf f(X) < \inf S < \infty$, then $\gamma = \inf S \in S$. We show this below.

Let n be an integer ≥ 1. Then by definition of $\inf S$, there exists $\beta \in S$ such that $\gamma \leq \beta < \gamma + \dfrac{1}{n}$ and $\mu\left(f^{-1}\left((\beta,\infty]\right)\right) = 0$. Since $f^{-1}\left((\gamma + \frac{1}{n},\infty]\right) \subseteq f^{-1}\left((\beta,\infty]\right)$ and $f^{-1}\left((\gamma + \frac{1}{n},\infty]\right)$ is measurable, $\mu\left(f^{-1}\left((\gamma + \frac{1}{n},\infty]\right)\right) = 0$. Therefore, since

$f^{-1}\left((\gamma,\infty]\right) = \bigcup_{n=1}^{\infty} f^{-1}\left((\gamma + \frac{1}{n},\infty]\right)$ and $f^{-1}\left((\gamma + \frac{1}{n},\infty]\right) \subseteq f^{-1}\left((\gamma + \frac{1}{n+1},\infty]\right)$,

$\mu\left(f^{-1}\left((\gamma,\infty]\right)\right) = \lim_{n\to\infty} \mu\left(f^{-1}\left((\gamma + \frac{1}{n},\infty]\right)\right) = 0$. Thus $\gamma = \inf S \in S$.

Take $N = f^{-1}(\gamma,\infty]$. Then $\mu(N) = 0$ and $\sup\left\{f(x) : x \notin N\right\} \leq \gamma$.

It follows that $\inf_{N\subseteq X,\mu(N)=0} \sup\left\{f(x) : x \notin N\right\} \leq \gamma$.

Now given any $\varepsilon > 0$, $\gamma - \varepsilon < \gamma = \inf S$. This means $\mu\left(f^{-1}(\gamma - \varepsilon,\infty]\right) > 0$. Hence, for any $N \subseteq X$ and $\mu(N) = 0$, $f^{-1}(\gamma - \varepsilon,\infty] \not\subset N$ and so $\sup\left\{f(x) : x \notin N\right\} \geq \gamma - \varepsilon$. Therefore, $\inf_{N\subseteq X,\mu(N)=0} \sup\left\{f(x) : x \notin N\right\} \geq \gamma - \varepsilon$. Since ε > 0 is arbitrarily chosen, $\inf_{N\subseteq X,\mu(N)=0} \sup\left\{f(x) : x \notin N\right\} \geq \gamma$. We can now conclude that $\inf_{N\subseteq X,\mu(N)=0} \sup\left\{f(x) : x \notin N\right\} = \inf S = \gamma$.

This completes the proof of Lemma 13.

We say f is *essentially bounded* if $\operatorname{ess\,sup}|f| < \infty$.

Note that if f is a non-negative function, then

$$\operatorname{ess\,sup} f = \inf\left\{\alpha \in \mathbb{R} : \mu\left(f^{-1}((\alpha,\infty])\right) = 0 \text{ and } \alpha \geq \inf f(X) \geq 0\right\} \geq 0.$$

Let $L^\infty(\mu) = \left\{f : X \to \mathbb{C};\, f \text{ measurable and } \operatorname{ess\,sup}|f| < \infty\right\}$. I.e., $L^\infty(\mu)$ is the space of essentially bounded measurable functions and we define $\|f\|_\infty = \operatorname{ess\,sup}|f|$ for f in $L^\infty(\mu)$.

We shall show that $L^\infty(\mu)$ is a linear space, i.e., a vector space over $\mathbb{C}$, and $\|\cdot\|_\infty$ is a norm on $L^\infty(\mu)$.

Note that for $f \in L^\infty(\mu)$, $0 \leq \|f\|_\infty = \inf\left\{\alpha \in \mathbb{R} : \mu\left(|f|^{-1}(\alpha,\infty]\right) = 0 \text{ and } \alpha \geq \inf|f| \geq 0\right\} < \infty$. Here we regard $|f|$ as a measurable function of X into $[0, \infty]$. Therefore, there exists a $\alpha_n \in \left\{\alpha \in \mathbb{R} : \mu\left(|f|^{-1}((\alpha,\infty])\right) = 0 \text{ and } \alpha \geq \inf|f| \geq 0\right\}$ such that $\alpha_n < \|f\|_\infty + \dfrac{1}{n}$. That is , $\mu\left(|f|^{-1}((\alpha_n,\infty])\right) = 0$. Let $E_n = |f|^{-1}((\alpha_n,\infty])$. Then $\mu(E_n) = 0$.

This means $|f(x)| \leq \|f\|_\infty + \dfrac{1}{n}$ for all x in $(E_n)^c$. If we let $A = \bigcup_{n=1}^{\infty} E_n$, then

$$|f(x)| \leq \|f\|_\infty \text{ for all } x \text{ in } A^c \text{ and } \mu(A) = 0.$$

We can now conclude that for any g in $L^\infty(\mu)$, there exists a set B of measure zero such that

$$|g(x)| \leq \|g\|_\infty \text{ for all } x \text{ in } B^c \text{ and } \mu(B) = 0.$$

Therefore,

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty$$

for all x in $(A \cup B)^c$ and $\mu(A \cup B) = 0$. Hence, $|f + g|^{-1}\left((\|f\|_\infty + \|g\|_\infty, \infty]\right) \subseteq A \cup B$ and so $\|f\|_\infty + \|g\|_\infty \in \left\{\alpha \in \mathbb{R} : \mu\left(|f + g|^{-1}((\alpha,\infty])\right) = 0, \alpha \geq \inf|f + g|\right\}$. It follows then from the definition of infimum that

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty < \infty. \quad \text{-------------------------} \quad (1)$$

Therefore, $f + g \in L^\infty(\mu)$.

Take any $f \in L^{\infty}(\mu)$ and $\lambda \neq 0$. There exists a set N of measure zero such that

$$|f(x)| \leq \|f\|_{\infty} \text{ for all } x \text{ in } N^c.$$

Hence, $|\lambda f(x)| \leq |\lambda| \|f\|_{\infty}$ for all x in N^c. Therefore, $\|\lambda f\|_{\infty} \leq |\lambda| \|f\|_{\infty} < \infty$. This means $\lambda f \in L^{\infty}(\mu)$.

Thus $L^{\infty}(\mu)$ is a linear space.

If $\|f\|_{\infty} = 0$, then there exists a set E of measure zero such that

$$|f(x)| \leq \|f\|_{\infty} = 0 \text{ for all } x \text{ in } E^c \text{ and } \mu(E) = 0$$

and so $f = 0$ almost everywhere on X. We have already shown that the triangle inequality (see (1)) holds. Hence $\|\cdot\|_{\infty}$ is a norm on equivalence classes of almost everywhere equal measurable functions in $L^{\infty}(\mu)$.

Note that if $(X, \mathscr{M}, \mu)$ is a measure space and $f : X \to \mathbb{R}$ is measurable, then $\operatorname{ess\,sup}|f| \leq \sup|f|$. Furthermore, if every non-empty measurable set in $\mathscr{M}$ has positive μ measure and f is measurable, then $\operatorname{ess\,sup}|f| = \sup|f|$.

Lemma 14. Suppose X is a topological space. If $f : X \to [0, \infty)$ is continuous and $(X, \mathscr{M}, \mu)$ is a measure space such that $\mathscr{M}$ contains all the Borel subsets of X and any non-empty open subset of X has non-zero measure *or* $X = \mathbb{R}^n$ and $(X, \mathscr{M}, \mu)$ is the Lebesgue measure on $\mathbb{R}^n$, then $\operatorname{ess\,sup} f = \sup f$.

Proof. Suppose $f : \mathbb{R}^n \to [0, \infty)$ is continuous. Note that $\inf f \geq 0$.

Suppose $\operatorname{ess\,sup} f = \infty$. Then $S = \left\{\alpha \in \mathbb{R} : \mu\left(f^{-1}((\alpha, \infty])\right) = 0 \text{ and } \alpha \geq \inf f(X)\right\} = \varnothing$.

This means for any real $\alpha \geq \inf f$, $\mu\left(f^{-1}((\alpha, \infty])\right) \neq 0$ and so $f^{-1}((\alpha, \infty]) \neq \varnothing$. Hence f is unbounded and so $\sup f = \infty$. Conversely, suppose $\sup f = \infty$. Then the function f is not bounded above. Therefore, for any real $\alpha \geq \inf f$, $f^{-1}((\alpha, \infty]) \neq \varnothing$ and since f is continuous $f^{-1}((\alpha, \infty])$ is open and non-empty and so $\mu\left(f^{-1}((\alpha, \infty])\right) \neq 0$. This means for any real $\alpha \geq \inf f$, $\alpha \notin S = \left\{\alpha \in \mathbb{R} : \mu\left(f^{-1}((\alpha, \infty])\right) = 0 \text{ and } \alpha \geq \inf f(X)\right\}$ and so $S = \varnothing$. Thus $\operatorname{ess\,sup} f = \infty$.

Now suppose $\operatorname{ess\,sup} f = \gamma < \infty$. Plainly $\operatorname{ess\,sup} f = \gamma \le \sup f$. As $\gamma \in S$, $\mu\left(f^{-1}((\gamma,\infty])\right) = 0$. Since $f^{-1}((\gamma,\infty])$ is open and of measure zero it must be an empty set as the only open set of measure zero in $\mathcal{M}$ is the empty set. Therefore, $\gamma \ge f(x)$ for all x in X. This means $\gamma \ge \sup f$ and so $\gamma = \sup f$.

Theorem 15. Suppose (X, $\mathcal{M}$, μ) is a measure space. $L^\infty(\mu)$ is a Banach space. More precisely, the equivalence classes of almost everywhere equal measurable functions on X, which are essentially bounded, is a Banach space with the metric induced by the essential sup norm.

Proof:

Let (f_n) be a Cauchy sequence in $L^\infty(\mu)$.

Let $A_n = \{x \in X : |f_n(x)| > \|f_n\|_\infty\}$. Then by definition of $\|f_n\|_\infty$, $\mu(A_n) = 0$. Let $B_{m,n} = \{x \in X : |f_n(x) - f_m(x)| > \|f_n - f_m\|_\infty\}$. Then we too have $\mu(B_{m,n}) = 0$.

Let $N = \bigcup_{n=1}^{\infty} A_n \cup \bigcup_{m=1,n=1}^{\infty} B_{m,n}$. Then by countable additivity, $\mu(N) = 0$. Thus, if x is in the complement of N, then $|f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty$. Now given any $\varepsilon > 0$, since (f_n) is a Cauchy sequence in $L^\infty(\mu)$, there exists an integer M such that $n, m \ge M$ implies that $\|f_n - f_m\|_\infty < \varepsilon$. Therefore, for all x in N^c,

$$|f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty < \varepsilon \text{ for } n, m \ge M. \text{ ----------------------------- (1)}$$

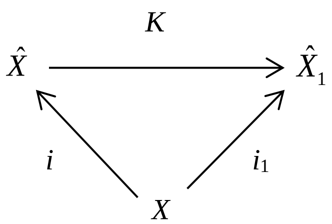
That is, $(f_n(x))$ is uniformly Cauchy on N^c. Moreover, for each x in N^c , $(f_n(x))$ is a Cauchy sequence of complex numbers and so it is convergent. Now for all x in N^c, $n \ge M$ implies that

$$|f_n(x)| = |f_n(x) - f_M(x) + f_M(x)| \le |f_n(x) - f_M(x)| + |f_M(x)| \le \|f_n - f_M\|_\infty + \|f_M\|_\infty < \varepsilon + \|f_M\|_\infty .$$

Therefore, for all x in N^c, $|f(x)| = \lim_{n\to\infty} |f_n(x)| \le \varepsilon + \|f_M\|_\infty$. Hence $f(x)$ is a bounded function on N^c. Now define $f(x) = 0$ for $x \in N$. Since N^c is measurable, f is a measurable function on N^c since it is the pointwise limit of measurable functions on N^c. As N is measurable, and so the extension to X by defining $f(x) = 0$ for $x \in N$, is measurable on X. (Note that each f_n restricted to N^c is measurable in the measure subspace (N^c, $\mathcal{M} \cap N^c$, $\mu|\mathcal{M} \cap N^c$) as $\mathcal{M} \cap N^c \subseteq \mathcal{M}$. Therefore, the pointwise limit, f , is $\mu|\mathcal{M} \cap N^c$ measurable and so is μ-measurable. Extending f to include N of μ-measure zero, plainly makes f a μ-measurable function on X as the preimage of f of any open set is either the preimage of the restriction of f to N^c or the union of this preimage with the μ-measurable set N, if 0 belongs to the open set. (The preimage of any open set of the restriction of f to N^c is μ-measurable and so the preimage of any open set of the extended function is μ-measurable.) Thus, f is measurable and $\|f\|_\infty = \text{ess}\sup|f| \le \varepsilon + \|f_M\|_\infty < \infty$. This means $f \in L^\infty(\mu)$. Observe that from (1) by letting m tend to infinity, for $n \ge M$, $|f_n(x) - f(x)| \le \varepsilon$ for all x in N^c. Therefore, $\sup\{|f_n(x) - f(x)| : x \notin N\} \le \varepsilon$ and as $\mu(N) = 0$, it follows that $\|f_n - f\|_\infty \le \varepsilon$. Hence $f_n \to f$ in $L^\infty(\mu)$. It follows that every Cauchy sequence in $L^\infty(\mu)$ is convergent and converges to a measurable function in $L^\infty(\mu)$ and so the equivalence classes of almost everywhere equal measurable functions in $L^\infty(\mu)$ with the induced metric is a Banach space.

4. A Special Dense Subspace of $L^p(\mu)$, $1 \le p < \infty$, Approximation by Continuous function, Lusin's Theorem.

If (X, d) is a metric space, then $(\hat{X}, \hat{d})$ is a *completion* of (X, d), if $(\hat{X}, \hat{d})$ is complete and there is an identification (isometry) of (X, d) with some dense subset of $(\hat{X}, \hat{d})$. That is to say, there is an isometric map $i : X \to \hat{X}$ such that every element of $\hat{X}$ is the limit of some sequence in $i(X)$ or that $i(X)$ is *dense* in $\hat{X}$. An *isometric map* between two metric spaces is a distance or metric preserving map, i.e., if (X_1, d_1) and (X_2, d_2) are metric spaces and $T : X_1 \to X_2$ is a map, then T is an isometry or an isometric map, if for all x and y in X_1, $d_2(T(x), T(y)) = d_1(x, y)$. Observe that any isometric map is injective for if $T(x) = T(y)$, then $d_1(x, y) = d_2(T(x), T(y)) = 0 \Rightarrow x = y$. Note that every metric space has a completion. The completion $(\hat{X}, \hat{d})$ is unique up to isometric isomorphism, a bijective isometry. For linear metric space (X, d), the completion $(\hat{X}, \hat{d})$ is also a linear space and the map $i : X \to \hat{X}$ should be a linear isometric map and the completion $(\hat{X}, \hat{d})$ is unique up to linear isometric isomorphism. More precisely, suppose $(\hat{X}_1, \hat{d}_1)$ is also a completion of (X, d) and $i_1 : X \to \hat{X}_1$ is an isometric embedding. Then there exists an isometric isomorphism $K : \hat{X} \to \hat{X}_1$ such that $K \circ i = i_1$.



Take $y \in \hat{X}$. Then there exists a sequence (x_n) in X such that $i(x_n) \to y$ in $(\hat{X}, \hat{d})$. Then $(i(x_n))$ is a Cauchy sequence in $(\hat{X}, \hat{d})$. This implies (x_n) is a Cauchy sequence in (X, d). Since $i_1 : X \to \hat{X}_1$ is an isometry, $(i_1(x_n))$ is also a Cauchy sequence in $(\hat{X}_1, \hat{d}_1)$ and so is convergent in $(\hat{X}_1, \hat{d}_1)$. Define $K(y) = \lim_{n\to\infty} i_1(x_n)$.

We now show that K is independent of the choice of the sequence (x_n).

Suppose (y_n) is another sequence in X such that $i(y_n) \to y$ in $(\hat{X}, \hat{d})$. Then

$$\hat{d}_1\left(i_1(y_k), \lim_{n\to\infty} i_1(x_n)\right) = \lim_{n\to\infty} \hat{d}_1\left(i_1(y_k), i_1(x_n)\right) = \lim_{n\to\infty} d\left(y_k, x_n\right) = \lim_{n\to\infty} \hat{d}\left(i(y_k), i(x_n)\right)$$

$$= \hat{d}\left(i(y_k), \lim_{n\to\infty} i(x_n)\right) = \hat{d}\left(i(y_k), y\right).$$

Therefore, $\lim_{k\to\infty} \hat{d}_1\left(i_1(y_k), \lim_{n\to\infty} i_1(x_n)\right) = \lim_{k\to\infty} \hat{d}\left(i(y_k), y\right) = 0$

and hence $\lim_{n\to\infty} i_1(y_n) = \lim_{n\to\infty} i_1(x_n)$.

Plainly, $K \circ i(x) = K(i(x)) = i_1(x)$ for all x in X.

K is an isometry.

$$\hat{d}_1\left(K(x), K(y)\right) = \hat{d}_1\left(\lim_{k\to\infty} i_1(x_k), \lim_{n\to\infty} i_1(y_n)\right), \text{ where } i(y_n) \to y \text{ and } i(x_k) \to x \text{ in } (\hat{X}, \hat{d})$$

$$= \lim_{k\to\infty} \hat{d}\left(i(x_k), \lim_{n\to\infty} i(y_n)\right) = \lim_{k\to\infty} \hat{d}\left(i(x_k), y\right) = \hat{d}\left(\lim_{k\to\infty} i(x_k), y\right) = \hat{d}\left(x, y\right).$$

K is plainly surjective. Any y in $(\hat{X}_1, \hat{d}_1)$ is the limit of a Cauchy sequence in $i_1(X)$. That is there is a sequence (y_n) in (X, d) such that $i_1(y_n) \to y$ in $(\hat{X}_1, \hat{d}_1)$. Then $(i(y_n))$ is a Cauchy sequence in $(\hat{X}, \hat{d})$. Therefore, $i(y_n)$ is convergent in $(\hat{X}, \hat{d})$. Let $x = \lim_{n\to\infty} i(y_n)$. Plainly, $K(x) = y$. Hence K is a surjective isometry and so is an isometric isomorphism. If (X, d) is a vector space, then the completion is also a vector space and $i : X \to \hat{X}$ is linear. It follows that K is a linear isometric isomorphism.

It is well known that every metric space has a completion that can be defined by equivalence classes of Cauchy sequences. For some function spaces as described below their completions are naturally also function spaces.

A measurable complex function on a measure space is *simple* if it takes on only finite number of complex values.

Proposition 16. Suppose $(X, \mathscr{M}, \mu)$ is a measure space. Let $S = \{s : X \to \mathbb{C}; s \text{ is a simple measurable function with } \mu(\{x : s(x) \neq 0\}) < \infty\}$ be endowed with $L^p(\mu)$ norm, i.e., $\|s\|_p = \left(\int_X |s|^p \, d\mu\right)^{\frac{1}{p}}$. Then $S \subseteq L^p(\mu)$ and is dense in $L^p(\mu)$ in the $L^p(\mu)$ metric. Hence the completion of $(S, \|\cdot\|_p)$ is $L^p(\mu)$.

Remark.

The collection of equivalence classes of almost everywhere equal measurable functions in $L^p(\mu)$ forms a normed linear space with the $\|\cdot\|_p$ norm. The $\|\cdot\|_p$ norm induces a metric on this collection of equivalence classes. By abuse of notation, we still denote this collection of equivalence classes by the same symbol, $L^p(\mu)$ and consider it as a metric space with the metric topology. When we say $S \subseteq L^p(\mu)$, we mean as a set S is a subset of

$$\{f : X \to \mathbb{C}; f \text{ is measurable and } \|f\|_p < \infty\}$$

and now by taking for each $s \in S$, the equivalence class it belongs to in $L^p(\mu)$, S considered as a collection of equivalence classes in $L^p(\mu)$ is a subset of all the equivalence classes of almost everywhere equal measurable functions in $L^p(\mu)$ and as a subset of the metric space $L^p(\mu)$ (equivalence classes), S (equivalence classes) is dense in $L^p(\mu)$ (equivalence classes). For simplicity of argument, we simply use the same symbol for the function and the equivalence class it belongs to in the proof and just use the representative before passing to the equivalence class without mentioning this last step.

Proof. Obviously, S is a vector space and $S \subseteq L^p(\mu)$. We shall show, for any f in $L^p(\mu)$, that given any $\varepsilon > 0$, there exists a function $s \in S$ such that $\|f - s\|_p < \varepsilon$.

Now for any f in $L^p(\mu)$, write f in its real and imaginary part, i.e., $f = u + i\,v$, where u = real part of f and v = imaginary part of f. Since f is measurable, u and v are measurable. Therefore, $u^+ = \sup\{u, 0\}$ and $u^- = \sup\{-u, 0\}$ are non-negative and measurable and the same is true for v^+ and v^-. Hence $f = u^+ - u^- + i\,v^+ - i\,v^-$.

So now we suppose $f \in L^p(\mu)$ and $f \geq 0$, i.e., f takes on only non-negative real values. Then there exists a sequence of measurable (non-negative) simple function, $(s_n : X \to \mathbb{R}^+)$, such that (s_n) is an increasing sequence of functions and $s_n \to f$ pointwise. Now as

$0 \le s_n \le f$, $\|s_n\|_p \le \|f\|_p < \infty$ and so $s_n \in S$ for all integer $n \ge 1$. Moreover since $|f - s_n|^p = (f - s_n)^p \le f^p$ and $\int_X f^p d\mu < \infty$, $(f - s_n)^p$ is dominated by the integrable function f^p. Therefore, by the Lebesgue Dominated Convergence Theorem, $\lim_{n\to\infty} \int_X |f - s_n|^p \, d\mu = 0$. Hence $\lim_{n\to\infty} \left(\int_X |f - s_n|^p \, d\mu \right)^{\frac{1}{p}} = \lim_{n\to\infty} \|f - s_n\|_p = 0$. This means that $s_n \to f$ in $L^p(\mu)$.

It now follows that for $f = u^+ - u^- + i\,v^+ - i\,v^-$, there exist sequences of measurable (non-negative) simple functions, $(u_n^+), (u_n^-), (v_n^+)$ and (v_n^-) such that $u_n^+ \to u^+, u_n^- \to u^-, v_n^+ \to v^+$ and $v_n^- \to v^-$ in $L^p(\mu)$. Therefore,

$$u_n^+ - u_n^- + i\,v_n^+ - i\,v_n^- \to u^+ - u^- + i\,v^+ - i\,v^- = f \quad \text{in} \quad L^p(\mu).$$

Plainly, $s_n = u_n^+ - u_n^- + i\,v_n^+ - i\,v_n^-$ is a measurable simple complex function and so $s_n \in S$ and $s_n \to f$ in $L^p(\mu)$.

Hence $L^p(\mu)$ is the closure of *S* in the equivalence classes of almost everywhere equal measurable functions in the metric topology induced by the L^p norm and so *S* is dense in $L^p(\mu)$. Now any Cauchy sequence (s_n) in $(S, \|\cdot\|_p)$ is also a Cauchy sequence in $L^p(\mu)$ and so as $L^p(\mu)$ is complete, (s_n) converges to a function in $L^p(\mu)$. It follows that the completion of *S* is $L^p(\mu)$. (Here for simplicity we use the same symbol to denote equivalence classes as well as its underlying space when considering the equivalence classes as metric space. More precisely, if $\mathscr{L}^p(\mu)$ denotes the equivalent classes of almost everywhere equal measurable functions in $L^p(\mu)$, then $\mathscr{L}^p(\mu)$ is a normed linear space with the $\|\cdot\|_p$ norm and is a metric space with the metric induced by the norm. If $\boldsymbol{\mathcal{S}}$ denotes the equivalence classes of almost everywhere equal measurable functions in $(S, \|\cdot\|_p)$, then we may embed $\boldsymbol{\mathcal{S}}$ as a subset of $\mathscr{L}^p(\mu)$ by simply assigning to each equivalence class in $\boldsymbol{\mathcal{S}}$, the extended equivalence class in $L^p(\mu)$. Thus in this way $\boldsymbol{\mathcal{S}}$ is a subset of the metric space $\mathscr{L}^p(\mu)$ and its closure in $\mathscr{L}^p(\mu)$ is $\mathscr{L}^p(\mu)$ and so $\boldsymbol{\mathcal{S}}$ is dense in $\mathscr{L}^p(\mu)$.) Note that for a complete metric space the completion of a dense subspace is the whole metric space itself. The above is just a simple verification of this fact.

When *X* is a special metric space, for example, the familiar $\mathbb{R}^n$ with the Euclidean metric, what is a suitable useful dense subspace of $L^p(\mu)$ other than the subspace *S* defined above? Note that $\mathbb{R}^n$ with the Euclidean metric is a topological space with the usual topology, it is locally compact and Hausdorff. So, we now assume that *X* is a locally compact Hausdorff space.

The set of continuous function on *X*, if endowed with the $L^p(\mu)$ norm, for some suitable measure space, (*X*, $\mathscr{M}$, μ), where $\mathscr{M}$ is a suitable σ algebra containing the Borel subsets of X and μ is some suitable positive measure on $\mathscr{M}$, is a normed linear space and the $L^p(\mu)$ norm gives rise to a metric on the equivalence classes of almost everywhere equal continuous functions.

Definition 17. Suppose *Y* is a topological space. Let $f : Y \to \mathbb{C}$ be a complex function. Then the *support* of f, denoted by supp f, is the closure of the set $\{x \in Y : f(x) \neq 0\}$, i.e., support f $= \overline{\{x \in Y : f(x) \neq 0\}}$. The collection of all continuous complex function on *Y* with compact support is denoted by $C_c(Y)$. That is to say, $f \in C_c(Y)$ if support of f is compact in *Y*. Then observe that $C_c(Y)$ is a vector space. We deduce this as follows. If f and *g* are in $C_c(Y)$, then

$$\{x \in Y : (f+g)(x) \neq 0\} \subseteq \overline{\{x \in Y : f(x) \neq 0\}} \cup \overline{\{x \in Y : g(x) \neq 0\}}$$

so that $\overline{\{x \in Y : (f+g)(x) \neq 0\}} \subseteq \overline{\{x \in Y : f(x) \neq 0\}} \cup \overline{\{x \in Y : g(x) \neq 0\}}$. This means

$$\text{supp}\,(f+g) \subseteq \text{supp}\, f \cup \text{supp}\, g.$$

As supp f and supp g are compact, supp $(f+g)$ is a closed subspace of a compact space and so is compact. Therefore, as $f+g$ is also continuous on *Y*, $f+g \in C_c(Y)$. For any $\alpha \neq 0$, $\beta \neq 0$, supp $(\alpha f + \beta g) \subseteq$ supp $f \cup$ supp g and so the same argument above shows that $\alpha f + \beta g \in C_c(Y)$. Thus $C_c(Y)$ is a vector pace over $\mathbb{C}$.

Now we elaborate on the measure space (*X*, $\mathscr{M}$, μ). It should ideally satisfy the following 6 properties:

(1) $\mathscr{M}$ is a σ algebra containing all the Borel subsets of X;

(2) μ is a positive measure on $\mathscr{M}$ satisfying:

(3) For all compact $K \subseteq X$, $\mu(K) < \infty$.

(4) For all $E \in \mathscr{M}$, $\mu(E) = \inf\{\mu(V) : V \supseteq E \text{ and } V \text{ is open}\}$ (Outer Regularity).

(5) For all $E \in \mathscr{M}$ such that either E is open or $\mu(E) < \infty$,

$$\mu(E) = \sup\{\mu(K) : K \subseteq E \text{ and } K \text{ is compact}\} \quad \text{(Inner Regularity).}$$

(6) $\mathscr{M}$ is *μ-complete*, i.e., for all $N \in \mathscr{M}$ such that $\mu(N)=0$, for all $E \subseteq N$, $E \in \mathscr{M}$.

Properties (1) to (6) is satisfied by the Lebesgue measure on $\mathbb{R}$.

If (X, $\mathscr{M}$, μ) is a measure space that satisfies properties (4) and (5) for all $E \in \mathscr{M}$ without any condition, then it is said to be *regular* . If (X, $\mathscr{B}(X)$, μ) is a Borel measure space, i.e., $\mathscr{B}(X)$ is the σ algebra generated by the Borel subsets of X and μ a positive measure on $\mathscr{B}(X)$, and if it satisfies (3), (4) and (5), then it is called a *Radon measure* on X.

Every measure space (X, $\mathscr{M}$, μ) has a completion (X, $\mathscr{M}^*$, μ^*), where $\mathscr{M}^*$ is complete and $\mu^* : \mathscr{M}^* \to [0, \infty]$ is a positive measure whose restriction to $\mathscr{M}$ is μ . Thus the completion of a Radon measure (X, $\mathscr{B}(X)$, μ) satisfies all six properties above.

For $X = \mathbb{R}^n$, the Lebesgue measure space ($\mathbb{R}^n$, $\mathscr{M}$, μ), where $\mathscr{M}$ consists of the Lebesgue measurable sets and μ is the n-dimensional Lebesgue measure, satisfies all of the above six properties.

The next result gives an approximation of measurable function on X by function in $C_c(X)$ and is the key or technical result needed in proving the density of a subspace of the function space to be introduced later.

Theorem 18. Lusin's Theorem.

Suppose X is a locally compact Hausdorff topological space and (X, $\mathscr{M}$, μ) is a measure space satisfying properties (1) to (5) above.

Suppose $f : X \to \mathbb{C}$ is a measurable function such that $\{x \in X : f(x) \neq 0\} \subseteq A$ and $A \in \mathscr{M}$, where $\mu(A) < \infty$. Then for any $\varepsilon > 0$, there exists $g \in C_c(X)$ such that

$$\mu\{x \in X : g(x) \neq f(x)\} < \varepsilon .$$

Moreover, we may arrange it so that $\|g\| = \sup_{x \in X} |g(x)| \leq \sup_{x \in X} |f(x)|$.

Before we prove this theorem, we shall recall some topological facts.

5. Topological Ideas

Recall a topological space is a set X with a collection $\mathscr{T}$ of subsets of X called open sets such that (i) $\varnothing$, $X \in \mathscr{T}$ and (ii) $\mathscr{T}$ is closed under arbitrary union and finite intersection, i.e., if $\mathscr{S}$ is any sub-collection of $\mathscr{T}$, then $\bigcup \{U: U \in \mathscr{S}\} \in \mathscr{T}$ and if U_1 , U_2 , … , $U_n \in \mathscr{T}$, then

$\bigcap_{i=1}^{n} U_i \in \mathscr{T}$. When we speak of a topological space, the topology or the collection $\mathscr{T}$ of open sets is understood to be given. A set is *open* if it belongs to $\mathscr{T}$. We do not normally specify the topology $\mathscr{T}$. A *neighbourhood* of a point *x* or a set *E* is a set *N* such that it contains an open set *U* containing the point *x* or set *E* respectively. A set is *closed* if its complement is open. A sequence $\{x_n\}_{n=1}^{\infty}$ in a topological space is said to converge to a point *x*, written $x_n \to x$, if for any open set *U* containing *x*, there exists an integer *M* such that $n \geq M \Rightarrow x_n \in U$. A topological space is *Hausdorff* if for any $x \neq y$, there are open neighbourhoods *U* and *V* of *x* and *y* respectively such that $U \cap V = \varnothing$. If a topological space *X* is Hausdorff, the limit of a sequence in *X* is unique. A point x_0 is a *limit point* of a subset *A* in a topological space if every open set containing x_0 also contains a point of *A* distinct from x_0. The closure $\overline{S}$ of a subset *S* in a topological space *X* is the smallest closed subset containing *S*, more precisely,

$$\overline{S} = \bigcap\{F : F \supseteq S,\ F \text{ is closed}\} .$$

Let S' be the set of limit points of *S*, then $\overline{S} = S \cup S'$. As a consequence of this, a point $x \in \overline{S}$ if and only if every neighbourhood *N* of *x* has $N \cap S \neq \varnothing$. Therefore, if there exists an open set *U* containing *x* such that $U \cap S = \varnothing$, then $x \notin \overline{S}$. It is to be noted that if *A* and *B* are subsets of a topological space *X* and $A \subseteq B$, then $\overline{A} \subseteq \overline{B}$ and that if *F* is closed in *X* and $S \subseteq F$, then $\overline{S} \subseteq F$. A subset *E* of a topological space *X* is said to be *dense* in *X* if $\overline{E} = X$. A topological space *X* is said to be *compact* if for any open cover $\mathscr{S}$ of *X*, i.e., $\mathscr{S}$ consists of a collection of open sets in *X* such that $\bigcup\{U \in \mathscr{S}\} = X$, there exists a finite subcover $\{U_1, U_2, \cdots, U_n\} \subseteq \mathscr{S}$ such that $U_1 \cup U_2 \cup \cdots \cup U_n = X$. A subset *A* of a topological space $(X, \mathscr{T})$ inherits the topology from *X* called the *subspace topology* or *relative topology* $\mathscr{T}|_A = \{U \cap A : U \in \mathscr{T}\}$. A subset of a topological space is said to be compact if it is compact with the relative topology. A subset *K* in $\mathbb{R}^n$ is compact if and only if it is closed and bounded (Generalized Heine-Borel Theorem). A map $f : (X_1, \mathscr{T}_1) \to (X_2, \mathscr{T}_2)$ between two topological spaces is said to be *continuous* if for any open set *U* in X_2, i.e., $U \in \mathscr{T}_2$, $f^{-1}(U)$ is open in X_1 or $f^{-1}(U) \in \mathscr{T}_1$. The continuous image of a compact set is compact. The *usual* or *metric topology* on a metric space (*M*, *d*), is the induced topology T_d on (*M*, *d*) generated by the family of open balls, i.e., open balls and arbitrary union of open balls in (*M*, *d*). In a metric space (*M*, *d*) the limit point of a subset *S* in *M* is precisely the limit of a sequence of distinct points in *S*. Therefore, if $(\hat{X}, \hat{d})$ is a *completion* of (*X*, *d*) and $i : X \to \hat{X}$ is the isometric embedding of *X*, then the closure $\overline{i(X)}$ of $i(X)$ in $\hat{X}$ is $\hat{X}$ or equivalently $i(X)$ is dense in $\hat{X}$.

(1) A topological space Y is said to be *locally compact* if each point y in Y has a compact *neighbourhood*. That is to say, there exists a compact subspace K of X and an open set such that $y \in V \subseteq K$. A subset $E \subseteq Y$ is said to be *relatively compact* if its closure $\overline{E}$ is compact.

(2) A compact subspace of a Hausdroff space Y is closed in Y.

(3) A closed subset of a compact set is compact.

(4) Therefore, for a Hausdorff space Y, Y is locally compact if for each point y in Y, there exists an open set V such that $y \in V$ and the closure $\overline{V}$ is compact.

(5) For a Hausdorff space Y, if $\{K_\alpha\}$ is a collection of compact sets in Y such that $\bigcap K_\alpha = \varnothing$, then some finite intersection also has empty intersection.

Proof.

Since Y is Hausdorff, each K_α is closed in Y. Therefore, the complement K_α^c of K_α is open in Y. Choose a member in $\{K_\alpha\}$, say K_1. Then since $K_1 \cap \bigcap_{\alpha \neq 1} K_\alpha = \varnothing$,

$K_1 \subseteq \left(\bigcap_{\alpha \neq 1} K_\alpha \right)^c = \bigcup_{\alpha \neq 1} K_\alpha^c$. Therefore, $\{K_\alpha^c : \alpha \neq 1\}$ is an open cover for K_1. As K_1 is compact, it has a finite subcover say $\{K_{\alpha_1}{}^c, K_{\alpha_2}{}^c, \cdots, K_{\alpha_n}{}^c\}$ and

$K_1 \subseteq \{K_{\alpha_1}{}^c \cup K_{\alpha_2}{}^c \cup \cdots \cup K_{\alpha_n}{}^c\} = \left(\bigcap_{j=1}^{n} K_{\alpha_j} \right)^c$. This means $K_1 \cap \bigcap_{j=1}^{n} K_{\alpha_j} = \varnothing$.

(6) For a Hausdorff space Y, we say Y is *regular* if for every closed set A in Y and $x \notin A$, they have disjoint neighbourhoods, i.e., there exists open set $N \supseteq A$ and open set V, $x \in V$ such that $N \cap V = \varnothing$. More precisely Y is regular Hausdorff or satisfies the T_3 separation axiom.

(7) For a regular Hausdorff space Y, for any $y \in Y$ and any neighbourhood N of y, there exists an open neighbourhood V of y such that $y \in V \subseteq \overline{V} \subseteq N$.

Proof.

Let $y \in Y$ and N be an open neighbourhood of y. Then the complement of N, N^c is closed in Y and $y \notin N^c$. Since Y is regular Hausdorff, there exist open sets V and U in Y such that $y \in V$, $N^c \subseteq U$ and $V \cap U = \varnothing$. Therefore, $V \subseteq U^c$, which is closed in Y. Hence $\overline{V} \subseteq U^c \subseteq N$. It follows that $y \in V \subseteq \overline{V} \subseteq N$.

(8) A Compact Hausdorff space is regular.

Proof.

Suppose *Y* is a compact Hausdorff topological space.

Take any closed set *A* in *Y* and $x \notin A$. Then *A* is compact and it follows that there exists open sets *N* and *V* such that $N \supseteq A$, $x \in V$ and $N \cap V = \varnothing$. Therefore, *Y* is regular.

(9) For a Hausdorff space *Y*, *Y* is locally compact if and only if for each point $y \in Y$ and for any neighbourhood *N* of *y*, there exists a neighbourhood *V* of *y* such that $\overline{V}$ is compact and $\overline{V} \subseteq N$.

Proof. ($\Leftarrow$) This is obvious. For any $y \in Y$, take any neighbourhood *N* of *y*. Just take $K = \overline{V}$.

($\Rightarrow$) Take any $y \in Y$ and any neighbourhood *N* of *y*. Since *Y* is locally compact, there exists a compact subspace *K* of *X* and an open set such that $y \in V \subseteq K$. Since *Y* is Hausdorff , *K* is closed. Therefore, $\overline{V} \subseteq K$ and *K* is compact and so by (3), $\overline{V}$ is compact. In particular, $\overline{V}$ is Hausdorff. Therefore, by (8) $\overline{V}$ is a regular topological space with the relative topology. As *N* is a neighbourhood of *y*, there exists an open set *U* in *Y* such that $y \in U \subseteq N$. Now $U \cap \overline{V}$ is an open set in $\overline{V}$ with the relative topology. This means $U \cap \overline{V}$ is a neighbourhood of *y* in $\overline{V}$ with the relative topology. Therefore, by (7), there exists a neighbourhood *W* of *y* open in $\overline{V}$ with the relative topology, such that $y \in W \subseteq$ closure of W in $\overline{V} \subseteq U \cap \overline{V}$.

Since *W* is relatively open in $\overline{V}$, there exists an open set *D* in *Y* such that $W = D \cap \overline{V}$. Let $E = D \cap V$. Then *E* is open and $E \subseteq W$ and $E \subseteq V$ and so $\overline{E} \subseteq \overline{W} \cap \overline{V}$. As closure of W in $\overline{V}$ is $\overline{W} \cap \overline{V}$ we have $y \in E \subseteq \overline{E} \subseteq U \cap \overline{V} \subseteq N$. Moreover $\overline{E}$ is compact.

(10) Suppose *Y* is a locally compact Hausdorff topological space. Suppose *U* is open in *Y* and *K* is a compact subset such that $K \subseteq U$. Then there is a relatively compact open set, i.e., an open set *V* where $\overline{V}$ is compact, such that $K \subseteq V \subseteq \overline{V} \subseteq U$.

Proof.

Suppose *K* is compact, $K \subseteq U$ and *U* is open in *Y*. For each point $y \in K$, *y* has a relatively compact open neighbourhood, V_y. Then the collection $\{V_y : y \in Y\}$ is an open covering for *K*. As *K* is compact a finite number of these open sets say, $\{V_{y_1}, V_{y_2}, \cdots, V_{y_n}\}$, cover *K*. Let $G = V_{y_1} \cup V_{y_2} \cup \cdots \cup V_{y_n}$. Then *G* is open, $K \subseteq G$ and $\overline{G}$ is compact.

If *U* is all of *Y*, then take *V* = *G* and $K \subseteq V \subseteq \overline{V} \subseteq U$.

Now assume *U* is not all of *Y*.

Now for $y \notin U$ and hence $y \notin K$ and as K is compact and Y is Hausdorff, there exists open neighbourhood V_y of K such that $K \subseteq V_y$ and open set E_y such that $y \in E_y$ and $E_y \cap V_y = \varnothing$. As $E_y \cap V_y = \varnothing$, $y \notin \overline{V_y}$.

Then the collection $\left\{U^c \cap \overline{G} \cap \overline{V_y} : y \in U^c\right\}$ is a collection of compact sets as each $U^c \cap \overline{G} \cap \overline{V_y}$ is compact for $\overline{G}$ is compact and $U^c \cap \overline{V_y}$ is closed (see (3)). Since $\bigcap\left\{\overline{V_y} : y \in U^c\right\} \cap U^c = \varnothing$ we have that $\bigcap\left\{U^c \cap \overline{G} \cap \overline{V_y} : y \in U^c\right\} = \varnothing$. Therefore by (5) a finite intersection of these compact sets is also empty. That is there exists $\left\{y_1, y_2, \cdots, y_n\right\} \subseteq U^c$ such that $\bigcap_{i=1}^{n} U^c \cap \overline{G} \cap \overline{V_{y_i}} = \varnothing$.

Let $V = G \cap V_{y_1} \cap V_{y_2} \cap \cdots \cap V_{y_n}$. Then $K \subseteq V$ and $\overline{V} = \overline{G \cap V_{y_1} \cap V_{y_2} \cap \cdots \cap V_{y_n}} \subseteq \overline{G} \cap \overline{V_{y_1}} \cap \overline{V_{y_2}} \cap \cdots \cap \overline{V_{y_n}}$. Thus $\overline{V}$ is a closed subset of $\overline{G}$, a compact set and so it is compact.

As $\bigcap_{i=1}^{n} U^c \cap \overline{G} \cap \overline{V_{y_i}} = \left(\overline{G} \cap \overline{V_{y_1}} \cap \overline{V_{y_2}} \cap \cdots \cap \overline{V_{y_n}}\right) \cap U^c = \varnothing$, $\overline{G} \cap \overline{V_{y_1}} \cap \overline{V_{y_2}} \cap \cdots \cap \overline{V_{y_n}} \subseteq U$. It follows that $K \subseteq V \subseteq \overline{V} \subseteq U$.

(11) A locally compact Hausdorff space Y is regular.

Proof. Take a closed set A in Y. Take any x not in A. Then $x \in A^c$. Now $U = A^c$ is open and so is a neighbourhood of x. Then by (9) there exists a neighbourhood V of x such that $\overline{V}$ is compact and $\overline{V} \subseteq U$. Then $\left(\overline{V}\right)^c$ is open and $\left(\overline{V}\right)^c \supseteq U^c = A$. Thus $\left(\overline{V}\right)^c$ is an open neighbourhood of A, V is a neighbourhood of x and $V \cap \left(\overline{V}\right)^c = \varnothing$. Hence Y is regular.

(12) Open subset of a locally compact Hausdorff space is locally compact and Hausdorff; closed subset of a locally compact Hausdorff space is locally compact and Hausdorff.

Proof.

Suppose X is locally compact. Then any subset A with the subspace topology is Hausdorff. For take any x and y in A with $x \neq y$. Then since X is Hausdorff, there are subsets U and V open in X such that $x \in U$, $y \in V$ and $U \cap V = \varnothing$. Suppose F is a closed subset of X. Take $x \in F$. Then there exists a compact subspace K of X and an open set V such that $x \in V \subseteq K$. Since X is Hausdorff, K is closed and as F is closed, $F \cap K$ is closed in X and is a subset of the Compact set K, it follows that $F \cap K$ is compact. Moreover $K \cap F \supseteq V \cap F$ and $V \cap F$ is relatively open in F. Hence $F \cap K$ is a compact neighbourhood of $x \in F$ in the relative topology. This means that F with the subspace topology is locally compact. It is of course

Hausdorff. Suppose H is an open subset of X. Take $x \in H$. Then H is an open neighbourhood of x in X. Then by (9), there exists a relatively compact open set V such that $x \in V$ and $V \subseteq \overline{V} \subseteq H$. As $\overline{V}$ is compact and $\overline{V} \subseteq H$, $\overline{V}$ is a compact neighbourhood of x in the subspace topology of H. Hence H is locally compact. As shown before, H is Hausdorff since X is.

We can approximate a measurable function by using characteristic functions of measurable sets. Characteristic function is not usually continuous but is nearly continuous, a fact that is very useful. It is semi-continuous. Characteristic functions are, in some sense, building blocks of measurable functions. We now discuss some properties of semi-continuity and its relation with continuity.

Lower and Upper Semi Continuous Functions

Definition 19.

Suppose X is a topological space.

Then a function $f : X \to \mathbb{R}$ or $\overline{\mathbb{R}}$ is said to be *lower semi-continuous* (abbreviated l.s.c.) if $\{x : f(x) > \alpha\}$ is open for all $\alpha \in \mathbb{R}$. The function f is said to be *upper semi-continuous* (abbreviated u.s.c.) if $\{x : f(x) < \alpha\}$ is open for all $\alpha \in \mathbb{R}$.

We say $f : X \to \mathbb{R}$ is *continuous* at $x_0 \in X$ if for all $\varepsilon > 0$, there exist an open set U_{x_0} containing x_0 such that for all $x \in U_{x_0}$, $-\varepsilon < f(x) - f(x_0) < \varepsilon$.

$f : X \to \mathbb{R}$ is *lower semi-continuous* at $x_0 \in X$ if for all $\varepsilon > 0$, there exist an open set U_{x_0} containing x_0 such that for all $x \in U_{x_0}$, $-\varepsilon < f(x) - f(x_0)$.

$f : X \to \mathbb{R}$ is *upper semi-continuous* at $x_0 \in X$ if for all $\varepsilon > 0$, there exist an open set U_{x_0} containing x_0 such that for all $x \in U_{x_0}$, $f(x) - f(x_0) < \varepsilon$.

It is clear that a real valued function f is *continuous* at $x_0 \in X$ if and only if f is *lower semi-continuous* at x_0 and *upper semi-continuous* at x_0.

Proposition 20.

(1) $f : X \to \mathbb{R}$ is u.s.c. if and only if f is u.s.c. at x for all $x \in X$.

(2) $f : X \to \mathbb{R}$ is l.s.c. if and only if f is l.s.c. at x for all $x \in X$.

Proof. It is clear that $f : X \to \mathbb{R}$ is u.s.c. (l.s.c.) implies f is u.s.c. (l.s.c) at x for all $x \in X$.

Suppose f is u.s.c. at x for all $x \in X$.

Take any $\alpha \in \mathbb{R}$. Let $E = \{x : f(x) < \alpha\}$. For each *x* in *E*, there exist an open set U_x containing *x* such that for all $y \in U_x$, $f(y) - f(x) < \varepsilon$. Take $\varepsilon = \dfrac{\alpha - f(x)}{2} > 0$. Then for all $y \in U_x$, $f(y) < \varepsilon + f(x) = \dfrac{\alpha + f(x)}{2} < \alpha$. Therefore, U_x is an open neighbourhood of *x* and $U_x \subseteq E$. Hence *E* is a neighbourhood of each of its points and so *E* is open. Hence *f* is u.s.c .

Suppose f is l.s.c. at *x* for all $x \in X$.

Take any $\alpha \in \mathbb{R}$. Let $E = \{x : f(x) > \alpha\}$. For each *x* in *E*, there exist an open set U_x containing *x* such that for all $y \in U_x$, $f(y) - f(x) > -\varepsilon$. Take $\varepsilon = \dfrac{f(x) - \alpha}{2} > 0$. Then for all $y \in U_x$, $f(y) > f(x) - \varepsilon = \dfrac{f(x) + \alpha}{2} > \alpha$. Therefore, U_x is an open neighbourhood of *x* and $U_x \subseteq E$. Hence *E* is a neighbourhood of each of its points and so *E* is open. Therefore, f is l.s.c .

Remark.

$f : X \to \mathbb{R}$ is continuous if and only if f is both l.s.c. and u.s.c.

The characteristic function χ_A is u.s.c if *A* is a closed set and is l.s.c. if *A* is an open set.

Proposition 21.

(1) If $\{f_\beta : X \to \mathbb{R}\}$ is a collection of lower semi-continuous functions, then $\sup_\beta f_\beta$ is lower semi-continuous.

(2) If $\{f_\alpha : X \to \mathbb{R}\}$ is a collection of upper semi-continuous functions, then $\inf_\alpha f_\alpha$ is upper semi-continuous.

Proof.

(1) Take any $\alpha \in \mathbb{R}$. Since each $f_\beta : X \to \mathbb{R}$ is lower semi-continuous, the set $U_\beta = \{x : f_\beta(x) > \alpha\}$ is open in *X*.

Observe that $E = \left\{x : \sup_\beta f_\beta(x) > \alpha\right\} = \bigcup_\beta \{x : f_\beta(x) > \alpha\} = \bigcup_\beta U_\beta$.

Plainly, $f_\beta(x) > \alpha \Rightarrow \sup_\beta f_\beta(x) > \alpha$ and so $U_\beta \subseteq E$ for each β. Hence $\bigcup_\beta U_\beta \subseteq E$.

Now if $x \in E$, then $\sup_{\beta} f_{\beta}(x) > \alpha$ and so there exists γ such that $\sup_{\beta} f_{\beta}(x) \geq f_{\gamma}(x) > \alpha$. Thus $x \in U_{\gamma}$. Therefore, $E \subseteq \bigcup_{\beta} U_{\beta}$ and so $E = \bigcup_{\beta} U_{\beta}$. Since U_{β} is open for each β, E is open in *X*. This means $\sup_{\beta} f_{\beta}$ is lower semi-continuous.

(2) Suppose each $f_{\beta} : X \to \mathbb{R}$ is upper semi-continuous Take any $\alpha \in \mathbb{R}$. Since each $f_{\beta} : X \to \mathbb{R}$ is upper semi-continuous, the set $V_{\beta} = \{x : f_{\beta}(x) < \alpha\}$ is open in *X*.

Observe that $D = \left\{x : \inf_{\beta} f_{\beta}(x) < \alpha\right\} = \bigcup_{\beta} \{x : f_{\beta}(x) < \alpha\} = \bigcup_{\beta} V_{\beta}$.

Plainly, $f_{\beta}(x) < \alpha \Rightarrow \inf_{\beta} f_{\beta}(x) < \alpha$ and so $V_{\beta} \subseteq D$ for each β. Hence $\bigcup_{\beta} V_{\beta} \subseteq D$.

Now if $x \in D$, then $\inf_{\beta} f_{\beta}(x) < \alpha$ and so there exists γ such that $\inf_{\beta} f_{\beta}(x) \leq f_{\gamma}(x) < \alpha$. Thus $x \in V_{\gamma}$. Therefore, $D \subseteq \bigcup_{\beta} V_{\beta}$ and it follows that $D = \bigcup_{\beta} V_{\beta}$. Since V_{β} is open for each β, *D* is open in *X*. This means $\inf_{\beta} f_{\beta}$ is upper semi-continuous.

We shall use the properties of some special complex function on a locally compact Hausdorff topological space to approximate a given function.

To describe these properties, we introduce the ideas in the definitions that follow.

Suppose *X* is a topological space and $C_c(X)$ is the vector space of all continuous complex functions with compact support.

We write $K \prec f$ and say f *dominates* K if

(i) *K* is compact,
(ii) $f \in C_c(X)$,
(iii) $0 \leq f \leq 1$ and
(iv) $f(x) = 1$ for all *x* in *K*.

We write $f \prec U$ and say *U dominates* f if

(i) $U \subseteq X$ is open,
(ii) $f \in C_c(X)$,
(iii) $0 \leq f \leq 1$ and
(iv) Support $f \subseteq U$.

We write $K \prec f \prec U$ if $K \prec f$ and $f \prec U$

Lemma 22. Urysohn's Lemma

Suppose X is a locally compact Hausdorff space, $U \subseteq X$ is open, K is compact with $K \subseteq U$. Then there exists $f \in C_c(X)$ such that $K \prec f \prec U$.

Remark. This means the characteristic functions χ_K and χ_U satisfy $\chi_K \le f \le \chi_U$.

Proof.

We shall make use of the rational number in [0, 1] to construct the Urysohn function f. Take an enumeration $r : \mathbb{N} \to [0,1]$ of the rational numbers, i.e., a bijective function of $\mathbb{N}$ onto [0, 1] such that $r_1 = r(1) = 0$ and $r_2 = r(2) = 1$. We denote the image $r(k)$ by r_k.

Suppose K is compact, $K \subseteq U$ and U is open.

Let $U_{r_1} = U_0$ be the relatively open compact set as given by (10) in the section on topological ideas, as X is locally compact and Hausdorff, such that

$$K \subseteq U_0 \subseteq \overline{U_0} \subseteq U \text{ ------------------ (1).}$$

Let $U_{r_2} = U_1$ be the relatively open compact set as given by (10) of *topological ideas* such that

$$K \subseteq U_1 \subseteq \overline{U_1} \subseteq U_0 \text{ ------------------- (2).}$$

We shall inductively define the relatively compact set U_{r_k}.

Suppose $U_{r_1}, U_{r_2}, \cdots, U_{r_n}$ have been chosen so that if $r_i < r_j$, $j \le n$, then $U_{r_j} \subseteq \overline{U_{r_j}} \subseteq U_{r_i}$. Then arrange $r_1, r_2, \cdots, r_n$ in increasing order. Suppose in this sequence $r_i < r_{n+1} < r_j$. Then using $\overline{U_{r_j}} \subseteq U_{r_i}$, by (10) choose relatively compact open $U_{r_{n+1}}$ such that

$$\overline{U_{r_j}} \subseteq U_{r_{n+1}} \subseteq \overline{U_{r_{n+1}}} \subseteq U_{r_i}. \text{ ----------------------------- (3)}$$

In this way we obtain a collection of relatively compact open sets $\{U_r : r \text{ rational } \in [0,1]\}$ satisfying $\overline{U_s} \subseteq U_r$ whenever $s > r$, $K \subseteq U_1$ and $\overline{U_0} \subseteq U$.

Define a collection of functions $\{f_r : r \text{ rational } \in [0,1]\}$ by defining $f_r : X \to [0,1]$ by

$$f_r(x) = \begin{cases} r, & \text{if } x \in U_r \\ 0, & \text{otherwise} \end{cases}$$

and a collection $\{g_s : s \text{ rational } \in[0,1]\}$ by defining $g_s : X \to [0,1]$ by

$$g_s(x) = \begin{cases} 1\,, & \text{if } x \in \overline{U_s} \\ s\,, & \text{otherwise} \end{cases} .$$

Note that $f_r = r\chi_{U_r}$. Since U_r is open for each rational $r \in [0,1]$, f_r is lower semi-continuous for each rational $r \in [0,1]$. Observe that $\{x : g_s(x) < \alpha\} = X$ if α > 1, $\{x : g_s(x) < \alpha\} = \left(\overline{U_s}\right)^c$ if $s < \alpha \le 1$ and $\{x : g_s(x) < \alpha\} = \varnothing$ if $\alpha \le s$. Thus g_s is upper semi-continuous for each rational $s \in [0,1]$.

Therefore, by Proposition 21, $f = \sup_r \{f_r : r \text{ rational } \in [0,1]\}$ is lower semi-continuous and $g = \inf_r \{g_r : r \text{ rational } \in [0,1]\}$ is upper semi-continuous.

We shall next show that $f = g$ and so f is both lower and upper semi-continuous and so f is continuous.

Firstly, we show that $f \le g$.

Suppose on the contrary, there exists *x* in *X* such that $f(x) > g(x)$. Then by the definition of supremum, there exists *r* in $\mathbb{Q} \cap [0,1]$ such that $f_r(x) > g(x)$. Next by the definition of infimum, there exists *s* in $\mathbb{Q} \cap [0,1]$ such that $f_r(x) > g_s(x)$. This can only happen if $x \notin \overline{U_s}$, $x \in U_r$ and $r > s$. But $r > s$ implies that $U_r \subseteq \overline{U_r} \subseteq U_s \subseteq \overline{U_s}$ and so $x \in \overline{U_s}$ and we have a contradiction. This proves that $f \le g$.

Next, we show that $f \ge g$.

Suppose on the contrary, there exists *x* in *X* such that $f(x) < g(x)$. Then by the density of the rational numbers we can find rational numbers *r* and *s* such that $f(x) < s < r < g(x)$.

Since $f(x) < s$, $x \notin U_s$ and since $g(x) > r$, $x \in \overline{U_r}$. As $s < r$, $U_r \subseteq \overline{U_r} \subseteq U_s \subseteq \overline{U_s}$ and so $x \in U_s$ and we arrived at a contradiction and so we have $f \ge g$. Hence $f = g$.

Plainly $0 \le f \le 1$. Now observe that $U_r \subseteq \overline{U_0}$ for all *r* in $\mathbb{Q} \cap [0,1]$. Therefore, $f(x) \ne 0 \Rightarrow x \in \overline{U_0}$ and it follows that supp $f = \overline{\{x : f(x) \ne 0\}} \subseteq \overline{U_0} \subseteq U$ and $\overline{U_0}$ is compact and so the support of f is compact. Hence $f \prec U$. As $K \subseteq U_r$, $f(x) = 1$ for all *x* in *K*. Therefore, $K \prec f$. It follows that $K \prec f \prec U$.

5. Proof of Lusin's Theorem.

For convenience, we state the theorem here.

Suppose X is a locally compact Hausdorff topological space and (X, $\mathcal{M}$, μ) is a measure space satisfying the following five properties:

(1) $\mathcal{M}$ is a σ algebra containing all the Borel subsets of X.

(2) μ is a positive measure on $\mathcal{M}$ satisfying:

(3) For all compact $K \subseteq X$, $\mu(K) < \infty$.

(4) For all $E \in \mathcal{M}$, $\mu(E) = \inf\{\mu(V) : V \supseteq E \text{ and } V \text{ is open}\}$ (Outer Regularity).

(5) For all $E \in \mathcal{M}$ such that either E is open or $\mu(E) < \infty$,

$\mu(E) = \sup\{\mu(K) : K \subseteq E \text{ and } K \text{ is compact}\}$ (Inner Regularity).

Suppose $f : X \to \mathbb{C}$ is a measurable function such that $\{x \in X : f(x) \neq 0\} \subseteq A$ and $A \in \mathcal{M}$, where $\mu(A) < \infty$. Then for any $\varepsilon > 0$, there exists $g \in C_c(X)$ such that

$$\mu\{x \in X : g(x) \neq f(x)\} < \varepsilon .$$

Moreover, we may arrange it so that $\|g\| = \sup_{x\in X}|g(x)| \leq \sup_{x\in X}|f(x)|$.

We prove first for the special case:

(1) *A* is compact and $0 \leq f < 1$.

Since f is non-negative and measurable, there exists an increasing sequence of measurable simple functions (s_n) converging pointwise to f .

We can construct the sequence (s_n) as follows, in a similar way for any non-negative measurable function. Then we shall specialize to a non-negative function whose values are strictly less than 1.

For each integer $n \geq 1$, divide the interval [0, n] into $n \times 2^n$ sub-intervals of length $\frac{1}{2^n}$.

Let $E_{n,i} = f^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)\right) \in \mathcal{M}$, $i = 1, 2, \cdots, n2^n$,

$$F_n = f^{-1}([n,\infty))$$

and $s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{n,i}} + n\chi_{F_n}$.

Since f is measurable, the sets $E_{n,i}$ and F_n are measurable.

Note that $E_{n,i} = E_{n+1,j} \cup E_{n+1,j+1}$, where $\frac{j-1}{2^{n+1}} = \frac{i-1}{2^n}$ or $j = 2i-1$. On the set $E_{n,i}$, $s_{n+1}(x)$ takes on the value $\frac{j-1}{2^{n+1}} = \frac{i-1}{2^n}$ when *x* is in $E_{n+1,j}$ and the value $\frac{j}{2^{n+1}} > \frac{i-1}{2^n}$ when *x* is in $E_{n+1,j+1}$. Observe also that

$$F_n = f^{-1}([n,\infty)) = f^{-1}([n+1,\infty)) \cup f^{-1}([n,n+1)) = F_{n+1} \cup f^{-1}([n,n+1))$$

and $f^{-1}([n,n+1)) = \bigcup\{E_{n+1,i} : i = n2^{n+1}+1 \text{ to } (n+1)2^{n+1}\}$.

Thus on the set F_{n+1} , $s_{n+1}(x)$ takes on the value $n+1$ when *x* is in $E_{n+1,j}$ and on the set $f^{-1}([n,n+1))$, $s_{n+1}(x)$ takes on values $\geq n$. Therefore, $s_{n+1} \geq s_n$.

Since $f(x) < \infty$, take an integer N such that $N > f(x)$, then for all $n \geq N$, $s_{n+1}(x) \leq N$ and so the sequence is pointwise convergence. Moreover, for each integer $n > f(x)$, $f(x)$ lies in $\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)$ for some i such that $1 \leq i \leq n2^n$ and so $s_n(x) \leq f(x)$. Furthermore, $s_n(x) \geq f(x) - \frac{1}{2^n}$. Hence $\lim_{n\to\infty} s_n(x) = f(x)$.

Let $t_1 = s_1$, $t_n = s_n - s_{n-1}$ for $n \geq 2$.

Then $f = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{n} t_k = \sum_{k=1}^{\infty} t_k$.

Now we specialize to the function f such that $0 \leq f < 1$.

For $0 \leq f < 1$, we investigate t_n and approximate it by a continuous function h_n so that $\sum_{k=1}^{\infty} h_k$ converges uniformly.

First of all, note that $F_n = \varnothing$ for all integer $n \geq 1$. For any integer $n > 1$,

$$E_{n,i} = f^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)\right) = \varnothing \quad \text{if} \quad 2^n + 1 \leq i \leq n2^n .$$

This means for $0 \leq f < 1$, we partition the interval [0,1] into 2^n sub-intervals each of length $\frac{1}{2^n}$.

For n =1, $t_1 = s_1 = \sum_{i=1}^{2} \frac{i-1}{2} \chi_{E_{1,i}} = \frac{1}{2} \chi_{E_{1,2}}$ since $F_1 = \varnothing$.

For $n \geq 2$, $t_n = s_n - s_{n-1} = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{n,i}} - \sum_{i=1}^{(n-1)2^{n-1}} \frac{i-1}{2^{n-1}} \chi_{E_{n-1,i}}$ since $F_n = \varnothing$ for all integer $n \geq 1$.

Note that $E_{n-1,i} = E_{n,j} \cup E_{n,j+1}$ where $\frac{j-1}{2^n} = \frac{i-1}{2^{n-1}}$ or $j = 2i - 1$.

Therefore, for $n \geq 2$, $s_{n-1} = \sum_{i=1}^{(n-1)2^{n-1}} \frac{i-1}{2^{n-1}} \chi_{E_{n-1,i}} = \sum_{i=1}^{(n-1)2^{n-1}} \frac{i-1}{2^{n-1}} (\chi_{E_{n,2i-1}} + \chi_{E_{n,2i}})$.

Hence, for $n \geq 2$,

$$t_n = s_n - s_{n-1} = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{n,i}} - \sum_{i=1}^{(n-1)2^{n-1}} \frac{i-1}{2^{n-1}} (\chi_{E_{n,2i-1}} + \chi_{E_{n,2i}})$$

$$= \sum_{i=1}^{n2^{n-1}} \frac{2i-2}{2^n} \chi_{E_{n,2i-1}} + \sum_{i=1}^{n2^{n-1}} \frac{2i-1}{2^n} \chi_{E_{n,2i}} - \sum_{i=1}^{(n-1)2^{n-1}} \frac{i-1}{2^{n-1}} \chi_{E_{n,2i-1}} - \sum_{i=1}^{(n-1)2^{n-1}} \frac{i-1}{2^{n-1}} \chi_{E_{n,2i}}$$

$$= \sum_{i=1}^{(n-1)2^{n-1}} \frac{1}{2^n} \chi_{E_{n,2i}} + \sum_{i=(n-1)2^{n-1}+1}^{n2^{n-1}} \frac{2i-1}{2^n} \chi_{E_{n,2i}} + \sum_{i=(n-1)2^{n-1}+1}^{n2^{n-1}} \frac{2i-2}{2^n} \chi_{E_{n,2i-1}}$$

$$= \sum_{i=1}^{(n-1)2^{n-1}} \frac{1}{2^n} \chi_{E_{n,2i}} + \sum_{i=(n-1)2^n+1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{n,i}}$$

But $E_{n,i} = f^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)\right) = \varnothing$ if $2^n + 1 \leq i \leq n2^n$ and as $(n-1)2^n + 1 \geq 2^n + 1$, we have that

$$t_n = \sum_{i=1}^{2^{n-1}} \frac{1}{2^n} \chi_{E_{n,2i}} \quad \text{for } n \geq 2.$$

Therefore, for $n \geq 2$, $2^n t_n = \sum_{i=1}^{2^{n-1}} \chi_{E_{n,2i}}$. Note that $2t_1 = \chi_{E_{1,2}}$.

If we let $T_n = \bigcup_{i=1}^{2^{n-1}} E_{n,2i}$ for integer $n \geq 1$, then T_n is measurable and $2^n t_n = \chi_{T_n}$.

Note that if $E_{n,2i} \neq \varnothing$, then $0 \notin f(E_{n,2i})$ for $1 \leq i \leq 2^{n-1}$ and so for all x in T_n , $f(x) \neq 0$.

Therefore, $T_n \subseteq A$. Since X is Hausdorff, A is closed. Then by Property 10 (Topological Spaces), since X is locally compact and Hausdorff, there exists a relatively compact open set V such that $A \subseteq V \subseteq \overline{V} \subseteq X$ and $\overline{V}$ is compact. Hence, $T_n \subseteq V$.

Since $\mu(A) < \infty$, $\mu(T_n) < \infty$. Thus, by the outer regularity of μ (Property (4)), $\mu(T_n) = \inf\{\mu(U) : U \supseteq T_n \text{ and } U \text{ is open}\}$. Hence given any $\varepsilon > 0$, there is a measurable open set V_n such that $V_n \supseteq T_n$ and $\mu(V_n) < \mu(T_n) + \frac{\varepsilon}{2^{n+1}}$. Furthermore, we may choose $V_n \subseteq V$. If V_n is not contained in V, we may replace V_n by $V_n \cap V$ and rename it as V_n .

By the inner regularity of μ (Property (5)), $\mu(T_n) = \sup\{\mu(K) : K \subseteq T_n \text{ and } K \text{ is compact}\}$. Hence, there is a compact set K_n such that $K_n \subseteq T_n$ and $\mu(K_n) > \mu(T_n) - \frac{\varepsilon}{2^{n+1}}$. Thus, we have $K_n \subseteq T_n \subseteq V_n \subseteq V$. Now $\mu(V_n) = \mu(V_n - K_n) + \mu(K_n)$ and so

$$\mu(V_n - K_n) = \mu(V_n) - \mu(K_n) < \mu(T_n) + \frac{\varepsilon}{2^{n+1}} - \left(\mu(T_n) - \frac{\varepsilon}{2^{n+1}}\right) = \frac{\varepsilon}{2^n} . \quad \text{---------} \ (1)$$

By Urysohn's Lemma, there exists $h_n \in C_c(X)$ such that $K_n \prec h_n \prec V_n$.

Let $g = \sum_{n=1}^{\infty} \frac{1}{2^n} h_n$. Since $0 \leq \frac{1}{2^n} h_n \leq \frac{1}{2^n}$ and $\sum_{n=1}^{\infty} \frac{1}{2^n} < \infty$, by the Weierstrass M-Test, $\sum_{n=1}^{\infty} \frac{1}{2^n} h_n$ is uniformly convergent. As each h_n is continuous, g is continuous. Since supp h_n $\subseteq V_n \subseteq V$ and $h_n \geq 0$, supp g $\subseteq V \subseteq \overline{V}$. Since $\overline{V}$ is compact, supp g is compact and so $g \in C_c(X)$.

Note that if $x \notin V_n - K_n$, $\frac{1}{2^n} h_n(x) = t_n(x)$. This is because $2^n t_n = \chi_{T_n}$ and for $x \notin V_n - K_n$, $\chi_{T_n}(x) = 1 = h_n(x)$, if $x \in K_n$ and $\chi_{T_n}(x) = 0 = h_n(x)$ if $x \in (V_n)^c$. Therefore,

$$\sum_{n=1}^{\infty} t_n(x) = g(x) \text{ except possibly for } x \text{ in } \bigcup_{n=1}^{\infty}(V_n - K_n) \in \mathscr{M} .$$

Now $\mu\left(\bigcup_{n=1}^{\infty}(V_n - K_n)\right) \le \sum_{n=1}^{\infty} \mu(V_n - K_n) < \sum_{n=1}^{\infty} \frac{1}{2^n}\varepsilon = \varepsilon$ and so $g(x) = \sum_{n=1}^{\infty} t_n(x) = f(x)$ except possibly for x in $\bigcup_{n=1}^{\infty}(V_n - K_n)$ of measure $< \varepsilon$. I.e., $\mu\{x \in X : g(x) \ne f(x)\} < \varepsilon$.

This proves the case for A compact and $0 \le f < 1$.

(2) Suppose A is compact and f takes on non-negative real values and is bounded, say by M. That is, $0 \le f < M$. Then applying (1) to $\frac{1}{M} f$, we have that, for any $\varepsilon > 0$, there exists $h \in C_c(X)$ such that $\mu\left\{x \in X : h(x) \ne \frac{1}{M} f(x)\right\} < \varepsilon$. Then let $g = M h$.

(3) Suppose A is not compact and $0 \le f < M$.

Since $\mu(A) < \infty$, by the inner regularity of μ, given any $\varepsilon > 0$, there exists compact K such that $K \subseteq A$, $\mu(K) > \mu(A) - \frac{\varepsilon}{2}$ so that $\mu(A - K) < \frac{\varepsilon}{2}$.

Consider $\chi_K f$. Then $\chi_K f$ is measurable, $0 \le \chi_K f < M$ and $\{x \in X : \chi_K f(x) \ne 0\} \subseteq K$, where $\mu(K) < \infty$. Then by part (2) for any $\varepsilon > 0$, there exists $g \in C_c(X)$ such that

$$\mu\{x \in X : g(x) \ne \chi_K f(x)\} < \frac{\varepsilon}{2} .$$

Let $U = \{x \in X : g(x) \ne \chi_K f(x)\}$. Observe that $g(x) = \chi_K f(x)$ for x in U^c and $f(x) = \chi_K f(x)$ for x in $K \cup A^c$. Therefore, $g(x) = f(x)$ for x in $U^c \cap (K \cup A^c)$. Now

$$\left(U^c \cap (K \cup A^c)\right)^c = U \cup (K \cup A^c)^c = U \cup (A \cap K^c) = U \cup (A - K) .$$

Therefore, $\{x \in X : g(x) \ne f(x)\} \subseteq U \cup (A - K)$ and so

$$\mu\{x \in X : g(x) \ne f(x)\} \le \mu(U) + \mu(A - K) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon .$$

(4) Suppose f is not bounded and $f \ge 0$.

Let $B_n = \{x : f(x) > n\}$. Then $\bigcap_{n=1}^{\infty} B_n = \varnothing$. Note that $B_n \subseteq A$ and since $\mu(A) < \infty$, $\mu(B_n) < \infty$. Therefore, $\mu(B_n) \to 0$. Hence, given ε > 0, there exists an integer *N* such that $n \geq N \Rightarrow \mu(B_n) < \frac{\varepsilon}{2}$.

Take a fixed *n* such that $n \geq N$. Consider $(1-\chi_{B_n})f$. Then $0 \leq (1-\chi_{B_n})f \leq n$. Apply part (3) to $(1-\chi_{B_n})f$. Note that

$$(1-\chi_{B_n})(x)f(x) \neq 0 \Rightarrow f(x) \neq 0 \text{ and } (1-\chi_{B_n})(x) \neq 0$$

$$\Rightarrow x \in A \text{ and } x \in (B_n)^c \Rightarrow x \in A \cap (B_n)^c.$$

Hence $\{x : (1-\chi_{B_n})(x)f(x) \neq 0\} \subseteq A$ and $\mu(A) < \infty$.

This gives by part (3) for any $\varepsilon > 0$, a function $g \in C_c(X)$ such that

$$\mu\{x \in X : g(x) \neq (1-\chi_{B_n})f(x)\} < \frac{\varepsilon}{2}.$$

Let $U = \{x \in X : g(x) \neq (1-\chi_{B_n})f(x)\}$. Then $g(x) = (1-\chi_{B_n})f(x)$ for *x* in U^c . $f(x) = (1-\chi_{B_n})f(x)$ for *x* in $(B_n)^c$. Hence $f(x) = g(x)$ for *x* in $(B_n)^c \cap U^c$.

Therefore, $\{x \in X : g(x) \neq f(x)\} \subseteq \left((B_n)^c \cap U^c\right)^c = B_n \cup U$. It then follows that

$$\mu(\{x \in X : g(x) \neq f(x)\}) \leq \mu(B_n) + \mu(U) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

(5) Suppose f is real valued. Since f is measurable, $f_+ = \sup\{f, 0\}$ and $f_- = \sup\{-f, 0\}$ are measurable and $f = f_+ - f_-$. Suppose $\{x \in X : f(x) \neq 0\} \subseteq A$ and $A \in \mathscr{M}$, where $\mu(A) < \infty$. Note that $f_+(x) \neq 0 \Rightarrow f_+(x) > 0 \Rightarrow f_+(x) > 0$ and $f_-(x) = 0 \Rightarrow f(x) \neq 0$ and so $\{x \in X : f_+(x) \neq 0\} \subseteq A$. Similarly, we can show that $\{x \in X : f_-(x) \neq 0\} \subseteq A$. Now we apply part (4) to f_+ and f_- . Given ε > 0, there exists $h, k \in C_c(X)$ such that

$$\mu\{x \in X : h(x) \neq f_+(x)\} < \frac{\varepsilon}{2} \quad \text{and}$$

$$\mu\{x \in X : k(x) \neq f_-(x)\} < \frac{\varepsilon}{2}.$$

Let $U_+ = \{x \in X : h(x) \neq f_+(x)\}$ and $U_- = \{x \in X : k(x) \neq f_-(x)\}$. Therefore,

$$h(x) - k(x) = f_+(x) - f_-(x) = f(x) \text{ for } x \text{ in } (U_+)^c \cap (U_-)^c.$$

Hence $\{x : h(x) - k(x) \neq f(x)\} \subseteq \left((U_+)^c \cap (U_-)^c\right)^c = U_+ \cup U_-$ and so

$$\mu\{x : h(x) - k(x) \neq f(x)\} \leq \mu(U_+) + \mu(U_-) < \varepsilon.$$

Let $g = h - k$. Then $g \in C_c(X)$, since support of $g = h - k$ lies in the union of support of *h* and the support of *k* and the union of two compact sets is compact.

(6) Suppose f is a complex function. Suppose $\{x \in X : f(x) \neq 0\} \subseteq A$ and $A \in \mathscr{M}$, where $\mu(A) < \infty$. Since f is measurable, Re(*f*) and Im(*f*) are measurable. Note that

$$\{x \in X : f(x) \neq 0\} = \{x \in X : \text{Re} f(x) + i \,\text{Im} f(x) \neq 0\}$$

$$= \{x \in X : \text{Re} f(x) \neq 0\} \cup \{x \in X : \text{Im} f(x) \neq 0\}.$$

Therefore, $\{x \in X : \text{Re} f(x) \neq 0\}, \{x \in X : \text{Im} f(x) \neq 0\} \subseteq A$. Thus applying part (5) to Re(*f*) and Im(*f*) we get $h, k \in C_c(X)$ such that

$$\mu\{x \in X : h(x) \neq \text{Re} f(x)\} < \frac{\varepsilon}{2} \quad \text{and} \quad \mu\{x \in X : k(x) \neq \text{Im} f(x)\} < \frac{\varepsilon}{2}.$$

Let $U = \{x \in X : h(x) \neq \text{Re} f(x)\}$ and $V = \{x \in X : k(x) \neq \text{Im} f(x)\}$. Therefore,

$$h(x) + ik(x) = \text{Re} f(x) + i \,\text{Im} f(x) = f(x) \text{ for } x \text{ in } (U)^c \cap (V)^c.$$

Hence $\{x : h(x) + ik(x) \neq f(x)\} \subseteq \left((U)^c \cap (V)^c\right)^c = U \cup V$ and so

$$\mu\{x : h(x) - ik(x) \neq f(x)\} \leq \mu(U) + \mu(V) < \varepsilon.$$

So, if we let $g = h + i\,k$, then $\mu\left(\{x : g(x) \neq f(x)\}\right) < \varepsilon$.

Note that $g = h + i\,k \in C_c(X)$ as the support of *g*, being a closed subset of the union of the support of *h* and *k*, is compact.

Now we come to the last conclusion.

If $\sup_{x\in X}|f(x)|=\infty$, then we have nothing to prove.

Suppose now $\sup_{x\in X}|f(x)|=M<\infty$

Define a complex function φ by $\varphi(z)=\begin{cases} z & \text{if } |z|\le M, \\ M\dfrac{z}{|z|} & \text{if } |z|>M \end{cases}$. Then φ is a continuous function on $\mathbb{C}$ mapping $\mathbb{C}$ onto the closed disk of radius *M*. Consider the composition $\varphi\circ g$. Then $\varphi\circ g(x)=f(x)$ if $g(x)=f(x)$ and so $\varphi\circ g(x)\ne f(x)$ implies $g(x)\ne f(x)$. Hence, $\{x:\varphi\circ g(x)\ne f(x)\}\subseteq\{x:g(x)\ne f(x)\}$ and so

$$\mu\left(\{x:\varphi\circ g(x)\ne f(x)\}\right)\le\mu\left(\{x:g(x)\ne f(x)\}\right)<\varepsilon .$$

Therefore, we may replace *g* by $g_1=\varphi\circ g$ since $\varphi\circ g$ is continuous and support of $\varphi\circ g$ is contained in the support of *g*. Moreover,

$$\|\varphi\circ g\|=\sup_{x\in X}|\varphi\circ g(x)|\le\sup_{x\in X}|f(x)| .$$

Theorem 23.

Suppose *X* is a locally compact Hausdorff topological space and $(X, \mathscr{M}, \mu)$ is a measure space satisfying properties (1) to (5) in Lusin's Theorem.

Give the space of all continuous functions on *X* with compact support $C_c(X)$ the (unusual) norm $\|f\|_p$, $1\le p<\infty$, for f in $C_c(X)$. Then $C_c(X)\subseteq L^p(\mu)$ and the completion of the metric space $\left(C_c(X),\|\cdot\|_p\right)$ is $L^p(\mu)$.

Proof.

Suppose $f\in C_c(X)$. Let *K* be the support of *f*. Since *f* is continuous, $|f|$ is continuous. Therefore, $\{|f(x)|:x\in X\}=\{|f(x)|:x\in K\}$ is a compact subset of $[0,\infty)$. Hence $\{|f(x)|:x\in X\}$ is bounded since a compact subset of $\mathbb{R}$ is bounded. Suppose $\{|f(x)|:x\in X\}$ is bounded by *M*. Plainly, any continuous function on *X* is μ-measurable since $\mathscr{M}$ contains the Borel subsets of *X*. Hence, *f* is measurable. Then for any $1\le p<\infty$,

$$\int_X|f|^p\,d\mu=\int_K|f|^p\,d\mu\le\int_K M^p d\mu=M^p\mu(K)<\infty$$

since by property (3) of the measure space (X, $\mathscr{M}$, μ), $\mu(K) < \infty$ as K is compact.

Therefore, $\| f \|_p = \left(\int_X | f |^p \, d\mu \right)^{\frac{1}{p}} < \infty$. Hence $f \in L^p(\mu)$ and it follows that $C_c(X) \subseteq L^p(\mu)$.

By Proposition 16, the set

$$S = \left\{ s : X \to \mathbb{C}; s \text{ is a simple measurable function with } \mu\left(\left\{ x : s(x) \neq 0 \right\}\right) < \infty \right\}$$

is dense in $L^p(\mu)$.

Each simple measurable function in S satisfies the condition of Lusin's Theorem.

Note that if $s \in S$, then s is bounded on X. Hence $\|s\| = \sup_{x \in X} |s(x)| < \infty$. By Lusin's Theorem, given $\varepsilon > 0$, if $s \in S$, then there exists $g \in C_c(X)$ such that $g(x) = s(x)$ except on a set of measure $< \frac{\varepsilon^p}{\left(2\|s\|+1\right)^p}$, i.e., $\mu\left(\left\{ x : g(x) \neq s(x) \right\}\right) < \frac{\varepsilon^p}{\left(2\|s\|+1\right)^p}$ and $\|g\| \leq \|s\|$.

Now

$$\left(\|g-s\|_p\right)^p = \int_X |g-s|^p \, d\mu \leq \int_{\{x: g(x)=s(x)\}} |g-s|^p \, d\mu + \int_{\{x: g(x) \neq s(x)\}} |g-s|^p \, d\mu \text{ and so}$$

$$\left(\|g-s\|_p\right)^p \leq 0 + \int_{\{x: g(x) \neq s(x)\}} \left(2\|s\|\right)^p d\mu \leq \left(2\|s\|\right)^p \mu\{x : g(x) \neq s(x)\}$$

$$< \left(2\|s\|\right)^p \frac{\varepsilon^p}{\left(2\|s\|+1\right)^p} < \varepsilon^p .$$

Hence, $\|g-s\|_p < \varepsilon$.

Since $\left(S, \|\cdot\|_p\right)$ is dense in $L^p(\mu)$, for any f in $L^p(\mu)$ and for any ε > 0, there exists $s \in S$ such that $\|s - f\|_p < \frac{\varepsilon}{2}$. By what we have just proved above using $\frac{\varepsilon}{2}$ instead of ε, there exists $g \in C_c(X)$ such that $\|g-s\|_p < \frac{\varepsilon}{2}$. Therefore,

$$\|g - f\|_p = \|g - s + s - f\|_p \leq \|g-s\|_p + \|s - f\|_p < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon .$$

It follows that $\left(C_c(X), \|\cdot\|_p\right)$ is dense in $L^p(\mu)$. Hence the completion of $\left(C_c(X), \|\cdot\|_p\right)$ is $L^p(\mu)$.

Remark.

The unit interval [0, 1] is compact and Hausdorff and so is locally compact. Hence the space of all continuous function on [0, 1], *C*[0,1], is the same as the space of all continuous function with compact support, $C_c[0,1]$. By Theorem 23, the completion of $C([0,1])$ with respect to the metric, $d(f,g)=\int_0^1|f-g|$, for f and g in $C([0,1])$, where the integral is the Lebesgue integral, is $L^1([0,1])$ and the measure is the Lebesgue measure on $\mathbb{R}$. Note that every continuous function on [0, 1] is Riemann integrable and is also Lebesgue integrable and both integrals are the same. $\left(C[0,1],\|\cdot\|_1\right)$ is a normed linear space and $\left(C[0,1],d\right)$ is a metric space because for any $f\in C[0,1]$, $\|f\|_1=0 \Leftrightarrow \int_0^1|f|d\mu=0 \Leftrightarrow f=0$ since a continuous non-negative function, whose integral over the interval [0, 1] is 0, must be the zero constant function. But the completion $L^1([0,1])$ is the equivalence classes of almost everywhere equal Lebesgue integrable functions on [0,1]. The norm on $\left(C[0,1],\|\cdot\|_1\right)$ is of course given by the Riemann integral $\|f\|_1=\int_0^1|f(x)|dx$ since it is the same as the Lebesgue integral but the norm on the completion is given by the Lebesgue integral. In some sense, we may regard the Lebesgue integral as a natural generalization of the Riemann integral.

For locally compact Hausdorff space *X* and for $(X, \mathcal{M}, \mu)$ satisfying properties (1) to (5) of Lusin's Theorem, although the underlying linear space $C_c(X)$ is the same, for different L^p norm, the completion of the space of all continuous function with compact support is in general different L^p space.

6. Space of Continuous function.

Suppose *X* is a locally compact Hausdorff topological space. Let $C(X)$ be the space of all continuous complex functions on *X*. A function f in $C(X)$ is said to *vanish at infinity* if for any $\varepsilon > 0$, there exists a compact set *K* in *X* such that $|f(x)|<\varepsilon$ for all $x\notin K$. Let $C_0(X)$ be the collection of continuous functions on *X* that vanish at infinity.

If f vanishes at infinity, we may say f is zero at infinity if we attach a point called 'infinity' to *X*. More precisely, if we take the one-point compactification $\tilde{X}=X\cup\{\infty\}$ of *X*. The open sets in $\tilde{X}$ are given by $\left\{\text{open sets in } X,\ K^c\cup\{\infty\}, K \text{ compact}\right\}$ and all those sets generated by this class. Then $C_0(X)$ is the ideal of functions in $C(\tilde{X})$ consisting of functions that are zero at infinity.

Let $BC(X)$ be the set of all bounded continuous complex functions on *X*. Then the sup norm, $\|f\|=\sup\left\{|f(x)|: x\in X\right\}$ is a norm on $BC(X)$ and moreover $\|f\|<\infty$ for all f in $BC(X)$.

Then we have:

Proposition 24.

(1) $C_c(X) \subseteq C_0(X) \subseteq BC(X) \subseteq C(X)$;

(2) If X is compact, then $C_c(X) = C_0(X) = BC(X) = C(X)$.

Proof.

Plainly $C_c(X) \subseteq C_0(X)$. If $f \in C_0(X)$, for any $\varepsilon > 0$, there exists a compact set K in X such that $|f(x)| < \varepsilon$ for all $x \notin K$. Since f is continuous, $f(K)$ is compact and so f is bounded say by M on K. Hence f is bounded by max$\{\varepsilon, M\}$. Therefore, $f \in BC(X)$. This proves (1)

Suppose X is compact. Take any $f \in C(X)$. Then the support of f is closed in X. Since X is compact, the support of f is compact. Hence $f \in C_c(X)$. Thus $C_c(X) = C(X)$ and (2) follows.

Proposition 25. Suppose X is a locally compact Hausdorff topological space.

(1) $C_0(X)$ and $BC(X)$ with the sup norm, $\|f\| = \sup_{x \in X} |f(x)|$, are Banach spaces.

(2) $\left(C_c(X), \|\cdot\|\right)$, with the sup norm $\|\cdot\|$, is a normed space, not usually complete.

(3) $\left(C_0(X), \|\cdot\|\right)$ is the completion of $\left(C_c(X), \|\cdot\|\right)$.

Proof:

(1) It is obvious that $\left(C_0(X), \|\cdot\|\right)$ is a normed space. Let (f_n) be a Cauchy sequence in $\left(C_0(X), \|\cdot\|\right)$. Then given any $\varepsilon > 0$, there exists an integer N such that

$$n, m \geq N \Rightarrow \|f_n - f_m\| = \sup_{x \in X} |f_n(x) - f_m(x)| < \frac{\varepsilon}{3} \text{ and so}$$

$$n, m \geq N \Rightarrow |f_n(x) - f_m(x)| \leq \sup_{x \in X} |f_n(x) - f_m(x)| = \|f_n - f_m\| < \frac{\varepsilon}{3} \text{ for all } x \text{ in } X.$$

This means $(f_n(x))$ is uniformly Cauchy. Since $\mathbb{C}$ is complete, $(f_n(x))$ converges uniformly. We claim that the pointwise limit, f, of (f_n) is a continuous function.

Note that for any $n \geq N$ and for all x in X,

$$|f_n(x) - f(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leq \lim_{m\to\infty} \|f_n - f_m\| \leq \frac{\varepsilon}{3} \quad \text{---------} \ (1)$$

Take x_0 in X.

For all x in X,

$$|f(x) - f(x_0)| = |f(x) - f_N(x) + f_N(x) - f_N(x_0) + f_N(x_0) - f(x_0)|$$

$$\leq |f(x) - f_N(x)| + |f_N(x_0) - f(x_0)| + |f_N(x) - f_N(x_0)|$$

$$\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + |f_N(x) - f_N(x_0)| = \frac{2\varepsilon}{3} + |f_N(x) - f_N(x_0)| . \quad \text{------------------} \ (2)$$

Since f_N is continuous at x_0, there exists an open neighbourhood U of x_0 such that $x \in U \Rightarrow |f_N(x) - f_N(x_0)| < \frac{\varepsilon}{3}$. It then follows from (2) that

$$x \in U \Rightarrow |f(x) - f(x_0)| < \varepsilon .$$

This means f is continuous at x_0. Hence, f is continuous on X.

Now we show that f vanishes at infinity. By (1), given $\varepsilon > 0$, we can find an integer N such that $n \geq N \Rightarrow |f_n(x) - f(x)| \leq \frac{\varepsilon}{3} < \frac{\varepsilon}{2}$ for all x in X. This means

$$n \geq N \Rightarrow \|f_n - f\| < \frac{\varepsilon}{2} .$$

Since f_N vanishes at infinity, there exists a compact set K such that

$$x \notin K \Rightarrow |f_N(x)| < \frac{\varepsilon}{2} .$$

Therefore, for $x \notin K$,

$$|f(x)| = |f(x) - f_N(x) + f_N(x)| \leq |f(x) - f_N(x)| + |f_N(x)| \leq \|f - f_N\| + |f_N(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon .$$

Hence, f vanishes at infinity and so $f \in C_0(X)$. Thus, any Cauchy sequence in $\left(C_0(X), \|\cdot\|\right)$ converges to a function in $\left(C_0(X), \|\cdot\|\right)$ and so $\left(C_0(X), \|\cdot\|\right)$ is complete and hence is a Banach space.

Next $\left(BC(X),\|\cdot\|\right)$ is a Banach space. Plainly it is a normed linear space.

Let (f_n) be a Cauchy sequence in $\left(BC(X),\|\cdot\|\right)$. We have already shown that $(f_n(x))$ is uniformly Cauchy and that $(f_n(x))$ converges uniformly to a continuous function f. We claim that f is bounded.

As $(f_n(x))$ converges uniformly to f given $\varepsilon > 0$, we can find in integer N such that $n \geq N \Rightarrow \| f_n - f \| < \varepsilon$.

Take $n = N$. Since f_N is bounded, $|f_N(x)| \leq M$ for some $M > 0$ and for all x in X.

Therefore, for all $x \in X$,

$$|f(x)| = |f(x) - f_N(x) + f_N(x)| \leq |f(x) - f_N(x)| + |f_N(x)| \leq \| f - f_n \| + M < M + \varepsilon \ .$$

Hence f is bounded by $M + \varepsilon$ and so $f \in BC(X)$. Thus $\left(BC(X),\|\cdot\|\right)$ is a Banach space.

(2) Plainly $\left(C_c(X),\|\cdot\|\right)$ is a normed linear space. $\left(C_c(X),\|\cdot\|\right)$ is usually not complete.

More precisely, if X is locally compact and Hausdorff but not compact and can be written as a strictly increasing sequence of relatively compact open sets, then $\left(C_c(X),\|\cdot\|\right)$ is not complete.

Suppose $U_1 \subseteq \overline{U_1} \subseteq U_2 \subseteq \overline{U_2} \subseteq \cdots U_n \subseteq \overline{U_n} \subseteq \cdots$

is a strictly increasing sequence of relatively compact sets such that $\bigcup_{n=1}^{\infty} U_n = X$.

By Urysohn's Lemma (Lemma 22), since $\overline{U_n} \subseteq U_{n+1}$ and $\overline{U_n}$ is compact, there exists $f_n \in C_c(X)$ such that $\overline{U_n} \prec f_n \prec U_{n+1}$ and so we have a sequence of functions (f_n) such that

(a) $f_n|_{\overline{U_n}} = 1$

(b) Support $f_n \subseteq U_{n+1}$ and

(c) $0 \leq f \leq 1$.

Let $g_n = \sum_{k=1}^{n} \frac{1}{k^2} f_k$. Then by the Weierstrass M Test, (g_n) converges uniformly to a continuous function g on X. Obviously (g_n) is a Cauchy sequence in $\left(C_c(X),\|\cdot\|\right)$. But

$g(x)=\sum_{k=1}^{\infty}\frac{1}{k^2}f_k(x)\neq 0$ for all *x*. This is because for any *x* in *X*, $x\in U_n$ for some *n* and

$f_n(x)=1$ so that $g(x)=\sum_{k=1}^{\infty}\frac{1}{k^2}f_k(x)\geq\frac{1}{n^2}f_n(x)=\frac{1}{n^2}>0$. Thus, support of *g* is *X*, which is not compact. Hence, $g\notin C_c(X)$ and so $\left(C_c(X),\|\cdot\|\right)$ is not complete.

For instance, $\left(C_c(\mathbb{R}),\|\cdot\|\right)$ is not complete. Take $U_n=(-n,n)$. We can define

$$f_n(x)=\begin{cases}1\text{ if } -n\leq x\leq n,\\ 2(n+\frac{1}{2}-x)\text{ if } n\leq x\leq n+\frac{1}{2},\\ 2\left(x+n+\frac{1}{2}\right)\text{ if } -n-\frac{1}{2}\leq x\leq -n,\\ 0\ \text{ otherwise.}\end{cases}.$$

Then $g(x)=\sum_{k=1}^{\infty}\frac{1}{k^2}f_k(x)$ is continuous but not in $C_c(\mathbb{R})$.

(3) $C_0(X)$ is the completion of $C_c(X)$.

Take any $g\in C_0(X)$. Then given any $\varepsilon > 0$, there exists compact set $K\subseteq X$ such that $|g(x)|<\frac{\varepsilon}{2}$ for all *x* not in *K*. Then by Property (10) (topological spaces), since *X* is locally compact and Hausdorff, there is a relatively compact open set, i.e., an open set *V*, where $\overline{V}$ is compact, such that $K\subseteq V\subseteq\overline{V}\subseteq X$.

By Urysohn's Lemma, there exists $f\in C_c(X)$, such that $K\prec f\prec V$, i.e., $0\leq f\leq 1$ and *f* (*x*) =1 for all *x* in *K* and the support of $f\subseteq V$. Since $\overline{V}$ is compact, the support of f is also compact. Therefore, $g\,f$ is continuous and the support of $g\,f$ is also compact, being a closed subset of the compact support of *f*. This means $g\,f\in C_c(X)$.

Now $g(x)-g(x)f(x)=0$ if *x* is in *K*. Thus since $|f(x)|\leq 1$,

$$|g(x)-g(x)f(x)|\leq 2|g(x)|<\varepsilon \text{ for all } x\in X.$$

Therefore, $\|gf-g\|=\sup_{x\in X}|g(x)-g(x)f(x)|\leq\varepsilon$. This means $C_c(X)$ is dense in $C_0(X)$ and so the completion of $\left(C_c(X),\|\cdot\|\right)$ is $\left(C_0(X),\|\cdot\|\right)$.

Remark:

Suppose X is a locally compact Hausdorff topological space and $(X, \mathcal{M}, \mu)$ is a measure space.

Then $L^\infty(\mu) = \{ f : X \to \mathbb{C} : f$ is measurable and $\operatorname{ess\,sup}|f| < \infty \}$. I.e., $L^\infty(\mu)$ is the space of essentially bounded measurable functions and the norm on $L^\infty(\mu)$ is given by $\|f\|_\infty = \operatorname{ess\,sup}|f|$ for f in $L^\infty(\mu)$. Theorem 15 asserts that $\left(L^\infty(\mu), \|\cdot\|_\infty\right)$ or more precisely, the equivalence classes of almost everywhere equal measurable functions in $L^\infty(\mu)$ is a Banach space and so $L^\infty(\mu)$ is complete.

Suppose X is locally compact and Hausdorff and $(X, \mathcal{M}, \mu)$ is a measure space where $\mathcal{M}$ contains all the Borel subsets of X. Suppose further that every non-empty open set in X has positive μ measure. Then every continuous function $f : X \to \mathbb{C}$ is measurable and for f in $C_0(X)$, $\|f\|_\infty = \operatorname{ess\,sup}|f| = \sup|f| = \|f\| < \infty$. Hence $\left(C_0(X), \|\cdot\|\right)$ is a subspace of $\left(L^\infty(\mu), \|\cdot\|_\infty\right)$. More precisely, $\left(C_0(X), \|\cdot\|\right) = \left(C_0(X), \|\cdot\|_\infty\right)$. We know $\left(C_0(X), \|\cdot\|_\infty\right)$ is complete by Proposition 25 part (1). Similarly, $\left(C_c(X), \|\cdot\|\right) = \left(C_c(X), \|\cdot\|_\infty\right)$. We have just shown that the completion of $\left(C_c(X), \|\cdot\|\right)$ is $\left(C_0(X), \|\cdot\|\right) = \left(C_0(X), \|\cdot\|_\infty\right)$ and so the completion of $\left(C_c(X), \|\cdot\|_\infty\right)$ is $\left(C_0(X), \|\cdot\|_\infty\right)$. Even though $\left(C_c(X), \|\cdot\|_\infty\right)$ is a subspace of $L^\infty(\mu)$ and $\left(L^\infty(\mu), \|\cdot\|_\infty\right)$ is complete, the completion of $\left(C_c(X), \|\cdot\|_\infty\right)$ in the metric induced by the essential supremum norm generally is not $L^\infty(\mu)$. More precisely, $L^\infty(\mu)$ may have measurable function f that is essentially bounded but not bounded or not continuous, i.e., $\|f\|_\infty < \infty$ but $\|f\| = \sup|f| = \infty$ or f is not continuous on X. An example is the Lebesgue measure space $(\mathbb{R}^n, \mathcal{M}, \mu)$, where $\mathcal{M}$ is the σ algebra of Lebesgue measurable subsets of $\mathbb{R}^n$, and $\mu : \mathcal{M} \to [0, \infty]$ is the Lebesgue measure on $\mathbb{R}^n$. The completion of $\left(C_c(\mathbb{R}^n), \|\cdot\|_\infty\right)$ is $\left(C_0(\mathbb{R}^n), \|\cdot\|_\infty\right)$ and not $L^\infty(\mathbb{R}^n)$. For $n = 1$, the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \begin{cases} x, & x \text{ is rational,} \\ 1, & x \text{ is irrational} \end{cases}$ is measurable, essentially bounded by 1 since $\|f\|_\infty = 1$ but plainly unbounded and also not continuous. If there is a sequence (g_n) in $\left(C_c(\mathbb{R}), \|\cdot\|_\infty\right)$ that tends to f in $L^\infty(\mathbb{R})$ then (g_n) is a Cauchy sequence in $L^\infty(\mathbb{R})$ and so is a Cauchy sequence in $\left(C_c(\mathbb{R}), \|\cdot\|_\infty\right)$, which is identically the same as $\left(C_c(\mathbb{R}), \|\cdot\|\right)$ and hence a Cauchy sequence in $\left(C_0(\mathbb{R}), \|\cdot\|_\infty\right)$. Thus its pointwise limit must be continuous and bounded. This contradicts that f is not bounded. Thus there does not exist a sequence in $\left(C_c(\mathbb{R}), \|\cdot\|_\infty\right)$ that tends to f in $L^\infty(\mathbb{R})$.

Remark.

We may replace all the results stated involving a locally compact Hausdorff topological space *X* by an open or closed non-empty subspace of *X or* non-empty intersection of a closed set and an open set, since such sets are also locally compact and Hausdorff and if measure is also involved, the measure space (X, $\mathcal{M}$, μ) is required to satisfy conditions (1) to (5) in Lusin's Theorem. For instance if X is $\mathbb{R}^n$, the measure space ($\mathbb{R}^n$, $\mathcal{M}$, μ) may be taken to be the Lebesgue measure on $\mathbb{R}^n$, which satisfies all the conditions (1) to (5) in Lusin's Theorem. Furthermore, the Lebesgue measure on $\mathbb{R}^n$ is complete.

Suppose *X* is a locally compact Hausdorff topological space and $\mathcal{B}(X)$ is the σ-algebra generated by the open sets of *X*. A Borel measure, i.e., a positive measure $\mu : \mathcal{B}(X) \to [0, \infty]$ on $\mathcal{B}(X)$, satisfying conditions (3), (4) and (5) in Lusin's Theorem but not necessary μ-complete is a Radon measure. We extend this Radon measure to a complete measure $\mathcal{M}$, that is, $\mathcal{M}$ is the collection of subsets *B* of *X* of the form $B = E \cup A$ where $E \in \mathcal{B}(X)$ and $A \subseteq C$ for some subset $C \in \mathcal{B}(X)$ with $\mu(C) = 0$. For such a set *B* define $\overline{\mu}(B) = \mu(E)$. Then ($X$, $\mathcal{M}$, $\overline{\mu}$) is a complete measure space satisfying all 5 conditions in Lusin's Theorem. For such measure, Theorem 23 holds true. Such a measure space, generalizes the Lebesgue measure on $\mathbb{R}^n$.