MATH 72051 & 2: Functions of Real Variables
Kent State University

Swapnamoy Kader
Fall 2025 - Spring 2026

This entire document is based on the coursework during the semester. This document is strictly meant to help students learn, recap and prepare for Abstract Measure theory.

Note*: This document is regularly updated to include corrections and new facts.

List of Contents

Continued...

Continued...

Continued...

Continued...

Definition 1 (Metric Space)**.** Let X be any nonempty set. A function $d : X \times X \to \mathbb{R}$ is called a metric on X if:

(i) $d(x, y) \geq 0$ for all $x, y \in X$, and $d(x, y) = 0$ if and only if $x = y$.

(ii) $d(x, y) = d(y, x)$ for all $x, y \in X$.

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$.

Then (X, d) is called a metric space equipped with the metric d.

Example 1
Let $X = \mathbb{R}$ and the metric be $d(x, y) = |x - y|$. Then $(\mathbb{R}, d)$ is a metric space.

Example 2
Let $X = \mathbb{R}^n$ and the metric be $d(x, y) = \max\limits_{1 \leq i \leq n} |x_i - y_i|$. Then $(\mathbb{R}^n, d)$ is a metric space.

Example 3
Let $X = C([0, 1])$ be the set of all continuous functions on $[0, 1]$ and define

$$d(f, g) = \int_0^1 |f(t) - g(t)|\, dt$$

Then (X, d) is a metric space.

Example 4 (Discrete Metric Space)
Let X be a nonempty set and define

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

Then (X, d) is a metric space, called discrete metric space. The constant 1 may be replaced by any positive constant.

Definition 2 (Open Ball)**.** Let (X, d) be a metric space. Let $x \in X$. Then an open ball centered at x with radius $r > 0$ is defined as:

$$B_r(x) = \{y \in X : d(x, y) < r\}$$

Definition 3 (Closed Ball)**.** Let (X, d) be a metric space. Let $x \in X$. Then a closed ball centered at x with radius $r > 0$ is defined as:

$$\overline{B}_r(x) = \{y \in X : d(x, y) \leq r\}$$

Definition 4 (Open Set)**.** Let (X, d) be a metric space. Let $U \subseteq X$. Then the set U is said to be open if, for all $x \in U$, there exists $r > 0$, such that, $B_r(x) \subseteq U$.

Theorem 1 (Open Set Properties)**.** *Let (X, d) be a metric space.*

(i) The empty set $\varnothing$ and the whole space X are open.

(ii) The arbitrary union of a collection of open sets is open.

(iii) The finite intersection of open sets is open.

Proof. We prove by using the definition of open set (Definition 4).

(i) The statement is vacuously true, because, the empty set contains no points, so the condition in the definition is automatically satisfied. Also, for every $x \in X$ and every $r > 0$, we have $B_r(x) \subseteq X$. Hence both $\varnothing$ and X are open.

(ii) Let $\{U_\alpha\}_{\alpha \in J}$ be an arbitrary collection of open sets in (X, d).
Let $U = \bigcup\limits_{\alpha \in J} U_\alpha$.
Take any point $x \in U$. By the definition of a union, $x \in U_{\alpha_0}$ for some specific $\alpha_0 \in J$. Because U_{α_0} is an open set, there exists an $r > 0$ such that the open ball satisfies:

$$B_r(x) \subseteq U_{\alpha_0}$$

Since $U_{\alpha_0} \subseteq \bigcup\limits_{\alpha \in J} U_\alpha = U$, it directly follows that:

$$B_r(x) \subseteq U$$

As this holds for every $x \in U$, the union U is an open set.

(iii) Let $\{U_1, U_2, \ldots, U_n\}$ be a finite collection of open sets in (X, d).
Let $V = \bigcap\limits_{i=1}^{n} U_i$.
Take any point $x \in V$. By the definition of an intersection, $x \in U_i$ for all $i \in \{1, 2, \ldots, n\}$. Since each individual U_i is open, there exist radii $r_i > 0$ such that:

$$B_{r_i}(x) \subseteq U_i \quad \text{for each } i = 1, 2, \ldots, n$$

Define $r = \min\{r_1, r_2, \ldots, r_n\}$. Because the collection is finite, the minimum of these strictly positive numbers is also strictly positive ($r > 0$).
For this chosen r, the open ball satisfies $B_r(x) \subseteq B_{r_i}(x)$ for every i. Therefore:

$$B_r(x) \subseteq U_i \quad \text{for all } i = 1, 2, \ldots, n$$

This implies that the ball is contained within the intersection:

$$B_r(x) \subseteq \bigcap_{i=1}^{n} U_i = V$$

As this holds for every $x \in V$, the intersection V is an open set.

$\square$

Theorem 2 (Openness of an Open Ball)**.** *An open ball is an open set.*

Proof. Let (X, d) be a metric space, and Let $B_r(x) = \{y \in X : d(x, y) < r\}$ be an open ball with center $x \in X$ and radius $r > 0$.
To prove that $B_r(x)$ is an open set, we must show that every point $y \in B_r(x)$ is an interior point. That is, for any $y \in B_r(x)$, there exists a radius $R > 0$ such that the open ball $B_R(y)$ is entirely contained within $B_r(x)$.
Let y be an arbitrary point in $B_r(x)$. By definition, the distance between x and y satisfies:

$$d(x, y) < r$$

We define the radius R for the neighborhood around y as:

$$R = r - d(x, y) > 0$$

Let z be any point in $B_R(y)$. By definition:

$$d(y, z) < R$$

Then the distance $d(x, z)$ using the triangle inequality:

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + R$$

Now, substitute $R = r - d(x, y)$ into the inequality:

$$d(x, z) < d(x, y) + (r - d(x, y)) = r$$

Since $d(x, z) < r$, the point z belongs to $B_r(x)$. Because this holds for all $z \in B_R(y)$, we have established that:

$$B_R(y) \subseteq B_r(x)$$

Since every point $y \in B_r(x)$ has a neighborhood $B_R(y)$ entirely contained within $B_r(x)$, the open ball $B_r(x)$ is an open set. $\square$

Definition 5 (Closed Set)**.** Let (X, d) be a metric space. Let $F \subseteq X$. A subset $F \subseteq X$ is said to be closed if its complement $F^c = X \setminus F$ is open.

Theorem 3 (Closed Set Properties)**.** *Let* (X, d) *be a metric space.*

(i) The empty set $\varnothing$ and the whole space X are closed.

(ii) The arbitrary intersection of a collection of closed sets is closed.

(iii) The finite union of closed sets is closed.

Proof. Let (X, d) be a metric space. The proof follows from Definition 5, Theorem 1, and De Morgan's laws.

(i) The complement of X is $X^c = \varnothing$. Since the empty set is vacuously open, its complement X is closed.
The complement of the empty set is $\varnothing^c = X$. Since the entire metric space X is open, its complement $\varnothing$ is closed.

(ii) Let $\{F_\alpha\}_{\alpha \in I}$ be an arbitrary collection of closed sets in X, where I is an index set. Let $F = \bigcap\limits_{\alpha \in I} F_\alpha$.
Taking the complement and applying De Morgan's Laws yields:

$$F^c = \left(\bigcap_{\alpha \in I} F_\alpha \right)^c = \bigcup_{\alpha \in I} F_\alpha^c$$

Since each F_α is closed, its complement F_α^c is open by definition. Because the arbitrary union of open sets is open, $\bigcup\limits_{\alpha \in I} F_\alpha^c$ is open.
Since the complement F^c is open, the intersection F is closed.

(iii) Let $\{F_i\}_{i=1}^n$ be a finite collection of closed sets in X. Let $F = \bigcup\limits_{i=1}^{n} F_i$.
Taking the complement and applying De Morgan's Laws yields:

$$F^c = \left(\bigcup_{i=1}^{n} F_i \right)^c = \bigcap_{i=1}^{n} F_i^c$$

Since each F_i is closed, its complement F_i^c is open by definition. Because the finite intersection of open sets is open, $\bigcap\limits_{i=1}^{n} F_i^c$ is open.
Since the complement F^c is open, the union F is closed.

□

Definition 6 (Interior Point)**.** Let (X, d) be a metric space and Let $A \subseteq X$. A point $x \in A$ is called an interior point of A if there exists $r > 0$ such that $B_r(x) \subseteq A$.

Definition 7 (Interior)**.** Let (X, d) be a metric space and Let $A \subseteq X$. The interior of A, denoted by A° or $\text{int}(A)$, is the set of all interior points of A. That is,

$$A^\circ = \{x \in A : \text{there exists } r > 0 \text{ such that } B_r(x) \subseteq A\}$$

Definition 8 (Closure)**.** Let (X, d) be a metric space and Let $A \subseteq X$. The closure of A, denoted by $\overline{A}$, is the set of all points $x \in X$ such that every open ball centered at x intersects A. That is,

$$\overline{A} = \{x \in X : B_r(x) \cap A \neq \varnothing \text{ for every } r > 0\}$$

Definition 9 (Boundary)**.** Let (X, d) be a metric space and Let $A \subseteq X$. The boundary of A, denoted by ∂A, is defined by

$$\partial A = \overline{A} \setminus A^\circ \iff \partial A = \overline{A} \cap \overline{X \setminus A}$$

Definition 10 (Dense Set)**.** Let (X, d) be a metric space and Let $A \subseteq X$. We say that A is dense in X if $\overline{A} = X$. Equivalently, A is dense in X if every nonempty open ball in X intersects A.

Definition 11 (Compact Set)**.** Let (X, d) be a metric space. A set $K \subseteq X$ is compact if every open cover of K has a finite sub-cover.
Mathematically, if $\{U_\alpha\}_{\alpha \in I}$ is an open cover of K, where I is any index set, then there exists a finite sub-cover, namely $\{U_{\alpha_1}, U_{\alpha_2}, \ldots, U_{\alpha_n}\}$ of K. That is,

$$K \subseteq \bigcup_{\alpha \in I} U_\alpha \quad \implies \quad K \subseteq \bigcup_{i=1}^{n} U_{\alpha_i}$$

Theorem 4 (Sequential Compactness of Compact Sets)**.** *A compact set K is sequentially compact.*

Proof. Let K be a compact set and $\{x_n\}$ be a sequence in K.
Assume that K is not sequentially compact; thus, $\{x_n\}$ has no convergent subsequence converging in K.
Then, $\forall x \in K$, $\exists\, r_x > 0$ such that $B_{r_x}(x)$ contains finitely many points of $\{x_n\}$. The collection $\{B_{r_x}(x) : x \in K\}$ is an open cover of K, and K being compact, this has a finite sub-cover. Thus,

$$K \subseteq \bigcup_{i=1}^{k} B_{r_{x_i}}(x_i), \qquad \text{where } x_i \in K$$

Continued...

As K is covered by finitely many balls and by assumption each ball has only finitely many elements of $\{x_n\}$, then $\{x_n\}$ being infinite in K has some elements in K that are not covered by that finite collection. This is a contradiction to the compactness of K.
Thus, our assumption is wrong. Then, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ which is convergent to some $x \in K$. The existence of x in K is guaranteed as a compact set contains all its limit points.
Thus, compactness implies sequential compactness. ☐

Theorem 5 (Total Boundedness of Sequentially Compact Sets). *A sequentially compact set is totally bounded.*

Proof. We claim that for $\varepsilon > 0$, the set of points of mutual distance $\geq \varepsilon$ is finite, in the sequentially compact set K.
If this is not true, then we can construct a sequence $\{x_n\}$ in K such that

$$d(x_n, x_m) \geq \varepsilon \quad \text{whenever } m \neq n$$

Clearly this sequence does not have any convergent subsequence. This contradicts the sequential compactness of K. Thus the claim is true.
Now choose $x_1 \in K$. Then, if possible choose $x_2 \in K$ such that $d(x_1, x_2) \geq \varepsilon$. If this is not possible, then the ball $B_\varepsilon(x_1)$ covers K and we are done.
In the same process, if possible choose $x_3 \in K$ such that $d(x_1, x_3) \geq \varepsilon$ and $d(x_2, x_3) \geq \varepsilon$. Again, if this is not possible, then K is covered by two balls of radius ε centered at x_1 and x_2, and we are done.
We continue this process of picking a point as long as possible. But, by the validity of the claim, this process must terminate after a finite number of steps. And thus K can be covered by a finite number of ε-balls.
This shows K is totally bounded.
Therefore, sequential compactness implies total boundedness. ☐

Lemma 1 (Lebesgue Covering Lemma). *If K is sequentially compact and $\{G_\alpha\}$ is any open cover of K, then there exists $r > 0$ such that for every $x \in K$, the ball $B_r(x)$ is contained in a single open set of the cover $\{G_\alpha\}$.*

Proof. Assume that this is not true. Then for every $n \in \mathbb{N}$ there exists a ball of radius $1/n$ which is not contained in any of the sets of $\{G_\alpha\}$.
Thus, for each $n \in \mathbb{N}$ there exists $x_n \in K$ such that $B_{1/n}(x_n)$ is not contained in any set of $\{G_\alpha\}$.
For each such $n \in \mathbb{N}$, using these x_n we construct the sequence $\{x_n\}_{n=1}^{\infty}$ in K. Since K is sequentially compact, the sequence $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$ converging to some $x \in K$.
Because $\{G_\alpha\}$ is an open cover of K and $x \in K$, there exists some index α_0 such that

$x \in G_{\alpha_0}$. As G_{α_0} is open, there exists $\delta > 0$ such that $B_\delta(x) \subseteq G_{\alpha_0}$.
Since $x_{n_k} \to x$, there exists $m_1 \in \mathbb{N}$ such that for all $k \geq m_1$,

$$d(x_{n_k}, x) < \frac{\delta}{2}$$

Also, since $n_k \to \infty$, there exists $m_2 \in \mathbb{N}$ such that for all $k \geq m_2$,

$$\frac{1}{n_k} < \frac{\delta}{2}$$

Let $m = \max\{m_1, m_2\}$. Then for every $k \geq m$ and every $z \in B_{1/n_k}(x_{n_k})$, we have

$$d(z, x) \leq d(z, x_{n_k}) + d(x_{n_k}, x) < \frac{1}{n_k} + \frac{\delta}{2} < \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

Hence,

$$B_{1/n_k}(x_{n_k}) \subseteq B_\delta(x) \subseteq G_{\alpha_0}$$

This contradicts the choice of x_{n_k}, since $B_{1/n_k}(x_{n_k})$ was assumed not to be contained in any member of the open cover $\{G_\alpha\}$.
This contradiction shows the assumption was false, and therefore there exists $r > 0$ with the required property. ☐

Theorem 6 (Compactness of Sequentially Compact Sets). *A sequentially compact set is compact.*

Proof. As $\{G_\alpha\}$ is any open cover of K, by Lemma 1, there exists $r > 0$ such that for all $x \in K$, $B_r(x)$ is contained in a single G_α.
By Theorem 5, as sequential compactness implies totally boundedness, thus, there exists $x_1, x_2, \ldots, x_t \in K$ such that

$$K \subseteq \bigcup_{i=1}^{t} B_r(x_i)$$

As each ball $B_r(x_i)$ is contained in a single set of $\{G_\alpha\}$, say G_{α_i}, then the collection $\{G_{\alpha_i}\}_{i=1}^{t}$ is a finite subcover of K. This shows that K is compact.
Thus, sequential compactness implies compactness. ☐

Theorem 7 (Compactness of Closed and Totally Bounded Sets). *Let X be a complete metric space. Every closed and totally bounded subset K of X is compact.*

Proof. Since K is closed in the complete metric space X, K is complete. Thus K is a closed, totally bounded and complete set. Let $\{x_n\}$ be a sequence in K. As K is

totally bounded, then, $\forall r > 0$, K can be covered by finitely many balls of radius r.
Consider $r = 1$. Then

$$K \subseteq \bigcup_{i=1}^{n_1} B_1\big(y_i^{(1)}\big), \qquad y_i^{(1)} \in K$$

As $\{x_n\}$ is an infinite sequence, at least one of the balls $B_1\big(y_i^{(1)}\big)$ must contain infinitely many elements of $\{x_n\}$.
Let $B^{(1)}$ denote that ball containing infinitely many terms, and choose the infinite subsequence of $\{x_n\}$ contained in $B^{(1)}$; denote it by $\{x_n^{(1)}\}$.
Next choose $r = \frac{1}{2}$ and cover $B^{(1)}$ by balls of radius $\frac{1}{2}$. This covering is again finite, thus one of the balls of radius $\frac{1}{2}$ will have infinitely many terms of $\{x_n^{(1)}\}$. Call this ball $B^{(2)}$ and the infinite subsequence contained in it $\{x_n^{(2)}\}$.
Continue this process inductively for all $k \geq 1$. We obtain an infinite nested chain of balls $B^{(k)}$ where $B^{(k)}$ has radius $r_k = 2^{-k+1}$ (or simply 2^{-k} depending on indexing), and an infinite subsequence $\{x_n^{(k)}\}$ contained in $B^{(k)}$. Moreover we may choose these subsequences so that the indices are strictly increasing: choose $n_1 < n_2 < \cdots$ with $x_{n_k} \in B^{(k)}$ (pick, for example, the first term of $\{x_n^{(k)}\}$ whose index is larger than n_{k-1}).
Now, consider the subsequence $\{x_n^{(n)}\}$ of the original sequence $\{x_n\}$. If $m \geq n$, then, using the triangle inequality,

$$d\big(x_m^{(m)}, x_n^{(n)}\big) \leq d\big(x_m^{(m)}, x_{m-1}^{(m)}\big) + \cdots + d\big(x_{n+1}^{(n+1)}, x_n^{(n)}\big)$$

Now, since

$$d\big(x_k^{(k)}, x_{k-1}^{(k-1)}\big) \leq 2^{2-k}$$

as both $\{x_k^{(k)}\}$ and $\{x_{k-1}^{(k-1)}\}$ are in $B^{\,k-1}$ by construction. Thus,

$$d\big(x_m^{(m)}, x_n^{(n)}\big) \leq 2^{2-m} + 2^{2-(m-1)} + \cdots + 2^{2-(n+1)} < 2^{2-n}$$

Hence, $\{x_n^{(n)}\}$ is a Cauchy sequence in K.
Then, as K is complete, every Cauchy sequence is convergent. Thus we have found a convergent subsequence $\{x_n^{(n)}\}$ of the sequence $\{x_n\}$.
Then, by closedness of K, the limit point to which the subsequence $\{x_n^{(n)}\}$ converges is in K.
Thus $\{x_n^{(n)}\}$ is a convergent subsequence of $\{x_n\}$ converging to a point in K, and $\{x_n\}$ being any arbitrary sequence.
This shows K is sequentially compact and hence compact, by Theorem 6. □

Theorem 8 (Closedness of Compact Sets). *A compact set is closed.*

Proof. Let K be a compact set in a metric space (X, d). Let $x \in K^c$. For each $y \in K$ define

$$r_y = \tfrac{1}{2}d(x, y)$$

Then $x \in B_{r_y}(x)$ and $y \in B_{r_y}(y)$, and clearly

$$B_{r_y}(x) \cap B_{r_y}(y) = \varnothing$$

Thus $\{B_{r_y}(y) : y \in K\}$ is an open cover of K. By compactness of K, there exist $y_1, \ldots, y_m \in K$ such that

$$K \subseteq \bigcup_{i=1}^{m} B_{r_{y_i}}(y_i)$$

Choose

$$r = \min\{r_{y_i} : i = 1, 2, \ldots, m\}$$

Clearly $r > 0$. For each i we have $B_r(x) \cap B_{r_{y_i}}(y_i) = \varnothing$ since $r \leq r_{y_i}$, so $B_r(x) \cap K = \varnothing$. Hence $B_r(x) \subseteq K^c$, which shows K^c is open. Therefore K is closed.
Thus compactness implies closedness. □

Definition 12 (Continuous Mapping in metric space)**.** Let (X, d_X) and (Y, d_Y) be two metric spaces. Then the mapping $f : (X, d_X) \to (Y, d_Y)$ is said to be continuous at a point $x \in X$, if for every $\varepsilon > 0$, there exists, $\delta > 0$, such that, $f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$. Equivalently, if for every $\varepsilon > 0$, there exists, $\delta > 0$, such that, $d_X(x, y) < \delta \implies d_Y(f(x), f(y)) < \varepsilon$.

Theorem 9 (Topological Characterization of Continuity)**.** *Let (X, d_X) and (Y, d_Y) be two metric spaces. The map $f : (X, d_X) \to (Y, d_Y)$ is continuous if and only if for every open set $V \subseteq Y$, the set $f^{-1}(V)$ is open in X.*

Proof. ($\implies$) Assume f is continuous on X. Let $V \subseteq Y$ be an open set. If $f^{-1}(V) = \emptyset$, it is open by definition. If $f^{-1}(V) \neq \emptyset$, choose an arbitrary point $x \in f^{-1}(V)$.
By definition of the preimage, $f(x) \in V$. Since V is open in Y, there exists an $\varepsilon > 0$ such that the open ball $B^Y_\varepsilon(f(x)) \subseteq V$. Because f is continuous at x, there exists a $\delta > 0$ such that for all $x' \in X$:

$$d_X(x, x') < \delta \implies d_Y(f(x), f(x')) < \varepsilon$$

This implies that $f(B^X_\delta(x)) \subseteq B^Y_\varepsilon(f(x)) \subseteq V$. Taking the preimage on both sides yields $B^X_\delta(x) \subseteq f^{-1}(V)$. Since every point $x \in f^{-1}(V)$ is an interior point, $f^{-1}(V)$ is open in X.

($\impliedby$) Assume that for every open set $V \subseteq Y$, the preimage $f^{-1}(V)$ is open in X. Let $x \in X$ and Let $\varepsilon > 0$ be given.

Consider the open ball $V = B_\varepsilon^Y(f(x))$, which is an open set in Y. By our assumption, its preimage $f^{-1}(B_\varepsilon^Y(f(x)))$ is open in X. Since $f(x) \in B_\varepsilon^Y(f(x))$, the point x belongs to this preimage.
Because the preimage is open, there exists a $\delta > 0$ such that the open ball $B_\delta^X(x) \subseteq f^{-1}(B_\varepsilon^Y(f(x)))$. This means that for any $x' \in X$:

$$d_X(x, x') < \delta \implies f(x') \in B_\varepsilon^Y(f(x)) \implies d_Y(f(x), f(x')) < \varepsilon$$

This matches the ε-δ definition of continuity at x. Since x was arbitrary, f is continuous on X. □

Theorem 10 (Preservation of Compactness)**.** *Let (X, d_X) and (Y, d_Y) be two metric spaces. Let the map $f : (X, d_X) \to (Y, d_Y)$ be continuous. Then the continuous image $f(K) \subseteq Y$ of a compact set $K \subseteq X$ is compact.*

Proof. Let $\{V_\alpha\}_{\alpha \in I}$ be an open cover of $f(K)$, meaning $f(K) \subseteq \bigcup\limits_{\alpha \in I} V_\alpha$ where each V_α is open in Y.
Because f is continuous, the preimage $f^{-1}(V_\alpha)$ is open in X for every $\alpha \in I$. By the properties of preimages, we have:

$$K \subseteq f^{-1}(f(K)) \subseteq f^{-1}\left(\bigcup_{\alpha \in I} V_\alpha\right) = \bigcup_{\alpha \in I} f^{-1}(V_\alpha)$$

Thus, $\{f^{-1}(V_\alpha)\}_{\alpha \in I}$ forms an open cover of K.
Since K is compact, this cover admits a finite subcover. There exist finitely many indices $\alpha_1, \ldots, \alpha_n \in I$ such that:

$$K \subseteq \bigcup_{i=1}^{n} f^{-1}(V_{\alpha_i})$$

Applying f to both sides yields:

$$f(K) \subseteq f\left(\bigcup_{i=1}^{n} f^{-1}(V_{\alpha_i})\right) = \bigcup_{i=1}^{n} f(f^{-1}(V_{\alpha_i})) \subseteq \bigcup_{i=1}^{n} V_{\alpha_i}$$

Hence, the arbitrary open cover $\{V_\alpha\}$ has a finite subcover $\{V_{\alpha_1}, \ldots, V_{\alpha_n}\}$, proving that $f(K)$ is compact. □

Definition 13 (Lipschitz Mapping)**.** Let (X, d_X) and (Y, d_Y) be two metric spaces. Then the mapping $f : (X, d_X) \to (Y, d_Y)$ is said to be Lipschitz if there exists a constant $A > 0$ such that, we have $d(f(x), f(y)) \leq Ad(x, y)$, for all $x, y \in X$.

Definition 14 (Product Metric)**.** Let (X, d_X) and (Y, d_Y) be two metric spaces. The product space $X \times Y$ consists of all ordered pairs (x, y) where $x \in X$ and $y \in Y$. To define a metric d on $X \times Y$ we can use the following; if $(x_1, y_1), (x_2, y_2) \in X \times Y$

- l_1-metric: $d((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2)$
- l_2-metric: $d((x_1, y_1), (x_2, y_2)) = \sqrt{d_X(x_1, x_2)^2 + d_Y(y_1, y_2)^2}$
- l_∞-metric: $d((x_1, y_1), (x_2, y_2)) = \max\{d_X(x_1, x_2), d_Y(y_1, y_2)\}$

Definition 15 (Open Sets in Product Metric)**.** An open set in the product space $V \in X \times Y$ is described as for $\tilde{x} = (x, y) \in X \times Y$, there exists $r > 0$ such that $B_r(\tilde{x}) \subseteq V$.
If we consider the l_∞-metric, where $\tilde{x} = (x_1, y_1)$ and $\tilde{y} = (x_2, y_2)$ then, $d(\tilde{x}, \tilde{y}) = \max\{d_X(x_1, x_2), d_Y(y_1, y_2)\}$ and so, $B_r(\tilde{x}) = B_r(x_1) \times B_r(y_1)$.

Theorem 11 (Finite Product of Compact Metric Spaces)**.** *Let (X, d_X) and (Y, d_Y) be two metric spaces. Let $K_1 \subseteq X$ and $K_2 \subseteq Y$ be compact. Then $K_1 \times K_2$ is compact in $X \times Y$.*

Proof. We will use the property of sequential compactness. Let $\{(x_n, y_n)\}_{n=1}^{\infty}$ be an arbitrary sequence in $K_1 \times K_2$. By definition of the Cartesian product, $x_n \in K_1$ and $y_n \in K_2$ for all $n \in \mathbb{N}$.
Since K_1 is compact, the sequence $\{x_n\}_{n=1}^{\infty}$ contains a subsequence $\{x_{n_i}\}_{i=1}^{\infty}$ that converges to some point $x \in K_1$:

$$\lim_{i \to \infty} x_{n_i} = x$$

We consider the corresponding subsequence $\{y_{n_i}\}_{i=1}^{\infty}$ in K_2. Since K_2 is compact, this subsequence contains a further subsequence (a sub-subsequence) $\{y_{n_{i_j}}\}_{j=1}^{\infty}$ that converges to some point $y \in K_2$:

$$\lim_{j \to \infty} y_{n_{i_j}} = y$$

Now, looking at the corresponding sub-subsequence of the first component, $\{x_{n_{i_j}}\}_{j=1}^{\infty}$. Since it is a subsequence of the convergent sequence $\{x_{n_i}\}_{i=1}^{\infty}$, it also converges to the same limit x:

$$\lim_{j \to \infty} x_{n_{i_j}} = x$$

By the definition of convergence in the product metric space, the sequence of pairs converges component-wise. Therefore, the sub-subsequence $\{(x_{n_{i_j}}, y_{n_{i_j}})\}_{j=1}^{\infty}$ converges to $(x, y) \in K_1 \times K_2$:

$$\lim_{j \to \infty} (x_{n_{i_j}}, y_{n_{i_j}}) = (x, y)$$

We have shown that every sequence in $K_1 \times K_2$ has a subsequence that converges to a point in $K_1 \times K_2$. Thus, $K_1 \times K_2$ is compact, by Theorem 6. □

Definition 16 (Topological Space)**.** Let X be any nonempty set. A topology on a set X may be defined as a collection τ of subsets of X, called open sets and satisfying the following axioms:

(i) The empty set $\varnothing$ and X belong to τ.

(ii) An arbitrary union of members of τ belong to τ.

(iii) The intersection of any finite number of members of τ belong to τ.

Then (X, τ) is called a topological space and the set τ of the open sets is commonly called a topology on X.

Definition 17 (Outer Measure)**.** Let X be any nonempty set. Then the map $\mu^* : \mathcal{P}(X) \to [0, \infty]$ is said to be an outer measure if:

(i) For the empty set $\varnothing$, we have $\mu^*(\varnothing) = 0$.

(ii) If $A \subseteq B$, where $A, B \in \mathcal{P}(X)$, then $\mu^*(A) \leq \mu^*(B)$.

(iii) For any sequence of sets $\{A_n\}$ from $\mathcal{P}(X)$, we have, $\mu^* \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

Remark 1 (Standard Construction of Outer Measure)**.** Let X be any nonempty set. Let $S \subseteq \mathcal{P}(X)$ and $h : S \to [0, \infty]$, then,

$$\mu^*(E) = \inf\{\sum_j h(A_j) : A_j \in S,\ E \subseteq \bigcup_j A_j\}$$

where j runs over a finite or countable index set. This forms an outer measure.

Definition 18 (σ-Algebra)**.** Let X be any nonempty set. Let $\mathcal{A} \subseteq \mathcal{P}(X)$. Then, $\mathcal{A}$ is called a σ-algebra if:

(i) The whole set X and the empty set $\varnothing$ must be in $\mathcal{A}$.

(ii) If a set A is in $\mathcal{A}$, its complement $A^c = X \setminus A$ is also in $\mathcal{A}$.

(iii) If sets $A_1, A_2, \ldots$ are in $\mathcal{A}$, then their union $\bigcup_{i=1}^{\infty} A_i$ is also in $\mathcal{A}$.

Definition 19 (Borel σ-algebra)**.** Let X be any nonempty metric (or topological) space. Then the σ-algebra generated by the family of all open sets is called the Borel σ-algebra on X and represented as $\mathcal{B}(X)$.

Definition 20 (Measure)**.** For a nonempty set X and a σ-algebra $\mathcal{A}$ over X, a measure μ is a function $\mu : \mathcal{A} \to [0, \infty]$ that satisfies two main properties:

(i) For the empty set $\varnothing$, we have $\mu(\varnothing) = 0$.

(ii) For every countable collection of pairwise disjoint sets $A_1, A_2, \dots$ in $\mathcal{A}$, we have,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Example 5 (Dirac Measure)
For a nonempty set X and the σ-algebra $\mathcal{P}(X)$, fix a point $X_0 \in X$, then the Dirac measure for all sets $E \in \mathcal{P}(X)$ is defined as:

$$\delta_{x_0}(E) = \begin{cases} 1 \quad , \quad x_0 \in E \\ 0 \quad , \quad x_0 \notin E \end{cases}$$

Definition 21 (Well Separating Set/ Carathéodory Measurable Set)**.** Let μ^* be an outer measure on $\mathcal{P}(X)$. A set $A \subseteq X$ is well-separating or Carathéodory measurable with respect to μ^* if,

$$\forall E \subseteq X, \qquad \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Theorem 12 (Carathéodory Theorem (I))**.** *Let μ^* be an outer measure on X. The collection of well separating sets is a σ-algebra. Furthermore, the restriction of μ^* to that collection is a complete measure.*

Proof. Let $\mathcal{A}$ be the collection of well separating sets w.r.t. μ^*. Then,

- Let $A = \varnothing$. Then $\forall E \subseteq X$,

  $$\mu^*(E) = 0 + \mu^*(E) = \mu^*(\varnothing) + \mu^*(E) = \mu^*(E \cap \varnothing) + \mu^*(E \cap X).$$

  As $\varnothing = X^c$, thus $\varnothing$, $X \in \mathcal{A}$.

- Let $A \in \mathcal{A}$. Then $\forall E \subseteq X$,

  $$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*(E \cap A^c) + \mu^*(E \cap (A^c)^c).$$

  Thus $A^c \in \mathcal{A}$.

- We next show that $\mathcal{A}$ is closed under finite unions. Let $A, B \in \mathcal{A}$. We shall prove that $A \cup B \in \mathcal{A}$.
  Then since $A \in \mathcal{A}$, for all $E \subseteq X$,
  $$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c).$$
  Again, as $B \in \mathcal{A}$, applying the same identity replacing E with $E \cap A$ and $E \cap A^c$ gives
  $$\mu^*(E \cap A) = \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c)$$
  $$\mu^*(E \cap A^c) = \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c)$$
  Therefore,
  $$\mu^*(E) = \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c)$$
  Now, we note that,
  $$E \cap A^c \cap B^c = E \cap (A \cup B)^c,$$
  and
  $$E \cap (A \cup B) = (E \cap A \cap B) \cup (E \cap A \cap B^c) \cup (E \cap A^c \cap B)$$
  By sub-additivity of μ^*,
  $$\mu^*(E \cap (A \cup B)) \leq \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B)$$
  Using the decomposition of $\mu^*(E)$ above we obtain
  $$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$$
  Then, by sub-additivity of μ^*,
  $$\mu^*(E) \leq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$$
  Therefore
  $$\mu^*(E) = \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$$
  Thus $A \cup B \in \mathcal{A}$.
  Since $\mathcal{A}$ is closed under complements and finite unions, it is also closed under finite intersections by De Morgan's laws.

- In particular, if $A, B \in \mathcal{A}$ and $A \cap B = \varnothing$, then we have
  $$E \cap A^c \cap B = E \cap B \qquad \text{and} \qquad E \cap A^c \cap B^c = (A \cup B)^c \cap E$$
  Then, as $A \in \mathcal{A}$, for every $E \subseteq X$,
  $$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

And as $B \in \mathcal{A}$, using E replaced by $E \cap A^c$, we get

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c)$$

hence

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap B) + \mu^*(E \cap (A \cup B)^c)$$

If we choose $E = A \cup B$, then

$$(A \cup B) \cap A = A, \quad (A \cup B) \cap B = B, \quad (A \cup B) \cap (A \cup B)^c = \varnothing$$

and thus

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$$

- Let $\{A_i\}$ be a countable collection of sets in $\mathcal{A}$. Let

$$A = \bigcup_{i=1}^{\infty} A_i, \qquad B_1 = A_1, \qquad B_n = A_n \setminus \bigcup_{i=1}^{n-1} A_i \quad (n \geq 2)$$

Then each $B_i \in \mathcal{A}$, and the collection $\{B_i\}$ is countable and pairwise disjoint. Also $B_n \subseteq A_n$ for every $n \geq 1$. Moreover,

$$A = \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$$

Now, consider the finite union

$$F_n = \bigcup_{i=1}^{n} B_i$$

Then $F_n \in \mathcal{A}, \quad \forall n \geq 1$

Also, since $\{B_i\}$ is a disjoint collection, $\{E \cap B_i\}$ is a disjoint collection for every $E \subseteq X$. Hence

$$\mu^*(E) = \mu^*(E \cap F_n) + \mu^*(E \cap F_n^c)$$

and

$$\mu^*(E \cap F_n) = \sum_{i=1}^{n} \mu^*(E \cap B_i)$$

Therefore, for every $n \geq 1$,

$$\mu^*(E) = \mu^*(E \cap F_n^c) + \sum_{i=1}^{n} \mu^*(E \cap B_i)$$

Now we note that F_n^c is a decreasing sequence and $F_n^c = \big(\bigcup\limits_{i=1}^{n} B_i\big)^c$, so

$$\bigcap_{n=1}^{\infty} F_n^c = \left(\bigcup_{i=1}^{\infty} B_i\right)^c = A^c$$

Since μ^* is monotone, the sequence $\mu^*(E \cap F_n^c)$ is decreasing and bounded below. Clearly, as;

$$F_n^c \downarrow \left(\bigcup_{i=1}^{\infty} B_i\right)^c \implies (E \cap F_n^c) \downarrow (E \cap \left(\bigcup_{i=1}^{\infty} B_i\right)^c) = (E \cap A^c)$$

and therefore

$$\mu^*(E \cap A^c) \ \leq \ \lim_{n\to\infty} \mu^*(E \cap F_n^c)$$

And by sub-additivity,

$$\mu^*(E \cap A) \leq \sum_{i=1}^{\infty} \mu^*(E \cap B_i)$$

By previous work,

$$\mu^*(E) = \sum_{i=1}^{n} \mu^*(E \cap B_i) + \mu^*(E \cap F_n^c)$$

As $n \to \infty$, we obtain,

$$\mu^*(E) = \lim_{n\to\infty} \left(\sum_{i=1}^{n} \mu^*(E \cap B_i) + \mu^*(E \cap F_n^c)\right)$$

And therefore,

$$\mu^*(E) \geq \sum_{i=1}^{\infty} \mu^*(E \cap B_i) + \mu^*(E \cap A^c)$$

Since

$$E \cap A = \bigcup_{i=1}^{\infty} (E \cap B_i)$$

sub-additivity gives

$$\mu^*(E \cap A) \leq \sum_{i=1}^{\infty} \mu^*(E \cap B_i)$$

Hence,

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

Continued...

On the other hand, by sub-additivity,

$$\mu^*(E) \le \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

Therefore,

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

Thus,

$$A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$$

Therefore $\mathcal{A}$ is a σ-algebra.

Next we show that the outer measure restricted to $\mathcal{A}$ is a complete measure. Let $\mu = \mu^*|_{\mathcal{A}}$.

- By definition of outer measure, $\mu^*(\varnothing) = 0$ and $\varnothing \in \mathcal{A}$. So $\mu(\varnothing) = 0$.
- Let $A, B \in \mathcal{A}$ with $A \subseteq B$. Then, as μ^* is monotone,

  $$\mu^*(A) \le \mu^*(B) \implies \mu(A) \le \mu(B)$$

  Thus μ is monotone.
- Let $\{B_i\}$ be a countable collection of disjoint sets in $\mathcal{A}$. Then, by a previous result we had, if $A, B \in \mathcal{A}$ and $A \cap B = \varnothing$, then

  $$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$$

  hence

  $$\mu(A \cup B) = \mu(A) + \mu(B)$$

  Then, by induction we get,

  $$S_n = \bigcup_{i=1}^{n} B_i \implies \mu(S_n) = \sum_{i=1}^{n} \mu(B_i)$$

  Let

  $$S = \bigcup_{i=1}^{\infty} B_i$$

  Since $\mathcal{A}$ is a σ-algebra, we have $S \in \mathcal{A}$. Applying the Carathéodory property to $S_n \subseteq S$ with $E = S$ gives,

  $$\mu(S) = \mu(S_n) + \mu(S \setminus S_n)$$

Therefore,

$$\mu(S) = \sum_{i=1}^{n} \mu(B_i) + \mu(S \setminus S_n)$$

Since $\mu(S \setminus S_n) \geq 0$,

$$\mu(S) \geq \sum_{i=1}^{n} \mu(B_i) \qquad \forall n$$

Hence,

$$\mu(S) \geq \sum_{i=1}^{\infty} \mu(B_i)$$

On the other hand, using countable sub-additivity of the outer measure,

$$\mu(S) = \mu^*(S) \leq \sum_{i=1}^{\infty} \mu^*(B_i) = \sum_{i=1}^{\infty} \mu(B_i)$$

Therefore,

$$\mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mu(B_i)$$

Thus μ is countably additive.

- To show that μ is complete it is enough to show that if $A \subseteq B$, $B \in \mathcal{A}$ and $\mu(B) = 0$, then $A \in \mathcal{A}$ and $\mu(A) = 0$. Since $A \subseteq B$ and $\mu(B) = 0$, monotonicity of the outer measure gives

  $$0 \leq \mu^*(A) \leq \mu^*(B) = \mu(B) = 0$$

  Hence, $\mu^*(A) = 0$.
  We now show that $A \in \mathcal{A}$. Let $E \subseteq X$ be arbitrary. By sub-additivity,

  $$\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

  For the reverse inequality, since $E \cap A \subseteq A$, we have,

  $$\mu^*(E \cap A) \leq \mu^*(A) = 0$$

  Hence,

  $$\mu^*(E \cap A) = 0$$

  Also, by monotonicity,

  $$\mu^*(E \cap A^c) \leq \mu^*(E)$$

  Therefore,

  $$\mu^*(E \cap A) + \mu^*(E \cap A^c) = 0 + \mu^*(E \cap A^c) \leq \mu^*(E)$$

Combining both inequalities, we obtain,

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

Since this holds for every $E \subseteq X$, we have $A \in \mathcal{A}$. Finally, because $A \in \mathcal{A}$,

$$\mu(A) = \mu^*(A) = 0$$

Thus, μ is a complete measure.

$\square$

Definition 22 (Approximating Hausdorff Measures)**.** Let (X, d) be any metric space. Fix any function h, such that

$$h : \mathcal{P}(X) \longrightarrow [0, \infty]$$

For any set $A \subseteq X$ we define the diameter as

$$\operatorname{diam}(A) = \sup\{d(x, y) : x, y \in A\} \quad \text{and} \quad \operatorname{diam}(\varnothing) = 0.$$

Take $\varepsilon > 0$ and set $E \subseteq X$. Let $\{E_j\}_{j=1}^{\infty}$ be a countable cover of E. Then we define

$$\mathcal{H}^{\varepsilon}(E) = \inf\Big\{ \sum_{j=1}^{\infty} h(E_j) \ : \ E \subseteq \bigcup_{j=1}^{\infty} E_j, \ \forall j, \ \operatorname{diam}(E_j) < \varepsilon \Big\}$$

This is called the Approximating Hausdorff Measure.

Definition 23 (Hausdorff Measures)**.** Assuming the existence of Approximating Hausdorff Measures, we define the Hausdorff measure by

$$\mathcal{H}(E) = \lim_{\varepsilon \to 0^+} \mathcal{H}^{\varepsilon}(E) = \sup_{\varepsilon > 0} \mathcal{H}^{\varepsilon}(E)$$

Theorem 13 (Outer Measure Validity of Approximating Hausdorff Measures)**.** *For each fixed $\varepsilon > 0$, the Approximating Hausdorff Measure, $\mathcal{H}^{\varepsilon}$ is an outer measure.*

Proof. Fix $\varepsilon > 0$, by definition,

- As $\operatorname{diam}(\varnothing) = 0$, then $h(\varnothing) = 0$ and $\varnothing$ is covered by countably many sets of diameter 0. Thus

  $$\mathcal{H}^{\varepsilon}(\varnothing) = \inf \sum h(\varnothing) = \inf 0 = 0$$

  Therefore $\mathcal{H}^{\varepsilon}(\varnothing) = 0$.

- If $0 < \varepsilon_1 < \varepsilon_2$, then for $E \subseteq X$ the collection of covers of E of diameter less than ε_1 is a subset of the collection of covers of E of diameter less than ε_2. Hence

  $$\inf_{\operatorname{diam}(E_j)<\varepsilon_2} \sum_{j=1}^{\infty} h(E_j) \leq \inf_{\operatorname{diam}(E_j)<\varepsilon_1} \sum_{j=1}^{\infty} h(E_j)$$

  so

  $$\mathcal{H}^{\varepsilon_2}(E) \leq \mathcal{H}^{\varepsilon_1}(E)$$

  Thus, $\mathcal{H}^{\varepsilon}(E)$ is decreasing as ε is increasing.

- If $A \subseteq B \subseteq X$, then any cover of B covers A. Thus

  $$\inf_{A} \sum_{j=1}^{\infty} h(E_j) \leq \inf_{B} \sum_{j=1}^{\infty} h(E_j)$$

  where $A \subseteq B \subseteq \bigcup_{j=1}^{\infty} E_j$. Hence $\mathcal{H}^{\varepsilon}(A) \leq \mathcal{H}^{\varepsilon}(B)$.

- Let $\{E_i\}$ be any countable collection of sets in X. Take $\delta > 0$. For each i choose a countable cover of E_i, say $\{O_{i,n}\}$, such that

  $$\sum_{n=1}^{\infty} h(O_{i,n}) \leq \mathcal{H}^{\varepsilon}(E_i) + \frac{\delta}{2^i}$$

  Then, since

  $$\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} \bigcup_{n=1}^{\infty} O_{i,n}$$

  we have

  $$\sum_{i=1}^{\infty} \sum_{n=1}^{\infty} h(O_{i,n}) \leq \sum_{i=1}^{\infty} \Big(\mathcal{H}^{\varepsilon}(E_i) + \frac{\delta}{2^i}\Big) = \sum_{i=1}^{\infty} \mathcal{H}^{\varepsilon}(E_i) + \delta$$

  Taking infimum over all sums and letting $\delta \to 0$, we have,

  $$\mathcal{H}^{\varepsilon}\Big(\bigcup_{i=1}^{\infty} E_i\Big) \leq \sum_{i=1}^{\infty} \mathcal{H}^{\varepsilon}(E_i)$$

Therefore, $\mathcal{H}^{\varepsilon}$ is an outer measure for fixed ε. □

Theorem 14 (Outer Measure Validity of Hausdorff Measures)**.** *The Hausdorff Measure $\mathcal{H}$ is an outer measure.*

Proof. Now, for $\mathcal{H}$, we have,

- $\mathcal{H}(\varnothing) = \lim\limits_{\varepsilon\to 0^+} \mathcal{H}^\varepsilon(\varnothing) = \lim\limits_{\varepsilon\to 0^+} 0 = 0.$
- If $A \subseteq B$ then $\mathcal{H}^\varepsilon(A) \leq \mathcal{H}^\varepsilon(B)$. Taking the limit $\varepsilon \to 0^+$ we get $\mathcal{H}(A) \leq \mathcal{H}(B)$.
- If $\{E_i\}$ is a countable collection of sets, then, for every fixed $\varepsilon > 0$,

  $$\mathcal{H}^\varepsilon\Big(\bigcup_{i=1}^{\infty} E_i\Big) \leq \sum_{i=1}^{\infty} \mathcal{H}^\varepsilon(E_i) \leq \sum_{i=1}^{\infty} \mathcal{H}(E_i)$$

  Now take the supremum over ε,

  $$\mathcal{H}\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sup_{\varepsilon>0} \mathcal{H}^\varepsilon\Big(\bigcup_{i} E_i\Big) \leq \sum_{i=1}^{\infty} \mathcal{H}(E_i)$$

Therefore, $\mathcal{H}$ is also an outer measure. □

Theorem 15 (Additivity of Hausdorff Measure for Positively Separated Sets)**.** *Let (X, d) be a metric space. If $A, B \subseteq X$ are positively separated sets, then the Hausdorff measure $\mathcal{H}$ is additive.*

Proof. Now, if $A, B \subseteq X$ are disjoint (or positively separated sets), then $d(A, B) = d$, i.e.

$$d = \inf\{d(x, y) : x \in A,\ y \in B\} > 0$$

Then for any $\varepsilon < d$, we have

$$\mathcal{H}^\varepsilon(A \cup B) = \mathcal{H}^\varepsilon(A) + \mathcal{H}^\varepsilon(B)$$

and consequently

$$\mathcal{H}(A \cup B) = \mathcal{H}(A) + \mathcal{H}(B)$$

To show this, consider $\{O_i\}$ to be a countable cover of $A \cup B$ such that $\forall i \geq 1$, $\operatorname{diam}(O_i) < \varepsilon < d$. Then each element of $\{O_i\}$ intersects either A or B by this construction as, $\operatorname{diam}(O_i) < \varepsilon < d(A, B)$.
If $\{O_i\}$ met both A and B, there would be $a \in A \cap O_i$, and $b \in B \cap O_i$ and therefore, $d(a, b) \leq \operatorname{diam}(O_i) < d(A, B)$, a contradiction.
Hence the cover splits into a cover of A and a cover of B.
Also, we have, $\forall \delta > 0$,

$$\sum_{i=1}^{\infty} h(O_i) \leq \mathcal{H}^\varepsilon(A \cup B) + \delta$$

By construction, Let the elements of $\{O_i\}$ which intersect with A be $\{O_i^A\}$ and those which intersect with B be $\{O_i^B\}$. Clearly,

$$\forall i,\ O_i^A \cap O_i^B = \varnothing \quad \text{and} \quad \{O_i\} = \{O_i^A\} \cup \{O_i^B\}$$

Then,

$$\mathcal{H}^{\varepsilon}(A) \leq \sum_{i=1}^{\infty} h(O_i^A) \qquad \text{and} \qquad \mathcal{H}^{\varepsilon}(B) \leq \sum_{i=1}^{\infty} h(O_i^B)$$

Hence,

$$\mathcal{H}^{\varepsilon}(A) + \mathcal{H}^{\varepsilon}(B) \leq \sum_{i=1}^{\infty} h(O_i^A) + \sum_{i=1}^{\infty} h(O_i^B) \leq \sum_{i=1}^{\infty} h(O_i) \leq \mathcal{H}^{\varepsilon}(A \cup B) + \delta$$

As $\delta \to 0$, we obtain

$$\mathcal{H}^{\varepsilon}(A) + \mathcal{H}^{\varepsilon}(B) \leq \mathcal{H}^{\varepsilon}(A \cup B)$$

But by sub-additivity,

$$\mathcal{H}^{\varepsilon}(A \cup B) \leq \mathcal{H}^{\varepsilon}(A) + \mathcal{H}^{\varepsilon}(B)$$

Therefore,

$$\mathcal{H}^{\varepsilon}(A \cup B) = \mathcal{H}^{\varepsilon}(A) + \mathcal{H}^{\varepsilon}(B)$$

Taking the limit $\varepsilon \to 0^+$ yields

$$\mathcal{H}(A \cup B) = \mathcal{H}(A) + \mathcal{H}(B)$$

□

Theorem 16 (Well-Separating Property of Borel Sets)**.** *Every Borel set is well separating with respect to the Hausdorff measure $\mathcal{H}$.*

Proof. As Borel sets are generated by open sets and Borel sets form a σ-algebra, it is sufficient to show that any open set $G \subseteq X$ is well separating with respect to $\mathcal{H}$.
Let $G \subseteq X$ be any open set. Then for every $E \subseteq X$, by subadditivity,

$$\mathcal{H}(E) \leq \mathcal{H}(E \cap G) + \mathcal{H}(E \cap G^c)$$

We assume that $\mathcal{H}(E) < \infty$, else the equality is trivial.
We consider the inner approximation of G. Define, for every $n \geq 1$,

$$G_n = \{x \in G : \ d(x, G^c) > \tfrac{1}{n}\}$$

Then $d(G_n, G^c) > \dfrac{1}{n} > 0$ and $G_1 \subseteq G_2 \subseteq G_3 \subseteq \cdots$ with $G = \displaystyle\bigcup_{i=1}^{\infty} G_i$. Now consider the partitions $D_0 = G_1$, and $D_n = G_{n+1} \setminus G_n$ for $n \geq 1$. Then

$$D_n = \left\{x \in G : \ \tfrac{1}{n+1} < d(x, G^c) \leq \tfrac{1}{n}\right\}$$

Then, $G = \bigcup\limits_{i=0}^{\infty} D_i$ and $\{D_i\}$ is a collection of pairwise disjoint sets.
Then for $i+2 \leq j$, D_i and D_j are clearly separated by a positive distance, as

$$d(D_i, D_j) \geq \frac{1}{i+1} - \frac{1}{j} > 0$$

Then, each of the sets of the subfamily of $\{D_i\}$ with odd index, i.e., the subfamily $\{D_{2k-1}\}$, and with even index, i.e., the subfamily $\{D_{2k}\}$, are pairwise disjoint and at a positive distance.
As each D_i are disjoint, then for every $E \subseteq X$ the collection $\{E \cap D_i\}$ is pairwise disjoint, and

$$\bigcup_{i=1}^{k} (E \cap D_i) \subseteq E$$

Now using the result for disjoint sets and monotonicity,

$$\mathcal{H}\left(\bigcup_{i=1}^{k} (E \cap D_i)\right) = \sum_{i=1}^{k} \mathcal{H}(E \cap D_i) \leq \mathcal{H}(E)$$

Then for the subfamily $\{D_{2k-1}\}$ we get

$$\sum_{k=1}^{n} \mathcal{H}(E \cap D_{2k-1}) \leq \mathcal{H}(E)$$

and for the subfamily $\{D_{2k}\}$ we get

$$\sum_{k=0}^{n} \mathcal{H}(E \cap D_{2k}) \leq \mathcal{H}(E)$$

Then,

$$\sum_{k=0}^{2n} \mathcal{H}(E \cap D_k) \leq 2\mathcal{H}(E) < \infty$$

Also,

$$G \setminus G_n = \bigcup_{i=n}^{\infty} D_i$$

Thus,

$$E \cap (G \setminus G_n) = E \cap \bigcup_{i=n}^{\infty} D_i = \bigcup_{i=n}^{\infty} (E \cap D_i)$$

Continued...

Then, by sub-additivity,

$$\mathcal{H}\big(E \cap (G \setminus G_n)\big) \leq \sum_{i=n}^{\infty} \mathcal{H}(E \cap D_i)$$

Now the series $\sum_{i=n}^{\infty} \mathcal{H}(E \cap D_i) < \infty$, thus it is convergent and all terms are positive, so the tail converges to 0. Hence

$$\lim_{n\to\infty} \mathcal{H}\big(E \cap (G \setminus G_n)\big) = \lim_{n\to\infty} \sum_{i=n}^{\infty} \mathcal{H}(E \cap D_i) = 0$$

As $d(G_n, G^c) > \dfrac{1}{n} > 0$, the sets $E \cap G_n$ and $E \cap G^c$ are disjoint and have a positive distance. Also, $(E \cap G_n) \cup (E \cap G^c) \subseteq E$. Thus,

$$\mathcal{H}(E \cap G_n) + \mathcal{H}(E \cap G^c) = \mathcal{H}\big((E \cap G_n) \cup (E \cap G^c)\big) \leq \mathcal{H}(E)$$

Also,

$$G = G_n \cup (G \setminus G_n) \implies E \cap G = (E \cap G_n) \cup \big(E \cap (G \setminus G_n)\big)$$

Then, by sub-additivity,

$$\mathcal{H}(E \cap G) \leq \mathcal{H}(E \cap G_n) + \mathcal{H}\big(E \cap (G \setminus G_n)\big)$$

Thus,

$$\mathcal{H}(E \cap G) + \mathcal{H}(E \cap G^c) \leq \mathcal{H}(E \cap G_n) + \mathcal{H}(E \cap G^c) + \mathcal{H}\big(E \cap (G \setminus G_n)\big)$$

But from the positive separation of G_n and G^c, we already proved,

$$\mathcal{H}(E \cap G_n) + \mathcal{H}(E \cap G^c) \leq \mathcal{H}(E)$$

Therefore,

$$\mathcal{H}(E \cap G) + \mathcal{H}(E \cap G^c) \leq \mathcal{H}(E) + \mathcal{H}\big(E \cap (G \setminus G_n)\big)$$

Since

$$\lim_{n\to\infty} \mathcal{H}\big(E \cap (G \setminus G_n)\big) = 0$$

Letting $n \to \infty$ yields,

$$\mathcal{H}(E \cap G) + \mathcal{H}(E \cap G^c) \leq \mathcal{H}(E)$$

Combining this with sub-additivity,

$$\mathcal{H}(E) \leq \mathcal{H}(E \cap G) + \mathcal{H}(E \cap G^c)$$

we obtain,

$$\mathcal{H}(E) = \mathcal{H}(E \cap G) + \mathcal{H}(E \cap G^c)$$

Thus, all open sets are well separating with respect to $\mathcal{H}$.
Therefore, every Borel set is well separating with respect to $\mathcal{H}$. □

Continued...

Definition 24 (Hausdorff Dimension)**.** Let (X, d) be any metric space. Fix the function h, such that

$$h : \mathcal{P}(X) \longrightarrow [0, \infty] \qquad ; \qquad h(E) = [\operatorname{diam}(E)]^{\alpha}$$

where for any set $A \subseteq X$ we define the diameter as

$$\operatorname{diam}(A) = \sup\{d(x, y) : x, y \in A\}$$

Take $\varepsilon > 0$ and set $E \subseteq X$. Let $\{E_j\}_{j=1}^{\infty}$ be a countable cover of E. Then we define,

$$\mathcal{H}_{\alpha}^{\varepsilon}(E) = \inf\left\{ \sum_{j=1}^{\infty} [\operatorname{diam}(E_j)]^{\alpha} \ : \ E \subseteq \bigcup_{j=1}^{\infty} E_j, \ \forall j, \ \operatorname{diam}(E_j) < \varepsilon \right\}$$

And therefore,

$$\mathcal{H}_{\alpha}(E) = \lim_{\varepsilon \to 0} \mathcal{H}_{\alpha}^{\varepsilon}(E)$$

The Hausdorff Dimension of a set $E \subseteq X$ is defined by;

$$\dim_{\mathcal{H}}(E) = \inf\{\alpha \geq 0 : \mathcal{H}_{\alpha}(E) = 0\} = \sup\{\alpha \geq 0 : \mathcal{H}_{\alpha}(E) = \infty\}$$

Remark 2 (Special cases of $\mathcal{H}_{\alpha}$)**.** Depending on the value of α we have the following:

(i) If $\alpha = 0$, then $\mathcal{H}_0$ is the counting measure.

(ii) If $\alpha = 1$, then $\mathcal{H}_1$ on $\mathbb{R}$ coincides with the Lebesgue measure (m), i.e., the length.

(iii) If $\alpha = n$, then $\mathcal{H}_n$ on $\mathbb{R}^n$ coincides with the n-dimensional Lebesgue measure (m_n) up-to a constant factor. Thus, $\mathcal{H}_n = c_n \cdot m_n$, where, $c_n = \dfrac{\pi^{n/2}}{2^n\,\Gamma(\frac{n}{2} + 1)}$.

(iv) If α is not an integer, then $\mathcal{H}_{\alpha}$ measures the α-dimensional size of fractals.

Remark 3 (Hausdorff Measure on [0,1])**.** Depending on the value of α we have the following:

(i) Consider the interval $I = [0, 1]$. We measure $\mathcal{H}_{\alpha}(I)$, where $\alpha > 1$. We fix $\alpha > 1$. The, for each integer n, we can cover I by using n adjacent closed intervals;

$$I_k = \left[\frac{k-1}{n}, \frac{k}{n}\right] \qquad \text{, where} \quad k = 1, 2, \ldots, n$$

Now, $\operatorname{diam}(I_k) = \frac{k}{n} - \frac{k-1}{n} = \frac{1}{n}$.

Thus, for any $\varepsilon \geq \frac{1}{n}$, the cover $\{I_k\}_{k=1}^{n}$ can be considered. Thus,

$$\mathcal{H}_\alpha^\varepsilon(I) \leq \sum_{k=1}^{n} [\operatorname{diam}(I_k)]^\alpha = n \cdot \left(\frac{1}{n}\right)^\alpha = n^{1-\alpha}$$

As $\mathcal{H}_\alpha^\varepsilon(I)$ is non-increasing in ε, the limit as $\varepsilon \longrightarrow 0$, for every n is,

$$\mathcal{H}_\alpha(I) = \lim_{\varepsilon \to 0} \mathcal{H}_\alpha^\varepsilon(I) < n^{1-\alpha}$$

As, $\alpha > 1$, we have, $1 - \alpha < 0$, so, $n^{1-\alpha} \longrightarrow 0$, as $n \longrightarrow \infty$. Then,

$$\mathcal{H}_\alpha(I) \leq \lim_{n \to \infty} n^{1-\alpha} = 0$$

As, Hausdorff measure is non-negative, then, $\mathcal{H}_\alpha(I) = 0$.

(ii) Consider the interval $I = [0, 1]$ and $\alpha = 1$, then we claim, $\mathcal{H}_1(I) = 1$.
Fix $\varepsilon > 0$, and we choose positive integer $n > \frac{1}{\varepsilon}$. Then we can partition I into n equal sub-intervals;

$$I_k = \left[\frac{k-1}{n}, \frac{k}{n}\right] \qquad \text{, where} \quad k = 1, 2, \ldots, n$$

Now, for each I_k, we have $\operatorname{diam}(I_k) = \frac{k}{n} - \frac{k-1}{n} = \frac{1}{n} < \varepsilon$, and $\{I_k\}_{k=1}^{n}$ covers I. Then,

$$\mathcal{H}_1^\varepsilon(I) \leq \sum_{k=1}^{n} [\operatorname{diam}(I_k)] = n \cdot \left(\frac{1}{n}\right) = 1$$

As, this holds for all $\varepsilon > 0$, then as $\varepsilon \longrightarrow 0$, we have, $\mathcal{H}_1(I) \leq 1$.
Now, Let $\varepsilon > 0$, and $\{E_i\}_{i \in \Lambda}$ be any cover of I, with $\operatorname{diam}(E_i) < \varepsilon$, for all $i \in \Lambda$. By compactness of I, we get a finite sub-cover, say, $\{E_{i_1}, E_{i_2}, \ldots, E_{i_m}\}$ and for each k, we define,

$$J_k = [\inf E_{i_k}, \sup E_{i_k}]$$

Then, $\operatorname{diam}(E_{i_k}) = |J_k|$, where $|J_k|$ represents the length of J_k and $I \subseteq \bigcup_{k=1}^{m} J_k$.

Then, $|\bigcup_{k=1}^{m} J_k| \leq \sum_{k=1}^{m} |J_k| \leq \sum_{k=1}^{m} \operatorname{diam}(E_{i_k})$.

As the length of the cover is at-least 1, hence, $1 \leq \sum_{k=1}^{m} \operatorname{diam}(E_{i_k})$.

This holds for every finite sub-cover and every cover with sets of diameter less than ε. Thus, $\mathcal{H}_1^{\varepsilon}(I) \geq 1$, for all $\varepsilon > 0$.
As $\varepsilon \longrightarrow 0$, we have, $\mathcal{H}_1(I) \geq 1$. Therefore, $\mathcal{H}_1(I) = 1$.

Definition 25 (Semi-ring)**.** Let X be an arbitrary set. A collection $\mathcal{S} \subseteq \mathcal{P}(X)$ is said to be a semi-ring if:

(i) $\varnothing \in \mathcal{S}$.

(ii) If $A, B \in \mathcal{S}$, then $A \cap B \in \mathcal{S}$.

(iii) If $A, B \in \mathcal{S}$, then there exists a finite collection of disjoint sets $R_i \in \mathcal{S}$, $i = 1, \dots, m$, such that

$$A \setminus B = \bigcup_{i=1}^{m} R_i$$

Definition 26 (Pre-measure)**.** Let $\mathcal{P}$ be a semi-ring and Let $\mu_0 : \mathcal{P} \to [0, \infty]$ be a function such that:

(i) $\mu_0(\varnothing) = 0$.

(ii) If $P \in \mathcal{P}$ and $P = \bigcup_{j=1}^{\infty} P_j$, where each $P_j \in \mathcal{P}$ for $j \geq 1$, and the family $\{P_j\}$ is pairwise disjoint, then

$$\mu_0(P) = \sum_{j=1}^{\infty} \mu_0(P_j)$$

Then μ_0 is a pre-measure on the semiring $\mathcal{P}$.

Lemma 2 (Monotonicity of the pre-measure)**.** *If $P \subseteq Q$ and $P, Q \in \mathcal{P}$, then by the property of the semiring $\mathcal{P}$,*

$$Q \setminus P = \bigcup_{j=1}^{m} R_j$$

where each $R_j \in \mathcal{P}$ and the R_j are pairwise disjoint. Hence

$$Q = P \cup (Q \setminus P) = P \cup \left(\bigcup_{j=1}^{m} R_j \right)$$

and therefore,

$$\mu_0(Q) = \mu_0(P) + \sum_{j=1}^{m} \mu_0(R_j) \geq \mu_0(P) \implies \mu_0(P) \leq \mu_0(Q)$$

Thus μ_0 is monotone.

Lemma 3 (Tiling Lemma)**.** *If* $P, Q_1, Q_2, \ldots, Q_m \in \mathcal{P}$*, then*

$$P \setminus \left(\bigcup_{j=1}^{m} Q_j \right) = \bigcup_{i=1}^{n} R_i$$

where the sets $R_i \in \mathcal{P}$ *and are disjoint.*

Proof. The base case for $m = 1$ is true by definition of $\mathcal{P}$.
Assume that the claim is true for $m = k$, i.e.,

$$P \setminus \left(\bigcup_{j=1}^{k} Q_j \right) = \left(\bigcup_{l=1}^{t} \widetilde{R}_l \right)$$

where $\widetilde{R}_l \in \mathcal{P}$ are disjoint. Then, for $m = k + 1$,

$$\begin{aligned} P \setminus \left(\bigcup_{j=1}^{k+1} Q_j \right) &= \left[P \setminus \left(\bigcup_{j=1}^{k} Q_j \right) \right] \setminus Q_{k+1} \\ &= \left(\bigcup_{l=1}^{t} \widetilde{R}_l \right) \setminus Q_{k+1} \\ &= \bigcup_{l=1}^{t} \left(\widetilde{R}_l \setminus Q_{k+1} \right) \\ &= \bigcup_{l=1}^{t} \left(\bigcup_{s=1}^{m} R'_s \right) \\ &= \bigcup_{i=1}^{n} R_i \end{aligned}$$

where $R'_s, R_i \in \mathcal{P}$ are disjoint. □

Lemma 4 (Finite Covering Lemma)**.** *If* $Q_1, \ldots, Q_m \in \mathcal{P}$ *are pairwise disjoint and* $Q_j \subseteq P \in \mathcal{P}$ *for each* j*, then,*

$$\mu_0(P) \geq \sum_{j=1}^{m} \mu_0(Q_j)$$

Proof. If $Q_1, Q_2, \ldots, Q_m$ are disjoint subsets of $P \in \mathcal{P}$, then

$$P = \left(\bigcup_{j=1}^{m} Q_j \right) \cup \left[P \setminus \left(\bigcup_{j=1}^{m} Q_j \right) \right]$$

By the tiling lemma (Lemma 3),

$$P = \Big(\bigcup_{j=1}^{m} Q_j \Big) \cup \Big(\bigcup_{i=1}^{n} R_i \Big)$$

where $Q_j, R_i \in \mathcal{P}$ and every set in the union is disjoint from each other. Hence

$$\mu_0(P) = \sum_{j=1}^{m} \mu_0(Q_j) + \sum_{i=1}^{n} \mu_0(R_i) \implies \mu_0(P) \geq \sum_{j=1}^{m} \mu_0(Q_j)$$

□

Lemma 5 (Countable Covering Lemma)**.** *If $\{Q_j\}$ is a countable collection of disjoint sets and $Q_j \subseteq P \in \mathcal{P}$, then*

$$\mu_0(P) \geq \sum_{j=1}^{\infty} \mu_0(Q_j)$$

Proof. If $\{Q_j\}$ is a countable collection of disjoint subsets of $P \in \mathcal{P}$, then by Finite Covering Lemma (Lemma 4),

$$\mu_0(P) \geq \sum_{j=1}^{m} \mu_0(Q_j) \quad \text{for every } m$$

Taking the limit $m \to \infty$ we obtain

$$\mu_0(P) \geq \sum_{j=1}^{\infty} \mu_0(Q_j)$$

□

Lemma 6 (General Covering Lemma)**.** *Let $P,\ Q_j \in \mathcal{P}$, such that $P \subseteq \bigcup_{j=1}^{\infty} Q_j$, then,*

$$\mu_0(P) \leq \sum_{j=1}^{\infty} \mu_0(Q_j)$$

Proof. Let $P,\ Q_j \in \mathcal{P}$, such that $P \subseteq \bigcup_{j=1}^{\infty} Q_j$, then, Let $Q'_j := P \cap Q_j \in \mathcal{P}$ by the property of $\mathcal{P}$. Also $Q'_j \subseteq Q_j$, and then by monotonicity $\mu_0(Q'_j) \leq \mu_0(Q_j)$.
Then, we have

$$P = \bigcup_{j=1}^{\infty} Q'_j = \bigcup_{j=1}^{\infty} \Big(Q'_j \setminus \bigcup_{i=1}^{j-1} Q'_i \Big) = \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{n_j} R_{j,k}$$

using the tiling lemma (Lemma 3), where $R_{j,k} \in \mathcal{P}$ are pairwise disjoint.
Therefore,

$$\mu_0(P) = \sum_{j=1}^{\infty}\sum_{k=1}^{n_j} \mu_0(R_{j,k}) = \sum_{j=1}^{\infty}\Big(\sum_{k=1}^{n_j} \mu_0(R_{j,k})\Big) \leq \sum_{j=1}^{\infty} \mu_0(Q'_j) \leq \sum_{j=1}^{\infty} \mu_0(Q_j).$$

Thus, we have our result;

$$\mu_0(P) \leq \sum_{j=1}^{\infty} \mu_0(Q_j)$$

□

Definition 27 (Outer Measure Induced by a Premeasure)**.** Let $\mathcal{P}$ be a semiring on X and Let μ_0 be a premeasure on $\mathcal{P}$. For every set $E \subseteq X$, define

$$\mu^*(E) = \inf\left\{\sum_{j=1}^{\infty} \mu_0(P_j) : P_j \in \mathcal{P}, ; E \subseteq \bigcup_{j=1}^{\infty} P_j\right\}$$

Then, the set function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ defined above is an outer measure induced by the premeasure μ_0.

Note*: This is precisely the standard outer-measure construction from countable covers. The proof is identical to the proof that the infimum construction yields an outer measure.

Theorem 17 (Carathéodory Existence Theorem)**.** *Let $\mathcal{P}$ be a semi-ring and μ_0 be a pre-measure on $\mathcal{P}$. Then there exists a measure μ on $\sigma(\mathcal{P})$, such that $\mu(P) = \mu_0(P)$, for every $P \in \mathcal{P}$.*

Proof. Let μ^* be the outer measure induced by μ_0 as in the previous definition. For any $P \in \mathcal{P}$, we have,

$$\mu^*(P) \leq \mu_0(P)$$

since P covers itself.
If, $P \subseteq \bigcup_{j=1}^{\infty} P_j$ with $P_j \in \mathcal{P}$, then by the General covering lemma (Lemma 6), $\mu_0(P) \leq \sum_{j=1}^{\infty} \mu_0(P_j)$. Hence

$$\mu_0(P) \leq \inf \sum_{j=1}^{\infty} \mu_0(P_j) = \mu^*(P) \implies \mu_0(P) \leq \mu^*(P)$$

Thus, we have $\mu_0(P) = \mu^*(P)$ for all $P \in \mathcal{P}$.
Lastly, we show that $\mathcal{P}$ is well separating with respect to μ^*. Let $P \in \mathcal{P}$ and $\forall E \subseteq X$, by sub-additivity, we get

$$\mu^*(E) \leq \mu^*(E \cap P) + \mu^*(E \cap P^c)$$

Fix $\varepsilon > 0$ and Let $E \subseteq \bigcup_{i=1}^{\infty} P_i$ where $P_i \in \mathcal{P}$. By the definition of outer measure, we may choose such a cover with

$$\sum_{i=1}^{\infty} \mu_0(P_i) \leq \mu^*(E) + \varepsilon$$

Now, each $P_i \in \mathcal{P}$. By the Tiling Lemma (Lemma 3), for every $i \geq 1$ we may write,

$$P_i \cap P = \bigcup_{k=1}^{m_i} A_{i,k}, \qquad P_i \setminus P = \bigcup_{\ell=1}^{n_i} B_{i,\ell}$$

where all $A_{i,k}, B_{i,\ell} \in \mathcal{P}$ are pairwise disjoint.
Since

$$P_i = (P_i \cap P) \cup (P_i \setminus P)$$

the pre-measure property gives,

$$\mu_0(P_i) = \sum_{k=1}^{m_i} \mu_0(A_{i,k}) + \sum_{\ell=1}^{n_i} \mu_0(B_{i,\ell})$$

Moreover, the family $\{A_{i,k}\}$ covers $E \cap P$ and the family $\{B_{i,\ell}\}$ covers $E \cap P^c$. Therefore,

$$\mu^*(E \cap P) \leq \sum_{i=1}^{\infty} \sum_{k=1}^{m_i} \mu_0(A_{i,k})$$

and,

$$\mu^*(E \cap P^c) \leq \sum_{i=1}^{\infty} \sum_{\ell=1}^{n_i} \mu_0(B_{i,\ell})$$

Adding these inequalities yields

$$\begin{aligned}\mu^*(E \cap P) + \mu^*(E \cap P^c) &\leq \sum_{i=1}^{\infty} \left(\sum_{k=1}^{m_i} \mu_0(A_{i,k}) + \sum_{\ell=1}^{n_i} \mu_0(B_{i,\ell}) \right) \\ &= \sum_{i=1}^{\infty} \mu_0(P_i) \leq \mu^*(E) + \varepsilon\end{aligned}$$

Letting $\varepsilon \to 0$ yields,

$$\mu^*(E) \geq \mu^*(E \cap P) + \mu^*(E \cap P^c)$$

Combining with the earlier inequality gives, for all $P \in \mathcal{P}$ and all $E \subseteq X$,

$$\mu^*(E) = \mu^*(E \cap P) + \mu^*(E \cap P^c)$$

Thus $\mathcal{P}$ is well separating with respect to μ^*.
By Carathéodory's Theorem (Theorem 12), the collection

$$\mathcal{M} = \{A \subseteq X : \ A \text{ is } \mu^*\text{-measurable}\}$$

forms a σ-algebra.
Since every $P \in \mathcal{P}$ is μ^*-measurable, we have

$$\mathcal{P} \subseteq \mathcal{M}$$

Therefore,

$$\sigma(\mathcal{P}) \subseteq \mathcal{M}$$

Define,

$$\mu := \mu^*|_{\sigma(\mathcal{P})}$$

Since $\sigma(\mathcal{P}) \subseteq \mathcal{M}$, the restriction μ is a measure on $\sigma(\mathcal{P})$. Finally, for every $P \in \mathcal{P}$,

$$\mu(P) = \mu^*(P) = \mu_0(P)$$

Hence,

$$\mu|_{\mathcal{P}} = \mu_0$$

Therefore μ is a measure on $\sigma(\mathcal{P})$ extending the pre-measure μ_0. □

Theorem 18 (Ring Generated by a Semi-ring)**.** *Let $\mathcal{P}$ be a semi-ring on X, and Let μ_0 be a pre-measure on $\mathcal{P}$. Define,*

$$\mathcal{R} = \left\{ \bigcup_{j=1}^{n} P_j : P_j \in \mathcal{P} \text{ are pairwise disjoint}, \ n \in \mathbb{N} \right\}$$

Then, $\mathcal{R}$ is a ring containing $\mathcal{P}$, and,

$$\sigma(\mathcal{R}) = \sigma(\mathcal{P})$$

Moreover, the function μ_0 extends to a pre-measure $\widetilde{\mu}_0$ on $\mathcal{R}$ by

$$\widetilde{\mu}_0\left(\bigcup_{j=1}^{n} P_j\right) = \sum_{j=1}^{n} \mu_0(P_j)$$

whenever $P_1, \dots, P_n \in \mathcal{P}$ are pairwise disjoint.

Proof. First, we show that $\mathcal{R}$ is a ring containing $\mathcal{P}$. Let $R_1, R_2 \in \mathcal{R}$ with

$$R_1 = \bigcup_{i=1}^{n} P_i, \qquad R_2 = \bigcup_{j=1}^{m} Q_j, \qquad P_i, Q_j \in \mathcal{P}$$

where the families $\{P_i\}_{i=1}^{n}$ and $\{Q_j\}_{j=1}^{m}$ are pairwise disjoint.

- Then,
  $$R_1 \cap R_2 = \left(\bigcup_{i=1}^{n} P_i\right) \bigcap \left(\bigcup_{j=1}^{m} Q_j\right) = \bigcup_{i=1}^{n}\bigcup_{j=1}^{m}(P_i \cap Q_j)$$
  Since $P_i, Q_j \in \mathcal{P}$ and $\mathcal{P}$ is a semiring, we have $P_i \cap Q_j \in \mathcal{P}$. Moreover, the family $\{P_i \cap Q_j\}_{i,j}$ is pairwise disjoint, because the P_i are pairwise disjoint and the Q_j are pairwise disjoint. Since the indices i, j run over finite sets, $R_1 \cap R_2$ is a finite union of disjoint sets from $\mathcal{P}$.
  Hence $R_1 \cap R_2 \in \mathcal{R}$.

- Also,
  $$R_1 \setminus R_2 = \left(\bigcup_{i=1}^{n} P_i\right) \setminus \left(\bigcup_{j=1}^{m} Q_j\right) = \bigcup_{i=1}^{n}\Big(P_i \setminus \bigcup_{j=1}^{m} Q_j\Big)$$
  By the tiling lemma (Lemma 3), for each $i = 1, \ldots, n$, there exist finitely many pairwise disjoint sets $R_{i,1}, \ldots, R_{i,k} \in \mathcal{P}$, such that,
  $$P_i \setminus \left(\bigcup_{j=1}^{m} Q_j\right) = \bigcup_{\ell=1}^{k} R_{i,\ell}$$
  Therefore,
  $$R_1 \setminus R_2 = \bigcup_{i=1}^{n}\bigcup_{\ell=1}^{k} R_{i,\ell}$$
  This is a finite union of sets from $\mathcal{P}$. Moreover, the union is pairwise disjoint because the P_i are pairwise disjoint and each family $\{R_{i,\ell}\}_{\ell=1}^{k_i}$ is pairwise disjoint inside P_i.
  Thus, $R_1 \setminus R_2 \in \mathcal{R}$ and similarly, $R_2 \setminus R_1 \in \mathcal{R}$.

- Finally,
  $$R_1 \cup R_2 = R_1 \cup (R_2 \setminus R_1)$$
  As $R_2 \setminus R_1 \in \mathcal{R}$ and $R_1 \in \mathcal{R}$, and R_1 and $R_2 \setminus R_1$ are disjoint, their union is again a finite union of pairwise disjoint sets from $\mathcal{P}$.
  Therefore, $R_1 \cup R_2 \in \mathcal{R}$.

Hence $\mathcal{R}$ is closed under finite unions, intersections and set difference.
Therefore $\mathcal{R}$ is a ring.
Now, $\mathcal{P} \subseteq \mathcal{R}$, because if $P \in \mathcal{P}$, then

$$P = \bigcup_{j=1}^{1} P \quad \Longrightarrow \quad P \in \mathcal{R}$$

By construction of $\mathcal{R}$, we have,

$$\mathcal{P} \subseteq \mathcal{R} \subseteq \sigma(\mathcal{P})$$

Let $\sigma(\mathcal{R})$ be the σ-algebra generated by $\mathcal{R}$ and hence the smallest σ-algebra containing $\mathcal{R}$.
Then, as $\sigma(\mathcal{R})$ is the smallest σ-algebra containing $\mathcal{R}$ and $\sigma(\mathcal{P})$ is a σ-algebra containing $\mathcal{R}$, we have

$$\sigma(\mathcal{R}) \subseteq \sigma(\mathcal{P})$$

Again, $\sigma(\mathcal{P})$ is the smallest σ-algebra containing $\mathcal{P}$ and $\sigma(\mathcal{R})$ is a σ-algebra containing $\mathcal{P}$ (as $\mathcal{R}$ is generated using $\mathcal{P}$), then we have

$$\sigma(\mathcal{P}) \subseteq \sigma(\mathcal{R})$$

Therefore,

$$\sigma(\mathcal{P}) = \sigma(\mathcal{R})$$

Now we define the extension of μ_0 to $\mathcal{R}$.
If $R \in \mathcal{R}$, then by definition of $\mathcal{R}$ there exist pairwise disjoint sets $P_1, \ldots, P_n \in \mathcal{P}$ such that,

$$R = \bigcup_{j=1}^{n} P_j$$

Define,

$$\widetilde{\mu}_0(R) = \widetilde{\mu}_0\left(\bigcup_{j=1}^{n} P_j\right) = \sum_{j=1}^{n} \mu_0(P_j)$$

We must show that this definition is well-defined.
Suppose that $R \in \mathcal{R}$ has two representations:

$$R = \bigcup_{j=1}^{n} P_j = \bigcup_{k=1}^{m} Q_k,$$

where $P_j, Q_k \in \mathcal{P}$ and both families are pairwise disjoint. We need to show that,

$$\sum_{j=1}^{n} \mu_0(P_j) = \sum_{k=1}^{m} \mu_0(Q_k)$$

For each fixed j, and for each fixed k, since

$$P_j \subseteq R = \bigcup_{k=1}^{m} Q_k, \quad Q_k \subseteq R = \bigcup_{j=1}^{n} P_j$$

Thus, we have,

$$P_j = P_j \cap R = P_j \cap \left(\bigcup_{k=1}^{m} Q_k \right) = \bigcup_{k=1}^{m} (P_j \cap Q_k)$$

and,

$$Q_k = Q_k \cap R = Q_k \cap \left(\bigcup_{j=1}^{n} P_j \right) = \bigcup_{j=1}^{n} (Q_k \cap P_j)$$

Since $P_j, Q_k \in \mathcal{P}$ and $\mathcal{P}$ is a semi-ring, $P_j \cap Q_k \in \mathcal{P}$. Also, for fixed j and fixed k, the sets $P_j \cap Q_k$ are pairwise disjoint because the P_j and the Q_k are pairwise disjoint. Therefore, by the pre-measure property of μ_0,

$$\mu_0(P_j) = \sum_{k=1}^{m} \mu_0(P_j \cap Q_k), \quad \mu_0(Q_k) = \sum_{j=1}^{n} \mu_0(Q_k \cap P_j)$$

Summing over $j = 1, \ldots, n$ and $k = 1, \ldots, m$ respectively, we get

$$\sum_{j=1}^{n} \mu_0(P_j) = \sum_{j=1}^{n} \sum_{k=1}^{m} \mu_0(P_j \cap Q_k), \quad \sum_{k=1}^{m} \mu_0(Q_k) = \sum_{k=1}^{m} \sum_{j=1}^{n} \mu_0(Q_k \cap P_j)$$

But,

$$Q_k \cap P_j = P_j \cap Q_k \implies \sum_{j=1}^{n} \sum_{k=1}^{m} \mu_0(P_j \cap Q_k) = \sum_{k=1}^{m} \sum_{j=1}^{n} \mu_0(Q_k \cap P_j)$$

Therefore,

$$\sum_{j=1}^{n} \mu_0(P_j) = \sum_{k=1}^{m} \mu_0(Q_k)$$

Thus $\widetilde{\mu}_0$ is well-defined.
Next, we show that $\widetilde{\mu}_0$ extends μ_0. If $P \in \mathcal{P}$, then $P \in \mathcal{R}$ and

$$P = \bigcup_{j=1}^{1} P$$

Hence, by definition,

$$\widetilde{\mu}_0(P) = \mu_0(P)$$

Therefore,

$$\widetilde{\mu}_0|_{\mathcal{P}} = \mu_0$$

Finally, we show that $\widetilde{\mu}_0$ is a pre-measure on $\mathcal{R}$. Let $\{R_j\}_{j=1}^{\infty}$ be a pairwise disjoint collection of sets in $\mathcal{R}$ such that

$$R = \bigcup_{j=1}^{\infty} R_j \in \mathcal{R}$$

We must prove that,

$$\widetilde{\mu}_0(R) = \sum_{j=1}^{\infty} \widetilde{\mu}_0(R_j)$$

Since $R \in \mathcal{R}$, there exist pairwise disjoint sets $P_1, \ldots, P_n \in \mathcal{P}$ such that

$$R = \bigcup_{i=1}^{n} P_i$$

Also, for each j, since $R_j \in \mathcal{R}$, there exist pairwise disjoint sets

$$Q_{j,1}, \ldots, Q_{j,m_j} \in \mathcal{P}$$

such that

$$R_j = \bigcup_{k=1}^{m_j} Q_{j,k}.$$

Since the sets R_j are pairwise disjoint, the family $\{Q_{j,k}\}_{j,k}$ is pairwise disjoint. Now,

$$R = \bigcup_{j=1}^{\infty} R_j = \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{m_j} Q_{j,k}$$

Therefore, for each $i = 1, \ldots, n$,

$$P_i = P_i \cap R = P_i \cap \left(\bigcup_{j=1}^{\infty} \bigcup_{k=1}^{m_j} Q_{j,k} \right) = \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{m_j} (P_i \cap Q_{j,k})$$

Each set $P_i \cap Q_{j,k}$ belongs to $\mathcal{P}$, since $\mathcal{P}$ is closed under finite intersections. Also, for fixed i, the sets $P_i \cap Q_{j,k}$ are pairwise disjoint because the sets $Q_{j,k}$ are pairwise disjoint. Hence, by the pre-measure property of μ_0 on $\mathcal{P}$,

$$\mu_0(P_i) = \sum_{j=1}^{\infty} \sum_{k=1}^{m_j} \mu_0(P_i \cap Q_{j,k})$$

Summing over $i = 1, \ldots, n$, we get

$$\widetilde{\mu}_0(R) = \sum_{i=1}^{n} \mu_0(P_i)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{\infty} \sum_{k=1}^{m_j} \mu_0(P_i \cap Q_{j,k})$$

$$= \sum_{j=1}^{\infty} \sum_{k=1}^{m_j} \sum_{i=1}^{n} \mu_0(P_i \cap Q_{j,k}).$$

For fixed j and k, since $Q_{j,k} \subseteq R$, we have

$$Q_{j,k} = Q_{j,k} \cap R = Q_{j,k} \cap \left(\bigcup_{i=1}^{n} P_i \right) = \bigcup_{i=1}^{n} (Q_{j,k} \cap P_i)$$

The sets $Q_{j,k} \cap P_i$ are pairwise disjoint and belong to $\mathcal{P}$. Therefore,

$$\mu_0(Q_{j,k}) = \sum_{i=1}^{n} \mu_0(Q_{j,k} \cap P_i)$$

Hence,

$$\widetilde{\mu}_0(R) = \sum_{j=1}^{\infty} \sum_{k=1}^{m_j} \mu_0(Q_{j,k}) = \sum_{j=1}^{\infty} \widetilde{\mu}_0(R_j)$$

Therefore, $\widetilde{\mu}_0$ is countably additive on $\mathcal{R}$ whenever the countable disjoint union belongs to $\mathcal{R}$. Also,

$$\widetilde{\mu}_0(\varnothing) = 0$$

because $\varnothing \in \mathcal{P}$ and $\mu_0(\varnothing) = 0$.
Hence $\widetilde{\mu}_0$ is a pre-measure on the ring $\mathcal{R}$ extending μ_0 from $\mathcal{P}$. □

Definition 28 (Monotone Class)**.** A collection $\mathcal{M} \subseteq \mathcal{P}(X)$ is called a monotone class if:

(i) whenever $E_1 \subseteq E_2 \subseteq \cdots$ and $E_j \in \mathcal{M}$ for all j, then $\displaystyle\bigcup_{j=1}^{\infty} E_j \in \mathcal{M}$

(ii) whenever $E_1 \supseteq E_2 \supseteq \cdots$ and $E_j \in \mathcal{M}$ for all j, then $\displaystyle\bigcap_{j=1}^{\infty} E_j \in \mathcal{M}$

Theorem 19 (Monotone Class Theorem)**.** *Let $\mathcal{R}$ be a ring of subsets of X. Assume that X is σ-finite with respect to $\mathcal{R}$, i.e.,*

$$X = \bigcup_{j=1}^{\infty} R_j$$

where $R_j \in \mathcal{R}$. Then the monotone class generated by $\mathcal{R}$ coincides with the σ-algebra generated by $\mathcal{R}$. That is,

$$\mathcal{M}(\mathcal{R}) = \sigma(\mathcal{R})$$

Proof. Let $\mathcal{M}$ be the least monotone class containing $\mathcal{R}$, and $\mathcal{A}$ be the least σ-algebra containing $\mathcal{R}$, i.e., $\mathcal{A} = \sigma(\mathcal{R})$ is the σ-algebra generated by $\mathcal{R}$.
By construction, $\mathcal{R} \subseteq \mathcal{M}$ and since every σ-algebra is a monotone class, and $\mathcal{A}$ contains $\mathcal{R}$, we have $\mathcal{M} \subseteq \sigma(\mathcal{R}) = \mathcal{A}$.
So it remains to show that $\mathcal{M}$ is a σ-algebra containing $\mathcal{R}$.

- Clearly $\varnothing \in \mathcal{P} \subseteq \mathcal{R} \subseteq \mathcal{M}$. So, $\varnothing \in \mathcal{M}$.
- By the σ-finiteness assumption, there exist sets $R_j \in \mathcal{R}$ such that,

  $$X = \bigcup_{j=1}^{\infty} R_j = \bigcup_{n=1}^{\infty} \left(\bigcup_{j=1}^{n} R_j \right) = \bigcup_{n=1}^{\infty} S_n$$

  Since $\mathcal{R}$ is a ring, each $S_n \in \mathcal{R}$. Hence $S_n \in \mathcal{M}$ for every n, and, $S_1 \subseteq S_2 \subseteq S_3 \subseteq \cdots$. Now, as $\mathcal{M}$ is a monotone class, it follows that,

  $$X \in \mathcal{M}$$

Define

$$\mathcal{M}' = \{A \in \mathcal{M} :\ A \cup R,\ A \cap R,\ A \setminus R,\ R \setminus A \in \mathcal{M} \text{ for all } R \in \mathcal{R}\}$$

Now, since $\mathcal{R} \subseteq \mathcal{M}$ and $\mathcal{R}$ is a ring, $\mathcal{R}$ is closed under finite unions, intersections, and set differences. Therefore, if $A \in \mathcal{R}$ and $R \in \mathcal{R}$, then, $A \cup R, \quad A \cap R, \quad A \setminus R, \quad R \setminus A$ are all in $\mathcal{R}$, and hence in $\mathcal{M}$. Therefore, we have $\mathcal{R} \subseteq \mathcal{M}'$. Also, by definition $\mathcal{M}' \subseteq \mathcal{M}$.
Let $\{E_j\}$ be an increasing sequence of sets in $\mathcal{M}'$, i.e., $E_1 \subseteq E_2 \subseteq \cdots$, and let

$$E = \bigcup_{j=1}^{\infty} E_j$$

Since each $E_j \in \mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{M}$ is a monotone class, we have $E \in \mathcal{M}$.
Fix $R \in \mathcal{R}$. Since $E_j \in \mathcal{M}'$, we have, for all j,

- $E_j \cup R \in \mathcal{M}$ and $E_j \cup R$ is increasing to $E \cup R$ in j in $\mathcal{M}$. Hence,

  $$\bigcup_{j=1}^{\infty} (E_j \cup R) \in \mathcal{M} \implies \Big(\bigcup_{j=1}^{\infty} E_j\Big) \cup R \in \mathcal{M} \implies E \cup R \in \mathcal{M}$$

- $E_j \cap R \in \mathcal{M}$ and $E_j \cap R$ is increasing to $E \cap R$ in j in $\mathcal{M}$. Hence,

$$\bigcup_{j=1}^{\infty}(E_j \cap R) \in \mathcal{M} \implies (\bigcup_{j=1}^{\infty} E_j) \cap R \in \mathcal{M} \implies E \cap R \in \mathcal{M}$$

- $E_j \setminus R \in \mathcal{M}$ and $(E_j \setminus R)$ is increasing to $E \setminus R$ in j in $\mathcal{M}$. Hence,

$$\bigcup_{j=1}^{\infty}(E_j \setminus R) \in \mathcal{M} \implies \Big(\bigcup_{j=1}^{\infty} E_j\Big) \setminus R \in \mathcal{M} \implies E \setminus R \in \mathcal{M}$$

- $R \setminus E_j \in \mathcal{M}$ and $(R \setminus E_j)$ is decreasing to $R \setminus E$ in j in $\mathcal{M}$. Hence,

$$\bigcap_{j=1}^{\infty}(R \setminus E_j) \in \mathcal{M} \implies R \setminus \left(\bigcup_{j=1}^{\infty} E_j\right) \in \mathcal{M} \implies R \setminus E \in \mathcal{M}$$

Therefore, $E = \bigcup_{j=1}^{\infty} E_j \in \mathcal{M}'$.
Let $\{E_j\}$ be a decreasing sequence of sets in $\mathcal{M}'$, i.e., $E_1 \supseteq E_2 \supseteq \cdots$, and let

$$E = \bigcap_{j=1}^{\infty} E_j$$

Since each $E_j \in \mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{M}$ is a monotone class, we have $E \in \mathcal{M}$.
Fix $R \in \mathcal{R}$. Since $E_j \in \mathcal{M}'$, we have, for all j,

- $E_j \cup R \in \mathcal{M}$ and $E_j \cup R$ is decreasing to $E \cup R$ in j in $\mathcal{M}$. Hence,

$$\bigcap_{j=1}^{\infty}(E_j \cup R) \in \mathcal{M} \implies (\bigcap_{j=1}^{\infty} E_j) \cup R \in \mathcal{M} \implies E \cup R \in \mathcal{M}$$

- $E_j \cap R \in \mathcal{M}$ and $E_j \cap R$ is decreasing to $E \cap R$ in j in $\mathcal{M}$. Hence,

$$\bigcap_{j=1}^{\infty}(E_j \cap R) \in \mathcal{M} \implies (\bigcap_{j=1}^{\infty} E_j) \cap R \in \mathcal{M} \implies E \cap R \in \mathcal{M}$$

- $E_j \setminus R \in \mathcal{M}$ and $(E_j \setminus R)$ is decreasing to $E \setminus R$ in j in $\mathcal{M}$. Hence,

$$\bigcap_{j=1}^{\infty}(E_j \setminus R) \in \mathcal{M} \implies \Big(\bigcap_{j=1}^{\infty} E_j\Big) \setminus R \in \mathcal{M} \implies E \setminus R \in \mathcal{M}$$

Continued...

- $R \setminus E_j \in \mathcal{M}$ and $(R \setminus E_j)$ is increasing to $R \setminus E$ in j in $\mathcal{M}$. Hence,

$$\bigcup_{j=1}^{\infty}(R \setminus E_j) \in \mathcal{M} \implies R \setminus \left(\bigcap_{j=1}^{\infty} E_j\right) \in \mathcal{M} \implies R \setminus E \in \mathcal{M}$$

Therefore, $E = \bigcap\limits_{j=1}^{\infty} E_j \in \mathcal{M}'$.
Therefore, $\mathcal{M}'$ is a monotone class containing $\mathcal{R}$, and $\mathcal{M}$ being the smallest monotone class containing $\mathcal{R}$, we get,

$$\mathcal{M} \subseteq \mathcal{M}'$$

Since by definition, $\mathcal{M}' \subseteq \mathcal{M}$, we conclude that, $\mathcal{M} = \mathcal{M}'$.
Therefore, for every $A \in \mathcal{M}$ and every $R \in \mathcal{R}$, $A \cup R, A \cap R, A \setminus R, R \setminus A \in \mathcal{M}$.
Now, define

$$\mathcal{M}'' = \{A \in \mathcal{M} :\ A \cup B,\ A \cap B,\ A \setminus B,\ B \setminus A \in \mathcal{M} \text{ for all } B \in \mathcal{M}\}$$

From, the previous result, as $\mathcal{M}' = \mathcal{M}$ and $\mathcal{R} \subseteq \mathcal{M}$, we get

$$\mathcal{R} \subseteq \mathcal{M}''$$

In a similar manner as above one proves that $\mathcal{M}''$ is a monotone class, simply by replacing $R \in \mathcal{R}$ by an arbitrary fixed $B \in \mathcal{M}$.
Thus, $\mathcal{M}''$ is a monotone class containing $\mathcal{R}$. Since $\mathcal{M}$ is the least monotone class containing $\mathcal{R}$, we have

$$\mathcal{M} \subseteq \mathcal{M}''$$

Since $\mathcal{M}'' \subseteq \mathcal{M}$ by definition, it follows that

$$\mathcal{M}'' = \mathcal{M}$$

Therefore,

$$\mathcal{M} = \mathcal{M}' = \mathcal{M}''$$

This concludes that $\mathcal{M}$ is closed under finite unions, finite intersections, and set differences.
Since $X \in \mathcal{M}$, for every $A \in \mathcal{M}$ we have,

$$A^c = X \setminus A \in \mathcal{M}$$

Thus, $\mathcal{M}$ is closed under complements.
Finally, Let $\{E_j\}$ is any countable collection of sets in $\mathcal{M}$, then

$$\bigcup_{j=1}^{\infty} E_j = \bigcup_{j=1}^{\infty}\Big(\bigcup_{k=1}^{j} E_k\Big) \qquad \text{and} \qquad \bigcap_{j=1}^{\infty} E_j = \bigcap_{j=1}^{\infty}\Big(\bigcap_{k=1}^{j} E_k\Big)$$

As $\mathcal{M}$ is a monotone class, then,

$$\{\bigcup_{k=1}^{j} E_k\}_{j\geq 1} \uparrow \text{ in } \mathcal{M} \implies \bigcup_{j=1}^{\infty}\Big(\bigcup_{k=1}^{j} E_k\Big) \in \mathcal{M} \implies \bigcup_{j=1}^{\infty} E_j \in \mathcal{M}$$

Similarly, the sequence

$$\{(\bigcap_{k=1}^{j} E_k\}_{j\geq 1} \text{ in } \mathcal{M} \implies \bigcap_{j=1}^{\infty}\Big(\bigcap_{k=1}^{j} E_k\Big) \in \mathcal{M} \implies \bigcap_{j=1}^{\infty} E_j \in \mathcal{M}$$

Therefore $\mathcal{M}$ is a σ-algebra containing $\mathcal{R}$.
Since $\sigma(\mathcal{R})$ is the smallest σ-algebra containing $\mathcal{R}$, we have, $\mathcal{A} = \sigma(\mathcal{R}) \subseteq \mathcal{M}$.
Earlier we already showed that, $\mathcal{M} \subseteq \sigma(\mathcal{R})$.
Consequently, $\mathcal{A} = \sigma(\mathcal{R}) = \mathcal{M}$.
Hence, we have proved that,

$$\mathcal{M}(\mathcal{R}) = \sigma(\mathcal{R})$$

□

Theorem 20 (Carathéodory Uniqueness Theorem). *Let $\mathcal{P}$ be a semiring on X and Let μ_0 be a premeasure on $\mathcal{P}$. Then there exists a measure μ on the σ-algebra $\sigma(\mathcal{P})$ generated by $\mathcal{P}$ such that $\mu|_{\mathcal{P}} = \mu_0$.*

Furthermore, if μ_0 is σ-finite on $\mathcal{P}$, then μ is unique, i.e., if $X = \bigcup_{j=1}^{\infty} P_j$ with $P_j \in \mathcal{P}$ and $\mu_0(P_j) < \infty$ for all j, then the extension μ is unique.

Proof. By the existence part, (Theorem 17) we already have a measure μ on $\sigma(\mathcal{P})$ with $\mu|_{\mathcal{P}} = \mu_0$.
Let μ_1 and μ_2 be two measures on $\mathcal{A} = \sigma(\mathcal{P})$, such that

$$\mu_1\big|_{\mathcal{P}} = \mu_2\big|_{\mathcal{P}} = \mu_0$$

Define the ring generated by the semi-ring $\mathcal{P}$ by,

$$\mathcal{R} = \left\{\bigcup_{j=1}^{n} P_j : P_j \in \mathcal{P} \text{ are pairwise disjoint, } n \in \mathbb{N}\right\}$$

By the theorem 18 on the ring generated by a semi-ring, $\mathcal{R}$ is a ring containing $\mathcal{P}$, and,

$$\sigma(\mathcal{R}) = \sigma(\mathcal{P})$$

Since μ_1 and μ_2 agree on $\mathcal{P}$, finite additivity implies that they agree on $\mathcal{R}$. Indeed, if

$$R = \bigcup_{j=1}^{n} P_j$$

where $P_j \in \mathcal{P}$ are pairwise disjoint, then,

$$\mu_1(R) = \sum_{j=1}^{n} \mu_1(P_j) = \sum_{j=1}^{n} \mu_0(P_j) = \sum_{j=1}^{n} \mu_2(P_j) = \mu_2(R)$$

Therefore,

$$\mu_1|_{\mathcal{R}} = \mu_2|_{\mathcal{R}}$$

By σ-finiteness, there exist sets $P_j \in \mathcal{P}$ such that, $X = \bigcup_{j=1}^{\infty} P_j$, where $P_j \in \mathcal{P}$ and $\mu_0(P_j) < \infty$, for all $j \geq 1$.
These sets need not be disjoint. Therefore, we replace them by disjoint sets in the ring $\mathcal{R}$. Define

$$R_1 = P_1, \quad \text{and} \quad R_j = P_j \setminus \bigcup_{i=1}^{j-1} P_i, \quad \forall j \geq 2$$

By the tiling lemma (Lemma 3), each set $P_j \setminus \bigcup_{i=1}^{j-1} P_i$ can be written as a finite disjoint union of sets from $\mathcal{P}$. Hence, $R_j \in \mathcal{R}$, for every $j \geq 1$.
The sets R_j are pairwise disjoint, and

$$X = \bigcup_{j=1}^{\infty} R_j.$$

Also, since $R_j \subseteq P_j$, and since μ_1 and μ_2 agree with μ_0 on $\mathcal{P}$ and agree on $\mathcal{R}$, we have

$$\mu_1(R_j) = \mu_2(R_j) < \infty.$$

More precisely,

$$\mu_1(R_j) = \mu_2(R_j) \leq \mu_0(P_j) < \infty$$

Now define,

$$\mathcal{M} = \{A \in \mathcal{A} : \ \mu_1(A \cap R_j) = \mu_2(A \cap R_j), \ \forall j \geq 1\}$$

We show that $\mathcal{M}$ is a monotone class.

- Let $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$ be an increasing sequence of sets in $\mathcal{M}$, and define $E = \bigcup_{k=1}^{\infty} E_k$.
  Then for each fixed $j \geq 1$, $E_k \cap R_j$ increases to $E \cap R_j$ in k. Since μ_1 and μ_2 are measures, continuity from below gives

  $$\mu_1(E \cap R_j) = \lim_{k \to \infty} \mu_1(E_k \cap R_j)$$

Since $E_k \in \mathcal{M}$ for every k, we have,

$$\mu_1(E_k \cap R_j) = \mu_2(E_k \cap R_j)$$

Hence,

$$\mu_1(E \cap R_j) = \lim_{k\to\infty} \mu_2(E_k \cap R_j)$$

Again by continuity from below,

$$\lim_{k\to\infty} \mu_2(E_k \cap R_j) = \mu_2(E \cap R_j)$$

Therefore, for every $j \geq 1$,

$$\mu_1(E \cap R_j) = \mu_2(E \cap R_j)$$

Thus, $E \in \mathcal{M}$.

- Similarly, Let $E_1 \supset E_2 \supset E_3 \supset \cdots$ be a decreasing sequence of sets in $\mathcal{M}$, and define, $E = \bigcap_{j=1}^{\infty} E_j$.

  Then for each fixed $j \geq 1$, $E_k \cap R_j$ decreases to $E \cap R_j$ in k. In order to use continuity from above, we need the first set to have finite measure. Since $E_1 \cap R_j \subseteq R_j$ and $\mu_1(R_j) = \mu_2(R_j) < \infty$, we have,

  $$\mu_1(E_1 \cap R_j) < \infty \quad \text{and} \quad \mu_2(E_1 \cap R_j) < \infty$$

  Therefore, by continuity from above,

  $$\mu_1(E \cap R_j) = \lim_{k\to\infty} \mu_1(E_k \cap R_j)$$

  Since each $E_k \in \mathcal{M}$,

  $$\mu_1(E_k \cap R_j) = \mu_2(E_k \cap R_j)$$

  Hence

  $$\mu_1(E \cap R_j) = \lim_{k\to\infty} \mu_2(E_k \cap R_j).$$

  Again by continuity from above,

  $$\lim_{k\to\infty} \mu_2(E_k \cap R_j) = \mu_2(E \cap R_j)$$

  Therefore, for every $j \geq 1$,

  $$\mu_1(E \cap R_j) = \mu_2(E \cap R_j)$$

  Thus, $E \in \mathcal{M}$.

Thus, $\mathcal{M}$ is a monotone class.
Next, we show that $\mathcal{R} \subseteq \mathcal{M}$.
Let $R \in \mathcal{R}$. Since $\mathcal{R}$ is a ring and $R_j \in \mathcal{R}$, we have, $R \cap R_j \in \mathcal{R}$, for every $j \geq 1$. Since μ_1 and μ_2 agree on $\mathcal{R}$, it follows that, for every $j \geq 1$,

$$\mu_1(R \cap R_j) = \mu_2(R \cap R_j)$$

Therefore,

$$R \in \mathcal{M}$$

Thus,

$$\mathcal{R} \subseteq \mathcal{M}$$

Since $\mathcal{M}$ is a monotone class containing $\mathcal{R}$, the monotone class generated by $\mathcal{R}$ is contained in $\mathcal{M}$. That is,

$$\mathcal{M}(\mathcal{R}) \subseteq \mathcal{M}$$

By the Monotone Class Theorem for rings (Theorem 19), using the σ-finiteness assumption,

$$\mathcal{M}(\mathcal{R}) = \sigma(\mathcal{R})$$

Hence, $\sigma(\mathcal{R}) \subseteq \mathcal{M}$.
But, $\sigma(\mathcal{R}) = \sigma(\mathcal{P}) = \mathcal{A}$. So, $\mathcal{A} \subseteq \mathcal{M}$.
Since by definition $\mathcal{M} \subseteq \mathcal{A}$, we conclude

$$\mathcal{M} = \mathcal{A}$$

Now Let $A \in \mathcal{A}$. Since $X = \bigcup_{j=1}^{\infty} R_j$ and the sets R_j are pairwise disjoint, we have,

$$A = A \cap X = A \cap \bigcup_{j=1}^{\infty} R_j = \bigcup_{j=1}^{\infty} (A \cap R_j)$$

where the sets $A \cap R_j$ are pairwise disjoint.
Since $A \in \mathcal{A} = \mathcal{M}$, for every $j \geq 1$, we have, $\mu_1(A \cap R_j) = \mu_2(A \cap R_j)$.
Therefore, by countable additivity,

$$\mu_1(A) = \sum_{j=1}^{\infty} \mu_1(A \cap R_j) = \sum_{j=1}^{\infty} \mu_2(A \cap R_j) = \mu_2(A)$$

Hence, for every $A \in \mathcal{A}$,

$$\mu_1(A) = \mu_2(A)$$

Therefore,

$$\mu_1 = \mu_2 \quad \text{on} \quad \sigma(\mathcal{P})$$

Thus the measure extending μ_0 from $\mathcal{P}$ to $\sigma(\mathcal{P})$ is unique. $\square$

Theorem 21 (Functional Generation of Borel Measurable Functions)**.** *Let (X, d) be a metric space and $\mathcal{F}$ be a family of functions from X to $\mathbb{R}$, such that:*

(i) $\mathcal{F} \supseteq C(X)$, where $C(X)$ is the space of continuous functions on X.

(ii) $\mathcal{F}$ is closed under addition and multiplication by constants.

(iii) $\mathcal{F}$ is closed under pointwise limits.

(iv) If $f \in \mathcal{F}$, then $|f| \in \mathcal{F}$.

Then $\mathcal{F}$ contains all Borel measurable functions.

Proof. Let $\mathcal{B}(X)$ be the Borel σ-algebra on X. Note that $\mathcal{B}(X)$ contains all closed sets. Define

$$\mathcal{M} = \{E \in \mathcal{B}(X) : \mathbb{1}_E \in \mathcal{F}\}$$

where $\mathbb{1}_E$ denotes the indicator (characteristic) function of E.
Clearly $\mathcal{M}$ is nonempty, since $\varnothing \in \mathcal{B}(X)$ and $\mathbb{1}_\varnothing = 0 \in C(X) \subseteq \mathcal{F}$. Thus $\varnothing \in \mathcal{M}$.
We first show that every closed set $K \subseteq X$ belongs to $\mathcal{M}$. Let K be closed; if $K = \varnothing$ this is already done, so assume $K \neq \varnothing$. Define

$$f(x) = d(x, K) = \inf_{y \in K} d(x, y), \qquad x \in X$$

Then, $\forall z \in X$, for all $\varepsilon > 0$, there exists $y \in K$ such that, $f(z) + \varepsilon \geq d(z, y)$.
Also, $\forall x \in X,\ \forall y \in K$,

$$f(x) \leq d(x, y) \leq d(x, z) + d(z, y) \leq d(x, z) + f(z) + \varepsilon$$

hence

$$f(x) \leq f(z) + d(x, z)$$

Similarly,

$$f(z) \leq f(x) + d(x, z)$$

Therefore,

$$|f(x) - f(z)| \leq d(x, z), \qquad \forall x, z \in X$$

Thus f is 1-Lipschitz, and in particular continuous. Thus, $f \in C(X) \subseteq \mathcal{F}$.
Now, $\forall x \notin K,\ f(x) > 0$, as K is closed, and $\forall x \in K,\ f(x) = 0$. Thus

$$f(x) = \begin{cases} 0, & x \in K, \\ > 0, & x \notin K \end{cases}$$

Now, for all $n \geq 1$, consider the construction of the function

$$f_n(x) = (1 - nf(x))_+ = \max\{0,\, 1 - nf(x)\}$$

Continued...

Since $f \in \mathcal{F}$ and $\mathcal{F}$ is closed under multiplication by constants and addition, we have, $1 - nf \in \mathcal{F}$. By assumption, if $g \in \mathcal{F}$, then $|g| \in \mathcal{F}$. Hence, $|1 - nf| \in \mathcal{F}$. Now observe that,

$$f_n(x) = (1 - nf(x))_+ = \max\{0, 1 - nf(x)\} = \frac{(1 - nf(x)) + |1 - nf(x)|}{2}$$

Therefore $f_n \in \mathcal{F}$ for every $n \geq 1$.
Then,

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \notin K \\ 1, & x \in K \end{cases} = \mathbb{1}_K$$

As $f_n(x) \in \mathcal{F}$ and $\mathcal{F}$ is closed under limits, so $\mathbb{1}_K \in \mathcal{F}$.
Therefore, $K \in \mathcal{M}$. Since K was an arbitrary closed subset of X, it follows that $\mathcal{M}$ contains all closed sets.
Now, Let $\{E_n\}$ be a family of sets in $\mathcal{M}$, such that

$$E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots \qquad \text{and} \qquad E = \bigcup_{n=1}^{\infty} E_n$$

Then $\mathbb{1}_{E_n} \in \mathcal{F}$ and also

$$\lim_{n\to\infty} \mathbb{1}_{E_n} = \mathbb{1}_E$$

Since $\mathcal{F}$ is closed under pointwise limits, we obtain,

$$\mathbb{1}_E \in \mathcal{F} \implies E \in \mathcal{M} \implies \bigcup_{n=1}^{\infty} E_n \in \mathcal{M}$$

Similarly, Let $\{B_n\}$ be a family of sets in $\mathcal{M}$, such that

$$B_1 \supset B_2 \supset B_3 \supset \cdots \qquad \text{and} \qquad B = \bigcap_{n=1}^{\infty} B_n$$

Then $\mathbb{1}_{B_n} \in \mathcal{F}$ and also

$$\lim_{n\to\infty} \mathbb{1}_{B_n} = \mathbb{1}_B$$

Since $\mathcal{F}$ is closed under pointwise limits, so $\mathbb{1}_B \in \mathcal{F} \implies B \in \mathcal{M} \implies \bigcap_{n=1}^{\infty} B_n \in \mathcal{M}$.
Therefore, $\mathcal{M}$ is a monotone class.
Now we show that $\mathcal{M}$ contains all Borel sets.
First note that $\mathcal{M}$ is closed under complements. Indeed, if $E \in \mathcal{M}$, then $\mathbb{1}_E \in \mathcal{F}$. Since $1 \in C(X) \subseteq \mathcal{F}$ and $\mathcal{F}$ is closed under addition and multiplication by constants,

$$\mathbb{1}_{E^c} = 1 - \mathbb{1}_E \in \mathcal{F}$$

Hence $E^c \in \mathcal{M}$.

Next, $\mathcal{M}$ is closed under finite unions. If $E, F \in \mathcal{M}$, then $\mathbb{1}_E, \mathbb{1}_F \in \mathcal{F}$. Since $\mathcal{F}$ is closed under addition, multiplication by constants, and absolute values, we have,

$$\mathbb{1}_{E\cup F} = \max\{\mathbb{1}_E, \mathbb{1}_F\} = \frac{\mathbb{1}_E + \mathbb{1}_F + |\mathbb{1}_E - \mathbb{1}_F|}{2} \in \mathcal{F}$$

Thus $E \cup F \in \mathcal{M}$.

Similarly,

$$\mathbb{1}_{E\cap F} = \min\{\mathbb{1}_E, \mathbb{1}_F\} = \frac{\mathbb{1}_E + \mathbb{1}_F - |\mathbb{1}_E - \mathbb{1}_F|}{2} \in \mathcal{F}$$

so $E \cap F \in \mathcal{M}$.

Let $\mathcal{R}$ be the ring generated by the closed subsets of X.

Since $\mathcal{M}$ is closed under finite unions, finite intersections, and complements, and since $\mathcal{M}$ contains every closed set, it follows that, $\mathcal{M}$ contains the ring $\mathcal{R}$ generated by the closed subsets of X. Therefore,

$$\mathcal{R} \subseteq \mathcal{M}$$

Since X is a metric space, it is closed in itself. Hence $X \in \mathcal{R}$. Therefore,

$$X = \bigcup_{n=1}^{\infty} X$$

so the σ-finiteness hypothesis of the Monotone Class Theorem (Theorem 19) is satisfied.

Since $\mathcal{M}$ is a monotone class containing $\mathcal{R}$, the Monotone Class Theorem (Theorem 19) gives

$$\sigma(\mathcal{R}) \subseteq \mathcal{M}$$

Since $\mathcal{R}$ contains every closed set,

$$\mathcal{B}(X) \subseteq \sigma(\mathcal{R})$$

On the other hand, every element of $\mathcal{R}$ is Borel, so,

$$\sigma(\mathcal{R}) \subseteq \mathcal{B}(X)$$

Therefore,

$$\sigma(\mathcal{R}) = \mathcal{B}(X)$$

Therefore,

$$\mathcal{B}(X) \subseteq \mathcal{M}$$

Thus, for every Borel set $E \subseteq X$, we have,

$$\mathbb{1}_E \in \mathcal{F}$$

Next we show that this implies all Borel measurable functions from $X \to \mathbb{R}$ are in $\mathcal{F}$.
Let f be any Borel measurable function from X to $\mathbb{R}$.
Consider $j = -n^2, -n^2+1, \dots, n^2-1$.
Let

$$f_n(x) = \begin{cases} n, & f(x) \geq n, \\ \dfrac{j}{n}, & f(x) \in \left[\dfrac{j}{n}, \dfrac{j+1}{n}\right), \\ -n, & f(x) < -n \end{cases},$$

Then each of the sets

$$\begin{aligned} E_1 &= \{f(x) \geq n\} \\ E_{n,j} &= \{f(x) \in \left[\frac{j}{n}, \frac{j+1}{n}\right)\} \\ E_3 &= \{f(x) < -n\} \end{aligned}$$

are Borel sets, and by the previous part we have shown that for every Borel set $E \subseteq X$ the indicator $\mathbb{1}_E \in \mathcal{F}$.
Hence

$$f_n(x) = \sum_{j=-n^2}^{n^2-1} \frac{j}{n} \mathbb{1}_{E_{n,j}} + n\,\mathbb{1}_{E_1} + (-n)\,\mathbb{1}_{E_3}$$

Since $\mathbb{1}_{E_{n,j}}, \mathbb{1}_{E_1}, \mathbb{1}_{E_3} \in \mathcal{F}$ and $\mathcal{F}$ is closed under addition and multiplication by constants, we have $f_n \in \mathcal{F}$.
Now, for each $x \in X$, we have,

$$|f(x)| < n \implies |f_n(x) - f(x)| \leq \frac{1}{n}$$

Thus, $f_n(x) \to f(x)$ pointwise as $n \to \infty$.
As $\mathcal{F}$ is closed under pointwise limits, $f \in \mathcal{F}$.
Therefore, $\mathcal{F}$ contains all Borel measurable functions from X to $\mathbb{R}$. $\square$

Theorem 22 (Regularity of Borel Measure)**.** *Let X be a metric space and $\mathcal{A}$ be the Borel σ-algebra on X. Let μ be a finite measure on $\mathcal{A}$, i.e. $\mu(X) < \infty$. For every $\varepsilon > 0$ and every $E \in \mathcal{A}$ there exist a closed set F and an open set G with*

$$F \subseteq E \subseteq G \qquad \text{and} \qquad \mu(G \setminus F) < \varepsilon$$

Proof. Define the collection of 'approximable' sets as:

$$\mathcal{A}' = \big\{E \in \mathcal{A} : \forall \varepsilon > 0, \exists \text{ closed } F \text{ and open } G \text{ with } F \subseteq E \subseteq G \text{ and } \mu(G \setminus F) < \varepsilon\big\}$$

We first show that every open set G belongs to $\mathcal{A}'$.
Let $E = G \subseteq X$ be open. If $G = \varnothing$, it is trivial. If not, define for each $n \in \mathbb{N}$ the inner-approximation sets:

$$G_n = \{x \in G : \, d(x, G^c) \geq \frac{1}{n}\}$$

Take $x \in G \setminus G_n$. Then $d(x, G^c) < \frac{1}{n}$.
So, $\exists\, y \in G^c \cap B_{1/n}(x)$, such that $d(x, y) < \frac{1}{n}$.
Let $d(x, y) = \delta$ and $r = \frac{1}{n} - \delta > 0$. Then, $\forall z \in B_r(x)$, we have,

$$d(y, z) \leq d(x, z) + d(x, y) = r + \delta = \frac{1}{n}$$

Therefore, as $y \in G^c$, and $d(z, y) < \frac{1}{n}$, it follows that $d(z, G^c) < \frac{1}{n}$.
Thus,

$$z \notin G_n \implies B_r(x) \subseteq G_n^c \implies G_n^c \quad \text{is open}$$

Therefore, G_n is closed.

Also, $\{G_n\}$ is an increasing sequence and $G = \bigcup_{n=1}^{\infty} G_n$, as if $x \in G$, then since G is open,

$$d(x, G^c) > 0$$

Choose n sufficiently large so that,

$$\frac{1}{n} < d(x, G^c)$$

Then, $x \in G_n$. Hence,

$$G = \bigcup_{n=1}^{\infty} G_n$$

Since μ is a finite measure, $\mu(G) < \infty$, and by continuity from below,

$$\mu(G) = \lim_{n \to \infty} \mu(G_n)$$

Then,

$$\mu(G \setminus G_n) = \mu(G) - \mu(G_n) \longrightarrow 0 \quad \text{as } n \to \infty$$

Hence, for every $\varepsilon > 0$, there exists n such that, $\mu(G \setminus G_n) < \varepsilon$, this means, we can approximate E from below by G_n for large n.
Setting $F = G_n$, we obtain, $F \subseteq G$.
Then $F = G_n \subseteq G$, where G_n is closed (left) and G is open (right), and $\mu(G \setminus G_n) \to 0$.
Therefore, $\forall \varepsilon > 0$, $\exists$ closed G_n and open G such that $G_n \subseteq G$ and $\mu(G \setminus G_n) < \varepsilon$.
Hence, every open set G is in $\mathcal{A}'$.
Now we show that $\mathcal{A}'$ is a σ-algebra:

- As the empty set $\varnothing$ is both open and closed, we consider $F = \varnothing$ and $G = \varnothing$. So we have,
$$\varnothing \subseteq \varnothing \subseteq \varnothing \qquad \text{and} \qquad \mu(\varnothing \setminus \varnothing) = 0$$
Thus $\varnothing \in \mathcal{A}'$.

- Let $E \in \mathcal{A}'$, then $\forall \varepsilon > 0$ there exist closed F and open G with $F \subseteq E \subseteq G$ and $\mu(G \setminus F) < \varepsilon$. Taking complements, we have,
$$G^c \subseteq E^c \subseteq F^c$$
where G^c is closed and F^c is open. Since $F^c \setminus G^c = G \setminus F$, we have,
$$\mu(F^c \setminus G^c) = \mu(G \setminus F) < \varepsilon$$
Thus, $E^c \in \mathcal{A}'$.

- *Let $E_1, E_2 \in \mathcal{A}'$. Then, by definition, we have $\forall \varepsilon > 0$:
  - $\exists$ closed F_1 and open G_1 such that $F_1 \subseteq E_1 \subseteq G_1$ with $\mu(G_1 \setminus F_1) < \dfrac{\varepsilon}{2}$.
  - $\exists$ closed F_2 and open G_2 such that $F_2 \subseteq E_2 \subseteq G_2$ with $\mu(G_2 \setminus F_2) < \dfrac{\varepsilon}{2}$.

  Now, Let $F := F_1 \cup F_2$, $E := E_1 \cup E_2$ and $G := G_1 \cup G_2$.
  Then, F, being union of two closed sets, is closed; and G, being union of two open sets, is open. Then,
$$F \subseteq E \subseteq G$$
  where F is closed and G is open.
  Now,
$$G \setminus F = (G_1 \cup G_2) \setminus (F_1 \cup F_2) \subseteq (G_1 \setminus F_1) \cup (G_2 \setminus F_2)$$
  Then
$$\begin{aligned}\mu(G \setminus F) &\leq \mu\big((G_1 \cup G_2) \setminus (F_1 \cup F_2)\big)\\ &\leq \mu(G_1 \setminus F_1) + \mu(G_2 \setminus F_2) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon\end{aligned}$$
  Thus, for $E = E_1 \cup E_2$ there exist closed F and open G such that $F \subseteq E \subseteq G$ and $\mu(G \setminus F) < \varepsilon$.
  Hence $E_1 \cup E_2 \in \mathcal{A}'$.

- *Let $\{E_j\}$ be a collection of sets in $\mathcal{A}'$ (not necessarily disjoint).
  Then, for each j, there exist closed F_j and open G_j such that
$$F_j \subseteq E_j \subseteq G_j, \qquad \text{and} \qquad \mu(G_j \setminus F_j) < \frac{\varepsilon}{2^{j+1}}$$

Let

$$F' = \bigcup_{j=1}^{\infty} F_j, \qquad G = \bigcup_{j=1}^{\infty} G_j, \qquad E = \bigcup_{j=1}^{\infty} E_j$$

Then G, being a countable union of open sets, is open, however, F' is not necessarily closed.
Clearly $F' \subseteq E \subseteq G$ and

$$G \setminus F' = \bigcup_{j=1}^{\infty} (G_j \setminus F_j)$$

Hence, by the subadditivity of μ,

$$\mu(G \setminus F') \leq \sum_{j=1}^{\infty} \mu(G_j \setminus F_j) \leq \sum_{j=1}^{\infty} \frac{\varepsilon}{2^{j+1}} = \frac{\varepsilon}{2}$$

However, since $F' = \bigcup\limits_{j=1}^{\infty} F_j$ and μ is finite, we have

$$\mu(\bigcup_{j=1}^{n} F_j) \to \mu(F') \qquad \text{as} \qquad n \to \infty$$

Therefore there exists $n_0 \in \mathbb{N}$ such that

$$\mu(F' \setminus \bigcup_{j=1}^{n_0} F_j) < \frac{\varepsilon}{2}$$

Let

$$F = \bigcup_{j=1}^{n_0} F_j$$

Then F is closed since it is a finite union of closed sets.
Then $F \subseteq F' \subseteq E \subseteq G$, and:

$$\mu(G \setminus F) \leq \mu(G \setminus F') + \mu(F' \setminus F) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Thus, for $E = \bigcup\limits_{j=1}^{\infty} E_j$, we have: for every $\varepsilon > 0$ there exist a closed set $F = \bigcup\limits_{j=1}^{n_0} F_j$ and an open set $G = \bigcup\limits_{j=1}^{\infty} G_j$ such that $F \subseteq E \subseteq G$ and $\mu(G \setminus F) < \varepsilon$.

Thus, $E = \bigcup\limits_{j=1}^{\infty} E_j \in \mathcal{A}'$.

Continued...

Therefore, $\mathcal{A}'$ is a σ-algebra.
All open sets are contained in $\mathcal{A}'$ and $\mathcal{A}' \subseteq \mathcal{A}$.
As, $\mathcal{A}'$ is a σ-algebra, and since $\mathcal{A}$ is the minimal σ-algebra containing the open sets (i.e. the Borel σ-algebra), $\mathcal{A} \subseteq \mathcal{A}'$.
Combined with $\mathcal{A}' \subseteq \mathcal{A}$ by definition, we have $\mathcal{A} = \mathcal{A}'$.
Therefore, every Borel set is approximable.
This proves that every finite Borel measure on a metric space is regular. □

Definition 29 (Pointwise Convergence)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. A sequence of functions $\{f_n\}$ is said to converge to f pointwise on X if $\forall \varepsilon > 0$ and $\forall x \in X$, there exists $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \varepsilon$ whenever $n \geq N$.

Definition 30 (Pointwise Convergence Almost Everywhere)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. A sequence of functions $\{f_n\}$ converges to f pointwise almost everywhere if $\exists$ a measurable set $E \subseteq X$ with measure $\mu(E) = 0$ such that $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in X \setminus E$.

Definition 31 (Uniform Convergence)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. A sequence of functions $\{f_n\}$ converges to f uniformly on X if, for every $\varepsilon > 0$, there exists an integer N such that for all $n \geq N$, we have, $|f_n(x) - f(x)| < \varepsilon$ for all $x \in X$.

Definition 32 (Convergence in Measure)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. A sequence of measurable functions $\{f_n\}$ is said to converge in measure to a function f if for every $\varepsilon > 0$, the measure of the set where f_n differs from f by more than ε vanishes as $n \to \infty$, i.e., $\lim\limits_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}) = 0$.

Theorem 23 (Uniform Implies Pointwise Almost Everywhere Convergence)**.** *If a sequence of functions $\{f_n\}$ converges to f uniformly on X, then $\{f_n\}$ converges to f pointwise almost everywhere in X.*

Proof. Suppose $\{f_n\}$ converges uniformly to f on X. By definition, for every $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ and for all $x \in X$:

$$|f_n(x) - f(x)| < \varepsilon$$

Fix an arbitrary $x_0 \in X$. For the same $\varepsilon > 0$, we can choose the same N such that for all $n \geq N$:

$$|f_n(x_0) - f(x_0)| < \varepsilon$$

This implies that $\lim\limits_{n\to\infty} f_n(x_0) = f(x_0)$ for every $x_0 \in X$. Since the sequence converges pointwise for all $x \in X$, it converges for all x except possibly on a set of measure zero.
Thus, $\{f_n\}$ converges to f pointwise almost everywhere in X. □

Theorem 24 (Uniform Implies Convergence in Measure)**.** *If a sequence of functions $\{f_n\}$ converges to f uniformly on X, then $\{f_n\}$ converges to f in measure.*

Proof. Suppose $\{f_n\}$ converges to f uniformly on X. To show $\{f_n\}$ converges to f in measure, we must show that for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}) = 0.$$

Given any $\varepsilon > 0$, since $\{f_n\}$ converges to f uniformly, there exists an integer N such that for all $n \geq N$:

$$\sup_{x\in X} |f_n(x) - f(x)| < \varepsilon.$$

This implies that for all $n \geq N$ and all $x \in X$, the inequality $|f_n(x) - f(x)| < \varepsilon$ holds. Consequently, the set of points where the difference is at least ε is empty:

$$\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\} = \varnothing \quad \text{for all } n \geq N.$$

Since the measure of the empty set is zero, we have:

$$\mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}) = 0 \quad \text{for all } n \geq N.$$

Taking the limit as $n \to \infty$, we obtain:

$$\lim_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| \geq \varepsilon\}) = 0.$$

Thus, f_n converges to f in measure. □

Theorem 25 (Convergence a.e. Implies Convergence in Measure)**.** *If a sequence of functions $\{f_n\}$ converges to f pointwise almost everywhere on X, and if, $\mu(X) < \infty$ then $\{f_n\}$ converges to f in measure.*

Proof. Assume that $f_n(x)$ converges to $f(x)$ for all $x \in X \setminus E$, where $\mu(E) = 0$.
Fix $\varepsilon > 0$ and let

$$B_n = \{x \in X : \exists m \geq n,\ |f_m(x) - f(x)| \geq \varepsilon\} = \bigcup_{m\geq n} \{x : |f_m(x) - f(x)| \geq \varepsilon\}$$

Then $\{B_n\}$ is a non-increasing (nested) sequence of measurable sets:

$$X \supseteq B_1 \supseteq B_2 \supseteq \cdots$$

Hence $\bigcap\limits_{n\geq 1} B_n$ is measurable.
Since $\mu(X) < \infty$ we have $\mu(B_n) \leq \mu(X) < \infty$ for every n, so $\lim_{n\to\infty} \mu(B_n)$ exists and is finite. Moreover

$$\bigcap_{n\geq 1} B_n \subseteq E$$

because if $x \notin E$ then $f_n(x) \longrightarrow f(x)$, so there is N with $|f_m(x) - f(x)| < \varepsilon$ for all $m \geq N$, hence $x \notin B_N$. Therefore,

$$\mu\Big(\bigcap_{n\geq 1} B_n\Big) \leq \mu(E) = 0$$

By continuity of measure from above,

$$\mu\Big(\bigcap_{n\geq 1} B_n\Big) = \lim_{n\to\infty} \mu(B_n) = 0$$

Finally, for each n,

$$\{x : |f_n(x) - f(x)| \geq \varepsilon\} \subseteq B_n$$

Thus,

$$\mu(\{x : |f_n - f| \geq \varepsilon\}) \leq \mu(B_n) \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty$$

Hence f_n converges to f in measure. □

Theorem 26 (Subsequential Extraction from Convergence in Measure)**.** *If a sequence of functions $\{f_n\}$ converges to f in measure, then there exists a subsequence $\{f_{n_k}\}$ such that f_{n_k} converges to f almost everywhere.*

Proof. For every $m \in \mathbb{N}$ choose $N(m)$ with $N(m) > N(m-1)$ such that

$$\mu\Big(\{\, |f_{N(m)} - f| > \frac{1}{m}\,\}\Big) < 2^{-m}$$

This is possible because f_n converges to f in measure, so,

$$\mu(\{\, |f_n - f| > \frac{1}{m}\,\}) \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty$$

Define

$$E_m = \{\, x \in X : |f_{N(m)}(x) - f(x)| > \frac{1}{m}\,\}$$

If x belongs to only finitely many of the sets E_m, then $f_{N(m)}(x) \longrightarrow f(x)$. Thus the set where convergence does not happen is

$$E = \bigcap_{m\geq 1} \bigcup_{k\geq m} E_k \subseteq \bigcup_{k\geq m} E_k$$

the set of points that belong to infinitely many E_k.
For each m,

$$\mu\Big(\bigcup_{k\geq m} E_k\Big) \leq \sum_{k\geq m} \mu(E_k) \leq \sum_{k\geq m} 2^{-k} = 2^{-m+1} \longrightarrow 0 \quad \text{as} \quad m \longrightarrow \infty$$

Hence $\mu(E) = 0$.
This shows convergence pointwise almost everywhere. □

Continued...

Theorem 27 (Egorov's Theorem). *If a sequence of functions $\{f_n\}$ converges to f almost everywhere and $\mu(X) < \infty$, then $\forall \varepsilon > 0$, $\exists E \subseteq X$, with $\mu(E) < \varepsilon$, such that, $\{f_n\}$ converges to f uniformly on $X \setminus E$.*

Proof. Fix m and N, and consider the set

$$\begin{aligned} E_{m,N} &= \{x \in X : \ \exists n \geq N \text{ with } |f_n(x) - f(x)| \geq \frac{1}{m}\} \\ &= \bigcup_{n \geq N} \{x \in X : \ |f_n(x) - f(x)| \geq \frac{1}{m}\} \end{aligned}$$

For every fixed m,

$$\bigcap_N E_{m,N} \subseteq \{x \in X : \ f_n(x) \not\to f(x)\}$$

But by assumption (pointwise a.e.) $\mu(\{x : \ f_n(x) \not\to f(x)\}) = 0$.
Also, for fixed m, the sets $E_{m,N}$ are decreasing in N, and since $\mu(X) < \infty$ the limit $\lim_{N\to\infty} \mu(E_{m,N})$ exists and equals $\mu\big(\bigcap_N E_{m,N}\big) = 0$.
Hence for each m there exists $N(m)$ such that

$$\mu\big(E_{m,N(m)}\big) \leq \frac{\varepsilon}{2^m}$$

Let

$$E = \bigcup_{m=1}^{\infty} E_{m,N(m)}$$

Then by subadditivity,

$$\mu(E) \leq \sum_{m=1}^{\infty} \mu\big(E_{m,N(m)}\big) \leq \sum_{m=1}^{\infty} \frac{\varepsilon}{2^m} = \varepsilon$$

If $x \in X \setminus E$, then for each m we have $x \notin E_{m,N(m)}$, so,

$$|f_n(x) - f(x)| < \frac{1}{m} \qquad , \qquad \forall n \geq N(m)$$

Thus for every $\delta > 0$ choose m with $\frac{1}{m} < \delta$ and then for all $n > \max_{1\leq k\leq m} N(k)$ we have $|f_n(x) - f(x)| < \delta$.
Therefore, f_n converges to f uniformly on $X \setminus E$. □

Example 6 (Pointwise $\not\Longrightarrow$ Convergence in measure)
Consider $f_n : \mathbb{R} \to \mathbb{R}$, $f_n(x) = \dfrac{|x|}{n}$, with Lebesgue measure μ on $\mathbb{R}$. Then

$$f(x) = \lim_{n\to\infty} f_n(x) = 0 \quad \text{for every } x \in \mathbb{R}$$

Continued...

So, $f_n \longrightarrow f$ pointwise.
However, for any fixed $\varepsilon > 0$,

$$\{x : |f_n(x) - f(x)| > \varepsilon\} = \{x : |x| > n\varepsilon\}$$

and $\mu(\{x : |x| > n\varepsilon\}) = \infty$ for every n.
Hence $\mu(\{x : |f_n - f| > \varepsilon\}) \not\longrightarrow 0$, so f_n does not converge to f in measure.

Example 7 (Convergence in measure $\not\Longrightarrow$ Pointwise)
Consider $[0,1]$ with Lebesgue measure. Define a sequence of measurable sets by

$$\begin{aligned} &E_1 = [0,1], \\ &E_2 = [0, \frac{1}{2}], \quad E_3 = [\frac{1}{2}, 1], \\ &E_4 = [0, \frac{1}{4}], \quad E_5 = [\frac{1}{4}, \frac{1}{2}], \quad E_6 = [\frac{1}{2}, \frac{3}{4}], \quad E_7 = [\frac{3}{4}, 1], \ \ldots \end{aligned}$$

(so the intervals get smaller and exhaust dyadic subintervals). Each set E_n has Lebesgue measure $\mu(E_n) \longrightarrow 0$ as $n \longrightarrow \infty$, and every point $x \in [0,1]$ belongs to infinitely many of the sets E_n (and in fact, to infinitely many sets of arbitrarily small length).
Define $f_n = n\mathbb{1}_{E_n}$.
Then $\mu(\{x : |f_n(x) - 0| > \varepsilon\}) = \mu(E_n) \longrightarrow 0$ for each $\varepsilon > 0$, so $f_n \longrightarrow 0$ in measure.
But for every $x \in [0,1]$,

$$\limsup_{n\to\infty} f_n(x) = \infty$$

because x is in infinitely many E_n and on those indices $f_n(x) = n \longrightarrow \infty$. Thus there is no pointwise convergence (not even on any point) and in particular f_n does not converge pointwise to 0.

Example 8 (Pointwise $\not\Longrightarrow$ Uniform Convergence)
Consider the set $[0,1)$ and the sequence of functions $f_n(x) = x^n$.
Then $f_n(x) \longrightarrow 0$ for all $x \in [0,1)$. But this is not convergent uniformly.
Indeed,

$$\sup_{x\in[0,1)} |x^n - 0| = 1 \qquad \forall n$$

since values arbitrarily close to 1 are attained on $[0,1)$.
(Note*: Changing $[0,1)$ to $[0, 1-\varepsilon]$ makes it uniformly convergent.)

Theorem 28 (Monotone Convergence Theorem)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f_n : X \to [0,\infty]$ be a monotone increasing sequence of nonnegative measurable functions. Let*

$$f(x) := \lim_{n\to\infty} f_n(x) \quad (\textit{possibly } f(x) = \infty)$$

Then f is measurable and

$$\lim_{n\to\infty}\int_X f_n\,d\mu = \int_X f\,d\mu$$

Proof. Since $f_n \le f$ for every n, by monotonicity of the integral we have

$$\int_X f_n\,d\mu \le \int_X f\,d\mu \qquad \text{for all } n$$

and hence,

$$\lim_{n\to\infty}\int_X f_n\,d\mu \le \int_X f\,d\mu$$

Let a be any real number with

$$a < \int_X f\,d\mu$$

Since

$$\int_X f\,d\mu = \sup\Big\{\int_X g\,d\mu : g \text{ simple},\ 0 \le g \le f\Big\}$$

there exists a nonnegative simple measurable function $g = \sum_{j=1}^{N} c_j\chi_{E_j}$, where $c_j > 0$ and the sets E_j are measurable and pairwise disjoint, such that $0 \le g \le f$ and

$$\int_X g\,d\mu > a$$

Let $q < 1$ be such that

$$q\int_X g\,d\mu = \sum_{j=1}^{N}(qc_j)\,\mu(E_j) > a$$

Let $c'_j = qc_j$. On E_j we have $0 < c'_j < f$.
Then, $\forall x \in E_j$,

$$\lim_{n\to\infty} f_n(x) = f(x) > c'_j$$

Hence $\exists n_0 \in \mathbb{N}$ such that $\forall n \ge n_0$, $f_n(x) \ge c'_j$.
Define

$$E_j^{(n)} = \{x \in E_j : f_n(x) > c'_j\}$$

Then $E_j^{(n)}$ increases in n for fixed j and

$$E_j = \bigcup_{n=1}^{\infty} E_j^{(n)}$$

Continued...

Thus, by continuity from below,

$$\mu(E_j) = \lim_{n\to\infty} \mu\big(E_j^{(n)}\big)$$

Let

$$g_n = \sum_{j=1}^{N} c_j'\, \chi_{E_j^{(n)}} \le f_n$$

We chose q so that $\displaystyle\sum_{j=1}^{N}(qc_j)\,\mu(E_j) > a$. By the convergence of $\mu(E_j^{(n)})$ to $\mu(E_j)$, there exists $N' \in \mathbb{N}$ such that for all $n \ge N'$,

$$\int_X f_n\, d\mu \ge \int_X g_n\, d\mu = \sum_{j=1}^{N} c_j'\, \mu\big(E_j^{(n)}\big) > a$$

Therefore for all $n \ge N'$ (for some $N' \in \mathbb{N}$) we have $\displaystyle\int_X f_n\, d\mu > a$.
Thus, for all sufficiently large n, we have,

$$\int_X f_n\, d\mu > a$$

Hence,

$$\liminf_{n\to\infty} \int_X f_n\, d\mu \ge a$$

Since $a < \int_X f\, d\mu$ was arbitrary, it follows that,

$$\liminf_{n\to\infty} \int_X f_n\, d\mu \ge \int_X f\, d\mu$$

On the other hand, we already showed that,

$$\int_X f_n\, d\mu \le \int_X f\, d\mu \qquad \text{for all } n$$

and therefore,

$$\limsup_{n\to\infty} \int_X f_n\, d\mu \le \int_X f\, d\mu$$

Hence,

$$\limsup_{n\to\infty} \int_X f_n\, d\mu \le \int_X f\, d\mu \le \liminf_{n\to\infty} \int_X f_n\, d\mu$$

Therefore,

$$\lim_{n\to\infty} \int_X f_n\, d\mu = \int_X f\, d\mu$$

□

Continued...

Lemma 7 (Fatou's Lemma)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $\{f_n\}$ be a sequence of nonnegative measurable functions. Then*

$$\int_X \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu$$

Proof. Define, for each $n \geq 1$,

$$g_n := \inf_{m\geq n} f_m$$

Each g_n is measurable, and the sequence $\{g_n\}$ is nondecreasing (since the index set in the infimum shrinks as n increases). Moreover,

$$\liminf_{n\to\infty} f_n = \sup_{n\geq 1} \inf_{m\geq n} f_m = \sup_{n\geq 1} g_n = \lim_{n\to\infty} g_n$$

For every $m \geq n$ we have $g_n \leq f_m$, hence integrating gives

$$\int_X g_n \, d\mu \leq \int_X f_m \, d\mu \quad \text{for all } m \geq n$$

Taking the infimum over $m \geq n$ on the right-hand side yields

$$\int_X g_n \, d\mu \leq \inf_{m\geq n} \int_X f_m \, d\mu$$

Now take the limit (supremum) as $n \to \infty$ of the left-hand side and the corresponding limit inferior on the right:

$$\lim_{n\to\infty} \int_X g_n \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu$$

By the Monotone Convergence Theorem (Theorem 28) (since $g_n \uparrow \liminf_n f_n$ and each $g_n \geq 0$),

$$\int_X \liminf_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_X g_n \, d\mu$$

Combining the last two displayed inequalities gives Fatou's lemma:

$$\int_X \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu$$

□

Theorem 29 (Dominated Convergence Theorem)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space, and Let $\{f_n\}$ be a sequence of measurable functions. Let f_n converges to f either* *pointwise a.e. or in measure, and suppose there exists a function* $\varphi : X \to [0,\infty)$ *such that* $|f_n| \leq \varphi$ *for all* n *and*

$$\int_X \varphi \, d\mu < \infty$$

Then

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu$$

Proof. First consider the case $f_n \to f$ pointwise a.e. Then f is measurable as the pointwise almost everywhere limit of measurable functions and, for a.e. $x \in X$,

$$|f(x)| = \lim_{n\to\infty} |f_n(x)| \leq \varphi(x)$$

So, $|f| \leq \varphi$ a.e. and hence,

$$\int_X |f| \, d\mu \leq \int_X \varphi \, d\mu < \infty$$

Therefore, f is integrable.
Consider $g_n = \varphi + f_n$ and $h_n = \varphi - f_n$.
As $|f_n| \leq \varphi$, we have $\varphi \pm f_n \geq 0$.
Then g_n and h_n are non-negative measurable functions.
Then by Fatou's lemma (Lemma 7),

$$\int_X \liminf_{n\to\infty} g_n \, d\mu \leq \liminf_{n\to\infty} \int_X g_n \, d\mu$$

But $\lim\limits_{n\to\infty} g_n = \varphi + f$ a.e., hence

$$\int_X (\varphi + f) \, d\mu \leq \liminf_{n\to\infty} \int_X (\varphi + f_n) \, d\mu = \int_X \varphi \, d\mu + \liminf_{n\to\infty} \int_X f_n \, d\mu$$

where in the last equality we use that $\int_X \varphi \, d\mu < \infty$ so the constant $\int_X \varphi \, d\mu$ may be taken outside the $\liminf$. Therefore

$$\int_X f \, d\mu \leq \liminf_{n\to\infty} \int_X f_n \, d\mu$$

Similarly by Fatou's lemma (Lemma 7),

$$\int_X \liminf_{n\to\infty} h_n \, d\mu \leq \liminf_{n\to\infty} \int_X h_n \, d\mu$$

As $\lim_{n\to\infty} h_n = \varphi - f$ a.e., we have

$$\int_X (\varphi - f)\, d\mu \leq \liminf_{n\to\infty} \left(\int_X \varphi\, d\mu - \int_X f_n\, d\mu \right)$$

Hence,

$$-\int_X f\, d\mu \leq \liminf_{n\to\infty} \left(-\int_X f_n\, d\mu \right)$$

which implies

$$\int_X f\, d\mu \geq \limsup_{n\to\infty} \int_X f_n\, d\mu$$

Combining this with the opposite inequality obtained previously, we get

$$\limsup_{n\to\infty} \int_X f_n\, d\mu \leq \int_X f\, d\mu \leq \liminf_{n\to\infty} \int_X f_n\, d\mu$$

and therefore

$$\int_X f\, d\mu = \lim_{n\to\infty} \int_X f_n\, d\mu$$

This proves the theorem in the case where $f_n \to f$ pointwise almost everywhere. Now suppose that $f_n \to f$ in measure. By Theorem 26, every subsequence of $\{f_n\}$ admits a further subsequence which converges to f pointwise almost everywhere. Applying the result already proved to such a subsequence, we obtain,

$$\int_X f_{n_k}\, d\mu \longrightarrow \int_X f\, d\mu$$

Therefore every subsequence of the sequence

$$\left\{ \int_X f_n\, d\mu \right\}$$

has a further subsequence converging to

$$\int_X f\, d\mu$$

Hence the whole sequence converges to the same limit, and so,

$$\lim_{n\to\infty} \int_X f_n\, d\mu = \int_X f\, d\mu$$

□

Definition 33 (Equi-Integrability)**.** A family of measurable functions $\{f_\alpha\}$ is equi-integrable if $\forall \varepsilon > 0$, there exists $\delta_\varepsilon > 0$ such that if $\mu(E) < \delta_\varepsilon$, then,

$$\int_E |f_\alpha|\, d\mu < \varepsilon \quad , \forall \alpha$$

Theorem 30 (Vitali Convergence Theorem Finite Measure Version)**.** *Let $\mu(X) < \infty$ and Let $f_n : X \to \mathbb{R}$ be an equi-integrable family such that*

$$\sup_n \int_X |f_n|\, d\mu < \infty$$

Let $f_n \to f$ in measure. Then f is integrable and

$$\int_X f_n\, d\mu \longrightarrow \int_X f\, d\mu$$

Proof. As $f_n \to f$ in measure and $\mu(X) < \infty$, there exists a subsequence $\{f_{n_k}\}$ which converges to f almost everywhere. Hence

$$|f| = \liminf_{k\to\infty} |f_{n_k}| \quad \text{a.e.}$$

By Fatou's lemma (Lemma 7),

$$\int_X |f|\, d\mu \le \liminf_{k\to\infty} \int_X |f_{n_k}|\, d\mu \le \sup_k \int_X |f_k|\, d\mu < \infty$$

Thus, f is integrable.
We now show that f also satisfies the same equi-integrability estimate. Let $\varepsilon > 0$ and choose $\delta > 0$ corresponding to $\varepsilon/2$ from the equi-integrability of $\{f_n\}$.
Let E be measurable with $\mu(E) < \delta$.
Since $f_n \to f$ in measure and $\mu(X) < \infty$, by Theorem 26 there exists a subsequence $\{f_{n_k}\}$ converging to f almost everywhere.
Applying Fatou's lemma (Lemma 7) to the nonnegative functions, $|f_{n_k}|\mathbb{1}_E$, we obtain,

$$\int_E |f|\, d\mu \le \liminf_{k\to\infty} \int_E |f_{n_k}|\, d\mu$$

Since $\mu(E) < \delta$, by equi-integrability,

$$\int_E |f_{n_k}|\, d\mu < \frac{\varepsilon}{2} \qquad \forall k$$

Continued...

Hence,

$$\int_E |f|\, d\mu \le \frac{\varepsilon}{2}$$

Therefore f satisfies the same equi-integrability condition.
Consequently, the family $\{f_n\} \cup \{f\}$ is equi-integrable.
Fix $\varepsilon > 0$. As $\mu(X) < \infty$, choose $\eta > 0$ such that

$$\eta\, \mu(X) < \frac{\varepsilon}{3}$$

By equi-integrability of $\{f_n\} \cup \{f\}$ we can pick $\delta > 0$ such that for every measurable set E with $\mu(E) < \delta$,

$$\sup_n \int_E |f_n|\, d\mu < \frac{\varepsilon}{6} \qquad \text{and} \qquad \int_E |f|\, d\mu < \frac{\varepsilon}{6}$$

As $f_n \to f$ in measure, there exists $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$,

$$\mu\big(\{x :\ |f_n(x) - f(x)| > \eta\}\big) < \delta$$

Then for $n \ge n_0$,

$$\int_X |f_n - f|\, d\mu = \int_{\{|f_n - f| \le \eta\}} |f_n - f|\, d\mu + \int_{\{|f_n - f| > \eta\}} |f_n - f|\, d\mu$$

The first term is bounded by

$$\int_{\{|f_n - f| \le \eta\}} |f_n - f|\, d\mu \le \eta\, \mu(X) < \frac{\varepsilon}{3}$$

by our choice of η.
Using the triangle inequality in the second term and, Let

$$A_n = \{\, |f_n - f| > \eta \,\}$$

where $\mu(A_n) < \delta$. Then,

$$\int_{A_n} |f_n - f|\, d\mu \le \int_{A_n} |f_n|\, d\mu + \int_{A_n} |f|\, d\mu \le \frac{\varepsilon}{6} + \frac{\varepsilon}{6} = \frac{\varepsilon}{3}$$

this is by choice of δ.
Hence, combining everything, for $n > n_0$,

$$\int_X |f_n - f|\, d\mu < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3} < \varepsilon$$

Thus,

$$\int_X |f_n - f|\, d\mu \longrightarrow 0$$

Continued...

and hence,

$$\int_X f_n \, d\mu \longrightarrow \int_X f \, d\mu$$

□

Lemma 8 (Urysohn's Lemma)**.** *Let (X, d) be a metric space. Let F_0 and F_1 be two disjoint closed subsets of X. Then there exists a continuous function $\Phi : X \to [0, 1]$ such that $\Phi|_{F_0} = 0$ and $\Phi|_{F_1} = 1$.*

Proof. Consider the sets F_0 and F_1 which are closed and disjoint. If $F_0 = \varnothing$ we define $\Phi \equiv 1$, and if $F_1 = \varnothing$ we define $\Phi \equiv 0$.
So we assume that both F_0 and F_1 are non-empty.
For any non-empty subset $Y \subseteq X$ and $x \in X$, define

$$d(x, Y) = \inf_{y \in Y} d(x, y)$$

As the distance is Lipschitz in x, the function $x \mapsto d(x, Y)$ is continuous. Further, if Y is closed then

$$d(x, Y) = \begin{cases} 0, & x \in Y, \\ > 0, & x \notin Y \end{cases}$$

Define

$$\Phi(x) = \frac{d(x, F_0)}{d(x, F_0) + d(x, F_1)}$$

Since F_0 and F_1 are closed and disjoint, for every $x \in X$ at least one of $d(x, F_0)$ and $d(x, F_1)$ is strictly positive. Hence,

$$d(x, F_0) + d(x, F_1) > 0$$

so Φ is well-defined.
Since $d(\cdot, F_0)$ and $d(\cdot, F_1)$ are continuous and the denominator is positive, Φ is continuous.
This defines a continuous map $\Phi : X \to [0, 1]$ with $\Phi|_{F_0} = 0$ and $\Phi|_{F_1} = 1$. □

Theorem 31 (Tietze Extension Theorem)**.** *Let X be a metric space and F be a closed set in X. Let $f_0 : F \to \mathbb{R}$ be continuous. Then there exists a continuous function $f : X \to \mathbb{R}$ such that $f|_F = f_0$.*

Proof. There are two cases that arise.

- **Case I:** f_0 is bounded.
  Then $\exists\, M > 0$ such that $|f_0| \leq M$.
  Define two sets:

$$F^+ = \left\{ x \in F : \; f_0(x) \in \left[\frac{M}{3}, M \right] \right\},$$

$$F^- = \left\{ x \in F : \ f_0(x) \in \left[-M, -\frac{M}{3} \right] \right\}.$$

Then F^+ and F^- are two closed and disjoint sets.
Thus, by the lemma (Lemma 8), there exists a continuous function $\Phi : X \to [0,1]$ such that

$$\Phi|_{F^+} = 1 \quad \text{and} \quad \Phi|_{F^-} = 0$$

Let

$$f_1 = \frac{2M}{3}\Phi - \frac{M}{3}$$

Then,

$$f_1|_{F^+} = \frac{M}{3}, \qquad f_1|_{F^-} = -\frac{M}{3}, \qquad |f_1| \le \frac{M}{3}$$

Hence,

$$\sup_{x \in F} |f_0(x) - f_1(x)| \le \frac{2M}{3}$$

Applying the same construction to $f_0 - f_1$ on F, we obtain a continuous function $f_2 : X \to \mathbb{R}$ such that,

$$|f_2| \le \frac{1}{3}\left(\frac{2M}{3}\right)$$

and

$$\sup_{x \in F} |f_0(x) - f_1(x) - f_2(x)| \le \left(\frac{2}{3}\right)^2 M$$

Proceeding inductively, we obtain continuous functions $f_k : X \to \mathbb{R}$ such that

$$|f_k| \le \frac{1}{3}\left(\frac{2}{3}\right)^{k-1} M$$

and

$$\sup_{x \in F} \left| f_0(x) - \sum_{i=1}^{k} f_i(x) \right| \le \left(\frac{2}{3}\right)^k M$$

Since

$$|f_k| \le \frac{1}{3}\left(\frac{2}{3}\right)^{k-1} M$$

the Weierstrass M-test implies that the series, $\sum_{k=1}^{\infty} f_k$ converges uniformly on X. Define,

$$f = \sum_{k=1}^{\infty} f_k$$

Since each f_k is continuous and the series converges uniformly, f is continuous on X. Also,

$$\sup_{x\in F}\left|f_0(x)-\sum_{i=1}^{k} f_i(x)\right| \leq \left(\frac{2}{3}\right)^k M \longrightarrow 0$$

Hence

$$f|_F = f_0$$

- **Case II:** f_0 is not necessarily bounded.
  Let $g_0 = \arctan(f_0)$, then $g_0 : F \to \mathbb{R}$ is continuous and $|g_0| < \dfrac{\pi}{2}$.
  By Case I, there exists a continuous function $g : X \to \mathbb{R}$ such that $g|_F = g_0$.
  Let
  $$F' = \left\{x \in X : \; |g(x)| \geq \frac{\pi}{2}\right\}$$
  Then F' is closed and $F' \cap F = \varnothing$, because on F we have, $|g(x)| = |g_0(x)| < \frac{\pi}{2}$.
  Applying Lemma 8 again, there exists a continuous function $\psi : X \to [0,1]$, such that,
  $$\psi|_F = 1 \qquad \text{and} \qquad \psi|_{F'} = 0$$
  - If $x \in F'$, then $\psi(x) = 0$, so $g(x)\psi(x) = 0 < \frac{\pi}{2}$.
  - If $x \notin F'$, then $|g(x)| < \frac{\pi}{2}$ and $0 \leq \psi(x) \leq 1$, so, $|g(x)\psi(x)| \leq |g(x)| < \frac{\pi}{2}$.

  Therefore,
  $$|g(x)\psi(x)| < \frac{\pi}{2} \qquad \forall x \in X$$
  Therefore, the function $f = \tan(g\psi)$ is well-defined and continuous on X.
  Finally, for $x \in F$, we have $\psi(x) = 1$ and $g(x) = g_0(x) = \arctan(f_0(x))$.
  Hence,
  $$f(x) = \tan(g(x)\psi(x)) = \tan(g_0(x)) = \tan(\arctan(f_0(x))) = f_0(x)$$
  Therefore,
  $$f|_F = f_0$$

□

Definition 34 (Completion of a Measure Space)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space and $(X, \widetilde{\mathcal{A}}, \widetilde{\mu})$ be its completion. Then, the σ-algebra $\widetilde{\mathcal{A}}$ consists precisely of all sets of the form $E = A \cup B$, where $A \in \mathcal{A}$, $B \subseteq N$, with $N \in \mathcal{A}$ such that $\mu(N) = 0$. For any such set E, $\widetilde{\mu}(E) = \mu(A)$.

Theorem 32 (Approximation of Complete Measurable Functions)**.** *Let μ be any measure on $(X, \mathcal{A})$ and Let $(X, \widetilde{\mathcal{A}}, \widetilde{\mu})$ be its completion. Let $\widetilde{f} : X \to \mathbb{R}$ be a function that is $\widetilde{\mathcal{A}}$-measurable. Then there exists $f : X \to \mathbb{R}$ that is $\mathcal{A}$-measurable such that $f(x) = \widetilde{f}(x)$ almost everywhere.*

Proof. We start using characteristic functions and eventually build up to general measurable functions.
Let $\widetilde{f} = \mathbb{1}_E$ for some $E \in \widetilde{\mathcal{A}}$, so $\widetilde{f}$ is $\widetilde{\mathcal{A}}$-measurable. Then $E = A \cup B$, where $A \in \mathcal{A}$, $B \subseteq N$ and $N \in \mathcal{A}$ with $\mu(N) = 0$. Let $f = \mathbb{1}_A$. As $A \in \mathcal{A}$, f is $\mathcal{A}$-measurable. Moreover,

$$\{x \in X : \mathbb{1}_E \neq \mathbb{1}_A\} \subseteq B \subseteq N$$

As $\tilde{\mu}(N) = 0$. Thus $f = \tilde{f}$ a.e.
Hence, for $\mathbb{1}_E$ which is $\widetilde{\mathcal{A}}$-measurable, there exists $\mathbb{1}_A$, which is $\mathcal{A}$-measurable, such that $f = \tilde{f}$ a.e.
Now consider simple functions. Let

$$\widetilde{f} = \sum_{k=1}^{n} c_k \mathbb{1}_{E_k}$$

where for $k = 1, 2, \ldots, n$, $E_k = A_k \cup B_k$ with $E_k \in \widetilde{\mathcal{A}}$, $A_k \in \mathcal{A}$, $B_k \subseteq N_k$ and N_k a null set with $\mu(N_k) = 0$.
Then for each $\mathbb{1}_{E_k}$ we obtain $\mathbb{1}_{A_k}$ by the previous step, and for $\widetilde{f}$ the $\mathcal{A}$-measurable function we get,

$$f = \sum_{k=1}^{n} c_k \mathbb{1}_{A_k}$$

which is $\mathcal{A}$-measurable and,

$$\{x \in X : \widetilde{f}(x) \neq f(x)\} \subseteq \bigcup_{k=1}^{n} B_k \subseteq \bigcup_{k=1}^{n} N_k$$

As $\tilde{\mu}(N_k) = 0$. So, $\tilde{\mu}\left(\bigcup_{k=1}^{n} N_k\right) = 0$. Thus $f = \tilde{f}$ a.e.
Now, Let $\widetilde{f}$ be any non-negative $\widetilde{\mathcal{A}}$-measurable function. Then $\widetilde{f}$ is the pointwise limit of simple functions $\widetilde{f_n}$ which are $\widetilde{\mathcal{A}}$-measurable. For each $\widetilde{f_n}$ we find an $\mathcal{A}$-measurable function f_n and a null set N_n such that

$$f_n = \widetilde{f_n} \quad \text{on } X \setminus N_n$$

Let $N = \bigcup_{n=1}^{\infty} N_n$. Then $\tilde{\mu}(N) = 0$. Thus, on the set $X \setminus N$,

$$f = \lim_{n \to \infty} f_n$$

and on the set N, we define $f \equiv 0$.
Then f is $\mathcal{A}$-measurable and $f = \tilde{f}$ a.e.
Finally, Let $\widetilde{f}$ be an arbitrary real-valued $\widetilde{\mathcal{A}}$-measurable function. Then we can write,

$$\widetilde{f} = \widetilde{f}^{+} - \widetilde{f}^{-}$$

By the previous step there exist $\mathcal{A}$-measurable functions f^+ and f^- such that,

$$f^+ = \widetilde{f}^+ \quad \text{a.e.} \qquad \text{and} \qquad f^- = \widetilde{f}^- \quad \text{a.e.}$$

Define,

$$f = f^+ - f^-$$

Then f is $\mathcal{A}$-measurable and $f = \widetilde{f}$ almost everywhere. □

Theorem 33 (Approximation by Continuous Functions)**.** *Let X be a metric space, $\mathcal{B}(X)$ be the Borel σ-algebra on X and μ be a finite regular Borel measure. Let $f : X \to \mathbb{R}$ be a measurable function. Then, $\forall \varepsilon > 0$, $\exists$ a continuous function $g : X \to \mathbb{R}$, s.t.*

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

Proof. Let $E \subseteq X$ be a Borel set. By regularity of the Borel measure, $\exists$ a closed set F and an open set G such that

$$F \subseteq E \subseteq G, \qquad \mu(G \setminus F) < \varepsilon$$

Then F is closed and disjoint from G^c.
By Urysohn's lemma (Lemma 8), $\exists$ a continuous function g such that $g|_F = 1$ and $g|_{G^c} = 0$ and $0 \leq g(x) \leq 1$ elsewhere.
Now for $f = \mathbb{1}_E$. On F, $\mathbb{1}_E = 1$ and $g|_F = 1$, and on G^c, $\mathbb{1}_E = 0$ and $g|_{G^c} = 0$. And,

$$\{x \in X :\ f(x) \neq g(x)\} \subseteq G \setminus F$$

and thus,

$$\mu\big(\{x \in X :\ f(x) \neq g(x)\}\big) \leq \mu(G \setminus F) < \varepsilon$$

Now consider the simple function

$$f = \sum_{k=1}^{n} c_k \mathbb{1}_{E_k}$$

where for $k = 1, 2, \ldots, n$. Then for each k, there exists a continuous function g_k such that

$$\mu\big(\{x \in X :\ g_k(x) \neq \mathbb{1}_{E_k}(x)\}\big) < \frac{\varepsilon}{n}$$

Let

$$g = \sum_{k=1}^{n} c_k g_k$$

Then,

$$\{x \in X :\ f(x) \neq g(x)\} \subseteq \bigcup_{k=1}^{n} \{x \in X :\ \mathbb{1}_{E_k}(x) \neq g_k(x)\}$$

and therefore,

$$\mu\big(\{x \in X :\ f(x) \neq g(x)\}\big) \leq \sum_{k=1}^{n} \frac{\varepsilon}{n} = \varepsilon$$

Now, Let f be any arbitrary real-valued measurable function. Then, there exists a sequence of real-valued simple measurable functions $\{f_n\}$ such that,

$$f = \lim_{n\to\infty} f_n$$

pointwise on X.
For each simple function f_n there exists a continuous function g_n such that

$$\mu\big(\{x \in X :\ f_n(x) \neq g_n(x)\}\big) < \frac{\varepsilon}{3 \cdot 2^n}$$

Let

$$E_1 = \bigcup_{n=1}^{\infty} \{x \in X :\ f_n(x) \neq g_n(x)\}$$

Then the pointwise limit of g_n exists and equals f on $X \setminus E_1$, and $\mu(E_1) < \dfrac{\varepsilon}{3}$.
Applying Egorov's theorem (Theorem 27), there exists E_2 such that $\mu(E_2) < \dfrac{\varepsilon}{3}$, where,

$$g_n \to f \quad \text{uniformly on } (X \setminus E_1) \setminus E_2$$

Let $\tilde{g} : (X \setminus E_1) \setminus E_2 \to \mathbb{R}$ be the limit of g_n. Then, $\tilde{g}$ is continuous on $(X \setminus E_1) \setminus E_2$. By regularity of the Borel measure, there exists an open set U such that

$$E_1 \cup E_2 \subseteq U \qquad \text{and} \qquad \mu(U) < \mu(E_1 \cup E_2) + \frac{\varepsilon}{3}$$

Let $F = X \setminus U$.
Then, F is closed and

$$F \subseteq X \setminus (E_1 \cup E_2)$$

Moreover,

$$\mu(X \setminus F) = \mu(U) < \mu(E_1 \cup E_2) + \frac{\varepsilon}{3} < \varepsilon$$

Since $F \subseteq (X \setminus E_1) \setminus E_2$, the restriction $\tilde{g}|_F$ is continuous on the closed set F. Then, by the Tietze extension theorem (Theorem 31), there exists a continuous function $g : X \to \mathbb{R}$ extending $\tilde{g}$ on F. Since g extends $\tilde{g}$ on F, we have $g = f$ on F. Hence,

$$\{x \in X :\ f(x) \neq g(x)\} \subseteq X \setminus F$$

Therefore,

$$\mu\big(\{x \in X :\ f(x) \neq g(x)\}\big) \leq \mu(X \setminus F) < \varepsilon$$

Thus, there exists a continuous $g : X \to \mathbb{R}$ such that

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

□

Theorem 34 (Density of Continuous Functions in L^1)**.** *Let X be a metric space and Let μ be a finite regular Borel measure on X. If f is integrable, i.e. $\int_X |f|\,d\mu < \infty$, then, for every $\varepsilon > 0$, there exists a continuous function $g : X \to \mathbb{R}$ such that, $\int_X |f - g|\,d\mu < \varepsilon$.*

Proof. Consider the truncations f_M defined by

$$f_M(x) = \begin{cases} M, & f(x) \geq M, \\ f(x), & |f(x)| \leq M, \\ -M, & f(x) \leq -M \end{cases}$$

Then $|f - f_M| \to 0$ pointwise as $M \to \infty$, and $0 \leq |f - f_M| \leq |f|$.
By the Dominated Convergence Theorem (Theorem 29), since $\int_X |f|\,d\mu < \infty$, we have

$$\int_X |f - f_M|\,d\mu \to 0$$

In particular, for any $\varepsilon > 0$ we can choose M large enough that

$$\int_X |f - f_M|\,d\mu < \frac{\varepsilon}{4}$$

By the Approximation by Continuous Functions Theorem (Theorem 33), there exists a continuous function g_M such that,

$$\mu\big(\{x \in X : f_M(x) \neq g_M(x)\}\big) < \frac{\varepsilon}{4M}$$

Define the truncation of g_M by

$$h_M(x) = \begin{cases} M, & g_M(x) \geq M, \\ g_M(x), & |g_M(x)| \leq M, \\ -M, & g_M(x) \leq -M \end{cases}$$

Since truncation is a continuous operation, h_M is continuous.
Moreover, if $f_M(x) = g_M(x)$, then $|f_M(x)| \leq M$, and therefore,

$$h_M(x) = g_M(x) = f_M(x)$$

Hence,

$$\{x \in X : \; f_M(x) \neq h_M(x)\} \subseteq \{x \in X : \; f_M(x) \neq g_M(x)\}$$

Also, $|f_M(x)| \leq M$ and $|h_M(x)| \leq M$ so, $|f_M(x) - h_M(x)| \leq 2M$.
Therefore,

$$\begin{aligned}\int_X |f_M - h_M|\, d\mu &= \int_{\{f_M \neq h_M\}} |f_M - h_M|\, d\mu \\ &\leq 2M\, \mu(\{f_M \neq h_M\}) \\ &\leq 2M\, \mu(\{f_M \neq g_M\}) \\ &< 2M \cdot \frac{\varepsilon}{4M} \\ &= \frac{\varepsilon}{2}.\end{aligned}$$

Let $g = h_M$. Then

$$\begin{aligned}\int_X |f - g|\, d\mu &\leq \int_X |f - f_M|\, d\mu + \int_X |f_M - g|\, d\mu \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{2} \\ &= \frac{3\varepsilon}{4} < \varepsilon.\end{aligned}$$

Thus there exists a continuous function $g : X \to \mathbb{R}$ such that

$$\int_X |f - g|\, d\mu < \varepsilon$$

□

Definition 35 (Product Space)**.** Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be two finite measure spaces. Then

$$X \times Y = \{(x, y) : x \in X, y \in Y\}$$

is a product space.

Definition 36 (Measurable Rectangle)**.** Let $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Then $A \times B$ is said to be a measurable rectangle.

Definition 37 (Elementary Set)**.** For some finite number of sets $A_i \in \mathcal{A}$ and $B_i \in \mathcal{B}$, a set of the form,

$$Q = \bigcup_{i=1}^{n} (A_i \times B_i)$$

where the rectangles $A_i \times B_i$ are pairwise disjoint, is called an elementary set.

Definition 38 (Least σ-Algebra)**.** $\Omega = \mathcal{A} \otimes \mathcal{B} = \sigma(\mathcal{P})$ is the least σ-algebra on $X \times Y$ containing $\mathcal{P} = \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$.

Theorem 35 (Semi-ring Existence)**.** *Let* $(X, \mathcal{A}, \mu)$ *and* $(Y, \mathcal{B}, \lambda)$ *be two finite measure spaces. Then, the collection of sets* $\mathcal{P} = \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$ *is a semi-ring.*

Proof. To prove that $\mathcal{P}$ is a semi-ring, we must verify three fundamental conditions:

- Since $\mathcal{A}$ and $\mathcal{B}$ are σ-algebra's, then, $\varnothing \in \mathcal{A}$ and $\varnothing \in \mathcal{B}$. By the definition of the Cartesian product:
  $$\varnothing = \varnothing \times \varnothing$$
  Since $\varnothing \in \mathcal{A}$ and $\varnothing \in \mathcal{B}$, it follows immediately that $\varnothing \in \mathcal{P}$.

- Let $P_1, P_2 \in \mathcal{P}$. So, we can express these sets as $P_1 = A_1 \times B_1$ and $P_2 = A_2 \times B_2$, where $A_1, A_2 \in \mathcal{A}$ and $B_1, B_2 \in \mathcal{B}$. Computing their intersection yields:
  $$P_1 \cap P_2 = (A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2)$$
  Because $\mathcal{A}$ and $\mathcal{B}$ are σ-algebra's, they are both closed under finite intersections. Therefore;
  $$A_1 \cap A_2 \in \mathcal{A}$$
  $$B_1 \cap B_2 \in \mathcal{B}$$
  Consequently, $(A_1 \cap A_2) \times (B_1 \cap B_2) \in \mathcal{P}$, proving $\mathcal{P}$ is closed under intersection.

- Let $P_1, P_2 \in \mathcal{P}$. We consider the difference between two rectangles $P_1 \setminus P_2$. Consider $P_1 = A_1 \times B_1$ and $P_2 = A_2 \times B_2$. We can write the difference using the following identity:
  $$(A_1 \times B_1) \setminus (A_2 \times B_2) = ((A_1 \setminus A_2) \times B_1) \cup ((A_1 \cap A_2) \times (B_1 \setminus B_2))$$
  Since $\mathcal{A}$ and $\mathcal{B}$ are σ-algebras,
  $$A_1 \setminus A_2 \in \mathcal{A}, \qquad A_1 \cap A_2 \in \mathcal{A}, \qquad B_1 \setminus B_2 \in \mathcal{B}$$
  Hence,
  $$(A_1 \setminus A_2) \times B_1 \in \mathcal{P}, \qquad (A_1 \cap A_2) \times (B_1 \setminus B_2) \in \mathcal{P}$$
  Moreover,
  $$\begin{aligned} &\big((A_1 \setminus A_2) \times B_1\big) \cap \big((A_1 \cap A_2) \times (B_1 \setminus B_2)\big) \\ &= \big((A_1 \setminus A_2) \cap (A_1 \cap A_2)\big) \times \big(B_1 \cap (B_1 \setminus B_2)\big) \\ &= \varnothing \times (B_1 \setminus B_2) = \varnothing \end{aligned}$$
  Therefore, $P_1 \setminus P_2$ is a finite union of pairwise disjoint sets from $\mathcal{P}$.

Therefore, the collection of sets $\mathcal{P} = \{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$ is a semi-ring. $\square$

Continued...

Theorem 36 (Pre-measure on the Semi-ring)**.** *Let* $(X, \mathcal{A}, \mu)$ *and* $(Y, \mathcal{B}, \lambda)$ *be two finite measure spaces. Then, there is a pre-measure* η_0 *on* $\mathcal{P}$ *such that* $\eta_0(A \times B) = \mu(A) \times \lambda(B)$*.*

Proof. Let us define the Set Function $\eta_0 : \mathcal{P} \to [0, \infty)$ by:

$$\eta_0(A \times B) = \mu(A) \cdot \lambda(B)$$

Since $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ are finite measure spaces, the product $\mu(A) \cdot \lambda(B)$ is a well-defined non-negative real number.
To show that η_0 is a pre-measure on the semiring $\mathcal{P}$, we must show that $\eta_0(\varnothing) = 0$ and that it is countably additive on pairwise disjoint decompositions in $\mathcal{P}$.

- The empty set $\varnothing$ in $\mathcal{P}$ can be written as $\varnothing \times \varnothing$. Using our definition:

  $$\eta_0(\varnothing) = \eta_0(\varnothing \times \varnothing) = \mu(\varnothing) \cdot \lambda(\varnothing) = 0 \cdot 0 = 0$$

- Let $A \times B = \bigcup_{i=1}^{\infty}(A_i \times B_i)$ where the rectangles $(A_i \times B_i)$ are pairwise disjoint members of $\mathcal{P}$.
  For each fixed $x \in A$, we have, $B = \bigcup_{\{i:\, x \in A_i\}} B_i$, because, if $y \in B$, then $(x, y) \in A \times B$, so there exists some i such that $(x, y) \in A_i \times B_i$. Hence $x \in A_i$ and $y \in B_i$.
  Moreover, the family $\{B_i : \ x \in A_i\}$ is pairwise disjoint because the rectangles $(A_i \times B_i)$ are pairwise disjoint.
  Since λ is a measure,

  $$\lambda(B) = \sum_{\{i:\, x \in A_i\}} \lambda(B_i) \implies \mathbb{1}_A(x)\lambda(B) = \sum_{i=1}^{\infty} \mathbb{1}_{A_i}(x)\lambda(B_i)$$

  Integrating over X with respect to μ, we obtain,

  $$\mu(A)\lambda(B) = \int_X \mathbb{1}_A(x)\lambda(B)\, d\mu(x) = \int_X \sum_{i=1}^{\infty} \mathbb{1}_{A_i}(x)\lambda(B_i)\, d\mu(x)$$

  Since the summands are nonnegative, the Monotone Convergence Theorem (Theorem 28) yields,

  $$\mu(A)\lambda(B) = \sum_{i=1}^{\infty} \int_X \mathbb{1}_{A_i}(x)\lambda(B_i)\, d\mu(x) = \sum_{i=1}^{\infty} \mu(A_i)\lambda(B_i)$$

  Therefore,

  $$\eta_0(A \times B) = \sum_{i=1}^{\infty} \eta_0(A_i \times B_i)$$

Therefore, η_0 is a pre-measure on the semi-ring $\mathcal{P}$. $\square$

Theorem 37 (Existence of a Measure)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be two finite measure spaces and Ω is the least σ-algebra on $X \times Y$. Then, there exists a unique measure η on Ω extending η_0.*

Proof. By Theorem 35, the collection, $\mathcal{P} = \{A \times B : \ A \in \mathcal{A}, \ B \in \mathcal{B}\}$ is a semi-ring.
By Theorem 36, the function $\eta_0(A \times B) = \mu(A)\lambda(B)$ defines a pre-measure on $\mathcal{P}$.
Also, by Definition 38, $\Omega = \sigma(\mathcal{P})$.
Therefore, by Carathéodory Existence Theorem (Theorem 17), there exists a measure η on Ω extending η_0.
Since $\eta_0(X \times Y) = \mu(X)\lambda(Y) < \infty$ the pre-measure η_0 is finite. Hence, by the Carathéodory Uniqueness Theorem (Theorem 20), this extension is unique.
Therefore, there exists a unique measure η on Ω extending η_0.
In particular, for every $A \in \mathcal{A}$ and $B \in \mathcal{B}$,

$$\eta(A \times B) = \eta_0(A \times B) = \mu(A)\lambda(B)$$

$\square$

Definition 39 (Slices)**.** Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be two finite measure spaces and Let $E \subseteq X \times Y$. For fixed $x \in X$; we define $E_x = \{y : (x, y) \in E\}$ and similarly, for fixed $y \in Y$, we define $E_y = \{x : (x, y) \in E\}$. These are called the x-slice and y-slice respectively.

Theorem 38 (Tonelli for Sets: Vertical Sections)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be two finite measure spaces and Ω is the least σ-algebra on $X \times Y$. For any set $E \in \Omega$, we have,*

(i) $\forall x \in X$, we have $E_x = \{y \in Y : (x, y) \in E\} \in \mathcal{B}$.

(ii) $x \mapsto \lambda(E_x)$ is $\mathcal{A}$ measurable.

(iii) $\eta(E) = \int_X \lambda(E_x) d\mu(x)$

Proof. We first check the Properties for rectangles in the semi-ring $\mathcal{P}$. Let $E = A \times B \in \mathcal{P}$ where $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

(i) For a fixed x, the section $E_x = \{y \in Y : (x, y) \in A \times B\}$.
If $x \in A$, then $E_x = B \in \mathcal{B}$.
If $x \notin A$, then $E_x = \varnothing \in \mathcal{B}$.
So, E_x is always in $\mathcal{B}$.

(ii) The function $x \mapsto \lambda(E_x)$ can be written as:

$$\lambda(E_x) = \begin{cases} \lambda(B) & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} = \lambda(B) \cdot \mathbb{1}_A(x)$$

Since $\mathbb{1}_A(x)$ is a measurable function (because $A \in \mathcal{A}$), the whole function is measurable.

(iii) By definition $\eta(E) = \eta(A \times B) = \mu(A)\lambda(B)$. And,

$$\int_X \lambda(E_x)\, d\mu(x) = \int_X \lambda(B)\mathbf{1}_A(x)\, d\mu(x) = \lambda(B) \int_X \mathbf{1}_A(x)\, d\mu(x) = \lambda(B)\mu(A)$$

Thus, $\eta(E) = \int_X \lambda(E_x) d\mu(x)$

Now Let $\{E_i\}$ be a disjoint collection of sets in $\mathcal{P}$. Let $E = \bigcup_{i=1}^{n} E_i$ where $E_i \cap E_j = \emptyset$ for $i \neq j$, and each E_i satisfies properties (i), (ii), and (iii). Then,

(i) As, the section of a union is the union of the sections:

$$E_x = \{y \in Y : (x, y) \in \bigcup_{i=1}^{n} E_i\} = \bigcup_{i=1}^{n} (E_i)_x$$

Since each E_i satisfies property (i), we know each $(E_i)_x \in \mathcal{B}$. Because $\mathcal{B}$ is a σ-algebra, it is closed under finite unions. Therefore:

$$E_x \in \mathcal{B}$$

(ii) Because the sets E_i are disjoint, their sections $(E_i)_x$ are also disjoint for any fixed x. The measure of a disjoint union is the sum of the measures:

$$\lambda(E_x) = \lambda\left(\bigcup_{i=1}^{n} (E_i)_x\right) = \sum_{i=1}^{n} \lambda((E_i)_x)$$

Each function $f_i(x) = \lambda((E_i)_x)$ is measurable by property (ii)) of E_i. The sum of a finite number of measurable functions is also a measurable function. Thus:

$$x \mapsto \lambda(E_x) \text{ is measurable.}$$

(iii) By additivity of the measure η:

$$\eta(E) = \eta\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \eta(E_i)$$

Since each E_i satisfies property (iii), we can replace $\eta(E_i)$ with its integral:

$$\sum_{i=1}^{n} \eta(E_i) = \sum_{i=1}^{n} \int_X \lambda((E_i)_x)\, d\mu(x)$$

By the linearity of the integral:

$$\sum_{i=1}^{n} \int_X \lambda((E_i)_x)\, d\mu(x) = \int_X \left(\sum_{i=1}^{n} \lambda((E_i)_x) \right) d\mu(x)$$

Substituting our result from Property (ii), which is $\sum \lambda((E_i)_x) = \lambda(E_x)$:

$$\eta(E) = \int_X \lambda(E_x)\, d\mu(x)$$

Thus, if the properties hold for individual disjoint sets, they hold for their finite union. By the Ring Generated by a Semi-ring Theorem (Theorem 18), every element of the ring $\mathcal{R}$ generated by $\mathcal{P}$ can be written as a finite disjoint union of elements of $\mathcal{P}$. Since we have shown that properties (i), (ii), and (iii) hold for elements of $\mathcal{P}$ and are preserved under finite disjoint unions, it follows that every element of $\mathcal{R}$ satisfies properties (i), (ii), and (iii).

Now, Let $\mathcal{M}$ be the collection of all defined as,

$$\mathcal{M} = \{E \in \Omega : \text{ properties (i), (ii), and (iii) hold}\}$$

- **Case I:** Let $E_1 \subseteq E_2 \subseteq \dots$ be sets in $\mathcal{M}$, and Let $E = \bigcup E_n$.

  (i) Since $E_x = \bigcup (E_n)_x$ and as each $(E_n)_x \in \mathcal{B}$ and $\mathcal{B}$ is a σ-algebra, so, $E_x \in \mathcal{B}$.

  (ii) By continuity of measure,

  $$\lambda(E_x) = \lambda(\bigcup (E_n)_x) = \lim_{n\to\infty} \lambda((E_n)_x)$$

  Since each function $x \mapsto \lambda((E_n)_x)$ is $\mathcal{A}$-measurable and measurable functions are closed under pointwise limits, it follows that, $x \mapsto \lambda(E_x)$ is $\mathcal{A}$-measurable.

  (iii) By the Monotone Convergence Theorem (Theorem 28):

  $$\begin{aligned} \int_X \lambda(E_x)\, d\mu &= \int_X \lim \lambda((E_n)_x)\, d\mu \\ &= \lim \int_X \lambda((E_n)_x)\, d\mu \\ &= \lim \eta(E_n) \\ &= \eta(E) \end{aligned}$$

- **Case II:** Let $E_1 \supseteq E_2 \supseteq \dots$ be a sequence of sets in $\mathcal{M}$, and Let $E = \bigcap\limits_{n=1}^{\infty} E_n$.

  Since the measure spaces are finite, so, $\mu(X) < \infty$ and $\lambda(Y) < \infty$.

  (i) For any fixed $x \in X$,

  $$E_x = \left(\bigcap_{n=1}^{\infty} E_n \right)_x = \bigcap_{n=1}^{\infty} (E_n)_x$$

  Since each $E_n \in \mathcal{M}$, we know $(E_n)_x \in \mathcal{B}$ for all n. Because $\mathcal{B}$ is a σ-algebra, it is closed under countable intersections. Therefore, $E_x \in \mathcal{B}$.

  (ii) Define $f_n(x) = \lambda((E_n)_x)$. Since $E_n \in \mathcal{M}$, each f_n is an $\mathcal{A}$-measurable function. Because E_n is a decreasing sequence of sets, the sections $(E_n)_x$ are also decreasing for every x. Since $\lambda(Y) < \infty$, we can use the continuity of measure for decreasing sequences:

  $$\lambda(E_x) = \lambda\left(\bigcap_{n=1}^{\infty} (E_n)_x \right) = \lim_{n\to\infty} \lambda((E_n)_x) = \lim_{n\to\infty} f_n(x)$$

  The pointwise limit of a sequence of measurable functions is measurable. Therefore, $x \mapsto \lambda(E_x)$ is $\mathcal{A}$-measurable.

  (iii) As, $(X \times Y, \Omega, \eta)$ is a finite measure space, the continuity of the product measure gives us:

  $$\eta(E) = \lim_{n\to\infty} \eta(E_n)$$

  Since $E_n \in \mathcal{M}$, we can replace $\eta(E_n)$ with its integral form:

  $$\eta(E) = \lim_{n\to\infty} \int_X \lambda((E_n)_x)\, d\mu(x)$$

  Let $g_n(x) = \lambda((E_n)_x)$. We know $g_n(x) \leq \lambda(Y)$ for all n and all x.
  Since the space is finite, the constant function $\lambda(Y)$ is integrable. By Dominated Convergence Theorem (Theorem 29),

  $$\eta(E) = \int_X \lim_{n\to\infty} \lambda((E_n)_x)\, d\mu(x) = \int_X \lambda(E_x)\, d\mu(x)$$

Therefore, $\mathcal{M}$ is a monotone class.
Since every element of $\mathcal{R}$ satisfies properties (i), (ii), and (iii), we have,

$$\mathcal{R} \subseteq \mathcal{M}$$

Hence, by the Monotone Class Theorem (Theorem 19),

$$\sigma(\mathcal{R}) \subseteq \mathcal{M}$$

Since $\mathcal{R}$ is the ring generated by $\mathcal{P}$, we have,

$$\sigma(\mathcal{R}) = \sigma(\mathcal{P}) = \Omega$$

Therefore,

$$\Omega \subseteq \mathcal{M}$$

But, by definition, $\mathcal{M} \subseteq \Omega$.
Hence,

$$\mathcal{M} = \Omega$$

Therefore, every set $E \in \Omega$ satisfies properties (i), (ii), and (iii). □

Theorem 39 (Tonelli for Sets: Horizontal Sections)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be two finite measure spaces and Ω is the least σ-algebra on $X \times Y$. For any set $E \in \Omega$, we have,*

(i) $\forall y \in Y$, we have $E_y = \{x \in X : (x, y) \in E\} \in \mathcal{A}$.

(ii) $y \mapsto \mu(E_y)$ is $\mathcal{B}$ measurable.

(iii) $\eta(E) = \int_Y \mu(E_y) d\lambda(y)$

Theorem 40 (Tonelli's Theorem: Vertical Sections)**.** *Let $f : X \times Y \to [0, \infty]$ be Ω measurable. Then,*

(i) $\forall x \in X$, we have $f(x, \cdot) : Y \to [0, \infty]$ is $\mathcal{B}$ measurable.

(ii) The function $x \mapsto \int_Y f(x, y) d\lambda(y)$ is $\mathcal{A}$ measurable.

(iii) $$\int_{X\times Y} f\, d\eta = \int_X \left(\int_Y f(x, y)\, d\lambda(y) \right) d\mu(x)$$

Proof. Let $f = \mathbb{1}_E$ be the characteristic function of a set $E \in \Omega$.

(i) For a fixed x, we have

$$f(x, y) = \mathbb{1}_{E_x}(y)$$

Since $E_x \in \mathcal{B}$ by Theorem 38, the function $f(x, \cdot)$ is $\mathcal{B}$-measurable.

(ii) For each $x \in X$,

$$\int_Y f(x, y)\, d\lambda(y) = \int_Y \mathbb{1}_{E_x}(y)\, d\lambda(y) = \lambda(E_x)$$

By Theorem 38, the function $x \mapsto \lambda(E_x)$ is $\mathcal{A}$-measurable.
Hence, $x \mapsto \displaystyle\int_Y f(x, y)\, d\lambda(y)$ is $\mathcal{A}$-measurable.

(iii) Since

$$\int_{X\times Y} \mathbb{1}_E \, d\eta = \eta(E)$$

and

$$\int_X \left(\int_Y \mathbb{1}_{E_x}(y) \, d\lambda(y) \right) d\mu(x) = \int_X \lambda(E_x) \, d\mu(x)$$

Theorem 38 gives,

$$\int_{X\times Y} \mathbb{1}_E \, d\eta = \int_X \left(\int_Y \mathbb{1}_{E_x}(y) \, d\lambda(y) \right) d\mu(x)$$

Therefore,

$$\int_{X\times Y} f \, d\eta = \int_X \left(\int_Y f(x,y) \, d\lambda(y) \right) d\mu(x)$$

Now Let f be a non-negative simple function on $X \times Y$. Then, we can write

$$f(x,y) = \sum_{i=1}^{n} c_i \mathbb{1}_{E_i}(x,y), \qquad c_i \geq 0$$

where $E_i \in \Omega$ are pairwise disjoint.
Since properties (i), (ii), and (iii) hold for each indicator function $\mathbb{1}_{E_i}$, they also hold for f by linearity of the integral.
This is because:

(i) Since

$$f(x,\cdot) = \sum_{i=1}^{n} c_i \mathbb{1}_{(E_i)_x}$$

so $f(x,\cdot)$ is $\mathcal{B}$-measurable for every $x \in X$.

(ii) Moreover,

$$\int_Y f(x,y) \, d\lambda(y) = \sum_{i=1}^{n} c_i \lambda((E_i)_x)$$

which is $\mathcal{A}$-measurable as a finite sum of $\mathcal{A}$-measurable functions.

(iii) Finally,

$$\begin{aligned}\int_{X\times Y} f \, d\eta &= \sum_{i=1}^{n} c_i \eta(E_i) \\ &= \sum_{i=1}^{n} c_i \int_X \lambda((E_i)_x) \, d\mu(x)\end{aligned}$$

$$= \int_X \sum_{i=1}^{n} c_i \lambda((E_i)_x)\, d\mu(x)$$

$$= \int_X \left(\int_Y f(x,y)\, d\lambda(y) \right) d\mu(x).$$

Now Let $f : X \times Y \to [0, \infty]$ be any Ω-measurable function. There exists an increasing sequence of non-negative simple functions $\{f_n\}$ such that $f_n \uparrow f$ pointwise.

(i) For a fixed $x \in X$, we have

$$f_n(x, \cdot) \uparrow f(x, \cdot)$$

Since each $f_n(x, \cdot)$ is $\mathcal{B}$-measurable and limits of measurable functions are measurable, $f(x, \cdot)$ is $\mathcal{B}$-measurable.

(ii) For each $x \in X$, the Monotone Convergence Theorem (Theorem 28) with respect to λ gives

$$\lim_{n\to\infty} \int_Y f_n(x,y)\, d\lambda(y) = \int_Y f(x,y)\, d\lambda(y)$$

Since each function

$$x \mapsto \int_Y f_n(x,y)\, d\lambda(y)$$

is $\mathcal{A}$-measurable, their pointwise limit

$$x \mapsto \int_Y f(x,y)\, d\lambda(y)$$

is also $\mathcal{A}$-measurable.

(iii) Applying the Monotone Convergence Theorem (Theorem 28) with respect to η,

$$\int_{X\times Y} f\, d\eta = \lim_{n\to\infty} \int_{X\times Y} f_n\, d\eta$$

Since the formula has already been proved for simple functions,

$$\int_{X\times Y} f_n\, d\eta = \int_X \left(\int_Y f_n(x,y)\, d\lambda(y) \right) d\mu(x)$$

Applying the Monotone Convergence Theorem (Theorem 28) again with respect to μ,

$$\begin{aligned}\int_{X\times Y} f\, d\eta &= \lim_{n\to\infty} \int_X \left(\int_Y f_n(x,y)\, d\lambda(y) \right) d\mu(x) \\ &= \int_X \lim_{n\to\infty} \left(\int_Y f_n(x,y)\, d\lambda(y) \right) d\mu(x) \\ &= \int_X \left(\int_Y f(x,y)\, d\lambda(y) \right) d\mu(x)\end{aligned}$$

Continued...

Therefore properties (i), (ii), and (iii) hold for every non-negative Ω-measurable function f. ☐

Theorem 41 (Tonelli's Theorem: Horizontal Sections)**.** *Let* $f : X \times Y \to [0, \infty]$ *be* Ω *measurable. Then,*

(i) $\forall y \in Y$*, we have* $f(\cdot, y) : X \to [0, \infty]$ *is* $\mathcal{A}$ *measurable.*

(ii) The function $y \mapsto \int_X f(x, y) d\mu(x)$ *is* $\mathcal{B}$ *measurable.*

(iii) $\displaystyle\int_{X\times Y} f \, d\eta = \int_Y \left(\int_X f(x, y) \, d\mu(x) \right) d\lambda(y)$

Theorem 42 (Fubini's Theorem: Vertical Sections)**.** *Let* $(X, \mathcal{A}, \mu)$ *and* $(Y, \mathcal{B}, \lambda)$ *be finite measure spaces. If* $f : X \times Y \to \mathbb{R}$ *is* Ω*-measurable and* $\int_{X\times Y} |f| \, d\eta < \infty$*, then:*

(i) For almost every $x \in X$*, the section* $f(x, \cdot)$ *is* λ*-integrable (meaning it is* $\mathcal{B}$*-measurable and its integral is finite).*

(ii) The function $x \mapsto \int_Y f(x, y) \, d\lambda(y)$ *is defined for almost every* $x \in X$*, and it is* $\mathcal{A}$*-measurable.*

(iii) The iterated integral exists and equals the integral over the product space:

$$\int_{X\times Y} f \, d\eta = \int_X \left(\int_Y f(x, y) \, d\lambda(y) \right) d\mu(x)$$

Proof. We define the positive and negative parts of f by:

$$f^+(x, y) = \max\{f(x, y), 0\}, \qquad f^-(x, y) = \max\{-f(x, y), 0\}$$

Then

$$f = f^+ - f^-, \qquad |f| = f^+ + f^-$$

Since

$$\int_{X\times Y} |f| \, d\eta < \infty$$

it follows that,

$$\int_{X\times Y} f^+ \, d\eta < \infty, \qquad \int_{X\times Y} f^- \, d\eta < \infty$$

(i) Since $|f|$ is a non-negative Ω-measurable function, Tonelli's theorem (Theorem 40) gives,

$$\int_{X\times Y} |f| \, d\eta = \int_X \left(\int_Y |f(x, y)| \, d\lambda(y) \right) d\mu(x)$$

The left-hand side is finite by assumption. Therefore,

$$\int_Y |f(x,y)|\, d\lambda(y) < \infty$$

for μ-almost every $x \in X$; otherwise the outer integral would be infinite. By Theorem 40(i), the section $f(x,\cdot)$ is $\mathcal{B}$-measurable for every x. Hence $f(x,\cdot)$ is λ-integrable for μ-almost every x.

(ii) Applying Theorem 40(ii) to f^+ and f^-, the functions

$$x \mapsto \int_Y f^+(x,y)\, d\lambda(y)$$

and

$$x \mapsto \int_Y f^-(x,y)\, d\lambda(y)$$

are $\mathcal{A}$-measurable.
By part (i), both integrals are finite for μ-almost every x. Hence,

$$x \mapsto \int_Y f(x,y)\, d\lambda(y) = \int_Y f^+(x,y)\, d\lambda(y) - \int_Y f^-(x,y)\, d\lambda(y)$$

is defined for μ-almost every x and is $\mathcal{A}$-measurable.

(iii) Applying Tonelli's theorem (Theorem 40) to f^+ and f^- gives,

$$\int_{X\times Y} f^+\, d\eta = \int_X \left(\int_Y f^+(x,y)\, d\lambda(y) \right) d\mu(x)$$

and

$$\int_{X\times Y} f^-\, d\eta = \int_X \left(\int_Y f^-(x,y)\, d\lambda(y) \right) d\mu(x)$$

Since all four integrals are finite, we may subtract the two equalities. Using linearity of the integral,

$$\begin{aligned} \int_{X\times Y} f\, d\eta &= \int_{X\times Y} (f^+ - f^-)\, d\eta \\ &= \int_X \left(\int_Y f^+(x,y)\, d\lambda(y) - \int_Y f^-(x,y)\, d\lambda(y) \right) d\mu(x) \\ &= \int_X \left(\int_Y f(x,y)\, d\lambda(y) \right) d\mu(x) \end{aligned}$$

Therefore properties (i), (ii), and (iii) hold. □

Theorem 43 (Fubini's Theorem: Horizontal Sections)**.** *Let* $(X, \mathcal{A}, \mu)$ *and* $(Y, \mathcal{B}, \lambda)$ *be finite measure spaces. If* $f : X \times Y \to \mathbb{R}$ *is* Ω*-measurable and* $\int_{X\times Y} |f|\, d\eta < \infty$*, then:*

(i) For almost every $y \in Y$*, the section* $f(\cdot, y)$ *is* μ*-integrable (meaning it is* $\mathcal{A}$*-measurable and its integral is finite).*

(ii) The function $y \mapsto \int_X f(x, y)\, d\mu(x)$ *is defined for almost every* $y \in Y$*, and it is* $\mathcal{B}$*-measurable.*

(iii) The iterated integral exists and equals the integral over the product space:

$$\int_{X\times Y} f\, d\eta = \int_Y \left(\int_X f(x, y)\, d\mu(x) \right) d\lambda(y)$$

Remark 4 (Completion of Product Measures)**.** Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \lambda)$ be finite measure spaces. If $\overline{\mu}$ and $\overline{\lambda}$ denote the completions of μ and λ, then the completion of the product measure $\mu \times \lambda$ is closely related to the product measure $\overline{\mu} \times \overline{\lambda}$.
(*A complete treatment requires a more careful analysis of null sets in product spaces and is beyond the scope of these notes.*)

Definition 40 (Signed Measure)**.** Let X be a set with σ-algebra $\mathcal{A}$. A function $\nu : \mathcal{A} \to \mathbb{R}$ is called a signed measure on $(X, \mathcal{A})$ if $\nu(\varnothing) = 0$ and whenever $\{E_j\}_{j=1}^{\infty} \subseteq \mathcal{A}$ are pairwise disjoint, we have,

$$\nu\left(\bigcup_{j=1}^{\infty} E_j \right) = \sum_{j=1}^{\infty} \nu(E_j)$$

Note*: In some texts, signed measures are allowed to take the values $\pm\infty$ (but not both). In these notes we restrict ourselves to finite signed measures, i.e. measures taking values only in $\mathbb{R}$.

Theorem 44 (Boundedness of the Range of Signed Measures)**.** *Let* X *be a set with* σ*-algebra* $\mathcal{A}$ *and Let* $\nu : \mathcal{A} \to \mathbb{R}$ *be a signed measure. Let* $S = \{\nu(E) : E \in \mathcal{A}\}$*. Then,* S *is bounded.*

Proof. Assume that S is unbounded. For $A \in \mathcal{A}$, define,

$$S_A = \{\nu(E) : E \in \mathcal{A},\ E \subseteq A\}$$

Since $S = S_X$ is unbounded, we start with $A_0 = X$.
We now construct pairwise disjoint sets $\{B_n\}_{n=1}^{\infty}$ such that, $|\nu(B_n)| > 1$, for every n.

Assume $A_{n-1} \in \mathcal{A}$ has the property that $S_{A_{n-1}}$ is unbounded. Then there exists a measurable set $C_n \subseteq A_{n-1}$, such that,

$$|\nu(C_n)| > |\nu(A_{n-1})| + 1$$

Now split

$$A_{n-1} = C_n \cup (A_{n-1} \setminus C_n)$$

where the union is disjoint. Since

$$\nu(A_{n-1}) = \nu(C_n) + \nu(A_{n-1} \setminus C_n)$$

we have,

$$|\nu(A_{n-1} \setminus C_n)| = |\nu(A_{n-1}) - \nu(C_n)| \geq |\nu(C_n)| - |\nu(A_{n-1})| > 1$$

At least one of the two collections S_{C_n} or $S_{A_{n-1} \setminus C_n}$ is unbounded. Otherwise, their union would imply that $S_{A_{n-1}}$ is bounded, since every measurable subset of A_{n-1} can be written as the disjoint union of a subset of C_n and a subset of $A_{n-1} \setminus C_n$.

- If S_{C_n} is unbounded, set

$$A_n = C_n, \qquad B_n = A_{n-1} \setminus C_n$$

- If $S_{A_{n-1} \setminus C_n}$ is unbounded, set

$$A_n = A_{n-1} \setminus C_n, \qquad B_n = C_n$$

In either case, S_{A_n} is unbounded and $|\nu(B_n)| > 1$.
Since $A_n \subseteq A_{n-1}$ and $B_n \subseteq A_{n-1} \setminus A_n$, the sets B_n are pairwise disjoint.
Let $E = \displaystyle\bigcup_{n=1}^{\infty} B_n$.
By countable additivity, $\nu(E) = \displaystyle\sum_{n=1}^{\infty} \nu(B_n)$.
Since $\nu(E) \in \mathbb{R}$, the series, $\displaystyle\sum_{n=1}^{\infty} \nu(B_n)$ converges. Hence, $\nu(B_n) \to 0$.
This contradicts $|\nu(B_n)| > 1$, for every n.
Therefore S is bounded. □

Theorem 45 (Continuity of Signed Measures)**.** *Signed measures are continuous, i.e., if* $\nu : \mathcal{A} \to \mathbb{R}$ *is a signed measure, then,* *(i) If* $E_1 \subseteq E_2 \subseteq \cdots$ *and* $E = \bigcup_{j=1}^{\infty} E_j$*, then,* $\nu(E_n) \to \nu(E)$*.*

(ii) If $E_1 \supseteq E_2 \supseteq \cdots$ *and* $E = \bigcap_{j=1}^{\infty} E_j$*, then,* $\nu(E_n) \to \nu(E)$*.*

Proof. We consider 2 cases:

- **Case I:** Let $E_1 \subseteq E_2 \subseteq \cdots$ and $E = \bigcup_{j=1}^{\infty} E_j$, where $E_j \in \mathcal{A}$ and $E \in \mathcal{A}$. Then,

  $$E = E_1 \cup (E_2 \setminus E_1) \cup (E_3 \setminus E_2) \cup \cdots$$

  and for each n,

  $$E_n = E_1 \cup (E_2 \setminus E_1) \cup \cdots \cup (E_n \setminus E_{n-1})$$

  Hence, using additivity of the signed measure ν on the disjoint union,

  $$\nu(E) = \nu(E_1) + \sum_{j=1}^{\infty} \nu\big(E_{j+1} \setminus E_j\big)$$

  and for each n,

  $$\nu(E_n) = \nu(E_1) + \sum_{j=1}^{n-1} \nu\big(E_{j+1} \setminus E_j\big)$$

  By definition the infinite series $\sum_{j=1}^{\infty} \nu\big(E_{j+1} \setminus E_j\big)$ must converge, so

  $$\lim_{n\to\infty} \sum_{j=1}^{n-1} \nu\big(E_{j+1} \setminus E_j\big) = \sum_{j=1}^{\infty} \nu\big(E_{j+1} \setminus E_j\big)$$

  Therefore,

  $$\begin{aligned}\lim_{n\to\infty} \nu(E_n) &= \lim_{n\to\infty} \Big[\nu(E_1) + \sum_{j=1}^{n-1} \nu\big(E_{j+1} \setminus E_j\big)\Big] \\ &= \nu(E_1) + \sum_{j=1}^{\infty} \nu\big(E_{j+1} \setminus E_j\big) \\ &= \nu(E)\end{aligned}$$

  Thus $\nu(E_n) \to \nu(E)$ as $n \to \infty$.

- **Case II:** Let $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$ and $E = \bigcap_{j=1}^{\infty} E_j$, where $E_j \in \mathcal{A}$ and $E \in \mathcal{A}$.

  Then,

  $$E_1 = \Big(\bigcap_{j=1}^{\infty} E_j\Big) \cup (E_1 \setminus E_2) \cup (E_2 \setminus E_3) \cup \cdots = E \cup (E_1 \setminus E_2) \cup (E_2 \setminus E_3) \cup \cdots$$

  and for each n,

  $$E_1 = (E_1 \setminus E_2) \cup (E_2 \setminus E_3) \cup \cdots \cup (E_{n-1} \setminus E_n) \cup E_n$$

  Thus, using additivity of the signed measure ν on the disjoint union,

  $$\nu(E_1) = \nu(E) + \sum_{j=1}^{\infty} \nu\big(E_j \setminus E_{j+1}\big)$$

  and for each n,

  $$\nu(E_1) = \nu(E_n) + \sum_{j=1}^{n-1} \nu\big(E_j \setminus E_{j+1}\big)$$

  By definition the infinite series $\sum_{j=1}^{\infty} \nu(E_j \setminus E_{j+1})$ must converge, so,

  $$\lim_{n\to\infty} \sum_{j=1}^{n-1} \nu\big(E_j \setminus E_{j+1}\big) = \sum_{j=1}^{\infty} \nu\big(E_j \setminus E_{j+1}\big)$$

  Hence,

  $$\nu(E_1) = \lim_{n\to\infty} \left[\nu(E_n) + \sum_{j=1}^{n-1} \nu\big(E_j \setminus E_{j+1}\big)\right] = \lim_{n\to\infty} \nu(E_n) + \sum_{j=1}^{\infty} \nu\big(E_j \setminus E_{j+1}\big)$$

  Comparing with $\nu(E_1) = \nu(E) + \sum_{j=1}^{\infty} \nu(E_j \setminus E_{j+1})$ we conclude,

  $$\lim_{n\to\infty} \nu(E_n) = \nu(E)$$

Therefore, in both cases, $\lim_{n\to\infty} \nu(E_n) = \nu(E)$.
Therefore, signed measures are continuous. □

Theorem 46 (Existence of Maximal Positive Subsets)**.** *Let X be a set with σ-algebra $\mathcal{A}$ and a signed measure ν. Let $E \in \mathcal{A}$. Then there exists, $E_+ \subseteq E$, $E_+ \in \mathcal{A}$, such that,*

$$s = \nu(E_+) = \sup\{\nu(E') : \ E' \in \mathcal{A}, \ E' \subseteq E\} < \infty$$

Proof. Choose a sequence of epsilons $\{\varepsilon_j\}_{j=1}^{\infty}$ with $\varepsilon_j \downarrow 0$ and $\sum_{j=1}^{\infty} \varepsilon_j < \infty$.
For each corresponding index k choose $E_k \subseteq E$, $E_k \in \mathcal{A}$, such that,

$$\nu(E_k) \geq s - \varepsilon_k$$

Thus, we have a sequence of subsets of E, $\{E_k\}_{k=1}^{\infty}$, with,

$$\nu(E_k) \geq s - \varepsilon_k, \qquad \text{for all } k = 1, 2, \ldots$$

Consider any two sets E_k and E_ℓ, and Let

$$A := E_k \setminus E_\ell, \qquad B := E_k \cap E_\ell, \qquad C := E_\ell \setminus E_k$$

Then, $A, B, C, A \cup B, A \cup C, B \cup C, A \cup B \cup C \subseteq E$. Therefore,

$$\begin{align} \nu(A), \ \nu(B), \ \nu(C) &\leq s \tag{1} \\ \nu(A) + \nu(B) + \nu(C) &\leq s \tag{2} \\ \nu(A) + \nu(B) &\geq s - \varepsilon_k \tag{3} \\ \nu(B) + \nu(C) &\geq s - \varepsilon_\ell \tag{4} \end{align}$$

From equation (2) and equation (3), $\nu(C) \leq \varepsilon_k$.
From equation (2) and equation (4), $\nu(A) \leq \varepsilon_\ell$.
From equation (1) and equation (3), $\nu(A) \geq s - \varepsilon_k - \nu(B) \geq -\varepsilon_k$.
From equation (1) and equation (4), $\nu(C) \geq -\varepsilon_\ell$.
Therefore, the measures $\nu(A)$ and $\nu(C)$ can't be large positive or large negative. Thus, we have,

$$-\varepsilon_k \leq \nu(A) \leq \varepsilon_\ell \qquad \text{and} \qquad -\varepsilon_\ell \leq \nu(C) \leq \varepsilon_k$$

Then,

$$\nu(B) = \nu(A) + \nu(B) - \nu(A) \geq s - \varepsilon_k - \varepsilon_\ell$$

Thus, for any k and ℓ, if $E_k, E_\ell \subseteq E$ with $\nu(E_k) \geq s - \varepsilon_k$ and $\nu(E_\ell) \geq s - \varepsilon_\ell$, then,

$$\nu(E_k \cap E_\ell) \geq s - \varepsilon_k - \varepsilon_\ell$$

Thus, more generally, if $P, Q \subseteq E$ are measurable sets with $\nu(P) \geq s - \alpha$ and $\nu(Q) \geq s - \beta$, then,

$$\nu(P \cap Q) \geq s - \alpha - \beta$$

Now, consider the sets $E_1, \ E_1 \cap E_2, \ E_1 \cap E_2 \cap E_3, \ldots$. Then, by induction,

$$\nu(E_1) \geq s - \varepsilon_1,$$

$$\nu(E_1 \cap E_2) \geq s - \varepsilon_1 - \varepsilon_2,$$
$$\nu(E_1 \cap E_2 \cap E_3) \geq s - \varepsilon_1 - \varepsilon_2 - \varepsilon_3$$

Now as the sequence goes down then, let

$$E^{(1)} = \bigcap_{j=1}^{\infty} E_j \implies \nu\big(E^{(1)}\big) \geq s - \Big(\sum_{j=1}^{\infty} \varepsilon_j\Big)$$

Similarly, let

$$E^{(2)} = \bigcap_{j=2}^{\infty} E_j \implies \nu\big(E^{(2)}\big) \geq s - \Big(\sum_{j=2}^{\infty} \varepsilon_j\Big)$$

Similarly, if

$$E^{(n)} = \bigcap_{j=n}^{\infty} E_j \implies \nu\big(E^{(n)}\big) \geq s - \Big(\sum_{j=n}^{\infty} \varepsilon_j\Big)$$

Then,

$$E^{(1)} \subseteq E^{(2)} \subseteq E^{(3)} \subseteq \cdots \subseteq E^{(n)} \subseteq \cdots$$

and Let

$$E_+ = \bigcup_{m=1}^{\infty} E^{(m)} \subseteq E$$

Then,

$$\nu(E_+) \leq s$$

And,

$$\nu(E_+) = \lim_{m\to\infty} \nu\big(E^{(m)}\big) \geq \lim_{m\to\infty} \Big(s - \sum_{j=m}^{\infty} \varepsilon_j\Big) = s$$

Since the tail sums $\displaystyle\sum_{j=m}^{\infty} \varepsilon_j \longrightarrow 0$ as $m \longrightarrow \infty$.

Therefore,

$$s \leq \nu(E_+) \leq s$$

Therefore, we have found $E_+ \subseteq E$ such that $\nu(E_+) = s$. ☐

Lemma 9 (Maximal Subset is Positive)**.** *Let E_+ be as in Theorem 46. Then E_+ is a positive set, i.e., $\nu(A) \geq 0$, for every $A \in \mathcal{A}$ with $A \subseteq E_+$.*

Proof. Let $A \in \mathcal{A}$ with $A \subseteq E_+$. Suppose, for contradiction, that $\nu(A) < 0$.
Since $A \subseteq E_+$, we can write,

$$E_+ = A \cup (E_+ \setminus A)$$

where the union is disjoint. Hence,

$$\nu(E_+) = \nu(A) + \nu(E_+ \setminus A)$$

Therefore,

$$\nu(E_+ \setminus A) = \nu(E_+) - \nu(A) > \nu(E_+) = s$$

But, $E_+ \setminus A \subseteq E$ and $E_+ \setminus A \in \mathcal{A}$, which contradicts the definition of

$$s = \sup\{\nu(E') : \ E' \in \mathcal{A}, \ E' \subseteq E\}$$

Thus, no such A exists, and therefore, $\nu(A) \geq 0$, for every measurable $A \subseteq E_+$. $\square$

Theorem 47 (Hahn-Jordan Decomposition)**.** *Let $(X, \mathcal{A})$ be a measurable space and ν be a signed measure on this space. Then there exist two sets X_+ and X_- in $\mathcal{A}$, and two unique positive measures ν_+ and ν_- such that;*

(i) $X = X_+ \cup X_-$;

(ii) $X_+ \cap X_- = \varnothing$;

(iii) for every $E \subseteq X_+$ *with* $E \in \mathcal{A}$, $\nu(E) \geq 0$;

(iv) for every $E \subseteq X_-$ *with* $E \in \mathcal{A}$, $\nu(E) \leq 0$;

(v) $\nu = \nu_+ - \nu_-$;

(vi) $\nu_+(X_-) = \nu_-(X_+) = 0$.

Proof. Let

$$s = \sup\{\nu(E) : E \in \mathcal{A}, \ E \subseteq X\}$$

Then, by Theorem 46, there exists $X_+ \in \mathcal{A}$ such that $\nu(X_+) = s$.
Define $X_- = X \setminus X_+$. For any $E \in \mathcal{A}$ with $E \subseteq X_-$, suppose that, $\nu(E) > 0$.
Since $E \cap X_+ = \varnothing$, we have,

$$\nu(X_+ \cup E) = \nu(X_+) + \nu(E) > s$$

which contradicts the definition of

$$s = \sup\{\nu(F) : F \in \mathcal{A}, \ F \subseteq X\}$$

Therefore,

$$\nu(E) \leq 0 \qquad \text{for every measurable } E \subseteq X_-$$

Also, by Lemma 9,

$$\nu(E) \geq 0 \qquad \text{for every measurable } E \subseteq X_+$$

Now if ν is a signed measure then $\forall A \in \mathcal{A}$, $\nu_A(E) = \nu(A \cap E)$ is a signed measure. Thus, define

$$\nu_+(E) = \nu(E \cap X_+) \qquad \text{and} \qquad \nu_-(E) = -\nu(E \cap X_-)$$

Since X_+ is a positive set and X_- is a negative set, both ν_+ and ν_- are non-negative set functions. Moreover, Since $E \mapsto \nu(E \cap A)$ is a signed measure whenever ν is a signed measure, both ν_+ and ν_- are measures.
Also,

$$\begin{aligned} \nu(E) &= \nu(E \cap X_+) + \nu(E \cap X_-) \\ &= \nu_+(E) - \nu_-(E) \end{aligned}$$

Therefore, $X = X_+ \cup X_-$ and $X_+ \cap X_- = \varnothing$, $X_+, X_- \in \mathcal{A}$ and, $\nu_+(X_-) = \nu_-(X_+) = 0$.

To prove uniqueness, suppose

$$\nu = \nu_+ - \nu_- = \nu'_+ - \nu'_-$$

are two decompositions satisfying

$$\nu_+(X_-) = \nu_-(X_+) = 0$$

and

$$\nu'_+(X'_-) = \nu'_-(X'_+) = 0$$

Let $E \in \mathcal{A}$. Since

$$\nu(E \cap X_+) = \nu_+(E \cap X_+) - \nu_-(E \cap X_+) = \nu_+(E \cap X_+)$$

and

$$\nu_+(E \cap X_+) = \nu_+(E)$$

we obtain,

$$\nu_+(E) = \nu(E \cap X_+)$$

Similarly,

$$\nu'_+(E) = \nu\big(E \cap X'_+\big)$$

Now,

$$E \cap X_+ = \big(E \cap X_+ \cap X'_+\big) \cup \big(E \cap X_+ \cap X'_-\big)$$

Since $X_+ \cap X'_-$ is both a positive set and a negative set, every measurable subset of it has measure 0. Indeed, if $A \subseteq X_+ \cap X'_-$ is measurable, then $\nu(A) \geq 0$ because $A \subseteq X_+$ and $\nu(A) \leq 0$ because $A \subseteq X'_-$, hence $\nu(A) = 0$.
Hence,

$$\nu\big(E \cap X_+ \cap X'_-\big) = 0$$

Continued...

Therefore,

$$\nu_+(E) = \nu\big(E \cap X_+ \cap X'_+\big)$$

By symmetry,

$$\nu'_+(E) = \nu\big(E \cap X_+ \cap X'_+\big)$$

Thus,

$$\nu_+(E) = \nu'_+(E) \qquad \forall E \in \mathcal{A}$$

Similarly,

$$\nu_-(E) = \nu'_-(E) \qquad \forall E \in \mathcal{A}$$

Hence the decomposition $\nu = \nu_+ - \nu_-$ is unique. □

Theorem 48 (Lebesgue Decomposition Theorem)**.** *Let μ and η be two finite measures on $(X, \mathcal{A})$. Then there exist a nonnegative measurable function f and a nonnegative measure η_s and a set $X_s \in \mathcal{A}$ such that*

(i) $\eta = f\,\mu + \eta_s$,

(ii) $\mu(X_s) = 0 \quad \textit{and} \quad \eta_s(X_s^c) = 0.$

The measure η_s is called the singular part of η with respect to μ, and $\eta_a := f\mu$ is called the absolutely continuous part and the measures η_a and η_s are unique.

Proof. **Existence:** Let $f \geq 0$ be a measurable function. We say f is admissible if

$$\forall E \in \mathcal{A}, \qquad \int_E f\,d\mu \leq \eta(E)$$

The existence of such a function is guaranteed since the zero function is admissible. Let f_1 and f_2 be two admissible functions and set

$$f = \max\{f_1, f_2\}$$

Then, $\forall E \in \mathcal{A}$, by splitting E into the two measurable pieces $E \cap \{f_1 \geq f_2\}$ and $E \cap \{f_1 < f_2\}$, we get,

$$\begin{aligned}\int_E f\,d\mu &= \int_{E\cap\{f_1\geq f_2\}} f\,d\mu + \int_{E\cap\{f_1<f_2\}} f\,d\mu \\ &= \int_{E\cap\{f_1\geq f_2\}} f_1\,d\mu + \int_{E\cap\{f_1<f_2\}} f_2\,d\mu\end{aligned}$$

Using admissibility of f_1 and f_2 we get

$$\int_E f\,d\mu \leq \eta\big(E \cap \{f_1 \geq f_2\}\big) + \eta\big(E \cap \{f_1 < f_2\}\big) = \eta(E)$$

Hence $f = \max\{f_1, f_2\}$ is admissible.
Therefore finite maxima of admissible functions are admissible.
Let $\{f_j\}$ be an increasing sequence of admissible functions. Let

$$f = \sup_{j\geq 1} f_j = \lim_{j\to\infty} f_j$$

Since we can replace any sequence with its partial maxima (g_n), we may assume without loss of generality that $\{f_j\}$ is an increasing sequence. Thus, for every $E \in \mathcal{A}$, using Monotone Convergence Theorem (Theorem 28)

$$\int_E f\, d\mu = \int_E \big(\lim_{j\to\infty} f_j\big)\, d\mu = \lim_{j\to\infty} \int_E f_j\, d\mu \leq \lim_{j\to\infty} \eta(E) \leq \eta(E)$$

Therefore $f = \lim\limits_{j\to\infty} f_j$ is admissible.
Consider a sequence of admissible functions $\{f_k\}$. Let

$$s = \sup\Big\{\int_X \widetilde{f}\, d\mu : \widetilde{f} \text{ is admissible}\Big\}$$

Choose $\{f_k\}$ such that $\displaystyle\int_X f_k\, d\mu \longrightarrow s$ as $k \longrightarrow \infty$. For each n define,

$$g_n = \max\{f_1, f_2, \ldots, f_n\}$$

Since each f_i is admissible and the maximum of admissible functions is admissible, each g_n is admissible. Moreover, for every $i = 1, 2, \ldots, n$,

$$\int_X g_n\, d\mu \geq \int_X f_i\, d\mu$$

Also, g_n is an increasing sequence of admissible functions. Let $f = \lim\limits_{n\to\infty} g_n$.
By the previous argument, f is admissible. Hence,

$$\int_X f\, d\mu \leq s$$

On the other hand, since $g_n \geq f_n$, we have,

$$\int_X f\, d\mu \geq \int_X g_n\, d\mu \geq \int_X f_n\, d\mu$$

Taking $n \to \infty$, we obtain,

$$\int_X f\, d\mu \geq s$$

Therefore,

$$s \leq \int_X f\, d\mu \leq s$$

and hence,

$$\int_X f\,d\mu = s$$

Now, take $\varepsilon > 0$ and define $f_\varepsilon := f + \varepsilon$. Also, define,

$$\eta_\varepsilon := \eta - f_\varepsilon\,\mu$$

This defines a finite signed measure. Then, by the Hahn-Jordan Decomposition theorem (Theorem 47),

$$\eta_\varepsilon = \eta_\varepsilon^+ - \eta_\varepsilon^-$$

and

$$X = X_\varepsilon^+ \cup X_\varepsilon^-, \qquad X_\varepsilon^+ \cap X_\varepsilon^- = \varnothing, \quad \text{with} \quad \eta_\varepsilon^\pm(X_\varepsilon^\mp) = 0$$

We claim that $\mu(X_\varepsilon^+) = 0$.
Assume not. So, consider,

$$g_\varepsilon := f + \varepsilon\,\mathbb{1}_{X_\varepsilon^+}$$

Then, for any measurable subset $A \subseteq X_\varepsilon^+$, $g_\varepsilon = f_\varepsilon$ on A, and for any measurable subset $B \subseteq X_\varepsilon^-$, $g_\varepsilon = f$ on B. Then, $\forall E \in \mathcal{A}$,

$$\int_E g_\varepsilon\,d\mu = \int_{E\cap X_\varepsilon^+} g_\varepsilon\,d\mu + \int_{E\cap X_\varepsilon^-} g_\varepsilon\,d\mu = \int_{E\cap X_\varepsilon^+} f_\varepsilon\,d\mu + \int_{E\cap X_\varepsilon^-} f\,d\mu$$

Now, $\eta_\varepsilon := \eta - f_\varepsilon\,\mu = \eta_\varepsilon^+ - \eta_\varepsilon^-$, and as $E \cap X_\varepsilon^+ \subseteq X_\varepsilon^+$, we have $\eta_\varepsilon^-(E \cap X_\varepsilon^+) = 0$. Thus,

$$\eta(E \cap X_\varepsilon^+) - (f_\varepsilon\,\mu)(E \cap X_\varepsilon^+) = \eta_\varepsilon^+(E \cap X_\varepsilon^+) \geq 0$$

Hence,

$$(f_\varepsilon\,\mu)(E \cap X_\varepsilon^+) \leq \eta(E \cap X_\varepsilon^+) \implies \int_{E\cap X_\varepsilon^+} f_\varepsilon\,d\mu \leq \eta(E \cap X_\varepsilon^+)$$

And, by admissibility of f, we have,

$$\int_{E\cap X_\varepsilon^-} f\,d\mu \leq \eta(E \cap X_\varepsilon^-)$$

Therefore,

$$\int_E g_\varepsilon\,d\mu \leq \eta(E \cap X_\varepsilon^+) + \eta(E \cap X_\varepsilon^-) = \eta(E)$$

Thus, g_ε is admissible. Therefore, by the maximality of f,

$$\int_X g_\varepsilon\,d\mu \leq \int_X f\,d\mu$$

But,

$$\int_X g_\varepsilon\,d\mu = \int_X f\,d\mu + \varepsilon\,\mu(X_\varepsilon^+)$$

Hence,

$$\int_X f\,d\mu + \varepsilon\,\mu(X_\varepsilon^+) \le \int_X f\,d\mu$$

Therefore,

$$\mu(X_\varepsilon^+) = 0$$

Now, if $E \subseteq X_\varepsilon^-$, then E is a negative set for the signed measure,

$$\eta_\varepsilon = \eta - (f+\varepsilon)\mu$$

Therefore,

$$\eta_\varepsilon(E) \le 0$$

Hence,

$$\eta(E) - \int_E (f+\varepsilon)\,d\mu \le 0$$

Thus,

$$\eta(E) \le \int_E f\,d\mu + \varepsilon\,\mu(E) \le \int_E f\,d\mu + \varepsilon\,\mu(X)$$

Again, f being admissible,

$$\int_E f\,d\mu \le \eta(E)$$

Then, choosing a sequence of epsilons $\{\varepsilon_n\}$ such that $\varepsilon_n \downarrow 0$ and taking

$$X_s = \bigcup_{n=1}^{\infty} X_{\varepsilon_n}^+$$

we get

$$\mu(X_s) \le \sum_{n=1}^{\infty} \mu\big(X_{\varepsilon_n}^+\big) = 0$$

Therefore, $\mu(X_s) = 0$.
Now if $E \subseteq X_s^c$, then,

$$\eta(E) \ge \int_E f\,d\mu$$

and for every n,

$$\eta(E) \le \int_E f\,d\mu + \varepsilon_n\,\mu(X)$$

So, as $n \to \infty$,

$$\eta(E) \le \int_E f\,d\mu \le \eta(E)$$

Therefore,

$$\eta(E) = (f\mu)(E)$$

Continued...

Now, for every $E \in \mathcal{A}$, define

$$\eta_s(E) = \eta(E) - (f\mu)(E) = \eta(E) - \int_E f\,d\mu$$

Since f is admissible, we have,

$$\int_E f\,d\mu \leq \eta(E) \qquad \forall E \in \mathcal{A}$$

Hence,

$$\eta_s(E) \geq 0 \qquad \forall E \in \mathcal{A}$$

Therefore, η_s is a non-negative measure.
In particular, if $E \subseteq X_s^c$, then,

$$\eta(E) = \int_E f\,d\mu$$

and hence,

$$\eta_s(E) = 0$$

And,

$$\eta_s(X_s^c) = \eta(X_s^c) - (f\mu)(X_s^c) = \eta(X_s^c) - \eta(X_s^c) = 0$$

Therefore, $\eta = \ f\mu + \eta_s$, where X_s is the singular set. In particular $\mu(X_s) = 0$ and $\eta_s(X_s^c) = 0$.
This proves the existence of the decomposition.

Uniqueness: Suppose

$$\eta = f_1\mu + \eta_s^{(1)} = f_2\mu + \eta_s^{(2)}$$

Let $\eta_a^{(1)} = f_1\mu$ and $\eta_a^{(2)} = f_2\mu$. Then,

$$\eta = \eta_a^{(1)} + \eta_s^{(1)} = \eta_a^{(2)} + \eta_s^{(2)}$$

Let $X_s^{(1)}$ and $X_s^{(2)}$ be singular sets for $\eta_s^{(1)}$ and $\eta_s^{(2)}$, respectively. Define,

$$X_s = X_s^{(1)} \cup X_s^{(2)}$$

Then,

$$\mu(X_s) = 0$$

Also,

$$\eta_s^{(1)}(X_s^c) = 0 \qquad \text{and} \qquad \eta_s^{(2)}(X_s^c) = 0$$

Let $E \in \mathcal{A}$. Since $\eta_a^{(1)} \ll \mu$ and $\eta_a^{(2)} \ll \mu$, we have,

$$\eta_a^{(1)}(E \cap X_s) = 0 \qquad \text{and} \qquad \eta_a^{(2)}(E \cap X_s) = 0$$

Continued...

Therefore,

$$\begin{aligned}\eta_s^{(1)}(E) &= \eta_s^{(1)}(E \cap X_s) \\ &= \eta(E \cap X_s) \\ &= \eta_s^{(2)}(E \cap X_s) \\ &= \eta_s^{(2)}(E)\end{aligned}$$

Hence,

$$\eta_s^{(1)} = \eta_s^{(2)}$$

Consequently,

$$\eta_a^{(1)} = \eta - \eta_s^{(1)} = \eta - \eta_s^{(2)} = \eta_a^{(2)}$$

Thus the absolutely continuous part and the singular part are unique.

□

Theorem 49 (Radon–Nikodym Density Theorem)**.** *Let μ and η be two finite measures on $(X, \mathcal{A})$ such that, $\eta \ll \mu$. Then there exists a nonnegative measurable function f such that,*

$$\eta(E) = \int_E f \, d\mu \qquad \forall E \in \mathcal{A}$$

The function f is unique up to equality μ-almost everywhere.

Proof. By Theorem 48, there exist a nonnegative measurable function f, a measure η_s, and a measurable set X_s such that,

$$\eta = f\mu + \eta_s$$

where,

$$\mu(X_s) = 0 \qquad \text{and} \qquad \eta_s(X_s^c) = 0$$

Since $\eta \ll \mu$ and $\mu(X_s) = 0$, we get,

$$\eta(X_s) = 0$$

But η_s is supported on X_s, so,

$$\eta_s(X) = \eta_s(X_s) = 0$$

Thus $\eta_s = 0$, and hence

$$\eta = f\mu$$

Let f_1 and f_2 be two nonnegative measurable functions. Assume that for every $E \in \mathcal{A}$

$$\int_E f_1 \, d\mu = \int_E f_2 \, d\mu < \infty$$

In particular $\int f_1\,d\mu = \int f_2\,d\mu < \infty$, so each f_i is finite μ–a.e.; hence, $\mu\{f_1 = \infty\} = \mu\{f_2 = \infty\} = 0$.
Fix $\varepsilon > 0$ and consider the measurable set

$$E_\varepsilon := \{x \in X : \; f_1(x) \geq f_2(x) + \varepsilon\}$$

Then,

$$\int_{E_\varepsilon} f_1\,d\mu \geq \int_{E_\varepsilon} (f_2 + \varepsilon)\,d\mu = \int_{E_\varepsilon} f_2\,d\mu + \varepsilon\,\mu(E_\varepsilon)$$

But by assumption $\int_{E_\varepsilon} f_1\,d\mu = \int_{E_\varepsilon} f_2\,d\mu$, hence, $\varepsilon\,\mu(E_\varepsilon) = 0$, and so $\mu(E_\varepsilon) = 0$.
Take a sequence $\varepsilon_n \downarrow 0$. Then,

$$\{f_1 > f_2\} = \bigcup_{n=1}^{\infty}\{f_1 \geq f_2 + \varepsilon_n\} = \bigcup_{n=1}^{\infty} E_{\varepsilon_n}$$

Thus, $\mu\{f_1 > f_2\} = 0$ as a countable union of null sets. By symmetry the same argument gives $\mu\{f_2 > f_1\} = 0$. Therefore,

$$\mu\{f_1 \neq f_2\} = \mu(\{f_1 > f_2\} \cup \{f_2 > f_1\}) = 0$$

i.e. $f_1 = f_2$ μ–almost everywhere. This proves uniqueness up to a set of μ–measure zero. $\square$

Remark 5 (Change of the Singular Set)**.** Assume that the Lebesgue Decomposition Theorem (Theorem 48) holds. Then for a given decomposition, Let $X_s^{(1)}$ and $X_s^{(2)}$ be two singular sets. Let $\eta = f\,\mu + \eta_s$ be the decomposition. Then,

$$\mu\big(X_s^{(1)}\big) = \mu\big(X_s^{(2)}\big) = 0$$

Now,

$$X_s^{(1)} \Delta X_s^{(2)} = \big(X_s^{(1)} \setminus X_s^{(2)}\big) \cup \big(X_s^{(2)} \setminus X_s^{(1)}\big) \subseteq X_s^{(1)} \cup X_s^{(2)}$$

Hence,

$$\mu\big(X_s^{(1)} \Delta X_s^{(2)}\big) = 0$$

In particular,

$$X_s^{(1)} \setminus X_s^{(2)} \subseteq X_s^{(1)} \quad \text{so} \quad \mu\big(X_s^{(1)} \setminus X_s^{(2)}\big) = 0$$

and therefore,

$$(f\mu)\big(X_s^{(1)} \setminus X_s^{(2)}\big) = 0$$

Thus,

$$\eta\big(X_s^{(1)} \setminus X_s^{(2)}\big) = (f\mu)\big(X_s^{(1)} \setminus X_s^{(2)}\big) + \eta_s\big(X_s^{(1)} \setminus X_s^{(2)}\big) = \eta_s\big(X_s^{(1)} \setminus X_s^{(2)}\big) = 0$$

since $X_s^{(1)} \setminus X_s^{(2)} \subseteq (X_s^{(2)})^c$ and $\eta_s\big((X_s^{(2)})^c\big) = 0$.
Similarly $\eta_s\big(X_s^{(2)} \setminus X_s^{(1)}\big) = 0$, and therefore,

$$\eta\big(X_s^{(1)} \Delta X_s^{(2)}\big) = 0$$

Therefore the singular set X_s is unique up to a set that is null for both μ and η.

Remark 6 (Lebesgue Decomposition Theorem σ-finite version)**.** As μ and η are σ-finite, then

$$X = \bigsqcup_{i=1}^{\infty} X_i' = \bigsqcup_{i=1}^{\infty} Y_i$$

where $\mu(X_i') < \infty$ and $\eta(Y_i) < \infty,\ \forall i = 1, 2, \ldots$.
Considering all intersections between every X_i' and Y_i of every possible combination of indices, we get

$$X = \bigsqcup_{j=1}^{\infty} X_j \quad \text{s.t.} \quad \mu(X_j),\ \eta(X_j) < \infty$$

Let $\eta_j(E) = \eta(E \cap X_j)$ and $\mu_j(E) = \mu(E \cap X_j)$ (so μ_j and η_j are the restrictions to X_j).
By the finite version of Lebesgue decomposition Theorem (Theorem 48) on each X_j we obtain

$$\eta_j = f_j \mu_j + (\eta_j)_s \quad \text{on } X_j$$

where f_j is μ_j-integrable on X_j and $(\eta_j)_s$ is singular with respect to μ_j.
Define $f = f_j$ on X_j,

$$\eta_s(E) = \sum_{j=1}^{\infty} (\eta_j)_s(E \cap X_j) \qquad X_s = \bigcup_{j=1}^{\infty} (X_j)_s$$

where $(X_j)_s$ denotes a supporting set for $(\eta_j)_s$ (so $\mu_j((X_j)_s) = 0$).
Since μ_j is the restriction of μ to X_j, integration over subsets of X_j agrees for μ_j and μ. Also, by definition, $f = f_j$ on X_j. Hence, for every measurable $E \in \mathcal{A}$,

$$\int_{E \cap X_j} f_j \, d\mu_j = \int_{E \cap X_j} f_j \, d\mu = \int_{E \cap X_j} f \, d\mu$$

Then for any measurable $E \subseteq X$,

$$\begin{aligned}
\eta(E) &= \sum_{j=1}^{\infty} \eta_j(E \cap X_j) \\
&= \sum_{j=1}^{\infty} \left(\int_{E \cap X_j} f_j \, d\mu_j + (\eta_j)_s(E \cap X_j) \right) \\
&= \sum_{j=1}^{\infty} \left(\int_{E \cap X_j} f \, d\mu + (\eta_j)_s(E \cap X_j) \right) \\
&= \sum_{j=1}^{\infty} \int_{E \cap X_j} f \, d\mu + \sum_{j=1}^{\infty} (\eta_j)_s(E \cap X_j)
\end{aligned}$$

$$= \int_E f\, d\mu + \eta_s(E)$$

Now, each $(X_j)_s$ satisfies

$$\mu_j((X_j)_s) = 0$$

Since $(X_j)_s \subseteq X_j$,

$$\mu((X_j)_s) = \mu_j((X_j)_s) = 0$$

Therefore,

$$\mu(X_s) = \mu\Big(\bigcup_{j=1}^{\infty}(X_j)_s\Big) \leq \sum_{j=1}^{\infty} \mu((X_j)_s) = 0$$

Thus, $\mu(X_s) = 0$.
And,

$$\begin{aligned} \eta_s\big(X \setminus X_s\big) &= \eta_s\Big(X \setminus \bigcup_{j=1}^{\infty}(X_j)_s\Big) \\ &= \sum_{k=1}^{\infty}(\eta_k)_s\Big(\big(X \setminus \bigcup_{j=1}^{\infty}(X_j)_s\big) \cap X_k\Big) \\ &= \sum_{k=1}^{\infty}(\eta_k)_s\big(X_k \setminus (X_k)_s\big) \\ &= 0 \end{aligned}$$

Therefore,

$$\eta = f\,\mu + \eta_s$$

with,

$$\mu(X_s) = 0 \qquad \text{and} \qquad \eta_s(X \setminus X_s) = 0$$

Hence the Lebesgue decomposition theorem remains valid for σ-finite measures.

Theorem 50 (Partition of Unity)**.** *Let X be a metric space and Let $K \subseteq X$ be compact. Let $\{G_j\}_{j=1}^{N}$ be a finite open cover of K. Then there exist continuous functions $\varphi_j : X \to [0,1]$, $j = 1, \ldots, N$, such that,* $\operatorname{supp}\varphi_j \subseteq G_j$ *and* $\displaystyle\sum_{j=1}^{N} \varphi_j \equiv 1$ *on K.*

Proof. Let X be a metric space, $K \subseteq X$ with a finite open cover $\{G_j\}_{j=1}^{N}$. We show that we can shrink $\{G_j\}$ and still cover K. So, we aim to find for all j subsets of $\{G_j\}_{j=1}^{N}$, such that;

- $F_j \subseteq G_j$ is closed.

- $H_j \subseteq G_j$ is open
- $F_j \subseteq H_j$,
- $\text{Cl}(H_j) \subseteq G_j$, and
- $K \subseteq \bigcup_j F_j$.

By finding such sets, we can apply Urysohn's lemma (Lemma 8) to get functions which take values between 0 and 1, are identically 1 on F_j, and identically 0 on H_j^c. This gives us functions whose support is (strictly) contained in G_j; some further manipulations (overlap in the sum on K and continuity) give us the desired functions φ_j.

Consider these open sets:

$$H_{j,m} = \left\{ x \in G_j : \text{dist}(x, G_j^c) > \frac{1}{m} \right\}$$

Then, $H_{j,m} \subseteq G_j$, for all m.

Further, for all m, Let

$$\text{Cl}(H_{j,m}) = \overline{H}_{j,m} = \left\{ x \in G_j : \text{dist}(x, G_j^c) \geq \frac{1}{m} \right\} \subseteq G_j$$

Then, we observe that, $H_{j,m}$ are increasing with m. Now,

$$K \subseteq \bigcup_{j=1}^{N} G_j = \bigcup_{m} (\bigcup_{j=1}^{N} H_{j,m})$$

This is an open cover of the compact set K; therefore, it can be refined to a finite sub-cover. So,

$$K \subseteq \bigcup_{j=1}^{N} G_j = \bigcup_{m=1}^{M} (\bigcup_{j=1}^{N} H_{j,m})$$

By monotonicity in m, as $H_{j,m}$ are nested increasing sequence of sets, there exists some $M \in \mathbb{N}$ such that,

$$K \subseteq \bigcup_{j=1}^{N} H_{j,M}$$

Henceforth, for all j, we denote $H_{j,M}$ as H_j for the M determined above. Now, let

$$F_{j,m} = \{x \in H_j : \text{dist}(x, H_j^c) \geq \frac{1}{m}\} \quad \text{and} \quad E_{j,m} = \{x \in H_j : \text{dist}(x, H_j^c) > \frac{1}{m}\}$$

Then, each set $F_{j,m}$ are closed and each set $E_{j,m}$ are open. Further,

$$\mathrm{Cl}(E_{j,m}) = F_{j,m} \quad \text{and} \quad E_{j,m} \subseteq F_{j,m}$$

Then,

$$K \subseteq \bigcup_{j=1}^{N} H_j \subseteq \bigcup_{m} (\bigcup_{j=1}^{N} E_{j,m})$$

This is an open cover of the compact set K; therefore, it can be refined to a finite sub-cover. Thus,

$$K \subseteq \bigcup_{m=1}^{L} (\bigcup_{j=1}^{N} E_{j,m})$$

By monotonicity in m, as $E_{j,m}$ are nested increasing sequence of sets, there exists some $L \in \mathbb{N}$, such that,

$$K \subseteq \bigcup_{j=1}^{N} E_{j,L}$$

But, as $E_{j,m} \subseteq F_{j,m}$, therefore,

$$K \subseteq \bigcup_{j=1}^{N} F_{j,L}$$

Henceforth, for all j, we denote $F_{j,L}$ as F_j for the L determined above.
By construction, $F_j \subseteq H_j$; and $\mathrm{Cl}(H_j) \subseteq G_j$; and $\{F_j\}_{j=1}^{N}$ covers K.
So we have found the desired sets $H_j,\ F_j$. We can now construct functions φ_j.
Consider, for all j, the closed sets $F_j,\ H_j^c$. Since $F_j \subseteq H_j$, we get $F_j \cap H_j^c = \varnothing$.
So, by Urysohn's lemma (Lemma 8), there exists a continuous function $\psi_j : X \to [0,1]$ such that $\psi_j|_{F_j} \equiv 1$, and $\psi_j|_{H_j^c} \equiv 0$.
Since $\psi_j = 0$ on H_j^c, we have,

$$\operatorname{supp} \psi_j = \overline{\{x \in X : \psi_j(x) \neq 0\}} \subseteq \overline{H_j} \subseteq G_j$$

Let

$$\Psi = \sum_{j=1}^{N} \psi_j \quad \text{and} \quad \Omega = \{x \in X : \Psi(x) > \frac{1}{2}\}$$

Clearly, Ω is open and so Ω^c is closed. Further, on K for at least one j we have $\psi_j \equiv 1$, so $K \subseteq \Omega$.
Then, again by Urysohn's lemma (Lemma 8), we have a continuous function $\eta : X \to [0,1]$, such that, $\eta|_{\Omega^c} \equiv 0$ and $\eta|_K \equiv 1$. Let

$$\varphi_j = \begin{cases} \frac{\psi_j\ \eta}{\Psi} & \text{if } \Psi \neq 0 \\ 0 & \text{if } \Psi = 0 \end{cases}$$

Then for $\{\varphi_j\}_{j=1}^{N}$, we have that for all j,

- $\varphi_j : X \to [0,1]$ since it is defined on all $x \in X$ and $\Psi \geq \psi_j \geq \psi_j\ \eta \geq 0$ on X;
- φ_j is continuous, because, on the open set $\{\Psi > 0\}$, we have,

  $$\varphi_j = \frac{\psi_j \eta}{\Psi}$$

  which is continuous.
  On the set $\{\Psi = 0\}$, we have $\eta = 0$, because $\{\Psi = 0\} \subseteq \Omega^c$ and $\eta|_{\Omega^c} \equiv 0$.
  Hence $\psi_j \eta = 0$ on $\{\Psi = 0\}$, and therefore φ_j extends continuously by 0 across $\{\Psi = 0\}$;
- Since $\varphi_j = 0$ wherever $\psi_j = 0$, we have,

  $$\operatorname{supp} \varphi_j \subseteq \operatorname{supp} \psi_j \subseteq \overline{H_j} \subseteq G_j$$

  Thus,

  $$\operatorname{supp} \varphi_j \subseteq G_j$$
- Finally, on K, we have $\eta \equiv 1$ and $\Psi > 0$. Therefore,

  $$\begin{aligned}\sum_{j=1}^{N} \varphi_j &= \sum_{j=1}^{N} \frac{\psi_j \eta}{\Psi} \\ &= \frac{\eta}{\Psi} \sum_{j=1}^{N} \psi_j \\ &= \frac{1}{\Psi} \Psi \\ &= 1\end{aligned}$$

  Thus, the functions $\{\varphi_j\}_{j=1}^{N}$ are the desired functions.

□

Theorem 51 (Riesz-Markov-Kakutani Representation Theorem). *Let X be a compact metric space. If I is any non-negative linear functional, i.e. $I : C(X) \to \mathbb{R}$, then there exists a unique non-negative finite regular Borel measure μ on X such that;*

$$I(f) = \int_X f\, d\mu, \qquad \forall f \in C(X)$$

Proof. For any open set $G \subseteq X$, define the set function

$$\mu(G) = \sup \big\{ I(\varphi) :\ \varphi \in C(X),\ 0 \leq \varphi \leq 1,\ \operatorname{supp}(\varphi) \subseteq G \big\}$$

We aim to show that μ is a non-negative finite Borel measure on X.

- Consider an open set G. As I is a non-negative linear functional, for every $f \in C(X)$ with $f \geq 0$ and $\operatorname{supp}(f) \subseteq G$ we have $I(f) \geq 0$. Thus,

  $$\mu(G) \geq 0, \qquad \forall \text{ open } G \subseteq X$$

  Thus, μ is non-negative.

- Also, for every admissible function φ in the definition of $\mu(X)$, we have

  $$0 \leq \varphi \leq \mathbb{1}_X$$

  Hence,

  $$\mathbb{1}_X - \varphi \geq 0$$

  Since I is non-negative,

  $$I(\mathbb{1}_X) - I(\varphi) = I(\mathbb{1}_X - \varphi) \geq 0$$

  Therefore,

  $$I(\varphi) \leq I(\mathbb{1}_X)$$

  Taking the supremum over all admissible φ gives,

  $$\mu(X) \leq I(\mathbb{1}_X) < \infty$$

  On the other hand, the constant function $\mathbb{1}_X$ belongs to $C(X)$, satisfies

  $$0 \leq \mathbb{1}_X \leq 1$$

  and,

  $$\operatorname{supp}(\mathbb{1}_X) = X$$

  Hence, $\mathbb{1}_X$ is admissible in the definition of $\mu(X)$. Therefore,

  $$\mu(X) \geq I(\mathbb{1}_X)$$

  Combining the two inequalities, we obtain,

  $$\mu(X) = I(\mathbb{1}_X) < \infty$$

  Thus, μ is finite.

- Consider the countable collection of open sets $\{G_i\}$ and Let $G = \bigcup\limits_{i=1}^{\infty} G_i$. Then G is open. Let

  $$f \in C(X), \qquad 0 \leq f \leq 1, \qquad \operatorname{supp}(f) \subseteq G$$

Let $K = \operatorname{supp}(f)$. By definition,

$$\mu(G) = \sup\{I(f) : f \in C(X),\ 0 \leq f \leq 1,\ \operatorname{supp}(f) \subseteq G\}$$

Now K is a closed subset of the compact space X. Thus, K is compact, and $K \subseteq G = \bigcup_{i=1}^{\infty} G_i$. Hence, by compactness, there exists n such that,

$$K \subseteq \bigcup_{i=1}^{n} G_i$$

By a partition of unity (Theorem 50), there exist $h_1, h_2, \ldots, h_n$ with $h_i \in C(X)$ such that,

$$0 \leq h_i \leq 1, \qquad \operatorname{supp}(h_i) \subseteq G_i \quad \text{for } i = 1, 2, \ldots, n$$

and

$$\sum_{i=1}^{n} h_i(x) = 1 \quad \text{for all } x \in K$$

Define $f_i = f\, h_i$, for $i = 1, 2, \ldots, n$. Then,

$$f = \sum_{i=1}^{n} f_i$$

because $\sum_{i=1}^{n} h_i = 1$ on $K = \operatorname{supp}(f)$ and both sides vanish outside K.
Then,

$$\operatorname{supp}(f_i) \subseteq \operatorname{supp}(h_i) \subseteq G_i \quad \text{with} \quad 0 \leq f_i \leq 1$$

By linearity,

$$I(f) = I\Big(\sum_{i=1}^{n} f_i\Big) = \sum_{i=1}^{n} I(f_i)$$

and since each f_i is admissible for $\mu(G_i)$, we have $I(f_i) \leq \mu(G_i)$, for $i = 1, 2, \ldots, n$. Then,

$$I(f) \leq \sum_{i=1}^{n} \mu(G_i)$$

Also, as all $\mu(G_i)$ are non-negative, for every $f \in C(X)$ with $\operatorname{supp}(f) \subseteq G$ we have,

$$I(f) \leq \sum_{i=1}^{\infty} \mu(G_i)$$

Since this holds for every $f \in C(X)$ satisfying, $0 \leq f \leq 1$ and $\operatorname{supp}(f) \subseteq G$, taking the supremum over all such choices yields,

$$\mu(G) = \sup I(f) \leq \sum_{i=1}^{\infty} \mu(G_i)$$

Thus, μ is countably sub-additive on open sets.

Therefore, μ is a finite, non-negative and countably sub-additive set function on the open subsets of X.
For an arbitrary set $E \subseteq X$, define,

$$\mu^*(E) = \inf\{\mu(G) : G \supseteq E,\ G \text{ open}\}$$

- Consider the empty set $\varnothing$. Since $\varnothing$ is open and $\mu(\varnothing) = 0$, we have,

  $$\mu^*(\varnothing) \leq \mu(\varnothing) = 0$$

  Since $\mu^* \geq 0$, it follows that,

  $$\mu^*(\varnothing) = 0$$

- Let $A, B \subseteq X$ with $A \subseteq B$. Then, every open set containing B also contains A. Hence,

  $$\{G :\ G \supseteq B,\ G \text{ open}\} \subseteq \{G :\ G \supseteq A,\ G \text{ open}\}$$

  Taking infima gives,

  $$\mu^*(A) \leq \mu^*(B)$$

- Let $E \subseteq X$ is open. Since E is open and $E \subseteq E$, the definition of μ^* gives

  $$\mu^*(E) = \inf\{\mu(G) : G \supseteq E,\ G \text{ open}\} \leq \mu(E)$$

  On the other hand, if G is any open set containing E, then every admissible function for $\mu(E)$ is also admissible for $\mu(G)$. Hence,

  $$\mu(E) \leq \mu(G)$$

  Since this holds for every open set $G \supseteq E$, we obtain,

  $$\mu(E) \leq \inf\{\mu(G) : G \supseteq E,\ G \text{ open}\} = \mu^*(E)$$

  Therefore, for every open set $E \subseteq X$,

  $$\mu^*(E) = \mu(E)$$

- Let $\{E_i\}$ be a countable collection of sets in X.
  Fix $\varepsilon > 0$. Then, $\forall i$, $\exists$ an open set G_i, s.t. $E_i \subseteq G_i$ and by property of infimum,

  $$\mu(G_i) \leq \mu^*(E_i) + \frac{\varepsilon}{2^i}$$

Also, as $E_i \subseteq G_i,\ \forall i$, then, $\bigcup\limits_{i=1}^{\infty} E_i \subseteq \bigcup\limits_{i=1}^{\infty} G_i$ and as a union of open sets is open, $\bigcup\limits_{i=1}^{\infty} G_i$ is open. Then,

$$\begin{aligned} \mu^*\Big(\bigcup_{i=1}^{\infty} E_i\Big) &\leq \mu^*\Big(\bigcup_{i=1}^{\infty} G_i\Big) \\ &= \mu\Big(\bigcup_{i=1}^{\infty} G_i\Big) \\ &\leq \sum_{i=1}^{\infty} \mu(G_i) \\ &\leq \sum_{i=1}^{\infty} \mu^*(E_i) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary,

$$\mu^*\Big(\bigcup_{i=1}^{\infty} E_i\Big) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$$

Therefore, μ^* is an outer measure and is finite as μ is finite.

We now show that every open set is μ^*-measurable.
Let $V \subseteq X$ be open. We need to show that for every $A \subseteq X$,

$$\mu^*(A) = \mu^*(A \cap V) + \mu^*(A \cap V^c)$$

By subadditivity of μ^*, we already have,

$$\mu^*(A) \leq \mu^*(A \cap V) + \mu^*(A \cap V^c)$$

Thus, it remains to prove the reverse inequality.
Let U be an open set such that $A \subseteq U$. Then

$$A \cap V \subseteq U \cap V \qquad \text{and} \qquad A \cap V^c \subseteq U \cap V^c$$

Hence,

$$\mu^*(A \cap V) + \mu^*(A \cap V^c) \leq \mu(U \cap V) + \mu^*(U \cap V^c)$$

We claim that for every pair of open sets $U, V \subseteq X$,

$$\mu(U) \geq \mu(U \cap V) + \mu^*(U \cap V^c)$$

Fix $\varepsilon > 0$. By the definition of $\mu(U \cap V)$, there exists $f_1 \in C(X)$ such that, $0 \le f_1 \le 1$, $\operatorname{supp}(f_1) \subseteq U \cap V$ and,

$$I(f_1) > \mu(U \cap V) - \varepsilon$$

Let $K = \operatorname{supp}(f_1)$. Then, K is compact and $K \subseteq V$.
Hence, $K \cap V^c = \varnothing$, and $U \setminus K$ is open with $U \cap V^c \subseteq U \setminus K$.
Therefore,

$$\mu^*(U \cap V^c) \le \mu(U \setminus K)$$

By the definition of $\mu(U \setminus K)$, there exists $f_2 \in C(X)$ such that, $0 \le f_2 \le 1$, $\operatorname{supp}(f_2) \subseteq U \setminus K$, and

$$I(f_2) > \mu(U \setminus K) - \varepsilon$$

Since $\operatorname{supp}(f_1) = K$ and $\operatorname{supp}(f_2) \subseteq U \setminus K$, the supports of f_1 and f_2 are disjoint. Hence,

$$0 \le f_1 + f_2 \le 1 \qquad \text{and} \qquad \operatorname{supp}(f_1 + f_2) \subseteq U$$

Thus, $f_1 + f_2$ is admissible for $\mu(U)$, and so

$$\begin{aligned} \mu(U) &\ge I(f_1 + f_2) \\ &= I(f_1) + I(f_2) \\ &> \mu(U \cap V) - \varepsilon + \mu(U \setminus K) - \varepsilon \\ &\ge \mu(U \cap V) + \mu^*(U \cap V^c) - 2\varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary,

$$\mu(U) \ge \mu(U \cap V) + \mu^*(U \cap V^c)$$

Therefore,

$$\mu^*(A \cap V) + \mu^*(A \cap V^c) \le \mu(U)$$

for every open set U with $A \subseteq U$. Taking the infimum over all such open U, we obtain,

$$\mu^*(A \cap V) + \mu^*(A \cap V^c) \le \mu^*(A)$$

Combining this with subadditivity gives

$$\mu^*(A) = \mu^*(A \cap V) + \mu^*(A \cap V^c)$$

Hence every open set V is μ^*-measurable.
Since the collection of μ^*-measurable sets is a σ-algebra by Carathéodory's Theorem (Theorem 12) and contains all open sets, it contains $\mathcal{B}(X)$. Therefore, $\mu := \mu^*|_{\mathcal{B}(X)}$ is a finite Borel measure on X.
Moreover, by the construction of μ^* from open covers, μ is outer regular. Since X is compact metric, finite Borel measures are regular. Hence μ is a finite regular Borel measure.

Let $f \in C(X)$ and as X is compact, we have f is bounded. Fix $\varepsilon > 0$. Since f is bounded, there exists $M \in \mathbb{Z}^+$ such that,

$$f(x) \leq M\varepsilon \qquad \forall x \in X$$

WLOG, assume $f \geq 0$. Define, for $k = 0, 1, 2, \ldots, M$,

$$f_k = \min\{(k+1)\varepsilon,\ \max\{f, k\varepsilon\}\} - k\varepsilon$$

For each $x \in X$, exactly one of the following occurs:

- If $f(x) \leq k\varepsilon$, then $f_k(x) = 0$.
- If $k\varepsilon < f(x) < (k+1)\varepsilon$, then, $f_k(x) = f(x) - k\varepsilon$.
- If $f(x) \geq (k+1)\varepsilon$, then, $f_k(x) = \varepsilon$.

Hence, $f_k \in C(X)$ and $0 \leq f_k \leq \varepsilon$. We define

$$g_k = \frac{f_k}{\varepsilon}$$

then $g_k \in C(X)$ and $0 \leq g_k \leq 1$. Let

$$G_k = \{x \in X :\ f(x) > k\varepsilon\}$$

and $G_{-1} = X$, $G_M = \varnothing$, and $G_i = \varnothing$ for all $i > M$. Then each G_k is an open set.
Also, $\operatorname{supp}(g_k) \subseteq G_{k-1}$, hence by definition, $\mu(G_{k-1}) \geq I(g_k)$.
Again, $g_k = 1$ on G_{k+1}, and thus $\operatorname{supp}(g_k) \supset G_{k+1}$. Now,

$$\mu(G_{k+1}) = \sup\left\{I(\varphi) :\ \varphi \in C(X),\ 0 \leq \varphi \leq 1,\ \operatorname{supp}(\varphi) \subseteq G_{k+1}\right\}$$

Then any such φ satisfies $\varphi \leq 1$ on G_{k+1}, while $g_k = 1$ on G_{k+1}.
Also, since $\operatorname{supp}(\varphi) \subseteq G_{k+1}$, we have $\varphi = 0$ outside G_{k+1}. Hence,

$$\varphi \leq g_k \qquad \text{on all of } X$$

Since I is non-negative,

$$I(g_k) - I(\varphi) = I(g_k - \varphi) \geq 0$$

Therefore,

$$I(\varphi) \leq I(g_k)$$

Taking supremum, we get

$$\mu(G_{k+1}) \leq I(g_k)$$

Therefore,

$$\mu(G_{k+1}) \leq I(g_k) \leq \mu(G_{k-1})$$

Moreover, for every $x \in X$, either $f(x) = M\varepsilon$, or there exists an integer m with $0 \leq m < M$ such that

$$m\varepsilon \leq f(x) < (m+1)\varepsilon$$

In the second case, $f_k(x) = \varepsilon$ for $k < m$, $f_m(x) = f(x) - m\varepsilon$ and $f_k(x) = 0$, for $k > m$.
Hence,

$$\sum_{k=0}^{M} f_k(x) = m\varepsilon + \big(f(x) - m\varepsilon\big) = f(x)$$

If $f(x) = M\varepsilon$, then $f_k(x) = \varepsilon$ for $k = 0, \ldots, M-1$ and $f_M(x) = 0$, so again,

$$\sum_{k=0}^{M} f_k(x) = f(x)$$

Therefore,

$$f = \sum_{k=0}^{M} f_k$$

Then,

$$I(f) = \sum_{k=0}^{M} I(f_k) = \varepsilon \sum_{k=0}^{M} I(g_k)$$

Thus,

$$\varepsilon \sum_{k=0}^{M} \mu(G_{k+1}) \leq I(f) \leq \varepsilon \sum_{k=0}^{M} \mu(G_{k-1})$$

Also,

$$\int_X f \, d\mu = \int_X \Big(\sum_{k=0}^{M} f_k \Big) \, d\mu = \varepsilon \sum_{k=0}^{M} \Big(\int_X g_k \, d\mu \Big)$$

Since

$$G_{k+1} \subseteq \operatorname{supp}(g_k) \subseteq G_{k-1}$$

and $g_k = 1$ on G_{k+1} and $0 \leq g_k \leq 1$ with $\operatorname{supp}(g_k) \subseteq G_{k-1}$, we have,

$$\mathbb{1}_{G_{k+1}} \leq g_k \leq \mathbb{1}_{G_{k-1}}$$

Thus, integrating over X, we have,

$$\mu(G_{k+1}) \leq \int_X g_k \, d\mu \leq \mu(G_{k-1})$$

Therefore,

$$\varepsilon \sum_{k=0}^{M} \mu(G_{k+1}) \leq \int_X f \, d\mu \leq \varepsilon \sum_{k=0}^{M} \mu(G_{k-1})$$

Then,

$$\begin{aligned}\Big|I(f) - \int_X f\,d\mu\Big| &\le \varepsilon\Big[\sum_{k=0}^{M}\mu(G_{k-1}) - \sum_{k=0}^{M}\mu(G_{k+1})\Big]\\ &= \varepsilon\big[\mu(G_{-1}) + \mu(G_0) - \mu(G_M) - \mu(G_{M+1})\big]\\ &\le \varepsilon\big[\mu(X) + \mu(X) - 0 - 0\big]\end{aligned}$$

Thus,

$$\left|I(f) - \int_X f\,d\mu\right| \le 2\varepsilon\,\mu(X)$$

As, $\mu(X) < \infty$ and as $\varepsilon \longrightarrow 0$, we get,

$$\left|I(f) - \int_X f\,d\mu\right| \longrightarrow 0$$

Therefore,

$$I(f) = \int_X f\,d\mu$$

This proves the existence of a non-negative finite Borel measure μ on X such that;

$$I(f) = \int_X f\,d\mu, \qquad \forall f \in C(X)$$

Now, we prove the uniqueness. Assume that there exists a non-negative finite regular Borel measure μ on X such that for all $f \in C(X)$,

$$I(f) = \int_X f\,d\mu$$

Let $G \subseteq X$ be an open set. Consider the class of functions on G,

$$\mathcal{F}_G = \{\varphi \in C(X)\ :\ 0 \le \varphi \le 1,\ \operatorname{supp}(\varphi) \subseteq G\}$$

If $\varphi \in \mathcal{F}_G$, then $\varphi \le \mathbb{1}_G$. Hence

$$I(\varphi) = \int_X \varphi\,d\mu \le \int_X \mathbb{1}_G\,d\mu = \mu(G)$$

Therefore,

$$\sup_{\varphi\in\mathcal{F}_G} I(\varphi) \le \mu(G)$$

Now, Let $d(x) = \operatorname{dist}(x, G^c)$. As G is open, G^c is closed and $d(x) > 0 \iff x \in G$. We define,

$$\varphi_n(x) = \min\{(n\,d(x) - 1)_+,\ 1\}$$

As distance and minimum are continuous functions, we have,

$$\varphi_n \in C(X) \quad , \quad 0 \leq \varphi_n \leq 1 \quad , \quad \operatorname{supp}(\varphi_n) \subseteq \left\{ x : d(x) \geq \frac{1}{n} \right\} \subseteq G$$

Hence $\varphi_n \in \mathcal{F}_G$.
Also,

$$\varphi_n(x) \uparrow \mathbb{1}_G(x) \qquad \text{for every } x \in X$$

By the Monotone Convergence Theorem (Theorem 28),

$$\lim_{n\to\infty} I(\varphi_n) = \lim_{n\to\infty} \int_X \varphi_n \, d\mu = \int_X \mathbb{1}_G \, d\mu = \mu(G)$$

Therefore,

$$\mu(G) = \sup_{\varphi\in\mathcal{F}_G} I(\varphi)$$

Now, assume that μ is not unique. Then there exist μ_1 and μ_2 that satisfy,

$$I(f) = \int_X f \, d\mu_1 = \int_X f \, d\mu_2, \qquad \forall f \in C(X)$$

As the measure of any open set is determined only by the functional I, for every open set G, we can write,

$$\mu_1(G) = \sup\{I(\varphi) : \varphi \in \mathcal{F}_G\} \quad \text{and} \quad \mu_2(G) = \sup\{I(\varphi) : \varphi \in \mathcal{F}_G\}$$

where $\mathcal{F}_G$ denotes the family of admissible test functions associated to G. Therefore,

$$\mu_1(G) = \mu_2(G) \quad \text{for all open sets } G$$

Finally, by regularity of Borel measurable sets, for any Borel set E, we have,

$$\mu_1(E) = \inf\{\mu_1(G) : E \subseteq G,\ G \text{ open}\}, \quad \mu_2(E) = \inf\{\mu_2(G) : E \subseteq G,\ G \text{ open}\}$$

Since $\mu_1(G) = \mu_2(G)$ for all open G, their infima agree, and hence,

$$\mu_1(E) = \mu_2(E) \quad \text{for all Borel sets } E$$

Therefore, μ is the unique non-negative finite regular Borel measure. □

Lemma 10 (Finite Vitali Covering Lemma)**.** *Let X be a metric space. Let $\{B_{r_i}(x_i)\}$ be a finite family of open balls in X, where $x_i \in X$, $r_i > 0$ and $i \in I$ for some finite index set I. Then there exists a finite sub-collection of pairwise disjoint balls $\{B_{r_{i_k}}(x_{i_k})\}$, where $x_{i_k} \in X$, $r_{i_k} > 0$ and $i_k \in J \subseteq I$ such that;*

$$\bigcup_{i\in I} B_{r_i}(x_i) \subseteq \bigcup_{i_k\in J} B_{3r_{i_k}}(x_{i_k})$$

Proof. Without loss of generality, Let the finite index set I be $\{1, 2, \cdots, n\}$ and we can reorder the balls in the family $\{B_{r_i}(x_i)\}$ in a way such that;

$$r_1 \geq r_2 \geq \cdots \geq r_n$$

We proceed with greedy algorithm.
We pick the ball with largest radius, i.e., $B_{r_1}(x_1)$. And we rename the index as $i_1 = 1$. Thus, at this stage $J = \{i_1\} = \{1\} \subseteq I$.
Next we pick the next ball with the largest radius such that it is disjoint from $B_{r_1}(x_1)$, say, $B_{r_m}(x_m)$ and we rename the index as, $i_2 = m$. Then, at this stage $J = \{i_1, i_2\} = \{1, m\} \subseteq I$.
Again we pick the next ball with the largest radius such that it is disjoint from $B_{r_1}(x_1)$ and $B_{r_m}(x_m)$, say, $B_{r_t}(x_t)$ and we rename the index as, $i_3 = t$. Then, at this stage $J = \{i_1, i_2, i_3\} = \{1, m, t\} \subseteq I$.
And we continue this process. This process terminates as the initial family had a finite number of balls. Thus, we obtain a finite sub-collection of pairwise disjoint balls $\{B_{r_{i_k}}(x_{i_k})\}$, where $x_{i_k} \in X$, $r_{i_k} > 0$ and $i_k \in J \subseteq I$.
Consider any ball $B_{r_j}(x_j)$ in the initial family $\{B_{r_i}(x_i)\}$. Then there are two possibilities that can occur:

- **Case I:** The ball $B_{r_j}(x_j)$ was chosen into the subfamily. Then $j = i_k$ for some k. And clearly, $B_{r_j}(x_j) \subseteq B_{3r_j}(x_j) \subseteq B_{3r_{i_k}}(x_{i_k})$
- **Case II:** The ball $B_{r_j}(x_j)$ was not chosen into the subfamily. Assume that this ball was eliminated after the k-th step where the ball $B_{r_t}(x_t)$ was chosen into the subfamily. So, $t = i_k$.
  Then $r_{i_k} = r_t \geq r_j$, and, $B_{r_j}(x_j) \cap B_{r_{i_s}}(x_{i_s}) = \varnothing$, for all, $s < k$.
  And, also, $B_{r_j}(x_j) \cap B_{r_{i_k}}(x_{i_k}) \neq \varnothing$.
  If, $y \in B_{r_j}(x_j) \cap B_{r_{i_k}}(x_{i_k})$, then,

  $$d(x_j, x_{i_k}) \leq d(x_j, y) + d(y, x_{i_k}) < r_j + r_{i_k} \leq 2r_{i_k}$$

  Let $x \in B_{r_j}(x_j)$. Then,

  $$\begin{aligned} d(x, x_{i_k}) &\leq d(x, x_j) + d(x_j, x_{i_k}) \\ &\leq r_j + 2 \cdot r_{i_k} \\ &\leq 3 \cdot r_{i_k} \end{aligned}$$

  Thus, $B_{r_j}(x_j) \subseteq B_{3r_{i_k}}(x_{i_k})$

Therefore,

$$\bigcup_{i \in I} B_{r_i}(x_i) \subseteq \bigcup_{i_k \in J} B_{3r_{i_k}}(x_{i_k})$$

□

Lemma 11 (Countable Maximal Disjoint Subfamily Lemma)**.** *Let (X, d) be a separable metric space. Let $\mathcal{F}$ be any family of non-empty open subsets of X. Then there exists an at-most countable subfamily $\mathcal{G} \subseteq \mathcal{F}$ such that:*

(i) the elements of $\mathcal{G}$ are pairwise disjoint;

(ii) $\mathcal{G}$ is maximal with respect to this property, i.e., if $U \in \mathcal{F}$, then either $U \in \mathcal{G}$, or U intersects some element of $\mathcal{G}$.

Proof. Since X is separable, X has a countable basis. Let $\{V_1, V_2, V_3, \dots\}$ be a countable basis for the topology of X.
We now construct $\mathcal{G}$ inductively.

- At the first step, consider V_1. If there exists $U \in \mathcal{F}$ such that, $V_1 \subseteq U$, choose one such set and call it U_1. If no such set exists, choose nothing.
- Next consider V_2. If there exists $U \in \mathcal{F}$ such that, $V_2 \subseteq U$ and U is disjoint from all previously chosen sets, choose one such set. If no such set exists, choose nothing.
- Continuing this process, at the n-th step, consider V_n. If there exists $U \in \mathcal{F}$ such that, $V_n \subseteq U$ and U is disjoint from all previously chosen sets, choose one such set. If no such set exists, choose nothing.

Let $\mathcal{G}$ be the collection of all sets chosen in this way. Since at most one set is chosen at each step, $\mathcal{G}$ is at most countable. Also, by construction, the elements of $\mathcal{G}$ are pairwise disjoint.
It remains to show that $\mathcal{G}$ is maximal.
Assume not. Then there exists $U \in \mathcal{F}$ such that U is disjoint from every element of $\mathcal{G}$. Since U is a non-empty open set, and $\{V_n\}_{n=1}^{\infty}$ is a basis, there exists some V_m such that, $V_m \subseteq U$.
At the m-th step of the construction, the set U was available, because it contains V_m and is disjoint from all previously chosen sets. Therefore, at that step, we would have chosen some set $U_m \in \mathcal{F}$ such that,

$$V_m \subseteq U_m$$

and U_m is disjoint from all previously chosen sets.
But then,

$$V_m \subseteq U \cap U_m$$

Hence,

$$U \cap U_m \neq \varnothing$$

This contradicts the assumption that U is disjoint from every element of $\mathcal{G}$.
Therefore, no such U exists.
Hence, $\mathcal{G}$ is maximal. □

Theorem 52 (Infinite Vitali Covering Theorem)**.** *Let X be a separable metric space. Let $\{B_\alpha\}$ be any family of balls in X, where $\alpha \in I$, for some index set I, such that, the supremum of radii, $\sup r_\alpha < \infty$. Then there exists an at-most countable subfamily of pairwise disjoint balls $\{B_\delta\}$, where $\delta \in J \subseteq I$, and J is a countable index set, such that,*

$$\bigcup_{\alpha\in I} B_\alpha \subseteq \bigcup_{\delta\in J} 5B_\delta$$

where $5B_\delta$ denotes the ball with the same center as B_δ and radius $5r_\delta$.

Proof. Let X be a separable metric space. Let $\{B_\alpha\}$ be any family of balls in X, where $\alpha \in I$, for some index set I, such that, the supremum of radii,

$$R = \sup_{\alpha\in I} r_\alpha < \infty$$

Since $r_\alpha > 0$ for every $\alpha \in I$, if the family is non-empty, then $R > 0$.
If the family is empty, then there is nothing to prove.
So, assume that the family is non-empty. For each $n \in \mathbb{N}$, define,

$$\mathcal{F}_n = \left\{ B_\alpha = B_{r_\alpha}(x_\alpha) : \frac{R}{2^n} < r_\alpha \leq \frac{R}{2^{n-1}} \right\}$$

Since $0 < r_\alpha \leq R$, there exists $n \in \mathbb{N}$ such that,

$$\frac{R}{2^n} < r_\alpha \leq \frac{R}{2^{n-1}}$$

Therefore, every ball B_α in the original family belongs to exactly one of the families $\mathcal{F}_n$.
We now choose the desired subfamily inductively.
From $\mathcal{F}_1$, using Lemma 11, choose an at-most countable maximal pairwise disjoint subfamily. Denote this chosen subfamily by $\mathcal{G}_1$.
Now suppose that $\mathcal{G}_1, \ldots, \mathcal{G}_{n-1}$ have already been chosen. Let

$$\mathcal{H}_n = \{B \in \mathcal{F}_n : B \text{ is disjoint from every ball in } \mathcal{G}_1 \cup \cdots \cup \mathcal{G}_{n-1}\}$$

Since every element of $\mathcal{F}_n$ is an open ball, every element of $\mathcal{H}_n$ is a non-empty open subset of X.
If $\mathcal{H}_n = \varnothing$, then we set $\mathcal{G}_n = \varnothing$.
otherwise, using Lemma 11, choose an at-most countable maximal pairwise disjoint

subfamily $\mathcal{G}_n$ of $\mathcal{H}_n$.
Continuing this process, we obtain a family,

$$\mathcal{G} = \bigcup_{n=1}^{\infty} \mathcal{G}_n$$

By construction, the balls in $\mathcal{G}$ are pairwise disjoint.
Since each $\mathcal{G}_n$ is at-most countable by Lemma 11, and

$$\mathcal{G} = \bigcup_{n=1}^{\infty} \mathcal{G}_n$$

it follows that $\mathcal{G}$ is at-most countable.
Therefore, we may write,

$$\mathcal{G} = \{B_\delta : \delta \in J\}$$

where $J \subseteq I$ is an at-most countable index set.
It remains to show that

$$\bigcup_{\alpha \in I} B_\alpha \subseteq \bigcup_{\delta \in J} 5B_\delta$$

Consider any ball $B_\alpha = B_{r_\alpha}(x_\alpha)$ from the original family. Then $B_\alpha \in \mathcal{F}_n$ for some $n \in \mathbb{N}$.
There are two possibilities.

- **Case I:** The ball B_α was chosen into $\mathcal{G}$. Then $B_\alpha = B_\delta$ for some $\delta \in J$. Therefore,

  $$B_\alpha \subseteq 5B_\delta$$

- **Case II:** The ball B_α was not chosen into $\mathcal{G}$.
  Since B_α was not chosen into $\mathcal{G}$, there are two possibilities. Either B_α is not disjoint from some ball in $\mathcal{G}_1 \cup \cdots \cup \mathcal{G}_{n-1}$, or $B_\alpha \in \mathcal{H}_n$.
  In the latter case, since $\mathcal{G}_n$ is maximal in $\mathcal{H}_n$ and $B_\alpha \notin \mathcal{G}_n$, the ball B_α must intersect some ball in $\mathcal{G}_n$.
  Hence, in either case, there exists a chosen ball $B_\tau = B_{r_\tau}(x_\tau) \in \mathcal{G}_1 \cup \cdots \cup \mathcal{G}_n$ such that, $B_\alpha \cap B_\tau \neq \varnothing$.
  Since $B_\tau \in \mathcal{G}_m$ for some $m \leq n$, we have, $B_\tau \in \mathcal{F}_m$. Hence,

  $$r_\tau > \frac{R}{2^m}$$

  Also, since $B_\alpha \in \mathcal{F}_n$,

  $$r_\alpha \leq \frac{R}{2^{n-1}}$$

Since $m \leq n$, we have,

$$\frac{R}{2^{n-1}} \leq \frac{2R}{2^m}$$

Therefore,

$$r_\alpha \leq \frac{R}{2^{n-1}} \leq \frac{2R}{2^m} < 2r_\tau$$

So,

$$r_\alpha < 2r_\tau$$

Now choose, $y \in B_\alpha \cap B_\tau$. Then,

$$d(x_\alpha, x_\tau) \leq d(x_\alpha, y) + d(y, x_\tau) < r_\alpha + r_\tau$$

Let $x \in B_\alpha$. Then,

$$\begin{aligned} d(x, x_\tau) &\leq d(x, x_\alpha) + d(x_\alpha, x_\tau) \\ &< r_\alpha + (r_\alpha + r_\tau) \\ &= 2r_\alpha + r_\tau \\ &< 2(2r_\tau) + r_\tau \\ &= 5r_\tau \end{aligned}$$

Hence,

$$x \in 5B_\tau$$

Since $x \in B_\alpha$ was arbitrary, we get $B_\alpha \subseteq 5B_\tau$.

Therefore, every ball in the original family is contained in $5B_\tau$ for some chosen ball $B_\tau \in \mathcal{G}$. Hence,

$$\bigcup_{\alpha \in I} B_\alpha \subseteq \bigcup_{B_\tau \in \mathcal{G}} 5B_\tau = \bigcup_{\delta \in J} 5B_\delta$$

$\square$

Theorem 53 (Finite Besicovitch Covering Theorem)**.** *Let $X = \mathbb{R}^n$ be a metric space. Let $\{B_{r_i}(x_i)\}$ be a finite family of open balls in X, where $x_i \in X$, $r_i > 0$ and $i \in I$ for some finite index set I. Let the set containing all the centers be E, i.e., $E = \{x_i : i \in I\}$.*

(i) Then there exists a finite sub-collection of balls $\{B_{r_{i_k}}(x_{i_k})\}$, where $x_{i_k} \in X$, $r_{i_k} > 0$ and $i_k \in J \subseteq I$ such that;

$$E \subseteq \bigcup_{i_k \in J} B_{r_{i_k}}(x_{i_k})$$

Continued...

(ii) Every $x \in X = \mathbb{R}^n$ *is contained in at-most* $N(n)$ *balls in the subfamily* $\{B_{r_{i_k}}(x_{i_k})\}$.

Proof. Without loss of generality, Let the finite index set I be $\{1, 2, \cdots, N\}$ and we can reorder the balls in the family $\{B_{r_i}(x_i)\}$ in a way such that;

$$r_1 \geq r_2 \geq \cdots \geq r_N$$

We proceed with greedy algorithm.
We pick the ball with largest radius, i.e., $B_{r_1}(x_1)$. And we rename the index as $i_1 = 1$. Thus, at this stage $J = \{i_1\} = \{1\} \subseteq I$.
Next we pick the next ball with the largest radius such that its center is not contained in $B_{r_1}(x_1)$, say, $B_{r_m}(x_m)$ and we rename the index as, $i_2 = m$. Then, at this stage $J = \{i_1, i_2\} = \{1, m\} \subseteq I$.
Again we pick the next ball with the largest radius such that its center is not contained in $B_{r_1}(x_1)$ and $B_{r_m}(x_m)$, say, $B_{r_t}(x_t)$ and we rename the index as, $i_3 = t$. Then, at this stage $J = \{i_1, i_2, i_3\} = \{1, m, t\} \subseteq I$.
Continuing inductively, at each stage we select a ball of largest radius among those whose centers are not contained in any previously selected ball. Since the original family is finite, the process terminates after finitely many steps.
Let $\{B_{r_{i_k}}(x_{i_k})\}_{i_k \in J}$ be the resulting subfamily. By construction, every center corresponding to a selected ball belongs to the union

$$\bigcup_{i_k \in J} B_{r_{i_k}}(x_{i_k})$$

Moreover, if $j \in I \setminus J$, then the ball $B_{r_j}(x_j)$ was not selected, and hence at the stage when the algorithm terminated its center x_j was already contained in one of the previously selected balls. Therefore,

$$x_j \in \bigcup_{i_k \in J} B_{r_{i_k}}(x_{i_k})$$

Consequently,

$$E = \{x_i : i \in I\} \subseteq \bigcup_{i_k \in J} B_{r_{i_k}}(x_{i_k})$$

Let $x \in \mathbb{R}^n$ be arbitrary and Let

$$P = \{\, i_p \in J : x \in B_{r_{i_p}}(x_{i_p}) \,\}$$

If $x = x_{i_p}$ for some $i_p \in P$, then P has only one element. Indeed, if another selected ball $B_{r_{i_q}}(x_{i_q})$ also contained x, then either $B_{r_{i_q}}(x_{i_q})$ was selected before $B_{r_{i_p}}(x_{i_p})$, in which case the center $x_{i_p} = x$ would already lie in a previously selected ball, or $B_{r_{i_q}}(x_{i_q})$ was selected after $B_{r_{i_p}}(x_{i_p})$, in which case the center x_{i_q} would already lie

in the previously selected ball $B_{r_{i_p}}(x_{i_p})$. Both contradict the selection rule.
Hence, we may assume that $x \neq x_{i_p}$ for every $i_p \in P$.
Consider two balls $B_{r_{i_m}}(x_{i_m})$ and $B_{r_{i_t}}(x_{i_t})$ from the selected subfamily such that,

$$x \in B_{r_{i_m}}(x_{i_m}) \cap B_{r_{i_t}}(x_{i_t})$$

Without loss of generality, assume that

$$r_{i_m} \geq r_{i_t}$$

and if the radii are equal, assume that $B_{r_{i_m}}(x_{i_m})$ was selected no later than $B_{r_{i_t}}(x_{i_t})$. Since $B_{r_{i_t}}(x_{i_t})$ was selected after $B_{r_{i_m}}(x_{i_m})$, its center does not belong to $B_{r_{i_m}}(x_{i_m})$. Therefore,

$$d(x_{i_m}, x_{i_t}) \geq r_{i_m}$$

Since x belongs to both balls, we also have,

$$d(x, x_{i_m}) < r_{i_m} \qquad \text{and} \qquad d(x, x_{i_t}) < r_{i_t} \leq r_{i_m}$$

Let

$$a = d(x, x_{i_m}), \qquad b = d(x, x_{i_t}), \qquad c = d(x_{i_m}, x_{i_t})$$

Then,

$$c \geq r_{i_m} > a \qquad \text{and} \qquad c \geq r_{i_m} \geq r_{i_t} > b$$

Hence c is strictly the largest side of the triangle with vertices x, x_{i_m}, x_{i_t}. Therefore the angle at x, which is the angle opposite the side of length c, is strictly the largest angle of the triangle. Since the largest angle of any triangle is at least 60°, and here it is strictly larger than the other two angles, we get,

$$\angle x > 60^\circ$$

Now consider the unit sphere centered at x, $S^{n-1}(x)$.
Let y_{i_m} and y_{i_t} be the intersection points of the rays $\overrightarrow{xx_{i_m}}$ and $\overrightarrow{xx_{i_t}}$ with $S^{n-1}(x)$, respectively. Then,

- $\angle(y_{i_m}\, x\, y_{i_t}) = \angle(x_{i_m}\, x\, x_{i_t}) > 60^\circ$
- $|x - y_{i_m}| = |x - y_{i_t}| = 1$

Applying the law of cosines in the triangle $\{x, y_{i_m}, y_{i_t}\}$,

$$|y_{i_m} - y_{i_t}|^2 = 1^2 + 1^2 - 2\cos\big(\angle(y_{i_m} x y_{i_t})\big)$$

Since $\angle(y_{i_m} x y_{i_t}) > 60^\circ$, we have,

$$\cos\big(\angle(y_{i_m} x y_{i_t})\big) < \frac{1}{2}$$

Continued...

Therefore,

$$|y_{i_m} - y_{i_t}|^2 > 1 \implies |y_{i_m} - y_{i_t}| > 1$$

Consequently,

$$B_{1/2}(y_{i_m}) \cap B_{1/2}(y_{i_t}) = \varnothing$$

For each $i_p \in P$, Let y_{i_p} be the intersection of the ray $\overrightarrow{x\,x_{i_p}}$ with the unit sphere $S^{n-1}(x)$.

By the previous argument, for any distinct $i_p, i_q \in P$,

$$B_{1/2}(y_{i_p}) \cap B_{1/2}(y_{i_q}) = \varnothing$$

Furthermore, since $|x - y_{i_p}| = 1$, if $z \in B_{1/2}(y_{i_p})$, then by the triangle inequality,

$$|z - x| \leq |z - y_{i_p}| + |y_{i_p} - x| < \frac{1}{2} + 1 = \frac{3}{2}$$

Hence,

$$B_{1/2}(y_{i_p}) \subseteq B_{3/2}(x) \qquad \text{for every } i_p \in P$$

Therefore,

$$\bigcup_{i_p \in P} B_{1/2}(y_{i_p}) \subseteq B_{3/2}(x)$$

Since the balls $\{B_{1/2}(y_{i_p})\}_{i_p \in P}$ are pairwise disjoint,

$$\sum_{i_p \in P} \mathrm{Vol}\big(B_{1/2}(y_{i_p})\big) = |P|\,\mathrm{Vol}(B_{1/2}) \leq \mathrm{Vol}(B_{3/2}(x))$$

Consequently,

$$|P| \leq \frac{\mathrm{Vol}(B_{3/2}(x))}{\mathrm{Vol}(B_{1/2})} = \frac{(3/2)^n}{(1/2)^n} = 3^n$$

Since $x \in \mathbb{R}^n$ was arbitrary and $|P|$ is precisely the number of selected balls containing x, every point of $\mathbb{R}^n$ belongs to at most 3^n balls of the selected subfamily. Therefore, the theorem holds with $N(n) = 3^n$. □

Theorem 54 (Modified Besicovitch Covering Theorem)**.** *Let $X = \mathbb{R}^n$ be a metric space. Let $\{B_{r_i}(x_i)\}$ be a finite family of open balls in X, where $x_i \in X$, $r_i > 0$ and $i \in I$ for some finite index set I. Fix $0 < \varepsilon < 1$. Then there exists $J \subseteq I$ such that,*

(i) Then there exists a finite sub-collection of balls $\{B_{r_{i_k}}(x_{i_k})\}$, where $x_{i_k} \in X$, $r_{i_k} > 0$ and $i_k \in J \subseteq I$ such that,

$$\bigcup_{i \in I} B_{(1-\varepsilon)r_i}(x_i) \subseteq \bigcup_{i_k \in J} B_{r_{i_k}}(x_{i_k})$$

(ii) Every $x \in X = \mathbb{R}^n$ *is contained in at-most* $N(n, \varepsilon)$ *balls in the subfamily* $\{B_{r_{i_k}}(x_{i_k})\}$.

Proof. Without loss of generality, Let the finite index set I be $\{1, 2, \cdots, N\}$ and reorder the balls in the family $\{B_{r_i}(x_i)\}$ in such a way that

$$r_1 \geq r_2 \geq \cdots \geq r_N$$

We proceed with a greedy algorithm. We pick the ball with largest radius, namely $B_{r_1}(x_1)$, and rename the index as $i_1 = 1$. Thus, at this stage, $J = \{i_1\} = \{1\} \subseteq I$. After choosing $B_{r_{i_1}}(x_{i_1})$, we remove from consideration every ball $B_{r_j}(x_j)$ such that,

$$B_{(1-\varepsilon)r_j}(x_j) \subseteq B_{r_{i_1}}(x_{i_1})$$

Next, from the balls still remaining, choose the ball with largest radius, say $B_{r_m}(x_m)$, and rename the index as $i_2 = m$. Thus, at this stage, $J = \{i_1, i_2\} = \{1, m\} \subseteq I$. After choosing $B_{r_{i_2}}(x_{i_2})$, remove from consideration every remaining ball $B_{r_j}(x_j)$ such that,

$$B_{(1-\varepsilon)r_j}(x_j) \subseteq B_{r_{i_2}}(x_{i_2})$$

Continuing inductively, at each stage we choose a ball of largest radius among the balls still under consideration. If this chosen ball is $B_{r_{i_k}}(x_{i_k})$, we add i_k to J, and then remove every remaining index j such that,

$$B_{(1-\varepsilon)r_j}(x_j) \subseteq B_{r_{i_k}}(x_{i_k})$$

Since the original family contains only finitely many balls, this process terminates after finitely many steps. Let

$$J = \{i_1, i_2, \ldots, i_M\} \subseteq I$$

be the resulting set of selected indices. By construction,

$$r_{i_1} \geq r_{i_2} \geq \cdots \geq r_{i_M}$$

Now we show that the selected family covers the union of the shrunken balls. Let $i \in I$. If $i \in J$, then

$$B_{(1-\varepsilon)r_i}(x_i) \subseteq B_{r_i}(x_i) \subseteq \bigcup_{i_k \in J} B_{r_{i_k}}(x_{i_k})$$

If $i \notin J$, then $B_{r_i}(x_i)$ was removed at some stage. Therefore, there exists some selected index $i_k \in J$ such that,

$$B_{(1-\varepsilon)r_i}(x_i) \subseteq B_{r_{i_k}}(x_{i_k})$$

Hence, in either case,

$$B_{(1-\varepsilon)r_i}(x_i) \subseteq \bigcup_{i_k \in J} B_{r_{i_k}}(x_{i_k})$$

Since $i \in I$ was arbitrary, we obtain,

$$\bigcup_{i \in I} B_{(1-\varepsilon)r_i}(x_i) \subseteq \bigcup_{i_k \in J} B_{r_{i_k}}(x_{i_k})$$

Fix any $x \in \mathbb{R}^n$. Define,

$$P = \{i_k \in J : x \in B_{r_{i_k}}(x_{i_k})\} \subseteq J$$

We will show that $|P|$ is bounded above by a constant depending only on n and ε.
For each $i_k \in P$, define

$$\rho_{i_k} = \frac{d(x, x_{i_k})}{r_{i_k}}$$

Since $x \in B_{r_{i_k}}(x_{i_k})$, we have

$$0 \leq \rho_{i_k} < 1$$

If $x \neq x_{i_k}$, Let D_{i_k} denote the unit vector in the direction from x to x_{i_k}.
If $x = x_{i_k}$, choose D_{i_k} to be any fixed unit vector. Thus $D_{i_k} \in S^{n-1}$. Let $\delta = \dfrac{\varepsilon}{3}$.
Since S^{n-1} is compact, there exist finitely many points $c_1, \ldots, c_L \in S^{n-1}$ such that,

$$S^{n-1} \subseteq \bigcup_{\ell=1}^{L} B_\delta(c_\ell)$$

Also cover $[0, 1]$ by finitely many intervals, $I_1, \ldots, I_R$, each having length at most δ.
Now suppose that $i_m, i_t \in P$ are distinct and,

$$x \in B_{r_{i_m}}(x_{i_m}) \cap B_{r_{i_t}}(x_{i_t})$$

Without loss of generality, assume that, $r_{i_m} \geq r_{i_t}$, and, in the case $r_{i_m} = r_{i_t}$, assume also that $B_{r_{i_m}}(x_{i_m})$ was selected no later than $B_{r_{i_t}}(x_{i_t})$.
Suppose D_{i_m} and D_{i_t} belong to the same δ-ball of the covering of S^{n-1}, i.e., $D_{i_m}, D_{i_t} \in B_\delta(c_\ell)$, for some ℓ, and $\rho_{i_m}, \rho_{i_t} \in I_s$, for some s.
We claim that this is impossible.
For simplicity, we write,

$$r' = r_{i_m}, \qquad x' = x_{i_m}, \qquad \rho' = \rho_{i_m}, \qquad D' = D_{i_m}$$

and

$$r'' = r_{i_t}, \qquad x'' = x_{i_t}, \qquad \rho'' = \rho_{i_t}, \qquad D'' = D_{i_t}$$

Since ρ' and ρ'' belong to the same interval of length at most δ, we have,

$$|\rho' - \rho''| \le \delta$$

Move x'' along the ray from x in the direction D'' to the point

$$x_1'' = x + \rho' r'' D''$$

Then,

$$d(x'', x_1'') = |\rho'' r'' - \rho' r''| = |\rho'' - \rho'| r'' \le \delta r''$$

Next define,

$$x_2'' = x + \rho' r'' D'$$

Since D' and D'' lie in the same ball $B_\delta(c_\ell)$, we have,

$$|D' - D''| < 2\delta$$

Therefore,

$$d(x_1'', x_2'') = \rho' r'' |D'' - D'| < 2\delta \rho' r'' \le 2\delta r''$$

Hence,

$$d(x'', x_2'') \le d(x'', x_1'') + d(x_1'', x_2'') < 3\delta r''$$

Now the balls $B_{r''}(x_2'')$ and $B_{r'}(x')$ are homothetic with respect to x, since their centers lie on the same ray from x and,

$$\frac{d(x, x_2'')}{r''} = \rho' = \frac{d(x, x')}{r'}$$

Since $r'' \le r'$, we have,

$$B_{r''}(x_2'') \subseteq B_{r'}(x')$$

Moreover, because,

$$d(x'', x_2'') < 3\delta r''$$

we get,

$$B_{(1-3\delta)r''}(x'') \subseteq B_{r''}(x_2'')$$

Thus,

$$B_{(1-3\delta)r''}(x'') \subseteq B_{r'}(x')$$

Since $\delta = \varepsilon/3$, this gives,

$$B_{(1-\varepsilon)r''}(x'') \subseteq B_{r'}(x')$$

By our assumption, either $r' > r''$, or $r' = r''$ and $B_{r'}(x')$ was selected no later than $B_{r''}(x'')$. Thus $B_{r'}(x')$ was selected no later than $B_{r''}(x'')$. Therefore, when $B_{r'}(x')$ was selected, the ball $B_{r''}(x'')$ would have been removed from consideration. Hence $B_{r''}(x'')$ could not also have been selected.

Continued...

This contradiction shows that, among the selected balls containing x, there is at most one selected ball containing x corresponding to each pair,

$$B_\delta(c_\ell) \times I_s$$

Therefore,

$$|P| \leq LR$$

The numbers L and R depend only on n and ε. Hence there exists a constant $N(n, \varepsilon)$ such that

$$|P| \leq N(n, \varepsilon)$$

Since $x \in \mathbb{R}^n$ was arbitrary, every point of $\mathbb{R}^n$ is contained in at most $N(n, \varepsilon)$ balls from the selected subfamily.
To make the constant explicit, choose the spherical cover so that,

$$L \leq \left(1 + \frac{2}{\delta}\right)^n$$

Also, choose the interval cover of $[0, 1]$ so that,

$$R \leq \left\lceil \frac{1}{\delta} \right\rceil + 1$$

Therefore,

$$|P| \leq LR \leq \left(1 + \frac{2}{\delta}\right)^n \left(\left\lceil \frac{1}{\delta} \right\rceil + 1\right)$$

Since

$$\delta = \frac{\varepsilon}{3}$$

we may take

$$N(n, \varepsilon) = \left(1 + \frac{6}{\varepsilon}\right)^n \left(\left\lceil \frac{3}{\varepsilon} \right\rceil + 1\right)$$

Hence,

$$|P| \leq N(n, \varepsilon)$$

□

Theorem 55 (Infinite Besicovitch Covering Theorem)**.** *Let $X = \mathbb{R}^n$ be a separable metric space. Let $0 < \varepsilon < 1$ and $\{B_\alpha\}$ be any family of balls in X, where $\alpha \in I$, for some index set I, such that, the supremum of radii, $\sup r_\alpha < \infty$.*

(i) Then there exists an at-most a countable subfamily of $\{B_\delta\}$, where $\delta \in J \subseteq I$, and J is a countable index set, such that,

$$\bigcup_{\alpha \in I} B_{(1-\varepsilon) r_\alpha}(x_\alpha) \subseteq \bigcup_{\delta \in J} B_{r_\delta}(x_\delta)$$

(ii) Every $x \in X = \mathbb{R}^n$ is contained in at-most $N(n, \varepsilon)$ balls in the subfamily of $\{B_\delta\}$.

Proof. Let $X = \mathbb{R}^n$. Let $0 < \varepsilon < 1$ and Let $\{B_\alpha\}_{\alpha \in I}$ be any family of open balls in X,such that,

$$R = \sup_{\alpha \in I} r_\alpha < \infty$$

If the original family is empty, then there is nothing to prove.
So assume that the family is non-empty. Then $R > 0$.
Define the open set

$$U = \bigcup_{\alpha \in I} B_{(1-\varepsilon)r_\alpha}(x_\alpha)$$

Since $X = \mathbb{R}^n$ is separable, it has a countable basis. Hence every open cover of an open subset of X admits a countable subcover by the same countable-basis argument used in Lemma 11. Thus, the open cover

$$U = \bigcup_{\alpha \in I} B_{(1-\varepsilon)r_\alpha}(x_\alpha)$$

admits a countable subcover. Therefore, there exists a countable subfamily

$$\{B_{\alpha_j}\}_{j=1}^{\infty} \subseteq \{B_\alpha\}_{\alpha \in I}$$

such that,

$$U = \bigcup_{j=1}^{\infty} B_{(1-\varepsilon)r_{\alpha_j}}(x_{\alpha_j})$$

Thus, it is enough to choose a suitable subfamily from the countable family $\{B_{\alpha_j}\}_{j=1}^{\infty}$.
Set,

$$\lambda = 1 - \frac{\varepsilon}{4}$$

For each $m \in \mathbb{N}$, define,

$$\mathcal{F}_m = \left\{ B_{\alpha_j} = B_{r_{\alpha_j}}(x_{\alpha_j}) : \lambda^m R < r_{\alpha_j} \leq \lambda^{m-1} R \right\}$$

Since $0 < \lambda < 1$ and $\lambda^m R \longrightarrow 0$, every ball in the countable family $\{B_{\alpha_j}\}_{j=1}^{\infty}$ belongs to exactly one of the families $\mathcal{F}_m$.
We now choose the desired subfamily. We process the families $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \ldots$ in this order.
Inside each $\mathcal{F}_m$, we use the order inherited from the enumeration $\{B_{\alpha_j}\}_{j=1}^{\infty}$. When a ball $B_{\alpha_j} = B_{r_{\alpha_j}}(x_{\alpha_j})$ is considered, we select it if, $B_{(1-\varepsilon)r_{\alpha_j}}(x_{\alpha_j})$ is not already contained in the union of the balls selected earlier. If it is already contained in the union of the previously selected balls, then we do not select it.

Let $\mathcal{G}$ be the family of all balls selected by this process. Since we selected from a countable family, $\mathcal{G}$ is at-most countable. Therefore, we may write

$$\mathcal{G} = \{B_\delta : \delta \in J\}$$

for some at-most countable set $J \subseteq I$.
By construction, for every $j \in \mathbb{N}$, either B_{α_j} was selected, or, $B_{(1-\varepsilon)r_{\alpha_j}}(x_{\alpha_j})$ was already contained in the union of previously selected balls. Hence,

$$B_{(1-\varepsilon)r_{\alpha_j}}(x_{\alpha_j}) \subseteq \bigcup_{\delta \in J} B_{r_\delta}(x_\delta) \qquad \forall j \in \mathbb{N}$$

Taking the union over all j, we get,

$$\bigcup_{j=1}^{\infty} B_{(1-\varepsilon)r_{\alpha_j}}(x_{\alpha_j}) \subseteq \bigcup_{\delta \in J} B_{r_\delta}(x_\delta)$$

Since

$$\bigcup_{j=1}^{\infty} B_{(1-\varepsilon)r_{\alpha_j}}(x_{\alpha_j}) = U = \bigcup_{\alpha \in I} B_{(1-\varepsilon)r_\alpha}(x_\alpha)$$

we obtain,

$$\bigcup_{\alpha \in I} B_{(1-\varepsilon)r_\alpha}(x_\alpha) \subseteq \bigcup_{\delta \in J} B_{r_\delta}(x_\delta)$$

It remains to prove the bounded overlap property. Fix $x \in \mathbb{R}^n$ and define,

$$P = \{\delta \in J : \ x \in B_{r_\delta}(x_\delta)\}$$

We claim that $|P|$ is bounded above by a constant depending only on n and ε.
The bounded overlap argument follows the same idea as the bounded overlap argument in the proof of the finite Modified Besicovitch Covering Theorem (Theorem 54), with one minor change.
In Theorem 54, when two selected balls

$$B' = B_{r'}(x') \qquad \text{and} \qquad B'' = B_{r''}(x'')$$

both contain x and B' was selected before B'', the ordering gave, $r' \geq r''$.
In the present construction, the balls are selected layer-by-layer. Hence, if B' was selected before B'', then either B' lies in an earlier layer or in the same layer as B''. In either case, by the definition of the layers,

$$r' > \lambda r'' = \left(1 - \frac{\varepsilon}{4}\right) r''$$

For a fixed point x, assign to every selected ball B_δ containing x the two quantities

- ratio: $\rho_\delta = \dfrac{|x - x_\delta|}{r_\delta} \in [0,1)$
- the direction: $D_\delta \in S^{n-1}$ from x to x_δ.

If $x = x_\delta$, choose D_δ to be any fixed unit vector.
Cover S^{n-1} by finitely many balls of radius $\varepsilon/12$, and cover $[0,1]$ by finitely many intervals of length at most $\varepsilon/12$.
Suppose two distinct selected balls containing x correspond to the same pair consisting of one spherical ball and one interval. Let the earlier one be, $B' = B_{r'}(x')$ and the later one be, $B'' = B_{r''}(x'')$. We claim that this is impossible.
Let

$$\delta = \frac{\varepsilon}{12}$$

As in the proof of Theorem 54, the fact that the two balls correspond to the same spherical ball and the same interval gives the estimate

$$|x' - x''| < 3\delta r'' + |r'' - r'|$$

We now show that this implies

$$B_{(1-\varepsilon)r''}(x'') \subseteq B_{r'}(x')$$

Let

$$z \in B_{(1-\varepsilon)r''}(x'')$$

Then,

$$|z - x''| < (1-\varepsilon)r''$$

First suppose $r' \geq r''$. Then, as $3\delta = \varepsilon/4 < \varepsilon$,

$$\begin{aligned} |z - x'| &\leq |z - x''| + |x'' - x'| \\ &< (1-\varepsilon)r'' + 3\delta r'' + (r' - r'') \\ &= r' - (\varepsilon - 3\delta)r'' \\ &< r' \end{aligned}$$

Now suppose $r' < r''$. Since

$$r' > \left(1 - \frac{\varepsilon}{4}\right) r''$$

we have,

$$r'' - r' < \frac{\varepsilon}{4} r''$$

Hence,

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

$$\begin{aligned}
&< (1-\varepsilon)r'' + 3\delta r'' + (r''-r') \\
&< (1-\varepsilon)r'' + \frac{\varepsilon}{4}r'' + \frac{\varepsilon}{4}r'' \\
&= \left(1-\frac{\varepsilon}{2}\right) r'' \\
&< \left(1-\frac{\varepsilon}{4}\right) r'' \\
&< r'
\end{aligned}$$

Therefore, in both cases, $z \in B_{r'}(x')$.
Thus,

$$B_{(1-\varepsilon)r''}(x'') \subseteq B_{r'}(x')$$

But then, when B'' was considered, its shrunken ball $B_{(1-\varepsilon)r''}(x'')$ was already contained in a previously selected ball. Hence B'' would not have been selected, which is a contradiction.
Therefore, for each fixed $x \in \mathbb{R}^n$, there is at most one selected ball containing x corresponding to each pair. Since the number of spherical balls and intervals depends only on n and ε, it follows that

$$|P| \leq N(n,\varepsilon)$$

For example, one may take,

$$N(n,\varepsilon) = \left(1+\frac{24}{\varepsilon}\right)^n \left(\left\lceil \frac{12}{\varepsilon} \right\rceil + 1\right)$$

Since $x \in \mathbb{R}^n$ was arbitrary, every point of $\mathbb{R}^n$ belongs to at most $N(n,\varepsilon)$ balls of the selected subfamily. ☐

Definition 41 (Maximal Function)**.** Let X be a separable metric space. Let μ be a locally finite Borel measure on X, i.e., $\forall x \in X,\ \exists r > 0$ such that $\mu(B_r(x)) < \infty$. Let f be a non-negative measurable function. Let ν be a Borel measure on X such that $\nu = f\,\mu$. Then, the maximal function of ν w.r.t. μ is,

$$M_\mu \nu(x) = \sup_{r>0} \frac{\nu\big(B_r(x)\big)}{\mu\big(B_r(x)\big)}$$

This is also called the Hardy-Littlewood maximal function.
We use the convention that if the denominator is 0, the supremum is 0.

Theorem 56 (Borel Measurability of the Maximal Function)**.** *Let X be a separable metric space. Let μ be a locally finite Borel measure on X, f be a non-negative measurable function and ν be a Borel measure on X. Then, the maximal function of ν w.r.t. μ, $M_\mu\nu$ is Borel measurable.*

Proof. To prove this, we aim to show that, for every $t > 0$, the set,

$$E_t = \{x \in X : \ M_\mu \nu(x) > t\}$$

is open, hence a Borel set.
If $x \in E_t$, then there exists $r > 0$ such that, $\mu(B_r(x)) > 0$ and $M_\mu \nu(x) > t$, hence,

$$\frac{\nu\big(B_r(x)\big)}{\mu\big(B_r(x)\big)} > t \quad \Longrightarrow \quad \nu\big(B_r(x)\big) > t\,\mu\big(B_r(x)\big)$$

Since

$$B_r(x) = \bigcup_{\rho<r} B_\rho(x)$$

Now, for ρ increasing to r, i.e., $\rho \uparrow r$ we have, by continuity of measures (from below),

$$\nu\big(B_\rho(x)\big) \longrightarrow \nu\big(B_r(x)\big)$$

Thus, there exists, $\rho < r$, such that,

$$\nu\big(B_\rho(x)\big) > t\,\mu\big(B_r(x)\big)$$

Let $y \in B_\delta(x)$ and $d(x, y) < \delta$.
Let $\delta = \dfrac{r-\rho}{2} > 0$. Define, $R = \rho + \delta = r - \delta$.
Let $z \in B_\rho(x)$. Then,

$$d(z, y) \leq d(z, x) + d(x, y) < \rho + \delta = R$$

Thus, $z \in B_R(y)$. So, $B_\rho(x) \subseteq B_R(y)$.
Again, as $z \in B_R(y)$. So,

$$d(z, x) \leq d(z, y) + d(y, x) < R + \delta = r$$

Thus, $z \in B_r(x)$. So, $B_R(y) \subseteq B_r(x)$.
Therefore, we have,

$$B_\rho(x) \subseteq B_R(y) \subseteq B_r(x)$$

This results in,

$$\nu\big(B_R(y)\big) \geq \nu\big(B_\rho(x)\big) > t\,\mu\big(B_r(x)\big) \geq t\,\mu\big(B_R(y)\big)$$

Therefore,

$$\nu\big(B_R(y)\big) > t\,\mu\big(B_R(y)\big)$$

Since this holds for every $y \in B_\delta(x)$,we have, $B_\delta(x) \subseteq E_t$.
Thus, E_t is open, for all $t > 0$.
Therefore, $M_\mu \nu$ is Borel measurable. □

Continued...

Theorem 57 (Weak Type 1-1 bound for Maximal Functions)**.** *Let X be a separable metric space. Let μ be a locally finite Borel measure on X and ν be a finite Borel measure on X. Assume one of the following holds:*

(i) μ is doubling, i.e.

$$\mu\big(B_{2r}(x)\big) \leq c\,\mu\big(B_r(x)\big)$$

with $c < \infty$ and c independent of x and r.

(ii) $X = \mathbb{R}^n$ or any Besicovitch satisfying space.

Then, for all $t > 0$,

$$\mu\big(\{x \in X : \ M_\mu \nu(x) > t\}\big) \leq \frac{C}{t}\,\nu(X)$$

where $C = c^3$ in the doubling case and $C = N$ in the Besicovitch case.

Proof. Let

$$E = \{x \in X : \ M_\mu \nu(x) > t\}$$

Then, for every $x \in E$, there exists $r'_x > 0$, such that,

$$\nu\big(B_{r'_x}(x)\big) > t\,\mu\big(B_{r'_x}(x)\big)$$

So we consider the family of balls $\{B_{r'_x}(x) : x \in E\}$.
For $R > 0$ define the truncated maximal ratio,

$$M_\mu^R \nu(x) = \sup_{0<r<R} \frac{\nu\big(B_r(x)\big)}{\mu\big(B_r(x)\big)}$$

Let

$$E_R = \{x \in X : \ M_\mu^R \nu(x) > t\}$$

Then, for every $x \in E_R$, there exists $r_x > 0$ (with $r_x < R$), such that,

$$\nu\big(B_{r_x}(x)\big) > t\,\mu\big(B_{r_x}(x)\big)$$

(i) Assuming μ is doubling, the collection $\{B_{r_x}(x) : x \in E_R\}$ is an arbitrary collection of balls covering E_R and by Vitali Covering Theorem (Theorem 52), $\exists$ a countable subfamily $\{B_{r_{x_j}}(x_j)\}$ of pairwise disjoint balls, such that,

$$E_R \subseteq \bigcup_{x \in E_r} B_{r_x}(x) \subseteq \bigcup_{j=1}^{\infty} B_{5r_{x_j}}(x_j)$$

Continued...

Then,

$$
\begin{aligned}
\mu(E_R) &\leq \mu\Big(\bigcup_{j=1}^{\infty} B_{r_{x_j}}(x_j)\Big) && \text{(by monotonicity)}\\
&\leq \sum_{j=1}^{\infty} \mu\big(B_{5r_{x_j}}(x_j)\big) && \text{(by sub-additivity)}\\
&\leq \sum_{j=1}^{\infty} \mu\big(B_{8r_{x_j}}(x_j)\big) && \text{(since } B_{5r_{x_j}}(x_j) \subseteq B_{8r_{x_j}}(x_j))\\
&\leq c^3 \sum_{j=1}^{\infty} \mu\big(B_{r_{x_j}}(x_j)\big) && \text{(by applying doubling three times)}\\
&< \frac{c^3}{t} \sum_{j=1}^{\infty} \nu\big(B_{r_{x_j}}(x_j)\big) && \text{(by hypothesis)}\\
&= \frac{c^3}{t}\, \nu\left(\bigcup_{j=1}^{\infty} B_{r_{x_j}}(x_j)\right) && \text{(by disjointness)}\\
&\leq \frac{c^3}{t}\, \nu(X)
\end{aligned}
$$

Therefore,

$$\mu(E_R) \leq \frac{c^3}{t}\, \nu(X)$$

where c^3 is independent of R.

(ii) Fix $0 < \varepsilon < 1$. Since

$$x \in B_{(1-\varepsilon)r_x}(x)$$

for every $x \in E_R$, we have,

$$E_R \subseteq \bigcup_{x \in E_R} B_{(1-\varepsilon)r_x}(x)$$

Assuming, $X = \mathbb{R}^n$ or any Besicovitch satisfying space, by Besicovitch Covering Theorem (Theorem 55), there exists a countable subfamily $\{B_{r_{x_j}}(x_j)\}_{j=1}^{\infty}$ such that,

$$E_R \subseteq \bigcup_{x \in E_R} B_{(1-\varepsilon)r_x}(x) \subseteq \bigcup_{j=1}^{\infty} B_{r_{x_j}}(x_j)$$

and every point of X is covered by at most N balls of this subfamily.
Thus,

$$\mu(E_R) \leq \mu\Big(\bigcup_{j=1}^{\infty} B_{r_{x_j}}(x_j)\Big) \qquad \text{(by monotonicity)}$$

$$
\begin{aligned}
&\leq \sum_{j=1}^{\infty} \mu\big(B_{r_{x_j}}(x_j)\big) && \text{(by sub-additivity)}\\
&< \frac{1}{t}\sum_{j=1}^{\infty} \nu\big(B_{r_{x_j}}(x_j)\big) && \text{(by hypothesis)}\\
&= \frac{1}{t}\sum_{j=1}^{\infty} \int_X \chi_{B_{r_{x_j}}(x_j)}\, d\nu && \text{(by maximal function)}\\
&= \frac{1}{t}\int_X \Big(\sum_{j=1}^{\infty} \chi_{B_{r_{x_j}}(x_j)}\Big)\, d\nu && \text{(by Tonelli's theorem)}\\
&\leq \frac{1}{t}\int_X N\, d\nu && \text{(by bounded overlap)}\\
&= \frac{N}{t}\,\nu(X)
\end{aligned}
$$

Therefore,

$$\mu(E_R) \leq \frac{N}{t}\,\nu(X)$$

where N is independent of R.

By construction, $\{E_R\}$ is a nested increasing family of sets. Thus, we have $E_R \uparrow E$ and $M_\mu^R \uparrow M_\mu \nu$.
In either case, there exists a constant C independent of R such that

$$\mu(E_R) \leq \frac{C}{t}\nu(X)$$

Therefore, by continuity from below,

$$\mu(E) = \lim_{R\to\infty} \mu(E_R) \leq \lim_{R\to\infty} \frac{C}{t}\nu(X) = \frac{C}{t}\nu(X)$$

Therefore,

$$\mu\big(\{x \in X :\ M_\mu\nu(x) > t\}\big) \leq \frac{C}{t}\nu(X)$$

□

Theorem 58 (Lebesgue Differentiation Theorem)**.** *Let X be a separable metric space and μ and ν be locally finite measures on $\mathcal{B}(X)$. Assume the weak type $(1,1)$ bound holds and the Lebesgue decomposition*

$$\nu = f\,\mu + \nu_s$$

holds, where f is a nonnegative measurable function and X_s is the corresponding singular set, such that,

$$\mu(X_s) = \nu_s(X_s^c) = 0$$

Then,

(i) *For μ-a.e. x, and for $x \in G$, we have*

$$f(x) = \lim_{r \to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))}$$

where G is any set on which μ and ν are finite.

(ii) *For ν_s-a.e. $x \in G$,*

$$\lim_{r \to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} = \infty$$

Proof. (i) To prove this we first consider the case $\nu_s = 0$, i.e. $\nu = f\mu$. Then two possible cases occur.

- **Case I:** f is continuous.
  Consider the ball $B_r(x)$.
  As $\nu = f\,\mu$, then
  $$\nu(B_r(x)) = \int_{B_r(x)} f\,d\mu$$
  And also,
  $$f(x)\,\mu(B_r(x)) = \int_{B_r(x)} f(x)\,d\mu$$
  Then,
  $$\begin{aligned}\left|\frac{\nu(B_r(x))}{\mu(B_r(x))} - f(x)\right| &= \left|\frac{\nu(B_r(x)) - f(x)\,\mu(B_r(x))}{\mu(B_r(x))}\right| \\ &= \frac{1}{\mu(B_r(x))}\left|\int_{B_r(x)} f\,d\mu - \int_{B_r(x)} f(x)\,d\mu\right| \\ &= \frac{1}{\mu(B_r(x))}\left|\int_{B_r(x)} (f - f(x))\,d\mu\right| \\ &\leq \frac{1}{\mu(B_r(x))}\int_{B_r(x)} |f - f(x)|\,d\mu\end{aligned}$$

  By continuity at x of f, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$. So, if $r < \delta$, then,

  $$\begin{aligned}\left|\frac{\nu\big(B_r(x)\big)}{\mu\big(B_r(x)\big)} - f(x)\right| &\leq \frac{1}{\mu\big(B_r(x)\big)}\int_{B_r(x)} |f - f(x)|\,d\mu \\ &\leq \frac{1}{\mu\big(B_r(x)\big)}\int_{B_r(x)} \varepsilon\,d\mu\end{aligned}$$

Continued...

$$\leq \frac{\varepsilon \cdot \mu\big(B_r(x)\big)}{\mu\big(B_r(x)\big)} \leq \varepsilon.$$

Therefore,

$$f(x) = \lim_{r\to 0} \frac{\nu\big(B_r(x)\big)}{\mu\big(B_r(x)\big)}, \qquad \text{for } \mu\big(B_r(x)\big) \neq 0$$

Now, when $\mu\big(B_r(x)\big) = 0$, we claim that in a separable space X, if,

$$E = \{x \in X : \exists r > 0,\ \mu(B_r(x)) = 0\}$$

then $\mu(E) = 0$.
As X is separable, there exists a countable dense subset $D \subseteq X$.
Assume a ball $B_r(x)$ such that $\mu\big(B_r(x)\big) = 0$.
Take $\varepsilon > 0$. Since D is dense, there exists $y \in D$ with $d(x, y) < \varepsilon$.
If $\rho = r - \varepsilon > 0$, then, $B_\rho(y) \subseteq B_r(x)$ and $x \in B_\rho(y)$, whenever $\varepsilon < r/2$ and, as $\mu\big(B_r(x)\big) = 0$, we have $\mu\big(B_\rho(y)\big) = 0$.
Thus, it is enough to show that $\mu(F) = 0$, where,

$$F = \{x \in G : \ \exists \rho \in \mathbb{Q}^+,\ y \in D,\ x \in B_\rho(y),\ \mu(B_\rho(y)) = 0\}$$

Now,

$$F \subseteq \bigcup_{\substack{y\in D\,,\,\rho\in\mathbb{Q}^+,\\ \mu(B_\rho(y))=0,\\ \text{countable}}} B_\rho(y)$$

Hence,

$$\mu(F) \leq \sum_{\substack{y\in D\,,\,\rho\in\mathbb{Q}^+,\\ \text{countable}}} \mu(B_\rho(y)) = 0$$

Therefore $\mu(E) = 0$ in separable space X.
Therefore, for continuous function f, $x \in G$, and μ-a.e. we have,

$$f(x) = \lim_{r\to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))}$$

This proves the continuous case.

- **Case II:** f is non-negative measurable and integrable on G.
  By the density of continuous functions in L^1 (Theorem 34), for every $\varepsilon > 0$, there exists a continuous function g on G such that,

$$\int_G |f - g|\, d\mu < \varepsilon$$

Let $h = f - g$. Then, $f = g + h$.
Consider the ball $B_r(x) \subseteq G$. Since $\nu = f\mu$ in the present case, we have,

$$\frac{\nu(B_r(x))}{\mu(B_r(x))} = \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f \, d\mu$$

Therefore,

$$\begin{aligned}
\left| \frac{\nu(B_r(x))}{\mu(B_r(x))} - f(x) \right| &= \left| \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f(y) \, d\mu(y) - f(x) \right| \\
&= \left| \frac{1}{\mu(B_r(x))} \int_{B_r(x)} (f(y) - f(x)) \, d\mu(y) \right| \\
&\leq \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |f(y) - f(x)| \, d\mu(y) \\
&\leq \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |g(y) - g(x)| \, d\mu(y) \\
&\quad + \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |h(y) - h(x)| \, d\mu(y) \\
&= I_1(x, r) + I_2(x, r).
\end{aligned}$$

By **Case I**, g being continuous, $I_1(x, r) \to 0$ as $r \to 0$.
Also,

$$\begin{aligned}
I_2(x, r) &= \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |h(y) - h(x)| \, d\mu(y) \\
&\leq \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |h(y)| \, d\mu(y) + |h(x)|
\end{aligned}$$

Now fix $t > 0$ and consider,

$$E = \{x \in G : \ \limsup_{r \to 0} I_2(x, r) > t\}$$

By the estimate for I_2, we have

$$E \subseteq E_1 \cup E_2$$

where,

$$E_1 = \left\{ x \in G : \ \sup_{\substack{r > 0 \\ B_r(x) \subseteq G}} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |h| \, d\mu > \frac{t}{2} \right\}$$

and,

$$E_2 = \left\{ x \in G : |h(x)| > \frac{t}{2} \right\}$$

Continued...

Since

$$E_1 \subseteq \left\{ x \in G : M_\mu(|h|\mu)(x) > \frac{t}{2} \right\}$$

Hence, by the weak type $(1,1)$ bound (Theorem 57), we get,

$$\mu(E_1) \leq \frac{2C}{t}(|h|\mu)(G) = \frac{2C}{t}\int_G |h|\, d\mu < \frac{2C\varepsilon}{t}$$

Also, by Chebyshev's inequality,

$$\mu(E_2) = \mu\left(\left\{ x \in G : |h(x)| > \frac{t}{2} \right\}\right) \leq \frac{2}{t}\int_G |h|\, d\mu < \frac{2\varepsilon}{t}$$

Therefore,

$$\mu(E) \leq \frac{2(C+1)\varepsilon}{t}$$

Since $\varepsilon > 0$ was arbitrary. Then for every $t > 0$,

$$\mu(E) = 0$$

Taking $t = 1/m$, $m \in \mathbb{N}$, (letting $t \longrightarrow 0$) and using countable subadditivity, we get,

$$\limsup_{r \to 0} I_2(x,r) = 0 \qquad \text{for } \mu\text{-a.e. } x \in G$$

Since $I_1(x,r) \to 0$ by the continuous case, it follows that,

$$\lim_{r \to 0} \left| \frac{\nu(B_r(x))}{\mu(B_r(x))} - f(x) \right| = 0$$

for μ-a.e. $x \in G$.
Therefore,

$$f(x) = \lim_{r \to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))}$$

for μ-a.e. $x \in G$.

This proves the absolutely continuous case.
Now we prove when ν also has a singular part ν_s and thus a singular set X_s exists, s.t. $\mu(X_s) = \nu_s(X_s^c) = 0$.
As μ and ν are locally finite we consider small chunks G of X. Then $\mu(G)$, $\nu_s(G) < \infty$.
Also, by separability of X, we have already shown that

$$\mu\big(\{x \in G : \ \exists r > 0, \ \mu(B_r(x)) = 0\}\big) = 0$$

Thus, we show that, for μ-a.e. then,

$$\lim_{r\to 0}\frac{\nu_s(B_r(x))}{\mu(B_r(x))}=0$$

Take $\varepsilon > 0$. By regularity of the ν_s on G, there exists a closed set $F \subseteq X_s \cap G$ such that,

$$\nu_s((X_s \cap G)\setminus F) < \varepsilon$$

Define,

$$\nu_s' = \nu_s \cdot \mathbb{1}_{F^c \cap G}$$

Then,

$$\nu_s'(X) < \varepsilon$$

If $x \in G \setminus F$, then since F is closed, there exists $r_x > 0$ such that

$$B_r(x) \cap F = \varnothing \qquad \text{for all } 0 < r < r_x$$

Therefore, for such r,

$$\nu_s(B_r(x)) = \nu_s'(B_r(x))$$

Fix $t > 0$ and define,

$$E_t = \left\{ x \in G \setminus F : \limsup_{r\to 0}\frac{\nu_s(B_r(x))}{\mu(B_r(x))} > t \right\}$$

Then,

$$E_t \subseteq \{x \in G\setminus F : \sup_{r\in(0,r_x)}\frac{\nu_s'(B_r(x))}{\mu(B_r(x))} > t\} \subseteq \{x \in G : M_\mu \nu_s'(x) > t\}$$

By the weak type $(1,1)$ bound,

$$\mu(E_t) \le \frac{C}{t}\nu_s'(X) < \frac{C\varepsilon}{t}$$

Since $\varepsilon > 0$ was arbitrary. Thus,

$$\mu(E_t) = 0$$

Taking $t = 1/m$, $m \in \mathbb{N}$, (letting $t \longrightarrow 0$) and using countable subadditivity, we conclude that,

$$\mu(E_t) = 0$$

Hence for μ-a.e. $x \in G$,

$$\lim_{r\to 0}\frac{\nu_s(B_r(x))}{\mu(B_r(x))}=0$$

Combining the absolutely continuous part and the singular part, we have

$$\frac{\nu(B_r(x))}{\mu(B_r(x))} = \frac{(f\mu)(B_r(x))}{\mu(B_r(x))} + \frac{\nu_s(B_r(x))}{\mu(B_r(x))}$$

The first term tends to $f(x)$ for μ-a.e. $x \in G$, and the second term tends to 0 for μ-a.e. $x \in G$. Hence,

$$f(x) = \lim_{r \to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))}$$

for μ-a.e. $x \in G$, where G is any set on which μ and ν are finite.
This proves part (i).

(ii) We show that,

$$\lim_{r \to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} = \infty \qquad \text{for } \nu_s\text{-a.e. } x \in G$$

Since $\nu_s \perp \mu$, there exists a singular set X_s such that,

$$\mu(X_s) = 0 \qquad \text{and} \qquad \nu_s(X_s^c) = 0$$

Let $\varepsilon > 0$. By regularity of μ, choose an open set U such that,

$$X_s \cap G \subseteq U \qquad \text{and} \qquad \mu(U) < \varepsilon$$

Define, $\mu' = \mu \cdot \mathbb{1}_U$.
Then

$$\mu'(X) = \mu(U) < \varepsilon$$

Fix $t > 0$ and consider

$$A_t = \left\{ x \in X_s \cap G : \liminf_{r \to 0} \frac{\nu_s(B_r(x))}{\mu(B_r(x))} < t \right\}$$

If $x \in A_t$, then for arbitrarily small $r > 0$,

$$\nu_s(B_r(x)) < t\,\mu(B_r(x))$$

Since $x \in U$ and U is open, for all sufficiently small $r > 0$ we have,

$$B_r(x) \subseteq U$$

Hence, for such r,

$$\mu(B_r(x)) = \mu'(B_r(x))$$

Therefore, for arbitrarily small $r > 0$,

$$\frac{\mu'(B_r(x))}{\nu_s(B_r(x))} > \frac{1}{t}$$

By separability we have,

$$\nu_s\big(\{x \in G : \exists\, r > 0,\ \nu_s(B_r(x)) = 0\}\big) = 0$$

Thus, ignoring the ν_s-null set of points for which $\nu_s(B_r(x)) = 0$ for some sufficiently small ball, we get,

$$A_t \subseteq \left\{x \in X : M_{\nu_s}\mu'(x) > \frac{1}{t}\right\}$$

Applying the weak type $(1, 1)$ bound (Theorem 57), with ν_s as the reference measure and μ' as the finite measure, we obtain,

$$\nu_s(A_t) \leq Ct\,\mu'(X) \leq Ct\,\varepsilon$$

Since $\varepsilon > 0$ was arbitrary, we have,

$$\nu_s(A_t) = 0$$

Taking $t = m$, $m \in \mathbb{N}$, $(t \longrightarrow \infty)$ and using

$$\left\{x \in X_s \cap G : \liminf_{r \to 0} \frac{\nu_s(B_r(x))}{\mu(B_r(x))} < \infty\right\} = \bigcup_{m=1}^{\infty} A_m$$

we obtain,

$$\nu_s\left(\left\{x \in X_s \cap G : \liminf_{r \to 0} \frac{\nu_s(B_r(x))}{\mu(B_r(x))} < \infty\right\}\right) = 0$$

Therefore,

$$\liminf_{r \to 0} \frac{\nu_s(B_r(x))}{\mu(B_r(x))} = \infty \qquad \text{for } \nu_s\text{-a.e. } x \in X_s \cap G$$

Since the ratios are nonnegative, thus,

$$\liminf_{r \to 0} \frac{\nu_s(B_r(x))}{\mu(B_r(x))} = \infty \implies \lim_{r \to 0} \frac{\nu_s(B_r(x))}{\mu(B_r(x))} = \infty$$

Hence,

$$\lim_{r \to 0} \frac{\nu_s(B_r(x))}{\mu(B_r(x))} = \infty \qquad \text{for } \nu_s\text{-a.e. } x \in G$$

Since

$$\nu = f\mu + \nu_s$$

we have,

$$\frac{\nu(B_r(x))}{\mu(B_r(x))} = \frac{(f\mu)(B_r(x))}{\mu(B_r(x))} + \frac{\nu_s(B_r(x))}{\mu(B_r(x))}$$

The first term is nonnegative, and the second term tends to ∞ for ν_s-a.e. $x \in G$. Therefore,

$$\lim_{r\to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} = \infty \qquad \text{for } \nu_s\text{-a.e. } x \in G$$

This proves part (ii).

$\square$

Definition 42 (Push-forward Measure)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space and $(Y, \mathcal{B})$ be a measurable space. Let $T : X \to Y$ be a measurable function. The push-forward measure $T_*\mu$ on Y is defined for any set $B \in \mathcal{B}$ by

$$(T_*\mu)(B) = \mu\big(T^{-1}(B)\big)$$

Theorem 59 (Validity of the Push-Forward Measure)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space and $(Y, \mathcal{B})$ be a measurable space. Let $T : X \to Y$ be a measurable function. The push-forward measure $T_*\mu$ on Y is a measure on $\mathcal{B}$.*

Proof. To prove that $T_*\mu$ is a measure on $\mathcal{B}$, it is enough to show non-negativity and countable additivity.

- Consider the empty set $\varnothing$. Then,

  $$(T_*\mu)(\varnothing) = \mu\big(T^{-1}(\varnothing)\big)$$

  The pre-image of the empty set is always empty, so $T^{-1}(\varnothing) = \varnothing$, as no points of X can be mapped to a set that has nothing, and hence $\mu(\varnothing) = 0$. Therefore,

  $$(T_*\mu)(\varnothing) = 0$$

  Also, for every $B \in \mathcal{B}$,

  $$(T_*\mu)(B) = \mu(T^{-1}(B)) \geq 0$$

  which shows non-negativity at the empty set.

- Now Let $\{B_i\}_{i=1}^{\infty}$ be a sequence of disjoint sets in $\mathcal{B}$. Then, by property of preimages,

  $$T^{-1}\Big(\bigcup_{i=1}^{\infty} B_i\Big) = \bigcup_{i=1}^{\infty} T^{-1}(B_i)$$

  As each B_i are disjoint in Y, their pre-image $T^{-1}(B_i)$ must be disjoint in X. Else, if x were in both $T^{-1}(B_i)$ and $T^{-1}(B_j)$, then $T(x) \in B_i$ and $T(x) \in B_j$,

hence, contradicting disjointness.
Then,

$$\begin{aligned}(T_*\mu)\Big(\bigcup_{i=1}^{\infty} B_i\Big) &= \mu\Big(T^{-1}\big(\bigcup_{i=1}^{\infty} B_i\big)\Big)\\ &= \mu\Big(\bigcup_{i=1}^{\infty} T^{-1}(B_i)\Big)\\ &= \sum_{i=1}^{\infty} \mu\big(T^{-1}(B_i)\big)\\ &= \sum_{i=1}^{\infty} (T_*\mu)(B_i)\end{aligned}$$

Thus, $T_*\mu$ is countably additive.

Therefore, $T_*\mu$ is a measure on $\mathcal{B}$. ☐

Theorem 60 (Abstract Change of Variable)**.** *Let $T : X \to Y$ be a measurable function and $f : Y \to \mathbb{R}$ be a $\mathcal{B}$-measurable function. Then f is integrable with respect to $T_*\mu$ if and only if $f \circ T$ is integrable with respect to μ. Mathematically,*

$$\int_Y f\, d(T_*\mu) = \int_X (f \circ T)\, d\mu$$

Proof. We first prove the result for non-negative measurable functions, i.e., $f \geq 0$.

- Consider first the case $f = \mathbb{1}_B$ where $B \in \mathcal{B}$ (the characteristic function of B). Then,

  $$\int_Y \mathbb{1}_B\, d(T_*\mu) = (T_*\mu)(B) = \mu\big(T^{-1}(B)\big)$$

  On the other hand, since $\mathbb{1}_B(T(x)) = 1$ if and only if $T(x) \in B$, i.e. $x \in T^{-1}(B)$. Hence,

  $$\mathbb{1}_B \circ T = \mathbb{1}_{T^{-1}(B)}$$

  Thus,

  $$\int_X (\mathbb{1}_B \circ T)\, d\mu = \int_X \mathbb{1}_{T^{-1}(B)}\, d\mu = \mu\big(T^{-1}(B)\big)$$

  Therefore, for characteristic functions,

  $$\int_Y f\, d(T_*\mu) = \int_X (f \circ T)\, d\mu$$

- Considering simple functions as

  $$f = \sum_{k=1}^{n} c_k \mathbb{1}_{B_k}$$

  Then, for each k, the statement holds for $\mathbb{1}_{B_k}$. Then,

  $$\begin{aligned}\int_Y \Big(\sum_{k=1}^{n} c_k \mathbb{1}_{B_k}\Big)\, d(T_*\mu) &= \sum_{k=1}^{n} c_k \int_Y \mathbb{1}_{B_k}\, d(T_*\mu) \\ &= \sum_{k=1}^{n} c_k \int_X (\mathbb{1}_{B_k} \circ T)\, d\mu \\ &= \int_X \Big((\sum_{k=1}^{n} c_k \mathbb{1}_{B_k}) \circ T \Big)\, d\mu\end{aligned}$$

  Therefore, for simple functions f,

  $$\int_Y f\, d(T_*\mu) = \int_X (f \circ T)\, d\mu$$

- Now Let f be any non-negative measurable function. Then f is the pointwise limit of an increasing sequence of simple functions. Let $f = \lim_{n\to\infty} f_n$ with $f_n \uparrow f$. Then, for each n, the statement holds for each simple function f_n, i.e.,

  $$\int_Y f_n\, d(T_*\mu) = \int_X (f_n \circ T)\, d\mu \qquad (n = 1, 2, \dots)$$

  Taking limits on both sides and applying the Monotone Convergence Theorem (Theorem 28), we obtain,

  $$\lim_{n\to\infty} \int_Y f_n\, d(T_*\mu) = \int_Y f\, d(T_*\mu), \qquad \lim_{n\to\infty} \int_X (f_n \circ T)\, d\mu = \int_X (f \circ T)\, d\mu$$

  Hence,

  $$\int_Y f\, d(T_*\mu) = \int_X (f \circ T)\, d\mu$$

Now Let $f : Y \to \mathbb{R}$ be measurable. Applying the non-negative case to $|f|$, we get

$$\int_Y |f|\, d(T_*\mu) = \int_X |f \circ T|\, d\mu$$

Therefore,

$$f \in L^1(T_*\mu) \quad \Longleftrightarrow \quad f \circ T \in L^1(\mu)$$

If these equivalent conditions hold, then applying the non-negative case to f^+ and f^- gives

$$\int_Y f^+ \, d(T_*\mu) = \int_X (f^+ \circ T) \, d\mu$$

and

$$\int_Y f^- \, d(T_*\mu) = \int_X (f^- \circ T) \, d\mu$$

Since

$$f = f^+ - f^-$$

and

$$(f \circ T)^+ = f^+ \circ T, \qquad (f \circ T)^- = f^- \circ T$$

we obtain

$$\int_Y f \, d(T_*\mu) = \int_X (f \circ T) \, d\mu$$

Therefore, f is integrable with respect to $T_*\mu$ if and only if $f \circ T$ is integrable with respect to μ. $\square$

Continued...