1. σ-ALGEBRAS

Definition 1. Let X be any set and let $\mathcal{F}$ be a collection of subsets of X. We say that $\mathcal{F}$ is a σ-algebra (on X), if it satisfies the following.

(1) $X \in \mathcal{F}$.
(2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.
(3) If $A_1, A_2, \cdots \in \mathcal{F}$, a countable collection, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

In the above situation, the elements of $\mathcal{F}$ are called *measurable sets* and X (or more precisely $(X, \mathcal{F})$) is called a *measurable space.*

Example 1. For any set X, $\mathcal{P}(X)$, the set of all subsets of X is a σ-algebra and so is $\mathcal{F} = \{\emptyset, X\}$.

Theorem 1. *Let X be any set and let $\mathcal{G}$ be any collection of subsets of X. Then there exists a σ-algebra, containing $\mathcal{G}$ and smallest with respect to inclusion.*

Proof. Let S be the collection of all σ-algebras containing $\mathcal{G}$. This set is non-empty, since $\mathcal{P}(X) \in S$. Then $\mathcal{F} = \cap_{\mathcal{A} \in S} \mathcal{A}$ can easily be checked to have all the properties asserted in the theorem. $\square$

Remark 1. In the above situation, $\mathcal{F}$ is called the σ-algebra *generated* by $\mathcal{G}$.

Definition 2. Let X be a topological space (for example, a metric space) and let $\mathcal{B}$ be the σ-algebra generated by the set of all open sets of X. The elements of $\mathcal{B}$ are called Borel sets and $(X, \mathcal{B})$, a Borel measurable space.

Definition 3. Let $(X, \mathcal{F})$ be a measurable space and let Y be any topological space and let $f : X \to Y$ be any function. We say that f is measurable, if for any open set $U \subset Y$, $f^{-1}(U) \in \mathcal{F}$.

Example 2.
(1) If $(X, \mathcal{F})$ is a measurable space with X a topological space and $\mathcal{B} \subset \mathcal{F}$, then any continuous function from X to any topological space is measurable.
(2) If $f : X \to Y$ is measurable and $g : Y \to Z$ is continuous, $g \circ f$ is measurable.

Theorem 2. *Let $(X, \mathcal{F})$ be a measurable space and let $f : X \to Y$ be a function, where Y is any set.*

(1) *If $\mathcal{G}$ is defined as the collection of all subsets A of Y such that $f^{-1}(A) \in \mathcal{F}$, then $\mathcal{G}$ is a σ-algebra on Y.*
(2) *If $f : X \to [-\infty, \infty]$ is any function, f is measurable if and only if $f^{-1}((a, \infty]) \in \mathcal{F}$ for any $a \in [-\infty, \infty]$.*

Proof. The first part is mere checking. For the second part, notice that since $(a, \infty]$ is open in $[-\infty, \infty]$, the condition is clearly necessary. To check sufficiency, one notes that $[-\infty, b) = \cup_n [-\infty, b - \frac{1}{n}]$ for any b and $[-\infty, c] = (c, \infty]^c$. So, one gets that $f^{-1}([-\infty, b)) \in \mathcal{F}$ too for any b and thus we get $f^{-1}((a, b)) \in \mathcal{F}$ for any a, b, since $(a, b) = [-\infty, b) \cap (a, \infty]$ and these generate all open sets. $\square$

Theorem 3. *Let $f_n : X \to [-\infty, \infty]$ be a sequence of measurable functions. Then, $\sup f_n, \inf f_n, \limsup f_n, \liminf f_n$ are all measurable.*

Proof. Let $g = \sup f_n$. Then $g^{-1}((a, \infty]) = \cup_n f_n^{-1}((a, \infty])$ and so we are done by the previous theorem. A similar proof applies to infimum. Let $g_n = \sup_{k \geq n} f_k$. Then all the g_n's are measurable by the previous part. Then $\limsup f_n = \inf g_n$ and hence it too is measurable. Similar argument applies to $\liminf$. $\square$

We get as a corollary the following results.

Corollary 1. *Let $f, g : X \to [-\infty, \infty]$ be measurable. Then so are $\max\{f, g\}$ and $\min\{f, g\}$. In particular, writing $f^+ = \max\{f, 0\}$ and $f^- = -\min\{f, 0\}$, we see that these are measurable.*

Remark 2. In all the above, we have dealt with $[-\infty, \infty]$ which has the advantage that supremum, $\limsup$ etc. make sense. But, notice that if $a, b \in [-\infty, \infty]$, then $a + b$ may not make sense. On the other hand, if we worked with $\mathbb{R}$, then the latter obviously makes sense, but not necessarily the former. In most cases, it should be clear which space we are working in and why.

Definition 4. A function $f : X \to [0, \infty)$ is called simple, if it takes only finitely many values.

Example 3.

1. If $A \subset X$, the characteristic function χ_A defined as $\chi_A(x) = 1$ if $x \in A$ an zero otherwise, is a simple function.
2. More generally, if s is a simple function, let $a_1, \ldots, a_n$ be its finitely many values and let $A_i = s^{-1}(a_i)$. Then $s = \sum_{i=1}^n a_i \chi_{A_i}$. Notice that $\cup_i A_i = X$ and $A_i \cap A_j = \emptyset$ if $i \neq j$ (pairwise disjoint).
3. If s is a simple function as above from a measurable space X, s is measurable if and only if all the A_i's are measurable.
4. If $s = \sum a_i \chi_{A_i}$ and $t = \sum b_j \chi_{B_j}$ are simple (measurable) functions, so is $s + t$. Letting $C_{ij} = A_i \cap B_j$, one can see that $s + t = \sum (a_i + b_j) \chi_{C_{ij}}$.

Theorem 4. *Let $f : X \to [0, \infty]$ be a measurable function. Then there exists simple measurable functions $s_1 \leq s_2 \leq \cdots \leq f$ such that $\lim s_n = f$.*

Proof. For any n and an i such that $1 \leq i \leq n2^n$, let $E_{n,i} = f^{-1}([\frac{i-1}{2^n}, \frac{i}{2^n}))$ and let $F_n = f^{-1}([n,\infty])$. Then these are all measurable sets, pairwise disjoint for a fixed n and cover all of X. So, $s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n}\chi_{E_{n,i}} + n\chi_{F_n}$ is a simple measurable function. Easy to check that $s_n \leq s_{n+1} \leq f$ for all n and $\lim s_n = f$. $\square$

2. Measures

Definition 5. Let $(X, \mathcal{F})$ be a measurable space. A (positive) measure on X is a function $\mu : \mathcal{F} \to [0,\infty]$ which is countably additive. That is, if $A_1, A_2, \ldots \in \mathcal{F}$ is a countable collection of pairwise disjoint sets, then $\mu(\cup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu(A_n)$.

Since $\mu(A) = \infty$ for all $A \in \mathcal{F}$ is trivially a measure by the above, to avoid this case, we will always assume that there exists some $A \in \mathcal{F}$ with $\mu(A) < \infty$ for our definition.

Example 4.
1. Let X be any set and consider the function $\mu : \mathcal{P}(X) \to [0,\infty]$ given by $\mu(A)$ is the cardinality of A if A is finite and ∞ otherwise. Then μ is a positive measure. This is called the counting measure.
2. Again consider $(X, \mathcal{P}(X))$ and let $a \in X$, fixed. Define $\mu(A) = 1$ if $a \in A$ and zero otherwise. This is a positive measure.

Theorem 5. *Let μ be a positive measure on $(X, \mathcal{F})$.*
1. $\mu(\emptyset) = 0$.
2. *If $A_1, A_2, \ldots, A_n \in \mathcal{F}$, pairwise disjoint, then $\mu(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mu(A_i)$. (Finite additivity)*
3. *If $A \subset B$ are measurable, $\mu(A) \leq \mu(B)$. (Monotonicity)*
4. *If $A_1 \subset A_2 \subset \cdots$ are measurable sets and $A = \cup_{n=1}^\infty A_n$, then $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.*
5. *If $A_1 \supset A_2 \supset \cdots$ and $A = \cap_{n=1}^\infty A_n$ with $\mu(A_1) < \infty$, then $\lim \mu(A_n) = \mu(A)$.*

Proof.
1. Let A be such that $\mu(A) < \infty$. Let $A_1 = A$ and $A_n = \emptyset$ for $n > 1$. Then these are pairwise disjoint and its union is A. So, we get $\mu(A) = \mu(A) + \sum \mu(\emptyset)$. Rest is clear, since we can cancel $\mu(A) < \infty$ from bot sides.
2. This is clear by taking $A_m = \emptyset$ for $m > n$ and using the first.
3. We can write $B = A \cup (B - A)$, pairwise disjoint and thus $\mu(B) = \mu(A) + \mu(B - A) \geq \mu(A)$.
4. Let $B_n = A_n - A_{n-1}$ for $n > 1$ and $B_1 = A_1$. Then $\{B_n\}$ are pairwise disjoint and $\cup_{n=1}^\infty B_n = A$. Also, $\cup_{i=1}^n B_i = A_n$. So, we get $\mu(A_n) = \sum_{i=1}^n \mu(B_i)$ and $\mu(A) = \sum_{n=1}^\infty \mu(B_n)$.

(5) For this, let $C_n = A_1 - A_n$. Then $C_n \subset C_{n+1}$ and $\cup_{n=1}^{\infty} C_n = A_1 - A$. So, from the previous part, we have $\lim \mu(C_n) = \mu(A_1 - A) = \mu(A_1) - \mu(A)$, since $A \subset A_1$. We also have by the same reason, $\mu(C_n) = \mu(A_1) - \mu(A_n)$. So, taking limits, we get, $\lim \mu(C_n) = \mu(A_1) - \lim(A_n) = \mu(A_1) - \mu(A)$ and canceling $\mu(A_1) < \infty$, we are done.

$\square$

3. Lebesgue Integral

Definition 6. Let $(X, \mathcal{F}, \mu)$ be a measure space and let $s : X \to [0, \infty)$ be a simple positive measurable function. Writing $s = \sum a_i \chi_{A_i}$, define $\int_X s d\mu = \sum a_i \mu(A_i)$, where we use the convention, $0 \cdot \mu(A) = 0$ even if $\mu(A) = \infty$. If $0 \leq f$ is any positive measurable function, we define $\int_X f d\mu$ to be the supremum of $\int_X s d\mu$ for simple measurable functions $0 \leq s \leq f$.

Remark 3. The above gives two possible definitions for the integral of a simple measurable function, but it is clear that the definitions coincide.

We collect a few easy facts which follow from the definitions.

Theorem 6. *Let $(X, \mathcal{F}, \mu)$ be a measure space. All functions mentioned below are measurable and so are all the subsets of X.*

(1) *If $0 \leq f \leq g$, then $\int_X f d\mu \leq \int_X g d\mu$.*
(2) *If $A \subset B$ in $\mathcal{F}$, for $0 \leq f$ on X, we have, $\int_A f d\mu \leq \int_B f d\mu$.*
(3) *If $0 \leq f$ and $0 \leq c < \infty$, then $\int_X c f d\mu = c \int_X f d\mu$.*
(4) *If $f(x) = 0$ for all $x \in A$, $\int_A f d\mu = 0$, even if $\mu(A) = \infty$.*
(5) *If $\mu(A) = 0$, for any measurable positive f, $\int_A f d\mu = 0$ even if $f(x) = \infty$ for some or all $x \in A$.*
(6) *$\int_X f \chi_A d\mu = \int_A f d\mu$.*
(7) *If s, t are simple functions, then $\int_X (s+g) d\mu = \int_X s d\mu + \int_X t d\mu$*

Proof. The only part which requires a proof is the last one.

If $s = \sum a_i \chi_{A_i}$ and $t = \sum b_j \chi_{B_j}$, we have $\int_X s d\mu = \sum a_i \mu(A_i)$ and $\int_X t d\mu = \sum b_j \mu(B_j)$. On the other hand, we have $s+t = \sum (a_i + b_j) \chi_{C_{ij}}$, where $C_{ij} = A_i \cap B_j$. Thus

$$\begin{aligned}\int_X (s+t)d\mu &= \sum_{i,j}(a_i + b_j)\mu(C_{ij}) \\ &= \sum_i a_i \sum_j \mu(C_{ij}) + \sum_j b_j \sum_i \mu(C_{ij}) \\ &= \sum a_i\mu(A_i) + \sum b_j\mu(B_j) \\ &= \int_X sd\mu + \int_X td\mu\end{aligned}$$

□

As a corollary, we get,

Corollary 2. *Let s be a simple measurable function on X and define $\phi(A) = \int_A sd\mu$ for A measurable. Then ϕ is a measure.*

Proof. Let $s = \sum_{i=1}^n a_i\chi_{A_i}$. If $B_1, B_2, \ldots$ is a countable collection of pairwise disjoint measurable sets and B is their union, we have,

$$\begin{aligned}\phi(B) = \sum_{i=1}^n a_i\mu(A_i \cap B) &= \sum_{i=1}^n a_i \sum_{j=1}^\infty \mu(A_i \cap B_j) \\ = \sum_{j=1}^\infty \sum_{i=1}^n a_i\mu(A_i \cap B_j) &= \sum_{j=1}^\infty \phi(B_j)\end{aligned}$$

Since $\phi(\emptyset) = 0 < \infty$, we see that ϕ is indeed a measure in our sense. □

Next we come to an important theorem.

Theorem 7 (Lebesgue Monotone Convergence Theorem)**.** *Let $0 \le f_1 \le f_2 \le \cdots \le f$ be a sequence of positive measurable functions converging to f. Then $\lim_{n\to\infty}\int_X f_n d\mu = \int_X f d\mu$.*

Proof. Let $\alpha = \lim_{n\to\infty}\int_X f_n d\mu$. Since $\int_X f_n d\mu \le \int_X f d\mu$ for all n, we get that $\alpha \le \int_X f d\mu$.

Next we prove the reverse inequality. For this, let $0 \le s \le f$ be a simple measurable function. It suffices to prove that $\alpha \ge \int_X sd\mu$. Let $0 < c < 1$ and consider the sets $A_n = \{x \in X | f_n(x) \ge cs(x)\}$. Then A_n is measurable, $A_1 \subset A_2 \subset \cdots$ and $\cup A_n = X$. Also, we have $\int_X f_n d\mu \ge \int_{A_n} f_n d\mu \ge c\int_{A_n} sd\mu$ for all n. Thus, $\alpha = \lim \int_X f_n d\mu \ge \lim c\int_{A_n} sd\mu$. Noting that the latter is a measure from the previous corollary and using an earlier theorem, we get that it is equal to $c\int_X sd\mu$. Since this is true for any $0 < c < 1$, we see that $\alpha \ge \int_X sd\mu$, finishing the proof. □

Corollary 3. *If $f, g : X \to [0,\infty]$ are measurable, $\int_X (f+g)d\mu = \int_X f d\mu + \int_X g d\mu$.*

Proof. By arguments as in the homework, it is immediate that $f+g$ is measurable. Let $0 \leq s_1 \leq s_2 \leq \cdots \leq f$ be simple function such that $\lim s_n = f$ (see Theorem 4). Similarly, $0 \leq t_1 \leq t_2 \leq \cdots \leq g$. Then $s_n + t_n$ is simple and $0 \leq s_1 + t_2 \leq s_2 + t_2 \leq \cdots \leq f+g$ with $\lim(s_n+t_n) = f+g$. So, by Theorem 7, we get $\lim \int_X (s_n+t_n)d\mu = \int_X (f+g)d\mu$. But we have $\int_X (s_n+t_n)d\mu = \int_X s_n d\mu + \int_X t_n d\mu$, since these are simple functions. We have $\lim \int_X s_n d\mu = \int_X f d\mu$ and $\lim \int_X t_n d\mu = \int_X g d\mu$ by Theorem 7 and so we are done. $\square$

Corollary 4. *Let $f_n : X \to [0,\infty]$ be a sequence of measurable functions and let $f = \sum_{n=1}^{\infty} f_n$. Then f is measurable and $\int_X f d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu$.*

Proof. Let $g_n = \sum_{k=1}^{n} f_k$. So, we have $\lim g_n = f$ with $0 \leq g_1 \leq g_2 \leq \cdots \leq f$ and thus by Theorem 7, $\lim \int_X g_n d\mu = \int_X f d\mu$. On the other hand, by a simple induction from the previous corollary, we have, $\int_X g_n d\mu = \sum_{k=1}^{n} \int_X f_k d\mu$ and thus $\lim \int_X g_n d\mu = \sum_{k=1}^{\infty} \int_X f_k d\mu$. $\square$

Corollary 5. *If $a_{ij} \geq 0$ are real numbers, $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}$.*

Proof. Consider $X = \mathbb{N}$ with its power set as a σ-algebra and μ, the counting measure. Let $f_n : X \to [0,\infty)$ be defined as $f_n(m) = a_{mn}$. Then f_n is measurable (if the σ-algebra is the power set, any function is measurable) and so we have, from the previous corollary,

$$\int_X \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu.$$

For any $g : X \to [0,\infty]$, we also have $\int_X g d\mu = \sum_{n=1}^{\infty} g(n)$. For this, consider simple functions $s_n : X \to [0,\infty]$ defined as $s_n(k) = g(k)$ if $k \leq n$ and zero other wise. Then $0 \leq s_1 \leq s_2 \leq \cdots \leq g$ and $\lim s_n = g$. So, $\int_X g d\mu = \lim \int_X s_n d\mu$. $s_n = \sum_{k=1}^{n} g(k)\chi_{\{k\}} + 0\chi\{k > n\}$ and thus $\int_X s_n d\mu = \sum_{k=1}^{n} g(k)\mu(\{k\}) = \sum_{k=1}^{n} g(k)$. Rest is clear. $\square$

Theorem 8 (Fatou's lemma)**.** *Let $f_n : X \to [0,\infty]$ be a sequence of measurable functions. Then,*

$$\int_X \liminf f_n d\mu \leq \liminf \int_X f_n d\mu.$$

Proof. Notice first, that we have proved $\liminf f_n$ is measurable, so the left hand side above makes sense.

Consider $g_k = \inf_{n \geq k} f_n$. Then g_k is measurable and we have $0 \leq g_1 \leq g_2 \leq \cdots \leq \liminf f_n$ and $\lim g_k = \liminf f_n$. Thus, by Theorem

7, we get the left hand side in the theorem to be $\lim \int_X g_k d\mu$. But $g_k \le f_n$ for all $n \ge k$ and thus $\int_X g_k d\mu \le \int_X f_n d\mu$ for all $n \ge k$. So, $\int_X g_k d\mu \le \inf_{n\ge k} \int_X f_n d\mu$. So,

$$\lim \int_X g_k d\mu \le \lim_k \inf_{n\ge k} \int_X f_n d\mu = \liminf \int_X f_n d\mu.$$

□

Theorem 9. *Let $f : X \to [0,\infty]$ be measurable. The the function $\phi : \mathcal{F} \to [0,\infty]$ given by $\phi(A) = \int_A f d\mu$ is a positive measure. If g is any positive measurable function, we have $\int_X g d\phi = \int_X g f d\mu$*

Proof. We have $\phi(\emptyset) = 0$. If $A_1, A_2, \dots$ is a countable collection of measurable disjoint sets with $A = \cup A_n$, we have $\chi_A = \sum \chi_{A_i}$. Thus, $\phi(A) = \int_A f d\mu = \int_X \chi_A f d\mu = \sum \int_X \chi_{A_i} f d\mu$, the last by the corollary to Theorem 7. Since $\int_X \chi_{A_i} f d\mu = \int_{A_i} f d\mu = \phi(A_i)$, we get countable additivity for ϕ proving that it is indeed a measure.

For the last part, we consider $g = \chi_A$ for some $A \in \mathcal{F}$. Then,

$$\int_X g d\phi = \phi(A) = \int_A f d\mu = \int_X g f d\mu.$$

Thus, the same holds for any simple measurable function and then Theorem 7 finishes the proof for any g, since any positive measurable function is the limit of an increasing sequence of simple functions. □

4. L^1 SPACES

In this section, we will consider functions $f : X \to \mathbb{R}$, not necessarily positive, but only taking finite values. If f is measurable, we have seen that both f^+, f^- are measurable and thus so is $|f| = f^+ + f^-$.

Definition 7. A measurable function $f : X \to \mathbb{R}$ belongs to the set $L^1(\mu)$ if $\int_X |f| d\mu < \infty$.

If $f \in L^1(\mu)$, then clearly $|f| \in L^1(\mu)$. Also, both $\int_X f^+ d\mu, \int_X f^- d\mu$ are finite and hence it makes sense to define,

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu.$$

For a measurable function f, not necessarily positive, $\int_X f d\mu$ may not make sense, since the only reasonable way we could do this was to write $f = f^+ - f^-$ and attempt to define $\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu$. The subtraction above will make sense only if at least on those integrals were finite.

Theorem 10. *The set $L^1(\mu)$ is a vector space over $\mathbb{R}$. That is, given $f, g \in L^1(\mu)$, $a, b \in \mathbb{R}$, $af + bg \in L^1(\mu)$.*

Proof. We already know that $af + bg$ is measurable (since they take finite values). So,

$$\int_X |af+bg|d\mu \leq \int_X (|a||f|+|b||g|)d\mu \leq |a|\int_X |f|d\mu + |b|\int_X |g|d\mu < \infty.$$

□

Theorem 11. *If $f \in L^1(\mu)$, $|\int_X f d\mu| \leq \int_X |f|d\mu$.*

Proof. If $a = \int_X f d\mu$, then $|\int_X f d\mu| = \theta a$ where $\theta = \pm 1$. So, $|\int_X f d\mu| = \int_X \theta f d\mu$. But, $\theta f \leq |\theta f| = |f|$ and so we are done. □

Theorem 12 (Lebesgue Dominated Convergence Theorem)**.** *Let $f_n : X \to \mathbb{R}$ be a sequence of measurable functions converging pointwise to a function $f : X \to \mathbb{R}$. Assume that there exists a $g \in L^1(\mu)$ such that $|f_n| \leq g$ for all n. Then, $\lim \int_X f_n d\mu = \int_X f d\mu$.*

Proof. Clearly, we have $|f| \leq g$ and thus $|f_n - f| \leq 2g$ for all n. So, $h_n = 2g - |f_n - f|$ are positive measurable functions. Thus we can apply Fatou's lemma, to conclude, $\int_X \liminf h_n d\mu \leq \liminf \int_X h_n d\mu$. Clearly $\liminf h_n = \lim h_n = 2g$. On the other hand, $\int_X h_n d\mu = \int_X 2g d\mu - \int_X |f_n - f|d\mu$ and thus,

$$\begin{aligned}\liminf \int_X h_n d\mu &= \int_X 2g d\mu + \liminf(-\int_X |f_n - f|d\mu) \\ &= \int_X 2g d\mu - \limsup \int_X |f_n - f|d\mu\end{aligned}$$

So, we get,

$$\int_X 2g d\mu \leq \int_X 2g d\mu - \limsup \int_X |f_n - f|d\mu.$$

Since $\int_X 2g d\mu < \infty$, we can cancel these to get, $\limsup \int_X |f_n - f|d\mu \leq 0$. Easy to see that this implies $\lim \int_X |f_n - f|d\mu = 0$. Since $|\int_X f_n - f d\mu| \leq \int_X |f_n - f|d\mu$ from the previous result, we get that $\lim |\int_X f_n - f d\mu| = 0$ and thus $\lim \int_X f_n d\mu = \int_X f d\mu$. □

5. Sets of measure zero

If $(X, \mathcal{F}, \mu)$ is a measure space, any $A \in \mathcal{F}$ with $\mu(A) = 0$ is naturally called a set of measure zero. If $B \subset A$, we would like to say that $\mu(B) = 0$, but B may not be in $\mathcal{F}$. So, we want to enlarge $\mathcal{F}$ to $\mathcal{G}$ so that, if $A \in \mathcal{F}$ has measure zero, and $B \subset A$, then $B \in \mathcal{G}$ and of course $\mathcal{F} \subset \mathcal{G}$. This may not be a σ-algebra, but of course we can replace $\mathcal{G}$ by the smallest σ-algebra containing $\mathcal{G}$. Then it is not clear that we can extend μ to this bigger σ-algebra, since we do not have an explicit enough description of it. All this can be achieved by the following construction, which can be easily checked to be the same as above.

Consider $\mathcal{G}$ to be the collection of all subsets $E \subset X$ such that there exists $A, B \in \mathcal{F}$ with $A \subset E \subset B$ and $\mu(B - A) = 0$. Then using simple set theory, one can easily check that $\mathcal{G}$ is a σ-algebra containing $\mathcal{F}$ and if we define $\mu(E) = \mu(A) = \mu(B)$, it is well defined. We usually call such a $\mathcal{G}$ (or more precisely $(X, \mathcal{G}, \mu)$) a *complete* measure space.

We will often say 'something happens' *almost everywhere*, abbreviated to a. e. to mean that it happens outside a set of measure zero. For example, we may write $f = g$ a. e. for two functions f, g on X when they coincide outside some unspecified set of measure zero. It is also common to use the expression *for almost all* $x \in X$ to mean the same.

Remark 4. Here is an important remark. Let $(X, \mathcal{F}, \mu)$ be a complete measure space. Let f be a measurable function defined a. e. This means, f is defined on $X - E$ where E is a set of measure zero and measurable on $X - E$ (with respect to the restricted σ-algebra). Then extending f to all of X by assigning arbitrary values at points of E, one easily checks that this function is measurable on all of X. So, one can be loose about the definition of f on a set of measure zero.

Many of the earlier results we proved for functions on all of X can thus be extended to functions defined a. e.

To illustrate, here is a corollary to Theorems 7 and 12

Theorem 13. *Suppose* $f_n : X \to \mathbb{R}$ *(or* $\mathbb{C}$*) be a sequence of measurable functions defined a. e. on* X *such that* $\sum_{n=1}^{\infty} \int_X |f_n| d\mu < \infty$. *Then the series* $f(x) = \sum_{n=1}^{\infty} f_n(x)$ *converges a. e. to a function* $f \in L^1(\mu)$ *and* $\int_X f d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu$.

Proof. Let E_n be the set where f_n is defined. Then letting $E = \cap E_n$, we see that $\mu(E^c) = 0$. Let $\phi(x) = \sum_{n=1}^{\infty} |f_n(x)|$. Then $\phi(x) : X \to [0, \infty]$ is a measurable function and $\int_X \phi d\mu = \sum_{n=1}^{\infty} \int_X |f_n| d\mu < \infty$. Let $A = \phi^{-1}(\infty)$. Then $\int_X \phi d\mu \geq \int_A \phi d\mu = \infty \cdot \mu(A)$. This implies that $\mu(A) = 0$. So, outside the measure zero set $E^c \cup A$, ϕ is a function to $\mathbb{R}$ and thus $f(x) = \sum f_n(x)$ converges (absolutely) on $E \cap A^c$. Further $|f(x)| \leq \phi(x)$ on this set and thus by Theorem 12, we see that $f \in L^1(\mu)$ and the rest is clear too by Theorem 12. □

Below are some similar results, which are easy to prove.

Theorem 14. (1) *If $f : X \to [0, \infty]$ is measurable, $A \in \mathcal{F}$ and $\int_A f d\mu = 0$, then $f = 0$ a. e. on A.*

(2) *If $f \in L^1(\mu)$ and $\int_A f d\mu = 0$ for every $A \in \mathcal{F}$, then $f = 0$ a. e. on X.*

(3) *If $f \in L^1(\mu)$ and $|\int_X f d\mu| = \int_X |f| d\mu$, then there is a constant a such that $af = |f|$ a. e. on X.*

Theorem 15. *Suppose $\mu(X) < \infty$ and $f \in L^1(\mu)$ and let $T \subset \mathbb{R}$ (or $\mathbb{C}$) be a closed subset. If the averages $A_E(f) = \frac{1}{\mu(E)} \int_E f d\mu$ lie in T for every $E \in \mathcal{F}$ with $\mu(E) > 0$, then $f(x) \in T$ for almost all $x \in X$.*

Proof. We can cover $\mathbb{R} - T$ by countably many open intervals of the form $I = (a - r, a + r) \subset \mathbb{R} - T$ where $a \notin T$ and $r > 0$. So, suffices to prove that $E = f^{-1}(I)$ has measure zero.

If $\mu(E) > 0$, then,

$$|A_E(f) - a| = \frac{1}{\mu(E)} \left| \int_E (f - a) d\mu \right| \leq \frac{1}{\mu(E)} \int_E |f - a| d\mu < r.$$

But $A_E(f) \in T$ and thus $|A_E(f) - a| \geq r$. This is a contradiction. □

Theorem 16. *Let $A_n \in \mathcal{F}$ such that $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. Then almost all x lie in at most finitely many A_ns.*

Proof. Consider $\phi(x) = \sum_{n=1}^{\infty} \chi_{A_n}$. Then $\phi : X \to [0, \infty]$ is measurable and $\int_X \phi d\mu = \sum \int_X \chi_{A_n} d\mu$ by Theorem 7. The latter is just $\sum \mu(A_n) < \infty$, by assumption. Thus as before, $\phi^{-1}(\infty)$ has measure zero. If an x belonged to infinitely many A_n's, then $\phi(x) = \infty$. This finishes the proof. □

6. Riesz Representation Theorem

We start with some notation. We will always use K for a compact set in a topological space X and V for an open set. We also write $C_c(X)$ to denote continuous functions on X with compact support- that is, functions vanishing outside a compact set. If $K \subset V \subset X$, and if $f \in C_c(X)$ with $f = 1$ on K and $f = 0$ outside V, with $0 \leq f \leq 1$, we write $K \prec f \prec V$. The notation $K \prec f$ would mean $f = 1$ on K, where $f \in C_c(X)$ and similarly $f \prec V$ would mean $f = 0$ outside V with $f \in C_c(X)$. The theorems in this section are valid for any locally compact Hausdorff topological space, but we will work with $\mathbb{R}^k$ (or subsets of it) , which is one such and we will denote it by X.

We start with a slight variation of Tietze extension theorem.

Theorem 17. *Let $K \subset V \subset X$. Then there exists an $f \in C_c(X)$ such that $K \prec f \prec V$.*

Proof. Notice that we may replace V by any open set containing K and contained in V. Since K is compact, it is bounded and hence contained in a ball of radius R with center the origin and $R > 0$. Thus, we may replace V by $V \cap B(0, R)$ and assume V is bounded. In particular, if $K \prec f \prec V$, with f continuous, it is in $C_c(X)$, since $f = 0$ outside the closed and bounded (and hence compact) set, the closure of V.

If we take the function which is 1 on K and zero on $X - V$, then it is continuous on the closed set $K \cup (X - V)$ and then by Tietze, we can extend this to a continuous function on X with $|f| \leq 1$. Clearly, we may replace f with $|f|$ and so we are done. $\square$

Remark 5. Some of you might have guessed that this could be done much more easily, since we are working on a metric space with the usual metric, which let me call d. Letting $K' = X - V$ above, we can explicitly describe such a function by taking $f(x) = \frac{d(x,K')}{d(x,K')+d(x,K)}$.

Theorem 18 (Partition of Unity)**.** *Let K be a compact set contained in $\cup_{i=1}^{n} V_i$ where V_is are open. Then there exists $f_i \in C_c(X)$ such that $f_i \prec V_i$ and $\sum f_i(x) = 1$ for all $x \in K$.*

Proof. For any point $x \in K$, we can find an open neighborhood W_x such that the closure of W_x is compact and this compact neighborhood is contained in V_i for some i. Since these cover K, by compactness, we can find $W_{x_1}, \ldots, W_{x_m}$ which cover K. Collecting the ones contained in V_i, thus we get $H_i \subset V_i$, H_i compact and $K \subset \cup H_i$.

By the previous result, we can find g_i such that $H_i \prec g_i \prec V_i$. Let $h_i = 1 - g_i$. Notice that $h_i \in C_c(X)$ and $0 \leq h_i \leq 1$. Now, consider f_i defined as,

$$\begin{aligned} f_1 &= 1 - h_1 \\ f_2 &= h_1 - h_1 h_2 \\ f_3 &= h_1 h_2 - h_1 h_2 h_3 \\ &\cdots\cdots \\ f_n &= h_1 \cdots h_{n-1} - h_1 \cdots h_n \end{aligned}$$

Clearly, all the $f_i \in C_c(X)$ and $0 \leq f_i \leq 1$, since each h_i is. If $x \notin V_i$, then $h_i(x) = 1$ and thus,

$$f_i(x) = h_1 \cdots h_{i-1}(x) - h_1 \cdots h_i(x) = 0.$$

So, $f_i \prec V_i$. If $x \in K$, then $x \in H_i$ for some i and then $h_i(x) = 0$. But, $\sum f_i = 1 - h_1 \cdots h_n$ and so $\sum f_i(x) = 1$. $\square$

Theorem 19 (Riesz Representation Theorem)**.** *Let $L : C_c(X) \to \mathbb{R}$ be linear map such that $L(f) \geq 0$ if $f \geq 0$. (Usually referred to as a positive linear functional.) Then there exists a σ-algebra $\mathcal{F}$, containing all open sets and a unique positive measure μ such that for any $f \in C_c(X)$,*

$$L(f) = \int_X f d\mu.$$

Further, we have,

(1) *For any compact set K, $\mu(K) < \infty$.*
(2) *If V is any open set, then $\mu(V) = \sup\{\mu(K) | K \subset V\}$, K compact.*
(3) *If $E \in \mathcal{F}$, then $\mu(E) = \inf\{\mu(V) | E \subset V\}$, V open.*
(4) *$(X, \mathcal{F}, \mu)$ is a complete measure space.*

We will not a give a complete proof, though it is not difficult, it is a bit tedious.

Proof. First, we show that μ is unique. If $(\mathcal{F}_1, \mu_1), (\mathcal{F}_2, \mu_2)$ are two such, we will show that they agree on $\mathcal{F}_1 \cap \mathcal{F}_2$. So, the measures are essentially the same.

It is clear that we only need to show that $\mu_1(K) = \mu_2(K)$ for any compact set. Given $\epsilon > 0$, by property 3) above, there exists an open set V containing K such that $\mu_1(K) \leq \mu_1(V) \leq \mu_1(K) + \epsilon$. Choose f such that $K \prec f \prec V$. Then we have,

$$\begin{aligned} \mu_2(K) &= \int_X \chi_K d\mu_2 \leq \int_X f d\mu_2 = L(f) \\ &= \int_X f d\mu_1 \leq \int_X \chi_V d\mu_1 = \mu_1(V) \leq \mu_1(K) + \epsilon \end{aligned}$$

Since ϵ was arbitrary, we get $\mu_2(K) \leq \mu_1(K)$ and reversing the roles of μ_1, μ_2, we get equality.

Notice that since $\mu(K) = \mu_2(K) \leq L(f) < \infty$, we have Property 1) of the theorem, if we have the equation connecting L and the integral.

Next, we define $\mu : \mathcal{P}(X) \to [0, \infty]$ as follows. For any open set V, define

$$\mu(V) = \sup\{L(f) | f \prec V\}.$$

It is clear that if $U \subset V$ are open, then $\mu(U) \leq \mu(V)$. Next define for any set E,

$$\mu(E) = \inf\{\mu(V) | E \subset V, V \text{open}\}.$$

Notice that this definition agrees with the previous definition for open sets. It looks as though we have defined μ on all subsets of X. But, we will construct a suitable σ-algebra, only on which all the properties will be checked-mainly the countable additivity.

For this, we first look at $\mathcal{F}'$ consisting of all subsets of $E \subset X$ such that $\mu(E) < \infty$ and

$$\mu(E) = \sup\{\mu(K) | K \subset E, K\text{compact}\}. \tag{1}$$

(Those of you who have seen Lebesgue measure using outer and inner measure would see where one is going.)

Finally, define $\mathcal{F}$ to be the set of all E such that $E \cap K \in \mathcal{F}'$ for any compact set K.

First some remarks. Clearly $\mu(A) \leq \mu(B)$ if $A \subset B$. If $\mu(E) = 0$, then $E \in \mathcal{F}', E \in \mathcal{F}$. So, we see that property 2) and property 4) both hold for our μ.

Next we write down steps which will finish the proof.

(1) μ is countably sub-additive. That is, if $E_1, E_2, \ldots$ are arbitrary subsets of X, $\mu(\cup E_n) \leq \sum \mu(E_n)$.
(2) $\mathcal{F}'$ contains every compact set.
(3) $\mathcal{F}'$ contains every open set with finite measure. Further, they satisfy equation 1 above.
(4) μ is countably additive on $\mathcal{F}'$.
(5) If $E \in \mathcal{F}'$ and $\epsilon > 0$, there exists a compact K and open V such that $K \subset E \subset V$ with $\mu(V - K) < \epsilon$.
(6) If $E_1, E_2 \in \mathcal{F}'$, then so do $E_1 - E_2, E_1 \cup E_2, E_1 \cap E_2$.
(7) $\mathcal{F}$ is a σ-algebra containing all open sets.
(8) $\mathcal{F}'$ consists of precisely those elements of $\mathcal{F}$ with finite measure.
(9) μ is a measure on $\mathcal{F}$.
(10) For every $f \in C_c(X)$, $L(f) = \int_X f d\mu$.

□

7. L^p-SPACES

Definition 8. A function $\phi : (a, b) \to \mathbb{R}$ is called convex if for any $a < x, y < b$ and any $t \in [0, 1]$, $\phi((1-t)x + ty) \leq (1-t)\phi(x) + t\phi(y)$.

Lemma 1. *A convex function is continuous.*

Lemma 2. *If ϕ is differentiable and ϕ' is monotonically increasing, then ϕ is convex.*

Example 5. The above lemma implies $\phi(x) = e^x$ is convex on $\mathbb{R}$. Similarly, $\phi(x) = x^p$ is convex on $(0, \infty)$ if $p > 0$.

Theorem 20 (Jensen's inequality)**.** *Let $(X, \mathcal{F}, \mu)$ be a measure space with $\mu(X) = 1$. Assume $f \in L^1(X)$ with $f(X) \subset (a, b)$. If ϕ is convex* *on* (a,b), *one has,*

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

As an immediate corollary, we get that geometric mean is less than the arithmetic mean.

Corollary 6. *Let* $x_1, x_2, \ldots, x_n$ *be* n *positive real numbers. Then,*

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

Proof. Consider a space $X = \{p_1, \ldots, p_n\}$ consisting of n elements with measure $\mu(\{p_i\}) = \frac{1}{n}$ for all i. Consider the function $f : X \to \mathbb{R}$ given by $f(p_i) = \log x_i$ and apply Jensen's inequality with the convex function exp. $\square$

Definition 9. If p, q are positive real numbers with $\frac{1}{p} + \frac{1}{q} = 1$, we say that p, q are conjugate.

Notice that the above condition forces $1 \le p, q \le \infty$ and $p = 1$ if and only if $q = \infty$.

Theorem 21. *Let* p, q *be conjugate with* $1 < p, q < \infty$ *and let* (X, μ) *be a measure space. Let* f, g *be positive measurable functions on* X. *Then,*

(1) *Hölder's inequality:*

$$\int_X f d\mu \le \left(\int_X f^p d\mu\right)^{\frac{1}{p}} \left(\int_X g^q d\mu\right)^{\frac{1}{q}}$$

(2) *Minkowski's inequality:*

$$\left(\int_X (f+g)^p d\mu\right)^{\frac{1}{p}} \le \left(\int_X f^p d\mu\right)^{\frac{1}{p}} + \left(\int_X g^p d\mu\right)^{\frac{1}{p}}$$

Definition 10. Let $0 < p < \infty$ and let f be measurable on (X, μ). Define $||f||_p^p = \int_X f^p d\mu$ and let $L^p(X)$ consists of all measurable functions f with $||f||_p < \infty$.

We define $||f||_\infty$ somewhat differently. Let f be any positive measurable function and let $S = \{a \in \mathbb{R} | f^{-1}((a, \infty])\text{has measure zero}\}$. If $S = \emptyset$, define $||f||_\infty = \infty$ and otherwise define $||f||_\infty = \inf S$. Notice that S is bounded below by zero (unless $\mu(X) = 0$, which is a trivial case and we will not usually consider it). Define $L^\infty(X)$ to be the set of all measurable functions f such that $g = |f|$ has $||g||_\infty < \infty$. The previous theorem (Hölder's and Minkowski's inequalities) immediately give the following.

Theorem 22. *Let p, q be conjugates with $1 \leq p, q \leq \infty$.*

(1) *If $f \in L^p(X), g \in L^q(X)$ then $fg \in L^1(X)$ and,*
$$||fg||_1 \leq ||f||_p ||g||_q.$$
(2) *If $f, g \in L^p(X)$, then,*
$$||f + g||_p \leq ||f||_p + ||g||_p.$$

One immediately sees that $L^p(X)$ is a vector space and $|| \cdot ||$ is a norm. In particular, for $f, g \in L^p(X)$, we can define a metric by $d(f, g) = ||f - g||_p$, the only difficulty is that it is possible $d(f, g) = 0$ even if $f \neq g$. But it is clear that $d(f, g) = 0$ if and only if $f = g$ a. e. and thus identifying this we still get a new vector space with actually a metric on it. One usually sweeps this under the carpet, since it is easier to think of elements of $L^p(X)$ as functions and not equivalence class of functions, and the context will make clear which of these we are dealing with.

The fundamental theorem about these spaces is,

Theorem 23. *$L^p(X)$ is a complete metric space for $1 \leq p \leq \infty$.*

We finally state some density properties. Notice that if s is a simple measurable functions with $\mu\{x | s(x) \neq 0\} < \infty$, then $s \in L^p$ for all $1 \leq p < \infty$.

Theorem 24. *Let S be the set of all measurable functions satisfying the above property. Then S is dense in $L^p(X)$ for $1 \leq p < \infty$.*

Next, we consider the case X to be an open or closed subset of $\mathbb{R}^n$ (or more generally a locally compact Hausdorff space, so that Riesz representation theorem is applicable) and the Lebesgue measure on it. Then, $C_c(X) \subset L^p(X)$ for $1 \leq p \leq \infty$.

Theorem 25. *$C_c(X)$ is dense in $L^p(X)$ for $1 \leq p < \infty$.*