Modern Analysis

Lectures by Shimon Brooks (brookss@math.biu.ac.il)
Summary by Ari Feiglin (ari.feiglin@gmail.com)

Contents

1 Introduction

The Riemann Integral and its Faults

Recall the definition of the Riemann integral: given a function $f(x)$ on an interval $[a,b]$ we take a partition of this interval and representatives from each partition, x_i. Then the Riemann sum of this function over this partition is $\sum f(x_i)\Delta x_i$. Then if this sum converges to a value as $\sup \Delta x_i$ converges to 0, this function is *Riemann integrable* and has an integral represented by $\int_a^b f(x)\,dx$. Previously we have shown that a function is Riemann integrable if and only if it is almost always continuous, in particular all continuous functions are continuous.

But take for example $\mathbb{Q}$'s indicator function

$$\chi_{\mathbb{Q}}(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}.$$

This function is nowhere continuous and thus is not Riemann integrable. But we argue that it *should* be integrable.

Any theory if an integral should have the following two basic properties:

(**1**) *monotonicity*: if $f(x) \leq g(x)$ for every x in the domain then $\int f(x)\,dx \leq \int g(x)\,dx$.

(**2**) *linearity*: $\int (\alpha f(x) + \beta g(x))\,dx = \alpha \int f(x)\,dx + \beta \int g(x)\,dx$.

Let $\{q_n\}_{n=1}^\infty$ be an enumeration of the rationals and $\varepsilon > 0$. Let us define $E_n := \left(q_n - \frac{\varepsilon}{2^{n+1}}, q_n + \frac{\varepsilon}{2^{n+1}}\right)$, then we should think that

$$\int \sum_{n=1}^\infty \chi_{E_n}(x)\,dx = \sum_{n=1}^\infty \int \chi_{E_n}(x)\,dx = \sum_{n=1}^\infty \frac{\varepsilon}{2^n} = \varepsilon \tag{1}$$

Now $\chi_{\mathbb{Q}}(x) \leq \sum_{n=1}^\infty \chi_{E_n}(x)$ for all x and so $\int \chi_{\mathbb{Q}}(x)\,dx \leq \varepsilon$ for every $\varepsilon > 0$ and so we should think that $\int \chi_{\mathbb{Q}}(x)\,dx = 0$.

Now obviously $\chi_{\mathbb{Q}}$ is not Riemann-integrable, and so there is an issue with the above argument. In fact there are two: firstly in (1) we assumed that $\int \sum_{n=1}^\infty \chi_{E_n} = \sum_{n=1}^\infty \int \chi_{E_n}$, which only holds if the sum converges uniformly. Secondly, $\sum_{n=1}^\infty \chi_{E_n}(x)$ takes on infinite values (for every rational number, the sum is infinite) and so it is not even Riemann integrable.

So we want a theory of integration which allows for two things: 1) the ability to deal with convergence of an integral without necessarily needing uniform convergence, 2) the ability to deal with functions which are not Riemann-integrable. Lebesgue's theory of integration is based on the following observation: partitioning the domain of the function will necessarily require some form of continuity, so instead try partitioning the *range* of the function. So given the partition $y_0 < \cdots < y_n$ we can imagine some Lebesgue sum of this partition to be

$$\sum_{i=0}^n y_i \cdot |E_i|, \qquad E_i = \{x \in [a,b] \mid y_i \leq f(x) \leq y_{i+1}\}$$

where $|S|$ is some notion of the "width" of the set S, *which we have not yet defined* (it is not the cardinality of the set). For an interval this should be the length of the interval, but for arbitrary sets it becomes harder to understand how we should approach defining it. And in order to define the integral of $\chi_{\mathbb{Q}}$, it is necessary to define this width for arbitrary sets, or at least for a larger family of sets than just intervals, since computing the Lebesgue sums of $\chi_{\mathbb{Q}}$ will involve terms containing $|\mathbb{Q}|$.

This "Lebesgue sum" will not be precisely how we define Lebesgue integration, but it does give us a starting point: how do we define the "width" of a set $E \subseteq \mathbb{R}$.

2 Lebesgue Integration

2.1 The Lebesgue Measure

We would like a function m which measures the width of arbitrary sets $E \subseteq \mathbb{R}$. Such a function would ideally satisfy the following properties:

(**m1**) m is a function $m\colon \mathcal{P}(\mathbb{R}) \longrightarrow [0, \infty]$,

(**m2**) for every interval I, $m(I)$ is the length of I,

(**m3**) the measure of a set is preserved under movement, ie. $m(E + \alpha) = m(E)$ for every $\alpha \in \mathbb{R}$ where $E + \alpha = \{x + \alpha \mid x \in E\}$,

(**m4**) if $\{E_n\}_{n=1}^{\infty}$ is a sequence of disjoint sets then $m(\dot{\bigcup}_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} m(E_n)$. This is called *$\sigma$-additivity*.

2.1.1 Definition (The Outer Measure)

Let $E \subseteq \mathbb{R}$, then we define E's **outer measure** to be

$$m^*(E) := \inf\left\{ \sum_{n=1}^{\infty} |I_n| \;\middle|\; E \subseteq \bigcup_{n=1}^{\infty} I_n \text{ where } I_n \text{ are all open intervals} \right\}$$

This set is nonempty since $\mathbb{R}$ can be covered by open intervals (eg. $\{(n, n+2)\}_{n\in\mathbb{Q}}$) and therefore so can every subset. Obviously we have that $m^*(\varnothing) = 0$ since we can take arbitrarily small arbitrary intervals. Notice that for every $\varepsilon > 0$, we showed in the previous section there exists open intervals $\{E_n\}_{n=1}^{\infty}$ such that $\mathbb{Q} \subseteq \bigcup_{n=1}^{\infty} E_n$ and $\sum_{n=1}^{\infty} |E_n| = \varepsilon$. Thus we have that $m^*(\mathbb{Q}) \leq \varepsilon$ for every $\varepsilon > 0$ and so $m^*(\mathbb{Q}) = 0$ as well.
Notice that we can also take a finite set of I_ns, as we can add infinitely many I_ns of arbitrarily small width (eg. add I_ns of length $\frac{\varepsilon}{2^n}$) and this will add ε to the sum, and so the infimum remains the same. Thus we have that

$$m^*(E) = \inf\left\{ \sum_{n=1}^{\infty} |I_n| \;\middle|\; E \subseteq \bigcup_{j\in J} I_j \text{ where } I_n \text{ are all open intervals and } J \text{ is countable} \right\}$$

Now, does the outer measure satisfy the conditions m1 through m4?

(**1**) m^* is indeed a function from $\mathcal{P}(\mathbb{R})$ to $[0, \infty]$,

(**2**) $m^*(I) = |I|$ for all intervals I (proven below).

(**3**) $m^*(E + \alpha) = m^*(E)$ as there is a width-preserving bijection between collections of intervals covering E and $E + \alpha$ (in particular $I \mapsto I + \alpha$).

But m^* is not σ-additive, and we can only ensure σ-subadditivity:

2.1.2 Theorem

Let $\{E_n\}_{n=1}^{\infty}$ be a sequence of (not necessarily disjoint) subsets of $\mathbb{R}$, then

$$m^*\left(\bigcup_{n=1}^{\infty} E_n \right) \leq \sum_{n=1}^{\infty} m^*(E_n)$$

Proof: for every E_n, let us take a cover for E_n of open intervals $\left\{ I_k^{(n)} \right\}_{k=1}^{\infty}$ such that $m^*(E_n) \geq \sum_{k=1}^{\infty} \left| I_k^{(n)} \right| + \frac{\varepsilon}{2^n}$. Then we have that

$$\sum_{n=1}^{\infty} m^*(E_n) \geq \sum_{n=1}^{\infty} \left(\sum_{k=1}^{\infty} \left| I_k^{(n)} \right| + \frac{\varepsilon}{2^n} \right) = \sum_{n,k=1}^{\infty} \left| I_k^{(n)} \right| + \varepsilon$$

Since $\left\{I_k^{(n)}\right\}_{n,k=1}^{\infty}$ is a cover of $\bigcup_{n=1}^{\infty} E_n$, we have that

$$\sum_{n=1}^{\infty} m^*(E_n) \geq \sum_{n,k=1}^{\infty} \left|I_k^{(n)}\right| + \varepsilon \geq m^*\left(\bigcup_{n=1}^{\infty} E_n\right) + \varepsilon$$

Since $\varepsilon > 0$ is arbitrary, we get the desired inequality (by taking $\varepsilon \to 0$, the above inequality is preserved). ■

Of course this does not prove that m^* is not σ-additive, we have simply proven a weaker condition.
This helps us show that $m^*(I) = |I|$ for all intervals I. For open intervals this is trivial, and we have that $m^*(E \cup \{a\}) \leq m^*(E) + m^*(\{a\}) = m^*(E)$ (the measure of a singleton is obviously zero) and $m^*(E \cup \{a\}) \geq m^*(E)$ by monotonicity. All intervals are obtained by adding a finite number of points to an open interval, so for example $m^*([a,b]) = m^*\big((a,b) \cup \{a,b\}\big) = m^*((a,b)) = b - a$ as required.

2.1.3 Proposition (The Vitali Set)

There exists no function which satisfies properties m1 through m4.

Proof: let us define an equivalence relation on $\mathbb{R}$ as follows: $x \sim y \iff x - y \in \mathbb{Q}$. Notice that the equivalence classes of this relation are of the form $x + \mathbb{Q}$ for some $x \in \mathbb{R}$ and thus they are all countable and dense. For every equivalence class choose a single representative in $[0,1]$ to form the set $E \subseteq [0,1]$. Notice then that

$$\mathbb{R} = \bigcupdot_{q \in \mathbb{Q}} (E + q) = E + \mathbb{Q} = \{x + q \mid x \in E, q \in \mathbb{Q}\}$$

this union is disjoint since if $x \in E + q \cap E + p$ then $x = x_1 + q$ and $x = x_2 + p$ so $x \sim x_1 \sim x_2$ but we put only a single representative of each equivalence class into E. For every $y \in [0,1]$, there exists an $x \in E$ such that $x \sim y$ so $x - y \in \mathbb{Q} \cap [-1,1]$, and so if $y \in E + q$ for some $q \in [-1,1]$. Thus

$$[0,1] \subseteq \bigcupdot_{q \in \mathbb{Q} \cap [-1,1]} E + q = E + \mathbb{Q} \cap [-1,1] \subseteq [-1,2]$$

If m is a function which satisfies the four properties above, it is also monotonic (which can be derived from σ-additivity: if $A \subseteq B$ then $m(B) = m(A) + m(B \setminus A) \geq m(A)$), and so

$$m([0,1]) = 1 \leq \sum_{q \in \mathbb{Q} \cap [-1,1]} m(E) \leq 3$$

Since $\mathbb{Q} \cap [-1,1]$ is infinite, we must have that $m(E) = 0$ as otherwise the sum is infinite. But then the sum is 0 and not greater than 1, which is a contradiction. So the *Vitali set* $\bigcup_{q \in \mathbb{Q} \cap [-1,1]} E + q$ cannot be measurable. ■

So we must weaken one of the conditions. m^* is already an example of a function which satisfies m1 through m3 so they are not contradictory (also notice that this means m^* cannot be σ-additive). But σ-additivity is extremely important to Lebesgue's theory of integration, so we will instead weaken m1 to be

$$\textbf{(m1)} \quad m \text{ is a function } \mathcal{L}(\mathbb{R}) \longrightarrow [0,\infty]$$

where $\mathcal{L}(\mathbb{R})$ is the set of all Lebesgue-measurable sets of $\mathbb{R}$:

2.1.4 Definition (Carathéodory)

A set $E \subseteq \mathbb{R}$ is called **Lebesgue measurable** (or just *measurable*) if for every $A \subseteq \mathbb{R}$,

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

Notice that by subadditivity $m^*(A) \leq m^*(A \cap E) + m^*(A \cap E^c)$, to show that E is Lebesgue measurable it is sufficient to show the other direction ($\geq$). This implies that all sets E with zero outer measure are Lebesgue measurable: $m^*(A \cap E) + m^*(A \cap E^c) = m^*(A \cap E^c) \leq m^*(A)$ as required ($m^*(A \cap E) \leq m^*(E) = 0$). Further notice that E is measurable if and only if E^c is by symmetry of the definition.

2.1.5 Theorem

Every open interval of the form (a,∞) is measurable.

Proof: let A be a subset of $\mathbb{R}$. Let us define

$$A^+ := A\cap(a,\infty), \qquad A^- := A\cap(a,\infty)^c = A\cap(-\infty,a]$$

We must show that $m^*(A^+)+m^*(A^-)\leq m^*(A)$. Let $\varepsilon>0$ and let $\{I_n\}_{n=1}^\infty$ be a cover of A consisting of open intervals where $\sum_{n=1}^\infty |I_n| \leq m^*(A)+\varepsilon$. Then for every n, define $I_n^+ = I_n\cap(a,\infty)$ and $I_n^- = I_n\cap(a,\infty)^c$, so $\{I_n^+\}$ is a cover for A^+ and $\{I_n^-\}$ is a cover for A^-. And since these are intervals we have $|I_n| = |I_n^+|+|I_n^-|$. So we have

$$m^*(A)\geq \sum_{n=1}^\infty |I_n| - \varepsilon = \sum_{n=1}^\infty |I_n^+| + \sum_{n=1}^\infty |I_n^-| + \varepsilon \geq m^*(A^+)+m^*(A^-)+\varepsilon$$

And since $\varepsilon>0$ is arbitrary, we get $m^*(A)\geq m^*(A^+)+m^*(A^-)$ as required. ■

2.1.6 Proposition

The set of all measurable sets is invariant under movement, meaning if E is measurable so too is $E+\alpha$.

Proof: let $A\subseteq\mathbb{R}$ then notice that $x-\alpha\in A\cap(E+\alpha)-\alpha \iff x\in A\cap(E+\alpha) \iff x\in A, E+\alpha \iff x-\alpha\in A-\alpha, E$. So $A\cap(E+\alpha)-\alpha = (A-\alpha)\cap E$, and so we get

$$\begin{aligned} m^*(A\cap(E+\alpha)) + m^*(A\cap(E+\alpha)^c) &= m^*(A\cap(E+\alpha)-\alpha) + m^*(A\cap(E+\alpha)^c-\alpha) \\ &= m^*((A-\alpha)\cap E) + m^*((A-\alpha)\cap E^c) = m^*(A-\alpha) = m^*(A) \end{aligned}$$

as required. ■

2.1.7 Lemma

The finite union of measurable sets is measurable. And if $\{E_k\}_{k=1}^n$ are disjoint and measurable, then

$$m^*\left(\dot{\bigcup_{k=1}^n} E_k\right) = \sum_{k=1}^n m^*(E_k)$$

Proof: it is sufficient to prove this for two measurable sets and then to proceed inductively for arbitrary n. So suppose E,F are measurable, we must show that $E\cup F$ is. Let $A\subseteq\mathbb{R}$, then

$$\begin{aligned} &m^*(A\cap(E\cup F)) + m^*(A\cap(E\cup F)^c) = m^*((A\cap E)\cup(A\cap F)) + m^*(A\cap E^c\cap F^c) \\ &\quad= m^*((A\cap E)\,\dot\cup\,(A\cap F\cap E^c)) + m^*(A\cap E^c\cap F^c) \leq m^*(A\cap E) + m^*(A\cap E^c\cap F) + m^*(A\cap E^c\cap F) \\ &\quad= m^*(A\cap E) + m^*(A\cap E^c) = m^*(A) \end{aligned}$$

The final equality is due to F being measurable, and so it satisfies Carathéodory's definition for $A\cap E^c$.

And we similarly prove additivity by induction on n. For two disjoint measurable sets E,F, we have by Carathéodory on E:

$$m^*(E\,\dot\cup\,F) = m^*((E\,\dot\cup\,F)\cap E) + m^*((E\,\dot\cup\,F)\cap E^c) = m^*(E)+m^*(F)$$

■

Since the complement of a set is measurable, and the empty set is measurable, this proposition tells us that $\mathcal{L}(\mathbb{R})$, the set of all Lebesgue measurable sets, is an algebra of sets (non-empty, closed under complements and unions/intersections). Intersections are of course measurable since $\bigcap_{k=1}^n E_k = \left(\bigcap_{k=1}^n E_k^c\right)^c$, and so are differences since $E\setminus F = E\cap F^c$. Notice that since $m^*(E) = m^*(E\cap F)+m^*(E\cap F^c)$, we have

$$m^*(E\setminus F) = m^*(E\cap F^c) = m^*(E) - m^*(E\cap F)$$

But stronger than this, $\mathcal{L}(\mathbb{R})$ is in fact a σ-algebra (to be defined later).

2.1.8 Lemma

Suppose $E_1, \dots, E_n$ are measurable and disjoint, then for every $A \subseteq \mathbb{R}$,

$$m^*\left(A \cap \dot{\bigcup_{k=1}^{n}} E_k\right) = \sum_{k=1}^{n} m^*(A \cap E_k)$$

Proof: We prove this by induction on n; for $n = 1$ this is trivial. If this holds for n, then using the inductive assumption for $E_1, \dots, E_{n-1}, E_n \mathbin{\dot\cup} E_{n+1}$

$$m^*\left(A \cap \dot{\bigcup_{k=1}^{n+1}} E_k\right) = \sum_{k=1}^{n-1} m^*(A \cap E_k) + m^*(A \cap (E_n \mathbin{\dot\cup} E_{n+1}))$$

Now,

$$\begin{aligned} m^*(A \cap (E_n \mathbin{\dot\cup} E_{n+1})) = m^*(A \cap (E_n \mathbin{\dot\cup} E_{n+1}) \cap E_n) = m^*(A \cap (E_n \mathbin{\dot\cup} E_{n+1}) \cap E_n^c) \\ = m^*(A \cap E_n) + m^*(A \cap E_{n+1}) \end{aligned}$$

and so we get that indeed

$$m^*\left(A \cap \dot{\bigcup_{k=1}^{n+1}} E_k\right) = \sum_{k=1}^{n+1} m^*(A \cap E_k)$$

■

2.1.9 Theorem

Let $\{E_n\}_{n=1}^{\infty}$ be Lebesgue measurable, then $\bigcup_{n=1}^{\infty} E_n$ is also Lebesgue measurable (equivalently $\mathcal{L}(\mathbb{R})$ is a σ-algebra). Furthermore if the sets are disjoint, then we have σ-additivity:

$$m^*\left(\dot{\bigcup_{n=1}^{\infty}} E_n\right) = \sum_{n=1}^{\infty} m^*(E_n)$$

Proof: we can assume from the outset that $\{E_n\}$ are disjoint as we can define $F_n := E_n \setminus \bigcup_{k=1}^{n-1} E_k$. This is measurable by the previous lemmas, and $\bigcup_{n=1}^{\infty} E_n = \dot{\bigcup}_{n=1}^{\infty} F_n$. Now define

$$E := \dot{\bigcup_{n=1}^{\infty}} E_n, \qquad G_n := \dot{\bigcup_{k=1}^{n}} E_k$$

Let $A \subseteq \mathbb{R}$, and by G_n's measurability we have that

$$m^*(A) = m^*(A \cap G_n) + m^*(A \cap G_n^c) \geq m^*(A \cap G_n) + m^*(A \cap E^c)$$

By the above lemma this means $m^*(A) \geq \sum_{k=1}^{n} m^*(A \cap E_k) + m^*(A \cap E^c)$, and this inequality is preserved by taking the limit $n \to \infty$ so

$$m^*(A) \geq \sum_{k=1}^{\infty} m^*(A \cap E_k) + m^*(A \cap E^c) \tag{1}$$

By subadditivity, we have that $m^*(A \cap E) = m^*(\dot{\bigcup}_{k=1}^{\infty}(A \cap E_k)) \leq \sum_{k=1}^{\infty} m^*(A \cap E_k)$, so

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

so E is indeed measurable, as required. If we let $A = E = \dot{\bigcup}_{n=1}^{\infty} E_n$, then we get that $A \cap E_k = E_k$ and so by (**1**) we get that

$$m^*(E) \geq \sum_{k=1}^{\infty} m^*(E_k)$$

and $\leq$ is given by subadditivity so we have equality. ■

Notice that the Lebesgue outer measure defines a pseudometric on $\mathcal{P}(\mathbb{R})$ by $d(A, B) := m^*(A \mathbin{\triangle} B)$.

2.1.10 Theorem

$E \subseteq \mathbb{R}$ is Lebesgue measurable if and only if forall $\varepsilon > 0$ there exists a countable union of open intervals $\bigcup_{n=1}^\infty I_n$ such that $d(\bigcup_{n=1}^\infty I_n, E) < \varepsilon$.

Proof: suppose E is measurable, then for every $\varepsilon > 0$ there exists a cover $E \subseteq \bigcup I_n$ of open intervals such that $m^*(E) + \varepsilon > m^*(\bigcup I_n)$. Now

$$d\Big(\bigcup I_n, E\Big) = m^*\Big(\bigcup I_n \setminus E\Big) = m^*\Big(\bigcup I_n\Big) - m^*(E) < \varepsilon$$

since all the sets in play are measurable, as required. Now suppose that the condition holds, we must prove that E is measurable. We can assume without loss of generality that $E \subseteq \bigcup I_n$ (as we can cover the symmetric difference with open intervals). Now let us take a sequence of such covers U_k such that $m^*(U_k \setminus E) < \frac{1}{k}$, and define $U = \bigcap U_k$. And $m^*(U \setminus E) \leq m^*(U_k \setminus E) < \frac{1}{k}$ so $m^*(U \setminus E) = 0$, and this means that $U \setminus E$ is measurable. Since U_k are all measurable, so is U and since $E = U \setminus (U \setminus E)$, E is measurable. ■

In fact what we have shown is that every measurable set E is of the form $G \setminus N$, where G is a countable intersection of open sets (a G_δ set), and N is a set of zero measure. This theorem implies that every open interval is Lebesgue measurable, as just define $I_n = I$. Recall that $\mathbb{R}$ is a second-countable topological space, meaning it has a countable basis, namely the basis $\{(p, q) \mid p < q \in \mathbb{Q}\}$. This means that every open set in $\mathbb{R}$ is the countable union of open intervals, and therefore every open set is contained in $\mathcal{L}(\mathbb{R})$. And so closed sets, which are complements of open sets, are also measurable. And so too are G_δ sets as the countable intersection of measurable sets, and so sets of the form $G \setminus N$ where G is G_δ and $m^*(N) = 0$ are measurable. So we have proven

2.1.11 Proposition

E is measurable if and only if there exists a G_δ set G and a zero-measure set N such that $E = G \setminus N$.

2.2 General Measure Spaces, Briefly

2.2.1 Definition

Let X be a set, then a non-empty collection $\Sigma \subseteq \mathcal{P}(X)$ is called a ***σ*-algebra** if it satisfies

(1) $S \in \Sigma \iff S^c \in \Sigma$

(2) if $\{S_n\}_{n=1}^\infty \subseteq \Sigma$ then $\bigcup_{n=1}^\infty S_n \in \Sigma$

By (1) and (2) we get that σ-algebras are also closed under intersections, and so $\varnothing = S \cap S^c \in \Sigma$ and $X = \varnothing^c \in \Sigma$.

If X is a set and Σ a σ-algebra over X, then (X, Σ) is called a **measurable space**. And a **measure space** is a triplet (X, Σ, μ) where μ is a σ-algebra over X, and μ is a **measure** over Σ:

(1) $\mu\colon \Sigma \longrightarrow [0, \infty]$,

(2) μ is σ-additive: if $\{S_n\}_{n=1}^\infty \subseteq \Sigma$ are disjoint $\mu(\dot\bigcup_{n=1}^\infty S_n) = \sum_{n=1}^\infty \mu(S_n)$

(3) $\mu(\varnothing) = 0$.

2.2.2 Example

Let m be the **Lebesgue measure** over $\mathbb{R}$, the restriction of m^* to the collection of Lebesgue measurable sets: $m := m^*\big|_{\mathcal{L}(\mathbb{R})}$. Then $(\mathbb{R}, \mathcal{L}(\mathbb{R}), m)$ is a measure space.

Measures are obviously monotonic: if $E, F \in \Sigma$ and $E \subseteq F$ then $\mu(F) = \mu((F \setminus E) \cup E) = \mu(F \setminus E) + \mu(E) \geq \mu(E)$ (since recall that σ-algebras are closed under differences, as they are just the intersections of the complement). And measures are subadditive:

2.2.3 Theorem

Let (X, Σ, μ) be a measure space, then for any $\{A_n\}_{n=1}^{\infty} \subseteq \Sigma$ (not necessarily disjoint):

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

Proof: define $B_n := A_n \setminus \bigcup_{k=1}^{n-1} A_k$, these are all disjoint and $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. And since $B_n \subseteq A_n$, $\mu(B_n) \leq \mu(A_n)$. Thus

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\dot{\bigcup_{n=1}^{\infty}} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n) \qquad \blacksquare$$

One of the most important properties of measure spaces is *continuity from below* and *continuity from above*:

2.2.4 Theorem (Continuity of Measures)

Let (X, Σ, μ) be a measure space, then

(1) **Continuity from below**: if $E_1 \subseteq E_2 \subseteq \cdots$ is an increasing sequence in Σ, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n)$$

(2) **Continuity from above**: if $E_1 \supseteq E_2 \supseteq \cdots$ is a decreasing sequence in Σ such that $\mu(E_1) < \infty$, then

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n)$$

Proof: notice that the limits in both cases exist since the sequences are monotonic.

(1) Let $E := \bigcup_{n=1}^{\infty} E_n$, then $E \supseteq E_n$ so $\mu(E) \geq \mu(E_n)$ and so $\mu(E) \geq \lim_{n\to\infty} \mu(E_n)$. Define $F_n := E_n \setminus E_{n-1}$, so that F_n are disjoint and $\dot{\bigcup}_{k=1}^{n} F_k = E_n$ for all n, and so $\dot{\bigcup}_{k=1}^{\infty} F_k = E$, so

$$\mu(E) = \sum_{k=1}^{\infty} \mu(F_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(F_k) = \lim_{n\to\infty} \mu\left(\dot{\bigcup_{k=1}^{n}} F_k\right) = \lim_{n\to\infty} \mu(E_n)$$

as required.

(2) Define $E := \bigcap_{n=1}^{\infty} E_n$, $F_n := E_1 \setminus E_n$, $F := \bigcup_{n=1}^{\infty} F_n$. Thus $E = E_1 \setminus F$, and so by above

$$\mu(E) = \mu(E_1) - \mu(F) = \mu(E_1) - \lim_{n\to\infty} \mu(F_n) = \mu(E_1) - \lim_{n\to\infty} \big(\mu(E_1) - \mu(E_n)\big) = \lim_{n\to\infty} \mu(E_n)$$

$\blacksquare$

2.3 Measurable Functions

2.3.1 Theorem

Let (X, Σ) be a measureable space, and $f\colon X \longrightarrow \overline{\mathbb{R}}$ $(= \mathbb{R} \cup \{\pm\infty\})$ be an extended real function. Then the following are equivalent:

(1) for every $\alpha \in \mathbb{R}$, $\{x \in X \mid f(x) > \alpha\} \in \Sigma$,

(2) for every $\alpha \in \mathbb{R}$, $\{x \in X \mid f(x) \geq \alpha\} \in \Sigma$,

(3) for every $\alpha \in \mathbb{R}$, $\{x \in X \mid f(x) < \alpha\} \in \Sigma$,

(4) for every $\alpha \in \mathbb{R}$, $\{x \in X \mid f(x) \leq \alpha\} \in \Sigma$.

Proof: notice that (1) $\iff$ (4) and (2) $\iff$ (3) are trivial since $\{x \in X \mid f(x) > \alpha\}^c = \{x \in X \mid f(x) \leq \alpha\}$ and similar, and S is measurable if and only if S^c is. Now we prove (1) $\iff$ (2), notice that

$$\{x \in X \mid f(x) \geq \alpha\} = \bigcap_{n=1}^{\infty} \left\{ x \in X \;\middle|\; f(x) > \alpha - \frac{1}{n} \right\}$$

so if (1) then the right side is the countable intersection of measurable sets and is therefore measurable, so (1) $\implies$ (2). And

$$\{x \in X \mid f(x) > \alpha\} = \bigcup_{n=1}^{\infty} \left\{ x \in X \;\middle|\; f(x) \geq \alpha + \frac{1}{n} \right\}$$

so anlogously (2) $\implies$ (1). Similarly for (3) $\iff$ (4). ■

2.3.2 Definition

Let (X, Σ) be a measurable space and $f\colon X \longrightarrow \overline{\mathbb{R}}$ an extended real function. If any of the above equivalent conditions hold, then f is a **Σ-measurable function** (if Σ is understood, just measurable).

Notice that constant functions $f\colon x \mapsto c$ are measurable: for $\alpha \geq c$, $\{x \in X \mid f(x) > \alpha\} = \varnothing$ which is measurable. And for $\alpha < c$, $\{x \in X \mid f(x) > \alpha\} = X$ which is measurable.

2.3.3 Corollary

Let f be an extended real function, then

(1) if f is measurable, for every $x_0 \in \overline{\mathbb{R}}$, $f^{-1}(x_0) \in \Sigma$,

(2) if f is measurable, for every interval $I \subseteq \overline{\mathbb{R}}$, $f^{-1}(I) \in \Sigma$,

(3) if $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ is continuous, then f is Lebesgue measurable (ie. measurable in $(\mathbb{R}, \mathcal{L}(\mathbb{R}))$).

Proof:

(1) Notice that

$$f^{-1}(x_0) = \{x \in X \mid f(x) \geq x_0\} \cap \{x \in X \mid f(x) \leq x_0\}$$

which is the intersection of measurable sets.

(2) For an interval of the form $I = [a, b)$,

$$f^{-1}(I) = \{x \in X \mid f(x) \geq a\} \cap \{x \in X \mid f(x) < b\}$$

which is the intersection of measurable sets. All other intervals are done similarly.

(3) Let $\alpha \in \mathbb{R}$, then $E = (\alpha, \infty)$ is open and so $f^{-1}(E)$ is open and thus Lebesgue measurable. And $f^{-1}(E) = \{x \in \mathbb{R} \mid f(x) > \alpha\} \in \mathcal{L}(\mathbb{R})$ as required. ■

2.3.4 Theorem

Let (X, Σ) be a measurable space, $f, g\colon X \longrightarrow \mathbb{R}$ and $c \in \mathbb{R}$. Then $f \pm g, c \cdot f, f \cdot g$ are measurable.

(1) Let $\alpha \in \mathbb{R}$, then $f(x)+g(x)<\alpha$ if and only if $f(x)<\alpha-g(x)$ if and only if there exists a $\beta \in \mathbb{Q}$ such that $f(x)<\beta<\alpha-g(x)$. So

$$\{x \in X \mid f(x)+g(x)<\alpha\} = \bigcup_{\beta \in \mathbb{Q}} \big(\{x \in X \mid f(x)<\beta\} \cap \{x \in X \mid g(x)<\alpha-\beta\}\big)$$

which is the countable union of the intersection of measurable sets, which is measurable. For $f-g$, this is due to (2) which states that $-g$ is measurable.

(2) Notice that for $\alpha \in \mathbb{R}$, if $c>0$ then $\{x \in X \mid c \cdot f(x)<\alpha\} = \left\{x \in X \;\middle|\; f(x)<\frac{\alpha}{c}\right\}$ which is measurable. If $c<0$ then $\{x \in X \mid c \cdot f(x)<\alpha\} = \left\{x \in X \;\middle|\; f(x)>\frac{\alpha}{c}\right\}$. If $c=0$, then $c \cdot f=0$, and constant functions are measurable always. This combined with (1) show us that all linear combinations of measurable functions are measurable.

(3) First we will prove that f^2 is measurable. Notice that for $\alpha \geq 0$, the set $\left\{x \in X \;\middle|\; f(x)^2 \geq \alpha\right\} = \{x \in X \mid f(x) \geq \sqrt{\alpha}\}$ is measurable. And for $\alpha<0$, $\{x \in X \mid f(x) \geq \alpha\} = X$ is measurable. And

$$f \cdot g = \frac{1}{4}\big((f+g)^2-(f-g)^2\big)$$

We just showed that $f \pm g$ are measurable, and so too are $(f \pm g)^2$ and any of their linear combinations, so $f \cdot g$ is measurable. ■

Notice that for $f, g\colon X \longrightarrow \overline{\mathbb{R}}$, $f(x)+g(x)$ is undefined when $f(x)=\infty$ and $g(x)=-\infty$ or vice versa. We generally take that $\pm\infty \cdot 0 := 0$, but $f(x) \cdot g(x)$ is not necessarily defined when $f(x)=\pm\infty$ and $g(x)=0$ or vice versa. But if we define $f \pm g$ and $f \cdot g$ to be any arbitrary constant (eg. 0) at these problematic points, we can similarly show that $f \pm g$ and $f \cdot g$ are measurable.

2.3.5 Theorem

Let (X, Σ) be a measurable space and $\{f_n\}_{n=1}^{\infty}$ a sequence of extended real measurable functions. Then $\sup_n f_n(x), \inf_n f_n(x), \limsup f_n(x), \liminf f_n(x)$, and $\lim f_n(x)$ if it exists are measurable.

Proof: first we prove it for $f(x)=\sup_n f_n(x)$. Let $\alpha \in \mathbb{R}$, then $f(x) \leq \alpha$ if and only if $f_n(x) \leq \alpha$ for all α, so

$$\{x \in X \mid f(x) \leq \alpha\} = \bigcap_{n=1}^{\infty} \{x \in X \mid f_n(x) \leq \alpha\}$$

which is the countable intersection of measurable sets. Similar for $g(x)=\inf_n f_n(x)$.

Now define $h(x)=\limsup_n f_n(x)$, then $h(x)=\inf_k \sup_{n \geq k} f_n(x)$. Let us define $g_k(x)=\sup_{n \geq k} f_n(x)$, so $h(x)=\inf_k g_k(x)$. Since g_k is the supremum of a sequence of measurable functions, it is measurable. And so h is the infimum of measurable functions, meaning it too is measurable. We use an analogous proof for the case $\liminf_n f_n(x)=\sup_k \inf_{n \geq k} f_n(x)$. If $\lim_n f_n(x)$ exists, then it is just equal to $\limsup_n f_n(x)=\liminf_n f_n(x)$. ■

2.4 The Lebesgue Integral

2.4.1 Definition

Let (X, Σ, μ) be a measure space, and $A \in \Sigma$ be measurable. Then we define the **Lebesgue integral** of χ_A to be

$$\int_X \chi_A \, d\mu = \mu(A)$$

And if the measurable function $\varphi\colon X \longrightarrow [0, \infty)$ is **simple**, meaning it takes on only a finite number of values $a_1, \ldots, a_n$, then define $A_i = \varphi^{-1}\{a_i\}$ such that $\varphi = \sum_{i=1}^{n} a_i \chi_{A_i}$ where A_i are disjoint, define its Lebesgue integral to be

$$\int_X \varphi \, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$

If f is a function whose Lebesgue integral is defined, and $E \in \Sigma$ is measurable, then f's Lebesgue integral over E is defined to be

$$\int_E f\,d\mu = \int_X f \cdot \chi_E\,d\mu$$

2.4.2 Proposition

For φ, ψ simple,

(1) $\int_X (\varphi + \psi) = \int_X \varphi + \int_X \psi$,

(2) if $c \geq 0$ then $\int_X c\varphi = c\int_X \varphi$,

(3) if E and F are disjoint measurable sets then $\int_{E\cup F} \varphi = \int_E \varphi + \int_F \varphi$,

(4) if $\varphi \leq \psi$ (pointwise, meaning $\varphi(x) \leq \psi(x)$ for every $x \in X$), then $0 \leq \int_X \varphi \leq \int_X \psi$,

(5) if $m \leq \varphi \leq M$ and E is measurable, then $m\mu(E) \leq \int_E \varphi \leq M\mu(E)$.

Proof:

(1) Firstly, notice that if $\varphi = \sum_{i=1}^n a_i\chi_{A_i}$ where A_i are not necessarily disjoint, still $\int_X \varphi = \sum_{i=1}^n a_i\mu(A_i)$. This is since $a\chi_A + b\chi_B = a\chi_{A\setminus B} + b\chi_{B\setminus A} + (a+b)\chi_{A\cap B}$ and so the integral is $a\mu(A\setminus B) + b\mu(B\setminus A) + (a+b)\mu(A\cap B) = a\mu(A) + b\mu(B)$. Then (1) is a direct result of this: write φ and ψ as linear combinations of characteristic functions.

(2) This is direct from definition: $c\varphi = \sum ca_i\chi_{A_i}$ so its integral is $c\sum a_i\mu(A_i)$.

(3) This is since $\chi_{E\cup F} = \chi_E + \chi_F$ and so

$$\int_{E\cup F} \varphi = \int(\varphi\chi_E + \varphi\chi_F) = \int \varphi\chi_E + \int \varphi\chi_F = \int_E \varphi + \int_F \varphi$$

(4) Notice that integrals are non-negative by definition. And then $\psi = \varphi + (\psi - \varphi)$, and $\psi - \varphi$ is still simple, so $\int \psi = \int \varphi + \int(\psi - \varphi) \geq \int \varphi$ as required.

(5) This is due to $\int_E c = \int_X c\chi_E = c\mu(E)$. ■

2.4.3 Definition

Let $f\colon X \longrightarrow [0,\infty]$ be measurable, then define its **Lebesgue integral**

$$\int_X f\,d\mu = \sup_{\substack{0\leq\varphi\leq f\\ \varphi \text{ simple}}} \int_X \varphi\,d\mu$$

Notice that we define the integral of a nonnegative function using the supremum of the set of simple functions bound by f. Why not use the infimum of the simple functions which bound f? ie, why not have

$$\int_X f\,d\mu \stackrel{?}{=} \inf_{\substack{\varphi\geq f\\ \varphi \text{ simple}}} \int_X \varphi\,d\mu$$

Take for example $f(x) = e^{-x^2}$ which has a finite (Riemann) integral, we would like its Lebesgue integral to be finite as well. But every simple function φ which bounds it must have an infinite integral: it must have a minimum value m, and this minimum value m cannot be 0 since f is never zero. And so $\int \varphi \geq \int m = \infty$ for every φ which bounds f, so by this definition $\int f = \infty$, which is not ideal.

Note that if f itself is simple, then every $\varphi \leq f$ has an integral $\int \varphi \leq \int f$ by the above theorem. And since $f \leq f$, $\int f$ is in the set over which we take the supremum, so this definition of the integral does not contradict the previous definition.

2.4.4 Theorem

For f, g measurable and nonnegative,

(1) if $0 \leq f \leq g$ then $0 \leq \int_X f \leq \int_X g$,

(2) if $E \subseteq F$ are measurable then $\int_E f \leq \int_F f$,

(3) if $f = 0$ almost always on E then $\int_X f = 0$,

(4) if $\mu(E) = 0$ then $\int_E f = 0$,

(5) if $m \leq f \leq M$ on E then $m\mu(E) \leq \int_X f \leq M\mu(E)$.

Proof:

(1) This is since simple functions $\varphi \leq f$ are also bound by g.

(2) This is since $\int_E f = \int_X f \cdot \chi_E$ and $f \cdot \chi_E \leq f \cdot \chi_F$, so this is a direct result of (1).

(3) Let $E_1 = f^{-1}\{0\}$ and $E_2 = E \setminus E_1$. Then for $\varphi \leq f$, φ must be zero on E_1 and so $\varphi = \varepsilon\chi_{E_2}$. So $\int \varphi = \varepsilon\mu(E_2) = 0$ since f is zero almost everywhere, meaning $\mu(E_2) = 0$.

(4) This is since $f \cdot \chi_E$ is zero almost everywhere.

(5) This is since $\int_E c = c\mu(E)$ still. ■

2.4.5 Theorem (The Monotone Convergence Theorem)

Let $f_n\colon X \longrightarrow [0, \infty]$ be a monotonically increasing sequence of measurable functions (meaning $0 \leq f_1(x) \leq f_2(x) \leq \cdots$ for every $x \in X$), define $f = \lim_{n\to\infty} f_n(x)$. Then

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

meaning we can swap the integral and limit, if the sequence is increasing.

At this point in time, the lecturer drew a tree on the board to make some point about Lebesgue basing his theory of integration off of this theorem. It was a nice tree, I didn't really listen too much to his explanation of why he drew the tree, but the tree itself was nice.

Proof: Notice that $f_n \leq f$ since f_n is increasing, so $\int_X f_n \leq \int_X f$, and so $\lim \int_X f_n \leq \int_X f$. So all that remains to show is the other direction: $\lim \int_X f_n \geq \int_X f$. By definition, this is equivalent to

$$\sup_{0 \leq \varphi \leq f} \int_X \varphi \leq \lim_{n\to\infty} \int_X f_n$$

Let $0 \leq \varphi \leq f$, set $\alpha \in (0, 1)$, and define $E_n = \{x \in X \mid f_n(x) \geq \alpha\varphi(x)\}$. Then $\{E_n\}$ is increasing (since f_n is), and since if $x \in X$ then $f(x) > \alpha\varphi(x)$ so there is an n such that $f_n(x) \geq \alpha\varphi(x)$, since $\alpha < 1$. Meaning there is an n such that $x \in E_n$, so $\bigcup E_n = X$.

Now, suppose $\varphi = \sum_i a_i \chi_{A_i}$, then since $E_n \cap A_i$ is increasing for every n, $\lim_n \mu(E_n \cap A_i) = \mu(A_i)$ since the union of E_n is X. So

$$\lim \int_{E_n} \varphi = \lim_n \sum_i a_i \chi_{A_i} \chi_{E_n} = \lim_n \sum_i a_i \mu(A_i \cap E_n) = \sum_i a_i \cdot \lim_n \mu(E_n \cap A_i) = \sum_i a_i \cdot \mu(A_i) = \int_X \varphi$$

So we have that $f_n(x) \geq \alpha\varphi(x)$ on E_n, so $\int_X f_n \geq \int_{E_n} f_n \geq \int_{E_n} \alpha\varphi = \alpha \int_{E_n} \varphi$. Taking the limit gives us $\lim_n \int_X f_n \geq \alpha \int_X \varphi$. Since this is true for all $\alpha < 1$, it is true for $\alpha = 1$, so $\lim_n \int_X f_n \geq \int_X \varphi$, so taking the supremum over all $\varphi \leq f$ gives $\lim_n \int_X f_n \geq \int_X f$ as required. ■

2.4.6 Example

Let $f_n = \chi_{(n,\infty)}$, then f_n is monotonically decreasing to 0. But $\int f_n = \infty$ which does not converge to $\int 0 = 0$. So the monotone convergence theorem does not hold in the case that the sequence is monotonically decreasing.

2.4.7 Theorem (Fatou's Lemma)

Let $f_n\colon X \longrightarrow [0,\infty]$ be measurable, then

$$\int_X (\liminf f_n)\, d\mu \leq \liminf \int_X f_n$$

Proof: Recall that $\liminf f_n(x) = \lim_n \inf_{k>n} f_k(x)$, which is the limit of an increasing sequence. So by The Monotone Convergence Theorem,

$$\int_X \liminf f_n\, d\mu = \lim_n \int_X \inf_{k>n} f_k\, d\mu = \liminf_n \int_X \inf_{k\geq n} f_k\, d\mu \leq \liminf_n \int_X f_n\, d\mu \qquad \blacksquare$$

2.4.8 Lemma

For every nonnegative measurable function $f\colon X \longrightarrow [0,\infty]$ there exists a sequence of nonnegative simple functions $\{\varphi_n\}_{n=1}^\infty$ which increases pointwise to f. And so by the monotone convergence theorem, $\int_X f\, d\mu = \lim_{n\to\infty} \int_X \varphi_n\, d\mu$.

Proof: define $A_{n,k} = f^{-1}\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]$, then define $\varphi_n(A_{n,k}) = \frac{k}{2^n}$. For $x \in A_{n,k} \cap A_{n,k+1}$ we can set $\varphi_n(x) = \frac{k+1}{2^n}$. We define this for $0 \leq k \leq 2^{2n}$. This obviously converges to f pointwise and is increasing. $\blacksquare$

2.4.9 Theorem

(1) if $f, g\colon X \longrightarrow [0,\infty]$ are measurable then

$$\int_X (f+g)\, d\mu = \int_X f\, d\mu + \int_X g\, d\mu$$

(2) if $f_n\colon X \longrightarrow [0,\infty]$ is a countable sequence of nonnegative measurable functions then

$$\int_X \left(\sum_n f_n\right) d\mu = \sum_n \int_X f_n\, d\mu$$

(3) if $E = \dot\bigcup_n E_n$ are all measurable then

$$\int_E f\, d\mu = \sum_n \int_{E_n} f\, d\mu$$

(4) for every $E \in \Sigma$, define $\nu_f(E) = \int_E f\, d\mu$. Then ν_f is a measure on (X, Σ).

Proof:

(1) By above let φ_n be a sequence of simple functions which increase to f and ψ_n be to g. Then $\varphi_n + \psi_n$ increases to $f+g$ and so by the monotone convergence theorem

$$\int_X (f+g) = \lim_n \int_X (\varphi_n + \psi_n) = \lim_n \int_X \varphi_n + \int_X \psi_n = \int_X f + \int_X g$$

(2) Define $S_n = \sum_{k=1}^n f_k$, then S_n increases to $\sum_n f_n$ so by the monotone convergence theorem and (1),

$$\int_X \sum_n f_n = \lim_n \int_X S_n = \lim_n \sum_{k=1}^n \int_X f_k = \sum_n \int_X f_n$$

(3) Define $f_n = f \cdot \chi_{E_n}$ so that $f \cdot \chi_E = \sum_n f_n$, then this follows from (2).

(4) This is a direct result of (3). ■

2.4.10 Definition

We say that a trait occurs **almost everywhere** (concisely, ae) if the set of all points where it doesn't occur has measure zero.

So for example $\chi_{\mathbb{Q}} = 0$ ae since it is not zero only on $\mathbb{Q}$ which has measure zero.

2.4.11 Theorem

Let $f, g\colon X \longrightarrow [0, \infty]$ be measurable.

(1) if $f(x) = g(x)$ almost everywhere, then $\int_X f = \int_X g$.

(2) $\int_X f = 0$ if and only if $f = 0$ almost everywhere.

(3) if $\int_X f < \infty$ then $f < \infty$ almost everywhere.

(1) Let $E = \{x \in X \mid f(x) \neq g(x)\}$ then $\mu(E) = 0$ and so $\int_E f = 0$ by a previous theorem, and $f(x) = g(x)$ for $x \in E^c$,

$$\int_X f = \int_E f + \int_{E^c} f = \int_{E^c} g = \int_X g$$

(2) If $f = 0$ almost everywhere then by above its integral is zero. Otherwise define $E_n = \left\{x \in X \;\middle|\; f(x) > \frac{1}{n}\right\}$ then

$$0 = \int_X f \geq \int_{E_n} f \geq \int_{E_n} \frac{1}{n} = \frac{1}{n}\mu(E_n)$$

so $\mu(E_n) = 0$. And $f(x) > 0$ if and only if $x \in E_n$ for some n, thus $E = \{x \in X \mid f(x) > 0\} = \bigcup_n E_n$ and so $\mu(E) = \lim_n \mu(E_n) = 0$ by continuity of μ.

(3) Set $E = \{x \in X \mid f(x) = \infty\}$, then

$$\infty > \int_X f \geq \int_E f = \infty\mu(E)$$

So $\mu(E)$ must be zero. ■

2.4.12 Definition

Let $f\colon X \longrightarrow [-\infty, \infty]$ be measurable, then define

$$f^+(x) := \begin{cases} f(x) & f(x) \geq 0 \\ 0 & f(x) < 0 \end{cases}, \qquad f^-(x) := \begin{cases} 0 & f(x) \geq 0 \\ -f(x) & f(x) < 0 \end{cases}$$

so that $f(x) = f^+(x) - f^-(x)$ for all $x \in X$. Notice that $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = \max\{-f(x), 0\}$ so that f^+ and f^- are measurable.

2.4.13 Definition

Let $f\colon X \longrightarrow [-\infty, \infty]$ be measurable, then f is **Lebesgue integrable** if

$$\int_X f^+ \, d\mu < \infty \quad \text{and} \quad \int_X f^- \, d\mu < \infty$$

and if so, define

$$\int_X f \, d\mu := \int_X f^+ \, d\mu - \int_X f^- \, d\mu$$

Notice that $|f| = f^+ + f^-$ so that

$$\int_X |f| \, d\mu := \int_X f^+ \, d\mu + \int_X f^- \, d\mu$$

And in particular f is integrable if and only if $|f|$ is. It is also direct from the triangle inequality that $\left|\int_X f\right| \leq \int_X |f|$.

2.4.14 Theorem

Let $f, g\colon X \longrightarrow \overline{\mathbb{R}}$ be integrable, then

(1) if $h\colon X \longrightarrow \overline{\mathbb{R}}$ is measurable and satisfies $|h(x)| \leq |f(x)|$ almost everywhere, then h is integrable.

(2) f is integrable on every measurable set, and

$$\int_{A \cup B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu$$

(3) $|f(x)| < \infty$ almost everywhere,

(4) if $\mu(E) = 0$ then $\int_E f \, d\mu = 0$,

(5) $c \cdot f$ is integrable and $\int_X (cf) \, d\mu = c \int_X f \, d\mu$,

(6) $f + g$ is integrable and satisfies $\int_X (f+g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$,

(7) if $f(x) \leq g(x)$ almost everywhere then $\int_X f \, d\mu \leq \int_X g \, d\mu$.

For (1): $|h|$ is then integrable and so therefore is h. (5) results in $f+g = (f+g)^+ - (f+g)^- = f^+ - f^- + g^+ - g^-$ and $|f+g| = (f+g)^+ + (f+g)^- = f^+ + f^- + g^+ + g^-$. So $(f+g)^+ = f^+ + g^+$ and $(f+g)^- = f^- + g^-$ are integrable. The rest of the statements are proven utilizing properties of nonnegative measurable functions. ■

2.4.15 Theorem (Dominated Convergence Theorem)

Let $f_n\colon X \longrightarrow \overline{\mathbb{R}}$ be a sequence of measurable functions which converge pointwise to f. Now suppose there exists an integrable function g such that $|f_n(x)| \leq g(x)$ for every $x \in X$ and $n \in \mathbb{N}$. Then

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

Proof: since $|f_n| \leq g$, f_n is integrable by the previous theorem. Furthermore $|f| = \lim_n |f_n| \leq g$, so f is also integrable. Now, we know $-g \leq f_n \leq g$ so $f_n + g \geq 0$ and $g - f_n \geq 0$. By Fatou:

$$\int_X (f+g) = \int_X \lim_n (f_n + g) \leq \liminf_n \int_X (f_n + g) = \liminf_n \int_X f_n + \int_X g$$

So $\int_X f \le \liminf_n \int_X f_n$. And by Fatou on $g - f_n$,

$$\int_X (g-f) = \int_X \lim_n (g-f_n) \le \liminf_n \int_X (g-f_n) = \int_X g + \liminf_n - \int_X f_n = \int_X g - \limsup_n \int_X f_n$$

So we have

$$\limsup_n \int_X f_n \le \int_X f \le \liminf_n \int_X f_n$$

so $\int_X f = \lim_n \int_X f_n$ as required. ■

2.4.16 Theorem (Bounded Convergence Theorem)

Let $E \subseteq X$ be measurable such that $\mu(E) < \infty$, and let us define measurable functions $f_n \colon E \longrightarrow \mathbb{R}$ which converge to a function f and there exists an $M > 0$ such that $|f_n(x)| \le M$ for every $x \in E$. Then

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

Proof: define $g(x) = M\chi_E$ then $|f_n\chi_E| \le g$ and $g(x)$ is integrable, so the result follows from the dominated convergence theorem. ■

2.4.17 Theorem (Equivalence of Riemann and Lebesgue Integrals)

Let $f \colon [a,b] \longrightarrow \mathbb{R}$ be bound. Then f is Riemann-integrable if and only if f is continuous almost everywhere. In such a case, f is Lebesgue integrable and the two integrals agree.

Proof: we will utilize Darboux's theorem: for every sequence of partitions $P_n \to 0$ of $[a,b]$,

$$\overline{S}(f, P_n) \longrightarrow \overline{\int}(f) = \inf_P (\overline{S}(P))$$
$$\underline{S}(f, P_n) \longrightarrow \underline{\int}(f) = \sup_P (\underline{S}(P))$$

Let P_n be a partition of $[a,b]$ into 2^n intervals $I_{n,k}$ of equal length, then every upper/lower Riemann sum is just the integral of a simple function,

$$\overline{S}(f, P_n) = \int_{[a,b]} \varphi_n$$
$$\underline{S}(f, P_n) = \int_{[a,b]} \psi_n$$

where φ_n is the supremum of f on each of the intervals $I_{n,k}$ and ψ_n is the infimum (meaning $\varphi_n(x_0) = \sup_{x\in I_{n,k}} f(x)$ for $x_0 \in I_{n,k}$). So $\psi_n \le \varphi_n$, $\varphi_{n+1} \le \varphi_n$, and $\psi_n \le \psi_{n+1}$. So by the convergence theorems, their integrals converge to the integral of their limits. Now, the limits are

$$\lim_{n\to\infty} \varphi_n(x_0) = \lim_{n\to\infty} \sup_{x\in I_{n,k}} f(x) = \max\left\{ f(x_0), \limsup_{x\to x_0} f(x) \right\}$$
$$\lim_{n\to\infty} \psi_n(x_0) = \lim_{n\to\infty} \inf_{x\in I_{n,k}} f(x) = \min\left\{ f(x_0), \limsup_{x\to x_0} f(x) \right\}$$

we denote the limit of φ_n by f^U and of ψ_n by f^L. Since all these functions are bound, by the bounded convergence theorem

$$\overline{\int_a^b} f = \lim_{n\to\infty} \varphi_n = \int_a^b f^U$$
$$\underline{\int_a^b} f = \lim_{n\to\infty} \psi_n = \int_a^b f^L$$

Thus we get that f is Riemann integrable if and only if the upper and lower integrals are equal, which is if and only if $\int_a^b f^U - f^L = 0$. Since $f^U \ge f^L$, this means that f is Riemann integrable if and only if $f^U = f^L$ almost everywhere. This means that for almost every x_0, $\max\{f(x_0), \limsup_{x\to x_0} f(x)\} = \min\{f(x_0), \liminf_{x\to x_0} f(x)\}$ which happens if and only if $\limsup f(x) = \liminf f(x) = f(x_0)$ since if $f(x_0)$ is the maximum then $\liminf f(x) = f(x_0)$ and so $\limsup f(x) \geq f(x_0)$ so there is equality. This is equivalent to saying that f is continuous at x_0. Thus we have that f is Riemann integrable if and only if f is continuous almost everywhere.

Now if f is Riemann integrable on $[a, b]$ then $f(x) = f^U(x) = \lim_{n\to\infty} \varphi_n(x)$ almost everywhere and denoting the Riemann integral by $R\int$, using the monotone convergence theorem

$$\int f = \lim_{n\to\infty} \int \varphi_n = \lim_{n\to\infty} R\int \varphi_n = \lim_{n\to\infty} \overline{S}(f, P_n) = R\int f$$

since the Riemann and Lebesgue integrals of φ_n coincide as they are both integrals of simple functions of the form $\sum a_i \chi_{I_i}$ where I_i is an interval; both integrals give $\sum a_i |I_i|$. ■

2.5 Product Spaces and Fubini-Tonelli's Theorem

2.5.1 Definition

Suppose $(X, \Sigma_1, \mu), (Y, \Sigma_2, \nu)$ are two measure spaces. We define **product space** $(X \times Y, \Sigma, \mu \times \nu)$ as follows: for a *measurable rectangle* $A \times B \subseteq X \times Y$ where $A \in \Sigma_1$ and $B \in \Sigma_2$ define

$$|A \times B| = \mu(A)\nu(B)$$

Define the outer measure similar to in $\mathbb{R}$:

$$w^*(E) := \inf\left\{ \sum_{n=1}^{\infty} |R_n| \;\middle|\; R_n \text{ are measurable rectangles such that } E \subseteq \bigcup_{n=1}^{\infty} R_n \right\}$$

We say that a set $E \subseteq X \times Y$ is **w-measurable** if for every $A \subseteq X \times Y$, $w^*(A) = w^*(A \cap E) + w^*(A \cap E^c)$. The set of all w-measurable sets forms a σ-algebra Σ (using a similar proof for $\mathbb{R}$) and so taking $w = \mu \times \nu$ to be the restriction of w^* to Σ, we get the product space (X, Σ, w).

2.5.2 Definition

A measure space (X, Σ, μ) is **σ-finite** if there exists a countable subset $\{X_n\}_{n=1}^{\infty} \subseteq \Sigma$ such that $\mu(X_n) < \infty$ and $X = \bigcup_{n=1}^{\infty} X_n$. And it is **complete** if for every $\mu(E) = 0$ if $F \subseteq E$ then $F \in \Sigma$ as well.

We will always assume that for product measures, both spaces are σ-finite (the product space itself is then σ-finite and complete as well). The following are a few traits of product measures which are provable in a manner similar to $\mathbb{R}$:

(1) If $w^*(E) = 0$ then E is w-measurable.

(2) The countable union of measurable rectangles (R_σ sets) and the countable intersection of R_σ sets ($R_{\sigma\delta}$ sets) are all measurable.

(3) For every measurable set E with finite measure, there exists $G \in R_{\sigma\delta}$ and $w(F) = 0$ such that $E = G \setminus F$.

2.5.3 Theorem (Fubini's Theorem)

Suppose $f \colon X \times Y \longrightarrow \overline{\mathbb{R}}$ is w-integrable. Then

(1) the function $f_x \colon Y \longrightarrow \overline{\mathbb{R}}$ defined by $f_x(y) = f(x, y)$ is ν-integrable for μ-almost every x,

(2) the function $f_y(x) = f(x, y)$ is μ-integrable for ν-almost every y,

(3) the function $x \mapsto \int_Y f_x \, d\nu$ is μ-integrable,

(4) the function $y \mapsto \int_X f_y \, d\mu$ is ν-integrable,

(5)

$$\int_Y \left[\int_X f_y \, d\mu\right] d\nu = \int_{X\times Y} f \, dw = \int_X \left[\int_Y f_x \, d\nu\right] d\mu$$

2.5.4 Theorem (Tonelli's Theorem)

Let $f\colon X \times Y \longrightarrow [0,\infty]$ be non-negative and w-measurable. Then

(1) f_x is ν-measurable for μ-almost every x,

(2) f_y is μ-measurable for ν-almost every y,

(3) the function $x \mapsto \int_Y f_x \, d\nu$ is μ-measurable,

(4) the function $y \mapsto \int_X f_y \, d\mu$ is ν-measurable,

(5)

$$\int_Y \left[\int_X f_y \, d\mu\right] d\nu = \int_{X\times Y} f \, dw = \int_X \left[\int_Y f_x \, d\nu\right] d\mu$$

Combining these together, we get the following:

2.5.5 Theorem (Fubini-Tonelli)

Let $f\colon X \times Y \longrightarrow \overline{\mathbb{R}}$ be measurable, then

$$\int_Y \left[\int_X |f_y| \, d\mu\right] d\nu = \int_{X\times Y} |f| \, dw = \int_X \left[\int_Y |f_x| \, d\nu\right] d\mu$$

and if any of these are finite,

$$\int_Y \left[\int_X f_y \, d\mu\right] d\nu = \int_{X\times Y} f \, dw = \int_X \left[\int_Y f_x \, d\nu\right] d\mu$$

Proof: the equality of the absolute values is due to Tonelli's theorem, since $|f|$ is nonnegative and $|f|_x = |f_x|$. If the terms are finite, then f is integrable (since f is integrable if and only if $|f|$ is)and so the second set of equalities follows from Fubini's theorem. ■

2.5.6 Definition

Let $f\colon X \longrightarrow \overline{\mathbb{R}}$ be a function, then its **support** is $\operatorname{supp}(f) := \{x \in X \mid f(x) \neq 0\}$. Define $C_0(\mathbb{R})$ to be the set of continuous real functions with compact support.

2.5.7 Theorem

Let $f\colon \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ be Lebesgue integrable. Then for every $\varepsilon > 0$ there exists a $g \in C_0(\mathbb{R})$ such that $\int_{\mathbb{R}} |f-g| \, dx < \varepsilon$.

Proof: in steps.

(1) If $f = \chi_{(a,b)}$ then add to the edges of f linear segments which go from $(a-\varepsilon, 0)$ to $(a,1)$ and $(b,1)$ to $(b+\varepsilon, 0)$ to form g. Then $g - f$ are just two triangles of height 1 and width ε which has an integral of ε.

(**2**) If $f = \chi_{\mathcal{U}}$ for $\mathcal{U}$ open then $\mathcal{U} = \dot{\bigcup}_{n=1}^{\infty}(a_n, b_n)$ a countable disjoint union of open intervals. Then $\chi_{\mathcal{U}} = \sum_{n=1}^{\infty} \chi(a_n, b_n)$, and using (1) we can approximate each $\chi(a_n, b_n)$ by a continuous g_n whose difference in integral is $\frac{\varepsilon}{2^n}$. Then the total difference is bound by $\sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$.

(**3**) If $f = \chi_E$ for E measurable and bound, then there exists an open $E \subseteq \mathcal{U}$ such that $m(\mathcal{U} \setminus E) < \frac{\varepsilon}{2}$. So using (2) approximate χ_E by g for $\frac{\varepsilon}{2}$. Then

$$\int |g - f| \leq \int |g - \chi_{\mathcal{U}}| + \int |\chi_{\mathcal{U}} - \chi_E| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

(**4**) If f is a simple function with compact support, then f is the finite linear combination of functions in (3) (since compact implies bound), and it follows.

(**5**) If f is integrable with compact support, then f can be approximated by a simple function φ with the same support such that $\int |f - \varphi| < \frac{\varepsilon}{2}$. We will approximate φ by g up to $\frac{\varepsilon}{2}$, so

$$\int |f - g| \leq \int |f - \varphi| + \int |\varphi - g| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

(**6**) If f is a general integrable function, then define $f_n = f \cdot \chi_{[-n,n]}$. Then $|f_n - f| \leq |f_n| + |f|$ which is integrable so by the dominated convergence theorem $\int |f_n - f| \longrightarrow 0$. So take an n where $\int |f_n - f| < \frac{\varepsilon}{2}$ and approximate f_n by g up to $\frac{\varepsilon}{2}$, so

$$\int |f - g| \leq \int |f - f_n| + \int |f_n - g| < \varepsilon$$

■

2.5.8 Theorem

Let $f \colon \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ be Lebesgue integrable. Then

$$\lim_{n \to \infty} \int_{\mathbb{R}} f(x) \sin(nx) = 0$$

Proof: first let us assume that f is continuously differentiable and supported in the closed interval $[a, b]$. We use integration by parts (since Lebesgue and Riemann integration are the same for continuous functions). Notice that by assumption $f(a) = f(b) = 0$ (or we can extend the bounds outside the support) so this just becomes

$$\lim_{n \to \infty} \left| \int_{\mathbb{R}} f(x) \sin(nx) \right| = \lim_{n \to \infty} \left| \int_{\mathbb{R}} f'(x) \frac{\cos(nx)}{n} \right|$$

Now $|f'|$ is bound by $\|f'\|_\infty = \sup_{x \in [a,b]} |f'|$ which exists since f' is continuous, and so

$$\leq \|f'\|_\infty \cdot \lim_{n \to \infty} \int_a^b \frac{1}{n} = \|f'\|_\infty (b - a) \cdot 0 = 0$$

We then approximate f by continuously differentiable g up to $\frac{\varepsilon}{2}$. Then we can take an n such that we have $\left| \int_{\mathbb{R}} g(x) \sin(nx) \right| < \frac{\varepsilon}{2}$ so

$$\left| \int_{\mathbb{R}} f(x) \sin(nx) \right| \leq \left| \int_{\mathbb{R}} g(x) \sin(nx) \right| + \left| \int_{\mathbb{R}} (f - g) \sin(nx) \right| \leq \frac{\varepsilon}{2} + \left| \int_{\mathbb{R}} f - g \right| < \varepsilon$$

so the limit is zero, as required. ■

2.6 L^p-spaces

2.6.1 Definition

Let (X, μ) be a measure space, define the equivalence relation $\sim$ on functions $f\colon X \longrightarrow \overline{\mathbb{R}}$ by $f \sim g$ if $f = g$ almost everywhere. Then $L^1(X, \mu)$ is the quotient space of integrable functions $X \longrightarrow \overline{\mathbb{R}}$ with respect to this relation.

Define for $p > 0$

$$L^p(X, \mu) = \left\{f\colon X \longrightarrow \overline{\mathbb{R}} \mid |f|^p \in L^1(X, \mu)\right\}/\sim$$

again the space is quotiented by $\sim$.

Define the L^p-norm (we have not yet proven that it is a norm, and it isn't for all p):

$$\|f\|_p := \left(\int_X |f|^p\right)^{1/p}$$

$L^p(X, \mu)$ is a vector space: since if $f, g \in L^p(X, \mu)$ then

$$\int_X |f + g|^p \, d\mu \le \int_X (2 \max |f|, |g|)^p \le 2^p \int_X (|f|^p + |g|^p) < \infty$$

and obviously $\alpha f \in L^p(X, \mu)$.

Notice that the L^1-norm is indeed a norm: $\|f\|_1 = 0$ if and only if $\int |f| = 0$ which is if and only if $f = 0$ almost everywhere (so $f \sim 0$). Obviously $\|\alpha f\|_1 = |\alpha| \|f\|_1$, and

$$\|f + g\| = \int |f + g| \le \int |f| + \int |g| = \|f\| + \|g\|$$

So it satisfies the triangle inequality.

Notice that we showed before that $C_0(\mathbb{R})$ is dense in $L^p(\mathbb{R})$. We know from calculus that $C[a, b]$ (the space of continuous functions on $[a, b]$) is complete relative to the norm $\|f\|_\infty := \max_{x \in [a,b]} |f(x)|$. Recall the definition of completeness:

2.6.2 Definition

A normed space $(X, \|\bullet\|)$ is **complete** if every Cauchy sequence relative to the norm $\|\bullet\|$ converges relative to the norm $\|\bullet\|$.

We will show later that $L^p(X, \mu)$ are complete for $p \ge 1$.

2.6.3 Definition

Let (X, μ) be a measure space, and $f\colon X \longrightarrow \overline{\mathbb{R}}$ measurable, then define

$$\|f\|_\infty := \inf\{M \in \mathbb{R} \mid |f(x)| \le M \text{ almost everywhere}\}$$

and define the L^∞ space:

$$L^\infty(X, \mu) = \left\{f\colon X \longrightarrow \overline{\mathbb{R}} \text{ measurable} \;\middle|\; \|f\|_\infty < \infty\right\}$$

2.6.4 Lemma

For every $a, b \ge 0$ and $0 < \lambda < 1$ we have $a^\lambda b^{1-\lambda} \le \lambda a + (1 - \lambda b)$ with equality if and only if $a = b$.

Proof: this is equivalent to

$$\left(\frac{a}{b}\right)^\lambda \le \lambda \frac{a}{b} + (1 - \lambda)$$

Define $t = \frac{a}{b}$, so we must prove $\varphi(t) = \lambda t - t^\lambda + (1-\lambda) \geq 0$ for every $t \geq 0$. Differentiating gives $\varphi'(t) = \lambda - \lambda t^{\lambda-1} = \lambda(1 - t^{\lambda-1})$. This is zero when $t = 1$, $\varphi' < 0$ when $t < 1$, and $\varphi' > 0$ when $t > 1$. So the minimum is obtained when $t = 1$, ie. $a = b$. But $\varphi(1) = 0$ so $\varphi' \geq 0$ with equality when $a = b$. ■

2.6.5 Theorem (Hölder's Inequality)

Let $1 \leq p \leq \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$ and $f \in L^p(X,\mu), g \in L^q(X,\mu)$ then $fg \in L^1(X,\mu)$ and

$$\|fg\|_1 \leq \|f\|_p \|g\|_q$$

Proof: define $\lambda = \frac{1}{p}$ so $1 - \lambda = \frac{1}{q}$, $a = |f(x)|^p$ and $b = |g(x)|^q$. So we have that

$$|fg| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q$$

for every $x \in X$. First let us assume $\|f\|_p = 1 = \|g\|_q$, so we have that

$$\|fg\|_1 = \int |f(x)g(x)| \leq \frac{1}{p}\int |f| + \frac{1}{q}\int |g| = \frac{1}{p}\|f\|_p^p + \frac{1}{q}\|f\|_q^q = \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p \|g\|_q$$

Now otherwise, define $f' = \frac{f}{\|f\|_p}$ and $g' = \frac{g}{\|g\|_q}$ so we have shown that

$$1 \geq \|f'g'\|_1 = \frac{1}{\|f\|_p \cdot \|g\|_q}\|fg\|_1 \implies \|fg\|_1 \leq \|f\|_p \|g\|_q$$ ■

2.6.6 Theorem (Minkoswki's Inequality)

Let $f, g \in L^p(X,\mu)$ then

$$\|f+g\|_p \leq \|f\|_p + \|g\|_q$$

Proof: let $\frac{1}{p} + \frac{1}{q} = 1$, meaning $q = \frac{p}{p-1}$ then

$$\begin{aligned} \|f+g\|_p^p = \int |f+g|^p = \int |f+g|^{p-1}|f+g| &\leq \int |f+g|^{p-1}|f| + \int |f+g|^{p-1}|g| \\ &= \left\| |f+g|^{p-1}|f| \right\|_1 + \left\| |f+g|^{p-1}|g| \right\|_1 \end{aligned}$$

Notice that $p + q = pq$ and so $\left(|f+g|^{p-1}\right)^q = |f+g|^{pq-q} = |f+g|^p$, meaning $|f+g|^{p-1} \in L^q(X,\mu)$, so by Hölder:

$$\leq \left\| |f+g|^{p-1} \right\|_q \|f\|_p + \left\| |f+g|^{p-1} \right\|_q \|g\|_p$$

Now

$$\left\| |f+g|^{p-1} \right\|_q = \left(\int_X |f+g|^p \right)^{1/q} = \|f+g\|_p^{p/q}$$

So we have shown that

$$\|f+g\|_p^{p-p/q} \leq \|f\|_p + \|g\|_p$$

but $p - p/q = 1$ so we have $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ as required. ■

This means that for $p \geq 1$, $\|\bullet\|_p$ is indeed a norm on $L^p(X,\mu)$. Minkowski's inequality proves the triangle inequality, if $\|f\|_p = 0$ implies $\int_X |f|^p = 0$ and since $|f|^p \geq 0$ this means $|f| = 0$ almost everywhere so it is zero in $L^p(X,\mu)$, and the norm is trivially nonnegative.

Now we want to prove that $L^p(X,\mu)$ is complete.

2.6.7 Lemma

Let $(X, \|\bullet\|)$ be a normed vector space. Then X is complete if and only if every series which converges absolutely also converges. Meaning if $\sum_{n=1}^{\infty}\|x_n\|$ converges, so too does $\sum_{n=1}^{\infty} x_n$.

Proof: suppose that absolute convergence implies convergence. Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence. Define n_k to be the index such that for every $n > n_k$, $\|x_{n_k} - x_n\| < \frac{1}{2^k}$, and so let $n > n_k, n_{k+1}$:

$$\|x_{n_k} - x_{n_{k+1}}\| \leq \|x_{n_k} - x_n\| + \|x_n - x_{n_{k+1}}\| < \frac{1}{2^k} + \frac{1}{2^{k+1}} = \frac{3}{2^{k+1}}$$

and so $\sum\|x_{n_k} - x_{n_{k+1}}\|$ converges, meaning $\sum(x_{n_k} - x_{n_{k+1}})$ converges to a value. Notice that

$$x_{n_k} = x_{n_1} + \sum_{k=1}^{\infty}(x_{n_k} - x_{n_{k+1}})$$

and so x_{n_k} converges to some $x_0 \in X$, ie. $x_{n_k} \xrightarrow{k\longrightarrow\infty} x_0$. So $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence with a convergent subsequence, which we knows means that x_n itself is convergent.
For the converse, all we must do is show that $\sum x_n$ is Cauchy. Notice that for $N > M$,

$$\left\|\sum_{n=1}^{N} x_n - \sum_{n=M}^{M} x_n\right\| = \left\|\sum_{n=M+1}^{N} x_n\right\| \leq \sum_{n=M+1}^{N} \|x_n\| \leq \sum_{n=M+1}^{\infty} \|x_n\|$$

and since $\sum\|x_n\|$ is convergent, its tail converges to zero. Thus for any $\varepsilon > 0$ we can find a n_0 such that for every $N > M > n_0$ the above expression is less than ε and we have the desired. ∎

2.6.8 Theorem

The normed vector spaces $L^p(X,\mu)$ are complete for $p \geq 1$.

Proof: let $\sum_{n=1}^{\infty} f_n$ be a series that converges absolutely where $f_n \in L^p$. Let us define

$$g(x) = \sum_{n=1}^{\infty}|f_n(x)|$$

which exists for every $x \in X$ (in a general sense, it can be ∞). Let us define the partial sums

$$g_N(x) = \sum_{n=1}^{N}|f_n(x)|$$

So we have that

$$\|g_N\|_p \leq \sum_{n=1}^{N}\|f_n\|_p \leq \sum_{n=1}^{\infty}\|f_n\|_p = M < \infty$$

since $\sum f_n$ converges absolutely. And g_N monotonically increase to g, and so $|g_N|^p$ does to $|g|^p$ as well, meaning by the monotone convergence theorem

$$\int_X |g|^p = \lim_{N\to\infty}\int_X |g_N|^p = \lim_{N\to\infty}\|g_N\|_p^p \leq M^p$$

This means that $g \in L^p(X,\mu)$, and in particular $g(x) < \infty$ almost everywhere. Thus $\sum_{n=1}^{\infty} f_n(x)$ converges absolutely almost everywhere (as a series of real numbers), and thus converges almost everywhere to $f(x) = \sum_{n=1}^{\infty} f_n(x)$. Since $|f(x)| \leq g(x)$ we have that $f \in L^p(X,\mu)$.
Now we need to show that in $L^p(X,\mu)$, $\sum_{n=1}^{N} f_n \longrightarrow f$. Let us define $S_N = \sum_{n=1}^{N} f_n$, then also $|S_N| \leq g$ so $|f - S_N| \leq |f| + |S_N| \leq |f| + g$ which is integrable. So by the dominated convergence theorem since $f - S_N \longrightarrow 0$,

$$\lim_{N\to\infty}\|f - S_N\|^p = \lim_{N\to\infty}\int_X |f - S_N|^p = \int_X \lim_{N\to\infty}|f - S_N|^p = \int_X 0 = 0$$ ∎

2.6.9 Theorem

Let $\{f_n\}_{n=1}^{\infty}$ be a convergent sequence of functions in $L^p(X,\mu)$. Then there exists a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ which converges pointwise almost everywhere.

Proof: let n_k be an index such that for every $n > n_k$, $\|f_n - f_{n_k}\|_p < 2^{-k}$ and so as before $\sum\|f_{n_{k+1}} - f_{n_k}\|$ is convergent. Now we showed in the above proof that we can define $f(x) = f_{n_1} + \sum_{k=1}^{\infty}\left(f_{n_{k+1}} - f_{n_k}\right)$ and this converges pointwise almost everywhere. Noticing once again that $f_{n_k} = f_{n_1} + \sum_{i=1}^{k-1}\left(f_{n_{i+1}} - f_{n_i}\right)$, this means f_{n_k} converges pointwise almost everywhere to f as required. ∎

2.6.10 Definition

Let $(X, \|\bullet\|_X)$ and $(Y, \|\bullet\|_Y)$ be two normed vector spaces. Further let $T\colon X \longrightarrow Y$ be a linear transform. Define the **operator norm** to be

$$\|T\|_{\text{op}} := \sup_{0\neq x\in X} \frac{\|Tx\|_Y}{\|x\|_X}$$

Notice that if $\|x\|_X = \alpha$ then

$$\left\|T\frac{x}{\alpha}\right\|_Y = \frac{\|Tx\|_Y}{\|x\|_X}$$

and thus

$$\|T\|_{\text{op}} = \sup_{\|x\|_X=1} \|Tx\|_Y$$

2.6.11 Theorem

Let $(X, \|\bullet\|_X)$ and $(Y, \|\bullet\|_Y)$ be two normed vector spaces, and $T\colon X \longrightarrow Y$ a linear transform. Then the following are equivalent:

(1) T is bound: $\|T\|_{\text{op}} < \infty$,

(2) T is uniformly continuous on X,

(3) T is continuous at a single point $x_0 \in X$,

(4) T is continuous at $0 \in X$.

Proof: (1) $\implies$ (2): we know that $\|T\| \geq \frac{\|Tx_1 - Tx_2\|_Y}{\|x_1 - x_2\|_X}$ so $\|Tx_1 - Tx_2\|_Y \leq \|T\|_{\text{op}}\|x_1 - x_2\|_X = M\|x_1 - x_2\|_X$ where $M = \|T\|_{\text{op}}$. Thus for every $\varepsilon > 0$ take $x_1, x_2 \in X$ such that $\|x_1 - x_2\| < \frac{\varepsilon}{M}$ and then $\|Tx_1 - Tx_2\|_Y \leq \varepsilon$, so T is uniformly continuous. (2) $\implies$ (3): trivial. (3) $\implies$ (4): let $x_n \to 0$ then $x_n + x_0 \to x_0$ and so $T(x_n + x_0) \to Tx_n + Tx_0 \to Tx_0$ thus $Tx_n \to 0$. Meaning T is continuous at 0. (4) $\implies$ (1): since T is continuous at 0, so there exists a $\delta > 0$ such that $\|x\| < \delta$ implies $\|Tx\| < 1$. So for $z \in X$ define $x := \frac{\delta z}{2\|z\|}$ so that $\|x\| < \delta$ and thus $\|Tx\| = \frac{\delta}{2\|z\|}\|Tz\| < 1$, meaning $\frac{\|Tz\|}{\|z\|} < \frac{2}{\delta}$. Thus $\|T\|_{\text{op}} \leq \frac{2}{\delta}$ as required. ∎

For example, let $f \in L^p(X,\mu)$ and $\frac{1}{p} + \frac{1}{q} = 1$, define

$$T_f\colon L^q(X,\mu) \longrightarrow \mathbb{R}, \qquad T_f g = \int_X fg\, d\mu$$

This is obviously a linear transform. By Hölder,

$$|T_f g| \leq \|f\|_p\|g\|_q$$

which means that $\|T_f\| \leq \|f\|_p$. Now let us take $g = \operatorname{sign} f \cdot |f|^{p-1}$ and notice that $|g|^q = |f|^{q(p-1)} = |f|^p$ so $g \in L^q(X,\mu)$ and

$$T_f g = \int_X |f|^p = \|f\|_p^p$$

and since $\|g\|_q^q = \|f\|_p^p$, we have that $\|g\|_q \|f\|_p = \|f\|_p^{p/q} \|f\|_p = \|f\|_p^{p/q+1} = \|f\|_p^p$

$$T_f g = \|f\|_p^p = \|g\|_q \|f\|_p$$

And thus our previous inequality is strict,

$$\|T_f\|_{\text{op}} = \|f\|_p$$