Analysis, Probability, and Stochastic Calculus

Andrew Lin

July 2019

1 Zorm lemma

Definition. *X: a set with a relation $\leqslant$ on X is a partial order if*
(1) $\forall x \in X, x \leqslant x$
(2) $\forall x, x' \in X[x \leqslant x'$ *and* $x' \leqslant x \Rightarrow x = x']$
(3) $\forall x, x', x''[x \leqslant x' \leqslant x'' \Rightarrow x \leqslant x''$.

Definition. *A poset* $(X, \leqslant)$ *is a chain or totally ordered set if* $x, x' \in X[x \leqslant x'$ *or* $x' \leqslant x$.

Definition. *b is a maximal element in X if* $\forall x \in x, b \leqslant x$, *then* $b = x$.

Definition. $(X, \leqslant)$ *a chain. We say that* $(X, \leqslant)$ *is a well-ordered set if* $\forall A \subseteq X[A \neq \emptyset \Rightarrow A$ *has a least element*].

Theorem. *(a version of Bourbaki's fixed point theorem)* $(X, \leqslant)$ *a poset in which every well-ordered subset has least upper bound.* $X \overset{f}{X}$ *a map s.t.* $x \leqslant f(x)$ *for every* $x \in X$. $\exists a \in X, f(a) = a$.

Theorem. *(1) For any X,* $\exists P_0(X) \xrightarrow{f} X, \forall S \in P_0(X), f(S) \in S$.
(2) If X is a poset in which every well-ordered subset has a least upper bound in X, then X has a maximal element.
(3) Every poset has a maximal chain.
(4) If X is a poset, in which every chain has an upper bound in X, then X has a maximal element.
(5) Every set has a well-order.
(6) $\forall$ *surjection* $X \xrightarrow{f} Y, \exists Y \xrightarrow{g} X$ *s.t.* $f \circ g = id_Y$.
(7) Given sets S_α, *there exists* $A \xrightarrow{f} \cup_{\alpha \in A} S_\alpha$ *s.t.* $f(\alpha) \in S_\alpha$ *for all* $\alpha \in A$.

Proof. (7) $\Rightarrow$ (1) Let $P_0(X) = \{S | S \in P_0(X)\}$. Then (7) $\Rightarrow \exists A \xrightarrow{f} \cup_{\alpha \in A} S_\alpha$ s.t. $f(\alpha) \in S_\alpha$ then $S \in f(S)$.
(1) $\Rightarrow$ (2) Assume that X has no maximal element, $\forall a \in X, X_a = \{x \in X | a < x\} \neq \emptyset$. By (1), define a map k $a \longrightarrow X_a$ and $\exists P_0(X) \xrightarrow{f} X$ s.t. $f(S) \in S$ for all $S \subseteq X \neq \emptyset$. Let $g = f \circ k$. $\forall a \in X, g(a) = f(X_a) \in X_a$ and $a < g(a)$.

Contradictory to Boubaki's fixed point theorem.
(2) $\Rightarrow$ (3) Consider $X = \{C | C$ is a chain in P w.r.t. $\leqslant\}$ and thus is a poset with respect to $\subseteq$. We claim a stronger result: any totally ordered set in X has a lub in X. If $T \subseteq X$ is a totally ordered set, $\cup_{C\in T} C$ is $lub_x T$. By (2), X has the maximal element, i.e. a maximal chain in P.
(3) $\Rightarrow$ (4) By (3), $\exists$ maximal chain C. By assumption, C has an upper bound, say a, in X. Then a is a maximal element in X, otherwise, $\exists x \in X, a < x$, and hence $X \cup \{x\}$ is a chain. Contradictory!
(4) $\Rightarrow$ (5) Let Y be a set. Consider $X = \{A | A = (S_A, \overset{\leqslant}{A})$ where $S_A \subseteq Y$ and $\overset{\leqslant}{A}$ is a well-ordering on $S_A\}$. We define a relation $\preceq$ on X: $A \preceq A' \Leftrightarrow A = A'$ or A is an initial segment of A', i.e. $a' \in S_{A'}, S_A = \{x \in S_{A'} | x < a'\}$ and $\forall x_1, x_2 \in S_A, x_1 \overset{\leqslant}{A} x_2 \Leftrightarrow x_1 \overset{\leqslant}{A'} x_2$. It is direct to see that $\preceq$ is a partial order. We claim X has a maximal element w.r.t. $\preceq$ and a maximal element in X w.r.t. $\preceq$ is of the form $(Y, \preceq)$. We first verify the latter. If $(Y_0, \preceq)$ is a maximal element in X w.r.t. $\preceq$ and $Y_0 \neq Y$, then $\exists y \in Y \setminus Y_0$, and $Y_0 \cup \{y\}$ admits a well-ordering which makes an initial segment. We apply (4) to the former. Let C be a chain in X w.r.t. $\preceq$. Let $A_0 = (S_{A_0}, \overset{\leqslant}{A_0})$ where $S_{A_0} = \cup_{A\in C} S_A$ and $\overset{\leqslant}{A_0}$: For any $x_1, x_2 \in S_{A_0}$, find $A \in C$, s.t. $x_1, x_2 \in S_A$ we say that $x_1 \overset{\leqslant}{A_0} x_2$ if $x_1 \overset{\leqslant}{A} x_2$. Such A exists since C is a chain. In addition, $\overset{\leqslant}{A_0}$ is a total order and a well-ordering since let $T \subseteq S_{A_0} \neq \emptyset$. $T = T \cap S_{A_0} = T \cap \cup S_A = \cup (T \cap S_A$, $\exists A \in C, T \cap S_A \neq \emptyset$. $(S_A, \overset{\leqslant}{A})$ is a well-ordered set, $T \cap S_A$ has a least element in S_A, say t. Then t is also the least element of T in S_{A_0} w.r.t. $\overset{\leqslant}{A_0}$. Thus, A_0 is in X. And check A_0 is the upper bound of X.
(5) $\Rightarrow$ (6) Choose a well ordering $\leqslant$ on X. $For y \in Y$, define $g(y) =$ the least element of $f^{-1}(y)$. $f \circ g(y) = y$.
(6) $\Rightarrow$ (7) Consider $S = \cup_{\alpha \in A} S_\alpha$. Let $X = \{(s, \alpha) \in S \times A | s \in S_\alpha\}$. Let $X \xrightarrow{f} A$ and $X \xrightarrow{p} S$. f is surjective. By (6) $\exists A \xrightarrow{g} X$ s.t. $f(g(\alpha)) = \alpha$ for all $\alpha \in A$. Let $h = p \circ g$. Since $g(\alpha) \in X$ and $f(g(\alpha)) = \alpha$. $(s(\alpha), \alpha) \in X \Rightarrow s(\alpha) \in S_\alpha$. $h(\alpha) = p(g(\alpha)) = p(s(\alpha), \alpha) = s(\alpha) \in S_\alpha$ □

Theorem. *Every orthonormal set B in a Hilbert space H is contained in a maximal orthonormal set in H.*

Proof. Let P be the class of all orthonormal sets in H which contain the given set B. Partially ordered P by set inclusion. Since $B \in P, P \neq \emptyset$, P contains a maximal chain Ω. Let S be the union of all members of Ω. It is clear that $B \subset S$. We then claim that S is a maximal orthonormal set: If $u_1, u_2 \in S$, then $u_1 \in A_1$ and $u_2 \in A_2$ for some A_1 and $A_2 \in \Omega$. Since Q is totally ordered, $A_1 \subset A_2$, SO THAT $u_1 \in A_2$ and $u_2 \in A_2$. Since A_2 is orthonormal, $< u_1, u_2 >= 0$ if $u_1 \neq u_2, < u_1, u_2 >= 1$ if $u_1 = u_2$. Thus S is an orthonormal set.
Suppose that S is not maximal. Then S is a proper subset of an orthonormal

set S^*. Clearly, S^* not in Ω. We may adjoin S^* to Ω and still have a total order. Contradictory! $\square$

Theorem. *Let H be a Hilbert space, and let F be an orthogonal set in H. The following are equivalent:*
F is maximal among all the orthogonal subsets of H.
spanF is dense in X.

2 Topology

Definition. *A topological space $X = (\underset{-}{X}, T_X)$ consists of a set $\underset{-}{X}$, called the underlying space of X, and a family T_X of subsets of $\underset{-}{X}$ s.t. (1) $\underset{-}{X}$ and $\varphi \in T_X$, (2) $U_\alpha \in T_X(\alpha \in A), \cup_{\alpha \in A} U_\alpha \in T_X$, (3) $U, U' \in T_X, U \cap U' \in T_X$, T_X called a topology on $\underset{-}{X}$.*

Theorem. *Any two norms on a finite-dimensional vector space are equivalent.*

Lemma. *Let $(K, ||\cdot||)$ be a non-trivially valued field and V be a K-vector space. Two norms are equivalent if and only if there are constants $A > 0$ and $B > 0$ such that $A||v||' \leqslant ||v|| \leqslant B||v||'$ for all $v \in V$.*

Proof. (Proof of lemma)
The lemma is obvious if $V = \{0\}$, so assume that V is not $\{0\}$.
($\Leftarrow$) First assume that there are positive A and B such that $A||v||' \leqslant ||v|| \leqslant B||v||'$ for all $v \in V$. Then for any open set $U \subset V$ w.r.t. $||\cdot||$ and $v \in U$, there is an $\epsilon > 0$ s.t. the open ϵ-ball w.r.t $||\cdot||$ around V is contained in U: $\{\omega \in V : ||\omega - v|| < \epsilon\} \subset U$. Since $||\omega - V||' < \frac{\epsilon}{B} \Rightarrow ||\omega - v|| < \epsilon \Rightarrow$ any open $||\cdot||$-ball around v contains an open $||\cdot||'$-ball around v, so U is open w.r.t. $||\cdot||$. The contrary relation holds if using $||v||' \leqslant \frac{1}{A}||v||$.
($\Rightarrow$) Assume that $||\cdot||$ and $||\cdot||'$ are equivalent. Then the open unit ball around origin in V relative to $||\cdot||$ is open relative to $||\cdot||'$ and the open unit ball around the origin in V relative to $||\cdot||'$ is open relative to $||\cdot||$, so there are $r > 0, s > 0$ such that $\{v \in V : ||v||' < r\} \subset \{v \in V : ||v|| < 1\}$, $\{v \in V : ||v|| < s\} \subset \{v \in V : ||v||' < 1\}$. Therefore, for each nonzero $v \in V$, there exists $\gamma \in K$ s.t. $|\gamma|^n \leqslant \frac{1}{s}||v|| \leqslant |\gamma|^{n+1}$. Then $||\frac{1}{\gamma^{n+1}}v|| = \frac{1}{|\gamma^{n+1}|}||v|| < s$, so $||\frac{v}{\gamma^{n+1}}||' < 1$. Thus $||v||'|| < |\gamma|^{n+1} \leqslant \frac{|\gamma|}{s}||v||$. By setting $B = |\gamma|/s$, we have $||v||' < B||v||$ for all nonzero $v \in V$, so $||v||' \leqslant B||v||$ for all v. The similar conclusion holds using r in replace of s. In that result, $A = \frac{r}{|\gamma|}$.
(Proof of theorem)
Choose arbitrary two norms $||\cdot||_a, ||\cdot||_b$.
We can claim that it is sufficient to consider $||\cdot||_b$ equivalent to $||\cdot||_1$ by transitivity. First define an L_1-style norm by $||x||_1 = \sum_{i=1}^n |a_i|$. Suppose both $||\cdot||_a$ and $||\cdot||_{a'}$ are equivalent to $||\cdot||_1$ for constants $0 < C_1 \leqslant C_2$ and $0 < C_1' \leqslant C_2'$,

respectively:

$$C_1||x||_1 \leqslant ||x||_a \leqslant C_2||x||_1,$$
$$C_1'||x||_1 \leqslant ||x||_a \leqslant C_2'||x||_1,$$

It immediately follows that

$$\frac{C_1'}{C_2}||x||_a \leqslant ||x||_{a'} \leqslant \frac{C_2'}{C_1}||x||_a,$$

and hence $||\cdot||_a$ and $||\cdot||_{a'}$ are equivalent.
Next, we claim that it is sufficient to consider only x with $||x||_1 = 1$ since the vector space is equipped with scalar multiplication.
Next, we claim that any norm $||\cdot||_a$ is continuous under $||\cdot||_1$. By the triangle inequality on $||\cdot||_a$, it follows that $|||x'||_a - ||x||_a| \leqslant ||x' - x||_a$. And applying the triangle inequality again, and writing $x = \sum_{i=1}^n a_i e_i$ and $x' = \sum_{i=1}^n a_i' e_i$, we can obtain

$$||x - x'||_a \leqslant \sum_{i=1}^n |a_i - a_i'| \cdot ||e_i||_a \leqslant ||x - x'||_1(\max_i ||e_i||_a).$$

Therefore, if we choose $\delta = \frac{\epsilon}{\max_i ||e_i||_a}$, it immediately follows that

$$||x - x'||_1 < \delta \Rightarrow |||x||_a - ||x'||_a| \leqslant ||x - x'||_a < \epsilon.$$

It is a standard theorem of analysis, the extreme value theorem, that a continuous functin on compact set must achieve a maximum and minimum value on the set. Let

$$C_1 = \min_{||u||_1=1} ||u||_a,$$
$$C_2 = \max_{||u||_1=1} ||u||_a,$$

Since $u \neq 0$ for $||u||_1 = 1$, it follows that $C_2 \geqslant C_1 > 0$ and $C_1 \leqslant ||u||_a \leqslant C_2$ as required by the previous step. □

Definition. *Let X and Y be topological spaces and $\underset{-}{X} \xrightarrow{f} \underset{bar}{Y}$ a map. We say that f is continuous at point $x_0 \in X$ if $\forall V \in T_Y, \exists U \in T_X$ s.t. $F(U) \subseteq V$*

Lemma. *f: continuous$\Leftrightarrow \forall V \in T_Y, f^{-1}(V) \in T_X$*

Definition. *X: a top. space, $K \subseteq \underset{bar}{X}$. K is compact in X if $\forall U_\alpha \subset_{open} X(\alpha \in A)$, if $K \subseteq \cup_{\alpha \in A} U_\alpha$, $\exists$ finite set $S \subseteq A$ s.t. $K \subseteq \cup_{\alpha \in S} U_\alpha$.*

Proposition 1. *$X \xrightarrow{f} Y$ and f is continuous, K is compact in X, $f(K) \subseteq Y$ is compact.*

Proof. For $V_\alpha \subseteq_{open} Y(\alpha \in A)$ s.t. $f(K) \subseteq \cup_{\alpha\in A} V_\alpha$, we have $K \subseteq f^{-1}(f(K)) \subseteq f^{-1}(cup_{\alpha\in A} V_\alpha) = \cup_{\alpha\in A} f^{-1}(V_\alpha)$. By the above lemma, $f^{-1}(V_\alpha) \subseteq_{open} X$. Since K is compact, $\exists$ finite set $S \subseteq A$ s.t. $K \subseteq \cup_{\alpha\in S} f^{-1}(V_\alpha) \Rightarrow f(K) \subseteq f(f^{-1}(\cup_{\alpha\in S} V_\alpha)) = \cup_{\alpha\in S} V_\alpha$. □

Theorem. *(Heine Borel theorem) K: cpt in X. K is bounded and closed in X.*

Proof. $K \subseteq \cup_{r>0} B_r(a)(\forall a \in X$. Hence, K is bounded. Fix any x not in K. For any y in K, $\exists U_y$ which includes y open subset to X and V_y which includes X and open subset to X s.t. $U_y \cap V_y = \varphi$. $K = \cup_{y\in K}\{y\} \subseteq \cup_{y\in K} U_y \Rightarrow \exists$ finite set $S \subseteq K$ s.t. $K \subseteq \cup_{y\in S} U_y$. Let $V = \cap_{y\in S} V_y, x \in V \subseteq_{open} X$. $V \cap K \subseteq \cap_{y\in S} V_y \cap (\cup_{z\in S} U_z) = \cup_{z\in S}(\cap_{y\in S} V_y \cap U_z) = \varphi$. Conclusion: $forallx$ not in K, $existsV_x \subseteq_{open} X$ s.t. $V_x \subseteq X \setminus K$. $X \setminus K = \cup_{x\in X\setminus K} V_x \subseteq_{open} X$. Thus K is closed. □

Corollary. *Finite-dimensional vector subspaces of a normed vector space are all closed.*

Proof. For definiteness, assume that E is a real n.v.s. with norm $||\cdot||$. Consider any finite-dimensional vector subspace F of E, put n=dimension of F and choose a basis $v_1, ..., v_n$ of F. Define a new norm $||\cdot||'$ on F as follows: for $u = \sum_{j=1}^n \alpha_j v_j$ where $\alpha_1, \alpha_2, ..., \alpha_n$ are real numbers, let $||u||' = (\sum_{j=1}^n \alpha_j^2)^{\frac{1}{2}}$. Clearly, $||\cdot||'$ is a norm on F. Let T be the linear map from the Euclidean space R^n onto F, defined by $Tx = \sum_{j=1}^n x_j v_j$ for $x = (x_1, ..., x_n)$. If we denote the Euclidean norm by $|\cdot|$, then $||Tx||' = |x|$. Then by the previous theorem, there is $c > 0$ such that $c||u||' \leqslant ||u|| \leqslant c^{-1}||u||'$ for $u \in F$. Consequently, $||Tx|| \leqslant c^{-1}|||Tx||' = c^{-1}|x|$ for $x \in R^n$. and hence T is a continuous map from R^n into E. We arbitrarily choose a sequence u_k in F that converges in E. Since the sequence converges, it is bounded, say $||u_k|| \leqslant A$ for all k for some A¿0. Now write $u_k = \sum_{j=1}^n \alpha_j^{(k)} v_j$ and put $\alpha^{(k)} = (\alpha_1^{(k)}, ..., \alpha_n^{(k)})$, then $u_k = T\alpha^{(k)}$ and $|\alpha^{(k)}| = ||u_k||' \leqslant c^{-1}||u_k|| \leqslant c^{-1}A$ for each k. Thus u_k is contained in the image $K \subset F$ of the closed ball $x \in R^n : |x| \leqslant c^{-1}A$ under T. Since closed ball in R^n are compact, K is compact by the above proposition and is therefore closed in E. Now $u_k \subset K$ implies that its limit is in $K \subset F$. This shows that F is closed. □

Definition. *A topological space X is*
first countable if $\forall x \in X, \exists$ *countable local basis at x. Ex: X: metric space. Choose* $B_r(x)|r > 0, r \in Q$*. Hence, it is easily to see that a metric space is first countable.*
second countable if $\exists$ *countable basis of the topology.*
separable if it contains a countable dense subset.
Lindelöf if every open cover has a countable subcover.

The 2nd countability can deduce to 1st countability, separability and Lindelöf.

3 Sequential descriptions of several notions in metric spaces

(1) Let (X, d) be a metric space. For any $A \subseteq X$, we have
$\bar{A} = \{x \in X | \exists$ sequence $x_n \in A(n \in N)$ s.t. $x_n \to x$ as $n \to \infty\}$.

Proof. (Necessary condition) Suppose there exists a sequence $x_n \in A(n \in N)$ s.t. $x_n \to x$ as $n \to \infty\}$. Let $\epsilon > 0$. Then by definition: $\exists N \in N : \forall n > N :$ $x_n \in B_\epsilon(x)$. Since $\forall n : x_n \in A$, it follows that: $\forall \epsilon > 0 : B_\epsilon(x) \cap A \neq \varphi$. Hence $x \in \bar{A}$. (Sufficient condition) Now suppose $x \in \bar{A}$. By definition of closure: $\forall n \in N : \exists x_n \in A \cup B_{\frac{1}{n}}(x)$. Thus clearly x_n converges to x. □

(2) all limit points of A in X=$\{x \in X | \exists$ sequence $a_n \in A \setminus \{n\}(n \in N)$ s.t. $a_n \to x$ as $n \to \infty$.
(3) Let $X \xrightarrow{f} Y$ be a map between metric spaces and $x_0 \in X$. We have f is continuous$\Leftrightarrow \forall x_n \in X(n \in N), x_n \to x$ as $n \to \infty \Rightarrow f(x_n) \to f(x)$.

Remark. *If X is a topological space instead of a metric space, $x \in X$ and $x_n \in X(n \in N)$, we may define $x_n \to x$ as $n \to \infty$ to mean that $\forall$ open neighborhood U of $x \in X$, $\exists N$ s.t. $\forall n \in N, n \geqslant N \Rightarrow x_n \in U$. Then in (1)(2)(3), the left can deduce the right. The converse holds if X is first countable. Regard to (1), first countability ensures that every point has a countable local basis $U_1, U_2 ....$ Then we can construct a decreasing set sequence $U_1, U_1 \cap U_2, ...$ And by closure property, we can choose a point from each intersection of A and the element of the above set sequence. Thus, we construct the point sequence as required.*

Definition. *(X,d): metric space.*
(1) (X,d) is sequentially compact if every sequence has a convergent subsequence.
(2) (X,d) is totally bounded if $\forall \epsilon > 0, \exists$ finite set $S \subseteq X$ s.t. $X = \cup_{s \in S} B_\epsilon(s)$.

Remark. *Total boundedness⇒separability.*

Proof. More precisely, for any $n \in N$, there exists a finite set $S_n \subseteq X$ s.t. $X = \cup_{s \in S_n} B_{\frac{1}{n}}(s)$. Then $S := \cup_{n=1}^{\infty} S_n$ which is a countable dense subset in X w.r.t. d. □

Theorem. *A space X is compact if and only if every collection of closed subsets of X satisfying the finite intersection property has non-empty intersection.*

Proof. $\Rightarrow$ Let X be compact. Let C be a collection of closed subsets of X. We show that if C has the finite intersection property, then it has non-empty intersection. Suppose that $\cap C = \varphi$. Then $U = \{X - C : C \in C\}$ is an open cover of X. By the compactness of X, U has a finite subcover $\{X - C_0, X - C_1, ..., X - C_n\}$. Contradictory.
$\Leftarrow$ Let U be an open cover of X such that it has no finite subcover. Note that C has the finite intersection property but $\cap C = \varphi$. Contradictory. □

Proposition 2. *(X,d): metric space. The following are equivalent.*
(1) X is compact
(2) X is sequentially compact
(3) X is totally bounded and complete.

Proof. (1)⇒(2) Suppose that $\exists x_n (n \in N), \forall x \in X$, x is not the limit of any subsequence of x_n. Thus for any $x \in X, \exists$ open neighborhood U_x of x in X s.t. $n \in N | x_n \in U_x$ is finite. $X = \cup_{x \in X} U_x \overset{X:cpt}{\Longrightarrow} \exists p_1, ..., p_m$, $X = U_{p_1} \cup ... \cup U_{p_m}$. N=$n \in N | x_n \in X$=$\cup_{j=1}^{m} \{n \in N | x_n \in U_{p_j}\}$. Contradictory!
(2)⇒(3) (Proof of completeness) A Cauchy sequence converges to x if and only if it has a subsequence that converges to x. (Necessary condition) If a Cauchy sequence x_n converges to x, it trivially follows that x_n is a subsequence to itself that converges to x. (Sufficient condition) Suppose that x_{n_k} is a subsequence of x_n that converges to x. Let $\epsilon > 0$. By the definition of a Cauchy sequence, there exists a positive integer M such that: $\forall i, j \in N : i, j \geqslant M \Rightarrow d(x_i, x_j) < \frac{\epsilon}{2}$. By the definition of convergence, there exists a positive integer N such that: $\forall k \in N : k \geqslant N \Rightarrow d(x_{n_k}, x) < \frac{\epsilon}{2}$. There exists a natural number $K > \max\{M, N\}$. Therefore, by the triangular inequality: $forallm \in N : m > K \Rightarrow d(x_m, x) \leqslant d(x_m, x_K) + d(x_K, x) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. That is x_n converges to x. Thus, it soon follows that if (X,d) is sequentially compact, then it is complete.
(Proof of totally-boundedness) If X is not totally bounded, there exists $\epsilon > 0$ and $x_1, ...$ s.t. $d(x_i, x_j) \geqslant \epsilon$ if $i \neq j$. Futhermore, any subsequence of the above x_n does not converge. Contradictory!
(3)⇒(2) Let $x_n (n \in N)$ be a sequence in X. (X,d) totally bounded⇒For any given $n \in N$, X can be covered by finitely many $\frac{1}{n}$-balls.⇒ $\exists$ 1-ball B_1 s.t. $n \in N | x_n \in B_1$ is infinite.⇒.....⇒ subsequence $x_{n_k} \in B_1 \cap ... \cap B_k$ for every $k \in N$. In fact, for every k and $l, l' \geqslant k$, we have $d(x_{n_l}, x_{n_{l'}})$. Thus, the subsequence is Cauchy and complete.
(2)⇒(1) Let F be a family of closed subsets of X which satisfies the FIP. We need to show that $\cap F \neq \varphi$. Suppose not $\cap F = \varphi \Rightarrow \{X \setminus C | C \in F\}$ is an open cover of X⇒ $\exists C_1, C_2, ... \in F$ s.t. $X \setminus C_n | n \in N$ still covers X⇒ $\cap_{n=1}^{\infty} C_n = \varphi$. Contradictory. We can choose $x_1 \in C_1, ...$ By sequentially compact, there exists a subsequence x_{n_k} that converges. Therefore, $\cap_{n=1}^{\infty} C_n \neq \varphi$. □

Definition. *Let X be a topological space and Y be a metric space. A family F of maps from X to Y is equicontinuous at a point $x_0 \in X$ if $\forall \epsilon > 0, \exists$ open neighborhood U of x_0 s.t. $\forall f \in F$ and $x \in X$ $x \in U \Rightarrow d(f(x), f(x_0)) < \epsilon$.*

Theorem. *If T is a continuous map a compact metric space M_1 into a metric space M_1, then T is uniformly continuous on M_1.*

Proof. Let $\epsilon > 0$ be given, and $x \in M_1$. Since T is continuous at x, there is $\delta_x > 0$ s.t. $\rho_2(Ty, Tx) < \frac{\epsilon}{2}$ if $\rho_1(y, x) < \delta_x$. Consider $B_{\frac{\delta_x}{2}}(x)$ which is an open cover of M_1. Since M_1 is compact, $\exists$ a finite subcover $B_{\frac{\delta_{x_1}}{2}}(x_1), ..., B_{\frac{\delta_{x_l}}{2}}(x_l)$. Choose $\delta = \frac{1}{2} \min\{\delta_{X_1}, ...\delta_{x_l}\}$. Suppose that $x, y \in M_1$ with $\rho_1(x, y) < \delta$. Since $x \in M_1$, $x_1 \in B_{\frac{\delta_{x_j}}{2}}(x_j), 1 \leqslant j \leqslant l$. Then $\rho_1(y, x_j) \leqslant \rho_1(y, x) + \rho_1(x, x_j) = \delta + \frac{\delta_{x_j}}{2} < \delta_{x_j}$.

Therefore, $\rho_2(Ty, Tx_j) < \frac{\epsilon}{2}$. Then, $\rho_2(Tx, Ty) \leqslant \rho_2(Tx, Tx_j) + \rho_2(Ty, Tx_j) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. ☐

Theorem. *Let f_n be a sequence of real-valued continuous functions defined on a compact metric space M such that $f_1(x) \leqslant f_2(x)$... and converges to a finite real number f(x) for each $x \in M$. If further, f is continuous on M, then the sequence f_n converges uniformly to f on M. If, further, f is continuous on M, then the sequence f_n converges uniformly to f on M.*

Proof. Given $\epsilon > 0$ and $x \in M$, there is $k_x \in N$ such that $0 \geqslant f(x) - f_{k_x}(x) < \frac{\epsilon}{3}$. Because both f and f_{k_x} are continuous, there is an open ball B(x) centered at x such that $|f(y) - f(x)| < \frac{\epsilon}{3}$ and $|f_{k_x}(y) - f_{k_x}(x)| < \frac{\epsilon}{3}$ whenever $y \in B(x)$. As the result, we have $0 \leqslant f(y) - f_{k_x}(y) \leqslant |f(y) - f(x)| + |f(x) - f_{k_x}(x)| + |f_{k_x}(x) - f_{k_x}(y)| < \epsilon$ whenever $y \in B(x)$. Now $B(x) : x \in M$ is an open covering of M. There exists a finite subfamily that also covers M, say $B_(x_1), ..., B_(x_l)$. Let $k_0 = \max\{k_{x_1}, ..., k_{x_l}\}$. Then for $y \in M$ and $k \geqslant k_0$, it follows that $0 \leqslant f(y) - f_k(y) < \epsilon$. ☐

Lemma. *Let f_n be a sequence of continuous functions defined on a metric space M. Suppose that f_n converges uniformly to a function f on M, then f is continuous on M.*

Proof. There exists n_0 s.t. $|f_{n_0}(x) - f(x)| < \frac{\epsilon}{3}$ for all x in M. In addition, f_{n_0} is continuous, for x,y in M with $\rho(x, y) < \delta$, $|f_{n_0}(y) - f_{n_0}(x)| < \frac{\epsilon}{3}$. $|f(x) - f(y)| \leqslant |f(x) - f_{n_0}(x)| + |f_{n_0}(y) - f_{n_0}(x)| + |f(y) - f_{n_0}(y)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$. ☐

Definition. *C(X): All continuous valued functions defined on a compact metric space X with norm given by $||f|| = \sup_{x \in X} |f(x)|$.*

Proposition 3. *C(X) is a Banach space.*

Proof. Let f_n be a Cauchy sequence in C(X). $|f_n(x) - f_m(x)| \leqslant ||f_n - f_m||$ for $x \in X$. Given $\epsilon > 0$, there is $n_0 \in N$ such that $||f_n - f_m|| < \epsilon$ whenever $n, m \geqslant n_0$, hence $|f_n(x) - f_m(x)| < \epsilon$ for all x in X and $n, m \geqslant n_0$, and thus $|f_n(x) - f(x)| < \epsilon$ for all x in X and $n, m \geqslant n_0$, and thus $|f_n(x) - f(x)| \leqslant \epsilon$ for all x in X if $n \geqslant n_0$, by letting $m \to \infty$. It follows that $f \in C(X)$. In addition, $\sup_{x \in X} |f_n(x) - f(x)| \leqslant \epsilon$ when $n \geqslant n_0$, or $||f_n - f|| \leqslant \epsilon$ when $n \geqslant n_0$. Thus, $\lim_{n \to \infty} ||f_n - f|| = 0$. ☐

Theorem. *(A generalization of Ascoli's theorem) Let X be a topological space and F a family of real-valued functions on X. If (1) X is separable, (2) F is equicontinuous everywhere on X and (3) $\forall x \in X$ $f(x)|f \in F$ is a bounded subset of R, then every sequence in F has a subsequence which converges compactly, i.e. uniformly on every compact subset of X.*

Proof. Let $A = \{a_1, a_2, ...\}$ be a countable dense subset. Suppose that $f_n(n \in N)$ be a sequence in F.
Claim 1: $\exists$ subsequence $f_{n_m}(m \in N)$ which converges pointwise on A such that $f_n(a_1)|n \in N \subseteq f(a_1)|f \in F \subseteq_{bounded} R. \Rightarrow n_m^{(1)}(m \in N)$ s.t. $f_{n^1}(a_1)$ converges.

Inductively, we can construct n^j_m s.t. (1) n^j strictly increasing, (2) $n^{(j)}_m \subseteq n^{(j-1)}_m$, and (3) $f_{n^{(j)}_m}(a_j)$ converges. Let $n_m := n^{(m)}_m$. Then $f_{n_m}(m = k, k+1, ...)$ is a subsequence of $f_{n^{(k)}_m}$, and hence $f_{n_m}(a_k)$ converges as $m \to \infty$.

Claim 2: $\forall \epsilon$ and $x \in X$, $\exists$ open neighborhood U_x of x and a number $N_x > 0$ s.t. if $x' \in U_x$ and $k, l \geqslant N_x$, $|f_{n_k}(x') - f_{n_l}(x')| < \epsilon$. F is equicontinuous at x, for any $\epsilon > 0$, $\exists$ open neighborhood U_x of x s.t. $|f(z) - f(x)| < \frac{\epsilon}{6} \forall z \in U_x$. Since A is dense in X, $\exists a \in U_x \cap A$. FOR any $x' \in U_x$, we have $|f_{n_k}(x') - f_{n_l}(x')| \leqslant |f_{n_k}(x') - f_{n_k}(x)| + |f_{n_k}(x) - f_{n_k}(a)| + |f_{n_k}(a) - f_{n_l}(a)| + |f_{n_l}(a) - f_{n_l}(x)| + |f_{n_l}(x) - f_{n_l}(x')| \leqslant |f_{n_k}(a) - f_{n_l}(a)| + \frac{2\epsilon}{3} < \epsilon$.
Claim 3: $\forall K$ is compact in X. $f_{n_m}|_K$ converges uniformly. For an given $\epsilon > 0$, we have found U_x and N_x as in Claim 2. $K = \cup_{x \in K}$x$\subseteq \cup_{x \in K} U_x \Rightarrow \exists x_1, ..., x_p \in K, K \subseteq U_{x_1} \cup ... \cup U_{x_p}$. Let $N = \max\{N_{x_1}, ..., N_{x_p}\}$. Then for any $q \in K$ and $k, l \geqslant N$, we have $|f_{n_k}(q) - f_{n_l}(q)| < \epsilon$. □

Definition. *Let X be a topological space and $A \subseteq X$. A is relatively compact if $\bar{A}$ is compact.*

By Ascoli's theorem, $F \subseteq C(x, R)$ is equicontinuous and uniformly bounded $\Rightarrow$ F is relatively compact in $(C(X, R), d_{sup})$.

Theorem. *A subset A of a metric space M is totally bounded if and only if every sequence in A has a Cauchy subsequence. In particular, compact sets are totally bounded.*

Proof. Suppose A is totally bounded and let x_n be a sequence in A. There is a $\frac{1}{2}$ net for A and hence ne of its balls contains a subsequence $x^{(1)}_n$. After the subsequnce is chosen, we then choose a
$frac14$ net for A and construct $x^{(2)}_n$ from $x^{(1)}_n$. Now, $x^{(n)}_n$ is a subsequence of x_n. For each positive integer n_0, if $n > m \geqslant n_0$, both $x^{(n)}_n, x^{(m)}_m$ are in a ball of radius 2^{-n_0}, hence $d(x^{(n)}_n, x^{(m)}_m) \leqslant 2^{-n_0+1}$, form which it follows that $x^{(n)}_n$ is a Cauchy subsequence of x_n.
Next, suppose that each sequence of A has a Cauchy subsequence. Suppose to the contrary that for some $\epsilon_0 > 0$, no ϵ_0 net for A exists. Then, we could find a sequence whose points are distanced from each other farther than ϵ_0. Hence, the sequence has no Cauchy subsequence. Contradictory. □

Theorem. *(Arzelà-Ascoli theorem) The converse of Ascoli's theorem is true.*

Proof. Suppose K is relatively compact. Since C(X) is complete, K is totally bounded. Let $\epsilon > 0$ and let $f_1, ..., f_n$ be the center of an $\frac{\epsilon}{3}$ net for K. Since X is a compact space, $f_1, ..., f_n$ are uniformly continuous on X, there is $\delta > 0$ such that $|f_i(x) - f_i(y)| < \frac{\epsilon}{3}$ for i=1,...,n when $d(x, y) < \delta$. Consider now $f \in K$, there exists $j \in \{1, ...n\}$ s.t. $\sup_{x \in X} |f(x) - f_j(x)| < \frac{\epsilon}{3}$. $|f(x) - f(y)| \leqslant |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_j(y) - f(y)| < \epsilon$. So K is equicontinuous. Since K is totally bounded, it is hence bounded. □

4 Partitions of unity and paracompactness

Definition. *For any R-valued function f on X, we define its support* $suppf := \{x \in X | \bar{f}(x) \neq 0\}$*.* $X \setminus suppf = \{x \in X | F(X) = 0\}$*; in other words, for any U open in X and* $f|_U = 0 \Leftrightarrow U \subseteq X \setminus U$ *i.e.* $U \cap suppf = \emptyset$*.*

Temporary notation
For R-valued functions f on A and g on B with $A, B \subseteq X$, we define $(f \cdot g)(x) = f(x)g(x)$ if $x \in A \cap B$, $= 0$ if $x \in X \setminus (A \cap B)$. Note that $supp(f \cdot g) \subseteq suppf \cap suppg$.

Lemma. *Given* $U \subseteq_{open} X, f \in C(U)_R$ *and* $\rho \in C(X)_R$*, if* $supp\rho \subseteq U$*, then* $\rho \cdot f \in C(X)$*.*

Proof. $supp\rho \subseteq U \Leftrightarrow \{U, X \setminus supp\rho\}$ is an open cover of X. $(\rho \cdot f)|_U = \rho|_U \cdot f \in C(U)_R$. $(\rho \cdot f)|_{X \setminus supp\rho} = 0 \in C(X \setminus supp\rho)$. □

The setting can be expanded to smoothness or continuous differentiability.

Definition. *(refinements of open covers/paracompactness/partitions of unity)*
(1) Let $U_k (k \in K)$ *and* $V_j (j \in J)$ *be an open cover of X. We say that V is a refinement of U if* $\exists$ *a map* $J \xrightarrow{k(\cdot)} K$ *s.t.* $\forall j \in J, V_j \subseteq U_{k(j)}$*.*
(2) A family S_j *of subsets of X is strongly locally finite.* $\forall p \in X, \exists$ *neighborhood W of p in X s.t.* $j \in J | W \cap S_j \neq \emptyset$ *is finite.*
(3) X is paracompact if X is Hausdorff and $\forall$ *open cover of X,* $\exists$ *strongly locally finite refinement V of U.*
(4) Let U_k *be an open cover of X and* $X_j \xrightarrow{\rho_j} R$ *a family of functions on X. We say that* ρ_j *is a continuous partition of unity on X subordinate to U if* $\exists$ *strongly locally finite refinement* V_j *of U, s.t.* $supp\rho_j \subseteq V_j$ *for every* $j \in J$*. In addition,* $\rho_j \geqslant 0$*. And* $\sum_{j \in J} \rho_j = 1$ *for every* $p \in X$*.*

Proposition 4. *X: paracompact,* $A \subseteq_{closed} X \Rightarrow$*A: paracompact.*

Proof. If W: w_k is an open cover of A, say $W_k = U_k \subseteq_{open} X$, U_k together with $X \setminus A$ form an open cover of X. Thus, $\exists$ a finite refinement V_j. Then $V_j \cap A$ is an expected refinement of W. □

Proposition 5. *Paracompactness implies normality.*

Proof. (Proof of regularity)
By Hausdorffness, we can define, for every $y \in A$ (A is a closed set of X and does not contain x), open subsets U_y, V_y s.t. $x \in U_y, y \in V_y$, and $U_y \cap V_y = \emptyset$. Then, $A \subseteq \cup_{y \in A} V_y$. The sets V_y and $X \setminus A$ form an open cover of X. Thus, by paracompactness of X, there is a locally finite open refinement. Throwing out from this any open subset not intersecting A, we still get a locally finite collection P of open subsets, each contained in some V_y that cover A. There exists an open set W containing x such that there are only finitely many members of P that intersect W. Let T be a finite subset of A that contains, for each of this finite list of members of P, a point y s.t. that member is contained in V_y.

Define $U = W \cap \cap_{y\in T} U_y$ and V to be the union of all the members of P. Then, $x \in U, A \subseteq V$, and U and V are disjoint: For this, note that all the members of P that intersect W are contained in V_ys, which are disjoint from the corresponding U_ys. So, U is disjoint from V.
(Proof of normality)
For every $a \in A$, there exist open sets $a \in U_a$, V_a containing B, such that U_a and V_a are disjoint. The U_as form a collection of open subsets of X covering A. Along with $X \setminus A$, these form an open cover of X. This has a locally finite open refinement. Throwing out from this any open subset not intersecting A, we still get a locally finite collection Q of open subsets, each contained in some U_a, that cover A. Let C be the union of all members of Q. For any $b \in B$, there exists an open subset D_b around b that does not intersect C: First, there exists an open subset W_b around b intersecting finitely many members of Q. Let T be a finite subset of A that contains for each of this finite list of members of Q, a point a such that that is contained in U_a. Then, $D_b = W_b \cap \cap_T V_a$ works. □

Theorem. *If X is a Hausdorff space, X is paracompact ⇔ every open cover U of X admits a partition of unity subordinate to it.*

Proof. (⇒) Possibly by replacing U with a strongly locally finite refinement, we may assume U to be such. We first set up some terminology: a U-admissible collection is a family of functions $\varphi_U | U \in J$ s.t. (1) the index set $J \subseteq U$, (2) $X \xrightarrow[continuous]{\varphi_U} [0,1]$ for every $U \in J$, $supp\varphi_U \subseteq U$, and (3) $\varphi_U^{-1}(0,1]$ together with $U \setminus J$ form a open cover of X. Let A=all U-admissible collection be partially ordered by $\subseteq$. Then A admits a maximal chain $C \subseteq A$. Let $C_0 = \cup C$, saying we may write $C_0 = g_U | U \in J_0$, for some $U_0 \subseteq U$. We then claim that C_0 is still a U-admissible collection.
It is straightforward that C_0 hold automatically. If C_0 is not in A, then (3) is violated, i.e. $\exists x \in X$ s.t. x not in $g_U^{-1}(0,1]$ and x not in U'. Let $U_1, U_2, ...U_n$ be the only members of U which contains x since U is locally finite. Then, $U_1, U_2, ...U_n$ must be in J_0. Since C is a chain, $\exists C \in C$ of the form $g_U | U \in J_1$ s.t. $U_1, ...U_n \in J_1$; in particular, $x \in X = (\cup_{U\in J_1} g_U^{-1}(0,1]) \cup \cup_{W \in U\setminus J_1} W$, and hence $x \in \cup_{U \in J_1} g_U^{-1}(0,1]$.
Then, we claim that $J_0 = U$.
Suppose that $\exists U_0 \in U \setminus J_0$ and let $Y := (\cup_{U\in J_0} g_U^{-1}(0,1]) \cup_{W \in U \setminus J_0 \setminus U_0} W$. Then $X = U_0 \cup Y$ by (3). Let $Z = X \setminus Y \subseteq U$, a closed subset. By the previous proposition, X is normal and hence $\exists$ open neighborhood V of Z in X s.t. $\bar{(}V) \subseteq U_0$. By Urysohn's construction, $\exists X \xrightarrow[continuous]{f} [0,1]$ s.t. $f|_Z = 1$ and $f|_{X\setminus V} = 0$. Then $Z \subseteq f^{-1}(0,1]$, and Hence X=$f^{-1}(0,1] \cup Y$. Let $g_{U_0} = f$. Then $g_U | U \in J_0 \cup U_0$. Contradiction.
In summary, $C_0 = g_U | U \in U$ is a U-admissible collection, and hence $X = \cup_{U \in U} g_U^{-1}(0,1]$. It suffices to take $\rho_U = \frac{g_U}{g}$ where $g = \sum_{U \in U} g_U$. □

Lemma. *If X is paracompact, then for every locally finite open cover U_j of X, there exists an open cover V_j s.t. $\bar{V_j} \subseteq U_j$.*

Proof. Choose a partition of unity $\rho_U|U \in U$. Let $V_j = \rho_{U_j}^{-1}(0,1]$. $\bar{V_j} \subseteq U_j$. Since $\sum_{j\in J} \rho_j = 1$, $X \subseteq \cup V_j$. □

Lemma. *Let X be a locally compact Hausdorff space. For any compact subset* $K \subseteq X$ *and open* $V_1 ... V_n$ *if* $K \subseteq V_1 \cup ... \cup V_n$*, then exists* ρ_j *s.t.* $supp\rho_j \subseteq V_j$ *and* $\forall x \in K$, $\sum_{j=1} \rho_j(x) = 1$. *K is compact,* $\exists x_1, ..., x_m \in K$*, s.t.* $K = W_{x_1} \cup ... \cup W_{x_m}$*. Let* $H_i = \cup \bar{W_{x_j}}$*, which is compact. By Urysohn's construction,* $\exists$ *a continuous* φ_i *s.t.* $\varphi_i|_{H_i} = 1$ *and* $supp\varphi_i \subseteq V_i$*. Let* $\rho_1 = \varphi_1, \rho_2 = (1-\varphi_1)\varphi_2 ....\rho_n = (1-\varphi_1)...(1-\varphi_{n-1})\varphi_n$*. Then,* $supp\rho_k \subseteq \varphi_k \subseteq V_k$*. In addition,* $\rho_1 + ...\rho_n = 1-(1-\varphi_1)...(1-\varphi_n)$*. For any* $x \in K$*, since* $X \in \bar{W_x} \subseteq\subseteq V_{i_x}, x \in H_{i_x}$*, and hence* $\varphi_{i_x} = 1$.

Proof. For each $x \in K$, choose $i_x \in 1, ..., n$ s.t. $x \in V_{i_x}$. X: locally compact Hausdorff⇒ $\exists$ open neighborhood W_x of x s.t. $\bar{W_x} \subseteq V_{i_x}$ □

5 Measure Theory and Probability

Definition. *A family A of subsets of* Ω *is called an algebra of* Ω *if* $\Omega \in A$*, if* $A \in A$*, then* $A^c \in A$*.* $A \cup B \in A$ *whenever* $A, B \in A$*.*

Definition. *A family P of subsets of* Ω *is called a* π*-system on* Ω *if* $A \cap B \in P$ *if A,B in P.*

Definition. *A family L of subsets of* Ω *is called a* λ*-system on* Ω *if* $\Omega \in L$*,* $A \in L$*, then* $A^c \in L$*, if* A_n *is a disjoint sequence in L, then* $\cup_n A_n \in L$*.*

Lemma. *A family is both a* π*-system and a* λ*-system is a* σ*-algebra.*

Proof. It suffices to show that the family, say S is closed under countable-unions. Let $A_1, ... \in S$. We want to prove that their union is in S. Let $B_1 = A_1$ and for $n \geqslant 1$, $B_n = A_n - (A_1 \cup A_2 \cup ... \cup A_{n-1}) = A_n \cap A_1^c \cap A_2^c \cap ... \cap A_n^c$ Thus S is a λ-system, each complement A_i^c is in S, and since S is a π-system it follows that B_n, which is a finite intersection of sets in S, is also in S. □

Theorem. *($\pi - \lambda$ theorem) If P is a* π*-system on* Ω*, then* $\lambda(P) = \sigma(P)$*.*

Proof. Let $L_0 = \lambda(P)$. If L_0 is a π-system, then L_0 is a σ-algebra, consequently $\sigma(P) \subset L_0$; since $L_0 = \lambda(P) \subset \sigma(P)$. It remains therefore to show that L_0 is a π-system.
For $A \in L_0$, let $L_A = \{B \subset \Omega | A \cap B\}$. To show that L_0 is a π-system is to show that $L_0 \subset L_A$ for every $A \in L_0$. Clearly, L_A is a λ-system. Observe then that if $B \in P$, since P is a π-system and hence L_B is a λ-system containing P. Therefore, $P \subset L_B$ if $B \in P$, this means that $A \cap B \in L_0$ and $B \in P$, or $P \subset L_A$ if $A \in L_0$. Since L_A is a λ-system, we then have $L_0 \subset L_A$ for $A \in L_0$. Thus, L_0 is a π-system and we proved the theorem. □

Lemma. *Let S be a set. Let I be a* π*-system on S, and let* $\Sigma := \sigma(I)$*. Suppose that* μ_1 *and* μ_2 *are measures on* (S, Σ) *such that* $\mu_1(S) = \mu_2(S) < \infty$ *and* $\mu_1 = \mu_2$ *on I. Then,* $\mu_1 = \mu_2$ *on* Σ*.*

Proof. Let $D = \{F \in \Sigma : \mu_1(F) = \mu_2(F)\}$. Then, D is a λ-system. Indeed, the fact that S in D is given. If $A, B \in D$, then $\mu_1(B \setminus A) = \mu_1(B) - \mu_1(A) = \mu_2(B) - \mu_2(A) = \mu_2(B \setminus A)$. So that $B \setminus A \in D$. Finally, if $F_n \in D$ and $F_n \to F$, then $\mu_1(F) = \lim \mu_1(F_n) = \lim \mu_2(F_n) = \mu_2(f)$ so that F is in D. Since D is a λ-system and $I \subset D$ by hypothesis, the Dynkin's theorem shows that $\sigma(I) = \Sigma \subseteq D$, and hence we proves the lemma. □

Definition. *Premeasure: A set function, taking nonnegative value, montone, continuous from below, with empty set taking measure zero.*

Definition. *Measure with the σ-additivity.*

Definition. *A measure is finite if $\mu(X) < \infty$.*
A measure is σ-finite if X is the countable union of measurable sets with finite measure.

5.1 Independence

Lemma. *Suppose that G and H are sub-σ-algebras of F, and that I and J are π-systems with $\sigma(I) = G, \sigma(J) = H$. Then G and H are independent if and only if I and J are independent in that $P(I \cap J) = P(I)P(J)$.*

Proof. Suppose that I and J are independent. For fixed I in I, the measures $H \to P(I \cap H)$ and $H \to P(I)P(H)$ on (Ω, H) have the same total mass $P(I)$, and agree on J. And thus the measures agree on H. Then, fix H in H. We can still deduce that the two measures have the same total mass P(H), and agree on I. Therefore, agree on G. □

Definition. *A sequence of events E_n happen infinitely often:*$= \lim sup E_n := \cap_m \cup_{n \geqslant m} E_n$*=*$\omega : \omega \in E_n$ *for infinitely many n.*

Lemma. *(First Borel-Cantelli lemma) Let E_n be a sequence of events such that $\sum_n P(E_n) < \infty$. Then $P(\lim sup E_n) = P(E_n, i.o.) = 0$*

Proof. Let $G_m = \cup_{n \geqslant m} E_n$. Then, we have $P(G) \geqslant P(G_m) \geqslant \sum_{n \geqslant} P(E_n)$. Drive $m \to \infty$. The result directly proves the lemma. □

Lemma. *(Second Borel-Cantelli lemma) If E_n is a sequence of independent events. Then, $\sum P(E_n) = \infty \Rightarrow P(E_n, i.o.) = 1$*

First, we have $(\lim sup E_n)^c = \lim inf E_n^c = \cup_m \cap_{n \geqslant m} E_n^c$. With p_n denoting $P(E_n)$, we have $P(\cup_{n \geqslant m} E_n^c) = (1 - p_m)....$ For $x > 0$, $1 - x \leqslant e^{-x}$, since $\sum p_n = \infty$, $(1 - p_m)... \leqslant e^{-\sum_{n \geqslant m}} = 0$ So, we proved the lemma.

Theorem. *(Komogorov's 0-1 law) Let X_n be a sequence of independent random variables, and let τ be the tail σ-algebra of X_n. Then, τ is P-trivial. That is,*
(i) $F \in \tau \Rightarrow P(F) = 0$ or $P(F) = 1$
(ii) if ϵ is a τ-measurable random variable, then ϵ is almost deterministic in that for some constant $c \in [-\infty, \infty]$, $P(\epsilon = c) = 1$.

Proof. Let $\chi_n = \sigma(X_1, ...X_n), \tau_n = \sigma(X_{n+1}, X_{n+2}, ...)$.
We first claim that χ_n and τ_n are independent. The class K of events of the form $\omega : X_i(\omega) \leqslant x_i : 1 \leqslant i \leqslant n$ is a π-system which generates χ_n. The class J of sets of forms $\omega : X_j(\omega) \leqslant x_j : n+1 \leqslant j \leqslant n+r$. The claim is proved since X_n is a sequence of independent r.v.
Because $\tau \subseteq \tau_n$. Thus, χ_n and τ are independent.
We claim that $\chi_\infty = \sigma(X_n)(n \in N)$ and τ are independent. Because $chi_n \subseteq \chi_{n+1}, \forall n$, the class $K_\infty = \cup \chi_n$ is a π-system which generates χ_∞. Moreover, K_∞ and τ are independent. Then, the claim is proved.
Since $\tau \subseteq \chi_{infty}$, τ is independent of τ. Hence, (i) is proved. (ii) can be proved from (i). □

5.2 The integration theory of Lebesgue

Definition. *X,Y: measurable space. A measurable map* $X \xrightarrow{f} Y$ *is a map* $\underset{-}{X} \xrightarrow{f} \underset{-}{Y}$

For a measure space (X,m,μ), we aim at defining $\int_A f(x)d\mu(x) = \int_A f(x)\mu(dx)$ for suitable measurable functions $X \xrightarrow{f} \bar{R}$ and for any $A \in m$ in a systematic manner.
(1) (Reduction to the [0,∞]-valued case) The idea is that $f = f^+ - f^-$ where $f^+ := \max\{+f, 0\}$, $f^- := \max\{-f, 0\}$. Since f,-f,0 are all measurable. f^+, f^- are measurable. We say that $\int_X f d\mu$ is defined if both $\int_X f^+ d\mu$ and $\int_X f^- d\mu$ are defined and not simultaneously ∞. If this is the case $\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu \in \bar{R}$. If furthermore, $\int_X f d\mu \in R$, we say that f is μ-integrable.
(Reduction to the case of measurable $[0, \infty)$-valued simple functions) For any measurable function $X \xrightarrow{f} [0, \infty]$, we define $\int_A f d\mu = sup \int_A s d\mu | S$ is a measurable $[0, \infty)$-valued simple functions on X s.t. $\forall x \in X, s(x) \leqslant f(x)$ where $\int_A s d\mu := \sum_{c \in [0,\infty)} c\mu(s^{-1}(c) \cap A)$. (Or, in more familiar terms, if $s = \sum_{j=1}^n \alpha_j \chi_{s^{-1}(\alpha_j)}$ with $\alpha_1, ...\alpha_n \in [0, \infty)$ distinct and $E_1, ...E_n \in m$ disjoint, then $\int_A s d\mu = \sum_{j=1}^n \alpha_j \mu(E_j \cap A)$. Again, for measurable non-negative simple functions s, the two definitions coincide.
Basic properties: for non-negative measurable functions, monotonicity, positive homogeneity holds and whole-spaced integrability ensures subset integrability.

Lemma. *Let s,t be non-negative measurable simple functions.*
(1) $m \xrightarrow{\nu} [0, \infty] (A \longrightarrow \int_A s d\mu)$ *is a positive measure.*
(2) $\int_X (s+t) d\mu = \int_X s d\mu + \int_X t d\mu$.

Proof. (1) Suppose that $A_n \in M$ be a disjoint family. Write $s = \sum_{\alpha \in [0,\infty)} \alpha \chi_{s^{-1}(x)}$.
$\nu(\cup_{n=1}^\infty A_n) = \int_{\cup_{n=1}^\infty A_n} s d\mu = \int_{\cup_{n=1}^\infty A_n} (\sum_{\alpha \in [0,\infty)} \alpha \chi_{s^{-1}(\alpha)}) d\mu = \sum_{\alpha \in [0,\infty)} \int_{\cup_{n=1}^\infty A_n} \alpha \chi_{s^{-1}(\alpha)} d\mu = \sum_{\alpha \in [0,\infty)} \alpha \int_{\cup_{n=1}^\infty A_n} \chi^{s^{-1}(\alpha)} d\mu = \sum_{\alpha \in [0,\infty)} \alpha \int_X \chi_{\cup_{n=1}^\infty A_n} \chi_{s^{-1}(\alpha)} d\mu = \sum_{\alpha \in [0,\infty)} \alpha \sum_{n=1}^\infty \chi_{a_n \cap s^{-1}(\alpha)} d\mu = \sum_{\alpha \in [0,\infty)} \alpha \sum_{n=1}^\infty \mu(s^{-1}(\alpha) \cap A_n) = \sum_{n=1}^\infty \sum_{\alpha \in [0,\infty)} \alpha \mu(s^{-1}(\alpha) \cap A_n) = \sum_{n=1}^\infty \int_{A_n} s d\mu = \sum_{n=1}^\infty \nu(A_n)$. Besides, $\nu(\emptyset) = 0$

(2) Write $t = \sum_{\beta\in[0,\infty)} \beta\chi_{t^{-1}(\beta)}$. $\int_X (s+t)d\mu = \nu((\cup_{\alpha\in[0,\infty)} s^{-1}(\alpha))\cap(\cup_{\beta\in[0,\infty)} t^{-1}(\beta)) = \nu(\cup_{\alpha,\beta\in[0,\infty)} s^{-1}(\alpha)\cap t^{-1}(\beta)) = \sum_{\alpha,\beta\in[0,\infty)} \nu(s^{-1}(\alpha)\cap t^{-1}(\beta)) = \sum_{\alpha,\beta} \int_{s^{-1}(\alpha)\cap t^{-1}(\beta)} (s+t)d\mu = \sum_{\alpha,\beta} \int_X \chi_{s^{-1}(\alpha)\cap t^{-1}(\beta)}(s+t)d\mu = \sum_{\alpha,\beta}(\alpha+\beta)\mu(s^{-1}(\alpha)\cap t^{-1}(\beta)) = \sum_\alpha \alpha \sum_\beta \mu(s^{-1}(\alpha)\cap t^{-1}(\beta)) + \sum_\beta \beta \sum_\alpha \mu(s^{-1}(\alpha)\cap t^{-1}(\beta)) = \int_X sd\mu + \int_X td\mu$. Note that additivity holds for countable functions, which can be seen by approximating measurable functions by simple functions and thereafter applying monotone convergence theorem. □

Theorem. *(Lebesgue's monotone convergence theorem) Let* $X \xrightarrow{f_n} [0,\infty](n \in N)$ *be sequence of non-decreasing measurable functions* f_n. *Then* $\int_X (\lim_{n\to\infty} f_n) = \lim_{n\to\infty} \int_x f_n d\mu$.

Proof. Let $f := \lim_{n\to\infty} f_n$. By the monotonicity we have $\lim_{n\to\infty}\int_X f_n d\mu = \alpha \leqslant \int_X f d\mu$. It remains to show that $sup\int_X sd\mu|s \leqslant f = \int_X fd\mu \leqslant \lim_{n\to\infty}\int_X f_n d\mu = \alpha$ and suffices to prove that $\alpha \geqslant \int_X sd\mu$ for all measurable $[0,\infty)$-valued simple function with $s \leqslant f$. Let $c \in (0,1)$ and consider $E_n = x \in X|f_n(x) \geqslant cs(x)$. $E_n \to \cup_{n=1}^\infty E_n = X$. Therefore, $\int_X f_n d\mu \geqslant \int_{E_n} f_n d\mu \geqslant \int_{E_n} csd\mu = c\int_{E_n} sd\mu (n \in N)$. Drivw $n \to \infty$. $\alpha \geqslant c\nu(\cup_{n=1}^\infty E_n) = c\int_X sd\mu$. Let $c \to 1^-$, we proved the theorem. □

Approximation by simple functions: for any measurable function $X \xrightarrow{f} [0,\infty]$, we let

$$s_n(x) = \begin{cases} \dfrac{k-1}{2^n} \quad if \quad \dfrac{k-1}{2^n} \leqslant f(x) < \dfrac{k}{2^n} \\ \qquad\qquad n \quad if \quad f(x) \geqslant n \end{cases} \tag{1}$$

Then $X \xrightarrow{s_n} [0,\infty)$ is a measurable simple function: $s_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n}\chi_{f^{-1}[\frac{k-1}{2^n},\frac{k}{2^n})} + n\chi_{f^{-1}[n,\infty)}$. In addition, $s_n \to f$ as $n \to \infty$. By monotone convergence theorem, $\int_X s_n d\mu$ converges to $\int_X f d\mu$.

Lemma. *(Fatou lemma) Let* $X \xrightarrow{f_n} [0,\infty](n \in N)$ *be measurable. Then* $\int_X (\lim_n inf f_n)d\mu \leqslant \lim_n inf \int_X f_n d\mu$.

Proof. Note $\lim_k inf f_k = \lim_{n\to\infty} inf_{m\geqslant n} f_m$. We let $inf_{m\geqslant n} = g_n$. For $m \geqslant n$, we have $f_m \geqslant g_n$. By monotonicity, $inf \int_X f_m d\mu \geqslant \int_X g_n d\mu$. Then, $\int_X (\lim_n inf f_n)d\mu = \int_X (\lim_{n\to\infty} g_n)d\mu = \lim_{n\to\infty}\int_X g_n d\mu \leqslant \lim_{n\to\infty} inf_{m\geqslant n}\int_X f_m d\mu$. □

Lemma. *(Reverse Fatou lemma) Let* $X \xrightarrow{f_n} [0,\infty](n \in N)$ *be measurable. We have* $f_n \leqslant g, \forall n$, *and* g *is integrable, then* $int_X \lim sup f_n d\mu \geqslant \lim sup \int_X f_n d\mu$.

We construct a new sequence of measurable functions $g_n = g - f_n$. By Fatou lemma, $\int_X \lim_n inf g_n d\mu \leqslant \lim_n inf \int_X g_n d\mu$. That is, $\int_X \lim_n inf(g - f_n)d\mu = \int_X gd\mu - \int_X \lim_n sup_{m\geqslant n} f_m d\mu \leqslant \int_X gd\mu - \lim_n sup_{m\geqslant n}\int_X f_m d\mu$

Lemma. *If* $f \in L^1(\mu)$, *then* $|f_X f d\mu| \leqslant \int_X |f|d\mu$

Proof. There exists $\alpha \in C$ s.t. $|\alpha| = 1$ and $\alpha \int_X f d\mu \in [0, \infty)$. $|\int_X f d\mu| = |\alpha \int_X f d\mu| = |\int_X \alpha f d\mu| = \int_X \alpha f d\mu = \int_X Re(\alpha f) d\mu \leqslant \int_X |\alpha f| d\mu =$ □

Theorem. *(Lebesgue's dominant convergence theorem) Let $f_n \in L^1(X, \Sigma, \mu)$ which converges pointwise to a function f. Suppose that there exists $g \in L^1(X, \Sigma, \mu)_R$, s.t. $\forall n \in N, |f_n| \leqslant g$. Then, $\lim_{n\to\infty} \int_X |f_n - f| = 0$, whence $\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu$.*

Proof. $|f_n| \leqslant g(n \in N) \Rightarrow |f| \leqslant g$. $2g - |f_n - f| \geqslant 0$. By Fatou lemma, $\int_X \lim_{n\to\infty} inf_{m \geqslant n}(2g - |f_m - f|) \leqslant \lim_{n\to\infty} inf_{m\geqslant n} \int_X 2g - |f_m - f| d\mu$. $\int_X 2g d\mu \leqslant \int_X 2g d\mu - \lim sup |f_n - f| d\mu \Rightarrow \lim sup |f_n - f| \geqslant \Rightarrow \lim_{n\to\infty} \int_X |f_n - f| d\mu = 0$. □

Definition. *An element F of Σ is called μ-null if $\mu(F) = 0$. An statement about points is said to hold almost everywhere if the condition that the statement is false is a μ-null set.*

It is possible to extend the result of the above theorems related to convergence with the concept of "almost everywhere".

Proposition 6. *If $X \xrightarrow{f} [0, \infty], A \in m$ and $\int_A f d\mu = 0$, then $f = 0$ μ-a.e. on A.*

Proof. By Chebyshev's inequality, $\forall c \in (0, \infty)$, $c\mu x \setminus A|f(x) \geqslant c \leqslant \int_A f d\mu$. In particular, for $n \in N$, $\frac{1}{n}\mu x \in A|f(x) \geqslant \frac{1}{n} \leqslant \int_A f d\mu$. So that $\mu x \in A|f(x) \geqslant \frac{1}{n} = \mu(A_n) = 0$. $A_n \to x \in A|f(x) > 0$ as $n \to \infty$. Then, $\mu x \in A|f(x) \neq 0 = 0$. Hence, we proved the proposition. □

Proposition 7. *For $f \in L^1(\mu)$, $|\int_X f d\mu| = \int_X |f| d\mu \Leftrightarrow \exists \alpha \in C$ s.t. $|\alpha| = 1$ and $|f| = |\alpha f|$.*

Proof. Let $\alpha \in C$ s.t. $|\alpha| = 1$ and $\alpha \int_X f d\mu = |\int_X f d\mu|$. Then, $|\int_X f d\mu| = \int_X \alpha f d\mu = Re \int_X \alpha f d\mu = \int_X Re(\alpha f) d\mu \leqslant \int_X |\alpha f| d\mu = \int_X |f| d\mu$ and by the assumption $|\int_X f d\mu| = \int_X |f| d\mu$. Therefore, $\int_X |\alpha f| - Re(\alpha f) d\mu = 0$. By the above proposition, $|\alpha f| = Re(\alpha f)$ μ-a.e. Therefore, $|f| = \alpha f$ μ-a.e. □

Proposition 8. *Given $X \xrightarrow{f_n} \bar{R}$ or C measurable. If $\sum_{n=1}^{\infty} \int_X |f_n| d\mu < \infty$, then*
(1) $\sum_{n=1}^{\infty} f_n$ converges absolutely μ-a.e.
(2) $\sum_{n=1}^{\infty} f_n \in L^1(\mu)$
(3) $\int_X (\sum_{n=1}^{\infty} f_n) d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu$.

Proof. By Chebyshev's inequality for every $m \in N$, $m\mu x \in X|\sum_{n=1}^{\infty} |f_n(x)| \geqslant m \leqslant \int_X \sum_{n=1}^{\infty} |f_n(x)| d\mu = \sum_{n=1}^{\infty} \int_X |f_n(x)| d\mu < \infty$. $m \in N$, $\mu x \in X|\sum_{n=1}^{\infty} |f_n(x)| \geqslant m$ decreases to $m \in N$, $\mu x \in X|\sum_{n=1}^{\infty} |f_n(x)| = \infty$ as $n \to 0$. Since $\mu x \in X|\sum_{n=1}^{\infty} |f_n(x)| \geqslant m \leqslant \frac{\sum_{n=1}^{\infty} \int_X |f_n(x)| d\mu}{m} \to 0$ as $m \to \infty$.
$|\sum_{n=1}^{\infty} f_n|$ is μ-a.e. measurable. $|\lim_{m\to\infty} \sum_{n=1}^{m} f_n| = \lim_{m\to\infty} |\sum_{n=1}^{m} f_n| \leqslant \lim_{m\to\infty} \sum_{n=1}^{m} |f_n| = \sum_{n=1}^{\infty} |f_n| \in L^1(\mu)$. By DCT, $\sum_{n=1}^{\infty} f_n \in L^1$. □

5.3 Constructions of measurable spaces

Definition. *A map $P(X) \xrightarrow{\mu} \bar{R}_+$ is an outer measure on X if $\mu(\emptyset) = 0$,$A \subseteq B \subseteq X \Rightarrow \mu(A) \leqslant \mu(B)$, and σ-subadditivity,$\forall A_n \subseteq X$,$\mu(\cup_{n=1}^{\infty} A_n) \leqslant \sum_{n=1}^{\infty} \mu(A_n)$.*

Theorem. *(Caratheòdory's construction of measures) A_μ :=$A \subseteq X | \forall E \subseteq X, \mu(E) = \mu(E \cap A) + \mu(E \setminus A)$. Then, $(X, A_\mu, \mu|_{A_\mu})$ is a complete measure space.*

Proof. It is clear that $\emptyset \in A_\mu$ and $A^c \in A_\mu \Leftrightarrow A \in A_\mu$. Now we prove that $\forall A_n \in A_\mu (n \in N), [\cap_{n=1}^{\infty} A_n \in A_\mu]$. For any $E \subseteq X$ s.t. $\mu(E) < \infty$, $\mu(E) = \mu(E \cap A_1) + \mu(E \cap A^c) = \mu(E \cap A_1 \cap A_2) + \mu(E \cap A_1 \cap A_2^c) + \mu(E \cap A_1^c) = \mu(E \cap \cap_{n=1}^{m} A_n) + \sum_{n=1}^{m} \mu(E \cap A_1 \cap ... \cap A_{n-1} \cap A_n^c) \geqslant \mu(E \cap \cap_{n=1}^{\infty} A_n) + \sum_{n=1}^{m} \mu(E \cap A_1 \cap ... \cap A_{n-1} \cap A_n^c)$. Then, $\sum_{n=}^{m} \mu(E \cap A_1 \cap ... \cap A_{n-1} \cap A_n^c) \leqslant \mu(E) - \mu(E \cap \cap_{n=1}^{\infty} A_n) \overset{m \to \infty}{\Rightarrow} \sum_{n=1}^{infty} \mu(E \cap A_1 \cap ... \cap A_{n-1} \cap A_n^c) \leqslant \mu(E) - \mu(E \cap \cap_{n=1}^{\infty} A_n)$. By σ-subadditivity, $\mu(E \cap (\cap_{n=1}^{\infty} A_n)^c) = \mu(E \cap A_1 \cap ... \cap A_{n-1} \cap A_n^c) \leqslant \mu(E) - \mu(E \cap \cap_{n=1}^{\infty} A_n)$. And again by *sigma*-subadditivity, $\mu(E \cap (\cap_{n=1}^{\infty} A_n)^c) + \mu(E \cap \cap_{n=1}^{\infty} A_n) \geqslant \mu(E)$. Hence, $\mu(E \cap (\cap_{n=1}^{\infty} A_n)^c) + \mu(E \cap \cap_{n=1}^{\infty} A_n) = \mu(E)$. Hence, we proved that A_μ is a σ-algebra on X. Now we show that $\mu|_{A_\mu}$ is a measure. Let $B_n \in A_\mu$ be a disjoint family. $\mu(\cup_{n=1}^{\infty} B_n) = \mu((\cup_{n=1}^{\infty} B_n) \cap B_1) + \mu((\cup_{n=1}^{\infty} B_n) \setminus B_1) = \mu(B_1) + \mu(\cup_{n=2}^{\infty} B_n) = ... = \sum_{n=1}^{m} + \mu(\cup_{n=m+1}^{\infty} B_n) \geqslant \sum_{n=1}^{m} \mu(B_n)$. As $m \to \infty$, $\mu(\cup_{n=1}^{infty} B_n) \geqslant \sum_{n=1}^{\infty} \mu(B_n)$. Again, by σ-subadditivity, the equality holds. In addition, completeness of measure can easily be seen. $\square$

Lemma. *(Creating outer measures via coverings) X: a set, $S \subseteq P(X), S \xrightarrow{\phi} \bar{R}_+$, a function. If $\emptyset \in S$ and $\phi(\emptyset) = 0$ (for simplicity), then $P(X) \xrightarrow{\mu_\phi} \bar{R}_+, A \longrightarrow inf \Phi_A$ is an outer measure where Φ_A =$\sum_{U \in U} \phi(U) | U \subseteq S$ is a countable cover of A.*

Proof. $0 = \phi(\emptyset) \in Phi_\emptyset \Rightarrow \mu_\phi(\emptyset) = inf \Phi_\emptyset = 0$.
$A \subseteq A' \Rightarrow \Phi_{A'} \subseteq \Phi_A \Rightarrow \mu_\phi(A) = inf \Phi_A \leqslant inf \Phi_{A'} = \mu_\phi(A')$.
Given $A_n \subseteq X$, by definition, for any $\epsilon > 0$ and $n \in N$, there exists a countable cover $U_n \subseteq S$ of A_n s.t. $\sum_{U \in U_n} \phi(U) \leqslant \mu_\phi(A_n) + \frac{\epsilon}{2^n}$. $U := \cup_{n=1}^{\infty}$ is a countable cover of $\cup_{n=1}^{\infty} A_n$. $\mu_\phi(\cup_n A_n) \leqslant_{U \in U} \phi(U) \leqslant \sum_{n=1}^{\infty} \sum_{U \in U_n} \phi(U) \leqslant \sum_{n=1}^{\infty} \mu_\phi(A_n) + \epsilon$. Hence, we proved σ-subadditivity. $\square$

Example. *X: metric space. $\alpha, \delta \in (0, \infty)$. Consider S =$S \subseteq X | diam S < \delta$.(diam:=sup$d(s_1, s_2) | s_1, s_2 \in S$)*

. Let $S \overset{\phi}{\bar{R}_+}, S \to (diam S)^\alpha$. Then $H_\delta^\alpha = \mu_\phi$ is an outer measure and $\forall A \subseteq X$ and $\delta' > \delta > 0, [H_{\delta'}^\alpha(A) \leqslant H_\delta^\alpha(A)]$. Define $H^{alpha} = sup H_\delta^\alpha$, called the α-dimensional Hausdorff outer measure when $X = R^d$ with the usual metric.

Remark. $\forall A \subseteq X$ *and* $\delta \in (0,1) [H_\delta^\alpha(A) \downarrow \quad as \quad \alpha \uparrow]$. *We call sup$\alpha > 0 | H^\alpha = \infty$ the Hausdorff dimension of A.*

Example. *In R^d a standard rectangle is a subset of the form $I_1 \times I_2 \times ... \times I_d$ with I_d an interval of R. Let S =all standard rectangles. $S \overset{vol}{\bar{R}_+}$, where $vol = l(I_1)...l(d)$. Then we call μ_{vol} the Lebesgue outer measure. We call μ_{vol} Lebesgue measurable sets.*

Definition. *Let X be a set and $S \subseteq P(Z)$.*
(1) S is a semiring on X if $\emptyset \in S$, $A \cap B \in S$ if $A, B \in S$, $\forall A, B \in S$,$exists S_1, ..., Sk \in S, B \setminus A = S_1 \cup ... \cup S_k$.
(2) A function $S \overset{\phi}{\bar{R}_+}$ is finitely/countable additive if for any finite/countable family of disjoint subsets A_n of X, $\cup_n A_n \in S$, then $\phi(\cup_n A_n) = \sum_n \phi(A_n)$. If furthermore $\phi(\emptyset) = 0$, we call ϕ a finitely/countably additive measure on the semiring S.

Lemma. *Let μ be finitely additive measure on a semiring $S \subseteq P(X)$.*
(1) All elements of S are $\mu\prime$-measurable.
(2) If μ is a countably additive measure on the semiring S, then $\mu\prime|_S = \mu$

Proof. Let $S \in S$. It follows immediately from the definition of the induced outer measure that $\mu\prime(S) \leqslant \mu(S)$. Therefore, it suffices to show that if $(A_n)_{n=0}^{\infty}$ is a countable cover of S, then $\mu(S) \leqslant \sum_{n=0}^{\infty} \mu(A_n)$ since $\mu(S) \leqslant inf \sum_{n=0}^{\infty} \mu(A_n) = \mu\prime$. Define, $\forall n \in N$: $B_n = A_n \setminus A_{n-1} \setminus ... \setminus A_0$. Using the mathematical induction, we will prove that for all natural numbers $m < n$, $B_{n,m} = A_n \setminus A_{n-1} \setminus ... \setminus A_{n-m}$ is the finite union of pairwise disjoint elements of S. The base case $m = 0$ is trivial. Now assume that the induction hypothesis that the above statement is true for some $m < n - 1$, and let $D_1, D_2, ..., D_N$ be pairwise disjoint elements of S such that: $B_{n,m} = \cup_{n=1}^{N} D_k$. Then, $B_{n,m+1} = B_{n,m} \setminus A_{n-m-1} = \cup_{n=1}^{N} D_k \setminus A_{n-m-1} = \cup_{n=1}^{N} (D_k \setminus A_{n-m-1})$. Hence $B_{n,m+1}$ is the finite union of pairwise disjoint elements of S, completing the induction step.
Using the above result, we can choose a finite set F_n of pairwise disjoint elements of S for which $B_n = \cup F_n$. Now, $x \in S$ if and only if $\exists n \in N$ such that $x \in S \cap A_n$. Taking the smallest such n, it follows that x not in $A_0, A_1, ...A_{n-1}$, and so $x \in S \cap B_n$. Therefore, $S = \cup_{n=0}^{\infty}(S \cap B_n)$. Hence, $\mu(S) = \mu(\cup_{n=0}^{\infty}(S \cap B_n)) = \mu(\cup_{n=0}^{\infty}(S \cap \cup_{T \in F_n} T)) = \mu(\cup_{n=0}^{\infty} \cup_{T \in F_n} (S \cap T) \leqslant \sum_{n=0}^{\infty} \sum_{T \in F_n} \mu(S \cap T) \leqslant \sum_{n=0}^{\infty} \sum_{T \in F_n} \mu(T) = \sum_{n=0}^{\infty} \mu(\cup F_n) = \sum_{n=0}^{\infty} \mu(B_n) \leqslant \sum_{n=0} =^{\infty} \mu(A_n)$. □

Remark. *Let S =all standard rectangles. Then S is a semiring. It is easy to check so. To state $\mu_{vol} = vol$, it suffices to show that Lebesgue measure on the standard rectangles is indeed measure.*

Lemma. *Lebesgue measure on the standard rectangles is indeed measure.*

Proof. It is known that $vol(\emptyset) = 0$. he only possibility for two disjoint half-open n-rectangles to constitute a single, large half-open n-rectangle is when they are of the form: $[[a..b))[[a'..b'))$ s.t. we have for some i with $1 \leqslant i \leqslant n$: $i = j \Rightarrow a'_j = b_j$. We can then see that vol is finitely additive.
Suppose that $[a_m..b_m) \downarrow \emptyset$. Then there exists at least $1 \leqslant j \leqslant n$ s.t.: $\lim_{m\to\infty} a_{m,j} =$

$\lim_{m\to\infty} b_{m,j}$. The fact that the sequence is decreasing means that, from the Cartesian product of subsets, $\forall m \in N$, and $\forall 1 \leqslant i \leqslant n$: $[a_{m,i}, ..., b_{m,i}) \subseteq [a_{1,i}, ...b_{1,i})$. Hence we have: $\lim_{m\to\infty} vol([a_m..b_m)) = \lim m \to \infty \Pi_{i=1}^{\infty}(b_{m,i} - a_{m,i}) \leqslant \lim m \to \infty (b_{m,j} - a_{m,j}) \Pi_{i=1,i\neq j}^{n}(b_{m,i} - a_{m,i}) = 0$. □

Theorem. *(Carathèodory's criterion of Borel measurability) Let X be a metric space and $P(X) \xrightarrow{\mu} \bar{R}_+$ an outer measure. If $\mu(A \cup B) = \mu(A) + \mu(B)$ for any $A, B \subseteq X$ s.t. d(A,B)¿0, then $B_x \subseteq A_\mu$*

Proof. It suffices to show that $C \in A_\mu$ for all C: closed in X. Let $C_k = x \in X | d(x,c) \leqslant \frac{1}{k}$. For any $E \subseteq X$ with $\mu(E) < \infty$, since $d(E \setminus C_k, E \cap C) \geqslant \frac{1}{k} > 0$. $\mu(E) \geqslant \mu((E \setminus C_k) \cap (E \cap C)) = \mu(E \setminus C_k) + \mu(E \cap C)$. We will show that $\lim_{k\to\infty} \mu(E \setminus C_k) = \mu(E \setminus C)$. $E \setminus C = (E \setminus C_k) \cup \cup_{j=k}^{\infty} D_j$ where $D_j = x \in E | \frac{1}{j+1} < d(x, C) \leqslant \frac{1}{j}$. Then $\mu(E \setminus C_k \leqslant \mu(E \setminus C) \leqslant \mu(E \setminus C_k) + \sum_{j=k}^{\infty} \mu(D_j$. Since $\sum_{j=1}^{\infty} \mu(D_j) = \sum_{j=1} \mu(D_{2j-1}) + \sum_{j=1} \mu(D_{2j}) \leqslant 2\mu(E) < \infty$. □

Definition. *(Regularity)*
X: top. space (X,M,μ):measurable space.$A \subseteq X$.
(1) A is outer regular w.r.t. μ if $\mu(A) = inf_{A\subseteq U \subseteq_{open} X} \mu(U)$.
(2) A is inner regular w.r.t. μ if $\mu(A) = sup_{C\subseteq A} \mu(C)$
(3) A is inner compact-regular w.r.t. μ if $\mu(A) = sup_{K\subseteq A} \mu(K)$.

5.4 Integration on product spaces

Definition. *Suppose that (X, A) and (Y, B) are measurable spaces. The product σ-algebra $A \otimes B$ is the σ-algebra on $X \times Y$ generated by the collection of all measurable rectangles, $A \otimes B = \sigma$ $(A \times B : A \in A, B \in B)$.*
The product of (X,A) and (Y,B) is the measurable space $(X \times Y, A \otimes B)$.

Definition. *Let Z be a set and $m \subseteq P(Z)$. We say that m is a monotone class if $\forall S_n \in m (n \in N)$, $S_n \uparrow S$ or $S_n \downarrow S$ and $S \in m$ as $n \to \infty$.*

Remark. *Every σ-algebra is a monotone class.*
For any $S \subseteq P(Z), m(S) := \cap_{S\subseteq m} m$ is clear the smallest monotone class containing S.

Theorem. *(monotone characterization of $A \otimes B$) Apply π-λ theorem. $\epsilon_{A,B}$ consists of all finite disjoint unions of measurable rectangles and thus a π-system. The smallest monotone class generated by $\epsilon_{A,B}$ is the smallest $\lambda(\epsilon_{A,B})$, and hence the smallest σ-algebra.*

Lemma. *Let (X, A) and Y, B be two measurable spaces. For any $E \in A \otimes B$ and $x \in X$(resp. $y \in Y$), we have $E_x \in B$(resp. $E_y \in A$).*

Proof. Consider $n = S \subseteq X \times Y | \forall x \in X$ and $y \in Y, S_x \in B$ and $S_y \in A$. Then n includes all measurable rectangles. Besides n is a σ-algebra. This can be seen by checking directly or noticing n is the final σ-algebra induced by all the maps i_x and i_y. Hence it contains $A \otimes B \subseteq n$. □

Theorem. *If (X, A, μ) and (Y, B, ν) are σ-finite measure spaces, then for any $Q \subseteq A \otimes B$,*
(1) the functions $X \xrightarrow{\phi_Q} \bar{R}_+, x \to \nu(Q_x)$ and $Y \xrightarrow{\psi_Q} \bar{R}_+, y \to \mu(Q_y)$ A and B measurable respectively
(2) $\int_X \phi_Q d\mu = \int_Y \psi_Q d\nu$.

Proof. Let n=$Q \subseteq A \otimes B|$(1) and (2) holds for Q.
(i) It is clear that n includes all measurable rectangles and all elementary sets, and hence a π-system.
(ii) We claim: For any Q_n, if $Q_n \uparrow Q$ as $n \to \infty$, then $Q \in n$. $Q_n \uparrow Q \Rightarrow (Q_n)_x \uparrow Q_x$ as $n \to \infty$. Thus, $\phi_{Q_n} \uparrow \phi_Q$ as $n \to \infty$ by monotonicity of measure. Hence ϕ_Q is A-measurable. $\int_X phi_Q = \int_X (\lim_{n\to\infty} \phi_{Q_n}) d\mu \overset{MON}{=} \lim_{n\to\infty} \int_X \phi_{Q_n} d\mu = \lim_{n\to} \int_Y \psi_{Q_n} d\nu = \int_Y \psi_Q d\nu$.
(iii) Next, we claim that: for any $Q_n \in n$, if $Q_n \downarrow Q$ as $n \to \infty$ and $\exists A \in A$ and $B \in B$ s.t. $Q_n \subseteq A \times B$ and $\mu(A), \nu(B) < \infty$, then $Q \in n$. $\phi_{Q_n} \downarrow \phi_Q$ and $\psi_{Q_n} \downarrow \psi_Q$ as $n \to \infty$ and hence ϕ_Q and ψ_Q are A and B measurable respectively. Besides, $\phi_{Q_n} \leqslant \nu(B)\chi_A$ and $\psi_{Q_n} \leqslant \mu(A)\chi_B$. Both $\nu(B)\chi_A$ and $\mu(A)\chi_B$ are in $L^1(\mu)$. By the dominated convergence theorem, $\int_X phi_Q = \int_X (\lim_{n\to\infty} \phi_{Q_n}) d\mu \overset{DCT}{=} \lim_{n\to\infty} \int_X \phi_{Q_n} d\mu = \lim_{n\to} \int_Y \psi_{Q_n} d\nu = \int_Y \psi_Q d\nu$.
(iv) We further claim: for any disjoint family $Q_n \in n$. $\cup_{n=1}^{\infty} Q_n \in n$. For any $x \in X$. $Q_x = \cup_{n=1}^{\infty} (Q_n)_x$ and hence $Q_x \in B$ and $\nu(Q_x) = \sum_{n=1}^{\infty} \nu(Q_{n x})$ and A measurable. $\int_X \phi_Q d\mu = \sum_{n=1}^{\infty} \int_X \phi_{Q_n} d\mu = \sum_{n=1}^{\infty} \int_Y \psi_{Q_n} d\nu = \int_Y \psi_Q d\nu$.
(v) We also claim: Let $n' = Q \in A \otimes B | \forall m, n [Q \cap (X_m \times Y_n) \in n]$, then $n' = A \otimes B$.
(i)+(iv)$\Rightarrow \epsilon_{A,B} \subseteq n'$.
(ii)+(iii)$\Rightarrow n'$ is a monotone class.
$A \otimes B \subseteq n'$. (vi) $n = A \otimes B$. For $Q \in A \otimes B = n'$, we have $Q \cap (X_m \times Y_n) \in n$. Since $Q \cap (X_m \times Y_n)$ is a disjoint countable family in n. By (iv), $Q = \cup Q \cap (X_m \times Y_n) \in n$. □

Remark. *(iv)+($n = A \otimes B$) $\Rightarrow$ σ-additivity.*
$\mu \otimes \nu(\emptyset \times \emptyset) = 0$.
$\mu \otimes \nu$ is called a positive product measure.

Theorem. *(Tonelli's theorem) If f is $\bar{R}_+$ valued, then ϕ_f is A measurable, and ψ_f is B measurable, and $\int_X \phi_f d\mu = \int_Y \psi_f d\nu$.*

Proof. The statement holds for $f = \phi_Q$, with $Q \in A \otimes B$, and hence every $\bar{R}_+$-valued $A \otimes B$-measurable simple functions. For a general $\bar{R}_+$-valued f, select a sequence s_n of $\bar{R}_+$-valued $A \otimes B$-measurable simple functions s.t. $s_n \uparrow f$ as $n \to \infty$. $\phi_{s_n}(x) = \int_Y (s_n)_x d\nu \uparrow \int_Y f_x d\nu = \phi_f(x)$ as $n \to \infty$. ϕ_{s_n} :A-measurable$\Rightarrow \phi_f$A-measurable. $\int_X \phi_f d\mu = \int_X (\lim_{n\to} \phi_{s_n} d\mu =$
$lim_{n\to\infty} \int_X \phi_{s_n} d\mu = \lim_{n\to\infty} \int_{X\times Y} s_n d(\mu \otimes \nu) = \int_{X\times Y} (\lim_{n\to\infty} s_n) d(\mu \otimes \nu) = \int_{X\times Y} f d(\mu \otimes \nu)$. □

Theorem. *(Fubini's theorem) Let f be a general measurable function.*
(1) If $\int_X \int_Y |f(x,y)| d\nu(Y) d\mu(x) < \infty$, then f is $\mu \otimes \nu$-integrable.

(2) f is $\mu \otimes \nu$-integrable, then f_x is ν-integrable for $\mu - a.e., x \in X$, f_y is μ-integrable for $\nu - a.e., y \in Y$.

Proof. (1) Applying Tonelli's theorem to $|f|$, we see that $\int_{X\times Y} |f| d(\mu \otimes \nu) = \int_X \int_Y |f(x,y)| d\nu(y) d\mu(x) < \infty$.
(2) It suffices to consider the $\bar{R}$-valued functions. By applying Tonelli's theorem to $|f|$, $\int_X \int_Y |f(x,y)| d\nu(y) d\mu(x) = \int_{X\times Y} |f| d(\mu \otimes \nu) < \infty$. Then $x \in X|\phi_{|f|}(x) = \infty$ is μ-null, in other words, $\phi_{|f|}(x) < \infty$ for ν-integrable for μ-a.e. and for all x. Thus, $\phi_f(x) = \int_Y f(x,y) d\nu$ is defined on a μ-big set $X \backslash x \in X|\phi_{|f|}(x) = \infty$. $\phi_f(x) = \int_Y f(x,y) d\nu = \int_Y f^+(x,y) d\nu - \int_Y f^-(x,y) d\nu = \phi_{f^+}(x) - \phi_{f^-}(x)$ at least on $X \backslash x \in X|\phi_{|f|}(x) = \infty$. By Tonelli's theorem, ϕ_{f^+}, ϕ_{f^-} are A-measurable functions on X. Thus, $\phi_f(x) \overset{\mu}{=} h(x) = \phi_{f^+}(x) - \phi_{f^-}(x)$ if $x \in X \backslash x \in X|\phi_{|f|}(x) = \infty$, 0 if $x \in x \in X|\phi_{|f|}(x) = \infty$. It is easy to check h is in L^1. □

Theorem. *(Egoroff's theorem) Let $f_n \in \mu(A)_k$. If (1) f_n takes value in C μ-a.e. for every $n \in N$, (2) f_n converges μ-a.e. as $n \to \infty$, and (3) $\mu(X) < \infty$, then $\forall \epsilon > 0, \exists A \in A$ s.t. $\mu(X \setminus A) < \epsilon$ and f_n converges uniformly in A.*

Proof. By (1)(2), $\exists N \in A$ with $\mu(N) = 0$ s.t. f_n takes values in C on $X \setminus N$ and $f_n \to f$ pointwise on $X \setminus N$ as $n \to \infty$. Let $g_n(x) = sup_{m \geqslant n} |f_m(x) - f(x)|$ for $x \in X \setminus N$. Then, $g_n(x) \downarrow 0$ on $X \setminus N$ as $n \to \infty$. By (3), $\mu(X) < \infty$. We have $g_n \xrightarrow{\mu} 0$ on $X \setminus N$ as $n \to \infty$. And hence $\forall \epsilon > 0$ and $k \in N$, $\exists n_k \in N$ s.t. $\mu x \in X \setminus N | g_{n_k} > \frac{1}{k} < \frac{\epsilon}{2^k}$. Let $B_k = x \in X \setminus N | g_{n_k} > \frac{1}{k}$ and $A = \cap((X \setminus N) \setminus B_k) = X \setminus (N \cup \cup_{k=1}^{\infty} B_k)$. So $\mu(X \setminus A) = \mu(N$ $cup \cup_{k=1}^{\infty} B_k) < \epsilon$ Besides, $x \in A \to \forall k \in N, x$ not in $N \cup B_k \Leftrightarrow sup_{m \geqslant n_k} |f_m - f| = g_{n_k} < \frac{1}{k}$, Therefore, f_n converges to f uniformly on A. □

5.5 Measures vs abstraction integration-Riesz's representation theorem

Let X be a topological space and $A \overset{\mu}{[} 0, \infty]$ a positive measure on a σ-algebra A which includes B_x on X. Then all the continuous maps on X are B_x-measurable. Besides, if $\mu(K) < \infty$ for every compact sets $K \subseteq X$, then $C_c(x) \subseteq L^1(\mu)$. If this is the case, we have a C-linear map $C_c(X) \to C$, $f \to \int_X f d\mu$.

Definition. *A C-linear map $C_c(X) \xrightarrow{\wedge} C$ is a positive functional if it maps $C_c(X)_{\geqslant 0}$ into $R_{\geqslant 0}$ or equivalently, $\forall f_1, f_2 \in C_c(X), f_1 - f_2 \geqslant 0 \Rightarrow \wedge f_1 \geqslant \wedge f_2$.*

So there is a question: given a positive functional $C_c(x) \xrightarrow{\wedge} C$, does there exists a measure $A \xrightarrow{\mu} [0, \infty] with B_x \subseteq A$ s.t. $\wedge f = \int_X f d\mu$ for all $f \in C_c(X)$?

Definition. *A demiregular measure on a topological space X is a measure on some σ-algebra which contains B_x s.t. every compact subsets of X is measure-finite, all open sets are inner compact regular, and the measure is outer regular.*

Notation. *Given a topological space X,*
$K \prec f$ *means that* K *is compact in X,* $f \in C_c(X), 0 \leqslant f \leqslant 1$*, and* $f|_K = 1$.
$f \prec V$ *means that* V *is open in X,* $f \in C_c(X), 0 \leqslant f \leqslant 1$*, and* $suppf \subseteq V$.

Theorem. *If* $K \subseteq U \subseteq X$*, then there exists* $f \in C_c(X)$ *such that* $0 \leqslant f \leqslant 1, f_K = 1$ *and support of f in U.*

Remark. *By Urysohn's lemma, X :locally compact Hausdorff*$\Rightarrow \forall K \subseteq V \subseteq X, \exists f$*, s.t.* $K \prec f \prec V$.

Lemma. *Let X be a locally compact Hausdorff space. Given a positive functional* $C_c(X) \hat{C}$ *and a* σ*-algebra A which includes* B_X *on X, there exists at most one demiregular measure* μ*-on A s.t.* $\forall f \in C_c(X), \wedge f = \int_X F d\mu$*. Suppose that* μ_1 *and* μ_2 *are two such measures on A. By inner compact regularity of open sets and outer regularity, it suffices to show that* $\mu_1(K) = \mu_2(K)$ *for all compact subsets K of X since agreeing on the compact subsets implies that they agree on the open subsets by inner regularity of open sets, and on every subsets in A by outer regularity.*
For any compact $K \subseteq X$ *and* $\epsilon > 0$*, by the property of demiregularity, there exists an open set* $V \supseteq k$ *s.t.* $\mu_2 < \mu_2(K) + \epsilon$*. By Urysohn's lemma,* $\exists f, K \prec f \prec V, f|_K = 1$*. Then* $\mu_1 = \int_X \chi_K d\mu_1 \leqslant \int_X f d\mu_1 = \wedge f = \int_X f d\mu_2 \leqslant \int_X \chi_v d\mu_2 = \mu_2(V) < \mu_2(K) + \epsilon$*. Let* $\epsilon \downarrow 0$*,* $\mu_1(K) \leqslant \mu_2(K)$*. By symmetry,* $\mu_1(K) = \mu_2(K)$*.*

Theorem. *(Riesz's representation theorem) Let X be a locally compact Hausdorff space. Given a positive functional* $C_c(X) \xrightarrow{\wedge} C$*, there exists a complete demiregular* μ *on some* σ*-algebra* $A \supseteq B_x$ *on X s.t.* $\wedge f = \int_X f d\mu$ *for every* $f \in C_c(X)$ *and every measure finite set is inner compact regular.*

Proof. For any $V \subseteq_{open} X$, we let $\mu(V) = sup\{\wedge f | f \prec V\}$. For any $A \subseteq X$, we let $\mu(A) = inf\{\mu(V) | A \subseteq V \subseteq X\}$. Note that if A in X open, the two definitions coincide. Let $A_F := \{A \subseteq X | \mu(A) < \infty$ and $\mu(A) = sup\{\mu(K) | K \subseteq A\}\}$. Finally, let $A = \{A \subseteq X | A \cap K \in A_F$, for every compact sets$\}$. We claim that $A \xrightarrow{\mu} [0, \infty]$ is an expected measure.
Step 1. $\forall A_j \subseteq X$. $\mu(\cup_j A_j) \leqslant \sum_j \mu(A_j)$.
We may assume that $\mu(A_j) < \infty$ for all $j \in N$. $\forall \epsilon > 0$,$\exists$ open $V_j \supseteq A_j$ s.t. $\mu(V_j) < \mu(A_j) + \frac{\epsilon}{2^j}$. Let $V = \cup V_j$. Recall that $\mu(V) = sup\{\wedge f | f \prec V\}$. For any $f \prec V$, since suppf is compact, $\exists m, suppf \subseteq V_1 \cup ... \cup V_m$ Take a partial partition of unity $\rho_1, ...\rho_m$ for suppf w.r.t. $V_1, ...V_m$. Then $f = \sum_{i=1}^m \rho_i f$, and hence $\wedge f = \sum_{i=1}^m \wedge(\rho_i f) \leqslant \sum_{i=1}^m \mu(V_i) \leqslant \sum_{i=1}^m \mu(V_i) \leqslant \sum_{i=1}^m (\mu(A_i) + \frac{\epsilon}{2^i}) \leqslant \sum_{m=1}^\infty \mu(A_i) + \epsilon$. This proves that $\mu(V) \leqslant \sum_{m=1}^\infty \mu(A_i) + \epsilon$. Let $\epsilon \downarrow 0$.
Step 2. $\forall$ compact $K \subseteq X$, $K \in A_F$ and $\mu(K) = inf\{\wedge f | k \prec f\}$. For any f s.t. $K \prec f$ and $0 < \alpha < 1$, we let $V_\alpha = \{x \in X | f(x) > \alpha\}$. Then $\alpha\mu(K) \leqslant \alpha\mu(V_\alpha) = sup\{\alpha \wedge g | g \prec V_\alpha\} \leqslant \wedge f$. Let $\alpha \uparrow 1$, $\mu(K) \leqslant \wedge f \Rightarrow \mu(K) \leqslant inf\{\wedge f | K \prec f\}$. On the other hand, $\forall \epsilon > 0$, $\exists$ open $V \supseteq K$ s.t. $\mu(V) < \mu(K) + \epsilon$. Choose h s.t. $K \prec h \prec V$. Then, $\wedge h \leqslant \mu(V) < \mu(K) + \epsilon$, and hence $inf\{\wedge f | K \prec f\} < \mu(K) + \epsilon$. Let $epsilon \downarrow 0$.
Step 3. $\forall$ open $V \subseteq X$. $\mu(V) = sup\{\mu(K) | K \subseteq V\}$.

It suffices to prove that $\mu(V) \leqslant sup\{\mu(K)|K \subseteq V\}$. For any $\beta < \mu(V)$, there exists $f \prec V$ s.t. $\beta < \wedge f$. Then $\beta < \wedge f \leqslant \mu(suppf)$: $\mu(suppf) = inf\mu(U)|suppf \subseteq U \subseteq X$ and $\wedge f \leqslant \mu(U)$ if $f \prec U$.
Step 4. $\forall$ disjoint $a_j \in A_F$. $\mu(\cup_j A_j) = \sum_j \mu(A_j)$. If furthermore, $\mu(\cup_j A_j) < \infty$, then $\cup_j A_j \in A_F$.
We claim that $\forall$ compact K and K' in X, $\mu(K \cup K') = \mu(K) + \mu(K')$. $\forall \epsilon > 0$, $\exists f$, $K \cup K' \prec f$ and $\wedge f < \mu(K \cup K') + \epsilon$ by step 2. Applying Urysohn's lemma, $\exists \rho$, $K \prec \rho \prec X \setminus K' \Rightarrow K \prec \rho f$ and $K' \prec (1-p)f$. $\mu(K) + \mu(K') \leqslant \wedge(\rho f) + \wedge((1-\rho)f) = \wedge f < \mu(K \cup K') + \epsilon$. Then, let $\epsilon \downarrow 0$. Now since $Aj \in A_F$, $\exists$ compact $K_j \subseteq A_j$ s.t. $\mu(A_j) - \frac{\epsilon}{2^j}$. $\sum_j \mu(A_j) - \epsilon \leqslant \sum_{j=1}^{\infty} \mu(K_j) = \lim_{m\to\infty} \sum_{j=1}^{m} \mu(K_j) = \lim_{m\to\infty} \mu(\cup_{j=1}^{m} K_j) \leqslant \mu(\cup_{j=1}^{\infty} A_j)$. Let $\epsilon \downarrow 0$. If furthermore $\mu(\cup_{j=1}^{\infty} A_j) < \infty$, then $\exists N$ s.t. $\mu(\cup_j A_j) - \epsilon < \sum_{j=1}^{N} \mu(A_j) < \sum_{j=1}^{N} (\mu(K_j + \frac{\epsilon}{2^j}) < \mu(\cup_{j=1}^{N} K_j) + \epsilon = \mu(\cup_{j=1}^{N} K_j) + \epsilon$. Hence we proved.
Step 5. $\forall A \in A_F$ and $\epsilon > 0$, $\exists K \subseteq A \subseteq V \subseteq X$, $\mu(V \setminus K) < \epsilon$. By the definitions of μ and A_F, $\exists K \subseteq A \subseteq V \subseteq X$ s.t. $\mu(V) < \mu(A) + \frac{\epsilon}{2}$ and $\mu(A) - \frac{\epsilon}{2} < \mu(K)$. Thus, $\mu(V) - \mu(K) < \epsilon$. Since $V \setminus K \subseteq X$ and $\mu(V \setminus K) < \mu(V)$. By step 3, $V \setminus K \in A_F$ and by step 2, $K \in A_F$. Thus, $\mu(V) = \mu(K) + \mu(V \setminus K)$ and hence $\mu(V \setminus K) < \epsilon$.
Step 6. $\forall A_1, A_2 \in A_F$, $A_1 \setminus A_2$, $A_1 \cap A_2$, and $A_1 \cup A_2 \in A_F$.
By step 5. $\forall \epsilon > 0$, $\exists K_i \subseteq A_i \subseteq V_i \subseteq X$ s.t. $\mu(V_i \setminus K_i) < \frac{\epsilon}{2}$. Then $A_1 \setminus A_2 \subseteq V_1 \setminus K_2 \subseteq (V_1 \setminus K_1) \cup (K_1 \setminus V_2) \cup (V_2 \setminus K_2)$, and hence $\mu(A_1 \setminus A_2) \leqslant \mu(V_1 \setminus K_1) + \mu(K_1 \setminus V_2) + \mu(V_2 \setminus K_2) < \mu(K_1 \setminus V_2) + \epsilon$. Note that $K_1 \setminus V_2$ is a compact and $\mu(A_1 \setminus A_2) < \infty$. Thus, $A_1 \setminus A_2 \in A_F$. $A_1 \cap A_2 = A_1 \setminus (A_1 \setminus A_2)$. $A_1 \cup A_2 = (A_1 \setminus A_2) \cup A_2$. Since $\mu(A_1 \cup A_2) \leqslant \mu(A_1) + \mu(A_2) < \infty$. Then, $A_1 \cup A_2 \in A_F$.
Step 7. A is a σ-algebra on X containing B_X.
Let $A \in A$, for any compact $K \subseteq X$, $(X \setminus A) \cap K = K \setminus (A \cap K) \in A_F$ by step 6. Under the complement operation, it is closed.
Let $A_j \in A$. For any compact $K \subseteq X$, set $B_1 = A_1 \cap K$ and $B_n = A_n \cap K \setminus \cup_{j=1}^{n-1} B_j$. Then B_j is a disjoint family in A_F, and hence $\cup_j A_j \cap K = \cup_j B_j \cap K$. Since $\mu(\cup_j (B_j \cap K)) \leqslant \mu(K)$. Therefore, by step 4, A_j is in A_F.
Let $C \subseteq_{closed} X$. For any compact $K \subseteq X$, $C \cap K$ compact$\Rightarrow C \in A_F$.
Step 8. $A_F = \{A \in A | \mu(A) < \infty\}$
$\subseteq$ Let $A \in A_F$. For any compact $K \subseteq X$, $A \cap K \subseteq K$, then $A \in A$.
($\supseteq$) Let $A \in A$ s.t. $\mu(A) < \infty$. By the definition of $\mu(A)$, $\exists$ open $V \supseteq A$ with $\mu(V) < \mu(A) + 1 < \infty$. By step 3, $V \in A_F$. Let $\epsilon > 0$, by step 5, $\exists K \subseteq V$ with $\mu(V \setminus K) < \epsilon$. $A \cap K \in A_F$ since $A \in A$. Thus $\exists$ compact $H \subseteq A \cap K$ with $\mu(A \cap k) < \mu(H) + \epsilon$. $A = (A \cap K) \cup (A \setminus K) \subseteq (A \cap K) \cup (V \setminus K)$. $\mu(A) \leqslant \mu(A \cap K) + \mu(V \setminus K) < \mu(H) + 2\epsilon$. $\mu(A) = sup\{\mu(H) | H \subseteq A\}$.
Step 9. $A \xrightarrow{\mu} [0, \infty]$ is a measure.
Given disjoint $A_j \in A$, if $\mu(A_j = \infty$ for some j, then $\mu(\cup_j A_j) = sum_j \mu(A_j)$, and if every $\mu(A_j) < \infty$, $A \in A_F$, by step 4, $\mu(\cup_j A_j) = sum_j \mu(A_j)$. The completeness of measure can be obtained by applying the results of previous steps.
Step 10. $\forall f \in C_c(X), \wedge f = \int_X f d\mu$.
It suffices to verify that $\wedge f \leqslant \int_X f d\mu$ for every $f \in C_c(X)_R$. Choose a¡b, s.t.

$a < f(x) \leqslant b$ for all $x \in X$ and for $\epsilon > 0$, choose $a = y_1 < ... < y_n = b$ s.t. $y_i - y_{i-1} < \epsilon$. Let $A_i = \{x \in X | y_{i-1} < f(x) \leqslant y_i\} \cap suppf$ which is Borel since f is continuous. Therefore, $A_i \in A$. $\exists$ open $V_i \supseteq A_i$ s.t. $\mu(V_i) < \mu(A_i) + \frac{\epsilon}{n}$. We assume that $f|_{v_i} < y_i + \epsilon$ by replacing V_i by $V_i \cap \{x \in X | f(x) < y_i + \epsilon\}$. Take a partial partition of unity, $\rho_1, ...\rho_n$ for suppf w.r.t. $V_1, ...V_n$. Then $f = \sum_{i=1}^n \rho_i f$ and $suppf \prec \sum_{i=1}^n \rho_i$. Then $\mu(suppf) \leqslant \wedge(\sum_{i=1}^n \rho_i)$. $\rho_i f \leqslant (y_i + \epsilon)\rho_i$. Then $\wedge f = \sum_{i=1}^n \wedge(\rho_i f) \leqslant \sum_{i=1}^n (y_i + \epsilon) \wedge \rho_i = \sum_{i=1}^n (|a| + y_i + \epsilon) \wedge \rho_i - |a| \sum_{i=1}^n \wedge \rho_i \leqslant \sum_{i=1}^n (|a| + y_i + \epsilon)(\mu(A_i + \frac{\epsilon}{n}) - |a|\mu(suppf) = \sum_{i=1}^n (y_i - \epsilon)\mu(A_i) + 2\epsilon \sum_{i=1}^n \mu(A_i) + \sum_{i=1}^n (|a| + y_i + \epsilon)\frac{\epsilon}{n} - |a|\mu(suppf) + |a| \sum_{i=1}^n \mu(A_i) \leqslant \int_X f d\mu + \epsilon(2\mu(suppf) + |a| + b + \epsilon)$. Let $\epsilon \downarrow 0$. □

Remark. *There is another approach of proving. In the above step 1 implies that $\mu_\wedge$ is an outer measure on X. By Carathéodory's construction $A_{\mu_\wedge} = \{A \subseteq X | \forall E \subseteq X, \mu_\wedge(E) = \mu_\wedge(E \cap A) + \mu(E \setminus A)\}$ is an σ-albegra and $\mu_\wedge$ is a complete measure. There can verify directly that $\sigma_{\mu_\wedge} \supseteq T_X$ and hence $A_{\mu_\wedge} \supseteq B_x$. Thus step 2 and the definition of $\mu_\wedge$ imply that $A_{\mu_\wedge} \xrightarrow{\mu_\wedge} [0, \infty]$ a demiregular measure on X. By Cohn 7.2.6, for a demiregular measure on a Hausdorff space every σ-finite measurable sets is $\mu_\wedge$-inner compact regular. Finally step 10 works, another proof is complete.*

6 L^p space

Let (X, A, μ) be a measure space. Recall that for $0 < p < \infty$ and $f \in \mu(A)$, $||f||_p = (\int_X |f(x)|^p d\mu(x))^{\frac{1}{p}} \in [0, \infty]$. $L^p(\mu)_K = \{f \in \mu(A)_K | ||f||_p < \infty\}$.

Definition. *$L^\infty(\mu)_K = \{$all K-valued bounded measurable functions$\}$. For $f \in L^\infty(\mu)_K$, $||f||_\infty = inf\{M | \{x \in X | |f(x)| > M\}$ is μ-finull$\}$.*

Theorem. *(Young's inequality) $\forall a, b \geqslant 0$ and $p, q \geqslant 0$, $\frac{1}{p} + \frac{1}{q} = 1 \Rightarrow ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}$.*

Proof. e^x is convex. $e^{\frac{1}{p}plna + \frac{1}{q}qlnb} \leqslant \frac{1}{p}e^{plna} + \frac{1}{q}e^{qlnb}$. □

Theorem. *(Hölder's inequality) For any conjugate exponents where $p, g \in [0, \infty] \frac{1}{p} + \frac{1}{q} = 1$. If f and g are extended real valued nonnegative A-measurable functions, then $\int X fg d\mu \leqslant ||f||_p ||g||_q$.*

Proof. $(1 < p < \infty)$ We may assume that $||f||_p$ and $||g||_q$ are both finite and nonzero. Let $F = \frac{f}{||f||_p}$ and $G = \frac{g}{||g||_q}$. Then by Young's inequality, $\int_X FG d\mu \leqslant \int_X (\frac{F^p}{p} + \frac{G^q}{q}) d\mu = 1$.
$(p = 1)$ Let $E = \{x \in X | f(x) > 0\}$, which is σ-finite. And $N = \{x \in X | g(x) > ||g||_\infty\}$ $E \cap N$ is μ-null. $f(x)g(x) \leqslant f(x)||g(x)||_\infty$ for all $x \in X \setminus E \cap N$. Then $\int_X |f(x)g(x)| d\mu = \int_{X \setminus E \cap N} |f(x)g(x)| d\mu \leqslant \int_{X \setminus E \cap N} f(x) d\mu ||g||_\infty$. □

Theorem. *(Minkowski's inequality) For any $1 \leqslant p \leqslant \infty$ and $f, g \in \mu(A)$, we may have $||f + g||_p \leqslant ||f||_p + ||g||_p$.*

Proof. ($p < \infty$) We may assume that both $||f||_p$ and $||g||_p$ are both finite. Then $||f+g||_p < \infty$, since $(\frac{|f+g|}{2})^p \leqslant (\frac{|f|+|g|}{2})^p \leqslant \frac{1}{2}|f|^p + \frac{1}{2}|g|^p$. $|f+g|^p \leqslant |f||f+g|^{p-1} + |g||f+g|^{p-1}$. Then by Hòlder's inequality, $\int_X |f+g|^p d\mu \leqslant ||f||_p(\int_X |f+g|^{(p-1)q})^{\frac{1}{q}} + ||g||_p(\int_X |f+g|^{(p-1)q})^{\frac{1}{q}} \Rightarrow (\int_X |f+g|^p d\mu)^{1-\frac{1}{q}} = (\int_X |f+g|^p d\mu)^{\frac{1}{p}} \leqslant ||f||_p + ||g||_p$.
($p = \infty$) $f, g \in L^\infty(\mu) \Rightarrow f+g \in L^\infty(\mu)$. Let $N_1 = \{x \in X | f(x) > ||f||_\infty\}$ and $N_2 = \{x \in X | g(x) > ||g||_\infty\}$. $N_1 \cup N_2$ is μ-finull. Then $|f(x)+g(x)| \leqslant |f(x)| + |g(x)| \leqslant ||f||_\infty + ||g||_\infty$ and hence $||f+g||_\infty \leqslant ||f||_\infty + ||g||_\infty$. □

Corollary. *$(L^p, ||\dot{}||_p)$ is a seminormed space.*

Definition. *$N^p(\mu) = \{f \in L^p(\mu) | ||f||_p = 0\} = \{f | f \underset{=}{\mu} 0\}$ for $0 < p < \infty$ and $\{f | \{x | f(x) > 0\}, \mu - finull\}$ for $p = \infty$ is a vector subspace of $L^p(\mu)$ by using Minkowski's inequality. Let ordinary $L^p = L^p(\mu) \setminus N^p(\mu)$.*
$(L^p(\mu), ||\dot{}||_{L^p})$ where $|| < f > ||_{L^p} = ||f||_p$.

Theorem. *Let (X, A, μ) be a measurable space and $1 \leqslant p \leqslant \infty$. $L^p(\mu)$ is a Banach space.*

Proof. ($1 \leqslant p < \infty$). Let $< f_n >$ be a Cauchy sequence in $L^p(\mu)$. Then f_n is Cauchy sequence w.r.t. the seminorm $||\dot{}||_p$, i.e. $\forall \epsilon > 0, \exists N, \forall n, m \geqslant N \Rightarrow ||f_n - f_m||_p < \epsilon$. By Chebyshev's inequality, for any $\epsilon > 0$, $\mu(\{x | f_n - f_m| > \epsilon\}) \leqslant \frac{1}{\epsilon^p} \int_X |f_n - f_m|^p d\mu = \frac{1}{\epsilon} ||f_n - f_m||_p{}^p$. Thus, $\lim_{N \to \infty, m, n \geqslant N} sup\mu(\{x||f_n - f_m| > \epsilon\}) \leqslant \lim_{n \to \infty, n, m \geqslant N} \frac{1}{\epsilon^p}(||f_n - f_m||_p)^p = 0$. Therefore, there exists a subsequence f_{n_k} wich converges μ-a.e. to some function f. By Fatou's lemma, $\int_X |f_{n_k} - f|^p d\mu = \int_X \lim_{l \to \infty} |f_{n_k} - f_{n_l}|^p d\mu \leqslant \lim_{l \to \infty} \int_X |f_{n_k} - f_{n_l}|^p d\mu \to 0$ as $k \to \infty$ and hence $||f_{n_k} - f||_p < \infty$ if k is sufficiently large since $||f||_p \leqslant ||f_{n_k} - f||_p + ||f_{n_k}||_p < \infty$ by Minkowski inequality. Therefore, $|| < f_{n_k} > - < f > ||_p \to 0$ as $k \to \infty$. Since the subsequence converges, the Cauchy sequence converges$\Rightarrow < f_n > \to < f >$ w.r.t. $||\dot{}||_p$ as $n \to \infty$.
($p = \infty$) Suppose that $< g_n \in L^\infty(\mu)$ s.t. $\sum_{n=1}^\infty ||g_n||_\infty < \infty$. For any $n \in N$, $N_n = \{x \in X | g_n(x) > ||g_n||_\infty\}$ is μ-finull. Since $(K, |\dot{}|)$ is complete, for every $x \in X \setminus \cup_{m=1}^\infty N_n$, the series $\sum_{n=1}^\infty g_n(x)$ converges absolutely. Let $s(x) = \sum_{n=1}^\infty g_n(x)$ if $x \in X \setminus \cup_{m=1}^\infty N_n, = 0$ if $x \in \cup_{n=1}^\infty N_n$. $|s(x) - \sum_{n=1}^m g_n(x)| \leqslant \sum_{n=m+1}^\infty |g_n(x)| \leqslant \sum_{n=m+1}^\infty ||g_n||_\infty$ if $x \in X \setminus \cup_{n=1}^\infty$. Thus, $||s - \sum_{n=1}^\infty g_n||_\infty \leqslant \sum_{n=m+1}^\infty ||g_n||_\infty \to 0$ as $m \to \infty$ and hence $s = \sum_{n=1}^\infty g_n$ in $L^\infty(\mu)$. □

Proposition 9. *The simple functions in $L^p(\mu)$ form a dense subspace of $L^p(\mu)$.*

Proof. We may assume K=R.
$1 \leqslant p < \infty$ For any $f \in L^p(\mu)$, there exits $[0, \infty)$-valued simple functions g_n and h_n s.t. $g_n \uparrow f^+$ and $h_n \uparrow f^-$. Let $f_n = g_n - hn$. $f_n \uparrow f$ pointwise. $|f_n(x) - f(x)|^p \leqslant |f(x)|^p$ for all $x \in X$ and $n \in N$. By LDCT, $\lim_{n \to \infty} ||f_n - f||_p = 0$
$p = \infty$ Given $f \in L^\infty(\mu)$ For any $\epsilon > 0$, there exists $a_0, ..a_k$ s.t. $a_0 < -||f||_\infty < a_1 < ... < a_k = ||f||_\infty$ and $\max_j a_j - a_{j-1} < \epsilon$. Let $g = \sum_{j=1}^k a_j \chi_{A_j}$ where $A_j = f^{-1}((a_{j-1}, a_j])$. Then $g \in L^\infty$. Then $||g - f||_\infty < \epsilon$. □

Corollary. *If $1 \leqslant p < \infty$, X is a LCH, and μ is a demiregular measure s.t. $A \in A$ with $\mu(A) < \infty$ is μ-inner compact regular, then $C_c X \underset{dense}{\subseteq} L^p(\mu)$.*

Proof. For any simple functions $g \in L^p(\mu)$, if we let $A = \{x \in X | g(x) \neq 0\}$, then $\mu(A) < \infty$. By Lusin's theorem, for any $\epsilon > 0$, $\exists h \in C_c(X)$ s.t. $\mu(\{x | h(x) \neq g(x)\}) < \epsilon$ and $sup_x |h| = sup_x |g|$. Then $||g - h||_p = (\int_X |g - h|^p d\mu)^{\frac{1}{p}} \leqslant \epsilon^{\frac{1}{p}} sup_x |g|$. □

6.1 Duality

Suppose that $\frac{1}{p} + \frac{1}{q} = 1$ with $p, q \in [1, \infty]$. We have natural pairing $L^p(\mu) \times L^q(\mu) \xrightarrow{T} K$ i.e. maps $(f, g) \longrightarrow \int_X fg d\mu$ which descends to a pairing $(< f >, < g >) \longrightarrow \int_X fg d\mu$ s.t. $|T(< f > . < g >)| \leqslant || < f > ||_p || < g > ||_q$. $T \in L(L^p, L^q, K)$. T induces a map $L^q(\mu) \xrightarrow{\varphi} L^p(\mu)^* = L(L^p(\mu), K)$. $||\varphi||_{L(L^q(\mu), L^p(\mu)^*)} = sup \frac{||\varphi_{<g>}||_{L^p(\mu)}}{||<g>||_q}$. $||\varphi_{<g>}||_{L^p(\mu)} = sup \frac{|\varphi_{<g>}(<f>)|_K}{||<f>||_p} = \frac{|T(<f>,<g>)|}{||<f>||_p} \leqslant || < g > ||_q$. Thus, $||\varphi||_{L(L^q(\mu), L^p(\mu)^*)} \leqslant 1$.

Theorem. *If $1 \leqslant p < \infty$, then $L^q(\mu) \xrightarrow{\varphi} L^p(\mu)^*$ preserves norms. Actually, ϕ is a surjection if $1 < p < \infty$, or $p = 1$ and μ is σ-finite, or $p = 1$, X is LCH, $(A, \mu) = (A_{\mu_\wedge}, \mu|_{A_{\mu_\wedge}})$, or p=1, G : LCH topological group, μ: demiregular measure on B_G s.t. $\mu(A) \Rightarrow$ inner compact regular.*

Proof. For any $z \in C$ we let

$$sign(z) = \begin{cases} \frac{z}{|z|}, & if \quad z \neq 0 \\ 0, & if \quad z = 0 \end{cases} \tag{2}$$

(1) p=1: Let $g \in L^\infty(\mu)_K$ with $||g||_\infty > 0$. Then for any $\epsilon > 0$, the set $\{x \in X | |g(x)| > ||g||_\infty - \epsilon\}$ is not *mu*-finull, and hence $\exists A \in A$ with $\mu(A) < \infty$ s.t. $B = \{x \in X | |g(x)| > ||g||_\infty - \epsilon\} \cap A, \mu(B) > 0$. If $f() = sig\bar{n}g()\chi_B$, then $f \in L^1(\mu)_K$, $||f||_1 \leqslant \mu(B)$ and $\varphi_{<g>}(< f >) = \int_X sig\bar{n}g()\chi_B(x) g(x) d\mu(x) = \int_B |g| d\mu \geqslant (||g||_\infty - \epsilon)\mu(B) \geqslant (||g||_\infty - \epsilon)||f||_1$. Then $||\varphi_{<g>}|| \geqslant ||g||_\infty - \epsilon$. Let $\epsilon \downarrow 0$.
(2) $1 < p < \infty$ Let $g \in L^q(\mu)_K$ and $f = sig\bar{n}g|g|^{q-1}$. Then $|f|^p = |g|^q \in L^1(\mu)_K$. $\varphi_{<g>}(< f >) = \int_X sig\bar{n}g|g|^{q-1} g d\mu = (||g||_q)^q = (||f||_p)^p$. $(||g||_q)^q = |\varphi_{<g>}(< f >)| \leqslant ||\varphi_{<g>}||_{L^p(\mu)^*} || < f > ||_p = ||\varphi_{<g>}||_{L^p(\mu)^*} (||g||_q)^{\frac{q}{p}}$. Then $||g||_q \leqslant ||\varphi_{<g>}||$. □

6.2 Signed measures and complex measures

Definition. *Let $A \xrightarrow{\mu} \bar{R}(resp.C)$ be a map.*
(1) μ is finitely or countably additive if $\forall$ disjoint $A_n \in A$, $\mu(\cup_n A_n) = \sum_n \mu(A_n)$
(2) μ is signed or complex measure on (X, A) if it is σ-additive and nontrivial $\mu(\emptyset) = 0$.
(3) μ is a finite signed neasure if it is a signed measure which take values in R.

Note that a signed measure is not monotone in general.

A signed measure μ cannot take both ∞ and $-\infty$ as values.

Definition. *Let μ be a signed measure on (X, A) and $A \in A$. We say that A is a positive set if $\forall E \in A, E \subseteq A \to \mu(E) \geqslant 0$.*
$\emptyset$ is both μ-positive and negative.
Countable union of μ-positive (resp. μ-negative) is μ-positive (resp. μ-negative).
μ is monotone on a μ-positive or μ-negative set

Lemma. *μ: signed measure on (X, A) and $A \in A$. If $-\infty < \mu(A) < 0$, then $\exists \mu$-negative set $B \subseteq A$ s.t. $\mu(B) \leqslant \mu(A)$.*

Proof. Let $\delta_1 = sup\{\mu(E)|E \in A \text{ and } E \subseteq A\} \geqslant \mu(\emptyset) = 0$
$\exists E_1 \in A$ s.t. $E_1 \subseteq A$ and $\mu(E_1) \geqslant \min\{\frac{\delta_1}{2}, 1\}$. Then define δ_n and E_n inductively. $\delta_n = sup\{\mu(E)|E \in A \text{ and } E \subseteq A \setminus (E_1 \cup ... \cup E_{n-1})\}$ and $E_n \in A$ s.t. $E_n \subseteq A \setminus (E_1 \cup ... \cup E_{n-1})$ and $\mu(E_n) \geqslant \min\{\frac{\delta_n}{2}, 1\}$. Let $A_\infty = \cup_{n=1}^{\infty} E_n$ and $B = A \setminus A_\infty$. It is easily to see that E_n are disjoint and $\mu(A_\infty) = \sum_{n=1}^{\infty} \mu(E_n) \geqslant 0$ and hence $\mu(B) \leqslant \mu(B) + \mu(A_\infty) = \mu(A)$. Now it remains to prove B is μ-negative. Since $\mu(A) \neq -\infty$ and $\mu(A) = \mu(B) + \mu(A_\infty) \Rightarrow \mu(A_\infty) \neq \infty$ and $\mu(A_\infty) = \sum_n \mu(E_n) < \infty \Rightarrow \lim_{n\to\infty} \mu(E_n) = 0 \Rightarrow \lim_{n\to\infty} \delta_n = 0$. For any $E \in A, E \subseteq B \subseteq A \setminus \cup_{n=1}^{\infty} E_n$, then $\mu(E) \leqslant \delta_n$ and hence $\mu(E) \leqslant 0$. $\square$

Theorem. *(Hahn's decomposition) For any signed measure μ on (X, A). There exists a μ-positive P and μ-negative N s.t. $X = P \cup N$ and $P \cap N = \emptyset$.*

Proof. We may assume that $\mu(E) \neq -\infty$ for every $E \in A$. Let $L = inf\{\mu(A)|$ A is μ-negative $\} \leqslant 0$ and choose a sequence N_m of μ-negative sets s.t. $\lim_{m\to\infty} \mu(N_m) = L$. We define $N = \cup_{m=1}^{\infty} N_m$. N is still a μ-negative set and $\mu(N) \neq -\infty$ by assumption. Since $L \leqslant \mu(N) \leqslant \mu(N_m)$ for all m, $\mu(N_m) \to L \Rightarrow \mu(N) = L$. Now define $P = X \setminus N$. It remains to show that P is μ-positive. Suppose not. $\exists A \in A$ s.t. $A \subseteq P$ and $\mu(A) < 0$. By the above lemma, $\exists\mu$-negative set $B \subseteq A$ s.t. $\mu(B) \leqslant \mu(A)$. Since $B \cap N = \emptyset$, $B \cup N$ is still a μ-negative set with $\mu(B \cup N) < \mu(N) = L$. Contradictory. $\square$

Remark. *The Hahn's decomposition is not unique. But the differences between $(P, N), (P', N')$ are some μ-trivial subsets in A. Jordan's decomposition of a signed measure. For a signed measure μ on (X, A), fix a Hahn decomposition (P,N) for μ. We define $A \xrightarrow{\mu^+} \mu(A \cap P)$ and $A \xrightarrow{\mu^-} -\mu(A \cap N)$. Then μ^+, μ^- are both positive measures on (X, A), at least one of which is finite, s.t. $\mu = \mu^+ - \mu^-$. Then (μ^+, μ^-) is called the Jordan's decomposition of μ, which is independent of the choice of (P,N).*

Proposition 10. *$\mu^{\pm}(A) = sup\{\pm\mu(E)|E \in A$ and $E \subseteq A\}$.*

Proof. For any $E \in A$ with $E \subseteq A$, $\mu(E) = \mu^+(E) - \mu^-(E) \leqslant \mu^+(E) \leqslant \mu^+(A)$ and hence $\geqslant$ holds. On the other hand, $\mu^+(A) = \mu(A \cap P)$, since $A \cap P \subseteq A \Rightarrow \leqslant$ holds. $\square$

Definition. *$|\mu|$ is called the variation of μ. $||\mu|| = |\mu|(X)$ is called the total variation.*

Definition. *Let μ be a complex measure on (X, A). For any $A \in A$, we define $|\mu|(A) = sup\{\sum_{j=1}^{n} |\mu(A_j)||A_1, ..., A_n$ form a partition of A,$n \in N\}$. We call $A \xrightarrow{|\mu|} [0, \infty]$ the variation and $||\mu|| = |\mu|(X)$ the total variation of μ.*

Proposition 11. *For any complex measure μ, $|\mu|$ is a finite positive measure. If $|\mu|(A) \leqslant \nu(A)$ where ν is a positive measure, then $|\mu| \leqslant \nu$.*

Proof. (1) $|\mu|(\emptyset) = 0$
(2) We first prove that $|\mu|$ is finitely additive. Let A and $A' \in A$ be disjoint. If $A \cup A' = \cup_{j=1}^{n} A_j$ where A_j is a partition of $A \cup A'$, then $\sum_{j=1}^{n} = \sum_{j=1}^{n} |\mu(A_j \cap A) + \mu(A_j \cap A'| \leqslant \sum_{j=1}^{n} |\mu(A_j \cap A)| + \sum_{j=1}^{n} |mu(A_j \cap A'| \leqslant |\mu|(A) + |\mu|(A')$ and hence $|\mu|(A \cup A') \leqslant |\mu|(A) + |\mu|(A')$. On the other hand, for any number $M < |\mu|(A) + |\mu|(A')$ we may choose $M_1, M_2 \in R$ s.t. $M = M_1 + M_2, M_1 < |\mu|(A)$ and $M_2 < |\mu|(A')$. Then $\exists$ partitions B_j, $A = \cup_{j=1}^{l} B_j$ and B'_k, $A' = \cup_{k=1}^{m} B'_k$ s.t. $M_1 < \sum_{j=1}^{l} |\mu(B_j)|$ and $M_1 < \sum_{k=1}^{m} |\mu(B'_k)|$. Therefore, $M = M_1 + M_2 < \sum_{j=1}^{l} |\mu(B_j)| + sum_{k=1}^{m} |\mu(B'_k)| \leqslant |\mu|(A \cup A')$ since B'_k, B_j together form a partition of $A \cup A'$. So $|\mu|(A) + |\mu|(A') \leqslant |\mu|(A \cup A')$.
Then, it suffices to show that for any $E_n \in A$, if $E_n \downarrow \emptyset$ then $|\mu|(E_n) \to 0$ as $n \to \infty$. $\forall A \in A$, $|\mu|(A) \leqslant |Re\mu|(A) + |Im\mu|(A)$ where $(Re\mu)() = Re(\mu()), (Im\mu)() = Im(\mu())$. It is easily to see that $|Re\mu|(A), |Im\mu|(A)$ are finite positive measure. Therefore, $|Im\mu|(E_n), |Re\mu|(E_n) \to 0$ as $n \to \infty$. Thus, $|mu|$ is σ-additive and finite. □

Remark. *If μ is a finite signed measure on (X, A), then the two definitions of $|\mu|$ coincide.*

Proposition 12. *$(M(X, A)_K, || \cdot ||)$ is a Banach space.*

Proof. It is easily to see that $(M(X, A)_K, || \cdot ||)$ is a normed space. We then want to prove every Cauchy sequence converges. Let $\mu_n \in M(X, A)_K$ be a Cauchy sequence w.r.t. total variation. Then, as a sequence of K-valued functions on A (σ-algebra), μ_n is uniformly Cauchy. For any $A \in A$, $|\mu_n(A) - \mu_m(A)| \leqslant |\mu_n - \mu_m|(A) \leqslant ||\mu_n - \mu_m||$. Thus, μ_n converge uniformly to a function $A \overset{\mu}{K}$. $\mu(\emptyset) = \lim_{n \to \infty} \mu_n(\emptyset) = 0$. We claim that μ is σ-additive. Then it suffices that for any $E_n \in A$, if $E_n \downarrow \emptyset$ then $\mu(E_n) \to 0$. (Noting that finite additivity is obvious)
$\forall \epsilon > 0, \exists N > 0, \forall A \in A, n \geqslant N$, by uniformly convergence, $|\mu(A) - \mu_n(A)| < \frac{\epsilon}{2}$. Since μ_N is a finite measure, $\exists M > 0$, $|\mu_N(E_m)| < \frac{\epsilon}{2}$ if $m \geqslant M$. Thus, $|\mu(E_m)| \leqslant |\mu(E_m) - \mu_N(E_m)| + |\mu_N(E_m)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ if $m \geqslant M$. In other words, $\mu(E_m) \to 0$ as $m \to \infty$. Therefore, $\mu \in M(X, A)_K$. Eventually, $||\mu_n - \mu|| \to 0$ as $n \to \infty$ can be easily proved, so we ignore. □

6.3 Hilbert space

Theorem. *(Orthogonal Projection Theorem) Let C be a closed convex set in a Hilbert space E and $x \in E$, then there is unique $y \in C$ such that $||y - x|| = int_{z \in C}||x - z|| = \min_{z \in C} ||x - z||$*
Furthermore, y is characterized by $x - y \in M^{\perp}$.

Proof. Let $\alpha = inf_{z \in M}||x - z||$. There is a sequence y_n in M such that $\alpha_2 \leqslant ||x - z|| \leqslant \alpha^2 + \frac{1}{n}$. By parallelogram property, $||(y_n - x) - (y_m - x)||^2 + ||(y_n - x) + (y_m - x)||^2 = 2(||y_n - x||^2 + ||y_m - x||^2) \leqslant 4\alpha^2 + \frac{2}{m} + \frac{2}{n}$. Therefore, $||y_n - y_m||^2 \leqslant 4\alpha^2 + \frac{2}{m} + \frac{2}{n} - 4||\frac{y_m + y_n}{2} - x||^2 \leqslant \frac{2}{n} + \frac{2}{m}$ since C is convex. Then y_n is a Cauchy sequence. Since M is a subspace of Hilbert space, then M is still complete and hence there is $y \in M$ such that $\lim_{n \to \infty} ||y_n - y|| = 0$.
Now let y be any element of M which satisfies $||x - y|| = \min_{z \in M} ||x - z||$. For $t \in R$ and $z \in M$, we have $||x - y - tz||^2 = ||x - y||^2 - 2Re(x - y, z)t + t^2||z||^2 \geqslant ||x - y||^2$. Then $-2Re(x-y, z)t + t^2||z||^2 \geqslant 0$. If $t > 0$, $-2Re(x-y, z) + t||z|| \geqslant 0$, then drive $t \downarrow 0$ and we have $-2Re(x-y, z) \leqslant 0$. If $t < 0$, $-2Re(x-y, z) + t||z|| \leqslant 0$, then drive $t \uparrow 0$ and we have $-2Re(x - y, z) \geqslant 0$. Then $Re(x - y, z) = 0$ By the similar way, we take iz and have $Im(x - y, z) = 0$. Thus, $x - y \perp z$.
Now we prove for uniqueness. Suppose that there are two $y, y' \in M$ with $\min_{z \in M} ||x - z|| = ||y - x|| = ||y' - x||$, $||y - y'||^2 = 2(||y - x||^2 + ||y' - x||^2) - ||y + y' - 2x||^2 \leqslant 0$. Then $y = y'$. □

Corollary. *If F is a closed vector subspace of E, then P_F is a linear map from E onto F with $||P_F|| = 1$ $F \neq \{0\}$ and the decomposition is unique.*

Theorem. *Let H_1 and H_2 be two Hilbert spaces with inner product $< \cdot >_1$ and $< \cdot >_2$ respectively. The following are equivalent.*
(1) $U : H_1 \longrightarrow H_2$ is an isometric isomorphism.
(2) $U : H_1 \longrightarrow H_2$ is a surjective isometry.
(3) $U : H_1 \longrightarrow H_2$ is surjective and $< Ux, Uy >_2 = < x, y >_1 \forall x, y \in H_1$.
(4) $U^* : H_2 \longrightarrow H_1$ s.t. $< Ux, y >_2 = < x, U^*y >_1, \forall x \in H_1$ and $\forall y \in H_2$ is the inverse of U.

Proof. $1 \Rightarrow 2$ is obvious.
$2 \Rightarrow 3$ follows from considering $< U(x+\alpha y), U(x+\alpha y) >_2 = < x+\alpha y, x+\alpha y >_1$ by isometry which leads to $< Ux, Ux >_2 = < Ux, U\alpha y >_2 + < U\alpha y, Ux >_2 + < U\alpha y, U\alpha y >_2 = < x, x >_1 +/, x\alpha y >_1 + < \alpha y, x >_1 + < \alpha y, \alpha y >_1$. Using the fact that U is isometric, we have $E(\alpha) = < Ux, \alpha Uy >_2 + < \alpha Uy, Ux >_2 = < x, \alpha y >_1 + < \alpha y, x >_1$. By considering $E(1)$ and $E(i)$ cases, we can find $< Ux, Uy >_2 = < x, y >_1 \forall x, y \in H_1$.
The proof of $3 \Rightarrow 4$ is straightforward: $\forall x, y \in H_1$, $< x, y >_1 - < Ux, Uy >_2 = < x, y >_1 - < U^*ux, y >_1 = < x - U^*Ux, y >_1$. But this means $U^*Ux = x$ for all $x \in H_1$. Since U is surjective, it implies U^*U and UU^* are equivalent to I. Therefore, U^* is a inverse map of U. Since U has an inverse it must be a bijection. Moreover, $||Ux||_2^2 = < Ux, Ux >_2 = < x, U^*Ux >_1 = < x, x >_1 = ||x||_2^2$. □

Proposition 13. *H: Hilbert space.*
$M \oplus M^{\perp} \xrightarrow{U} H$ *is an isomorphism of Hilbert spaces.*

Proof. By the previous theorem, it suffices to show that $M \oplus M^{\perp} \xrightarrow{U} H$ is surjective and $< U(m, m^{\perp}), U(n, n^{\perp}) >_H =< (m, m^{\perp}), < n, n^{\perp}) >_{M \oplus M^{\perp}}$.
First the surjectivity: Denote by P the orthogonal projection on M. Then for any $h \in H$, we have $h = Ph + (h - Ph)$. By definition of P, $Ph \in M$. Furthermore, by the orthogonal projection theorem, $h - Ph \perp M$; that is $h - Ph \in M^{\perp}$. It follows that $h = U(Ph, h - Ph)$, showing U is a surjection.
It remains to check U preserves the inner product: $< U(m, m^{\perp}), U(n, n^{\perp}) >_H =< m + m^{\perp}, n + n^{\perp} >_H =< m, n >_H + < m^{\perp}, n >_H + < m, n^{\perp} >_H + < M^{\perp}, n >_H =< m, n >_H + < m^{\perp}, n^{\perp} >_H =< (m, m^{\perp}), (n, n^{\perp}) >_{M \oplus M^{\perp}}$. ☐

Corollary. *If $L]inL(H, K)$, $\exists! y \in H$, s.t. $\forall x \in H, Lx = (x, y)_h$. If $Lx = 0$ for all x, then take y=0. Otherwise, define $M = \{x : Lx = 0\}$. The linearity of L shows that M is a subspace. The continuity of L shows that M is closed. Then by the orthogonal projection theorem, $M^{\perp}$ does not consist of 0 alone. Hence there exists $z \in M^{\perp}$, with $||z|| = 1$. Put $u = (Lx)z - (Lz)x$. Since $Lu = (Lx)(Lz) - (Lz)(Lx) = 0$, we have $u \in M$. Thus, (u,z)=0. Thus, $Lx = (Lx)(z, z) = (Lz)(x, z)$. Take $y = \alpha z$, where $\bar{\alpha} = Lz$. The uniqueness of y is easily proved, for if $(x, y) = (x, y')$ for all $x \in H$, set $z = y - y'$. Then, $(x, z) = 0$ for all $x \in H$; in particular, $(z, z) = 0$, hence $z = 0$.*

Corollary. $L^2(\mu)_K \overset{\Phi^2}{L} (\mu)^*_K$. Φ *is isomorphic.*

Theorem. *(Equivalences of definitions of closed linear span) Let H be a Hilbert space over K and let $A \subseteq H$. The following definitions of the concept of closed linear psan of A are equivalent:*
(1) $s\bar{p}an(A) = \cap M$, where M consists of all closed linear subspaces M of H with $A \subseteq M$.
(2) $s\bar{p}an(A)$ is the smallest linear subspace M of H s.t. $A \subseteq M$.
(3) $s\bar{p}an(A) = cl(\{\sum_{k=1}^{n} \alpha_k f_k : n \in N, \alpha_i \in F, f_i \in A\})$

Proof. Let the proposition (1) holds; assume that the closed linear subspace M' contains the set A, then because $M' \in M$, w e have $s\bar{p}an(A) \subseteq M'$. We claim that the intersection of arbitrary family of subspaces is a subspace. Suppose C is a family of subspaces. Denotes $\cap C = \{f \in H|$ for any $V \in C$, there $f \in V\}$. If $f \in \cap C$, then for any $V \in C$, $f \in V$, there $\alpha f \in V$ for $\alpha \in F$. If $f, g \in \cap C$, we have for any $V \in C, f + g \in V$. Therefore, $s\bar{p}an(A)$ is a subspace and in addition it is closed,as intersection of arbitrarily family of closed closed sets is closed.
Next if (2) holds. Since $A \subseteq s\bar{p}an(A), s\bar{p}an(A) \in M$. $s\bar{p}an(A)$ is the smallest one in M; hence $s\bar{p}an(A) = \cap M$. then, the equivalence between (1) and (2) are established. Finally we come to (3). We claim that $cl(span(A))$ is a subspace. It is easily to check by the definition of closure. Now we need to establish the equivalence of (2) and (3): $cl(span(A))$, the closed linear subspace, contains span(A) and thus contains A. For any closed linear subspace M which contains

A, $span(A) \subseteq M$ since the linear span of A is the smallest subspace that contains A. Because M is closed, $cl(span(A)) \subseteq M$. $cl(span(A))$ is the smallest closed linear subspace Mof H with $A \subseteq M$. Because arbitrayr intersection of closed sets is closed and arbitrary intersection of subspaces is a subspace, the smallestness is unique. $\square$

Theorem. *Let $S \subset H$ be any subset of a Hilbert space H. Then $s\bar{pan}S = (S^{\perp})^{\perp}$. That is, $y \in s\bar{pan}S$ if and only if y is perpendicular to everything that is perpendicular to S: $< y, z >= 0$ for all z such that $< x, z >= 0$ for all $x \in S$.*

Proof. Recall that a closed subspace Y satisfies $(Y^{\perp})^{\perp} = Y$. Thus,
$barspan(S) = (s\bar{pan}(S)^{\perp})^{\perp} = (S^{\perp})^{\perp}$. It suffices to show that $s\bar{pan}(S)^{\perp} = S^{\perp}$. Since $S \subset s\bar{pan}(S)$ we clearly have $(s\bar{pan}(S))^{\perp} \subseteq S^{\perp}$. On the other end, if $z \in S^{\perp}$. Thus z is perpendicular to span S and by continuity of the scalar product $z \perp sp\bar{a}nS =$
$barspanS$. Thus, $S^{\perp} \subset (s\bar{pan}S)^{\perp}$. $\square$

Lemma. *Let S be an orthonormal set of vectors in a Hilbert space H. Then the span S consists of all vectors of the for*

6.4 Absolute continuity and singularity

Definition. *Given $\mu \in M(X, A)_{\geqslant}$ and $\nu \in M(X, A)_K$, we say that ν is absolute continuous w.r.t. μ if $\forall A \in A$, $\mu(A) = 0 \Rightarrow \nu(A) = 0$. We denote this relation $\nu << \mu$.*

Definition. *Given a measure ν on (X, A) and $A \in A$, we say that ν is concentrated on A if $\forall E \in A$, $\nu(E) = \nu(E \cap A)$ or equivalently, $\nu(F) = 0, \forall F \in A$ with $F \cap A = \emptyset$.*

Definition. *The two measures are mutually singularly if $\exists A_1, A_2 \in A$, disjoint s.t. ν_j is concentrated on $A_j (j = 1, 2)$. Denote $\nu_1 \perp \nu_2$.*

Lemma. *For a positive finite measure μ on $(X, A), f \in L^1(\mu)$ and S is closed in C, if $A_E(f) = \frac{1}{\mu(E)} \int_E f d\mu \in S$ for every $E \in A$ with $\mu(E) > 0$, then $f(x) \in S$ for μ-a.e. $x \in X$.*

Proof. Write $C \setminus S = \cup_{n=1}^{\infty} B_{r_n}(a_n)$. If $\mu(f^{-1}(C \setminus S)) \neq 0$, then $\exists n \in N$ s.t. $\mu(E) > 0$ where $E = f^{-1}(B_{r_n}(a_n)$. $|A_E(f) - a_n| = |\frac{1}{\mu(E)} \int_E (f - a_n) d\mu| \leqslant \frac{1}{\mu(E)} \int_E |f - a_n| d\mu < r_n$. Contradictory! $\square$

Lemma. *$\forall\sigma$-finite positive measure μ on (X, A), $\exists w \in L^1$ s.t. $0 < w < 1$.*

Proof. Suppose $X = \cup_{n=1}^{\infty} A_n$ for some $A_n \in A$ with $\mu(A_n) < \infty$. We may make $w = \sum_{n=1}^{\infty} \frac{1}{2^n(1+\mu(A_n))} \chi_{A_n}$. $\square$

Theorem. *Let μ be a σ-finite positive measure and $\nu \in M(X, A)_C$.*
(1) (Lebesgue's decomposition) $\exists!(\nu_a, \nu_s) \in M(X, A)_C \times M(X, A)_C$ s.t. $\nu = \nu_a + \nu s$, $\nu_a << \mu$ and $\nu_s \perp \mu$. If furthermore, ν is a finite signed measure

(respectively finite measure), so are ν_a and ν_s.
(2) (The Radon-Nikodym theorem) $\exists! < h > \in L^1(\mu)$ s.t. $\nu_a(E) = \int_E h d\mu$ for every $E \in A$.

Proof. (von Neumann) Both the uniqueness parts of (1) and (2) are clear.
For the rest parts, we only need to deal with the case that ν be a finite positive measure. By the previous lemma, $\exists w \in L^1(\mu)$ with $0 < w < 1$. Consider the measure $\phi(E) = \nu(E) + \int_E w d\mu$, a finite positive measure. Then $\int_X f d\phi = \int_X f d\nu + \int_X f w d\mu$ for every nonnegative real extended measurable functions. (By the monotone convergence theorem)
We claim the map $f \in L^2(\phi) \longrightarrow \int_X f d\nu \in C$ is a continuous linear function w.r.t. $||\cdot||_{L^2(\phi)}$. $|\int_X f d\nu| \leqslant \int_X |f| d\nu \leqslant \int_X \nu(X)^{\frac{1}{2}} ||f||_{L^2(\nu)} \leqslant \nu(X)^{\frac{1}{2}} ||f||_{L^2(\phi)}$.
We further claim that $\exists$ measurable function with $0 \leqslant g \leqslant 1$ and $\int_X f d\nu = \int_X f g d\phi$ for every $f \in L^2(\phi)$. By Riesz representation theorem for Hilbert spaces, $\exists g \in L^2(\phi)$ s.t. $int_X f d\nu = \int_X f g d\phi$ for every $f \in L^2(\phi)$. For $E \in A$, then $\chi_E \in L^2(\phi)$, and hence $\nu(E) = \int_E g d\phi$. If $\phi(E) > 0$, then $\frac{1}{\phi(E)} \int_E g d\mu = \frac{\nu(E)}{\phi(E)} \in [0,1]$. By the previous lemma, $g(x) \in [0,1]$ for ϕ a.e. $x \in X$. Redefining $g(x)$ to be 0 if $g(x)$ not in $[0,1]$.
Now that $\int_X f d\nu = \int_X f g d\nu + \int_X f g w d\mu$. $\int_X f(1-g) d\nu = \int_X f g w d\mu$. We claim that if $A = \{x | 0 \leqslant g(x) < 1\}$ and $B = \{x | g(x) = 1\}$, then $\nu_a(E) = \nu(E \cap A)$ and $\nu_s(E) = \nu(E \cap B)$ form a Lebesgue's decomposition. $\nu = \nu_a + \nu_s$ is clear since A and B cover X and are disjoint. Invert $(1-g)$. Consider $1 + g + ... + g^n$. Setting $f = \chi_E(1+g+...g^n)$, $\int_E (1-g^{n+1}) d\nu = \int_E (1+g+...g^n) g w d\mu$. Note that $1 - g^{n+1} \uparrow \chi_A$ as $n \to \infty$. Besides on $A = X \backslash B$, $1+g+...g^n) g w \uparrow \frac{gw}{1-g}$ as $n \to \infty$. By the monotone convergence theorem, $\nu(E \cap A) = \int_E \chi_A d\nu = \int_{E \cap A} \frac{gw}{1-g} d\mu$ since setting $f = \chi_B$, $0 = \int_B w d\mu$, and hence $\mu(B) = 0$. In addition, we can see that $\nu_s \perp \mu$. Finally, let

$$h(x) = \begin{cases} \frac{g(x)w(x)}{1-g(x)}, & x \in A \\ 0, & x \in B \end{cases} \tag{3}$$

Then $\nu_a(E) = \int_{E \cap A} \frac{gw}{1-g} d\mu = \int_E h d\mu$. □

7 Differentiation

Theorem. *If $R \xrightarrow{F} R$ is monotone, F' exists a.e.*

Lemma. *If $R \xrightarrow{F} R$ is monotone nondecreasing,*
(1) $F(x^{\pm}) = \lim_{y \to x^{\pm}} F(y)$ both exist for $x \in R$
(2) $D = \{x \in R | F \quad is \quad discontinuous \quad at \quad x\}$ is countable
(3) $x \in R \xrightarrow{G} F(x^+)$ is nondecreasing and right continuous.

Proof. (1) Actually $F(x^+) = \inf_{x<y} F(y)$ and $F(x^-) = \sup_{x>y} F(y)$.
(2) For any $x_1 < x_2$, we have $F(x_1^+) \leqslant F(x_2^-)$. Note that $x \in D \Leftrightarrow F(x^-) < F(x^+)$. For each $x \in D$ we choose $r_x \in (F(x^-), F(x^+)) \cap Q$. The map is an

injection. Thus the cardinality of D is equal to or smaller than the cardinality of the rationals, and thus countable.
If $x_1 < x_2$, $F(x_1^+) \leqslant F(x_2^-) \leqslant F(x_2) \leqslant F(x_2^+)$. Thus, let $G(x) = F(x^+)$. G is monotone. $\forall M > G(x) = F(x^+) = \inf_{y>x} F(y)$, $\exists y_0 > x$ and $F(y_0) < M$. $G(y) = F(y^+) < F(y_0) < M$ if $y \in (x, y_0)$.
Proof of theorem
(Reduction to the right continuous case) We follow the notation in the lemma. Note that G coincides with F at least on $R \setminus D$. Besides, if $G(c) = F(c)$, $\frac{F(x)-F(c)}{x-c}$ lies between $\frac{G(x^-)-G(c)}{x-c}$ and $\frac{G(x)-G(c)}{x-c}$. In particular, if $G'(c)$ exists, then $\lim_{x\to c} \frac{G(x^-)-G(c)}{x-c} = G'(c)$.
(Reduction to the bounded case) Replace F by $F^{[-n,n]}(n \in N)$.
Now assume that F is bounded, monotone, and right continuous s.t. $F(-\infty) = 0$. Let $\mu = \mu_F$ (the finite Borel measure on R s.t. $\mu_F = F$). Take Lebesgue's decomposition $\mu = \mu_a + \mu_s = hd\lambda + \mu_s$ w.r.t. λ on R, $\frac{F(x_n)-F(c)}{x_n-c} = \frac{\mu_a((c,x_n])}{\lambda((c,x_n])} + \frac{\mu_s((c,x_n])}{\lambda((c,x_n])} \to h(c) + 0 = h(c)$. $\square$

Given a function $R \xrightarrow{F} R$ and an interval $[a, b]$, when can we conclude that (1) F' exists a.e. (2) $F' \in L^1([a, b])$ and (3) $F(x) - F(a) = \int_a^x F'(t)$ for all $x \in [a, b]$?

Find necessary condition for (1)(2)(3).

Suppose that F satisfies (1)+(2). Then $\nu(E) = \int_E F'd\lambda$ is a finite signed measure on $[a, b]$ s.t. $\nu << \lambda_{[a,b]}$. For any interval $[\alpha, \beta] \subseteq [a, b]$, $\nu([\alpha, \beta]) = \int_{[\alpha,\beta]} F'd\lambda$. If (3) holds, then $\nu([\alpha, \beta]) = F(\beta) - F(\alpha)$.
In summary, F satisfies (1)(2)(3) only if $\forall\epsilon > 0, \exists\delta > 0$, s.t. $\forall$ countable disjoint family $[\alpha_j, \beta_j] \subseteq [a, b]$,

$$\sum_j(\beta_j-\alpha_j) < \delta \Rightarrow \sum_j |F(\beta_j)-F(\alpha_j)| = \sum | \int_{(\alpha_j,\beta_j)} F'd\lambda| \leqslant \int_{\cup(\alpha_j,\beta_j)} |F'|d\lambda < \epsilon$$

by the property of the absolute continuity of measure.

Lemma. *($\epsilon - \delta$ characterization of absolute continuity of measures) $\mu << \nu \Leftrightarrow \forall\epsilon, \exists\delta$ s.t. $\forall A \in A, \nu(A) < \delta \Rightarrow \mu(A) < \epsilon$.*

Proof. ($\Leftarrow$) It is obvious. ($\Rightarrow$) Suppose not, then there exists ϵ s.t. for all $E_n(n \in N)$ with $\nu(E_n) < 2^{-n}$ and $\mu(E_n) \geqslant \epsilon$. Then let $F_k = \cup_{i=k}^{\infty} E_i$ and $F = \cap_{k=1}^{\infty} F_k$. $\nu(F_k) < 2^{1-k} \Rightarrow \nu(F) = 0$. However, $\mu(F) \geqslant \epsilon$ leads to contradiction! $\square$

Lemma. *Let $[a, b] \xrightarrow{F} R$ be continuous and nondecreasing. The followings are equivalent.*
(1) F is AC on $[a, b]$.
(2) F maps sets of measure 0 to sets of measure 0.
(3) F is differentiable a.e. and $F' \in L^1$. Besides $F(x) - F(a) = \int_a^x F'(t)dt$.

Proof. (1)⇒(2) Let $A \subseteq R$ be λ-null. To show that $F(A)$ is λ-null it is harmless to assume that $A \subseteq (a, b)$. For $\epsilon > 0$, let $\delta > 0$ be as given by the definition of AC functions. Since λ is outer regulat, there exists $V \subseteq (a, b)$, and V is open s.t. $A \subseteq V$ and $\lambda(V) < \lambda(A) + \delta = \delta$. Note that V may be written as the union of a countable disjoint family of open intervals $(\alpha_j, \beta_j)(j \in N)$. Then $\sum_j(\beta_j - \alpha_j) = \lambda(V) < \delta$, and hence $\lambda(F(A)) \leqslant \lambda(F(V)) \leqslant \lambda(\cup_j F([\alpha_j, \beta_j]) \leqslant \sum_j \lambda(F([\alpha_j, \beta_j])) = \sum_j |F(\alpha_j) - F(\beta_j)| < \epsilon$.
(2)⇒(3) First assume that F is strictly increasing. Let $m_{[a,b]} = \{A \in m_L | A \subseteq [a, b]\}$. We show that F maps $m_{[a,b]}$ into m_L. Let $A \in m_{[a,b]}$. Then $A = C \cup N$ for some σ-compact C and some λ-null N. Thus $F(A) = F(C) \cup F(N) \in m_L$. Now we define $\mu(A) = \lambda(F(A))(A \in m_{[a,b]})$ and is a measure by the infectivity of F. In addition, $\mu << \lambda_{[a,b]}$ by (2). By the Radon-Nikodym theorem, $\exists h \in L^1(\lambda_{[a,b]})$ s.t. $\mu = hd\lambda_{[a,b]}$. In particular, for a4Fny $x \in [a, b]$, we have $\int_{[[a,x]} hd\lambda = \mu([a, x]) = \lambda(F([a, x])) = F(x) - F(a)$. By the easy part of the fundamental theorem of calculus, we have proved the equivalence. □

If F is only strictly nondecreasing, we can extend the result further to the nondecreasing functions G(x) by letting $G(x) = x + f(x)$.

Lemma. *If* $[a, b] \xrightarrow{F} R$ *is AC on* $[a, b]$*, then so is* V_F*.*

Proof. For $\epsilon > 0$ let $\delta > 0$ be as given by the definitions of AC.Consider any non-overlapping family $[\alpha_j, \beta_j] \subseteq [a, b](j \in N)$. For any $\eta > 0$ and $j \in N$, choose $\alpha_j = t_0^{(j)} \leqslant ... \leqslant t_{k_j}^{(j)} = \beta_j$ s.t. $V_F[\alpha_j, \beta_j] - \frac{]eta}{2^j} < \sum_{l=1}^{k_j} |F(t_l^{(j)}) - F(t_{l-1}^{(j)})|$. Then we have $\sum_{j=1}^{\infty} |V(\beta_j) - V(\alpha_j)| = \sum_{j=1}^{\infty} V_F[\alpha_j, \beta_j] < \sum_{j=1}^{\infty} \sum_{l=1}^{k_j} |F(t_l^{(j)}) - F(t_{l-1}^{(j)})| + \eta$ since $[t_{l-1}^{(j)}, t_l^{(j)}]$ are non-overlapping and the choose of η is arbitrary. □

Corollary. *If* $[a, b] \xrightarrow{F} R$ *is AC on* $[a, b]$*, then* F' *exists a.e. and is integrable on* $[a, b]$ *and* $F(x) - F(a) = \int_a^x F'^{]}(t)dt$*.*

Proof. $F = \frac{1}{2}(V + F) - \frac{1}{2}(V - F)$. □

Lipschitz condition implies AC.

8 Tangent Vectors and Tangent Maps

Let M be a manifold diffeomorphic to dimension n and p ∈ M. Consider $M(p) := \{I \xrightarrow{\nu} M | I \ \ an \ \ open \ \ neighborhood \ \ of \ \ 0 \in R \ \ and \ \ r \ \ differentiable \ \ at \ \ t = 0 \ \ and \ \ r(0) = p\}$. For any $I_1 \xrightarrow{r_1} M$ and $I_2 \xrightarrow{r_2} M$ in M(p), we say that $r_1 \overset{\sim}{\sim}_p r_2$ if there exists a chart $\varphi \in \Phi_M$

9 Oriented manifolds and orientation

Definition. *Let M be a C^∞ manifold of dim m. Two charts $\varphi_\alpha, \varphi_\beta \in \Phi_M$ with coordinates $x_\alpha^1, ..., x_\alpha^m$ and $x_\beta^1, ..., x_\beta^m$ respectively are said to have compatible orientations if $det(\partial x_\alpha / \partial x_\beta) > 0$ for all $p \in U_\alpha \cup U_\beta$.*

10 Oriented integration of differential top forms on manifolds

C^∞ manifold of dimension m which is oriented by maximal compatible C^∞ atlas $\tilde{\Phi}_M$. We are talking about the notion $\int_M \omega$ for $\omega \in A^m(M)$

Definition. *Choose a C^∞ partition of unity ρ_j ($j \in J$) of M subordinate to the open cover $\{U_\varphi | \varphi \in \tilde{\Phi}_M\}$ (say $supp\rho_j \subseteq U_{\varphi_j}$ for some φ_j) so that $supp\rho_j$ is compact for every $j \in J$. For any $\omega \in A^m(M)$, we let $\omega_j = \omega_{\varphi_j} \in A^m(V_j)$. Thus $\rho_j\omega \in A_c^m(M)$ has local expression $(\rho_j \circ \varphi_j^{-1})(x_j) f_j(x_j)$ which can be viewed as an element of $A_c^m((-\infty, 0] \times R^{m-1})$. If $\sum \int_{(-\infty,0]\times R^{m-1}} (\rho_j \circ \varphi_j^{-1})(x_j) f_j(x) dx^1 ... dx^m$ exists and has the same value for all choices of such partitions of unity ρ_j and φ_j, we call this value the integral of ω on M denoted by $\int_M \omega$.*

Proposition 14. *If $\omega \in A_c^m(M)$, then $\int_M \omega$ exists.*

Proof. Suppose that ρ_j and ρ'_k are two smooth partitions of unity subordinate to $\{U_\varphi | \varphi \in \tilde{\Phi}_M\}$ so that $supp\rho_j \subseteq U_{\varphi_j}$ is compact and the same relation holds for ρ'_k where $\varphi_j, \varphi'_k \in \tilde{\Phi}_M$. Do $\sum_{j\in J} \int_{(-\infty,0]\times R^{m-1}} (\rho_j \circ \varphi_j^{-1})(x) f_j(x) dx^1 ... dx^m$ and $\sum_{k\in K} \int_{(-\infty,0]\times R^{m-1}} (\rho'_k \circ \varphi_k'^{-1})(x') f'_k(x) dx^1 ... dx^m$ exist and take the same value? Essentially, the sum is a finite sum since $supp\omega$ is compact and $supp\rho_j$ is strongly locally finite.
$\sum_{j\in J} \int_{(-\infty,0]\times R^{m-1}} (\rho_j \circ \varphi_j^{-1})(x) f_j(x) dx^1 ... dx^m =$
$\sum_{j\in J} \int_{(-\infty,0]\times R^{m-1}} \sum_{k\in K} (\rho'_k \circ \varphi_j^{-1})(x)(\rho_j \circ \varphi_j^{-1})(x) f_j(x) dx^1 ... dx^m =$
$\sum_{j\in J} \sum_{k\in K} \int_{(-\infty,0]\times R^{m-1}} (\rho'_k \rho_j \circ \varphi_j^{-1})(x) f_j(x) dx^1 ... dx^m =$
$\sum_{j\in J} \sum_{k\in K} \int_{(-\infty,0]\times R^{m-1}} (\rho'_k \rho_j \circ \varphi_j^{-1})(\varphi_j \varphi_k^{-1}(x')) f_j(\varphi_j \varphi_k^{-1}(x'))$
$|det(\partial x / \partial x')| dx'^1 ... dx'^m = \sum_{k\in K} \int_{(-\infty,0]\times R^{m-1}} (\rho'_k \circ \varphi_k'^{-1})(x') f'_k(x) dx^1 ... dx^m$
(by positive orientation) □

11 Stokes' theorem

11.1 Origin

Notation. *If Z is an oriented C^∞ manifold of dimension d and $Z \xrightarrow{f} M$ a C^∞ map to a C^∞ manifold M, for any $\omega \in A^d(M)$, we have $f^*\omega \in A^d(Z)$ and* *hence we can talk about whether $\int_Z f^*\omega$ exists. If $\int_Z f^*\omega$ exists, we often write $\int_Z \omega$ if f is clear in the context.*

For example, for any $\tau \in A^{dimM-1}(M)$, we define $\int_{\partial M} \tau = \int_{\partial M} i^*\tau$ where $\partial M \xrightarrow{i} M$.If M oriented, is there a natural orientation on ∂M?

11.2 Positively oriented manifold boundary

Let M be a C^∞ manifold of dim m. For any $\varphi \in \Phi_M$ which maps $p \in U_\varphi$ to $\varphi(p) = (x^1(p), ..., x^m(p)) \in (-\infty, 0] \times R^{m-1}$ ($p \in U_\varphi \cap \partial M \Leftrightarrow x^1(p) = 0$), we let $U_\varphi \cap \partial M \xrightarrow{\varphi^{\partial M}} (x^2(p), ..., x^m(p))$, which gives a topological chart of ∂M on $U_\varphi \cap \partial M$. $Phi^{\partial M} = (\varphi^{\partial M}\varphi \in \Phi_M)$ is a C^∞ atlas of ∂M which induces the unique C^∞ structure on ∂M so that $\partial M \xrightarrow{i} M$ is C^∞. And $\Phi^{\partial M}$ is a maximal c^∞ atlas. Now suppose that M is oriented and Φ_M a maximal compatible C^∞ atlas of M which determines the orientation of M. Then $\widetilde{\Phi^{\partial M}} = (\varphi^{\partial M} | \varphi \in \widetilde{\varphi_M})$ is also a compatible smooth atlas on ∂M, and hence determines an orientation on ∂M, which is called the positive orientation of ∂M induced by the orientation of M. Unless otherwise mentioned, we will always use ∂M to denote the positively oriented boundary.

Remark. *M: a C^∞ manifold of dim m. Let $A^K_\Phi(M) = ((\omega_\varphi | \varphi \in \Phi|$ $\omega_\varphi \in A^k(\varphi(U_\varphi))$ for all $\varphi \in \Phi$ so that * holds for every pair of charts) where $\Phi \in \Phi_M$ is a C^∞ atlas of M. If $\Phi_1 \subseteq \phi_2 ... \subseteq \Phi_M \Rightarrow A^k_{\Phi_2}(M) \xrightarrow{T^{\Phi_2}_{\Phi_1}} A^k_{\Phi_1}(M)$ where $T^{\Phi_2}_{\Phi_1}$ is a bijection.*

11.3 Stokes' theorem

If M is an oriented C^∞ manifold and $\omega \in A^{dimM-1}_c(M)$, then $\int_M d\omega = \int_{\partial M} \omega = \int_{\partial M} i^*\omega$

Proof. Choose an arbitrary C^∞ partition of unity $\rho_j (j \in J)$ subordinate to $(U_\varphi | \varphi \in \widetilde{\Phi_M})$ so that $supp\rho_j$ are all compact, say $supp\rho_j \subseteq U_{\varphi_j} = U_j$ for some $\varphi_j \in \widetilde{\Phi_M}$. Since $supp\omega$ is compact, $\omega = \sum_{j\in J} \rho_j\omega$ is essentially a finite sum, it suffice to show that if $\eta \in A^{dimM-1}_c(M)$ and $supp\eta \subseteq U_\varphi$ for some φ, then $\int_M d\eta = \int_{\partial M} \eta$.
Suppose that the coordinates induced by φ are $x^1, ..., x^m$ and the local expression of η is $\sum_{l=1}^m f_l dx^1 \wedge ... \wedge \overset{\wedge}{dx^l} ... \wedge dx^m$. f_l is a C^∞ function on half space with compact support. $d\eta$ has local expression on U_φ is $\sum_{l=1}^m \frac{\partial f_l}{\partial x^l} dx^l \wedge ... \wedge dx^m = (\sum_{l=1}^m (-1)^{l-1} \frac{\partial f_l}{\partial x^l}) dx^1 \wedge ... \wedge dx^m$. $\int_M d\eta = \int_{(-\infty,0]\times R^{m-1}} \sum_{l=1}^m (-1)^{l-1} (\frac{\partial f_l}{\partial x^l}) dx^1 \wedge ... \wedge dx^m$. There are two conditions.
For $l = 2, ..., m$, choose a suitable rectangle R and the integral is equal to 0 since $\int_R \frac{\partial f_l}{\partial x^l} dx^1 \wedge ... \wedge dx^m = \int_{R_l} (\int_{a_l}^{b_l} \frac{\partial f_l}{\partial x^l} dx^l) dx^1 \wedge ... \wedge \overset{\wedge}{dx^l} ... \wedge dx^m =$

$\int_{R_l}(f(x^1,...b_l,...,x^m)-f(x^1,...a_l,...,x^m))dx^1\wedge...\wedge \overset{\wedge}{dx^l}...\wedge dx^m=0.$
For $l=1$, the integral is equal to $\int_{(-\infty,0]\times R^{m-1}} frac\partial f_l\partial x^1 dx^1\wedge...\wedge dx^m = \int_{R_1}(\int_{a_1}^{b_1}\frac{\partial f_l}{\partial x^1}dx^1)\wedge...\wedge dx^m = \int_{R_1}((f(b_1,...,x^m)-f(a^1,...,x^m))dx^2\wedge,...,\wedge dx^m = \int_{R_1} f(b_1,...,x^m)dx^2\wedge,...,\wedge dx^m.$
Then integrate $i^*\eta$ on ∂M. $i^*\eta$ has local expression $f_l(0,x^2,...,x^m)dx^2\wedge,...,\wedge dx^m$.
Thus the proof is complete. □

12 Tangent vector fields

M: a c^k manifold of dim m, $V\in M$.

Definition. *A rule assigning to each point $p\in V$ with a tangent vector $X(p)\in T_pM$ is called a tangent vector field on V. More formally, we define the tangent bundle of M. $TM := \cup_{p\in M}T_pM$ and canonical projection. Then a tangent vector field X on V is exactly a map $V\subseteq M\xrightarrow{x} TM$ so that $\pi\circ X(p)=p$. X is called a section of $TM\xrightarrow{\pi} M$ over V.*

Definition. *Let X be a vector field on V. For any $\varphi\in\Phi_M$ defined on $U\subseteq_{open} M$ whose coordinates are $x^1,...,x^m$, and for any $p\in V\cap U$ since $(\frac{\partial}{\partial x^1})_p,...(\frac{\partial}{\partial x^m})_p$ forms a basis of T_pM. $x(p)$ can be written as $x^j(p)(\partial/\partial x^j)_p$ for a unique set of "components" $x^1(p),...$. $x^1,...,x^m$ are functions on $V\cap U$. $(\frac{\partial}{\partial x^1},...,\frac{\partial}{\partial x^m})$ is called the frame of TM on U induced by ϕ.*

Definition. *A vector field X on V is C^k near a point $p\in V$ if $\exists\varphi\in\Phi_M$ defined near p so that the components of X induced by φ are C^k functions, i.e. $x^1\circ\varphi^{-1},...,x^m\circ\varphi^{-1}$ are C^k functions on $\varphi(V\cap U)$. If X is C^k at or near every point of V, we can call X a C^k vector field.*

12.1 The quotient/gluing viewpoint

Idea. *A topological manifold X with a topological atlas Φ can be reconstructed by gluing $V_\alpha := \varphi_\alpha(U_\alpha)(\alpha\in A)$ along $V_{\alpha\beta} := \varphi_\alpha(U_\alpha\cap U_\beta)(\alpha\beta\in A)$ via $v_{\alpha\beta}\xleftarrow{\varphi_{\alpha\beta}} V_{\beta\alpha}$. Note that the data $(V_\alpha, v_{\alpha\beta}\xleftarrow{\varphi_{\alpha\beta}} V_{\beta\alpha})$ satisfy the following conditions:$\forall\alpha,\beta,\gamma\in A$, $V_{\alpha\alpha}=V_\alpha$,$V_{\alpha\beta}\subseteq_{open} V_\alpha, v_{\alpha\beta}\xleftarrow{\varphi_{\alpha\beta}} V_{\beta\alpha}$ is homeomorphic, $V_{\alpha\beta}\cap V_{\alpha\gamma}=\varphi_{\alpha\beta}(V_{\beta\alpha}\cap V_{\beta\gamma})$ and informally $\varphi_{\alpha\beta}\circ\varphi_{\beta\gamma}=\varphi_{\alpha\gamma}$.*

13 Multilinear algebra-tensors

V: usually finitely dimensional vector space over K. $V^* := (v\xrightarrow{f} K|f\quad K-linear)$. If $e_1,...,e_n$ form a basis of V, then we have the dual basis $e^1,...,e^n$ of V^* where $e^j(e_k)=\delta^j_k(Kronecker\quad\delta)=1(j=k)\quad or\quad =0(j\neq k)$.

13.1 Multilinear algebra

U, V, W: vector spaces over K. $U^* \otimes V^* \otimes W^* := (U \times V \times W \xrightarrow{f} K | f$ is multilinear). Does $U^* \otimes V^* \otimes W^*$ have a specific basis induced by these bases?

Definition. *$f \in U^*, g \in V^*, h \in W^*$, $U \times V \times W \overset{f\otimes g\otimes h}{\longrightarrow} K$.*
$\begin{pmatrix} \widetilde{u_r} \\ \widetilde{v_s} \\ \widetilde{w_t} \end{pmatrix}$ *another bases of* $\begin{pmatrix} U \\ V \\ W \end{pmatrix}$. $u^i \otimes v^j \otimes w^k = a_r^i b_s^r c_t^k \ \widetilde{u_r} \otimes \widetilde{v_s} \otimes \widetilde{w_t}$ *if*

$$\begin{cases} \widetilde{u_r} = a_r^i u^i \\ \widetilde{v_s} = b_s^r v_s \\ \widetilde{w_t} = c_t^k w_t \end{cases}$$

(by Einstein convention)

Definition. *V: vector space over k. $\wedge^k V^* := (V \times ... \times V \xrightarrow{f} K | f \quad k-linear$ and alternating). Ex: determinants. $V^* \otimes ... \otimes V^*$ can be denoted by $\otimes^k V^*$. An alternating mapping from $\otimes^k V^*$ to itself is an isomorphism.*
$\otimes^k V^* \xrightarrow{Alt} \wedge^k V^*$. It is easily to find that for any $f \in \otimes^k V^*$, $f \in \wedge^k V^* \Leftrightarrow Alt(f) = f$.
$\widetilde{Alt}$ is different from Alt by multiplying k!. $\widetilde{Alt}\ (\varphi_1 \otimes ... \otimes \varphi_k)(V_1, ..., V_k) = det \begin{pmatrix} \varphi_1(V_1) & ... & \varphi_1(V_k) \\ ... & ... & ... \\ \varphi_k(V_1) & ... & \varphi_k(V_k) \end{pmatrix}$.
$\varphi \in \wedge^k V^*, \psi \in \wedge^l V^*, \varphi \wedge \psi \overset{def}{=} \frac{(k+l)!}{k!l!} \times Alt(\varphi \otimes \psi) = \frac{1}{k!l!} \times \widetilde{Alt}\ (\varphi \otimes \psi)$.

Claim. *$e^{i_1} \wedge ... \wedge e^{i_k}$ forms a basis of $\wedge^k V^*$.*

Proof. $f \in \wedge^k V^*$. $f = f_{i_1,...,i_k} e^{i_1} \otimes ... \otimes e^{i_k} = \sum_{1 \leqslant i_1 < i_2 < ... < i_k \leqslant n} \sum_{\sigma \in S_k} f_{i_{\sigma 1},...,i_{\sigma k}} e^{i_{\sigma 1}} \otimes ... \otimes e^{i_{\sigma k}} = \sum_{1 \leqslant i_1 < i_2 < ... < i_k \leqslant n} f_{i_{\sigma 1},...,i_{\sigma k}} \sum_{\sigma \in S_k} (-1)^\sigma e^{i_{\sigma 1}} \otimes ... \otimes e^{i_{\sigma k}} = \sum_{1 \leqslant i_1 < i_2 < ... < i_k \leqslant n} f_{i_{\sigma 1},...,i_{\sigma k}}\ \widetilde{Alt}\ e^{i_1} \otimes ... \otimes e^{i_k}$ □

13.2 Tensor fields and differential forms

M is a C^∞ manifold of dim m. $TM = \cup_{p \in M} T_p M$. $T^*M = \cup_{p \in M} T_p^* M$. (cotangent bundle of m) $\otimes^k T^*M = \cup_{p \in M} \otimes^k T_p^* M$. (tensor bundle of m) $\wedge^k T^*M = \cup_{p \in M} \wedge^k T_p^* M$.

Definition. *A tensor field S on $V \subseteq_{open} M$ is a map $V \xrightarrow{S} \otimes^k T^*M$. Furthermore, S is called a k-tensor of M on V. S is called a differential k-form if $S_p \in \wedge^k T_p^* M$ for all $p \in M$.*

Definition. $(\frac{\partial}{\partial x^j})_p \overset{dualbasis}{\longrightarrow} (dx^j)_p$

Definition. *vector field X:* $X(p) = X^j(\frac{\partial}{\partial x^j})_p \in T_pM$
k-tensor S: $S(p) = S_{j_1,...,j_k}(p)(dx^{j_1})_p \otimes ... \otimes (dx^{j_k})_p$
k-form S: $S(p) = \sum_{j_1<...<j_k} S_{j_1,...,j_k}(p)(dx^{j_1})_p \wedge ... \wedge (dx^{j_k})_p$. *We say that S is* C^∞ *if all* $S_{j_1,...,j_k}$ *are smooth functions on* $V \cap U$.

13.3 Cartan's exterior differentiation

M: C^∞ manifold of dim m. $A^k(M) = (M \xrightarrow{\omega} \wedge^k T^*M | \omega(p) \in \wedge^k T_p^* M)$.

Definition. $A^k(M) \xrightarrow{d} A^{k+1}(M)$. $d\omega \in \wedge^{k+1}T^*M$. *For any* $U_\alpha \xrightarrow{\varphi_\alpha} \varphi_\alpha(U_\alpha) = V_\alpha \subseteq R^m \in \Phi_M$, *we can write* $\omega = \omega_{j_1,...,j_k} dx^{j_1} \wedge ... \wedge dx^{j_k}$. *We define* $d\omega$ *on* $U.p \in U_\alpha \Longrightarrow (d\omega)(p) = \frac{\partial(\omega_{j_1,...,j_k} \circ \varphi_\alpha^{-1})}{\partial x_\alpha^{j_0}}(\varphi_\alpha(p))(dx_\alpha^{j_0})_p \wedge (dx_\alpha^{j_1})_p \wedge ... \wedge (dx_\alpha^{j_k})_p$.

Beyond the definition above, we should ensure the $d\omega$ is still the same mapping through different charts. Suppose that on $U \cap V$ (V is another chart), $\eta = \eta_{l_1,...,l_k} dy^{l_1} \wedge ... \wedge dy^{l_k}$. By coordinate transformation,
$\eta_{l_1,...,l_k}(\frac{\partial y^{l_1}}{\partial x^{j_1}})(\frac{\partial y^{l_1}}{\partial x^{j_1}})...(\frac{\partial y^{l_k}}{\partial x^{j_k}}) = \omega_{j_1,...,j_k} \Longrightarrow$
$\frac{\partial \omega_{j_1,...,j_k}}{\partial x^{j_0}} = \frac{\partial \eta_{l_1,...,l_k}}{\partial y^{l_0}}(\frac{\partial y^{l_0}}{\partial x^{j_0}})(\frac{\partial y^{l_1}}{\partial x^{j_1}})...(\frac{\partial y^{l_k}}{\partial x^{j_k}}) + \eta_{l_1,...,l_k} \sum_{s=1}^k (\frac{\partial y^{l_1}}{\partial x^{j_1}})(\frac{\partial^2 y^{l_s}}{\partial x^{j_s} \partial x^{j_0}})...(\frac{\partial y^{l_k}}{\partial x^{j_k}})$
$\Longrightarrow \frac{\partial \omega_{j_1,...,j_k}}{\partial x^{j_0}} dx^{j_0} \wedge ... \wedge dx^{j_k} = (\frac{\partial \eta_{l_1,...,l_k}}{\partial y^{l_0}}(\frac{\partial y^{l_0}}{\partial x^{j_0}})(\frac{\partial y^{l_1}}{\partial x^{j_1}})...(\frac{\partial y^{l_k}}{\partial x^{j_k}}) + \eta_{l_1,...,l_k}$
$\sum_{s=1}^k (\frac{\partial y^{l_1}}{\partial x^{j_1}})(\frac{\partial^2 y^{l_s}}{\partial x^{j_s} \partial x^{j_0}})...(\frac{\partial y^{l_k}}{\partial x^{j_k}}))dx^{j_0} \wedge ... \wedge dx^{j_k} = \frac{\partial \eta_{l_1,...,l_k}}{\partial y^{l_0}} dy^{l_0} \wedge ... \wedge dy^{l_k}$. (the latter term is equal to zero)

14 Homology

Let M be a C^∞ manifold of dim m. We call $\longrightarrow A^{-1}(M) \longrightarrow A^0(M) \xrightarrow{d^0} A^1(M) ... \longrightarrow A^j(M) \xrightarrow{d^0} A^j(M) \longrightarrow ... \longrightarrow A^m(M)$ the deRham complex of M and let $H^j(M, C)$ =closed n-form on M/exact n-form on M, called the j-th deRham cohomology of M which is a C-vector space.

Terminology. *For* $\omega \in A^j(M)$, ω *is closed*$\Leftrightarrow$ $d\omega = 0(Z^j)$ *and is exact*$\Leftrightarrow$ $\exists \eta \in A^{j-1}(M), \omega = d\eta(B^j)$.

Example. $H^0(M, C) \simeq ker(A^0(M) \xrightarrow{d} A^1(M))$. *That is,* $f \in H^0(M, C)$ *is locally constant. Therefore, if we let* $\pi_0(M)$ *be the path connected components of M (each of which is open in M), then* $H^0(M, C) \xleftarrow{\sim} C^{x\pi_0(M)}$, *which is the canonical map.*

Example. *Let* $M = R^2/\{p\}$ *where* $p = (a, b) \in R^2$. *Let* $\omega_p = \frac{(y-b)dx-(x-a)dy}{(x-a)^2+(y-b)^2} \in Z^1(M) = ker(A^1(M) \xrightarrow{d} A^2(M))$. *For any closed* $\eta \in Z^1(M)$, *let* $c = \int_\gamma \eta$. *gamma is the path surrounding p. So what is* $H^1(M, C)$? $\eta - \frac{c}{2\pi}\omega_p \in B^1(M)$. $H^1(M, C) \simeq C$ *with basis* ω_p. *In addition,* $M = R^2/\{p_1, p_2\}(p_1 \neq p_2)$. *It is easy to find that* $H^1(M, C) \simeq C \oplus C$ *with two linear-independent bases* $\omega_{p_1}, \omega_{p_2}$.

Definition. *A sequence of homomorphisms of groups* $... \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow ...$ *is an exact at* B *if* $ker(g) = im(f)$. *It is called an exact sequence at every position.*

Theorem. *Given chain maps* $(A_*) \xrightarrow{(S_*)} (B_*) \xrightarrow{(T_*)}$. *If* $0 \longrightarrow (A_{j+1}) \xrightarrow{(S_{j+1})} (B_{j+1}) \xrightarrow{(T_{j+1})} 0$ *for every* $j \in Z$, *then* $\exists$ *homomorphisms* $H_j(C_j) \xrightarrow{\delta_j} H_{j-1}(A_{j-1})$ *forming an exact long chain.*

Remark. $0 \longrightarrow A \xrightarrow{f} B$ *is exact,* f *is injective.* $B \xrightarrow{f} C \longrightarrow 0$ *is exact,* f *is surjective.*

15 The deRham cohomologies of C∞ manifolds

Definition. *deRham cohomologies* $H^k(M, C) = ker(a^k(M) \xrightarrow{d} A^{k+1}(M))/im(a^{k-1}(M) \xrightarrow{d} A^k(M))$. *The elements of* $H^k(M, C)$ *are of the form* $\omega + dA^{k-1}(M)$ *with* ω *a closed k-form.l*

15.1 The cup product on cohomologies

For any $k, l \in Z$, we define the cup product map $H^k(M, C) \times H^l(M, C) \xrightarrow{\cup}= H^{k+l}(M, C), ([\omega], [\eta]) \longmapsto [\omega] \cup [\eta] = [\omega \wedge \eta]$. Recall the super-Leibniz rule: $\forall \alpha \in A^a(M)$ and $\beta \in A^b(M) \Longrightarrow d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^{\alpha} \alpha \wedge (d\beta) = 0$. Therefore $[\omega \wedge \eta]$ is defined. If $[\omega_1] = [\omega_2]$, then $\exists \tau \in A^{k-1}(M)$ s.t. $\omega_1 = \omega_2 + d\tau$ and hence $\omega_1 \wedge \eta - \omega_2 \wedge \eta = d\tau \wedge \eta = d(\tau \wedge \eta) \pm \tau \wedge (d\eta) = 0 \Longrightarrow [\omega_1 \wedge \eta] = [\omega_2 \wedge \eta]$. $\cup$ is a C-bilinear map.

Definition. $H^*(M, C) := \oplus_{k \in Z} H^k(M, C)$

Proposition 15. $(H^*(M, C), +, \cup)$ *is a supercommutative Z-gradient C-algebra.*

15.2 Pulling-back cohomology classes

Every C^∞ map $M \xrightarrow{f} N$ induces a cochain map between deRham complexes. Particularly, f induces C-linear maps $H^k(N, C) \xrightarrow{f^*} H^k(M, C)$. (well defined by basic algebra) Moreover,$H^*(N, C) \xrightarrow{f^*} H^*(M, C)$ is a homomorphism of C-linear algebra: $f^*([\eta] \cup [\eta']) = f^*[\eta \wedge \eta'] = [f^*(\eta \wedge \eta')] = [(f^*\eta) \wedge (f^*\eta')]$. If there is a chain $M \xrightarrow{f} N \xrightarrow{g} P \Rightarrow (g \circ f)^* = f^* \circ g^*$. And there exists a identity map between $H^*(M, C)$ and itself. So there exists a contravariant functor).

16 The long exact sequence of cohomologies induced by a short exact sequence of cochain complexes

Given a short sequence of chain complexes with all squares commutative all columns complexes and rows exact. We have the induced maps (homomorphisms) $H_k(A) \xrightarrow{f_{*k}} H_k(B) \overset{g_{*k}}{H}_k (C)$.. We can construct natural (functorial) homomorphisms $H_K(c) \overset{\partial_{*k}^{(f,g)}}{longrightarrow} H_{k-1}(A)$. and the induced long sequence which is an exact sequence. $\partial_{*k}(k \in Z)$ so constructed are called the connecting homomorphisms. (zig-zag)

Proof. (well-defined) Begin from choose any z, z', where $[z] = [z']$. It is straightforward to write $z - z' = dc(c \in C^{k-1})$. Since the mapping between B and C is onto.$\exists \tilde{b}$. $b - b' - d\tilde{b} \mapsto 0$. (where b, b' maps to c, c' by g. Next,$\exists \tilde{a} \mapsto b - b' - d\tilde{b}$. $d\tilde{a} \mapsto db - db' \mapsto 0$.Therefore, $a - a' - d\tilde{a}$. Hence, $a = a' + d\tilde{a} \Longrightarrow [a] = [a']$.

(exactness) Three conditions. For $H^k(A) \xrightarrow{f^{*k}} H^k(B) \overset{g^{*k^k}}{H} (C)$, $[z] \mapsto [f(z)] \mapsto [g(f(z))] = 0 \Longrightarrow imf^{*k} \subseteq kerg^{*k}$. For $H^k(B) \xrightarrow{g^{*k}} H^k(C) \xrightarrow{\delta^k} H^{k+1}(A)$, first $[\omega] \mapsto [g(\omega)]$, and $d\omega = 0$, and hence $\delta^k[g(\omega)] = [0] = 0 \Longrightarrow img^{*k} \subseteq ker\delta^k$. And for $H^k(C) \xrightarrow{\delta^k} H^{k+1}(A) \xrightarrow{f^{*k+1}} H^{k+1}(B)$, $[\omega] \in H^k(C) \mapsto [a] \mapsto [f(a)] = [db] = 0 \Longrightarrow im\delta^k \subseteq kerf^{*k+1}$. And proof for the opposite direction is similar. $\square$

17 Cohomologies of the simplest class of space

Definition. *Given an R-vec. space V, a subset $S \subseteq V$ is a star-shaped set if $\exists p \in V$, s.t. $\forall s \in S$ and $t \in [0,1][(1-t)p + ts \in S]$, such a point p is called a center of S. In most occassions, we may assume 0 is a center of S by translation.*

Now consider a star-shaped open $U \subseteq R^m$(center at 0). Then $H^0(U, C) \simeq C$ by the path-connectedness of U. So what is $H^n(U, C)$? Given a closed $\omega \in A^l(U)$ with $l \geqslant 1$, can one obtain an $\eta \in A^{l-1}(U)$ s.t. $\omega = d\eta$ by integration?

Notation. *$U' := ((x,t) \in R^m \times R | tx \in U)$. There is a map $U' \xrightarrow{H} U$. We consider a slightly more general setting. Given any set U and a subset $V \subseteq U \times R$, we say that all vertical slices of V are open intervals containing $0 \in R$ if $\forall x \in U, \exists -\infty \leqslant a_x < 0 < b_x \leqslant \infty$ s.t. $V \cap (x \times R) = X \times (a_X, b_x)$. When talking about such a V, we adopt the following definitions:*
for any $t \in R$, $V_t := (x \in U | (x,t) \in V)$, called the horizontal slice of V of height t; $V_t \xrightarrow{l_t} V$
$V \xrightarrow{\pi} U$
for any $W \subseteq U$, $W_0 := \pi^{-1}(W)$

Now consider the case $U \subseteq_{open} R^m$ and $V \subseteq_{open} U \times R$. What if we integrate a C^∞ differential form along t?
Every $\phi \in A^l(V)$ can be uniquely written in the form $\sum_{1 \leqslant j_1 < \ldots < j_l \leqslant m} A_{j_1,..,j_l}(x,t)$ $dx^{j_1} \wedge \ldots \wedge dx^{j_l} + \sum_{1 \leqslant k_1 < \ldots < k_{l-1} \leqslant m} B_{k_1,\ldots,k_{l-1}}(x,t) dt \wedge dx^{k_1} \wedge \ldots \wedge dx^{k_{l-1}}$. We define $A^l(V) \xrightarrow{I} A^{l-1}(V)$, $phi = \sum_{|J|=l} A_J(x,t) dx^J + \sum_{|k|=l-1} B_K(x,t) dt \wedge dx^K \mapsto$ $\sum_{|K|=l-1} (\int_0^t B_K(x,s) ds) dx^k$.
Then $dI\phi = \sum_{|K|=l-1} [(\int_0^t \frac{\partial B_K}{\partial x^k}(x,s) ds) dx^k \wedge dx^K + B_K(x,t) dt \wedge dx^K]$. Since $d\phi = \sum_{|J|=l} (\frac{\partial A_J}{\partial x^j})(x,t) dx^j \wedge dx^J + \frac{\partial A_J}{\partial t}(x,t) dt \wedge dx^J) + \sum |K| = l-1 \frac{\partial B_k}{\partial x^k}(x,t) dx^k \wedge$ $dt \wedge dx^K$.
$Id\phi = \sum_{|J|=l} (\int_0^t \frac{\partial A_J}{\partial s}(x,s) ds) dx^J - \sum_{|K|=l-1} (\int_0^t (\frac{\partial B_k}{\partial x^k}(x,s) ds) dx^k \wedge dx^K =$ $\sum_{|J|=l} A_J dx^J - \sum_{|J|=l} A_J(x,0) dx^J - \sum_{|K|=l-1} [(\int_0^t \frac{\partial B_K}{\partial x^k}(x,s) ds) dx^k \wedge dx^K]$.
Thus $dI\phi + Id\phi = \phi - (l_0 \circ \pi)^* \phi$ for all $\phi \in A^l(V)$.

Corollary. *(the Poincare lemma) $H^l(U,C) = C$(if l=0) or 0 if $l \neq 0$ if U is a star-shaped open subset of R^m.*

Proof. For any $[\omega] \in H^l(U,C)$ with $l \geqslant 1$, we have $dIH^*\omega + IdH^*\omega = H^*\omega - \pi^* l_0^* H^*\omega$ with $IdH^*\omega = IH^*d\omega = 0$ and $H \circ l_0(x) = H(x,0) = 0$. Therefore, $dIH^*\omega = H^*\omega$. Applying l_1^*, we have $dl_1^* IH^*\omega = l_1^* dIH^*\omega = l_1^* H^*\omega = \omega \Rightarrow$ $[\omega] = 0$. □

We may globalize the construction of I. Let M be a C^∞ manifold of dim m and $V \subseteq M \times R$ an open subset whose vertical slices are open intervals containing $0 \in R$. V has a natural C^∞ structure by the atlas Phi_0 which consists of the charts $U_{\varphi_0} := \xrightarrow{\varphi_0} \varphi_0(U_{\varphi_0} = \cup_{x \in U\varphi}(\varphi(x)0 \times (a_x, b_x)$.

Definition. *We define $A^l(V) \xrightarrow{I_V} A^{l-1}(V)$ to be unique map which is commute.*

18 Homotopy

Definition. *We say that two C^∞ maps f_1, f_0 from M to N are C^∞ homotopic to each other if $\exists C^\infty$ map $M \times I \xrightarrow{H} N$ where I is an open interval containing* $[0,1]$ *s.t.* $f_j = H \circ l_j$

Notation. *$f_1 \sim f_0$ denotes f_1, f_0 are homotopic.*

Corollary. *(homotopy invariance of cohomology maps) Given C^∞ maps f_1, f_0 S.T. $f_1 \sim f_0$, then $H^l(M,C) \xleftarrow{f_0^* = f_1*} H^l(M,C)$.*

Proof. For any $[\omega] \in H^l(N,C)$, we have $f_1^*\omega = (H \circ l_1)^*\omega = l_1^* H^*\omega = l_1)^*(dIh^*\omega +$ $IdH^*\omega + \pi^* l_0^* H^*\omega) = d(l_1^* IH^*\omega) + l_1^* \pi^* l_0^* H^*\omega = d(l_1^* IH^*\omega) + (id_M)^* f_0^*\omega$. □

Definition. *(homotopy equivalence) A C^∞ map $M \xrightarrow{f} N$ is a C^∞ homotopy equivalence if $\exists C^{infty}$ map $M \xleftarrow{g} N$ s.t. $g \circ f \sim id_M$ and $f \circ G \sim id_N$. Such* *a map g is called a homotopy invariance of f. Given C^∞ manifolds M and N, we say that they have the same homotopy type if $\exists C^\infty$ homotopy equivalence $M \xrightarrow{f} N$.*

Corollary. *Given C^∞ manifods M and N, if $M \xrightarrow{f} N$, then $H^l(M,C) \overset{f^*}{\simeq} H^l(N,C)$.*

Definition. *Given a C^∞ manifold M, a subset $A \subseteq M$ which itself has a smooth structure s.t. the inclusion map $A \xrightarrow{i} M$ is C^∞ and a C^∞ map $M \xrightarrow{r} A$, r is called a retraction if $r(p) = p$ when $p \in A$; r is called a deformation retraction if $r \circ i = id_A$ and $i\circ \sim id_M$.*

On cohomologies, if r is a C^∞ retraction$\Longrightarrow$ $H^l(A,C) \xrightarrow{r^*} H^l(M,C)$ a injection and $H^l(A,C) \xleftarrow{i^*} H^l(M,C)$ a surjection.
What does the condition $I \circ r \sim id_M$ mean under the assumption $r \circ i = id_A$? $\exists C^\infty$ map $M \times I \xrightarrow{H} M$ s.t. $\forall p \in M, H(p,0) = p$ and $H(p,1) = i \circ r(p) = r(p).t \in I \mapsto h(p,t) \in M$ is a smooth path. In practice, to construct a smooth deformation retraction, we first create a smooth retraction $M \xrightarrow{r} A$ and a smoth path $I \xrightarrow{\gamma_p} M$ s.t. $\gamma_p(0) = p$ and $\gamma_p(1) = r(p)$ for every $p \in M$, and prove that the map $M \times I \xrightarrow{"H"} M$ is C^∞.

Example. *(contractibe spaces) M is contractible if $\exists p \in M$ and a defromation retraction $M \xrightarrow{r} (p)$. For example, all star-shaped are contractible.*

Theorem. *(Brouwer's fixed point theorem) If $\bar{B} \xrightarrow{f} \bar{B}$ is a continuous map ($\bar{B} := \bar{B_1}(0) \subseteq R^n$), then $\exists x \in \bar{B}$, f(x)=x.*

Proof. Case 1. f is C^∞. Suppose that $\forall x \in \bar{B}, f(x) \neq x$. We can yield a map $\bar{B} \xrightarrow{r} \partial\bar{B} = S$. Then r is a smooth retraction. Contradictory! □

Remark. *(Lefschiz's fixed point theorem) Let M be a compact oriented C^∞ manifold and $M \xrightarrow{f} M$ be a C^{infty} map. If $L(f) := \sum_{l=0}^{\infty}(-1)^l tr(H^l(M,C) \xleftarrow{f^*} H^l(M,C) \neq 0$, then $\exists x \in M, f(x) = x$.*

19 Stochastic process and calculus

19.1 The continuity of sample paths

Definition. *Let $(X_t)_{t\in T}$ and $(\tilde{X}_t)_{t\in T}$ be two random processes indexed by the same index set T and with values in the same metric space E, We say that $\tilde{X}$ is a modification of X if $\forall t \in T, P(\tilde{X}_t = X_t) = 1$.*

Definition. *The process $\tilde{X}$ is said to be indistinguishable from X if there exists a negligible subset N of Ω such that $\forall \omega \in \Omega \setminus N, \forall t \in T$, $\tilde{X}_t(\omega) = X_t(\omega)$.*

Lemma. *(Kolmogorov's lemma) Let $X = (X_t)_{t \in I}$ be a random process indexed by a bounded interval I of R, and taking values in a complete metric space (E,d). Assume that there exists three reals $q, \epsilon, C > 0$ s.t. for every $s, t \in I$, $E[d(X_s, X_t)^q \leqslant C|t-s|^{1+\epsilon}$. Then there is a modification $\tilde{X}$ of X whose sample paths are Hölder continuous with component α for every $\alpha \in (0, \frac{\epsilon}{q})$. This means that for every $\omega \in \Omega$ and every $\alpha \in (0, \frac{\epsilon}{q})$, there exists a finite constant $C_\alpha(\omega)$ such that for every $s, t \in I$, $d(\tilde{X}_s(\omega), \tilde{X}_t(\omega)) \leqslant C_\alpha(\omega)|t-s|^\alpha$.*

Proof. To simplify the presentation, we take $I = [0,1]$ and then fix $\alpha \in (0, \frac{\epsilon}{q})$. By Chebyshev inequality and the assumption of the lemma, for $\alpha > 0, s, t \in I$, $P(d(X_s, X_t) \geqslant A = a) \leqslant a^{-q} E[d(X_s, X_t)^q] \leqslant Ca^{-q}|t-s|^{1+\epsilon}$. We apply this inequality to $s = (i-1)2^{-n}, t = i2^{-n}$ for $i \in \{1, ..., 2^n\}$ and $\alpha = 2^{-n\alpha}$: $P(d(X_{(i-1)2^{-n}}, X_{i2^{-n}}) \geqslant 2^{-n\alpha}) \leqslant C2^{nq\alpha}2^{-(1+\epsilon)n}$. By summing over i, $P(\cup_{i=1}^{2^n}\{d(X_{(i-1)2^{-n}}, X_{i2^{-n}}) \geqslant 2^{-n\alpha}\} \leqslant 2^n C 2^{nq\alpha-(1+\epsilon)n} = C2^{-n(\epsilon-q\alpha}$. By assumption, $\epsilon - q\alpha > 0$, summing over n, we obtain $\sum_{n=1}^\infty P(\cup_{i=1}^{2^n}\{d(X_{(i-1)2^{-n}}, X_{i2^{-n}}) \geqslant 2^{-n\alpha}\}) < \infty$, and by Borel-Cantelli lemma, with probability 1, we can find a finite integer $n_0(\omega)$ s.t. $\forall n \geqslant n_0(\omega), \forall i \in \{1, ..., 2^n\}$, $d(X_{(i-1)2^{-n}}, X_{i2^{-n}}) \geqslant 2^{-n\alpha}$. Then let s, t satisfy $0 < t - s < 2^{n_0(\omega)}$. Hence there exists $n \geqslant n_0(\omega)$ such that $2^{-(n+1)} \leqslant t - s < 2^{-n}$. Next, we claim that there is a constant $K_\alpha(\omega)$, such that $d(X_t(\omega), X_s(\omega)) \leqslant K_\alpha(\omega)|t-s|^\alpha, \forall s, t \in D, 0 < s - t < 2^{-n_0(\omega)}$. For the moment, we restrict to the set of $s, t \in \cup_{m \geqslant n+1} D_m$, with $0 < t - s < 2^{-n}$. By induction to $m \geqslant n+1$ we will first show that $d(X_t(\omega), X_s(\omega)) \leqslant 2\sum_{k=n+1}^m 2^{-\alpha k}$ if $s, t \in D_m$. Suppose that $s, t \in D_{n+1}$, then $t - s = 2^{-(n+1)}$. Therefore, $\exists k \in \{0, ..., 2^{(n+1)} - 1\}$, s.t. $t = \frac{k}{2^{n+1}}$ and $s = \frac{k+1}{2^{n+1}}$. Assume that the claim hols for some $m \geqslant n+1$. Put $s' = \min\{u \in D_m | u \geqslant s\}$ and $t' = \max\{u \in D_m | u \leqslant t\}$. By construction and the assumption, $s \leqslant s' \leqslant t' \leqslant t$ and $s' - s, t - t' \leqslant 2^{-(m+1)}$. $d(X_t(\omega), X_s(\omega)) \leqslant d(X_t(\omega), X_{t'}(\omega)) + d(X_{t'}(\omega), X_{s'}(\omega)) + d(X_{s'}(\omega), X_s(\omega)) \leqslant 2^{-\alpha(m+1)} + 2\sum_{k=n+1}^m 2^{-\alpha k} + 2^{-\alpha(m+1)} = 2\sum_{k=n+1}^{m+1} 2^{-\alpha k}$.
Now let $s, t \in D$ with $0 < t - s < 2^{-n_0(\omega)}$. As noted before, there exists $n \geqslant n_0(\omega)$ s.t. $2^{-(n+1)} \leqslant t - s < 2^{-n}$. Then there exists $m \leqslant n+1$ such that $t, s \in D_m$. Apply the previous result, we construct $K_\alpha(\omega) = \frac{2}{1-2^{-\alpha}}$. And now fix ω the mapping $t \to X_t(\omega)$ is Hölder continuous on D and hence uniformly continuous on D. Since (E, d) is complete, the mapping has a unique continuous extension. □

Corollary. *Let $B = (B_t)_{t \geqslant 0}$ be a pre-Brownian motion. The process B has a modification whose sample paths are continuous, and even locally Hölder continuous with exponent $\frac{1}{2} - \delta$ for every $\delta \in (0, \frac{1}{2})$.*

Proof. If $s < t$, the random variable $B_t - B_s$ is distributed as $N(0, t-s)$. For every $q > 0$, $E|B_t - B_s|^q = (t-s)^{\frac{q}{2}} E|U|^q$ where U $N(0,1)$. Taking $q > 2$, we apply the lemma and $\epsilon = \frac{q}{2} - 1$. It follows that B has a modification whose sample paths are locally Höolder continuous with exponent α for every $\alpha < \frac{q-2}{2q}$. □

19.2 Filtrations and Martingales

Definition. *A process X_t with values in a measurable space (E, ϵ) is said to be measurable if the mapping $(\omega, t) \to X_t(\omega)$ is measurable to $F \otimes B(R^+)$ adapted if for every $t \geqslant 0$, X_t is F_t-measurable.*
progressive if for every $t \geqslant 0$, X_t is $F_t \otimes B([0,t])$-measurable.

Proposition 16. *X_t is adapted and the sample paths are right or left continuous. Then X_t is progressive measurable.*

Proof. It suffices to show that it is the case for right continuity. Fix t¿0. For every $n \geqslant 1$ and $s \in [0,t]$, define a random variable X_s^n by setting $X_s^n = X_{\frac{kt}{n}}$ if $s \in [\frac{(k-1)t}{n}, \frac{kt}{n}), k \in \{1, ..., n\}$ and $X_t^n = X_t$. The right continuity of sample paths ensures $X_s(\omega) = \lim_{n \to} X_s^n(\omega)$. On the other hand, for every Borel subset A of E, $\{(\omega, s) \in \Omega \times [0,t] : X_s^n(\omega) \in A\} = (\{X_t \in A\} \times \{t\}) \cup (\cup_{k=1}^n(\{X_{\frac{kt}{n}} \in A\} \times [\frac{(k-1)t}{n}, \frac{kt}{n}))) \in F_t \otimes B([0,t])$. Hence, for every $n \geqslant 1$, the mapping $(\omega, s) \to X_s^n(\omega)$ is measurable for $F_t \otimes B([0,t])$. Since a pointwise limit of measurable functions is also measurable. Thus X is progressive. □

(Upcrossings, discrete version) The number $U_N[a,b](\omega)$ of upcrossings $[a,b]$ made by $n \to X_n(\omega)$ by time N is defined to be the largest k in Z^+ such that we can find $0 \leqslant s_1 < t_1 < s_2 < t_2 < ... < s_k < t_k \leqslant N$ with $X_{s_i}(\omega) < a, X_{t_i}(\omega) > b$. $C_n = I_{\{C_{n-1}=1\}} I_{\{X_{n-1} leqslant b\}} + I_{\{C_{n-1}=0\}} I_{\{X_{n-1}<a\}}$. Therefore, C_n is bounded, nonnegative, and previsible. We then have the following inequality: $Y_n(\omega) = C_n \dot{X}_n \leqslant (b-a)U_N[a,b](\omega) - [X_N(\omega) - a]^-$.

Theorem. *(Doob's upcrossings lemma, discrete version) Let X be a supermartingale. Let $U_N[a,b]$ be the number of upcrossings of $[a,b]$ by time N. Then, $(b-a)EU_N[a,b] \leqslant E[(X_N - a)^-]$.*

Proof. The process C is previsible, bounded and nonnegative, and $Y = C\dot{X}$. Hence Y is a supermartingale, and $E(Y_N) \leqslant 0$. □

Theorem. *(Martingale convergence theorem, discrete version) Let X be a supermartingale bounded in L^1. Then, a.s., $X_\infty = \lim X_n$ exists and is finite. For definiteness, we define $X^\infty(\omega) = \lim supX_n(\omega)$ s.t. X_∞ is F^∞ measurable and $X_\infty = \lim X_n$ a.s.*

Proof. Let $A = \{\omega : X_n(\omega)$ does not converge to a limit in $[-\infty, \infty]\} = \{\omega : \lim inf X_n(\omega) < \lim sup X_n(\omega)\} = \cup\{\omega : \lim inf X_n(\omega) < a < b < \lim sup X_n(\omega)\} \subseteq \{\omega : U_\infty[a,b](\omega) = \infty\}$ since $(b-a)EU_N[a,b] \leqslant |a| + E|X_N| \leqslant |a| + sup_n E|X_N|$ and MON can be applied. Thus, $P(A) = 0$. And since A is a countable union of $\{\omega : \lim inf X_n(\omega) < a < b < \lim sup X_n(\omega)\}$, $X_\infty = \lim X_n$ a.s.
But by Fatou's lemma, $E|X_\infty| = E(\lim inf|X_n|) \leqslant \lim inf E(|X_n|) \leqslant sup E(|X_n|) < \infty$. □

Theorem. *If M is a martingale and p¿1, then for all $n \in N$, $E(\max_{k \leqslant n} |M_k|^p) \leqslant (\frac{p}{1-p})^p E|M_n|^p$ provided that M is in L^p.*

Proof. Define $M^* = \max_{k \leqslant n} |M_k|$. We have for any $m \in NN$. $E(M^* \wedge m)^p = \int_\omega (M^*(\omega) \wedge m)^p dP(\omega) = \int_\omega \int_0^{M^* \wedge m} px^{p-1} dx dP(\omega) = \int_\omega \int_0^m px^{p-1} 1_{\{M^*(\omega) \geqslant x\}} ds dP(\omega) = \int_0^m px^{p-1} P\{M^* \geqslant x\} dx$. By maximal inequality, $P\{M^* \geqslant x\} \leqslant \frac{E(|M_n| 1_{\{M^* \geqslant x\}})}{x}$.
Then $E(M^* \wedge m)^p \leqslant \int_0^m px^{p-2} \frac{E(|M_n| 1_{\{M^* \geqslant x\}})}{x} dx = \int_0^m px^{p-2} \int_{\omega: M^* \geqslant x} |M_n(\omega)| dP(\omega) dx = p \int_\omega |M_n(\omega)| \int_0^{M^*(\omega) \wedge m} x^{p-2} dx dP(\omega) = \frac{p}{p-1} E(|M_n|(M^8 \wedge m)^{p-1})$. By Hölder's inequality, $E|M^* \wedge m|^p \leqslant \frac{p}{p-1} (E|M_n|^p)^{\frac{1}{p}} (E|M^* \wedge m|^p)^{\frac{p-1}{p}}$. Then $E|M^* \wedge m|^p \leqslant (\frac{p}{p-1})^p E|M_n|^p$. At the end, drive m to infinity. □

Theorem. *Let X_t be a supermartingale, and let D be a countable dense subset of R_+.*
(1) For almost every $\omega \in \Omega$, the restriction of the functions $s \to X_s(\omega)$ to the set D has a right-limit at every $t \in [0, \infty)$ and a left-limit at every $t \in (0, \infty)$.
(2) For every $t \in R_+, X_{t^+} \in L^1$ and $X_t \geqslant E[X_{t^+}|F_t]$, witj equality if the function $t \to E[X_t]$ is right-continuous. The process X_{t^+} is a supermartingale with respect to the filtration F_{t^+}. It is a martingale if X is a martingale.

Proof. (1) Fix $T \in D$. By the maximal inequality, $sup_{s \in S \cap [0,T]} |X_s| < \infty$ a.s. We then choose a swquence $(D_m)_{m \geqslant 1}$ of finite subsets of D that increase to $D \cap [0,T]$ and are such that $0, T \in D_m$. Upcrossing inequality then can be applied. $E[M_{ab}^X(D_m)] \leqslant \frac{1}{b-a} E[(X_T - a)^-]$. Then we drive $m \to \infty$. We thus have $M_{ab}^f([0,T] \cap D) < \infty$ a.s. Set $N = \cup_{T \in D}(\{sup_{t \in D \cap [0,T]} |X_t| = \infty\} \cup \{\cup_{a,b \in Q, b<a} \{M_{ab}^X(D \cap [0,T]) = \infty\})$. Then, the right and left limit exist.
(2) It follows from (1). We set

$$X_{t^+}(\omega) = \begin{cases} \lim_{s \downarrow t, s \in D} X_s(\omega) & \\ 0, & otherwise \end{cases} \tag{4}$$

With this definition, X_{t^+} is F_{t^+}-measurable.
Fix $t \geqslant 0$ and choose a sequence in D such that t_n decreases to t as $n \to$. Then by construction, we have a.s. $X_{t^+} = \lim_{n \to \infty} X_{t_n}$. Set $Y_k = X_{t_{-k}}$ for every $k \leqslant 0$. Then Y is a backward supermartingale with respect to the backward discrete filtration. Since $sup_{k \leqslant 0} E|Y_k| < \infty$, the backward convergence theorem can be applied and then $X_{t_n} \xrightarrow{L^1} X_{t^+}$. Thanks to L^1 convergence, $X_t \geqslant E[X_{t_n}|F_t] \Rightarrow X_t \geqslant \lim_{n \to \infty} E[X_{t_n}|F_t] = E[\lim_{n \to \infty} X_{t_n}|F_t] = E[X_{t^+}|F_t]$. Thanks again to L^1 convergence, we have $E[X_{t^+}] = \lim E[x_{T_n}$. Thus, if the function $s \to E[X_s]$ is right-continuous, we must have $E[X_t] = E[X_{t^+}] = E[E[x_{t^+}|F_t]]$, and the inequality $X_t \geqslant E[X_{t^+}|F_t]$ then forces $X_t = E[X_{t^+}|F_t]$. □

Theorem. *Assume that the filtration F_t is right-continuous and complete. Let X_t be a supermartingale, such that the function $t \to E[X_t]$ is right-continuous. Then X has a modification with cadlag sample paths, which is also an F_t-supermartingale.*

Proof. We can construct

$$Y_t(\omega) = \begin{cases} X_{t^+}(\omega), & \omega \quad not \quad in \quad N \\ 0, & \omega \in N \end{cases} \tag{5}$$

Then the sample paths of Y_t are cadlag.
The random variable X_{t^+} is F_{t^+}-measurable, and thus F_t-measurable since the filtration is right-continuous. As the negligible set N belongs to F_∞, the completeness of the filtration ensures Y_t is F_t-measurable. By the previous theorem, $X_t = E[X_{t^+}|F_t] = E[X_{t^+}|F_{t^+}] = X_{t^+} = Y(t)$ a.s. Consequently, Y_t is a modification of X_t. □

Definition. *A class C of random variables is called uniformly integrable if given $\epsilon > 0$, $\exists K \in [0, \infty)$ s.t. $\forall X \in C, E(|X|1_{|X|>K}) < \epsilon$.*

Theorem. *(An absolute continuity property of Lebesgue integral) Assume f is Lebesgue integrable on E. $\forall \epsilon > 0$, $\exists \delta$ s.t. if the Lebesgue measure of A is less than δ, the integral of $|f|$ over A is less than ϵ.*

Proof. Note that by DCT, we have that $\lim_{\lambda\to\infty} \int_{\{|f|>\lambda\}} |f| d\mu = 0$. Let $\epsilon > 0$, there exists λ s.t. $\int_{\{|f|>\lambda\}} |f| d\mu < \frac{\epsilon}{2}$. Choose $\delta \leqslant \frac{\epsilon}{2\lambda}$ and take any measurable A s.t. $\mu(A) < \delta$. Then $\int_A |f| d\mu = \int_{A\cap\{|f|>\lambda\}} |f| d\mu + \int_{A\cap\{|f|\leqslant\lambda\}} |f| d\mu \leqslant \int_{\{|f|>\lambda\}} |f| d\mu + \int_{A\cap\{|f|\leqslant\lambda\}} |f| d\mu \leqslant \frac{\epsilon}{2} + \delta\lambda \leqslant \epsilon$. □

Theorem. *(Bounded convergence theorem) Let X_n be a sequence of random variables, and let X be a random variable. Suppose that $X_n \xrightarrow{p} X$ and for some K is nonnegative and finite, we have for every n and ω, $|X_n(\omega)| \leqslant K$. Then $X_n \xrightarrow{L^1} X$.*

Proof. $P(|X| > K + k^{-1}) \leqslant P(|X - X_n| > k^{-1})$. $P(|X| > K) = P(\cup_k\{|X| > K + k^{-1}\}) = 0$. Let $\epsilon > 0$ be given. Choose n_0 s.t. $P(|X_n - X| > \frac{\epsilon}{3}) < \frac{\epsilon}{3K}$ when $N \geqslant n_0$. $E(|X_n - X|) = E(|X_n - X|1_{|X_n-X|>\frac{\epsilon}{3}}) + E(|X_n - X|1_{|X_n-X|\leqslant\frac{\epsilon}{3}}) \leqslant 2KP(|X_n - X| > \frac{\epsilon}{3}) + \frac{\epsilon}{3} \leqslant \epsilon$. □

Theorem. *Let X_n be a sequence in L^1, and let $X \in L^1$. Then $X_n \xrightarrow{L^1} X$ if and only if $X_n \xrightarrow{p} X$ and X_n is uniformly integrable.*

Proof. (Proof of if part) For $K \in [0, \infty)$, define a function $\varphi_K(x)$:

$$\varphi_K(x) = \begin{cases} K, & if \quad x > K \\ x, & if \quad |x| \leqslant K \\ -K, & if \quad x < -K \end{cases} \tag{6}$$

Let $\epsilon > 0$ as given. By the UI property of X_n, we can choose K s.t. $\forall n, E\{|\varphi_K(X_n) - X_n|\} < \frac{\epsilon}{3}$ and $E\{\varphi_K(X) - X|\} < \frac{\epsilon}{3}$. Since $P(|\varphi_K(X_n) - \varphi_K(X)| > \epsilon) \leqslant P(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$, $\varphi_K(X_n) \xrightarrow{p} \varphi_K(X)$. By the bounded convergence theorem, we have n_0 s.t. $\forall n \geqslant n_0$, $E\{|\varphi_K(X_n) - \varphi_K(X)|\} < \frac{\epsilon}{3}$. Since

$E(|X_n - X|) \leqslant E\{|\varphi_K(X_n) - X_n|\} + E\{|\varphi_K(X_n) - \varphi_K(X)|\} + E\{\varphi_K(X) - X|\} < \epsilon$

(Proof of only if part) Suppose $X_n \xrightarrow{L^1} X$. Let $\epsilon > 0$ be given. Choose N such that $n \geqslant N \Rightarrow E(|X_n - X|) < \frac{\epsilon}{2}$. By the absolute continuity of Lebesgue integral, we can choose $\delta > 0$ s.t. $P(F) < \delta$, we have $E(|X_n|1_F) < \epsilon \quad (1 \leqslant n \leqslant N)$ and $E(|X|1_F) < \frac{\epsilon}{2}$. Since X_n is bounded in L^1, we can choose K such that $K^{-1} sup_r E(|X_t|) < \delta$. Then for $n \geqslant N$, we have $P(|X_n| > K) < \delta$ and $E(|X_n|1_{\{|X_n > K\}}) \leqslant E(|X|1_{\{|X_n > K\}}) + E(|X - X_n|) < \epsilon$. For $n \leqslant N$, we have $P(|x_n| > K) < \delta$ and $E(|X_n|1_{\{|X_n|>K\}}) < \epsilon$. x_n is a UI family. The convergence in probability is directly implied by convergence in L^1. □

Definition. *A martingale X_t is said to be closed if there exists $Z \in L^1$ s.t. for every $t \geqslant 0$, $X_t = E[Z|F_t]$.*

Theorem. *Let X be a martingale with right-continuous sample paths. Then the following properties are equivalent.*
(1) X is closed.
(2) the collection X_t is uniformly integrable.
(3) X_t converges a.s. and in L^1.

Proof. (1) $\Rightarrow$ (2) Suppose $Z \in L^1$ closes X_t. Let $\epsilon > 0$ be given. Choose $\delta > 0$ s.t. $F \in F$, $P(F) < \delta$, then $E(|Z|1_F) < \epsilon$. Choose K s.t. $K^{-1}E(|Z|) < \delta$. Since $X_t = E(Z|F_t)$. By Jensen's inequality, $E|X_t| \leqslant E|z|$ and $KP(|X_t| > K) \leqslant E|x_t| \leqslant E|Z|$. Therefore, $P(|X_t| > K) < \delta$. Since $|X_t| > K$ is F_t-measurable and thus F measurable, $E(|X_t|1_{\{|X_t|>K\}}) \leqslant E(|Z|1_{\{|X_t|>K\}}) < \epsilon$.
(2) $\Rightarrow$ (3) It is easily seen by applying the martingale convergence theorem.
(3) $\Rightarrow$ (1) By simply take $Z = X_\infty$. □

Theorem. *(Optional stopping theorem, discrete version) Let X be a supermartingale. Let T be a stopping time. Then X_T is integrable and $E(X_T) \leqslant E(X_0)$ in each of the following situations:*
(1) T is bounded.
(2) X is bounded and T is a.s. finite.
(3) $E(T) < \infty$ and for some K, $|X_n(\omega) - X_{n-1}(\omega)| \leqslant K$.

Proof. We know that $E(X_{T\wedge n} - X_0) \leqslant 0$. For (1), we can take $n = N$.
For (2), we can let $n \rightarrow$ by using bounded convergence theorem.
For (3), $|X_{T\wedge n} - X_0| = |\sum_{k=1}^{T\wedge n}(X_k - X_{k-1})| \leqslant$, then by applying the DCT we have proved the theorem. □

Theorem. *If X_n is a uniformly integrable submartingale then for any stopping time N, $X_{N\wedge n}$ is uniformly integrable.*

Proof. X_n^+ is a submartingale, so $EX_{N\wedge n}^+ \leqslant EX_N^+$. Since X_n^+ is uniformly integrable, it follows that $sup_n EX_{N\wedge n}^+ \leqslant sup_n EX_n^+ < \infty$. By the martingale convergence theorem, $X_{N\wedge n} \rightarrow X_N$ a.s. and $E|X_N| < \infty$. $E(|X_{N\wedge n}|1_{\{|X_{N\wedge n}|>K\}} = E(|X_N|1_{\{|X_N|>K, N\leqslant n\}} + E(|X_N|1_{\{|X_N|>K, N>n\}} \leqslant E(|X_n|1_{\{|X_n|>K\}} + E(|X_N|1_{\{|X_N|>K\}}$. □

Theorem. *If X_n is a uniformly integrable submartingale then for any stopping time $N \leqslant \infty$, we have $EX_0 \leqslant EX_N \leqslant EX_\infty$.*

Proof. Letting $n \to \infty$ implies to $X_{N\wedge n} \xrightarrow{L^1} X_N$ and $X_n \xrightarrow{L^1} X_\infty$. □

Theorem. *(Levy's upward theorem) Let M_n be closed by ϵ. Then M is a UI martingale and $M_n \to E(\epsilon|F_\infty)$ a.s. and in L^1.*

Proof. It suffices to show that $M_\infty = E(\epsilon|F_\infty)$. Now consider the measures Q_1 and Q_2 on (Ω, F_∞), where $Q_1(F) = E(E(\epsilon|F_\infty)1_F)$ and $Q_2(F) = E(M_\infty)1_F)$, $F \in F_\infty$. If F is in F_n, then $E(E(\epsilon|F_\infty)1_F) = E(E(E(\epsilon|F_\infty)1_F)|F_n) = E(M_n 1_F) = E(M_\infty 1_F)$. Since F_n is a π-system generating F_∞, therefore Q_1 and Q_2 agree on F_∞. □

Theorem. *(Optional stopping theorem for uniformly integrable martingale) If $L \leqslant M$ are stopping times and $Y_{M\wedge n}$ is a uniformly integrable submartingale then $EY_L \leqslant EY_M$ and $Y_L \leqslant E(Y_M|F_L)$.*

Proof. Let $A \in F_L$ and define

$$N = \begin{cases} L, & on \quad A \\ M, & on \quad A^c \end{cases} \tag{7}$$

is a stopping time because $\{N = n\} = (\{L = n\} \cap A) \cup \cup_{m=1}^{n}(\{L = m\} \cap \{M = n\} \cap A^c)$. Since M=N on A^c and $EY_N = E[Y_N 1_A] + E[Y_N 1_{A^c}]$, it follows that $E[Y_L 1_A] = E[Y_N 1_A] \leqslant E[Y_M 1_A] = E[E[Y_M|F_L]1_A]$. In particular, if $\epsilon > 0$ and we let $A = \{Y_L - E[Y_M|F_L] > \epsilon\} \in F_L$, then $\epsilon P(A) \leqslant E[Y_L - E[Y_M|F_L]] \leqslant 0$ and so $P(A) = 0$. We have $Y_L \leqslant E(Y_M|F_L)$ a.s. □

19.3 Local martingales

Definition. *An adapted process $M(t)$ is called a local martingale if there exists a sequence of stopping time T_n s.t. $T_n \uparrow \infty$ and for each n the stopped process $M(t \wedge T_n)$ is a uniformly integrable martingale in t.*

Theorem. *Let $M(t)$ be a local martingale such that $|M(t)| \leqslant Y$, with $EY < \infty$. Then M is a uniformly integrable martingale.*

Proof. Let T_n be a localizing sequence. Then for any n and $s < t$.

$$E(M_{t\wedge T_n}|F_s) = M_{s\wedge T_n}. \tag{8}$$

M is clearly integrable. By dominated convergence of conditional expectations $\lim_{n\to\infty} E(M_{t\wedge T_n}|F_s) = E(M_t|F_s)$. Since $\lim_{n\to n} M_{s\wedge T_n} = M_s$, $\lim_{n\to\infty} E(M_{t\wedge T_n}|F_s) = E(M_t|F_s) = M(s)$. And the UI property is clear. □

Theorem. *A non-negative local martingale M_t is a supermartingale, that is $EM_t < \infty$, and for $s < t$, $E(M_t|F_s) \leqslant M_s$.*

Proof. Since $M_{t\wedge T_n} \geqslant 0$, by Fatou's lemma $E(\lim_{n\to\infty} inf M_{t\wedge T_n}) \leqslant \lim_{n\to\infty} inf E(M_{t\wedge T_n})$. Since the limit exists, $E(\lim_{n\to\infty} inf M_{t\wedge T_n}) = E(M_t) \leqslant \lim_{n\to\infty} inf E(M_{t\wedge T_n}) = EM_0$, so that M is integrable. Then applying Fatou's lemma again for conditional expectations, $E(\lim_{n\to\infty} inf M_{t\wedge T_n}|F_s) \leqslant \lim_{n\to\infty} inf E(M_{t\wedge T_n}|F_s) = M_{s\wedge T_n}$. Then drive $n \to \infty$. We obtain $E(M_t|F_s) \leqslant M_s$. □

Definition. *A process X is of Dirichet class D, if the family X_T is uniformly integrable.*

Theorem. *A local martingale is a uniformly integrable martingale if and only if it is of class D.*

Proof. Suppsoe that M is a local martingale of class D. Let T_n be a localizing sequence. Since $T_n \to \infty$, $M_{s\wedge T_n} \to M_s$ a.s. By class D property, $M_{s\wedge T_n} \to M_s$ also in L^1. Using the properties of conditional expectation, $E|E(M_{t\wedge T_n}|F_s) - E(M_t|F_s)| = E|E(M_{t\wedge T_n} - M_t|F_s) \leqslant E(E|M_{t\wedge T_n} - M_t||E_s) = E|M_{t\wedge T_n} - M_t|$. The latter converges to zero. This implies $E(M_{t\wedge T_n}|F_s) \to E(M_t|F_s)$ as $n \to \infty$. $\lim_{n\to} M_{s\wedge T_n} = M(s) = \lim_{n\to} E(M_{t\wedge T_n}|F_s) = E(M_t|F_s)$. □

Proposition 17. *For every $t \in (0,T]$, $\int_0^t |da(s)| = sup\{\sum_{i=1}^p |a(t_i) - a(t_{i-1})|\}$. Clearly, it is enough to treat the case $t = T$. $|a(t_i) - a(t_{i-1})| = |\mu((t_{i-1}, t_i])| \leqslant |\mu|((t_{i-1}, t_i])$. In order to show the reverse inequality, we will use a martingale argument, leaving aside the trivial case and introduce the probability space $\Omega = [0,T]$, which is equipped with the Borel σ-field $B[0,T]$ and the probability measure $P(ds) = (|\mu|([0,T]))^{-1}|\mu|(ds)$. On this probability space, we consider discrete filtration B_n s.t. for every integer $\geqslant 0$, B_n is the σ-field generated by the intervals $(t_{i-1}^n, t_i^n], 1 \leqslant i \leqslant p_n$. We then set $X(s) = 1_{D^+}(s) - 1_{D^-}(s) = \frac{d\mu}{d|\mu|}$ and for every n, $X_n = E[X|B_n]$ and is a constant. Since X_n is closed martingale and thus converges to X in L^1. In addition, since $|X(s)| = 1, |\mu|(ds)$ a.e., $\lim_{n\to\infty} E|X_n| = E|X| = 1$. Note that $E|X_n| = (|\mu|([0,T]))^{-1} \sum_{i=1}^{p_n} |a(t_i^n) - a(t_{i-1}^n)|$. Drive $n \to \infty$.*

Proposition 18. *Let A be a finite variation process, and let H be a progressive process such that $\forall t \geqslant 0, \forall \omega \in \Omega, \int_0^t |H_s(\omega)||dA_s(\omega)| < \infty$. Then the process $H \cdot A$ defined by $(H \cdot A)_t = \int_0^t H_s dA_s$ is also a finite variation process.*

Theorem. *Let M be a continuous local martingale. Assume that M is also a finite variation process, in particular $M_0 = 0$. Then $M_t = 0$ a.s.*

Proof. Set $T_n = inf\{t \geqslant 0 : \int_0^t |dM_s| \geqslant n\}$ for every integer $n \geqslant 0$. Fix $n \geqslant 0$ and set $N = M^{T_n}$. $|N_t| = |M_{t\wedge T_n}| \leqslant \int_0^{t\wedge T_n} |dM_s| \leqslant n$. N is a bounded martingale. Then, we have $E[N_t^2] = \sum_{i=1}^p E[(N_{t_i} - N_{t_{i-1}}])^2] \leqslant E[(sup_{1\leqslant p\leqslant p}|N_{t_i} - N_{t_{i-1}}|) \sum_{i=1}^p |N_{t_i} - N_{t_{i-1}}|] \leqslant nE[sup_{1\leqslant p\leqslant p}|N_{t_i} - N_{t_{i-1}}|]$. Since N is bounded and with continuous sample paths, $\lim_{k\to\infty} E[sup_{1\leqslant p\leqslant p}|N_{t_i^k} - N_{t_{i-1}^k}|] = 0$. Then, $E[N_t^2] = 0$, and hence $M_{t\wedge T_n} = 0$ a.s. Letting n tend to ∞, we get $M_t = 0$ a.s. □

Theorem. *(The quadratic variation of a continuous local martingale) Let M_t be a continuous local martingale. There exists an increasing process denoted by $< M, M >_t$, which is unique up to indistinguishability, such that $M_t^2 - < M, M >_t$ is a continuous local martingale. Furthermore, for every fixed t¿0, if $0 = t_0^n < t_1^n < ... < t_{p_n}^n = t$ is an increasing sequences of subdivisions of $[0, t]$ with mesh tending to 0, we have $< M, M >_t = \lim_{n\to\infty} \sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}^n})^2$ in probability.*

Proof. The proof is divided into two parts.
(Proof of uniqueness) If A and A' be two increasing processes satisfying the condition in the statement. Then $A_t - A'_t = (M_t^2 - A')t) - (M_t^2 - A_t$ is both the continuous local martingale and a finite variation process. $A - A' = 0$ a.s.
(Proof of existence) We consider first the case $M_0 = 0$ and M is bounded. Hence M is a true martingale. Fix $K > 0$ and an increasing sequence $0 = t_0^n < t_1^n < ... < t_{p_n}^n = K$ with mesh tending to 0.
We observe that, for every $0 \leqslant r < s$ and for every bounded F_r-measurable variable Z, the process $N_t = Z(M_{s\wedge t} - M_{r\wedge t})$ is a martingale since for $k < r$, $E(N_t|F_k) = 0 = N_k$ and for $r \leqslant k < s$, $E(N_t|F_k) = ZE(M_{s\wedge t} - M_{r\wedge t}|F_k) = ZE(M_{s\wedge t}|F_k) - ZEM_r = ZEM_{s\wedge k} - ZEM_r = N_k$, and for $k \geqslant s$, $E(N_t|F_k) = Z(M_s - M_r) = N_k$. It follows that for every n, $X_t^n = \sum_{i=1}^{p_n} M_{t_{i-1}^n}(M_{t_i^n\wedge t} - M_{t_{i-1}^n\wedge t})$ is a martingale. Then, $M_{t_j^n}^2 - 2X_{t_j^n}^n = M_{t_j^n}^2 - 2\sum_{i=1}^{j} M_{t_{i-1}^n}(M_{t_i^n} - M_{t_{i-1}^n}) = \sum_{i=1}^{j}(M_{t_i^n} - M_{t_{i-1}^n})^2$.
We then claim that $\lim_{n,m\to\infty} E[(X_K^m - X_K^n)^2] = 0$.
Let us fix $n \leqslant m$ and evaluate the product

$$\begin{aligned}
E(X_K^m X_K^n) &= E(\sum_{i=1}^{p_n}[M_{t_{i-1}^n}(M_{t_i^n} - M_{t_{i-1}^n})]\sum_{j=1}^{p_m}[M_{t_{j-1}^m}(M_{t_j^m} - M_{t_{j-1}^m})]) \\
&= \sum_{i=1}^{p_n}\sum_{j=1}^{p_m} E[M_{t_{i-1}^n}(M_{t_i^n} - M_{t_{i-1}^n})M_{t_{j-1}^m}(M_{t_j^m} - M_{t_{j-1}^m})].
\end{aligned}$$

In this double sum, the only terms that may be nonzero are those corresponding to indices i and j such that the interval $(t_{j-1}^m, t_j^m]$ is contained in $(t_{i-1}^n, t_i^n]$ since suppose that $t_i^n \leqslant t_{j-1}^m$ (the case $t_j^m \leqslant t_{i-1}^n$ can be treated analogically), then conditionally on $F_{t_{j-1}^m}$, we have

$$\begin{aligned}
&E[M_{t_{i-1}^n}(M_{t_i^n} - M_{t_{i-1}^n})M_{t_{j-1}^m}(M_{t_j^m} - M_{t_{j-1}^m})] \\
&= E(E[M_{t_{i-1}^n}(M_{t_i^n} - M_{t_{i-1}^n})M_{t_{j-1}^m}(M_{t_j^m} - M_{t_{j-1}^m})|F_{j-1}^m]) \\
&= E[E[M_{t_{i-1}^n}(M_{t_i^n} - M_{t_{i-1}^n})M_{t_{j-1}^m}E(M_{t_j^m} - M_{t_{j-1}^m}|F_{j-1}^m)] = 0
\end{aligned}$$

For every $j = 1, ..., p_m$ write $i_{n,m}(j)$ for the unique index i such that $(t_{j-1}^m, t_j^m] \subset (t_{i-1}^n, t_i^n]$. It follows from the previous considerations that

$$E(X_K^m X_K^n) = \sum_{1\leqslant j\leqslant p_m, i=i_{n,m}(j)} E[M_{t_{i-1}^n}(M_{t_i^n} - M_{t_{i-1}^n})M_{t_{j-1}^m}(M_{t_j^m} - M_{t_{j-1}^m})]$$

In each term $E[M_{t^n_{i-1}}(M_{t^n_i} - M_{t^n_{i-1}})M_{t^m_{j-1}}(M_{t^m_j} - M_{t^m_{j-1}})]$, we can decompose

$$M_{t^n_i} - M_{t^n_{i-1}} = \sum_{k:i_{n,m}(k)=i} (M_{t^m_k} - M_{t^m_{k-1}})$$

We observe that if $i_{n,m}(k) = i, k \neq j$, $E[M_{t^n_{i-1}}(M_{t^m_k} - M_{t^n_{k-1}})M_{t^m_{j-1}}(M_{t^m_j} - M_{t^m_{j-1}})] = 0$ (condition on $F_{t^m_{k-1}}$ if $k > j$ and on $F_{t^m_{j-1}}$ if $k < j$). The case that remains is $k = j$, we have thus obtained

$$E(X^m_K X^n_K) = \sum_{1\leqslant j\leqslant p_m, i=i_{n,m}(j)} E[M_{t^n_{i-1}}M_{t^m_{j-1}}(M_{t^m_j} - M_{t^m_{j-1}})^2]$$

As a special case of this relation, we have

$$E[(X^m_K)^2] = \sum_{1\leqslant j\leqslant p_m} E[M^2_{t^m_{j-1}}(M_{t^m_j} - M_{t^m_{j-1}})^2].$$

Furthermore,

$$\begin{aligned}
E[(X^n_K)^2] &= \sum_{1\leqslant i\leqslant p_n} E[M^2_{t^n_{i-1}}(M_{t^n_i} - M_{t^n_{i-1}})^2] \\
&= \sum_{1\leqslant i\leqslant p_n} E[M^2_{t^n_{i-1}}E(M_{t^n_i} - M_{t^n_{i-1}})^2|F_{t^n_{i-1}})] \\
&= \sum_{1\leqslant i\leqslant p_n} E[M^2_{t^n_{i-1}} \sum_{j:i_{n,m}(j)=i} E[(M_{t^m_j} - M_{t^m_{j-1}})^2|F_{t^n i-1}]] \\
&= \sum_{1\leqslant j\leqslant p_m, i=i_{n,m}(j)} E[M^2_{t^n_{i-1}}(M_{t^m_j} - M_{t^m_{j-1}})^2]
\end{aligned}$$

Then we combine the last three equation:

$$E[(X^n_K - X^m_K)^2] = E[\sum_{1\leqslant j\leqslant p_m, i=i_{n,m}(j)} (M_{t^n_{i-1}} - M_{t^m_{j-1}})^2(M_{t^m} - M_{t^m_{j-1}})^2].$$

Using Cauchy-Schwarz inequality, we then have

$$\begin{aligned}
E[(X^n_K - X^m_K)^2] \leqslant E[sup_{1\leqslant j\leqslant p_m, i=i_{n,m}(j)}](M_{t^n_{i-1}} - M_{t^m_{j-1}})^4]^{\frac{1}{2}} \\
\times E[(\sum_{1\leqslant j\leqslant p_m} (M_{t^m_j} - M_{t^m_{j-1}})^2)^2]^{\frac{1}{2}}.
\end{aligned}$$

By the continuity of sample paths and dominated convergence, we have

$$\lim_{n,m\to\infty, n\leqslant m} E[sup_{1\leqslant j\leqslant p_m, i=i_{n,m}(j)}](M_{t^n_{i-1}} - M_{t^m_{j-1}})^4] = 0$$

To complete the proof that $\lim_{n,m\to\infty} E[(X^n_K - X^m_K)^2] = 0$, it remains to show that there exists a constant C such that, for every m, $E[(\sum_{1\leqslant j\leqslant p_m}(M_{t^m_j} -$

$M_{t^m_{j-1}})^2)^2] \leqslant C$.
Let A be a constant such that $|M_t| \leqslant A$ for every $t \geqslant 0$.

$$E[(\sum_{1\leqslant j\leqslant p_m}(M_{t^m_j}-M_{t^m_{j-1}})^2)^2]$$
$$= E[\sum_{1\leqslant j\leqslant p_m}(M_{t^m_j}-M_{t^m_{j-1}})^4] + 2E[\sum_{1\leqslant j<k\leqslant p_m}(M_{t^m_j}-M_{t^m_{j-1}})^2(M_{t^m_k}-M_{t^m_{k-1}})^2]$$
$$\leqslant 4A^2E[\sum_{1\leqslant j\leqslant p_m}(M_{t^m_j}-M_{t^m_{j-1}})^2] + 2\sum_{j=1}^{p_m-1}E[(M_{t^m_j}-M_{t^m_{j-1}})^2E[\sum_{k=j+1}^{p_m}(M_{t^m_k}-M_{t^m_{k-1}})^2|F_{t^m_j}]]$$
$$= 4A^2E[\sum_{1\leqslant j\leqslant p_m}(M_{t^m_j}-M_{t^m_{j-1}})^2] + \sum_{j=1}^{p_m-1}E[(M_{t^m_j}-M_{t^m_{j-1}})^2E[(M_K-M_{t^m_j})^2|F_{t^m_j}]]$$
$$\leqslant 12A^2E[\sum_{1\leqslant j\leqslant p_m}(M_{t^m_j}-M^{t^m_{j-1}})^2] = 12A^2E[(M_K-M_0)^2] \leqslant 48A^4$$

Then thanks to Dob's inequality in L^2, we have

$$E[sup_{t\leqslant k}(X^n_t-X^m_t)^2] \leqslant 4E[(X^n_K-X^m_K)^2] \Rightarrow \lim_{n,m\to\infty} E[sup_{t\leqslant k}(X^n_t-X^m_t)^2] = 0$$

Therefore, for every $t \in [0, K]$, X^n_t is a Cauchy sequence in L^2 and thus converges in L^2. We want to argue that the limit yields a process Y index by $[0, K]$ with continuous sample paths. To see this, we choose a strictly increasing sequence $(n_k)_{k\geqslant 1}$ of positive integers such that for every $k \geqslant 1$, $E[sup_{t\leqslant K}(X^{n_k+1}_t - X^{n_k}_t)^2] \leqslant 2^{-k}$. This implies that

$$E[\sum_{k=1}^{\infty} sup_{t\leqslant K}|X^{n_k+1}_t - X^{n_k}_t|] < \infty$$

and thus

$$\sum_{k=1}^{\infty} sup_{t\leqslant K}|X^{n_k+1}_t - X^{n_k}_t| < \infty \quad , a.s.$$

Consequently, except on the negligible set N where the series in the last display diverges, the sequence of random functions $(X^{n_k}_t, 0 \leqslant t \leqslant K)$ converges uniformly on $[0, K]$ as $k \to \infty$, and the limiting random function is continuous by uniform convergence.
Since the filtration is complete, we can thus set

$$Y_t(\omega) = \begin{cases} \lim_{k\to\infty} X^{n_k}_t(\omega), & if \quad \omega \in \Omega \setminus N \\ 0, & if \quad \omega \in N \end{cases}$$

Furthermore, since the L^2-limit of X^n_t must coincide with the a.s. limit of a subsequence, Y_t is also the limit of X^n_t in L^2. Then, we can pass to the limit in the martingale property for X_n, to obtain $E[Y_t|F_s] = Y_s$ for every

$0 \leqslant s \leqslant t \leqslant K$. It follows that $(Y_{t\wedge K})_{t\geqslant 0}$ is a martingale with continuous sample paths.
On the other hand, the sample paths of $M_t^2 - 2X_t^n = \sum_{i=1}^{p_n}(M_{t_i^n} - M_{t_{i-1}^n})^2$ are nondecreasing, by passing to the limit $n \to \infty$ along the sequence n_k, we get the sample paths of $M_t^2 - 2Y_t$ are nondecreasing on $[0, K]$, except maybe on the negligible set N. For every $t \in [0, K]$, we set $A_t^{(K)} = M_t^2 - 2Y_t$ on $\omega \setminus N$, and $A_t^{(K)} = 0$ on N. Then $A_0^{(K)} = 0, A_t^{(K)}$ is F_t-measurable for every $t \in [0, K]$. By the uniqueness argument, for the case $M_0 = 0$ and M_t is bounded, the existence equality holds in L^2.
Let us consider the general case. Writing $M_t = M_0 + N_t$, s.t. $M_t^2 = M_0^2 + 2M_0N_t + N_t^2$, and noting that M_0N_t is a continuous local martingale, we see that we may assume that $M_0 = 0$. We then set $T_n = \{t \geqslant 0 : |M_t| \geqslant N\}$ and we can apply the bounded case to the stopped martingales M^{T_n}. Set $A[n] =< M^{T_n}, M^{T_n} >$. The uniqueness hows that the processes $A_{t\wedge T_n}^{[n+1]}$ and $A_t^{[n]}$ are distinguishable. It follows that there exists an increasing process A such that for every n, the processes $A_{t\wedge T_n}$ and $A_t^{[n]}$ are indistinguishable. By construction and the previous theorem, $M_{t\wedge T_n}^2 - A_{t\wedge T_n}$ is a martingale for every n, which precisely implies that $M_t^2 - A_t$ is a continuous local martingale. We take $< M, M >_t = A_t$. Finally, the previously bounded case holds if M and $< M, M >_t$ are replaced by M^{T_n} and $< M, m >_{t\wedge T_n}$. Then it is enough to observe that for every $t > 0$, $P(t \leqslant T_n)$ converges to 1 when $n \to \infty$. □

Theorem. *Let M be a continuous local martingale such that $M_0 = 0$. Then we have $< M, M >= 0$ if and only if $M = 0$.*

Proof. Suppsoe that $< M, M >_t = 0$. Then M_t^2 is a nonnegative continuous local martingale. And by the previous theorem, it is also a supermartingale. Hence, $E(M_t^2) \leqslant E(M_0^2)$. Then, $M_t = 0$ a.s. The converse is obvious. □

Proposition 19. *Let M be a continuous local martingale with $M_0 \in L^2$.*
(1) The following are equivalent:
M is a martingale bounded in L^2
$E[< M, M >_\infty] < \infty$ (2) The following are equivalent:
M is a martingale and $M_t \in L^2$ for every $t \geqslant 0$
$E[< M, M >_t] < \infty$ for every $t \geqslant 0$

Proof. We may assume $M_0 = 0$ in the proof.
(1) Let us first assume that M is a martingale bounded in L^2. By Doob's inequality, for every $T > 0$,

$$E[sup_{0\leqslant t\leqslant T} M_t^2] \leqslant 4E[M_T^2].$$

By letting T goes to infinity, we have

$$E[sup_{t\geqslant 0} M_t^2] \leqslant 4sup_{t\geqslant 0} E[M_t^2] = C < \infty.$$

Set $S_n = inf\{t \geqslant 0 :< M, M >_T \geqslant n\}$. Then the continuous local martingale $M_{t\wedge S_n}^2 - < M, M >_{t\wedge S_n}$ is dominated by the variable $sup_{s\geqslant 0} M_s^2 + n$, which is

integrable. Then $M^2_{t\wedge S_n} - < M, M >_{t\wedge S_n}$ is a uniformly integrable martingale. $E(M^2_{t\wedge S_n}) = E(< M, M >_{t\wedge S_n}) \leqslant C < \infty$. By letting n and then t tend to infinity, and using monotone convergence theorem, we get $E[< M, M >_\infty] \leqslant C < \infty$.
Conversely, assume that $E[< M, M >_\infty] < \infty$. Set $T_n = inf\{t \geqslant 0 : |M_t| \geqslant n\}$. Then the continuous local martingale $M^2_{t\wedge T_n}$ is dominated by $n^2 + < M, M >_\infty$, which is integrable. Hence, this continuous local martingale is a uniformly integrable martingale. Using Fatou's lemma,

$$
\begin{aligned}
&E[\lim_{n\to\infty} inf M^2_{t\wedge T_n}] \leqslant \lim_{n\to\infty} inf E[M^2_{t\wedge T_n}] \\
&= \lim_{n\to\infty} inf E[< M, M >_{t\wedge T_n}] \\
&= \lim_{n\to\infty} inf E[< M, M >_\infty] < \infty E[\lim_{n\to\infty} inf M^2_{t\wedge T_n}] \leqslant \lim_{n\to\infty} inf E[M^2_{t\wedge T_n}] \\
&= \lim_{n\to\infty} inf E[< M, M >_{t\wedge T_n}] \\
&= \lim_{n\to\infty} inf E[< M, M >_\infty] < \infty.
\end{aligned}
$$

So M_t is bounded in L^2. In addition, since the bound on $E[M^2_{[}t \wedge T_n]$ shows that the sequence is uniformly integrable, and therefore converges both a.s. and in L^1 to M_t, for every $t \geqslant 0$. Recalling that M^{T_n} is a martingale, $E[M_{t\wedge T_n}|F_s] = M_{s\wedge T_n}$, for $0 \leqslant s < t$. By the L^1 covergence, $E[\lim_{n\to\infty} M_{t\wedge T_n}|F_s] = \lim_{n\to\infty} M_{s\wedge T_n} = M_s$. Thus, M is a martingale. The uniformly integrable property is clear. □

Definition. *If M and N are two continuous local martingales, the brackets is the finite variation process defined by setting, for every* $t \geqslant 0$, $< M, N >_t = \frac{1}{2}(< M + N, M + N >_t - < M, M >_t - < N, N >_t)$.

Theorem. *Let M and N be continuous local martingales and let H and K be two measurable processes. Then, a.s.,* $\int_0^\infty |H_s||K_s||d < M, N >_s | \leqslant (\int_0^\infty H_s^2 d < M, M >_s)^{\frac{1}{2}} (\int_0^\infty K_s^2 d < N, N >_s)^{\frac{1}{2}}$

19.4 Stochastic integrals for martingales bounded in L^2

We denote H^2 for the space of all continuous martingales M which are bounded in L^2 and such that $M_0 = 0$. In addition, if $M, N \in H^2$, the random variable $< M, N >_\infty$ is well-defined and we have $E| < M, N >_\infty | < \infty$. This allows us to define a symmetric bilinear form on H^2 via the formula $(M, N)_{H^2} = E < M, N >_\infty = E[M_\infty N_\infty]$. Clearly, $(M, M)_{H^2} = 0$ if and only if $M_t = 0$. Then, the scalar product $(M, N)_{H^2}$ thus yields a norm on H^2 given by

$$||M||_{H^2} = (M, M)^{\frac{1}{2}}_{H^2} = E[< M, M >_\infty]^{\frac{1}{2}} = E[(m_\infty)^2]^{\frac{1}{2}}.$$

Proposition 20. *The space* H^2 *equipped with the scalar product* $(M, N)_{H^2}$ *is a Hilbert space.*

Proof. We need to verifty the completeness of the space. Let M^n be a sequence in H^2 which is Cauchy for that norm. We have then

$$\lim_{m,n\to\infty} E[(M^n_\infty - M^n_\infty)^2] = \lim_{m,n\to\infty}(M^n - M^m, M^n - M^m)_{H^2} = 0$$

Consequently, the sequence M^n_∞ converges in L^2 to a limit, which we denote by Z.
By Doob's inequality, $E[sup_{t\geqslant 0}(M^n_t - M^m_t)^2] \leqslant 4E[(M^n_\infty - M^m_\infty)^2]$. We thus obtained that $\lim_{m,n\to\infty} E[sup_{t\geqslant 0}(M^n_t - M^m_t)^2] = 0$. Hence for every $t \geqslant 0$, m^n_t converges in L^2.
Then, we want to argue that the limit yields a processs with continuous sample paths. We fisrdt choose an increasing $n_k \uparrow \infty$ s.t.

$$E[\sum_{k=1}^{\infty} sup_{t\geqslant 0}|M^{n_k}_t - M^{n_k+1}_t|] \leqslant \sum_{k=1}^{\infty} E[sup_{t\geqslant 0}(M^{n_k}_t - M^{n_k+1}_t)^2]^{\frac{1}{2}} < \infty.$$

The last display implies that, a.s. $\sum_{k=1}^{\infty} sup_{k\geqslant 1}|M^{n_k}_t - M^{n_k+1}_t| < \infty$, and thus the sequence converges uniformly on R^+ a.s. to a limit denoted by $(M_t)_{t\geqslant 0}$. On the negligible set where the uniform convergence does that hold, we take $M_t = 0$ for every $t \geqslant 0$. Clearly the limiting process has continuous sample paths and is adapted. Furthermore, by L^2 convergence, we can yield that M_t is a continuous martingale and is bounded in L^2, so that $M \in H^2$. The a.s. convergence of $(M^{n_k}_t)_{t\geqslant 0}$ to $(M_t)_{t\geqslant 0}$ then ensures $M_\infty = \lim M^{n_k}_\infty = Z$ a.s. Finally, the L^2 convergence of (M^n_∞) to Z shows that the sequence converges to M in H^2. $\square$

We denote the progreesive σ-field on $\Omega \times R^+$ by P and if $M \in H^2$, we let $L^2(M)$ be the set of all progressive processes such that

$$E[\int_0^\infty H^2_s d < M, M >_s] < \infty.$$

We can view $L^2(M)$ as an ordinary L^2 space, namely,

$$L^2(M) = L^2(\Omega \times R^+, P, dPd < M, M >_s)$$

where $dpd < M, M >_s$ refers to the finite measure on $(\omega \times R^+, P)$ that assigns the mass to a set $A \in P$

$$E[\int_0^\infty 1_A(\omega, s) d < M, M >_s]$$

Just like any L^2 space, $L^2(M)$ is a Hilbert space with the associated norm $||H||_{L^2} = (E[\int_0^\infty H^2_s d < M, M >_s])^{\frac{1}{2}}$.

Definition. *An elementary process is a progressive process of the form*

$$H_s(\omega) = \sum_{i=0}^{p-1} H_{(i)}(\omega) 1_{(t_i, t_{i+1}]}(s)$$

, where $0 = t_0 < t_1 < ... < t_p$ and for every $i \in \{0, 1, ..., p-1\}$, $H_{(i)}$ is a bounded F_{t_i} measurable random variable.
The set ε of all elementary processes forms a linear subspace of $L^2(M)$. To be precise, we should here say equivalence classes of elementary processes. (Recall that H and H' are identified in $L^2(M)$ if $||H - H'||_{L^2(M)} = 0$)

Proposition 21. *For,every $M \in H^2$, ε is dense in $L^2(M)$.*

Proof. It suffices to show that if $K \in L^2(M)$ is orthogonal to ε, then $K = 0$. Assume that $K \in L^2(M)$ is orthogonal to ε, and set for every $t \geqslant 0$,

$$X_t = \int_0^t K_u d < M, M >_u .$$

The integral on the ride hand side makes sense and is finite since by the Cauchy Schwarz inequality,

$$E[\int_0^t |K_u| < M, M >_u] \leqslant (E[\int_0^t (K_u)^2 d < M, M >_u])^{\frac{1}{2}} \times (E[< M, M >_\infty])^{\frac{1}{2}}$$

Therefore, we yield that a.s. $\forall t \geqslant 0, \int_0^t |K_u| d < M, M >_u < \infty$. Then, X_t is a finite variation process and bounded in L^1.
Next, we let the elementary process $H_r(\omega) = F(\omega) 1_{(s,t]}(r)$. Writing $(H \cdot M)_{L^M} = 0$, we get

$$\begin{aligned} 0 &= (H \cdot M)_{L^2(M)} \\ &= E[\int_0^\infty H_u K_u d < M, M >_u] \\ &= E[\int_s^t H_u K_u d < M, M >_u] \\ &= E[F \int_s^t K_u d < M, M >_u] \end{aligned}$$

It follows that $E[F(X_t - X_s)] = 0$ for every $s < t$ and every bounded F_s measurable variable F. Since the process X is adapted and we know that $X_r \in L^1$ for every $r \geqslant 0$, this implies that X is a martingale. On the other hand, x is a finite variation process. Thus $X = 0$ a.s. Then, $X_t = \int_0^t K_u d < M, M >_u = 0, \forall t \geqslant 0$ a.s. Thus, $K_u = 0, d < M, M >_u$ a.e. a.s. □

Theorem. *Let $M \in H^2$. For every $H \in \varepsilon$ of the form*

$$H_s = \sum_{i=0}^{p-1} H_{(i)}(\omega) 1_{(t_i, t_{i+1}]}(s)$$

the formula

$$(H \cdot M)_t = \sum_{i=0}^{p-1} H_{(i)}(M_{t_{i+1} \wedge t} - M_{t_i \wedge t})$$

defines a process $H \cdot M \in H^2$. The mapping $H \longrightarrow H \cdot M$ extends to an isometry from $L^2(M)$ into H^2. Furthermore, $(H \cdot M)$ is the unique martingale of H^2 that satisfies the property

$$< H \cdot M, N >= h\cdot < M, N >, \forall N \in H^2$$

If T is a stopping time, we have

$$(1_{[0,T]}H) \cdot M = (H \cdot M)^T = H \cdot M^T$$

We often use the notation

$$(H \cdot M)_t = \int_0^t H_s dM_s$$

Proof. It is easy to check that for every $i \in \{0, 1, ..., p-1\}$, set $M_t^i = H_{(i)}(M_{t_{i+1}\wedge t} - M_{t_i \wedge t})$ for every $t \geqslant 0$, M^i is a continuous martingale. Since $H_{(i)}$ is bounded, it follows that $H \cdot M = \sum_{i=0}^{p-1} H_{(i)}(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})$ is a martingale in H^2. In addition, M^i are orthogonal and their respective quadratic variations are given by

$$< M^i, M^i >_t = H_{(i)}^2(< M, M >_{t_{i+1}\wedge t} - < M, M >_{t_i\wedge t}).$$

We conclude that $< H \cdot M, H \cdot M >_t = \sum_{i=0}^{p-1} = H_{(i)}^2(< M, M >_{t_{i+1}\wedge t} - < M, M >_{t_i \wedge t}) = \int_0^t H_s^2 d < M, M >_s$. Consequently,

$$||H \cdot M||_{H^2}^2 = E[< H \cdot M, H \cdot M >_\infty] = E[\int_0^\infty H_s^2 d < M, M >_s] = ||H||_{L^2(M)}^2$$

Therefore, the mapping $H \longrightarrow H \cdot M$ makes sense from ε viewed as a subspace of $L^2(M)$ into H^2. The latter mapping is linear, and since it preserves the norm, it is an isometry from ε into H^2. Since ε is dense in $L^2(M)$ and H^2 is a Hilbert space, this mapping can be extended in a unique way to an isometry from $L^2(M)$ into H^2.
Next, we fix $N \in H^2$. We first note that, if $H \in L^2(M)$, the Kunita-Watanabe inequality shows that

$$E[\int_0^\infty |H_s||d < M, N >_s |] \leqslant ||H||_{L^2(M)}||N||_{H^2} < \infty$$

and thus the variable $\int_0^\infty H_s d < M, N >_S = (H\cdot < M, N >)_\infty$ is well defined and in L^1.
Consider first the case where H is an elementary process of the form given in the statement of the theorem, and define the continuous martingale $M^i, 0 \leqslant i \leqslant p-1$, as previously. Then, we have

$$< H \cdot M, N >= \sum_{i=0}^{p-1} < M^i, N >$$

It follows that

$$< H{\cdot}M, N >_t = \sum_{i=0}^{p_{n-1}} H_{(i)}(< M, N >_{t_{i+1}\wedge t} - < M, N >_{t_i \wedge t}) = \int_0^t H_s d < M, N >_s .$$

Hence, we prove that the property $< H \cdot M, N >= h\cdot < M, N >, \forall N \in H^2$ holds for $H \in \varepsilon$.
We then observe that the linear mapping $X \longrightarrow < X, N >_\infty$ is continuous from H^2 L^1 since again by Kunita-Watanabe inequality,

$$E[| < X, N >_\infty |] \leqslant E[< X, X >_\infty^{\frac{1}{2}} E[< N, N >_\infty]^{\frac{1}{2}} = ||N||_{H^2}||X||_{H^2}.$$

If H^n is a sequence in ε, such that $H_n \to H$ in $L^2(M)$, we have therefore

$$< H{\cdot}M, N >_\infty = \lim_{n\to\infty} < H^n{\cdot}M, N >_\infty = \lim_{n\to\infty} (H^n\cdot < M, N >)_\infty = (H\cdot < M, N >)_\infty$$

where the first equality holds for continuity, the secong equality holds for the property we have proved for elementary processes, and the third equality holds in L^1 since again by Kunita-Watanabe inequality,

$$E[| \int_0^\infty (H_s^n - H_s) d < M, N >_s |] \leqslant E[< N, N >_\infty]^{\frac{1}{2}} ||H^n - H||_{L^2(M)}.$$

Then, we can replace N by the stopped martingale N^t in this identity and yield $< H \cdot M, N >_t = (H\cdot < M, N >)_t$.
If $N \in H^2$,

$$\begin{aligned} < (H \cdot M)^T, N >_t &= < H \cdot M, N >_{t\wedge T} \\ &= (H\cdot < M, N >)_{t\wedge T} \\ &= (1_{[0,T]} H\cdot < M, N >)_t \\ &= < 1_{[0,T]} H \cdot M, N >_t \end{aligned}$$

$$< H \cdot M^T, N >= H\cdot < M^T, N >= H\cdot < M, N >^T = 1_{[0,T]} H\cdot < M, N >,$$

we proved that $(1_{[0,T]} H) \cdot M = (H \cdot M)^T = H \cdot M^T$. $\square$