not typed yet

Foundation of stochastic processes

Dedicated to Professor K. Yosida.

Foundations of stochastic processes

Notation

Chapter 1. ✓ analytic spaces and standard spaces ~~Preliminaries~~ ~~The advanced theory of measurability of measures~~

Chapter 2. ~~D. General facts~~ Fundamental ~~Notions~~ Concepts in probability theory

Chapter 3. Additive Processes.

Chapter 4. Markov Processes

Chapter 5. Stochastic ~~[illegible]~~ differential equations

Chapter 6. Stationary processes

Appendix

Notation

$\delta(A)$	the diameter of a subset A of a metric space.
$\#A$	the cardinal number of A
1_A	the indicator of A
A^c	the complement of A in a certain basic set.
$\cap$	intersection (of sets)
$\cup$	union (of sets)
Σ or $+$	disjoint union (of sets)
$\setminus$	difference (of sets)
$-$	proper difference (of sets)
■	Q.E.D.
$\prod$	Cartesian product
A^n	the Cartesian product of n copies of A (countably many copies if $n=\infty$)
$a \wedge b$	the minimum of a and b
$a \vee b$	the maximum of a and b
$c := a+b$	c is defined to be equal to $a+b$

$\{a, b, c\}$ = the set consisting of a, b and c

(a, b) open interval

$[a, b]$ closed interval

$\mathbb{N}$	the set of all natural numbers
[illegible]	[illegible]
$\mathbb{R}$	the set of all real numbers
$\mathbb{Q}$	the set of all rational numbers
$\mathbb{K}$	the Cantor set

~~chapter~~

Theorem 1. 2. 3	Theorem 3 of Section 2 of Chapter 1. ~~the section number is omitted in the same section or chapter~~
Theorem 2. 3	Theorem 3 of Section 2 in the same chapter.
Theorem 3	Theorem 3 in the same section.

Similarly for numbering lemmas and sections.

$U(a, r)$ (or $\bar{U}(a,r)$)	the open (or closed) ball with center a and radius r in a metric space
$\bar{A}$	the closure of A in a topological space

Chapter 1. Analytic spaces and standard spaces

A Hausdorff topological space S is called an analytic space if we can find a complete separable metric space P and a continuous surjection $f : P \to S$. If we can take a continuous bijection f in this definition, then S is called a **standard** space. These special topological spaces are important in the theory of stochastic processes. First, practically all function spaces useful in probability theory are analytic (even standard) spaces (See Section 7). Second, these spaces have many nice Borel properties which are not enjoyed by general Hausdorff topological spaces (See Sections 4.5.6). We assume the reader to be familiar with measure theory in general, but we will give a quick review of σ-algebras, Borel spaces and the analytic operation in Sections 1, 2 and 3 to standardize our terminology and notation. In Section 8 we will introduce standard (or analytic) Borel spaces and in Sections 9 and 10 we will discuss some properties of probability measures which may not be discussed in standard textbooks of measure theory.

1. σ-algebras.

Let S be a set. A class of subsets of S is often called a class on S. A non-empty class on S is called complementary (resp. multiplicative, additive, σ-additive) if it is closed under complements (resp. finite intersections, finite unions, countable unions). A complementary additive (resp. σ-additive) class on S is called an algebra (resp. a σ-algebra) on S.

A class closed under countable disjoint unions and proper differences is called a Dynkin class if it contains S (as a member).

Let $\mathfrak{A}$ be an arbitrary class on S. The smallest σ-algebra containing $\mathfrak{A}$ (as a subclass) is called the σ-algebra generated by $\mathfrak{A}$, denoted by $\sigma[\mathfrak{A}]$. The Dynkin class $\delta[\mathfrak{A}]$ generated by $\mathfrak{A}$ is defined in the same way.

Theorem 1. (The Dynkin class theorem) Let $\mathfrak{A}$ be a multiplicative class on S and $\mathcal{D}$ a Dynkin class on S. If $\mathcal{D} \supset \mathfrak{A}$, then $\mathcal{D} \supset \sigma[\mathfrak{A}]$.

Proof. It is obvious that $\mathcal{D} \supset \delta[\mathfrak{A}] \supset \mathfrak{A}$. To prove that $\mathcal{D} \supset \sigma[\mathfrak{A}]$ it is enough to check that $\delta[\mathfrak{A}]$ is a σ-algebra. Since $\delta[\mathfrak{A}]$ is a Dynkin class, we need only prove that $\delta[\mathfrak{A}]$ is multiplicative. Using the assumption that $\mathfrak{A}$ is multiplicative, we can check that the class

$$\mathcal{D}_1 = \{B \subset S : A \cap B \in \delta[\mathfrak{A}] \text{ for every } A \in \mathfrak{A}\}$$

is a Dynkin class containing $\mathfrak{A}$. Hence $\mathcal{D}_1 \supset \delta[\mathfrak{A}]$, i.e.

$$A \in \mathfrak{A},\ B \in \delta[\mathfrak{A}] \Rightarrow A \cap B \in \delta[\mathfrak{A}].$$

This implies that the class

$$\mathcal{D}_2 = \{A \subset S : A \cap B \in \delta[\mathfrak{A}] \text{ for every } B \in \delta[\mathfrak{A}]\}$$

is also a Dynkin class containing $\mathfrak{A}$. Hence $\mathcal{D}_2 \supset \delta[\mathfrak{A}]$, i.e.

$$A, B \in \delta[\mathcal{A}] \Rightarrow A \cap B \in \delta[\mathcal{A}].$$

Theorem 2. Let $\mathcal{A}$ be a complementary class on S and $\mathcal{B}$ a class on S closed under countable disjoint unions and countable intersections. If $\mathcal{B} \supset \mathcal{A}$, then $\mathcal{B} \supset \sigma[\mathcal{A}]$.

Proof. Let $\mathcal{B}_1 := \{A : A, A^c \in \mathcal{B}\}$. Since $\mathcal{A}$ is complementary,

$$\mathcal{A} \subset \mathcal{B}_1 \subset \mathcal{B}.$$

To prove that $\mathcal{B} \supset \sigma[\mathcal{A}]$ it is enough to check that $\mathcal{B}_1$ is a σ-algebra. If $A_n \in \mathcal{B}_1$, $n = 1,2,\ldots$, then

$$\bigcup_n A_n = \sum_n A_1^c \cap A_2^c \cap \cdots \cap A_{n-1}^c \cap A_n \in \mathcal{B} \qquad (\textstyle\sum : \text{disjoint union})$$

and $(\bigcup_n A_n)^c = \bigcap_n A_n^c \in \mathcal{B}$,

so $\bigcup A_n \in \mathcal{B}_1$. Hence $\mathcal{B}_1$ is σ-additive. Since $\mathcal{B}_1$ is obviously complementary, $\mathcal{B}_1$ is a σ-algebra.

Let us define several operations deriving new σ-algebras from given ones.

Let $\{\mathcal{B}_\lambda\}_{\lambda \in \Lambda}$ be a family of σ-algebras on S. The set-theoretical intersection $\bigcap_\lambda \mathcal{B}_\lambda$ is a σ-algebra on S, but the set-theoretical union $\bigcup_\lambda \mathcal{B}_\lambda$ is not. The σ-algebra $\sigma[\bigcup_\lambda \mathcal{B}_\lambda]$ is called the lattice union of the family $\{\mathcal{B}_\lambda\}$, denoted by $\bigvee_\lambda \mathcal{B}_\lambda$.

Let $f : S_1 \to S_2$. If $\mathcal{B}_2$ is a σ-algebra on S_2, then the inverse image

$$f^{-1}(\mathcal{B}_2) = \{f^{-1}(B) : B \in \mathcal{B}_2\}$$

is a σ-algebra on S_1.

Let T be a subset of S. If $\mathcal{B}$ is a σ-algebra on S, the class

$$\mathcal{B} \cap T = \{B \cap T : B \in \mathcal{B}\}$$

is a σ-algebra on T, called the trace σ-algebra of $\mathcal{B}$ on T.

Let $\mathcal{B}_\lambda$ be a σ-algebra on S_λ for $\lambda \in \Lambda$ and let $S := \Pi_\lambda S_\lambda$. Then $\pi_\lambda^{-1}(\mathcal{B}_\lambda)$ is a σ-algebra on S for each λ, where π_λ denotes the canonical projection from the product space S to its λ-component space S_λ. The lattice union $\bigvee_\lambda \pi_\lambda^{-1}(\mathcal{B}_\lambda)$ is called the product σ-algebra of $\mathcal{B}_\lambda$, $\lambda \in \Lambda$, denoted by $\Pi_\lambda \mathcal{B}_\lambda$. Note that $\Pi_\lambda \mathcal{B}_\lambda$ is not the set-theoretical product of $\mathcal{B}_\lambda$, $\lambda \in \Lambda$. If $S_\lambda = T$ and $\mathcal{B}_\lambda = \mathcal{F}$ for every λ, then $\Pi_\lambda S_\lambda$ and $\Pi_\lambda \mathcal{B}_\lambda$ are denoted by T^Λ and $\mathcal{F}^\Lambda$ respectively.

Let $\mathcal{B}_i$ be a σ-algebra on S_i for $i = 1,2$. A map $f : S_1 \to S_2$ is called measurable $\mathcal{B}_1/\mathcal{B}_2$ if $f^{-1}(\mathcal{B}_2) \subset \mathcal{B}_1$. If $\mathfrak{A}_i$ generates $\mathcal{B}_i$ for $i = 1.2$, then "$f^{-1}(\mathfrak{A}_2) \subset \mathfrak{A}_1$" implies that f is measurable $\mathcal{B}_1/\mathcal{B}_2$. Measurability is transitive in the obvious sense. It is easy to see that if $f^{-1}(\mathfrak{A}_2) \subset \mathfrak{A}_1$, then f is measurable $\sigma[\mathfrak{A}_1]/\sigma[\mathfrak{A}_2]$.

If $f_\lambda : S \to S_\lambda$ is measurable $\mathcal{B}/\mathcal{B}_\lambda$ for $\lambda \in \Lambda$, then the product map

$$\Pi_\lambda f_\lambda : S \to \Pi_\lambda S_\lambda, \qquad x \mapsto (f_\lambda(x))$$

is measurable $\mathcal{B}/\Pi_\lambda\mathcal{B}_\lambda$; use $f_\alpha = \pi_\alpha \circ (\Pi_\lambda f_\lambda)$ to prove this.

If $f_\lambda : S_\lambda \to T_\lambda$ is measurable $\mathcal{B}_\lambda/\mathcal{F}_\lambda$ for $\lambda \in \Lambda$, then the <u>bilateral product map</u>

$$\Pi^b_\lambda f_\lambda : \Pi_\lambda S_\lambda \to \Pi_\lambda T_\lambda \ , \quad (x_\lambda) \mapsto (f_\lambda(x_\lambda))$$

is measurable $\Pi_\lambda\mathcal{B}_\lambda/\Pi_\lambda\mathcal{F}_\lambda$, because this map is equal to the map $\Pi_\lambda f_\lambda \circ \pi_\lambda$.

2. Borel spaces.

A set S endowed with a σ-algebra $\mathcal{S}$ on S is called a <u>Borel space</u> (or a <u>measurable space</u>), denoted by $(S,\mathcal{S})$. A subset B of a Borel space $S = (S,\mathcal{S})$ is called a <u>Borel subset</u> of S, if $S \in \mathcal{S}$. A map $f : (S,\mathcal{S}) \to (T,\mathcal{T})$ is called a <u>Borel map</u> or a <u>measurable map</u> if it is measurable $\mathcal{S}/\mathcal{T}$.

Whenever it is necessary, a subset T of a Borel space $(S,\mathcal{S})$ is regarded as a Borel space with the trace σ-algebra $\mathcal{S} \cap T$, called a <u>Borel subspace</u> of $(S,\mathcal{S})$. Similarly the product $\Pi_\lambda S_\lambda$ of Borel spaces $S_\lambda = (S_\lambda,\mathcal{S}_\lambda)$, $\lambda \in \Lambda$, is regarded as a Borel space with the product σ-algebra $\Pi_\lambda\mathcal{S}_\lambda$, called the <u>Borel product</u> of $(S_\lambda,\mathcal{S}_\lambda)$, $\lambda \in \Lambda$. Every canonical projection is a Borel map.

A map $f : (S,\mathcal{S}) \to (T,\mathcal{T})$ is called <u>bimeasurable</u> if f is bijective and if $f(\mathcal{S}) = \mathcal{T}$. If there exists a bimeasurable map from $(S,\mathcal{S})$ to $(T,\mathcal{T})$, then $(S,\mathcal{S})$ is called <u>Borel isomorphic</u> to $(T,\mathcal{T})$,

$$(S,\mathcal{S}) \underset{B}{\sim} (T,\mathcal{T})$$

in notation. The relation $\underset{B}{\sim}$ is an <u>equivalence relation</u>. It is easy to check that this relation is preserved by forming Borel products.

<u>Theorem 1</u>. Let $(S,\mathcal{S})$ and $(T,\mathcal{T})$ be Borel spaces and suppose that

$$S = \sum_n S_n, \quad S_n \in \mathcal{S} \ (n=1,2,\ldots) \quad \text{and} \quad T = \sum_n T_n, \quad T_n \in \mathcal{T} \quad (n=1,2,\ldots).$$

Then

$$S_n \underset{B}{\sim} T_n \quad (n=1,2,\ldots) \rightarrow S \underset{B}{\sim} T.$$

<u>Proof</u>. If $f_n : S_n \to T_n$ is bimeasurable for each n, then the map

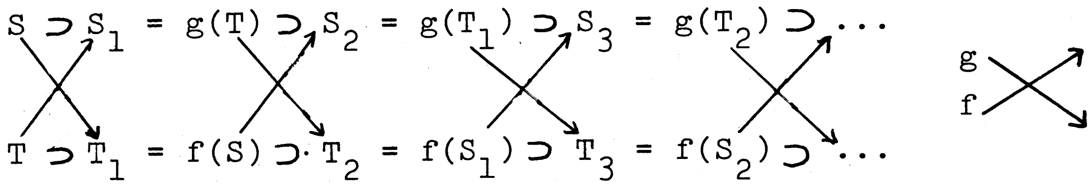
$$f : S \to T, \quad f(x) = f_n(x) \quad \text{for} \quad x \in S_n$$

is bimeasurable. ∎

<u>Theorem 2</u>. Let $S = (S,\mathcal{S})$ and $T = (T,\mathcal{T})$ be Borel spaces. Then

$$S \underset{B}{\sim} T_1 \in \mathcal{T} \quad \text{and} \quad T \underset{B}{\sim} S_1 \in \mathcal{S} \rightarrow S \underset{B}{\sim} T.$$

<u>Proof</u>. This is a Borel version of Benstein's theorem on equivalence of sets and can be proved by the same trick. Let $f : S \to T_1$ and $g : T \to S_1$ be bimeasurable. Define S_n and T_n for $n = 2,3,\ldots$ as follows.



Then

$$S = (S-S_1) + (S_1-S_2) + (S_2-S_3) + \dots + \bigcap_n S_n$$

$$T = (T-T_1) + (T_1-T_2) + (T_2-T_3) + \dots + \bigcap_n T_n .$$

Since the Borel spaces connected by lines are Borel isomorphic to each other, we can use Theorem 1 to conclude that $S \underset{B}{\simeq} T$.

Let S be a topological space. The σ-algebra generated by the open subsets of S is called the <u>topological σ-algebra</u> on S, denoted by $\mathcal{B}(S)$. If we want to clearly specify the topology τ on S, we use the notation (S,τ) and $\mathcal{B}(S,\tau)$ instead of S and $\mathcal{B}(S)$ respectively. Every topological space is regarded as a Borel space with the topological σ-algebra. Hence <u>Borel subsets</u>, <u>Borel maps</u> and <u>Borel isomorphisms</u> are defined for topological spaces. Open sets, closed sets, G_δ sets, F_σ sets, etc. are Borel subsets, continuous maps are Borel maps, and homeomorphic topological spaces are Borel isomorphic.

Let T be a subset of a topological space S. Then T is a topological space with the induced topology, so T is regarded as a Borel space $\mathcal{B}(T)$. Since S is a Borel space with $\mathcal{B}(S)$, T is also regarded as a Borel space with $\mathcal{B}(S) \cap T$. Since it is easy to check that

$$\mathcal{B}(T) = \mathcal{B}(S) \cap T,$$

it does not matter in which way T is regarded as a Borel space.

Let $S = \Pi_{\lambda\in\Lambda} S_\lambda$, where every S_λ is a topological space. Since S is a topological space with the product topology, it is regarded as a Borel space with $\mathcal{B}(S)$. Since every S_λ is a Borel space with $\mathcal{B}(S_\lambda)$, S is regarded as a Borel space with $\Pi_\lambda \mathcal{B}(S_\lambda)$. In general we have

$$\Pi_\lambda \mathcal{B}(S_\lambda) \subsetneq \mathcal{B}(\Pi_\lambda S_\lambda),$$

so we should clearly mention in which way we want to regard $\Pi_\lambda S_\lambda$ as a Borel space. But we have

Theorem 3. If S_n has a countable open base for $n = 1,2,\ldots,$ then $\Pi_n \mathcal{B}(S_n) = \mathcal{B}(\Pi_n S_n)$.

Proof. Since $\Pi_\lambda \mathcal{B}(S_\lambda) \subset \mathcal{B}(\Pi_\lambda S_\lambda)$ always holds, it is enough to prove the opposite inclusion relation. Let $\mathcal{U}_n$ be a countable open base in S_n. Then the class

$$\mathcal{U} := \{\cap_{i=1}^{n} \pi_i^{-1}(U_i) : n = 1,2,\ldots, \quad U_i \in \mathcal{U}_i\}$$

is a countable open base in the product space $S := \Pi_n S_n$. It is obvious that $\mathcal{U} \subset \mathcal{S} := \Pi_n \mathcal{B}(S_n)$. Every open subset of S belongs to $\sigma[\mathcal{U}]$, being a countable unions of sets in $\mathcal{U}$. Hence

$$\mathcal{B}(S) \subset \sigma[\mathcal{U}] \subset \mathcal{S}, \text{ i.e. } \mathcal{B}(\Pi_n S_n) \subset \Pi_n \mathcal{B}(S_n). \qquad \blacksquare$$

A topological space is called fully Lindelöf if every family of open subsets has a countable subfamily with the same union. Every topological space with a countable open base is fully Lindelöf. As a generalization of Theorem 3 we have

<u>Theorem 4</u>. If S_n is fully Lindelöf for every n, then $\Pi_n \mathcal{B}(S_n) = \mathcal{B}(\Pi_n S_n)$.

<u>Proof</u>. Essentially the same as above.

3. The analytic operation.

An indexed family of sets

$$A_{n_1 n_2 \ldots n_k} : \; k = 1,2,\ldots \; ; \; n_i = 1,2,\ldots \; (i=1,2,\ldots).$$

is called a <u>Souslin scheme</u>. With every Souslin scheme $\mathcal{S} = \{A_{n_1 n_2 \ldots n_k}\}$ we associate its <u>kernel</u>

$$K(\mathcal{S}) = \bigcup_{(n_i)} \bigcap_{k=1}^{\infty} A_{n_1 n_2 \ldots n_k},$$

where the union runs over all sequences $(n_i) \in \mathbb{N}^\infty$. The operation $\mathcal{S} \mapsto K(\mathcal{S})$ is called the <u>analytic operation</u>. Countable unions and countable intersections are special cases of the analytic operation, because

$$\bigcup_n B_n = K(\mathcal{S}) \qquad \text{for } \mathcal{S} = \{A_{n_1 n_2 \ldots n_k} := B_{n_1}\},$$

$$\bigcap_n B_n = K(\mathcal{S}) \qquad \text{for } \mathcal{S} = \{A_{n_1 n_2 \ldots n_k} := B_k\}.$$

Let $\mathfrak{A}$ be an arbitrary class of sets. The class of all sets obtained from sets in $\mathfrak{A}$ by the analytic operation is denoted by $\alpha[\mathfrak{A}]$. If $A \in \mathfrak{A}$, then

$$A = A \cup A \cup \ldots . \in \alpha[\mathfrak{A}].$$

Therefore

$$\mathfrak{A} \subset \alpha[\mathfrak{A}] \subset \alpha[\alpha[\mathfrak{A}]].$$

But we have

<u>Theorem 1</u>. $\mathfrak{A} \subset \alpha[\mathfrak{A}] = \alpha[\alpha[\mathfrak{A}]]$, if $\mathfrak{A}$ is multiplicative.

<u>Proof</u>. It is enough to prove that $\alpha[\alpha[\mathfrak{A}]] \subset \alpha[\mathfrak{A}]$, i.e. that

$$K(\mathcal{B}) \in \alpha[\mathfrak{A}] \quad \text{for every } \mathcal{B} = \{B_{n_1 n_2 \ldots n_k}\} \subset \alpha[\mathfrak{A}].$$

Let

$$B_{n_1 n_2 \ldots n_k} := \bigcup_{(m_j)} \bigcap_{r=1}^{\infty} A^{n_1 n_2 \ldots n_k}_{m_1 m_2 \ldots m_r}, \quad A^{\ldots}_{\ldots} \in \mathfrak{A}.$$

Then

$$K(\mathcal{B}) = \bigcup_{(n_i)} \bigcap_{k=1}^{\infty} \bigcup_{(m_j)} \bigcap_{r=1}^{\infty} A^{n_1 n_2 \ldots n_k}_{m_1 m_2 \ldots m_r}.$$

Using the general distributive law we can exchange $\bigcap_k$ and $\bigcup_{(m_j)}$ to obtain

$$K(\mathcal{B}) = \bigcup_{(n_i)} \underbrace{\bigcup_{(m_j^1),(m_j^2),\ldots}} \bigcap_{k=1}^{\infty} \bigcap_{r=1}^{\infty} A^{n_1 n_2 \ldots n_k}_{m_1^k m_2^k \ldots m_r^k}$$

$$= \underbrace{\bigcup_{(n_i),(m_j^1),(m_j^2)\ldots}} \bigcap_{p=1}^{\infty} \bigcap_{k=1}^{p} A^{n_1 n_2 \ldots n_k}_{m_1^k m_2^k \ldots m_{p+1-k}^k}$$

where the union runs over all indices n_i and m_k^j $(i,j,k = 1,2,\cdots)$. These indices can be arranged in a triangular array :

$$\begin{array}{lllll} n_1 & n_2 & n_3 & \cdots\cdots & n_p \;\cdots \\ m_1^1 & m_2^1 & m_3^1 & \cdots\cdots & m_p^1 \;\cdots \\ & m_1^2 & m_2^2 & \cdots\cdots & m_{p-1}^2 \cdots \\ & & m_1^3 & \cdots\cdots & m_{p-2}^3 \cdots \\ & & & & \cdots \\ & & & & m_1^p \;\cdots \\ & & & & \quad\;\cdots \end{array}$$

Since n_i and m_k^i move freely on $\mathbb{N}$, the p-th row of the array moves freely on $\mathbb{N}^{p+1}$ for every p. Since $\mathbb{N}^{p+1}$ is a countable infinite set, the p-th row can be indexed by a natural number ν_p. Since the last intersection $\bigcap_{k=1}^{p} A_{\cdots}^{\cdots}$ in the above expression of $K(\mathcal{S})$ depends only on the indices appearing in the first p rows of the array, it can be denoted by $D_{\nu_1\nu_2\cdots\nu_p}$. Hence we obtain

$$K(\mathcal{S}) = \bigcup_{(\nu_i)} \bigcap_{p=1}^{\infty} D_{\nu_1\nu_2\cdots\nu_p}.$$

But $D_{\nu_1\nu_2\cdots\nu_p} \in \mathfrak{A}$, because $\mathfrak{A}$ is multiplicative. Therefore $K(\mathcal{S})$ belongs to $\alpha[\mathcal{S}]$. ∎

A souslin scheme $\mathcal{S} = \{A_{n_1n_2\cdots n_k}\}$ is called <u>decreasing</u> if

$$A_{n_1} \supset A_{n_1n_2} \supset A_{n_1n_2n_3} \supset \cdots \quad \text{for every sequence } (n_i).$$

$\mathcal{S}$ is called disjoint if $A_{n_1n_2\cdots n_kn}$, $n = 1,2,\ldots,$ are disjoint for every k and every $(n_1,n_2,\ldots,n_k)$.

Since a non-countable union is involved in the analytic operation, $K(\mathscr{S}) \notin \sigma[\mathscr{S}]$ in general. But we have

Theorem 2.

(i) If $\mathscr{S}$ is decreasing and disjoint, then

$$K(\mathscr{S}) = \bigcap_{k=1}^{\infty} \bigcup_{(n_1,n_2,\ldots,n_k)} A_{n_1 n_2 \ldots n_k}$$

(ii) If $\mathscr{S}$ is disjoint, then

$$K(\mathscr{S}) \in \sigma[\mathscr{S}].$$

Proof.

(i) Using the general distributive law in set theory, we can express the right hand side R as follows:

$$R = \bigcup_{n_1^1, n_1^2, n_2^2, n_1^3, n_2^3, n_3^3, \ldots} A_{n_1^1} \cap A_{n_1^2 n_2^2} \cap A_{n_1^3 n_2^3 n_3^3} \cap \cdots,$$

where all indices move freely on $\mathbb{N}$. Since $\mathscr{S}$ is decreasing and disjoint, these countable intersections are empty unless

$$n_1^1 = n_1^2 = n_1^3 = \ldots \ (= n_1), \quad n_2^2 = n_2^3 = n_2^4 = \ldots \ (= n_2), \ldots .$$

Hence

$$R = \bigcup_{n_1, n_2, \ldots} A_{n_1} \cap A_{n_1 n_2} \cap A_{n_1 n_2 n_3} \cap \cdots = K(\mathscr{S}).$$

(ii) The Souslin scheme $\mathscr{S}' := \{A'_{n_1 n_2 \ldots n_k} := \bigcap_{i=1}^{k} A_{n_1 n_2 \ldots n_i}\}$

is decreasing and disjoint. Hence

$$K(\mathscr{S}') \in \sigma[\mathscr{S}'] \quad \text{by (i).}$$

By the definition of the analytic operation we have

$$K(\mathscr{S}) = K(\mathscr{S}') \in \sigma[\mathscr{S}'] \subset \sigma[\mathscr{S}].$$

┘

4. Polish spaces, standard spaces and analytic spaces.

Throughout this section a Hausdorff topological space is simply called a space. If the topology τ on S is given by a metric ρ, then the induced topology $\tau|_T$ on a subset T of S is given by the induced metric $\rho|_T$.

A metric ρ on a set S is called Polish if the metric space (S,ρ) is separable and complete. A Polish space is defined to be a space whose topology can be given by a Polish metric. A space is Polish if and only if it is homeomorphic to a complete separable metric space.

A space S is called standard (resp. analytic) if it is a continuous bijective (resp. surjective) image of a Polish space, i.e. if we can find a Polish space P and a continuous bijection (resp. surjection) $f:P \to S$. It is obvious that

$$\text{Polish} \Longrightarrow \text{standard} \Longrightarrow \text{analytic}.$$

A metrizable standard (resp. analytic) space is called a Lusin space (resp. Souslin space).

Theorem 1. Every analytic space is fully Lindelöf.

Proof. Let S be analytic. Then we can find a Polish space P and a continuous surjection $f:P \to S$. Let $\{G_\lambda\}_{\lambda\in\Lambda}$ be a family of open subset of S. Then $f^{-1}(G_\lambda)$ is also open for every λ. Since P has a countable open base, we can find a sequence $\{\lambda_n\} \subset \Lambda$ such that

$$\bigcup_n f^{-1}(G_{\lambda_n}) = \bigcup_\lambda f^{-1}(G_\lambda).$$

Since f is surjective, this implies that

$$\bigcup_n G_{\lambda_n} = \bigcup_\lambda G_\lambda .$$ ∎

Since the continuous bijective (or surjective) maps are closed under compositions, we have

<u>Theorem 2</u>. Every continuous bijective image of a standard space is standard and every continuous surjective image of an analytic space is analytic.

Since the identity map from a space to the same space with a weaker topology is a continuous bijection, we have

<u>Theorem 3</u>. The property of being standard (or analytic) is preserved by weakening the topology.

<u>Theorem 4</u>. The property of being Polish (or standard or analytic) is preserved by forming countable topological products.

<u>Proof</u>. Let $\{P_n\}$ be a sequence of Polish spaces and ρ_n a Polish metric defining the topology on P_n for each n. Then the product topology on $P := \Pi_n P_n$ is given by

$$\rho((x_n),(y_n)) := \sum_n 2^{-n}[\rho_n(x_n,y_n) \wedge 1],$$

It is easy to check that ρ is Polish. Hence P is Polish. This proves the assertion for Polish spaces.

Let S_n be standard for $n = 1,2,\dots$. Take a Polish space P_n and a continuous bijection $f_n : P_n \to S_n$ for each n. The bilateral product map $f := \Pi_n^b f_n$ is a continuous bijection from $P := \Pi_n P_n$ to $S := \Pi_n S_n$. Since P is Polish, S is standard. This proves the assertion for standard spaces. The same argument works for analytic spaces by using surjections instead of bijections. ∎

<u>Theorem 5</u>. If S_n, $n = 1,2,\ldots$, are analytic, then

$$\mathcal{B}(\Pi_n S_n) = \Pi_n \mathcal{B}(S_n).$$

<u>Proof</u>. $\Pi_n S_n$ is analytic (Theorem 4), so it is fully Lindelöf (Theorem 1). Now use Theorem 2.4. ∎

Let $\{S_\lambda, \lambda\in\Lambda\}$ be a family of spaces. For each $\lambda\in\Lambda$ we topologize the space $S_\lambda' := \{(x,\lambda): x\in S_\lambda\}$ so that the map $x \mapsto (x,\lambda)$ is bicontinuous. Next we topologize

$$S' := \sum_\lambda S_\lambda'$$

so that G is open in S' if and only if $G\cap S_\lambda'$ is open in S_λ' for every λ. The space S' thus defined is called the <u>topological sum</u> of $\{S_\lambda\}$, denoted by $\oplus_\lambda S_\lambda$.

If $f_\lambda: S_\lambda \to T_\lambda$ is continuous (resp. bijective, surjective), then the <u>sum map</u>

$$\bigoplus_\lambda f_\lambda : \oplus S_\lambda \to \oplus T_\lambda, \quad (x,\lambda) \mapsto (f_\lambda(x),\lambda) \quad \text{for } x\in S_\lambda$$

is continuous (resp. bijective, surjective).

Theorem 6. The property of being Polish (or standard or analytic) is preserved by forming countable topological sums.

Proof. If we prove the assertion for Polish spaces, then the argument in the proof of Theorem 4 will work for standard or analytic spaces by using sum maps instead of bilateral product maps.

Let P_n, $n = 1,2,\dots$ be Polish. Take a Polish metric ρ_n defining the topology on P_n for each n. Then the topology on $\oplus_n P_n$ is given by the metric

$$\rho((x,m),\ (y,n)) := \begin{cases} \rho_n(x,y)\wedge 1 & \text{if } m = n \\ 1 & \text{if } m \neq n \end{cases}$$

It is easy to check that ρ is Polish. Hence $\oplus_n P_n$ is Polish. This proves the assertion for Polish spaces.

Theorem 7. Every compact metrizable space is Polish.

Proof. Immediate from the fact that every metric defining the topology of a compact metrizable space is Polish.

Theorem 8. Every separable Banach space is Polish with respect to the norm topology and standard with respect to every Hausdorff topology weaker than the norm topoloty (for example, the week topology).

Proof. Immediate from the definition and Theorem 3.

5. Standard subsets and analytic subsets.

Let S be a space, i.e. a Hausdorff topological space. A subset A of S is called Polish if the set A endowed with the induced topology is a Polish space. Similarly we define standard subsets and analytic subsets.

We denote the class of all standard (resp. analytic, closed, open) subsets of S by $\mathcal{S}(S)$ (resp. $\mathcal{A}(S), \mathcal{F}(S), \mathcal{G}(S)$). It is obvious that

$$\mathcal{B}(S) = \sigma[\mathcal{G}(S)] = \sigma[\mathcal{F}(S)] = \sigma[\mathcal{G}(S) \cup \mathcal{F}(S)].$$

Since every restriction of a continuous map is continuous, we have

$$f(\mathcal{S}(S)) \subset \mathcal{S}(T) \quad \text{if} \quad f:S \to T \quad \text{is a continuous injection,}$$

and
$$f(\mathcal{A}(S)) \subset \mathcal{A}(T) \quad \text{if} \quad f:S \to T \quad \text{is continuous.}$$

If $f:P \to A (\subset S)$ is continuous, then $i_{A,S} \circ f:P \to S$ is also continuous, where $i_{A,S}$ is the canonical injection from A into S. Hence

$A \in \mathcal{A}(S)$ if and only if we can find a Polish space P (or equivalently a complete separable metric space (P,ρ)) and a continuous map $f:P \to S$ with $f(P) = A$. We can characterize $A \in \mathcal{S}(S)$ similarly.

If S is standard (or analytic), we can prove very simple relations among the classes $\mathcal{F}(S), \mathcal{G}(S), \mathcal{B}(S), \mathcal{S}(S)$ and $\mathcal{A}(S)$ (Theorem 7).

Lemma 1. Every closed or open subset of a Polish (resp. standard, analytic) space is Polish (resp. standard, analytic).

Proof. Let P be Polish and ρ a Polish metric defining the topology on P. If F is closed in P, then the induced metric $\rho|_F$ is also a Polish metric defining the induced topology on F. Hence F is Polish. Let G be open in P. If $G = P$, then G is Polish trivially. If $G \neq P$, then the function $f(x) := \rho(x, P-G)$ is continuous and $f(x) > 0$ if and only if $x \in G$. It is easy to check that

$$\rho_G(x,y) := \rho(x,y) + \left|\frac{1}{f(x)} - \frac{1}{f(y)}\right|, \quad x,y \in G$$

is a Polish metric defining the topology on G. Hence G is Polish.

Let S be standard and B a closed or open subset of S. Take a Polish space P and a continuous bijection $f: P \to S$. Then the inverse image $A := f^{-1}(B)$ is closed or open in P, so A is Polish. Since $f(A) = B$, B is standard. Similarly we can prove that every closed or open subset of an analytic space is analytic. ∎

Let $\{S_\lambda, \lambda \in \Lambda\}$ be a family of subspaces of a space S. The set

$$D := \{x \in \Pi_\lambda S_\lambda : \pi_\lambda(x) \text{ is independent of } \lambda\}$$

is called the <u>diagonal set</u> of $\Pi_\lambda S_\lambda$, denoted by $D(\Pi_\lambda S_\lambda)$. It is not hard to prove that D is closed in $\Pi_\lambda S_\lambda$ and homeomorphic to the intersection $\bigcap_\lambda S_\lambda$ ($\subset S$).

<u>Theorem 1</u>. $\mathcal{S}(S)$ is closed under countable disjoint unions and countable intersections.

<u>Remark</u>. Later we will prove that $\mathcal{S}(S)$ is closed under arbitrary countable unions (Theorem 8).

<u>Proof</u>. Suppose that $A_n \in \mathcal{S}(S)$, $n = 1,2,\ldots$. We will prove that

$$A := \sum_n A_n \in \mathcal{S}(S), \quad \text{and} \quad B := \bigcap_n A_n \in \mathcal{S}(S).$$

Let

$$A' := \bigoplus_n A_n$$

and consider the map

$$f: A' \to A, \ (x,n) \mapsto x \quad \text{for} \quad x \in A_n.$$

This map is a continuous bijection. Since A' is Polish (Theorem 4.6), A is standard, i.e. $A \in \mathcal{S}(S)$. Let D be the diagonal set of $\Pi_n A_n$. Then D is homeomorphic to B. Since $\Pi_n A_n$ is standard (Theorem 4.4) and since D is closed in $\Pi_n A_n$, D is standard (Lemma 1). Hence B is also standard. ∎

<u>Theorem 2</u>. $\mathcal{A}(S)$ is closed under the analytic operation.

<u>Proof</u>. By the argument used above we can prove that $\mathcal{A}(S)$ is closed under countable unions and countable intersections; note that if $A = \bigcup_n A_n$, $A_n \in \mathcal{A}(S)$, then the map $f: A' \to A$ used above is a continuous <u>surjection</u>.

Let $\mathfrak{A} = \{A_{n_1 n_2 \ldots n_k}\} \subset \mathcal{A}(S)$ be a Souslin scheme. We want to prove that

$$A := K(\mathcal{A}) \in \mathcal{A}(S).$$

Without loss of generality we can assume that $\mathcal{A}$ is decreasing, because

$$A'_{n_1n_2\dots n_k} = \bigcap_{i=1}^{k} A_{n_1n_2\dots n_i} \in \mathcal{A}(S)$$

and the Souslin scheme $\{A'_{n_1n_2\dots n_k}\}$ has the same kernel as the original scheme $\mathcal{A}$. Since $\mathbb{N}$ is Polish, $\mathbb{N}^\infty$ is Polish. Let $\mathcal{N}$ and $\mathcal{B}$ denote the Souslin schemes composed of

$$N_{n_1n_2\dots n_k} := \{\xi \in \mathbb{N}^\infty : \pi_i(\xi) = n_i,\ i = 1,2,\dots,k\} \subset \mathbb{N}^\infty$$

and $$B_{n_1n_2\dots n_k} := N_{n_1n_2\dots n_k} \times A_{n_1n_2\dots n_k} \subset \mathbb{N}^\infty \times S$$

respectively. Being closed in $\mathbb{N}^\infty$, every $N_{n_1n_2\dots n_k}$ is Polish (Lemma 1). Since $\mathcal{A} \subset \mathcal{A}(S)$, we have

$$B_{n_1n_2\dots n_k} \in \mathcal{A}(\mathbb{N}^\infty \times S) \quad \text{(Theorem 4.4)}.$$

Since $\mathcal{N}$ is decreasing and disjoint and since $\mathcal{A}$ is decreasing, $\mathcal{B}$ is decreasing and disjoint. Hence we can use Theorem 3.2 (i) to obtain

$$K(\mathcal{B}) = \bigcap_{k} \bigcup_{(n_1,n_2,\dots,n_k)} B_{n_1n_2\dots n_k}$$

Since $\mathcal{A}(\mathbb{N}^\infty \times S)$ is closed under countable unions and countable intersections, we have

$K(\mathcal{B}) \in \mathcal{A}(\mathbb{N}^\infty \times S)$.

From the definition of $K(\mathcal{B})$ we obtain

$$K(\mathcal{B}) = \bigcup_{(n_i)} \bigcap_{k=1}^{\infty} B_{n_1 n_2 \ldots n_k}$$

$$= \bigcup_{(n_i)} \{(n_i)\} \times \bigcap_{k=1}^{\infty} A_{n_1 n_2 \ldots n_k},$$

Let π_2 be the canonical projection from $\mathbb{N}^\infty \times S$ to S. Then

$$\pi_2(K(\mathcal{B})) = \bigcup_{(n_i)} \bigcap_{k=1}^{\infty} A_{n_1 n_2 \ldots n_k} = K(\mathcal{O}).$$

Since π_2 is continuous and since $K(\mathcal{B}) \in \mathcal{A}(\mathbb{N}^\infty \times S)$, we obtain $K(\mathcal{O}) \in \mathcal{A}(S)$. ┛

Let $\{A_n\}$ be a sequence of subsets of S. We say that $\{A_n\}$ <u>monotically converges</u> to a point $a \in S$ if

(i) $A_n \ni a$ for every n,

(ii) $A_1 \supset A_2 \supset \ldots$,

and

(iii) for every neighborhood $U(a)$ we can find an index n_0 such that $A_{n_0} \subset U(a)$, (i.e. $A_n \subset U(a)$ for every $n \geq n_0$ by (ii)).

<u>Lemma 2</u>. $A_n \downarrow a \Rightarrow \bigcap_n A_n = \bigcap_n \overline{A_n} = \{a\}$.

<u>Proof</u>. Let b be any point of S distinct from a. Then we can find disjoint neighborhoods $U(a)$ and $V(b)$. Take an index r such that $A_r \subset U(a)$. Then A_r and $V(b)$ are disjoint. Hence b does not belong to $\overline{A_r}$. This implies that

$$\{a\} \subset \bigcap_n A_n \subset \bigcap_n \bar{A}_n \subset \{a\},$$

so all these sets are the same. ∎

Lemma 3. If $f: S \to T$ is continuous, then

$$A_n \downarrow a \Rightarrow f(A_n) \downarrow f(a) \Rightarrow \{f(a)\} = \bigcap_n f(A_n) = \bigcap_n \overline{f(A_n)}.$$

Proof. Let V be any neighborhood of $f(a)$. Then $f^{-1}(V)$ is a neighborhood of a. Hence we can find an index r such that $A_r \subset f^{-1}(V)$. Then $f(A_r) \subset V$. It is obvious that

$$f(A_1) \supset f(A_2) \supset \cdots \quad \text{and} \quad f(A_n) \ni f(a) \ (n=1,2,\cdots).$$

Hence $f(A_n) \downarrow f(a)$. This proves the first implication. The second implication follows from the last lemma. ∎

Theorem 3. If S is analytic, then we can find a decreasing Souslin scheme $\mathcal{S} = \{S_{n_1 n_2 \cdots n_k}\} \subset \mathcal{A}(S)$ ($\mathcal{S} \subset \mathcal{F}(S)$ in case S is Polish) satisfying the following conditions.

(i) . $S = \bigcup_n S_n$ and $S_{n_1 n_2 \cdots n_k} = \bigcup_n S_{n_1 n_2 \cdots n_k n}$.

(ii) For every sequence $(n_i) \in \mathbb{N}^\infty$ the sequence $\{S_{n_1 n_2 \cdots n_k}\}_{k=1,2,\cdots}$ monotonically converges to a point in S.

Proof. First we consider the case where S is Polish. Let ρ be a Polish metric defining the topology on S. Take a sequence $\{x_n\}$ dense in S and let

$$S_n := \bar{U}(x_n, 2^{-1}), \ n=1,2,\cdots$$

where $\bar{U}(x, r)$ denotes the closed ball with center x and

ρ-radius r. Then S_n is a non-empty closed set and

$$S = \bigcup_n S_n.$$

Suppose that the non-empty closed sets

$$S_{n_1 n_2 \dots n_k},\ (n_1, n_2, \dots, n_k) \in \mathbb{N}^k$$

are defined. Take a sequence $\{y_n = y_n(n_1 n_2 \dots n_k)\}_{n=1,2,\dots}$ dense in $S_{n_1 n_2 \dots n_k}$ and let

$$S_{n_1 n_2 \dots n_k n} := S_{n_1 n_2 \dots n_k} \cap \bar{U}(y_n, 2^{-k-1}).$$

Then $S_{n_1 n_2 \dots n_k n}$ is a non-empty closed set. Thus we obtain a Souslin scheme $\mathcal{S} = \{S_{n_1 n_2 \dots n_k}\} \subset \mathcal{F}(S)$. We will verify (i) and (ii). (i) is obvious by the construction, (ii) follows from the Cantor intersection theorem. This proves our theorem for S Polish.

Let S be analytic. Take a Polish space P and a continuous surjection $f: P \to S$. Take a Souslin scheme $\mathcal{P} = \{P_{n_1 n_2 \dots n_k}\} \subset \mathcal{F}(P)$ satisfying (i) and (ii). Let

$$S_{n_1 n_2 \dots n_k} := f(P_{n_1 n_2 \dots n_k}).$$

Being closed in P, $P_{n_1 n_2 \dots n_k}$ is Polish, so $S_{n_1 n_2 \dots n_k}$ is analytic. Then $\mathcal{S} = \{S_{n_1 n_2 \dots n_k}\}$ is a Souslin scheme to be constructed, because (i) is obvious and (ii) follows from Lemma 3. ∎

<u>Theorem 4</u>. $\mathcal{A}(S) \subset \alpha[\mathcal{F}(S)]$

<u>Proof</u>. Let $A \in \mathcal{A}(S)$. Applying the last theorem to A we can

find a Souslin scheme $\mathcal{O}\mathcal{L} = \{A_{n_1 n_2 \ldots n_k}\}$ satisfying Conditions (i) and (ii). Let $\overline{\mathcal{O}\mathcal{L}} := \{\overline{A}_{n_1 n_2 \ldots n_k}\}$, where the bar means the closure in S (<u>not</u> in A). Then it is easy to see that

$$A = K(\mathcal{O}\mathcal{L}) \subset K(\overline{\mathcal{O}\mathcal{L}}).$$

For every $(n_i) \in \mathbb{N}^{\infty}$ we have

$$A_{n_1 n_2 \ldots n_k} \downarrow a \quad \text{in} \quad A \text{ (so in } S) \text{ for some } a \in A.$$

Hence Lamma 2 ensures that

$$\bigcap_k \overline{A}_{n_1 n_2 \ldots n_k} = \{a\} \subset A,$$

proving $K(\overline{\mathcal{O}\mathcal{L}}) \subset A$. Thus we have

$$A = K(\overline{\mathcal{O}\mathcal{L}}) \in \mathcal{A}[\mathcal{F}(S)]. \qquad ∎$$

If we can find a <u>disjoint</u> family $\{B_n\}$ such that

$$A_n \subset B_n \in \mathcal{B}(S) \quad \text{for every} \quad n,$$

then we say that $\{A_n\}$ is <u>Borel separated</u>.

<u>Lemma 4</u>. If $\{A_m, A'_n\}$ is Borel separated for each (m,n), then $\{\bigcup_m A_m, \bigcup_n A'_n\}$ is Borel separated.

<u>Proof</u>. For each (m,n) we can find $B_{mn}, B'_{mn} \in \mathcal{B}(S)$ such that

$$A_m \subset B_{mn}, \; A'_n \subset B'_{mn} \quad \text{and} \quad B_{mn} \cap B'_{mn} = \phi.$$

Then

$$\bigcup_m A_m \subset \bigcap_n \bigcup_m B_{mn} \in \mathcal{B}(S) \quad \text{and} \quad \bigcup_n A'_n \subset \bigcup_n \bigcap_m B'_{mn} \in \mathcal{B}(S).$$

In general, if $C_n \cap C'_n = \phi$ for every n, then

$$(\bigcup_n C_n) \cap (\bigcap_n C'_n) = \phi.$$

Using this fact twice, we can check that

$$(\bigcap_n \bigcup_m B_{mn}) \cap (\bigcup_n \bigcap_m B'_{mn}) = \phi.$$

Hence $\{ \bigcup A_m, \bigcup A'_n \}$ is Borel separated. ∎

Lemma 5. If $\{A_m, A_n\}$ is Borel separated for every (m,n) $(m \neq n)$, then $\{A_1, A_2, \ldots\}$ is Borel separated.

Proof. We can use Lemma 4 to find disjoint Borel sets

$$B_n \supset A_n \quad \text{and} \quad B'_n \supset \bigcup_{k>n} A_k$$

for each n. Then

$$B''_n := B'_1 \cap B'_2 \cap \ldots \cap B'_{n-1} \cap B_n \ (\supset A_n), \ n = 1,2,\ldots$$

are disjoint Borel sets. Hence the family $\{ A_n \}$ is Borel separated. ∎

If $\{ A_n \}$ is Borel separated, it is obviously disjoint. The converse is not true in general, but we have

Theorem 5 (The Borel sepatation theorem of Lusin). If $\{ A_n \}$ is a countable disjoint subclass of $\mathcal{A}(S)$, then it is Borel separated.

Proof. By virtue of the last lemma it is enough to prove that if A and B are disjoint analytic subsets of S, then $\{ A,B \}$ is Borel separated. Supposing that $\{A,B\}$ is not Borel separated, we will deduce a contradition. Applying Theorem 3 to the analytic spaces A and B we construct two Souslin schemes

$$\mathfrak{A} = \{A_{m_1 m_2 \ldots m_k}\} \subset \mathcal{A}(A) \quad \text{and} \quad \mathfrak{B} = \{B_{n_1 n_2 \ldots n_k}\} \subset \mathcal{A}(B).$$

Since $\{A, B\}$ is supposed to be not Borel separated and since $A = \bigcup_m A_m$ and $B = \bigcup_n B_n$, we can use Lemma 4 to find a pair $\{A_{m_1}, B_{n_1}\}$ which is not Borel separated. Since

$$A_{m_1} = \bigcup_m A_{m_1 m} \quad \text{and} \quad B_{n_1} = \bigcup_n B_{n_1 n},$$

we can again use Lemma 4 to find a pair $\{A_{m_1 m_2}, B_{n_1 n_2}\}$ which is not Borel separated. Continuing this procedure, we can find two sequences (m_i) and (n_i) such that $\{A_{m_1 m_2 \ldots m_k}, B_{n_1 n_2 \ldots n_k}\}$ is not Borel separated for each k. From Condition (ii) of Theorem 3 we have

$$A_{m_1 m_2 \ldots m_k} \downarrow a \in A \quad \text{and} \quad B_{n_1 n_2 \ldots n_k} \downarrow b \in B.$$

Since $A \cap B = \phi$, a must be distinct from b. Hence we can find two disjoint open sets

$$U(\ni a) \quad \text{and} \quad V(\ni b).$$

Then we have

$$A_{m_1 m_2 \ldots m_k} \subset U \quad \text{and} \quad B_{n_1 n_2 \ldots n_k} \subset V \quad \text{for some} \quad k = k_0.$$

This implies that $\{A_{m_1 m_2 \ldots m_k}, B_{n_1 n_2 \ldots n_k}\}$ is Borel separated, which is a contradiction. ■

Lemma 6. Every Borel subset of a Polish space is standard.

Proof. $\mathcal{S}(P)$ is closed under countable disjoint unions and countable intersections (Theorem 1) and

$$\mathcal{S}(P) \supset \mathcal{O} := \mathcal{G}(P) \cup \mathcal{F}(P) \quad \text{(Lemma 1)}.$$

Hence we can use Theorem 1.2 to conclude that

$$\mathcal{S}(P) \supset \sigma[\mathcal{O}] = \mathcal{B}(P).$$ ∎

Theorem 6. $\mathcal{S}(S) \subset \mathcal{B}(S)$

Proof. Let $A \in \mathcal{S}(S)$. Then we can find a Polish space P and a continuous bijection $f: P \to A$. Applying Theorem 3 to the Polish space P, we construct a Souslin scheme

$$\mathcal{P} = \{P_{n_1 n_2 \ldots n_k}\} \subset \mathcal{F}(P).$$

Let

$$Q_{n_1} := P_{n_1} - \bigcup_{n<n_1} P_n,$$

$$Q_{n_1 n_2} := P_{n_1 n_2} - \bigcup_{n<n_2} P_{n_1 n},$$

$$Q_{n_1 n_2 n_3} := P_{n_1 n_2 n_3} - \bigcup_{n<n_3} P_{n_1 n_2 n},$$

and so on. Then

$$Q_{n_1 n_2 \ldots n_k} \in \mathcal{B}(P) \subset \mathcal{S}(P) \text{ (Lemma 6)},$$

$$P = \sum_n P_n \quad \text{and} \quad Q_{n_1 n_2 \ldots n_k} \subset P_{n_1 n_2 \ldots n_k} = \sum_n Q_{n_1 n_2 \ldots n_k n}.$$

Hence every point $p \in P$ belongs to a unique intersection

$\bigcap_k Q_{n_1 n_2 \ldots n_k}$ and the Souslin scheme $Q = \{Q_{n_1 n_2 \ldots n_k}\}$ is disjoint. Since f is bijective, the Souslin scheme

$$\mathcal{A} := \{A_{n_1 n_2 \ldots n_k} := f(Q_{n_1 n_2 \ldots n_k})\}$$

is disjoint. Since $Q_{n_1 n_2 \ldots n_k} \in \mathcal{S}(P)$, $A_{n_1 n_2 \ldots n_k} \in \mathcal{S}(S)$. Also $\{A_{n_1 n_2 \ldots n_{k-1} n}\}_{n=1,2,\ldots}$ is disjoint. Hence Theorem 5 ensures the existence of disjoint Borel sets

$$B_{n_1 n_2 \ldots n_{k-1} n} \ (\supset A_{n_1 n_2 \ldots n_{k-1} n}), \ n = 1,2,\ldots .$$

Thus we obtain a disjoint Souslin scheme $\mathcal{B} = \{B_{n_1 n_2 \ldots n_k}\} (\subset \mathcal{B}(S))$. Then the Souslin scheme

$$\mathcal{A}' = \{A'_{n_1 n_2 \ldots n_k} := \bar{A}_{n_1 n_2 \ldots n_k} \cap B_{n_1 n_2 \ldots n_k}\} (\subset \mathcal{B}(S))$$

is also disjoint, the closure being taken in S (not in A). Hence Theorem 3.2 ensures that

$$K(\mathcal{A}') \in \sigma[\mathcal{B}(S)] = \mathcal{B}(S).$$

We will prove that

$$A = K(\mathcal{A}'),$$

which will complete the proof of our theorem.

Let $p \in P$. Then

$$p \in \bigcap_k Q_{n_1 n_2 \ldots n_k} \quad \text{for some } (n_i).$$

Hence we have

$$f(p) \in \bigcap_k f(Q_{n_1 n_2 \ldots n_k}) = \bigcap_k A_{n_1 n_2 \ldots n_k} \subset \bigcap_k A'_{n_1 n_2 \ldots n_k} \subset K(\mathcal{A}'),$$

proving that $A \subset K(\mathfrak{A}')$.

Let $a \in K(\mathfrak{A}')$. Then

$$a \in \bigcap_k A'_{n_1 n_2 \ldots n_k} \quad \text{for some} \quad (n_i).$$

Since $\{P_{n_1 n_2 \ldots n_k}\}_{k=1,2,\ldots}$ monotonically converges to a point $p \in P$ and f is regarded as a continuous map from P into S, Lemma 3 ensures that

$$\{f(p)\} = \bigcap_k \overline{f(P_{n_1 n_2 \ldots n_k})} \supset \bigcap_k \overline{f(Q_{n_1 n_2 \ldots n_k})}$$

$$= \bigcap_k \bar{A}_{n_1 n_2 \ldots n_k} \supset \bigcap_k A'_{n_1 n_2 \ldots n_k} \ni a.$$

Hence we obtain $a = f(p)$, proving that $K(\mathfrak{A}') \subset A$. ∎

Theorem 7.

(i) If S is standard, then $\mathcal{B}(S) = \mathcal{S}(S)$

(ii) If S is analytic, then

$$\mathcal{B}(S) = \{A : A, A^c \in \mathcal{A}(S)\} \subset \mathcal{A}(S) = \alpha[\mathcal{F}(S)] = \alpha[\mathcal{B}(S)]$$

$$= \alpha[\mathcal{A}(S)]$$

Proof.

(i) Since $\mathcal{S}(S) \subset \mathcal{B}(S)$ (Theorem 6), it is enough to prove that

$$\mathcal{B}(S) \subset \mathcal{S}(S).$$

Since S is standard, we can find a Polish space P and a continuous bijection $f: P \to S$. Let $B \in \mathcal{B}(S)$. Being continuous, f is Borel. Hence we have

$$A := f^{-1}(B) \in \mathcal{B}(P) \subset \mathcal{S}(P) \text{ (Lemma 6)},$$

which implies that $B = f(A) \in \mathcal{S}(S)$. This proves that $\mathcal{B}(S) \subset \mathcal{S}(S)$.

(ii) By the same argument as above we can check that

$$\mathcal{B}(S) \subset \mathcal{A}(S).$$

Hence

$$A \in \mathcal{B}(S) \Longrightarrow A,\ A^c \in \mathcal{B}(S) \Longrightarrow A,\ A^c \in \mathcal{A}(S).$$

If $A, A^c \in \mathcal{A}(S)$, then we can use Theorem 5 to find disjoint Borel sets

$$B_1 \supset A \quad \text{and} \quad B_2 \supset A^c.$$

Since $A + A^c = S$, we have $A = B_1 \in \mathcal{B}(S)$. Hence

$$A \in \mathcal{B}(S) \Longleftrightarrow A,\ A^c \in \mathcal{A}(S),$$

so

$$\mathcal{B}(S) = \{A: A, A^c \in \mathcal{A}(S)\} \subset \mathcal{A}(S).$$

But

$$\mathcal{A}(S) \subset \alpha[\mathcal{F}(S)] \text{ (Theorem 4) and } \alpha[\mathcal{A}(S)] \subset \mathcal{A}(S) \text{ (Theorem 2)}.$$

Hence we have

$$\mathcal{A}(S) \subset \alpha[\mathcal{F}(S)] \subset \alpha[\mathcal{B}(S)] \subset \alpha[\mathcal{A}(S)] \subset \mathcal{A}(S),$$

so all these classes must coincide. ∎

<u>Theorem 8</u>. $\mathcal{S}(S)$ is closed under countable unions for every

space S.

<u>Proof</u>. Let $A_n \in \mathcal{S}(S)$, $n = 1, 2, \ldots$. The union $A := \bigcup_n A_n$ is the disjoint union of the following disjoint sets:

$$B_1 := A_1$$
$$B_n := A_n - A_n \cap (B_1 + B_2 + \ldots + B_{n-1}), \quad n=2,3,\ldots .$$

Since a countable disjoint union of standard subsets of S is also standard (Theorem 1), it is enough to prove that every B_n is standard. It is trivial that B_1 is standard. Suppose that $B_1, B_2, \ldots, B_{n-1}$ are standard. Then $(B_1 + B_2 + \ldots + B_{n-1})$ is standard. Since A_n is standard, Theorems 1 and 7(i) ensures that

$$A_n \cap (B_1+B_2+\ldots+B_{n-1}) \in \mathcal{S}(A_n) = \mathcal{B}(A_n),$$

so

$$B_n \in \mathcal{B}(A_n) = \mathcal{S}(A_n).$$

∎

A space is called <u>σ-compact</u> if it is expressible as a countable union of compact subsets. Since every compact metrizable space is standard by Theorem 4.7, the last theorem implies

<u>Theorem 9</u>. Every σ-compact metrizable space is stnadard.

6. Borel maps in standard spaces and analytic spaces.

Let f be a map from S into T. Then the set

$$\{(x,y) \in S \times T : y = f(x)\}$$

is called the <u>graph</u> of f, denoted by $G(f)$.

<u>Theorem 1</u>. Let $f: S \to T$ be a Borel map, where S and T are analytic. Then

$$G(f) \in \mathcal{B}(S \times T) \subset \mathcal{A}(S \times T)$$

<u>Proof</u> Consider the map

$$g: S \times T \to T \times T, \quad (x,y) \mapsto (f(x),y)$$

and the diagonal set D of $T \times T$. It is obvious that

$$G(f) = g^{-1}(D).$$

Also g is measurable $\mathcal{B}(S) \times \mathcal{B}(T)/\mathcal{B}(T) \times \mathcal{B}(T)$, being the bilateral product map of the map $f: S \to T$ and the identity map $i: T \to T$. Since S and T are analytic, Theorem 4.5 ensures that

$$\mathcal{B}(S) \times \mathcal{B}(T) = \mathcal{B}(S \times T) \quad \text{and} \quad \mathcal{B}(T) \times \mathcal{B}(T) = \mathcal{B}(T \times T),$$

so the map g is Borel. Since D is closed,

$$G(f) = g^{-1}(D) \in \mathcal{B}(S \times T)$$

Since $S \times T$ is analytic, $\mathcal{B}(S \times T) \subset \mathcal{A}(S \times T)$ (Theorem 5.7(ii)). ∎

<u>Theorem 2</u>. Let $f: S \to T$ be a Borel map.

(i) If S and T are analytic, then

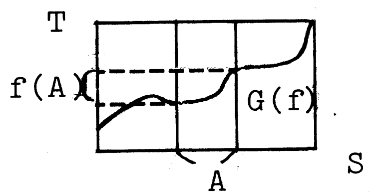
$$f(\mathcal{A}(S)) \subset \mathcal{A}(T) \text{ (especially } f(S) \in \mathcal{A}(T)) \text{ and } f^{-1}(\mathcal{A}(T)) \subset \mathcal{A}(S).$$

(ii) If S and T are standard and if f is injective, then

$$f(\mathcal{B}(S)) \subset \mathcal{B}(T) \text{ (especially } f(S) \in \mathcal{B}(T)).$$

<u>Proof</u>

(i) Let $\pi_2 : S \times T \to T$ be the canonical projection. Then



$$f(A) = \pi_2[(A \times T) \cap G(f)].$$

$G(f) \in \mathcal{A}(S \times T)$ (Theorem 1). Let $A \in \mathcal{A}(S)$. Then $A \times T \in \mathcal{A}(S \times T)$ (Theorem 4.4). Hence

$$(A \times T) \cap G(f) \in \mathcal{A}(S \times T) \quad \text{(Theorem 5.2).}$$

Since $\pi_2 : S \times T \to T$ is continuous, $f(A) \in \mathcal{A}(T)$. This proves that $f(\mathcal{A}(S)) \subset \mathcal{A}(T)$.

Let $B \in \mathcal{A}(T)$. Then $B \in \alpha[\mathfrak{F}(T)]$ (Theorem 5.7(ii)), so

$$B = \bigcup_{(n_i)} \bigcap_k F_{n_1 n_2 \ldots n_k} \quad \text{where} \quad F_{\ldots} \in \mathfrak{F}(T).$$

Hence

$$f^{-1}(B) = \bigcup_{(n_i)} \bigcap_k f^{-1}(F_{n_1 n_2 \ldots n_k}) \in \alpha[\mathcal{B}(S)] = \mathcal{A}(S)$$

by Theorem 5.7(ii). This proves that $f^{-1}(\mathcal{A}(T)) \subset \mathcal{A}(S)$.

(ii) Let $A \in \mathcal{B}(S)$. Then $A \times T \in \mathcal{B}(S \times T)$, so

$$A' := (A \times T) \cap G(f) \in \mathcal{B}(S \times T) = \mathcal{S}(S \times T),$$

because $S \times T$ is standard. Let π_2' denote the restriction of π_2 to A'. Then $\pi_2' : A' \to T$ is a continuous injection and

$$f(A) = \pi_2(A') = \pi_2'(A'),$$

so $f(A) \in \mathcal{S}(T) = \mathcal{B}(T)$, because T is standard. This proves that $f(\mathcal{B}(S)) \subset \mathcal{B}(T)$. ┘

Theorem 3. Let S and T be analytic spaces. If $f: S \to T$ is a Borel bijection, then f is bimeasurable, so S is Borel isomorphic to T.

Proof. Since f is bijective and since $f^{-1}(\mathcal{B}(T)) \subset \mathcal{B}(S)$, it is enough to prove that $f(\mathcal{B}(S)) \subset \mathcal{B}(T)$. Let $A \in \mathcal{B}(S)$. Then $A, A^c \in \mathcal{A}(S)$ (Theorem 5.7(ii)), so

$$f(A), f(A^c) \in \mathcal{A}(T).$$

Since f is bijective,

$$f(A)^c = f(A^c) \in \mathcal{A}(T).$$

Hence $f(A) \in \mathcal{B}(T)$ (Theorem 5.7(ii)). ┘

We denote the cardianl member of a set S by $\#S$.

<u>Theorem 4</u>. Let S be analytic. If $\#S > \aleph_0$, then S has a compact subset homeomorphic to the Cantor set $\mathbb{K}$ and also S has a Borel subset Borel isomorphic to $[0,1]$.

<u>Proof</u>. Take a complete separable metric space $P = (P,\rho)$ and a continuous surjection $f:P \to S$. Since $f^{-1}(x) \neq \phi$ for every $x \in S$, we can use the axiom of choice to find a subset A of $\mathbf{P}$ such that $g = f|_A : A \to S$ is bijective. Then

$$\#A = \#S > \aleph_0.$$

Let B denote the set of all $p \in A$ such that

$$\#(U \cap A) > \aleph_0 \text{ for every neighborhood } U \text{ of } p$$

and let $C := A - B$. Then every point $p \in C$ obviously has a neighborhood $U(p)$ such taht $U(p) \cap A$ is countable. Since $\mathbf{P}$ is fully Lindelöf, we can find a countable set $\{p_n\} \subset C$ such that

$$\bigcup_{p \in C} U(p) = \bigcup_n U(p_n).$$

Hence we have

$$C \subset \bigcup_{p \in C} U(p) \cap A = \bigcup_n U(p_n) \cap A,$$

implying that C is countable, so $\#B > \aleph_0$. Observing that

$$\#(U \cap A - U \cap B) \leq \#C \leq \aleph_0 ,$$

we have

(1) $\#(U \cap B) > \aleph_0$ for every neighborhood U of $p \in B$.

Now we will construct a family

$$U_{i_1 i_2 \ldots i_n} : n=1,2,\ldots; \quad i_\nu = 0,1,$$

each being a neighborhood of a point of B. Take two distinct points $p_0, p_1 \in B$. Since $B \subset A$, $f(p_0) \neq f(p_1)$. Hence there are disjoint neighborhoods $V(f(p_0))$ and $V(f(p_1))$. Since P is metrizable and since f is continuous, we can find disjoint neighborhoods $U_i = U(p_i)$, $i=0,1$ such that

$$f(\overline{U}_i) \subset V(f(p_i)), \quad i=0,1.$$

Then $f(\overline{U}_0) \cap f(\overline{U}_1) = \phi$. We can take U_i so that

$$\delta(\overline{U}_i) < 2^{-1}, \quad i=0,1 \quad (\delta = \rho\text{-diameter}).$$

Suppose that we have constructed $U_{i_1 i_2 \ldots i_n}$. Since $U_{i_1 i_2 \ldots i_n}$

is a neighborhood of a point of B, $\#(U_{i_1 i_2 \cdots i_n} \cap B) > \aleph_0$, so we can take two distinct points in $U_{i_1 i_2 \cdots i_n} \cap B$. Applying the same argument as above, we construct $U_{i_1 i_2 \cdots i_{n+1}}$, $i_{n+1} = 0, 1$ such that

$$f(\overline{U}_{i_1 i_2 \cdots i_n 0}) \cap f(\overline{U}_{i_1 i_2 \cdots i_n 1}) = \phi.$$

We can take $U_{i_1 i_2 \cdots i_n i}$ so that

$$U_{i_1 i_2 \cdots i_n i} \subset U_{i_1 i_2 \cdots i_n} \text{ and } \delta(\overline{U}_{i_1 i_2 \cdots i_n i_{n+1}}) < 2^{-n-1}.$$

Thus we obtain $U_{i_1 i_2 \cdots i_n}$ for every n and for every $(i_1, i_2, \ldots, i_n)$.

Let $\xi = (i_\nu) \in \{0,1\}^\infty$. Then $\overline{U}_{i_1 i_2 \cdots i_n}$ decreases as $n \uparrow \infty$ and $\delta(\overline{U}_{i_1 i_2 \cdots i_n}) \downarrow 0$. Hence the Cantor intersection theorem ensures that $\overline{U}_{i_1 i_2 \cdots i_n}$ monotonically converges to a point which will be denoted by p_ξ. Suppose that $\xi = (i_\nu) \neq \eta = (j_\nu)$. Then

$$i_1 = j_1, \; i_2 = j_2, \ldots, \; i_{n-1} = j_{n-1} \text{ and } i_n \neq j_n$$

for some n.

Hence $f(p_\xi) \in f(\overline{U}_{i_1 \cdots i_{n-1} i_n})$ and $f(p_\eta) \in f(\overline{U}_{i_1 i_2 \cdots i_{n-1} j_n})$. Since these two sets are disjoint according to the construction above, $f(p_\xi) \neq f(p_\eta)$. Thus the map

$$\varphi : \{0,1\}^\infty \to S, \quad \xi \mapsto f(p_\xi)$$

is injective. Since

$$\pi_\nu(\xi) = \pi_\nu(\eta) = i_\nu \, (\nu \le n) \Rightarrow p_\xi, p_\eta \in \overline{U}_{i_1 i_2 \ldots i_n} \Rightarrow \rho(p_\xi, p_\eta) < 2^{-n},$$

the map $\xi \mapsto p_\xi$ is continuous. Hence $\varphi : \xi \mapsto f(p_\xi)$ is also continuous. Since $\{0,1\}^\infty$ is compact, the image

$$E := \varphi(\{0,1\}^\infty)$$

is a compact subset homeomorphic to $\{0,1\}^\infty$.

Since $\{0,1\}^\infty$ is homeomorphic to the Cantor set $\mathbb{K}$ under the map

$$(i_\nu) \mapsto \sum_{\nu=1}^{\infty} 2i_\nu/3^\nu,$$

E is a compact subset of S homeomorphic to $\mathbb{K}$. This proves the first conclusion.

Let Γ be the set of all $(i_\nu) \in \{0,1\}^\infty$ such that <u>either</u> $i_\nu = 1$ for every ν <u>or</u> $i_\nu = 0$ for infinitely many ν. Since $\{0,1\}^\infty - \Gamma$ is countable, Γ is a Borel subset of $\{0,1\}^\infty$. Since

$$\psi : \Gamma \to [0,1], \quad (i_\nu) \mapsto \sum_{\nu=1}^{\infty} i_\nu/2^\nu$$

is a continuous bijection, $\Gamma \underset{B}{\simeq} [0,1]$ (Theorem 3). Since $\varphi : \{0,1\}^\infty \to E$ is bicontinuous and since $\Gamma \in \mathcal{B}(\{0,1\}^\infty)$, $F := \varphi(\Gamma) \in \mathcal{B}(E) \subset \mathcal{B}(S)$ and $F \underset{B}{\simeq} \Gamma$, so $[0,1] \underset{B}{\simeq} F \in \mathcal{B}(S)$. This proves the second conclusion.

∎

<u>Theorem 5</u>:

(i) Every analytic space is Borel isomorphic to an analytic subset of [0,1].

(ii) Every standard space is Borel isomorphic to one of [0,1], $\mathbb{N}=\{1,2,\dots\}$ and $\mathbb{N}_n = \{1,2,\dots n\}$ $(n = 1,2,\dots)$.

Remark. The second assertion implies that every standard space is Borel isomorphic to a compact subspace of [0,1], because $\mathbb{N}$ and $\mathbb{N}_n$ are Borel isomorphic to

$$\{2^{-1},2^{-2},\dots,0\} \text{ and } \{2^{-1},2^{-2},\dots,2^{-n}\}$$

respectively.

Proof of the theorem.

(i) Let S be analytic. First we prove that there is a sequence $\{U_n\} \subset \mathcal{B}(S)$ such that for every two distinct points $x,y \in S$ we have

$$1_{U_m}(x) \neq 1_{U_m}(y) \quad \text{for some } m.$$

Let D be the diagonal set of $S^2 := S \times S$. Then the set $G := S^2 - D$ is open in S^2. Being analytic, S^2 is fully Lindelöf. Therefore G can be expressed as

$$G = \bigcup_n U_n \times V_n, \quad U_n, V_n \text{ open in } S.$$

Since $G \cap D = \phi$,

$$U_n \cap V_n = \phi \quad \text{for every } n.$$

Let x and y be two distinct points in S. Then $(x,y) \in G$, so

$$(x,y) \in U_m \times V_m \quad \text{for some} \quad m.$$

Hence $x \in U_m$ and $y \in V_m \subset U_m^c$, so

$$1_{U_m}(x) = 1 \quad \text{and} \quad 1_{U_m}(y) = 0, \quad \text{i.e.} \quad 1_{U_m}(x) \neq 1_{U_m}(y).$$

Define a map $f : S \to [0,1]$ by

$$f(x) = \sum_{n=1}^{\infty} \frac{2}{3^n} 1_{U_n}(x).$$

Since $U_n \in \mathcal{B}(S)$, $n=1,2,\ldots,$ it is easy to check that f is Borel. If $x \neq y$, then $1_{U_m}(x) \neq 1_{U_m}(y)$ for some m, so $f(x) \neq f(y)$. Hence f is a Borel injection. Now use Theorems 3 and 2(i) to conclude that

$$S \underset{B}{\sim} f(S) \in \mathcal{A}([0,1]).$$

(ii) Let S be standard. If $\#S \leqq \aleph_0$, the conclusion is obvious. Suppose that there is an injective Borel map $f : S \to [0,1]$ by (i). Then

$$S \underset{B}{\sim} f(S) \in \mathcal{B}([0,1])$$

by Theorems 3 and 2(ii). Also Theorem 4 ensures that we can find a subset E of S such that

$$[0,1] \underset{B}{\sim} E \in \mathcal{B}(S).$$

Now use Theorem 2.2 to conclude that $S \underset{B}{\sim} [0,1]$. ∎

Let $f : S \to T$ be a <u>surjection</u>. A map $g : T \to S$ is called an <u>inverse map</u> if

$$f(g(y)) = y \quad \text{for every} \quad y \in T.$$

If $g : T \to S$ is an inverse map of $f : S \to T$, then the image

$$A := g(T)$$

is a subset of S satisfying the following condition.

(C) For every $y \in T$ $A \cap f^{-1}(y)$ consists of exactly one point, which will be denoted by $x_A(y)$.

Conversely, if A is a subset of S satisfying (C), then the map

$$g : T \to S, \quad y \mapsto x_A(y)$$

is a unique inverse map of f with the image $g(T) = A$. Hence there is a 1-1 correspondence between the inverse maps of f and the subsets of S satisfying (C).

<u>Theorem 6</u>. Let $f: S \to T$ be a Borel surjection, where S and T are analytic. Then f has an inverse $g : T \to S$ with the following properties.

(I.1) $g(T) \in \sigma[\mathcal{A}(S)]$

(I.2) g is measurable $\sigma[\mathcal{A}(T)]/\mathcal{B}(S)$.

<u>Proof.</u> First we will prove the theorem under the assumption that f is continuous. Take a decreasing Souslin scheme $\mathcal{S} = \{S_{n_1 n_2 \dots n_k}\}$ mentioned in Theorem 5.3 and denote the limit

of $\{S_{n_1 n_2 \ldots n_k}\}_{k=1,2,\ldots}$ by x_ξ where $\xi = (n_i)$. Consider a new Souslin scheme $\mathfrak{A}$ composed of

(1) $$A_{n_1 n_2 \ldots n_k} = S_{n_1 n_2 \ldots n_k} \setminus f^{-1}\Big(\bigcup_{n<n_k} f(S_{n_1 n_2 \ldots n_k})\Big)$$

and let A denote the kernel $K(\mathfrak{A})$. We will prove that (i) A satisfies (C) and (ii) the inverse map g of f corresponding to A satisfies (I.1) and (I.2).

Using the obvious relations

$$f\Big(\bigcup_\lambda C_\lambda\Big) = \bigcup_\lambda f(C_\lambda) \quad \text{and} \quad f[C \setminus f^{-1}(f(D))] = f(C) \setminus f(D),$$

we obtain

(2) $$f(A_{n_1 n_2 \ldots n_k}) = f(S_{n_1 n_2 \ldots n_k}) \setminus \bigcup_{n<n_k} f(S_{n_1 n_2 \ldots n_{k-1} n}).$$

Next we will prove that

(3) $$\bigcap_k f(A_{n_1 n_2 \ldots n_k} \cap F) = f\Big(\bigcap_k A_{n_1 n_2 \ldots n_k} \cap F\Big)$$

for every $F \in \mathcal{F}(S)$.

Denote these sets by L and R. $L \supset R$ is obvious. Let $y \in L$.

$$y \in \bigcap_k f(A_{n_1 n_2 \ldots n_k} \cap F) \subset \bigcap_k f(S_{n_1 n_2 \ldots n_k}) = \{f(x_\xi)\}$$

where $\xi = (n_i)$; the last equality follows from the continuity of f. Hence $y = f(x_\xi)$. Since

$$f(x_\xi) = y \in f(A_{n_1 n_2 \ldots n_k}),$$

we have

$$f(x_\xi) \notin \bigcup_{n<n_k} f(S_{n_1 n_2 \dots n_{k-1} n}) \qquad \text{by } (2)$$

so

$$x_\xi \notin f^{-1}(\bigcup_{n<n_k} f(S_{n_1 n_2 \dots n_{k-1} n})).$$

Since $x_\xi \in S_{n_1 n_2 \dots n_k}$ obviously, $x_\xi \in A_{n_1 n_2 \dots n_k}$. Suppose that $x_\xi \notin F$, i.e. $x_\xi \in F^c$. Since $S_{n_1 n_2 \dots n_k} \downarrow x_\xi$ and since F^c is open,

$$S_{n_1 n_2 \dots n_r} \subset F^c \quad \text{for some } r$$

so

$$A_{n_1 n_2 \dots n_r} \subset F^c, \text{ i.e. } A_{n_1 n_2 \dots n_r} \cap F = \phi.$$

Then L must be empty contrary to the assumption that $y \in L$. Therefore $x_\xi \in F$. Thus

$$x_\xi \in \bigcap_k A_{n_1 n_2 \dots n_k} \cap F, \quad \text{so} \quad y = f(x_\xi) \in R.$$

This proves $L \subset R$, which, combined with $L \supset R$, implies (3).

Now we prove that the set A satisfies (C). Let y be an arbitrary point of T. Then

$$y \in T = f(S) = \bigcup_n f(S_n) .$$

Let n_1 be the minimum of n for which $y \in f(S_n)$. Then

$$y \in f(S_{n_1}) = \bigcup_n f(S_{n_1 n}).$$

Let n_2 be the minimum of n for which $y \in f(S_{n_1 n})$. Repeating this, we determine n_i, $i = 1,2,\ldots$. Then

$$y \in \bigcap_k f(S_{n_1 n_2 \ldots n_k}) = \{f(x_\xi)\}, \quad \text{where} \quad \xi = (n_i),$$

so

$$y = f(x_\xi).$$

It is obvious that $x_\xi \in S_{n_1 n_2 \ldots n_k}$. By the choice of n_k

$$f(x_\xi) = y \notin \bigcup_{n<n_k} f(S_{n_1 n}), \quad \text{so} \quad x_\xi \notin f^{-1}[\bigcup_{n<n_k} f(S_{n_1 n})].$$

Hence $x_\xi \in A_{n_1 n_2 \ldots n_k}$. This implies that

$$x_\xi \in \bigcap_k A_{n_1 n_2 \ldots n_k} \subset A,$$

so we have

$$x_\xi \in f^{-1}(y) \cap A.$$

Suppose that $x \in f^{-1}(y) \cap A$. Then

$$f(x) = y \quad \text{and} \quad x \in A = \bigcup_{(m_i)} \bigcap_k A_{m_1 m_2 \ldots m_k}.$$

Then

$$x \in \bigcap_k A_{m_1 m_2 \ldots m_k} \quad \text{for some} \quad \eta = (m_1, m_2, \ldots),$$

so

$$x \in \bigcap_k S_{m_1 m_2 \dots m_k} \in \{x_\eta\}.$$

Hence $x = x_\eta$. Supose that $\eta \neq \xi$. Then we have

$$m_1 = n_1, \quad m_2 = n_2, \dots, \quad m_{r-1} = n_{r-1} \quad \text{and} \quad m_r \neq n_r.$$

Since $y = f(x) = f(x_\xi)$, we have

$$y \in f(A_{n_1 n_2 \dots n_{r-1} n_r}) \quad \text{and} \quad y \in f(A_{m_1 \dots m_{r-1} m_r})$$

$$= f(A_{n_1 n_2 \dots n_{r-1} m_r}) .$$

This is a contradiction, because $f(A_{n_1 n_2 \dots n_{r-1} n})$, $n = 1,2,\dots$ must be disjoint by virtue of (2). Thus we have $\eta = \xi$, so

$$x = x_\eta = x_\xi .$$

This proves that x_ξ is the only one element of $f^{-1}(y) \cap A$. Thus the set A satisfies (C).

Let g denote the inverse map of f corresponding to A. Then

$$g(T) = A = K(\mathcal{A}).$$

It follows from (1) that

$$A_{n_1 n_2 \dots n_k} \subset S_{n_1 n_2 \dots n_k} - \bigcup_{n < n_k} S_{n_1 n_2 \dots n_{k-1} n},$$

so the Souslin scheme $\mathcal{A} = \{A_{n_1 n_2 \dots n_k}\}$ is disjoint. Hence

$$K(\mathcal{A}) \in \sigma[\mathcal{A}] \text{ (Theorem 3.2(ii))}.$$

Using Theorem 2(i), we can check that

$$\mathfrak{A} \subset \sigma[\mathcal{A}(S)],$$

so

$$g(T) = A = K(\mathfrak{A}) \in \sigma[\mathfrak{A}] \subset \sigma[\mathcal{A}(S)].$$

This proves that g satisfies (I.1).

To prove (I.2) it is enough to show that

$$g^{-1}(F) \in \sigma[\mathcal{A}(T)] \quad \text{for every} \quad F \in \mathcal{F}(S)$$

i.e.

(4) $\quad f(A \cap F) \in \sigma[\mathcal{A}(T)] \quad$ for every $\quad F \in \mathcal{F}(S)$.

Since $A = K(\mathfrak{A})$,

(5) $$f(A \cap F) = \bigcup_{(n_i)} f(\bigcap_k A_{n_1 n_2 \ldots n_k} \cap F).$$

$$= \bigcup_{(n_i)} \bigcap_k f(A_{n_1 n_2 \ldots n_k} \cap F) \qquad \text{by (3).}$$

Since the Souslin scheme $\{f(A_{n_1 n_2 \ldots n_k})\}$ is disjoint by virtue of (2), the Souslin scheme $\{f(A_{n_1 n_2 \ldots n_k} \cap F)\}$ is disjoint. But

$$f(A_{n_1 n_2 \ldots n_k} \cap F) = f[S_{n_1 n_2 \ldots n_k} \cap F \setminus f^{-1}(\bigcup_{n<n_k} f(S_{n_1 n_2 \ldots n_{k-1} n}))]$$

$$= f(S_{n_1 n_2 \ldots n_k} \cap F) \setminus \bigcup_{n<n_k} f(S_{n_1 n_2 \ldots n_{k-1} n})$$

$$\in \sigma[\mathcal{A}(T)].$$

Hence we can use Theorem 3.2(ii) to conclude that

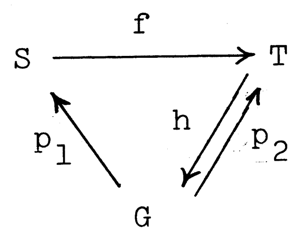
$$f(A \cap F) \in \sigma[\mathcal{A}(T)],$$

proving that g is measurable $\sigma[\mathcal{A}(T)]/\mathcal{B}(S)$. Thus our theorem is proved under the assumption that f is continuous.

Now we will discuss the general case where f is Borel measurable. The graph $G = G(f)$ is Borel in $S \times T$ (Theorem 1), so G is analytic (Theorem 5.7(ii)). Consider the canonical projections

$$p_1 : G \to S \quad \text{and} \quad p_2 : G \to T .$$

Since $f : S \to T$ is surjective, p_2 is a continuous surjection. Hence p_2 has an inverse map h with the properties:



(I'.1) $\quad h(T) \in \sigma[\mathcal{A}(G)]$

(I'.2) $\quad h$ is measurable $\sigma[\mathcal{A}(T)]/\mathcal{B}(G)$.

Since $p_1 : G \to S$ is a continuous bijection, p_1 is bimeasurable (Theorem 3), i.e.

$$p_1(\mathcal{B}(G)) = \mathcal{B}(S) \quad \text{and} \quad p_1^{-1}(\mathcal{B}(S)) = \mathcal{B}(G).$$

Keeping this in mind we will prove that the composition $g : p_1 \circ h : T \to S$ is an inverse map of f satisfying (I.1) and (I.2).

Let y be an arbitrary point in T. Since $h(y) \in G$, we have

$$h(y) = (x_y, y) \quad \text{where} \quad f(x_y) = y .$$

Hence

$$g(y) = p_1(h(y)) = x_y,$$

so

$$f(g(y)) = f(x_y) = y.$$

Hence g is an inverse map of f. Using Theorem 5.7(ii), we obtain

$$g(T) = p_1(h(T)) \in p_1[\sigma[\mathcal{A}(G)]] = p_1(\sigma[\alpha[\mathcal{B}(G)]]).$$

Since p_1 is bijective, we have

$$p_1(\alpha[\mathfrak{A}]) = \alpha[p_1(\mathfrak{A})] \quad \text{and} \quad p_1(\sigma[\mathfrak{A}]) = \sigma[p_1(\mathfrak{A})]$$
$$\text{for every} \quad \mathfrak{A} \subset 2^G,$$

so

$$p_1(\sigma[\alpha[\mathcal{B}(G)]]) = \sigma[\alpha[p_1(\mathcal{B}(G))]] = \sigma[\alpha[\mathcal{B}(S)]] = \sigma[\mathcal{A}(S)].$$

Therefore

$$g(T) \in \sigma[\mathcal{A}(S)].$$

Since

$$g^{-1}(\mathcal{B}(S)) = h^{-1}(p_1^{-1}(\mathcal{B}(S))) = h^{-1}(\mathcal{B}(G)) \subset \sigma[\mathcal{A}(T)],$$

g is measurable $\sigma[\mathcal{A}(T)]/\mathcal{B}(S)$. ∎

7. Function spaces.

Practically all function spaces appearing in probability theory are analytic (even standard). Here we will give some typical examples. For simplicity we consider only spaces of real functions on [0,1], but it is not difficult to extend the results to more general cases.

(a) $C = C[0,1]$ = the space of all continuous functions.

C is a separable real Banach space with the usual linear operation and the maximum norm:

$$\|f\| = \max_{0 \le t \le 1} |f(t)|.$$

Hence the space C with the norm topology is Polish. The same space with the weak topology is standard (Theorem 4.8).

(b) $D_* = D_*[0,1]$ = the space of all right continuous functions with left limits.

Let Φ be the set of all increasing continuous bijections $\varphi: [0,1] \to [0,1]$. The Skorohod topology τ_S on D_* is given by the metric

$$\rho_S(f,g) = \inf_{\varphi \in \Phi} [\|\varphi - i\| + \|f \circ \varphi - g\|]$$

where i is the identity map on $[0,1]$ and $\| \;\; \|$ denotes the supremum norm. The space D_* (with τ_S) is Polish. In fact the metric ρ_S itself is not Polish, but there are Polish metrics giving the topology τ_S. One of such metrics is the Billingsley metric ρ_B given as follows. Let

$$\beta(\varphi) := \sup_{t \neq s} \left|\log \frac{\varphi(t) - \varphi(s)}{t - s}\right|, \quad \varphi \in \Phi .$$

and let Ψ denote the set $\{\varphi \in \Phi : \beta(\varphi) < \infty\}$. The <u>Billingsley metric</u> ρ_B is given by

$$\rho_B(f,g) = \inf_{\psi \in \Psi} [\|f \circ \psi - g\| + \beta(\psi)].$$

See Billingsley [ ] for the details. Let

$$D = D[0,1] = \{\varphi \in D_* : \varphi(1-) = \varphi(1)\}, \quad \varphi(t-) = \lim_{s \uparrow t} \varphi(t).$$

The Skorohod topology τ_S on D is defined in the same way as above and the space D with τ_S is also Polish.

(c) $M^+ = M^+[0,1] =$ <u>the space of all finite measures on</u> $[0,1]$ (<u>defined for all Borel sets</u>).

The weak topology on M^+ is defined by the following neighborhood base

$$U(\mu; f_1, f_2, \ldots, f_n, \varepsilon) := \{\nu \in M^+ : |\langle f_i, \nu\rangle - \langle f_i ; \mu\rangle| < \varepsilon, \ i = 1,2,\ldots,n\}$$

$$\varepsilon > 0 ; \quad n = 1,2,\ldots ; \ f_i \in C[0,1],$$

where

$$\langle f, \mu\rangle = \int_{[0,1]} f d\mu .$$

The space M^+ (with the weak topology) is Polish. This is a special case of Prohorov's theorem (Appendix).

(d) $M = M[0,1]$ = the space of all signed measures on $[0,1]$ (defined for all Borel sets).

The weak topology on M is defined in the same way as above. The space M with the weak topology is standard. To prove this we first recall several known facts, M is the dual space of the Banach space C i.e. $M = C^*$ where the norm $\|\theta\|$ $(\theta \in M)$ is the total absolute variation of θ. Hence the weak topology on M should be called the weak-star topology in accordance with the Banach space terminology, but we will use the word "weak topology" for simplicity. Note that M is not separable in general, so we cannot use Theorem 4.8 to prove that the space M with the weak topology is standard. For $\mu,\nu \in M^+$ given the largest measure $\leq \mu, \nu$ is denoted by $\mu \wedge \nu$. Using the Radon-Nikodym densities $d\mu|d(\mu+\nu)$ and $d\nu|d(\mu+\nu)$, we can easily prove that

$$\|\mu \wedge \nu\| = \frac{1}{2}[\|\mu + \nu\| - \|\mu - \nu\|] , \quad \mu,\nu \in M^+.$$

Every $\theta \in M$ has a unique decomposition (the Jordan decomposition);

$$\theta = \mu-\nu , \quad \mu,\nu \in M^+, \qquad \|\mu \wedge \nu\| = 0.$$

In the discussion below we always consider M^+ and M with the weak topology. Since M^+ is Polish, $(M^+)^2$ is Polish. Since the map

$$\varphi : (M^+)^2 \longrightarrow M, \quad (\mu,\nu) \longmapsto \mu - \nu$$

is a continuous surjection, M must be analytic. The proof that M is standard is slightly harder. Let

$$\Delta := \{(\mu,\nu) \in (M^+)^2 : \|\mu \wedge \nu\| = 0\}.$$

Then the restinction $\psi = \varphi|_\Delta : \Delta \to M$ is a continuous bijection. Since $\|\mu\| = \sup_{f\in\Gamma} |\langle f,\mu\rangle|$ where Γ being any countable dense subset of C, $\mu \mapsto \|\mu\|$ is Borel. Since $(\mu,\nu) \mapsto \mu\pm\nu$ are continuous, the maps $(\mu,\nu) \mapsto \|\mu\pm\nu\|$ are Borel, so $(\mu,\nu) \mapsto \|\mu\wedge\nu\|$ is Borel. Thus Δ is a Borel subset of $(M^+)^2$. Since $(M^+)^2$ is Polish, Δ is standard (Theorem 5.7(i)). Since $\psi : \Delta \to M$ is a continuous bijection, M is also standard.

(e) $\mathbb{L}^0 = \mathbb{L}^0[0,1] =$ the space of all Lebesgue measurable functions, where equivalent functions are identified.

$\mathbb{L}^0$ is topologized by the following metric:

$$\rho_0(f, g) = \int_0^1 [|f(t) - g(t)| \wedge 1]\, dt.$$

This topology is often called the topology of convergence in measure, because

$$\rho_0(f_n,f) \to 0 \iff \lambda\{t \in [0,1] : |f_n(t) - f(t)| > \varepsilon\} \to 0\ ,\quad \forall \varepsilon > 0,$$

where λ denotes the Lebegue measure. Since ρ_0 is Polish, the space $\mathbb{L}^0$ with the ρ_0-topology is obviously Polish.

(f) $\mathbb{L}^p = \mathbb{L}^p[0,1] = \{f \in L^0 : \int_0^1 |f(t)|^p\, dt < \infty\}$ $(1\le p<\infty)$.

$\mathbb{L}^p$ is a separable Banach space with the usual linear operation and the p-norm:

$$\|f\|_p = \left(\int_0^1 |f(t)|^p dt\right)^{\frac{1}{p}} .$$

Hence the space $\mathbb{L}^p$ with the norm topology is Polish and the same space with the weak topology is standard. Suppose that $p > 1$. Since the dual space of $\mathbb{L}^p$ is $\mathbb{L}^q$ $(p^{-1} + q^{-1} = 1)$, the weak topology in L^p is given by the following neighborhood base

$$U(f; g_1, g_2, \ldots, g_n, \varepsilon) = \{h : |\langle g_i, h\rangle - \langle g_i, f\rangle| < \varepsilon,\ i=1,2,\ldots,n\}$$

$$\varepsilon > 0;\ n = 1,2,\ldots;\ g_i \in L^q.$$

The dual space of $\mathbb{L}^1$ is

$$\mathbb{L}^\infty := \{f \in L^0 : \operatorname*{ess.sup}_t |f(t)| < \infty\},$$

where the norm in $\mathbb{L}^\infty$ is defined by

$$\|f\|_\infty = \operatorname*{ess.sup}_t |f(t)|.$$

The space L^∞ with the norm topology is a non-separable Banach space.

Generalizing the notion L^p we can define

$$\mathbb{L}^p_\mu = \mathbb{L}^p([0,1], \mu) = \{f : \int_0^1 |f|^p \, d\mu < \infty\} \quad (1 \le p < \infty)$$

where the p-norm $\|f\|_p$ is defined similarly, the space $\mathbb{L}^p_\mu$ with the norm topology is also Polish.

(g) $\mathcal{D} = \mathcal{D}[0,1]$ = <u>the space of all C^∞ functions on [0,1], where the derivatives at 0 (or 1) are understood to be the right (or left) derivatives</u>.

$\mathcal{D}$ is a vector space with the usual linear operation. $\mathcal{D}$ is a topological vector space with the <u>Schwartz topology</u> defined by the collection of norms

$$\|\varphi\|_n = \left(\sum_{k=0}^{n} \int_0^1 |\varphi^{(k)}(t)|^2 dt\right)^{1/2}, \quad n = 1,2,3,\ldots .$$

Since this topology is given by the metric

$$\rho(\varphi,\psi) = \sum_{n=0}^{\infty} 2^{-n}[\|\varphi-\psi\|_n \wedge 1]$$

and since ρ is Polish, the space $\mathcal{D}$ with the Schwartz topology is Polish.

(h) $\mathcal{D}' = \mathcal{D}'[0,1]$ = <u>the space of Schwartz distributions on</u> $[0,1]$. This is the dual space of $\mathcal{D}$, i.e. the space of all continuous linear functionals on $\mathcal{D}$.

$\mathcal{D}'$ is a topological vector space with the usual linear operation and the strong topology τ_s or the weak topology τ_w. τ_s (resp. τ_w) is defined by the collection of semi-norms:

$$\|F\|_B = \sup_{\varphi\in B}|F(\varphi)| , \qquad B: \text{ bounded}$$

(resp. $\|F\|_\Phi = \sup_{\varphi\in\Phi}|F(\varphi)|$, Φ: finite),

where a subset B of $\mathcal{D}$ is called <u>bounded</u> if for every neighborhood U of 0 we can find n such that $B\subset nU$.

For any fixed n the normed space $(\mathcal{D}, \| \|_n)$ is a (real) pre-Hilbert space, because $\| \|_n$ is induced from an inner product:

$$\|\varphi\|_n^2 = (\varphi,\varphi)_n \quad \text{where} \quad (\varphi,\psi)_n = \sum_{k=0}^{n}\int_0^1 \varphi^{(k)}(t)\psi^{(k)}(t)dt.$$

Let $\mathcal{D}'_n$ be the dual space of this pre-Hilbert space, where the norm $\| \|_{-n}$ in $\mathcal{D}'_n$ is defined

$$\|F\|_{-n} = \sup_{\|\varphi\|\le 1} |F(\varphi)| .$$

Then $\mathcal{D}'_n$ is Polish, being an separable Hilbert space isomorphic to the completion of $\mathcal{D}$. It is known that

$$\mathcal{D}' = \bigcup_n \mathcal{D}'_n .$$

Also the topology on $\mathcal{D}'_n$ as a subspace of $(\mathcal{D}', \tau_s)$, i.e. the induced topology $\tau_s|_{\mathcal{D}'_n}$, coincides with the original topology on $\mathcal{D}'_n$. Since $\mathcal{D}'_n$ with the original topology is Polish, $\mathcal{D}'_n \in \mathfrak{P}(\mathcal{D}') \subset \mathcal{S}(\mathcal{D}')$. Hence $(\mathcal{D}', \tau_s)$ is standard (Theorem 5.8). Since τ_w is weaker than τ_s, $(\mathcal{D}', \tau_w)$ is also standard (Theorem 4.3). These facts are due to X.Fernique [ ].

Now we will investigate the relation among these spaces. It is obvious that

$$\mathcal{D} \subset C \subset D \subset \mathbb{L}^p \subset \mathbb{L}^r \subset M \subset \mathcal{D}', \quad 1 \leq r < p < \infty .$$

Denote these spaces by S_k, $k=1,2,\ldots,7$. Then the canonical injection $i_m: S_m \to S_{m+1}$ is continuous. Hence Theorem 5.7 (i) ensures that

$$S_m \in \mathcal{B}(S_{m+1}) ,$$

so

$$S_m \in \mathcal{B}(S_{m+k}) .$$

also Theorem 6.3 ensures that $\mathcal{B}(S_m) = \mathcal{B}(S_{m+k}) \cap S_m$, which, together with $S_m \in \mathcal{B}(S_{m+k})$, implies that $\mathcal{B}(S_m) \subset \mathcal{B}(S_{m+k})$.

The continuity of i_m $(m \neq 3)$ follows easily from the definitions. We will prove that $i_3: D \to \mathbb{L}^p$ is continuous. Suppose that

$$\rho_S(f_n, f) \to 0 \quad \text{where} \quad f_1, f_2, \ldots, f \in D .$$

Then we can find $\varphi_n \in \Phi$ such that

$$\varphi_n(t) \to t \quad \text{and} \quad f_n(t) - f(\varphi_n(t)) \to 0$$
uniformly in $t \in [0,1]$.

Hence

$$\| f_n - f \circ \varphi_n \|_p \to 0 \ .$$

Since $\varphi_n(t) \to t$,

$$f(\varphi_n(t)) \to f(t) \quad \text{at every continuity point } t \text{ of } f.$$

This implies that

$$\| f \circ \varphi_n - f \|_p \to 0,$$

because the discontinuity points of $f (\in D_*)$ form a countable set. Thus we have

$$\| f_n - f \|_p \to 0 \ ,$$

proving the continuity of i_3.

Similarly we have

$$\mathcal{D} \subset C \subset D_* \subset \mathbb{L}^p_\mu \subset \mathbb{L}^r_\mu \subset M \subset \mathcal{D}', \ 1 \leq r < p < \infty,$$

where μ is the sum of the Lebesgue measure and the <u>δ-measure</u> concentrated at 1. Denote these spaces by T_k, $k=1,2,\ldots,7$. Then

$$T_m \in \mathcal{B}(T_{m+k}) \ .$$

Note that $D \subset \mathbb{L}^p$ does not hold, because two functions taking the same values on $[0,1)$ and different values at 1 are distinct in D, though they are identified in $\mathbb{L}^p$.

8. Standard Borel spaces and analytic Borel spaces.

Standard spaces and analytic spaces are special topological spaces and have several nice properties that have been discussed in the previous sections. The corresponding notions for Borel spaces are standard Borel spaces and analytic Borel spaces.

A Borel space is called a standard Borel space or a Mackey space if it is Borel isomorphic to a standard space. Similarly a Borel space is called an analytic Borel space or a Blackwell space if it is Borel isomorphic to an analytic space. It is obvious that every standard Borel space is an analytic Borel space.

Every standard space is standard as a Borel space with the topological σ-algebra. Similarly for analytic spaces.

A subset F of a Borel space $(E,\mathcal{E})$ is called standard if the Borel space $(F, \mathcal{E} \cap F)$ is standard. The class of all standard subsets of $(E,\mathcal{E})$ is denoted by $\mathcal{S}(E,\mathcal{E})$. Similarly for analytic subsets of $(E,\mathcal{E})$ and the class $\mathcal{A}(E,\mathcal{E})$ of all analytic subsets of $(E,\mathcal{E})$.

Let S be a Hausdorff topological space. We have defined $\mathcal{A}(S)$ and $\mathcal{S}(S)$ in Section 5. Since S is regarded as a Borel space with $\mathcal{B}(S)$, both $\mathcal{S}(S, \mathcal{B}(S))$ and $\mathcal{A}(S, \mathcal{B}(S))$ are meaningful in the sense defined above. Since $\mathcal{B}(T) = \mathcal{B}(S) \cap T$ for $T \subset S$, it is obvious that

$$\mathscr{S}(S) \subset \mathscr{S}(S, \mathcal{B}(S)) \quad \text{and} \quad \mathcal{A}(S) \subset \mathcal{A}(S, \mathcal{B}(S)).$$

But we have

Theorem 1.

(i) $\mathscr{S}(S) = \mathscr{S}(S, \mathcal{B}(S))$ if S is a standard space.

(ii) $\mathcal{A}(S) = \mathcal{A}(S, \mathcal{B}(S))$ if S is an analytic space.

Proof. To prove (i), it is enough to check that $\mathscr{S}(S, \mathcal{B}(S)) \subset \mathscr{S}(S)$. Let $T \in \mathscr{S}(S, \mathcal{B}(S))$. Then T is Borel isomorphic to a standard space U. Therefore if follows easily from Theorem 6.2 (ii) that $T \in \mathscr{S}(S)$. The same argument works for the proof of (ii).

Theorem 2. Every countable Borel product of standard (resp. analytic) Borel spaces is standard (resp. analytic).

Proof. Let $(E_n, \mathcal{E}_n)$, $n = 1,2,\ldots$, be standard. Then we can find standard spaces S_n, $n = 1,2,\ldots$ such that

$$(E_n, \mathcal{E}_n) \underset{B}{\sim} (S_n, \mathcal{B}(S_n)).$$

Hence

$$(\Pi E_n, \Pi \mathcal{E}_n) \underset{B}{\sim} (\Pi S_n, \Pi \mathcal{B}(S_n)) = (\Pi S_n, \mathcal{B}(\Pi S_n));$$

the last equality follows from Theorem 2.4. Since ΠS_n is a standard space, $(\Pi E_n, \Pi \mathcal{E}_n)$ is a standard Borel space. This proves the assertion for standard Borel spaces. Similarly we can prove the assertion for analytic Borel spaces. ∎

Theorem 3.

(i) If $(E, \mathcal{E})$ is standard, then $\mathcal{B}(E,\mathcal{E}) = \mathcal{E}$.

(ii) If $(E, \mathcal{E})$ is analytic, then

$$\mathcal{E} = \{A \subset E : A, A^c \in \mathcal{A}(E,\mathcal{E})\} \subset \mathcal{A}(E,\mathcal{E}) = \alpha[\mathcal{E}].$$

Proof. Easy from Theorem 5.7.

Theorem 4. Let $f : (E,\mathcal{E}) \to (F,\mathcal{F})$ be a Borel map.

(i) If $(E,\mathcal{E})$ and $(F,\mathcal{F})$ are analytic, then

$$f(\mathcal{A}(E,\mathcal{E})) \subset \mathcal{A}(F,\mathcal{F}) \quad (\text{especially } f(E) \in \mathcal{A}(F,\mathcal{F}))$$

and $f^{-1}(\mathcal{A}(F,\mathcal{F})) \subset \mathcal{A}(E,\mathcal{E})$.

(ii) If $(E,\mathcal{E})$ and $(F,\mathcal{F})$ are standard and if f is injective, then $f(\mathcal{E}) \subset \mathcal{F}$ (especially $f(E) \in \mathcal{F}$).

Proof. Easy from Theorem 6.2.

Theorem 5. Let $(E,\mathcal{E})$ and $(F,\mathcal{F})$ be analytic. If $f : (E,\mathcal{E}) \to (F,\mathcal{F})$ is a Borel bijection, then f is bimeasurable, so $(E,\mathcal{E})$ is Borel isomorphic to $(F,\mathcal{F})$.

Proof. Easy from Theorem 6.3.

Theorem 6.

(i) Every analytic Borel space is Borel isomorphic to an analytic subset of [0,1].

(ii) Every standard Borel space is Borel isomorphic to one of [0,1], $\mathbb{N}$ and $\{1,2,\ldots,n\}$ $(n = 1,2,\ldots)$.

Proof. Easy from Theorem 6.5.

Theorem 7. Let $(S,\mathcal{S})$ be a Borel space and $(T,\mathcal{T})$ an analytic Borel space. If both f and g are Borel maps from $(S,\mathcal{S})$ into $(T,\mathcal{T})$, then

$$\{x \in S : f(x) = g(x)\} \in \mathcal{S}.$$

Proof. If $T \subset R$, then this is obvious. Hence our theorem follows, because every analytic Borel space is Borel isomorphic to an analytic subset of [0,1].

9. Probability meansures

Let S be a set and $\mathcal{F}$ a σ-algebra on S. A map $\mu: \mathcal{F} \to [0,1]$ is called a probability measure on S with domain $\mathcal{F}$ if μ is σ-additive, i.e.

$$\mu(\sum_n A_n) = \sum_n \mu(A_n) \quad \text{for disjoint } A_1, A_2, \ldots \in \mathcal{F}$$

and if $\mu(S) = 1$. $\mathcal{F}$ is denoted by $\mathcal{D}(\mu)$. A set A is called μ-measurable if $A \in \mathcal{D}(\mu)$ and $\mu(A)$ is called the μ-measure of A. A subset of μ-mearue 0 is called a μ-null set. A set S endowed with a probability measure μ on S is called a probability space (S,μ) or $(S,\mathcal{F},\mu)$ ($\mathcal{F} = \mathcal{D}(\mu)$).

A probability measure μ on S is called complete if

$\mu(N) = 0$ and $N' \subset N \Rightarrow N' \in \mathcal{D}(\mu)$ (so $\mu(N') = 0$).

Every probability measure μ can be extended to a complete probability measure, which is called a complete extension of μ. The least complete extension of μ is called the Lebesgue extension of μ, denoted by $\bar{\mu}$.

Let μ be a probability measure on S. The outer μ-measure μ^* and the inner μ-measure μ_* are defined by

$$\mu^*(A) = \inf_{\substack{B \supset A \\ B \in \mathcal{D}(\mu)}} \mu(B) \quad \text{and} \quad \mu_*(A) = \sup_{\substack{B \subset A \\ B \in \mathcal{D}(\mu)}} \mu(B) \quad \text{for } A \subset S.$$

The Lebesgue extension $\bar{\mu}$ of μ is characterized in terms of μ^* and μ_* as follows:

$$A \in \mathcal{D}(\bar{\mu}) \text{ and } \bar{\mu}(A) = \mu^*(A) \text{ if and only if } \mu^*(A) = \mu_*(A).$$

For every set $A \subset S$ we can find μ-measurable sets B_1 and B_2 such that

$$B_1 \subset A \subset B_2 \text{ and } \mu(B_1) = \mu_*(A) \leq \mu^*(A) = \mu(B_2).$$

We can use this fact to prove the following facts:

$$\mu^*(\bigcup_n A_n) \leq \sum_n \mu^*(A_n), \quad A_n \uparrow A \Rightarrow \mu^*(A_n) \uparrow \mu^*(A),$$

$$\mu_*(\sum_n A_n) \geq \sum_n \mu_*(A_n), \quad A_n \downarrow A \Rightarrow \mu_*(A_n) \downarrow \mu_*(A),$$

$$\mu^*(A) + \mu_*(A^c) = 1,$$

$$\mu^*(A) \leq \mu_*(A) \Rightarrow A \in \mathcal{D}(\mu), \text{ if } \mu \text{ is complete.}$$

<u>Theorem 1</u>. Let μ be a complete probability measure on S. Then $\mathcal{D}(\mu)$ is closed under the analytic operation.

<u>Proof</u>. Let $\mathfrak{A} = \{A_{n_1,n_2 \dots n_k}\}$ be a Souslin scheme and suppose that $\mathfrak{A} \subset \mathcal{D}(\mu)$. We want to prove that

$$A := \bigcup_{(n_i)} \bigcap_{k=1}^{\infty} A_{n_1 n_2 \dots n_k} \in \mathcal{D}(\mu).$$

We can assume without loss of generality that $\mathfrak{A}$ is decreasing, since $\mathcal{D}(\mu)$ is closed under finite intersections. Consider two Souslin schemes:

$$\bar{\mathfrak{A}} : \bar{A}_{n_1 n_2 \dots n_k} := \bigcup_{h_i \leq n_i (i \leq k)} A_{h_1 h_2 \dots h_k}$$

$$\underline{\mathcal{A}}: \quad \underline{A}_{n_1 n_2 \cdots n_k} := \bigcup_{\substack{h_i \le n_i \ (i \le k) \\ h_i \in \mathbb{N} \ (i > k)}} \bigcap_{j=1}^{\infty} A_{h_1 h_2 \cdots h_j}$$

Then we obtain the following facts:

(1) both $\overline{\mathcal{A}}$ and $\underline{\mathcal{A}}$ are decreasing Souslin schemes,

(2) $\underline{A}_{n_1 n_2 \cdots n_k} \subset \overline{A}_{n_1 n_2 \cdots n_k}$,

(3) $\overline{A}_{n_1 n_2 \cdots n_k} \in \mathcal{D}(\mu)$, but $\underline{A}_{n_1 n_2 \cdots n_k} \notin \mathcal{D}(\mu)$ in general,

(4) $\bigcap_k \overline{A}_{n_1 n_2 \cdots n_k} \subset A$.

(1), (2) and (3) are obvious. Let x be any point in the intersection on the left hand side of (4). Then we can find a triangular array of indices:

$$\begin{array}{llll} h_1^1, & h_1^2, & h_1^3, \cdots & \le n_1 \\ & h_2^2, & h_2^3, \cdots & \le n_2 \\ & & h_3^3, \cdots & \le n_3 \\ & & \cdots & \end{array}$$

such that

$$x \in A_{h_1^1} \cap A_{h_1^2 h_2^2} \cap A_{h_1^3 h_2^3 h_3^3} \cap \cdots .$$

Since $h_1^k \le n_1$ for $k = 1,2,\cdots$, we can find $r_1 \le n_1$ such that

$$h_1^k = r_1 \quad \text{for infinitely many } k\text{'s}.$$

Observing h_2^k for such k's, we can find $r_2 \le n_2$ such that

$$h_2^k = r_2 \quad \text{for infinitely many } k\text{'s}.$$

Repeating this procedure, we can find a sequence $r_i \leq n_i$, $i = 1,2,\cdots$, such that for each i we have

$$h_1^k = r_1,\ h_2^k = r_2, \cdots,\ h_i^k = r_i \quad \text{for infinitely many } k\text{'s}.$$

Taking a number $k = k(i)\ (\geq i)$ satisfying the above condition, we have

$$x \in A_{h_1^k h_2^k \cdots h_k^k} = A_{r_1 r_2 \cdots r_i\, h_{i+1}^k \cdots h_k^k} \subset A_{r_1 r_2 \cdots r_i}.$$

Since this holds for every i, we have

$$x \in \bigcap_i A_{h_1 h_2 \cdots h_i} \subset A,$$

proving (4).

Keeping (1), (2), (3) and (4) in mind, we will prove that $A \in \mathfrak{D}(\mu)$. Since $\underline{A}_n \uparrow A$, we have

$$\mu^*(\underline{A}_n) \uparrow \mu^*(A).$$

Similarly

$$\mu^*(\underline{A}_{n_1 n_2 \cdots n_k n}) \uparrow \mu^*(\underline{A}_{n_1 n_2 \cdots n_k}).$$

Hence we can find $m_i = m_i(\varepsilon)$ such that

$$
\begin{aligned}
\mu^*(A) &< \mu^*(\underline{A}_{m_1}) + 2^{-1}\varepsilon \\
&< \mu^*(\underline{A}_{m_1 m_2}) + 2^{-2}\varepsilon + 2^{-1}\varepsilon \\
&\cdots \\
&< \mu^*(\underline{A}_{m_1 m_2 \cdots m_k}) + 2^{-k}\varepsilon + 2^{-(k-1)}\varepsilon + \cdots + 2^{-1}\varepsilon \\
&\cdots
\end{aligned}
$$

This implies that

$$\mu^*(A) \leq \lim_k \mu^*(A_{m_1 m_2 \ldots m_k}) + \varepsilon$$

$$\leq \lim_k \mu(\overline{A}_{m_1 m_2 \ldots m_k}) + \varepsilon$$

$$= \mu(\bigcap_k \overline{A}_{m_1 m_2 \ldots m_k}) + \varepsilon$$

$$\leq \mu_*(A) + \varepsilon \qquad \text{by (4).}$$

Letting $\varepsilon \downarrow 0$ we have $\mu^*(A) \leq \mu_*(A)$, which implies that $A \in \mathcal{D}(\mu)$, because μ is complete. ∎

Let f be a map from a probability space $S = (S,\mu)$ into a set T. Define a probability measure ν on T by

$$\mathcal{D}(\nu) = \{B \subset T : f^{-1}(B) \in \mathcal{D}(\mu)\} \quad \text{and} \quad \nu(B) = \mu(f^{-1}(B)).$$

It is easy to check that ν is a probability measure on T, which will be called the <u>image measure</u> of μ under the map f, denoted by $f\mu$ or μf^{-1}. If μ is complete, then $f\mu$ is also complete. It is obvious that f is measureble $\mathcal{D}(\mu)/\mathcal{D}(f\mu)$. It $g : T \to U$ is another map, then $(g \circ f)\mu = g(f\mu)$, as we can easily check.

Let $(E,\mathcal{E})$ be a Borel space. The Lebesgue extension of a probability measure on E with domain $\mathcal{E}$ is called a <u>B-regular probability measure</u> on $(E,\mathcal{E})$. A probability measure μ on $E = (E,\mathcal{E})$ is B-regular if and only if (i) μ is complete, (ii) $\mathcal{D}(\mu) \supset \mathcal{E}$ and (iii) for every $A \in \mathcal{D}(\mu)$ there exists a subset B of A such that $B \in \mathcal{E}$ and $\mu(B) = \mu(A)$. A Borel space $E = (E,\mathcal{E})$ endowed with a B-regular probability measure μ is called a <u>B-regular probability</u>

space (E,μ) or $((E,\mathcal{E}),\mu)$.

A subset A of $E = (E,\mathcal{E})$ is called universally measurable if A is μ-measurable for every B-regular probability measure μ on $(E,\mathcal{E})$. The class $\mathcal{M}(E,\mathcal{E})$ of all universally measurable subsets of $(E,\mathcal{E})$ is a σ-algebra on E containing $\mathcal{E}$. A map $f : (E,\mathcal{E}) \to (F,\mathcal{F})$ is called universally measurable if it is measurable $\mathcal{M}(E,\mathcal{E})/\mathcal{F}$.

Theorem 2. Let $(E,\mathcal{E})$ be analytic. Then every analytic subset of $(E,\mathcal{E})$ is universally measurable.

Proof.
$$\mathcal{A}(E,\mathcal{E}) = \alpha[\mathcal{E}] \quad \text{(Theorem 8.3)}$$
$$\subset \mathcal{M}(E,\mathcal{E}) \quad \text{(Theorem 1)}.$$
∎

A map $f : S = (S,\mu) \to F = (F,\mathcal{F})$ is called μ-measurable if f is measurable $\mathcal{D}(\mu)/\mathcal{F}$. Then the image measure $f\mu$ is complete and

$$\mathcal{D}(f\mu) \supset \mathcal{F}.$$

Hence $f\mu$ is an extension of $\overline{f\mu|_{\mathcal{F}}}$ but these two measures are different in general, i.e. $f\mu$ is not always B-regular, as the following example shows.

Let λ be the Lebesgue measure on $[0,1]$ and S the well-known example of a non-measurable subset of $[0,1]$. Then it is easy to see that

$$0 = \lambda_*(S) < \lambda^*(S) = 1.$$

Define a probability measure μ on S by

$$\mathcal{D}(\mu) = \mathcal{D}(\lambda) \cap S \quad \text{and} \quad \mu(A) = \lambda^*(A).$$

It is easy to check that μ is a complete probability measure. Let $f : S \to [0,1]$ be the canonical injection. Then f is μ-measurable. But the image measure $\nu := f\mu$ is not B-regular. To check this, observe that $\nu = \lambda$ on $B[0,1]$ but $S \in \mathcal{D}(\nu) \setminus \mathcal{D}(\lambda)$; if ν were B-regular, then ν should coincide with λ.

<u>Theorem 3</u>. Let μ be a B-regular probability measure on $E = (E, \mathcal{E})$ and let $F = (F, \mathcal{F})$ be an analytic Borel space. Then every μ-measurable map $f : E \to F$ has the following properties.

(i) There exists a Borel map $g : (E, \mathcal{E}) \to (F, \mathcal{F})$ such that

$$f(x) = g(x) \quad \text{a.e.}(\mu), \text{ i.e. } \quad \mu\{x \in E : f(x) \neq g(x)\} = 0.$$

(ii) The image measure $f\mu$ is B-regular.

<u>Proof</u>.

(i) The assertion is well-known in the special case where $F \subset [0,1]$, from which the general case follows at once because every analytic Borel space is Borel isomorphic to an analytic subset of $[0,1]$.

(ii) If $f(x) = g(x)$ a.e.(μ), then it is obvious that $f\mu = g\mu$. Hence the assertion (i) ensures that we can assume without loss of generality that f is a Borel map from $(E, \mathcal{E})$ into $(F, \mathcal{F})$. Let $\nu := \overline{f\mu|_{\mathcal{F}}}$. Then $f\mu$ is an extension of ν . Let $A \in \mathcal{B}(f\mu)$. Then $f^{-1}(A) \in \mathcal{B}(\mu)$, so we have

$$f^{-1}(A) \supset E_1 \quad \text{and} \quad \mu(f^{-1}(A)) = \mu(E_1) \quad \text{for some } E_1 \in \mathcal{E} .$$

Since $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ is Borel and since both $(E, \mathcal{E})$ and $(F, \mathcal{F})$ are analytic,

$$f(E_1) \in \mathcal{A}(F, \mathcal{F}) \subset \mathcal{B}(\nu) \qquad \text{(Theorem 2)}.$$

Hence we have

$$f(E_1) \supset F_1 \quad \text{and} \quad \nu(f(E_1)) = \nu(F_1) \quad \text{for some } F_1 \in \mathcal{F} .$$

Therefore $A \supset f(E_1) \supset F_1$ and

$$(f\mu)(A) = \mu(f^{-1}(A)) = \mu(E_1) \leq \mu(f^{-1}(f(E_1))) = (f\mu)(f(E_1))$$

$$= \nu(f(E_1)) = \nu(F_1) = (f\mu)(F_1) \leq (f\mu)(A),$$

so we have

$$A \supset F_1 \in \mathcal{F} \quad \text{and} \quad (f\mu)(A) = (f\mu)(F_1),$$

proving that $f\mu$ is B-regular. ∎

Since a topological space is regarded as a Borel space with the topological σ-algebra, we can talk about B-regular measures on a topological space. Let S be a Hausdorff topological space, and let μ be a B-regular probability measure on S. A subset A of S is said to have inner K-regularity (with respect to μ) if

$$\mu(A) = \sup_{\substack{K \subset A \\ K:\text{compact}}} \mu(K) .$$

If every μ-measurable set has inner K-regularity, then μ is called a <u>K-regular measure</u> on S. This condition is equivalent to the condition that every Borel set has inner K-ragularity.

<u>Lemma 1.</u> Let μ be a B-regular probability measure on S and suppose that every open subset of S is expressible as a countable union of closed subsets; for example, every metrizable space has this property. Then μ is K-regular if every closed subset has inner K-regularity.

<u>Proof</u>. It is obvious that both open sets and closed sets have inner K-regularity. Since

$$\mu(\bigcup_n A_n \setminus \bigcup_n B_n) \leq \sum_n \mu(A_n \setminus B_n)$$

and

$$\mu(\bigcap_n A_n \setminus \bigcap_n B_n) \leq \sum_n \mu(A_n \setminus B_n) ,$$

we can easily check that the class of all sets having inner K-regularity is closed under countable unions and countable intersections. Hence every Borel set has inner K-regularity by virtue of Theorem 1.2. ∎

<u>Theorem 4.</u> Every B-regular probability measure μ on an analytic space S is K-regular.

<u>Proof</u>. First we will discuss the special case where S is a complete separable metric space with metric ρ. By virtue of Lemma 1 it is enough to show that every closed

subset has inner K-regularity. Let $\{a_n\}$ be a countable dense subset of S and let

$$B_{nk} := \overline{U}(a_n, 2^{-k}), \quad n,k = 1,2,\cdots.$$

Since $S = \bigcup_n B_{nk}$ for every k, we can find $N(k)$ such that

$$\mu(S-F_k) < 2^{-k} \quad \text{where} \quad F_k = \bigcup_{n=1}^{N(k)} B_{nk}.$$

Let

$$K_m := \bigcap_{k=m}^{\infty} F_k .$$

Since F_k has a finite 2^{-k+1}-covering, K_m has a finite 2^{-k+1}-covering for every $k \geq m$. Hence K_m is totally bounded. It is obvious that K_m is closed. Therefore K_m is compact. Also

$$\mu(S-K_m) \leq \sum_{k=m}^{\infty} \mu(S-F_k) < 2^{-m+1}.$$

If F is an arbitrary closed set, then $K_m \cap F$ is compact and

$$\mu(F - K_m \cap F) = \mu(F \cap (S-K_m)) \leq \mu(S-K_m) < 2^{-m+1},$$

so F has inner K-regularity. This proves that every B-regular measure on a complete separable metric space (or on a Polish space) is K-regular.

Now consider the general case. Take a Polish space P and a continuous surjection $f:P \to S$. Then there exists an inverse map $g:S \to P$ measurable $\sigma[\mathcal{A}(S)]/\mathcal{B}(P)$ (Theorem 6.6). Since $\mathcal{A}(S) = \alpha[\mathcal{B}(S)] \subset \mathcal{D}(\mu)$ (Theorems 5.7 (ii) and 1), g is μ-measurable. Hence Theorem 3 (ii) ensures that the image measure $\nu := g\mu$ is a B-regular probability measure on P, so ν is K-regular. Let $B \in \mathcal{B}(S)$. Then $f^{-1}(B) \in \mathcal{B}(P)$ by continuity of f. Since ν is K-regular, we can find compact sets $K_n \subset f^{-1}(B)$, $n = 1,2,\cdots,$ such that

$$\nu(f^{-1}(B) - K_n) < 2^{-n}.$$

Since $f \circ g : S \to S$ is the identity map,

$$f\nu = f(g\mu) = (f \circ g)\mu = \mu.$$

Since K_n is compact, $f(K_n)$ is also compact and

$$\mu(B - f(K_n)) = \nu(f^{-1}(B) - f^{-1}(f(K_n)))$$

$$\leq \nu(f^{-1}(B) - K_n) < 2^{-n}.$$

Hence every set $B \in \mathcal{B}(S)$ has inner K-regularity. ∎

<u>Theorem 5</u> (<u>The generalized Lusin theorem</u>). Let $f:(S, \mu) \to T$ be μ-measurable, where both S and T are analytic spaces and μ is a B-regular probability measure on S. For every μ-measurable subset A of S and for every $\varepsilon > 0$ we can find a compact subset $K = K(\varepsilon)$ of A such that the restriction $f|_K : K \to T$ is continuous.

<u>Proof.</u> First we consider the special case where $A = S$ and T is a complete separable metric space with metric ρ. Then T has the following decomposition for each $k = 1,2,\cdots$:

$$B = \sum_n B_{nk} \quad \text{where} \quad \delta(B_{nk}) < 2^{-k},\ n = 1,2,\cdots.$$

Let $A_{nk} := f^{-1}(B_{nk})$. Then

$$S = \sum_n A_{nk}, \quad k = 1,2,\cdots,$$

so $\mu(S - \sum_{n=1}^{N(k)} A_{nk}) < 2^{-k-1}\varepsilon$ for some $N(k)$. Using the last theorem, we can find a compact subset K_{nk} of A_{nk} such that

$$\mu(A_{nk} - K_{nk}) < 2^{-k-1}\varepsilon/N(k).$$

Then

$$\mu(S - \sum_{n=1}^{N(k)} K_{nk}) < 2^{-k}\varepsilon.$$

Let

$$K := \bigcap_k \sum_{n=1}^{N(k)} K_{nk}$$

and define a sequence of maps $f_k : K \to T$, $k = 1,2,\cdots$, as follows:

$$f_k(x) \equiv b_{nk} \quad \text{on} \quad K \cap K_{nk} \quad \text{if this set is non-empty,}$$

where b_{nk} is any point in $f(K \cap K_{nk})$. Since $K \cap K_{nk}$, $k = 1,2,\cdots,n$ are disjoint compact sets, f_k is continuous for every k. Since for every $x \in K \cap K_{nk}$ $(\neq \phi)$ we have

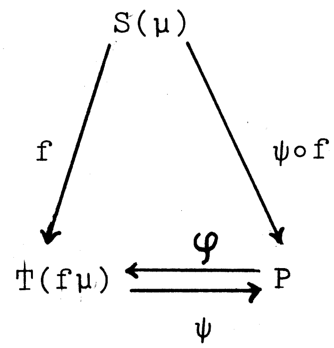
$$\rho(f_k(x),f(x)) \leq \delta_\rho(f(K \cap K_{nk})) \leq \delta_\rho(f(A_{nk})) \leq \delta_\rho(B_{nk}) < 2^{-k},$$

$f_k(x)$ converges to $f(x)$ uniformly on K. Hence the restriction $f|_K$ is continuous. Also

$$\mu(A-K) \leq \sum_k \mu(A-\sum_{n=1}^{N(k)} K_{nk}) < \varepsilon.$$

Thus our theorem is proved in case T is a complete separable metric space (or a Polish space) and $A = S$.

Now we consider the general case. Take a Polish space P and a continuous map $\varphi : P \to T$. Using Theorem 6.6 we can find a $f\mu$-measurable inverse map ψ of φ. Since f is measurable $\mathcal{B}(\mu)/\mathcal{B}(f\mu)$, the composite map $\psi \circ f : S \to P$ is μ-measurable. Since P is Polish, we can use the result proved above to find a compact subset H of S such that



$$\mu(S-H) < \frac{\varepsilon}{2} \quad \text{and} \quad \psi \circ f|_H \text{ is continuous.}$$

Since $\varphi : P \to T$ is continuous, $\varphi \circ (\psi \circ f|_H)$ is continuous. This means that $f|_H$ is continuous, because ψ is an inverse map of φ. Since μ is K-regular, every μ-measurable set A has a compact subset J such that $\mu(A-J) < \varepsilon/2$. Let $K := H \cap J$. Then $f|_K$ is continuous and

$$\mu(A-K) \leq \mu(A-H) + \mu(A-J) < \varepsilon. \quad \blacksquare$$

10. Standard probability spaces.

A probability measure μ on S is called standard if we can find a σ-algebra $\mathcal{S}$ on S and a subset S_1 with μ-measure 1 satisfying the following two conditions:

(S.1) μ is a B-regular probability measure on $(S,\mathcal{S})$,

(S.2) $(S_1, \mathcal{S} \cap S_1)$ is a standard Borel space.

A probability space (S,μ) is called standard, if μ is standard. A standard probability space is often called a probability space of type L or an L-space.

Every B-regular probability measure on a standard Borel space is obviously standard. More generally we have

Theorem 1. Every B-regular probability measure μ on an analytic Borel space $(E,\mathcal{E})$ is standard.

Proof. Since $E = (E,\mathcal{E})$ is an analytic Borel space, E is Borel isomorphic to an analytic subset of $[0,1]$. Hence we can assume without loss of generality that E is an analytic subset of $[\ ,1]$ and that $\mathcal{E} = \mathcal{B}[0,1] \cap E$. Let $i : E \to [0,1]$ be the canonical injection. Then the image measure $\nu := i\mu$ is B-regular (Theorem 9.3(ii)). It is obvious that

$$\nu(E) = \mu(i^{-1}(E)) = \mu(E) = 1.$$

Hence there exists a Borel subset B of $[0,1]$ such that

$$\mathcal{B} \subset E \quad \text{and} \quad \nu(B) = \nu(E) = 1.$$

Being a Borel subset of $[0,1]$, B is a standard space, so $(B, \mathcal{B}(B))$ is a standard Borel space. Since

1.10

$$\mathcal{B}(B) = \mathcal{B}([0,1]) \cap B = \mathcal{B}([0,1]) \cap E \cap B = \mathcal{E} \cap B,$$

$(B, \mathcal{E} \cap B)$ is a standard Borel space. Also

$$\mu(B) = \mu(i^{-1}(B)) = \nu(B) = 1 \qquad \blacksquare$$

Since the above proof works in case A is universally measurable, we have

<u>Theorem 2</u>. Let $(E,\mathcal{E})$ be a Borel space Borel isomorphic to a universally measurable subset of $[0,1]$. Then every B-regular probability measure μ on $(E,\mathcal{E})$ is standard.

<u>Theorem 3</u>. Let (S,μ) be a standard probability space and $(E,\mathcal{E})$ an analytic space. If $f : S \to E$ is μ-measurable, then the image measure $\nu := \varphi\mu$ is B-regular and (E,ν) is standard.

<u>Proof</u>. Take a σ-algebra $\mathcal{S}$ on S and a subset S_1 of S with μ-measure 1 satisfying (S.1) and (S.2). Then $S_1 = (S_1, \mathcal{S} \cap S_1)$ is a standard Borel space. It is easy to check that the restriction $\mu_1 := \mu|_{\mathcal{B}(\mu) \cap S_1}$ is a B-regular probability measure on $(S_1, \mathcal{S} \cap S_1)$. Also the restriction $\varphi_1 : \varphi|_{S_1} : S_1 \to E$ is μ_1-measurable. Hence the image measure $\varphi_1\mu_1$ is a B-regular measure on $(E,\mathcal{E})$. Since $\mu(S-S_1) = 0$, it is easy to check that $\varphi_1\mu_1 = \varphi\mu = \nu$. Hence ν is B-regular (Theorem 9.3(ii)), so (E,ν) is standard (Theorem 1). $\blacksquare$

Let (S,μ) and (T,ν) be probability spaces where μ and ν are complete. If there exists a bijective map $\varphi : S \to T$ such that

$$\varphi(\mathcal{B}(\mu)) = \mathcal{B}(\nu) \quad \text{and} \quad \mu(A) = \nu(\varphi(A)),$$

then (S,μ) is called <u>strictly isomorphic</u> to (T,ν), $(S,\mu) \approx (T,\nu)$ in notation. More generally, if we can find a subset S_1 with $\mu(S_1) = 1$ and a subset T_1 of T with $\nu(T_1) = 1$ such that

$$(S_1, \mu|_{\mathcal{B}(\mu) \cap S_1}) \approx (T_1, \nu|_{\mathcal{B}(\nu) \cap T_1}),$$

then (S,μ) is called <u>isomorphic</u> to (T,ν), $(S,\mu) \sim (T,\nu)$ in notation. Both $\approx$ and $\sim$ are equivalence relations.

Let μ be a B-regular probability measure on $(E,\mathcal{E})$ and let $f : (E,\mathcal{E}) \to (F,\mathcal{F})$ be bimeasurable. Then it is easy to check that $(E,\mu) \approx (F,f\mu)$.

<u>Theorem 4</u>. Every standard probability space is isomorphic to a probability space $[0,1]$ endowed with a B-regular probability measure.

<u>Proof</u>. Let (S,μ) be standard. Take a σ-algebra $\mathcal{S}$ on S and a subset S_1 of S with μ-measure 1 satisfying (S.1) and (S.2). Let μ_1 denote the restriction of μ to $\mathcal{B}(\mu) \cap S_1$. Then μ_1 is a B-regular probability measure on $S_1 = (S_1, \mathcal{S} \cap S_1)$. Since Theorem 6.5 ensures the existence of a bimeasurable map $f : S_1 \to B$ where $B \in \mathcal{B}([0,1])$, $(S_1,\mu_1) \approx (B, f\mu_1)$. But $f\mu_1$ can be extended to a B-regular probability measure ν on $[0,1]$ such that $\mathcal{B}(f\mu_1) = \mathcal{B}(\nu) \cap S_1$. Hence $(S,\mu) \sim ([0,1],\nu)$. ∎

Chapter 2. General concepts of probability theory

1. Sample spaces, events and random variables.

In the modern theory of probability we take a probability space (Ω,P), define random variables to be P-measurable functions and formulate all probabilistic facts such as independence, conditional probabilities, expectations etc. in terms of measures and integrals. This idea goes back to E. Borel [ ]. Also N. Wiener used measure theory to discuss Brownian motion [ ]. But Kolmogorov's celebrated work [ ]:

Grundbegriffe der Wahrscheinlichkeitsrechnung (1933) is the first systematic theory of probability presented in the framework of measure theory.

In application Ω represents the set of all possible outcomes of the random phenomenon in observation and $P(A)$ is the probability that the observed autcome be in the set A. Hence Ω may be a finite set, a countable set, $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^\infty$ or a function space according to the nature of the random phenomenon in consideration.

In this book we assume that

(A) (Ω,P) is a standard probability space. This assumption enables us to establish probability theory in a more natural way. Also (A) is not too strong, because all probability spaces useful in application satisfy (A), as we have seen in Chapter 1.

Let $S = (S,\mathcal{B})$ be a standard Borel space. A P-measurable (i.e. measurable $\mathcal{B}(P)/\mathcal{B}$) map $X : \Omega \to S$ is called an S-valued random variable. It is customary to denote by ω a generic point of Ω and an S-valued random variable $X : \Omega \to S$ by $X(\omega)$. The space $S = (S, \mathcal{S})$ is called the sample space of an S-valued random variable X.

Since every standard space T is regarded as a standard Borel space with the topological σ-algebra $\mathcal{B}(T)$, we can talk about T-valued random variables.

An S-valued random variable is called

a real random variable if $S = \mathbb{R}$,

a random vector if $S = \mathbb{R}^n$ $(n = 1,2,\ldots)$,

a random sequence if $S = \mathbb{R}^\infty$,

a random continuous function on $[0,1]$ if $S = C[0,1]$,

a random L^2 function on $[0,1]$ if $S = L^2[0,1]$,

a random distribution on $[0,1]$ if $S = \mathcal{D}'[0,1]$,

and so on.

Let $\alpha = \alpha(\omega)$ be a condition concerning a generic point $\omega \in \Omega = (\Omega, P)$. In probability theory it is called an event. The probability (of occurrence) of α is defined to be the P-measure of the set of all $\omega \in \Omega$ for which $\alpha(\omega)$ holds, if this set is P-measurable. Hence the probability of α is equal to $P(\{\omega : \alpha(\omega)\})$, which is simply denoted by $P(\alpha)$. If $P(\alpha) = 1$, i.e. $\alpha(\omega)$ a.e.(P), we say that α occurs almost surely, α a.s. in notation. In view of the well-known relation between conditions and their extensions (the set $\{\omega : \alpha(\omega)\}$ being called the extension of α in logics) we can reduce the properties of probabilities to those of P-measures. Let α_n, $n = 1,2,\ldots$ be a sequence of events. The event that $\alpha_n(\omega)$ holds for infinitely many n's is denoted by

$$\alpha_n \text{ i.o.} \quad (\text{i.o.} = \underline{\text{infinitely often}})$$

and the event that $\alpha_n(\omega)$ holds except for finitely many n's by

$$\alpha_n \text{ f.e.} \quad (\text{f.e.} = \underline{\text{with finite exceptions}}).$$

We obviously obtain

$$\{\omega : \alpha_n(\omega) \text{ i.o.}\} = \overline{\lim_{n\to\infty}} \{\omega : \alpha_n(\omega)\}$$

and

$$\{\omega : \alpha_n(\omega) \text{ f.e.}\} = \underline{\lim_{n\to\infty}} \{\omega : \alpha_n(\omega)\} .$$

The well-known Borel-Cantelli lemma claims that if $\sum_n P(\alpha_n) < \infty$, then $P(\alpha_n \text{ i.o.}) = 0$, i.e. $P(\neg\alpha_n \text{ f.e.}) = 1$, $\neg\alpha_n$ denoting the negation of α . Denoting $\{\omega : \alpha_n(\omega)\}$ by A_n , we can reduce this lemma to a measure-theoretical lemma that

$$\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(\overline{\lim_{n\to\infty}} A_n) = 0 \Leftrightarrow P(\underline{\lim_{n\to\infty}} A_n^c) = 1.$$

Let $X(\omega)$ be an S-valued random variable, where $S = (S, \mathcal{B})$ is standard. The image measure XP on S is called the $\underline{\text{probability law}}$ of X, denoted by P^X . Then

(1) $\{\omega : X(\omega) \in A\} \in \mathcal{B}(P) \Leftrightarrow A \in \mathcal{B}(P^X) \Rightarrow P(X \in A) = P^X(A)$,

which justifies the definition. Immediately from the definition we obtain

(2) $X : \Omega \to S$ is measurable $\mathcal{B}(P) / \mathcal{B}(P^X)$,

(3) $X(\Omega) \in \mathcal{B}(P^X)$ and $P^X(X(\Omega)) = 1$.

<u>Theorem 1</u>. P^X is a B-regular probability measure on $(S, \mathcal{S})$, and (S, P^X) is a standard probability space.

<u>Proof</u>. Immediate from Theorems 1.9.3(ii) and 1.10.1.

Two S-valued random variables are called <u>equivalent</u> if they are equal almost surely. Equivalence in this sense is an equivalence relation. Equivalent random variables have the same probability law, but not conversely.

Let $S = (S, \mathcal{S})$ and $T = (T, \mathcal{T})$ be standard Borel spaces. If $S \subset T$ and $\mathcal{S} \supset \mathcal{T} \cap T$, then the canonical injection $i : S \to T$ is Borel, so $S = i(S) \in \mathcal{S}$ (Theorem 1.8.4) and $\mathcal{S} = \mathcal{T} \cap T$ (Theorem 1.8.5). Let X be an S-valued random variable. Then it is obvious that $Y := i \circ X$ is a T-valued random variable. But

$$X(\omega) = Y(\omega) \qquad \text{for every} \quad \omega \in \Omega.$$

Hence every S-valued random variable X is regarded as a T-valued random variable $i \circ X$. In this sense we regard real random variables as complex random variables and random continuous functions on $[0,1]$ as random L^2 functions on $[0,1]$. This convention is commonly used without mentioning.

From now on $S, T, U \cdots$ stand for standard Borel spaces where the endowed σ-algebras are denoted by the corresponding script letters $\mathcal{S}, \mathcal{T}, \mathcal{U}, \cdots$

Let X be an S-valued random variable. Then (S,P^X) is a standard probability space. Let $f : S \to T$ be P^X-measurable. Then the map

$$Y := f \circ X : \Omega \to T$$

defines a T-valued random variable with $P^X = fP^X$. It is obvious that

$$(C) \qquad X(\omega_1) = X(\omega_2) \Rightarrow Y(\omega_1) = Y(\omega_2).$$

i.e. the value of Y is completely determined by that of X. Conversely we have

Theorem 2. Let X be an S-valued random variable. Then every T-valued random variable Y whose value is completely determined by the value of X is expressible as

$$Y = f \circ X \text{ , where } f : S \to T \text{ is } P^X\text{-measurable.}$$

Such a map f is uniquely determined on $S_1 := X(\Omega)$, where $P^X(S_1) = 1$.

Proof. Since X and Y satisfy (C), there exists a unique map $f_1 : S_1 \to T$ such that

$$f_1(x) = y \quad \text{if} \quad x = X(\omega) \quad \text{and} \quad y = Y(\omega) \quad \text{for some} \quad \omega \in \Omega.$$

It is obvious that

$$f_1(X(\omega)) = Y(\omega) \text{ , i.e. } f_1 \circ X = Y.$$

Let $f : S \to T$ be any extension of the map $f_1 : S_1 \to T$. Since $X(\Omega) = S_1$, $f \circ X = Y$. Hence

$$X^{-1}(f^{-1}(B)) = (f \circ X)^{-1}(B) = Y^{-1}(B) \in \mathcal{B}(P) \quad \text{for } B \in \mathcal{T}$$

i.e.. $\quad f^{-1}(B) \in \mathcal{B}(P^X)$ for $B \in \mathcal{T}$.

Hence f is a P^X-measurable map satisfying $Y = f \circ X$. Every such map must coincide with the map f_1 on S_1. ∎

Let Z be a real random variable. Then the integral

$$\int_A Z(\omega) P(d\omega)$$

is denoted by $E(Z,A)$, if it is well-defined. $E(Z,\Omega)$ is called the expectation of Z, denoted by $E(Z)$. Let X be an S-valued random variable. If $Z = f \circ X$, then

$$E(Z, f^{-1}(B)) = \int_B f(x) P^X(dx) \quad \text{and} \quad E(Z) = \int_S f(x) P^X(dx).$$

Remark 1. If we take a general probability space (Ω, P), then Theorem 1 does not hold even in case X is a real random variable. This was pointed out by Kolmogorov [ ]. To get rid of the trouble he assumed P to be perfect. This assumption is weaker than our assumption (A). Similarly for Theorem 2.

Remark 2. We may similarly define S-valued random variables in case S is an analytic Borel space; then both Theorems 1 and 2 also hold. But we will not consider such random variables in this book.

Remark 3. Let (Ω, P) be a probability space and N a

P-null set, i.e. a subset of Ω with P-measure 0 . Then the probabilistic nature of every random variable X on (Ω,P) is the same as that of the restriction of X to $\Omega - N$ on the probability space $(\Omega - N, P|_{\mathcal{D}(P) \cap (\Omega - N)})$.

~~Hence we can remove any P-null set from (Ω,P) without any essential charge of the results~~ Hence we can assume without loss of generality that Ω is a standard Borel space and P is a B-regular probability measure on Ω .

<u>Remark 4</u>. In many cases there is a random variable $X(\omega)$ with values in a standard space S such that we are only concerned with the random variables whose values are completely determined by $X(\omega)$. Since such random variables are expressible in the form $f(X)$ (f: P^X-measurable), they are regarded as random variables on the probability space (S, P^X) . This observation ensures that in many cases we can assume that Ω is a standard space and P is a B-regular measure on Ω .

The set $\mathcal{L}^0 = \mathcal{L}^0(\Omega,P)$ of all real random variables on (Ω,P) is a complete separable metric space with metric

$$\rho_0(X,Y) = E(|X-Y| \wedge 1),$$

where two equivalent functions are identified in $\mathcal{L}^0$. The ρ_0-topology coincides with the topology of convergence in probabilty , because

$$P(|X-Y| > \varepsilon) \leq \rho_0(X,Y) \leq P(|X-Y| > \varepsilon) + \varepsilon.$$

2. Joint random variables and extension theorems

Let $S_n = (S_n, \mathcal{S}_n)$, $n = 1,2,\cdots$, be standard Borel spaces. Then their Borel product

$$(S, \mathcal{S}) := (\prod_n S_n, \prod_n \mathcal{S}_n)$$

is also a standard Borel space. If X_n is an S_n-valued random variable on (Ω, P) for $n = 1,2,\cdots$, then

(1) $$X(\omega) = (X_1(\omega), X_2(\omega),\cdots)$$

defines an S-valued random variable on (Ω, P), called the joint (random) variable of $X_1, X_2,\cdots$. In fact the map $X:\Omega \to S$ defined by (1) is P-measurable, being the product map of $X_1, X_2,\cdots$ (Section 1.1). The probability law of $X = (X_1, X_2,\cdots)$ is called the joint probability law of $X_1, X_2,\cdots$.

If $S_1, S_2,\cdots$ are standard spaces, then the topological product $S = \prod_n S_n$ is also a standard space. In view of $\mathcal{B}(S) = \prod_n \mathcal{B}(S_n)$ (Theorem 1.4.5) we have that if X_n is an S_n-valued random viriable for each n, then the joint variable X of $X_1, X_2, \cdots$ is an S-valued random variable.

The joint variable of finitely many real random variables is a random vector and the joint variable of a sequence of real random variables is a random sequence.

It is obvious that

$$X_n = Y_n \text{ a.s. for each } n$$
$$\Rightarrow (X_1,X_2,\cdots) = (Y_1,Y_2,\cdots) \text{ a.s.} \Rightarrow P^{(X_1,X_2,\cdots)} = P^{(Y_1,Y_2,\cdots)}$$

Theorem 1. Let X and Y be S-valued random variables where $S = (S, \mathcal{S})$ is standard. Then

$$\{\omega: X(\omega) = Y(\omega)\} \in \mathcal{B}(P)$$

Proof. The joint variable (X, Y) is an $S \times S$-valued random variable and the above ω-set is $\{\omega: (X(\omega), Y(\omega)) \in \Delta\}$, where $\Delta = \{(x, y) \in S \times S: x = y\} \in \mathcal{S} \times \mathcal{S}$ (Theorem 1.8.7).∎

~~Since the Borel product of uncountably many standard Borel spaces is no longer standard in general, we cannot always define the joint variable of uncountably many random variables. However, there are some cases where it can be defined in a modified sense (the next section).~~

Let $S_n = (S_n, \mathcal{S}_n)$ be a sequence of standard Borel spaces. We consider the Borel products

$$T_n = (T_n, \mathcal{T}_n) := (\prod_{k=1}^{n} S_k, \prod_{k=1}^{n} \mathcal{S}_k), \quad n = 1,2,\cdots,\infty$$

and the following projections

$$\pi_{nm}: T_m \to S_n, \quad (x_1,x_2,\cdots x_m) \mapsto x_n,$$

$$p_{nm}: T_m \to T_n, \quad (x_1,x_2,\cdots,x_m) \mapsto (x_1,x_2,\cdots,x_n),$$

where $1 \le n \le m \le \infty$ and if $m = \infty$, $(x_1,x_2,\cdots x_m)$ should be replaced by $(x_1,x_2,\cdots)$. Every T_n is a standard Borel space and

$$p_{nm} \circ p_{m\ell} = p_{n\ell}$$

$$\pi_{nm} \circ p_{m\ell} = \pi_{n\ell}$$

where $1 \leq n \leq m < \ell \leq \infty$. We keep using this notation.

Suppose that X_n is an S_n-valued random variable on (Ω, P) for $n = 1,2,\cdots$ and let Y_n denote the joint variable of $X_1,X_2,\cdots,X_n$ for each n. Then the probability laws $\mu_n := P^{Y_n}$, $n = 1,2,\cdots$, satisfy the consistency condition:

$$(\mathcal{C}) \qquad \mu_n = p_{nm}\mu_m, \; n \leq m < \infty,$$

because

$$\mu_n = Y_n P = (p_{nm} \circ Y_m)P = p_{nm}(Y_m P) = p_{nm}\mu_m,$$

where $1 \leq n \leq m < \infty$.

Theorem 2. The joint probability law μ of $Y_\infty := (X_1,X_2,\cdots)$ is completely determined by $\mu_1, \mu_2,\cdots$.

Proof. Let $X_1', X_2',\cdots$ be another sequence of random variables on another probability space (Ω', P') such that the joint probability law of $X_1', X_2',\cdots,X_n'$ is μ_n for every n. We want to prove that the joint probability law μ' of $X_1',X_2',\cdots$ is equal to μ. Let $\mathcal{B}_n := p_{n\infty}^{-1}(\mathcal{T}_n)$, $n = 1,2,\cdots$. Then $\mathcal{B}_n$ is a σ-algebra on T_∞ and the union $\mathcal{A} := \bigcup_n \mathcal{B}_n$ generates the σ-algebra $\mathcal{T}_\infty$. Let Y_n' denote the joint variable of $X_1',X_2',\cdots,X_n'$ for every $n = 1,2,\cdots$, and Y' the joint variable of $X_1',X_2',\cdots$. Since $(\psi \circ \varphi)\nu = \psi(\varphi \nu)$, we have

$$p_{n\infty}\mu' = p_{n\infty}(Y'P) = (p_{n\infty} \circ Y')P = Y_n'P = \mu_n$$

and similarly

$$p_{n\infty}\mu = \mu_n,$$

so $$p_{n\infty}\mu' = p_{n\infty}\mu,$$

i.e $$\mu'(p_{n\infty}^{-1}(E_n)) = \mu(p_{n\infty}^{-1}(E_n)) \quad \text{for} \quad E_n \in \mathcal{T}_n.$$

This implies that

$$\mu' = \mu \quad \text{on} \quad \mathcal{B}_n, \; n = 1,2,\cdots,$$

so $$\mu' = \mu \quad \text{on} \quad \mathfrak{A}.$$

Let $\mathcal{B}$ be the class of all $B \in \mathcal{T}_\infty$ for which $\mu'(B) = \mu(B)$. Then $\mathcal{B}$ is a Dynkin class on T_∞ containing $\mathfrak{A}$. Since $\mathfrak{A}$ is multiplicative, the Dynkin class theorem ensures that $\mathcal{B} \supset \sigma[\mathfrak{A}] = \mathcal{T}_\infty$. This implies that $\mu' = \mu$ on $\mathcal{T}_\infty$, so $\mu' = \mu$ by virtue of B-regularity of μ and μ'. ∎

Theorem 3. (Kolmogorov's extension theorem). Let μ_n be a B-regular probability measure on $T_n (= \Pi_{k=1}^n S_k)$ for $n = 1,2,\cdots$, where $S_1, S_2, \cdots$ are standard Borel spaces. If $\{\mu_n\}$ satisfies the consistency condition (C), then we can construct a standard probability space (Ω, P) and a sequence of random variables $X_1, X_2, \cdots$ on (Ω, P) (each X_n being S_n-valued) so that μ_n is the joint probability law of $X_1, X_2, \cdots, X_n$ for every n.

Proof. It is enough to find a B-regular probability measure P on T_∞ such that

$$\mu_n = p_{n\infty}P, \quad n = 1,2,\cdots;$$

if this is done, $\Omega = T_\infty$, P and $X_n = \pi_{n\infty}$ will be what we want to construct.

First consider the special case where S_n is a compact subset of $[0, 1]$ for every n. Then T_∞ is a compact metrizable space with the product topology, because $T_\infty = \Pi_{n=1}^\infty S_n$. A <u>tame function</u> g on T_∞ is defined to be a real function of the form $g = g_n \circ p_{n\infty}$ where $g_n : T_n \to \mathbb{R}$ is Borel. The family F of all bounded tame functions on T_∞ forms a normed vector space with the usual linear operation and the supremum norm $\| \ \|_\infty$. The completion $\overline{F}$ of F is a Banach space consisting of all real functions f on T_∞ such that

$$\|f_n - f\|_\infty \to 0 \quad \text{for some sequence } \{f_n\} \subset F$$

Let C be the family of all continuous real functions on T_∞. We claim that $C \subset \overline{F}$. Let $f \in C$. Since T_∞ is compact, for every $\varepsilon > 0$ we can choose a finite number of neighborhoods $U_1, U_2, \cdots, U_\alpha$ from the usual open base of the product topology on T_∞ such that

$$\sup_{\omega_1, \omega_2 \in U_\alpha} |f(\omega_1) - f(\omega_2)| < \varepsilon, \ i = 1,2,\cdots,\alpha.$$

Let $E_i := U_i - \bigcup_{j<i} U_j$, $i = 1,2,\cdots,\alpha$, and define $g : T_\infty \to \mathbb{R}$ by

$$g(\omega) := \text{any fixed point } b_i \in f(E_i) \ \text{ if } \ \omega \in E_i.$$

Then

$$|g(\omega) - f(\omega)| < \varepsilon \quad \text{for every} \quad \omega \in \Omega.$$

Since each U_i is of the form $p_{n\infty}^{-1}(V_i)$ ($V_i \in \mathcal{T}_n$), we have

$$E_i = p_{N\infty}^{-1}(F_i) \quad (F_i \in \mathcal{T}_N),\ i = 1,2,\cdots,\alpha.$$

for a sufficiently large N independent of i. Hence g is expressed as

$$g = g_N \circ p_{N\infty} \quad \text{where} \quad g_N = \sum_{i:E_i \neq \phi} b_i 1_{F_i},$$

so $g \in F$. This proves that $C \subset \overline{F}$.

Now define a linear functional L on F by

$$L(f) = \int_{T_n} f_n d\mu_n \quad \text{if} \quad f = f_n \circ p_{n\infty}.$$

This is well-defined independently of the expression of f by virtue of the consistency condition (C). It is easy to check that

$$|L(f)| \leq \|f\|_\infty,\ L(f) \geq 0 \quad \text{for} \quad f \geq 0, \text{ and } \quad L(1) = 1.$$

Hence L can be extended to a linear functional on $\overline{F}$ with these properties. Since $C \subset \overline{F}$, we can use the Riesz representation of measures to prove the existence of a B-regular probability measure P on T_∞ such that

$$L(f) = \int_{T_\infty} f dP \quad \text{for every} \quad f \in C.$$

Let f_n be any continuous function on T_n. Then

$$f_n \circ p_{n\infty} \in C \cap T,$$

Hence

$$\int_{T_n} f_n d\mu_n = \int_{\Omega} (f_n \circ p_{n\infty}) dP = \int_{T_n} f_n d(p_{n\infty}P).$$

This proves that $\mu_n = p_{n\infty}P$.

Now we consider the general case. Since $S_n = (S_n, \mathcal{S}_n)$ is standard, Theorem 1.8.6 (ii) ensures that there exists a bimeasurable map from S_n to a compact subset S'_n of $[0, 1]$ for each n. Then the bilateral product map

$$\psi_n := \prod_{k=1}^{n}{}^{b} \varphi_k : T_n \to T'_n (:= \prod_{k=1}^{n} S'_k)$$

is also bimeasurable for each n. Let

$$\mu'_n := \psi_n \mu_n, \quad n < \infty.$$

Coresponding to $p_{nm}: T_m \to T_n$ we define $p'_{nm}: T'_m \to T'_n$. It is easy to check that

$$p'_{nm} = \psi_n \circ p_{nm} \circ \psi_m^{-1}, \quad n < m \leq \infty.$$

Since $\{\mu_n\}$ satisfies (C) and since $(\varphi \circ \psi)\nu = \varphi(\psi\nu)$,

$$p'_{nm}\mu'_m = (\psi_n \circ p_{nm} \circ \psi_m^{-1})(\psi_m \mu_m)$$

$$= \psi_n \mu_n = \mu'_n,$$

so we can find a regular probability measure P' on T'_∞ such that $\mu'_n = p'_{n\infty}P'$. Let $P := \psi_\infty^{-1}P'$. Then

$$p_{n\infty}P = p_{n\infty}(\psi_\infty^{-1}P') = (p_{n\infty} \circ \psi_\infty^{-1})P',$$

$$\mu_n = \psi_n^{-1}\mu'_n = \psi_n^{-1}(p'_{n\infty}P') = \psi_n^{-1}((\psi_n \circ p_{n\infty} \circ \psi_\infty^{-1})P')$$

$$= (p_{n\infty} \circ \psi_\infty^{-1})P'$$

and hence

$$p_{n\infty}P = \mu_n .$$ ∎

Let A be a countably infinite directed index set. We consider a family of standard Borel spaces $S_\alpha = (S_\alpha, \mathcal{S}_\alpha)$, $\alpha \in A$, and a family $\mathcal{F}$ of Borel maps $f_{\alpha\beta}: S_\beta \to S_\alpha$, $\alpha, \beta \in A$, $\alpha \leq \beta$, such that

$$f_{\alpha\beta} \circ f_{\beta\gamma} = f_{\alpha\gamma}, \ \alpha \leq \beta \leq \gamma, \text{ and } f_{\alpha\alpha} = \text{the identity map.}$$

Let $S = (S, \mathcal{S})$ be the Borel product of S_α, $\alpha \in A$ and $\pi_\alpha: S \to S_\alpha$, $\alpha \in A$, denote the canonical projections. The set

$$S' := \{x \in S: \pi_\alpha(x) = f_{\alpha\beta}(\pi_\beta(x)), \ \alpha < \beta\}$$

endowed with the trace σ-algebra $\mathcal{S}' := \mathcal{S} \cap S'$ is called the <u>projective</u> <u>limit</u> of S_α, $\alpha \in A$, relative to $\mathcal{F}$, denoted by

$$\overleftarrow{\lim}\, S_\alpha \quad \text{or} \quad \overleftarrow{\lim_{\mathcal{F}}}\, S_\alpha .$$

Being a countable Borel product of standard Borel spaces, $S = (S, \mathcal{S})$ is standard. Hence Theorem 1.8.7 ensures that

$$S'_{\alpha\beta} := \{x \in S: \pi_\alpha(x) = (f_{\alpha\beta} \circ \pi_\beta)(x)\} \in \mathcal{S}, \ \alpha \leq \beta,$$

so

$$S' = \bigcap_{\alpha \leq \beta} S'_{\alpha\beta} \in \mathcal{S} .$$

This implies that $S' = (S', \mathcal{S}')$ is a standard Borel space (Theorem 1.8.3 (i)). We keep using this notation below.

Let X_α be an S_α-valued random variable on (Ω, P) for every $\alpha \in A$, and suppose that they are related as follows:

(R) $X_\alpha(\omega) = f_{\alpha\beta}(X_\beta(\omega))$, $\omega \in \Omega$, $\alpha \leq \beta$.

Then it is obvious that the joint variable $X(\omega) :=$ $(x_\alpha(\omega), \alpha \in A)$ is an S'-valued random variable on (Ω, P).

Let ν_α be the probability law of X_α for $\alpha \in A$. Then it follows from (R) that the following consistency condition holds:

(C') $\nu_\alpha = f_{\alpha\beta}\nu_\beta$, $\alpha \leq \beta$.

Theorem 4. The probability law of X is completely determined by μ_α, $\alpha \in A$.

Proof. Similar to the proof of Theorem 1.

Theorem 5. (Bochner's extension theorem). Let ν_α be a B-regular probability measure on a standard Borel space S_α for each $\alpha \in A$. If $\{\nu_\alpha\}$ satisfies the consistency condition (C'), then we can construct a standard probability space (Ω, P) and a countable family of random variables X_α, $\alpha \in A$, on (Ω, P), each X_α being S_α-valued, so that

$$X_\alpha(\omega) = f_{\alpha\beta}(X_\beta(\omega)) \quad \text{and} \quad \nu_\alpha = P^{X_\alpha},$$

where $\alpha, \beta \in A$ and $\alpha \leq \beta$.

Proof. Since A is countable and directed, we can choose a sequence $\alpha_1 \leq \alpha_2 \leq \cdots$ in A such that for every $\alpha \in A$ we can find $\alpha_n \geq \alpha$. Denote S_{α_n}, ν_{α_n} and $f_{\alpha_n\alpha_m}$ by $\tilde{S}_n$,

$\tilde{\nu}_n$ and $\tilde{f}_{nm}$ respectively. Let

$$\tilde{T}_n := \prod_{k=1}^{n} \tilde{S}_n, \quad n = 1,2,\cdots,\infty$$

and

$$\tilde{p}_{nm} := \text{the canonical projection from } \tilde{T}_m \text{ to } \tilde{T}_n, \quad 1 \leq n \leq m \leq \infty.$$

Since the product map

$$\prod_{k=1}^{n} \tilde{f}_{kn} : \tilde{S}_n \to \tilde{T}_n$$

is Borel, the image measure on $\tilde{T}_n$ $(n < \infty)$:

$$\tilde{\mu}_n := \left(\prod_{k=1}^{n} \tilde{f}_{kn}\right)\tilde{\nu}_n$$

is B-regular for $n < \infty$.

We claim that $\{\tilde{\mu}_n\}$ satisfies the consistency condition:

$$\tilde{\mu}_n = \tilde{p}_{nm}\tilde{\mu}_m, \quad n \leq m < \infty.$$

Let $n \leq m$. Then

$$\tilde{\mu}_n = \left(\prod_{k=1}^{n} \tilde{f}_{kn}\right)\tilde{\nu}_n = \left[\left(\prod_{k=1}^{n} \tilde{f}_{kn}\right)\circ\tilde{f}_{nm}\right]\tilde{\nu}_m$$

$$= \left[\prod_{k=1}^{n} (\tilde{f}_{kn}\circ\tilde{f}_{nm})\right]\tilde{\nu}_m = \left(\prod_{k=1}^{n} \tilde{f}_{km}\right)\tilde{\nu}_m$$

$$= \left(\tilde{p}_{nm}\circ \prod_{k=1}^{m} \tilde{f}_{km}\right)\nu_m = \tilde{p}_{nm}\tilde{\mu}_m.$$

Using Kolmogorov's extension theorem, we can construct a standard probability space (Ω,P) and a sequence of random variables $\tilde{X}_n(\omega)$, $n = 1,2,\cdots$, on (Ω,P), each $\tilde{X}_n$

being $\tilde{S}_n$-valued, so that $\tilde{\mu}_n$ is the joint probability law of $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_n$ for every n.

It follows from the definition of $\tilde{\mu}_n$ that the joint probability law of the random variables

$$\tilde{X}_{kn}(s) := \tilde{f}_{kn}(s),\ s \in (\tilde{S}_n, \tilde{\nu}_n),\ k = 1,2,\cdots,n,$$

is $\tilde{\mu}_n$. Since the joint probability law of $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_n$ is also $\tilde{\mu}_n$, we have

$$P\{\tilde{X}_n(\omega) \in E\} = \tilde{\nu}_n\{\tilde{f}_{nn}(s) \in E\} = \tilde{\nu}_n(E)$$

and

$$P\{\tilde{X}_k(\omega) = \tilde{f}_{kn}(\tilde{X}_n(\omega)\} = \tilde{\nu}_n\{\tilde{f}_{kn}(s) = \tilde{f}_{kn}(\tilde{f}_{nn}(s))\} = 1$$

for $k \leq n$. The first equation means that the probability law of $\tilde{X}_n$ is $\tilde{\nu}_n$ and the second equation implies that

$$P\{\tilde{X}_k(\omega) = \tilde{f}_{kn}(\tilde{X}_n(\omega)),\ k \leq n\} = 1.$$

Removing a P-null set from Ω, (Remark 4 of the last section), we obtain

$$\tilde{X}_k(\omega) = \tilde{f}_{kn}(\tilde{X}_n(\omega)),\ k \leq n \quad \text{for every} \quad \omega \in \Omega.$$

Now define $X_\alpha(\omega)$ by

$$X_\alpha(\omega) := f_{\alpha\alpha_n}(\tilde{X}_n(\omega)) \quad \text{if} \quad \alpha < \alpha_n.$$

Since $f_{\alpha\beta} \circ f_{\beta\gamma} = f_{\alpha\gamma}$, $\alpha \leq \beta \leq \gamma$, X_α is well-defined independently of the choice of $\alpha_n \geq \alpha$. If $\alpha \leq \beta \leq \alpha_n$, then

$$X_\alpha = f_{\alpha\alpha_n} \circ \tilde{X}_n = (f_{\alpha\beta} \circ f_{\beta\alpha_n}) \circ \tilde{X}_n = f_{\alpha\beta} \circ X_\beta.$$

If $\alpha < \alpha_n$, then

$$P^{X_\alpha} = X_\alpha P = (f_{\alpha\alpha_n} \circ \tilde{X}_n)P = f_{\alpha\alpha_n}\tilde{\nu}_n = f_{\alpha\alpha_n}\nu_{\alpha_n} = \nu_\alpha .$$

This completes the proof of our theorem. ∎

3. Regularly measurable functions.

In Section 7 we introduced the space $\mathbb{L}^0 = \mathbb{L}^0[0,1]$ endowed with metric

$$(1) \quad \rho(f,g) = \int_0^1 [\,|f(t) - g(t)| \wedge 1\,]\, dt;$$

ρ satisfies all conditions of a metric except for the separation axiom which will hold only if equivalent functions are identified. The relation

$$(2) \quad C \subset D \subset \mathbb{L}^0$$

holds under such identification; see Section 1.7 for the definitions of C and D. To get rid of the trouble of identification of equivalent functions we choose from each equivalence class a single well-behaved function, called a <u>regularly measurable function</u>, and consider the space L^0 of all regularly measurable functions instead of $\mathbb{L}^0$ so that (2) may hold in the ordinary sense.

Let E be a (Lebesque) measurable subset of $[0,1]$. A point $t \in [0,1)$ is called a <u>right density point</u> of E, $t \in E^r$ in notation, if

$$\lim_{\varepsilon \downarrow 0} \frac{\lambda(E \cap [t,t+\varepsilon))}{\varepsilon} = 1 \qquad (\lambda = \text{Lebesgue measure}).$$

Similarly we define $t \in (0,1]$ to be a <u>left density point</u> of E, $t \in E^{\ell}$ in notation, if we have the same condition where the interval $[t,t+\varepsilon)$ is replaced by $(t-\varepsilon,t]$. The <u>upper regularization</u> $\bar{R}f$ of $f \in \mathbb{L}^0$ is defined by

$$\bar{R}f(t) := \begin{cases} \inf \{ a : t \in \{f \le a\}^r \} , & t \in [0,1) , \\ \inf \{ a : 1 \in \{f \le a\}^{\ell} \} , & t = 1, \end{cases}$$

where $\{f \le a\}$ denotes the set of all $s \in [0,1]$ such that $f(s) \le a$. By replacing 'inf' and '$f \le a$' by 'sup' and '$f \ge a$' respectively we define the <u>lower regularization</u> of f. It is obvious that

$$-\infty \le \underline{R}f(t) \le \bar{R}f(t) \le \infty \qquad \text{for every } t \in [0,1].$$

The <u>regularization</u> Rf of $f \in \mathbb{L}^0$ is defined by

$$Rf(t) = \begin{cases} \bar{R}f(t) & \text{if } \bar{R}f(t) = \underline{R}f(t) \quad (-\infty,\infty), \\ 0 & \text{otherwise.} \end{cases}$$

From now on we use the following notation:

$f = g$	$f(t) = g(t)$	everywhere on $[0,1]$
$f \sim g$	$f(t) = g(t)$	a.e. on $[0,1]$.

Theorem 1. $\bar{R}f \sim \underline{R}f \sim f$. Hence $Rf \sim f$.

Proof By the Lebesgue density theorem we have

$$\lambda(E \triangle E^r) = 0 \qquad (\triangle = \text{symmetric difference}).$$

Let $N_a := \{f \le a\} \triangle \{f \le a\}^r$ and $N := \bigcup_{a \in \mathbb{Q}} N_a$. Then $\lambda(N) = 0$. For every $t \in [0,1) - N$ and $a \in \mathbb{Q}$ we have

$$f(t) \le a \Leftrightarrow t \in \{f \le a\} \Leftrightarrow t \in \{f \le a\}^r \Rightarrow \bar{R}f(t) \le a$$

and

$\overline{R}f(t) < a \Rightarrow t \in \{f \le a\}^r \Leftrightarrow t \in \{f \le a\} \Leftrightarrow f(t) \le a.$

Hence $f(t) = \overline{R}f(t)$ for every $t \in [0,1]-N$, because $\mathbb{Q}$ is dense in $[-\infty, \infty]$. Since $\lambda(N) = 0$, we have $f \sim \overline{R}f$. Similarly we can prove that $f \sim \underline{R}f$. ∎

Theorem 1 ensures that Rf is a measurable function belonging to the same equivalence class as f and that

$R(Rf) = Rf,$

because it is obvious by the definition that $f \sim g \Rightarrow Rf = Rg$. Hence the function space

$L^0 := R(\mathbb{L}^0)$

consists of all functions $f \in L^0$ such that $Rf = f$. A function $f \in \mathbb{L}^0$ is called <u>regularly measurable</u> if $f \in L^0$, i.e. if $Rf = f$. From the observation above it is obvious that each equivalence class in $\mathbb{L}^0$ contains exactly one regularly measurable function. The space L^0 is a complete separable metric space with the metric given by (1).

Let $L^p := \mathbb{L}^p \cap L^0$ for $1 \le p < \infty$. Then L^p is a complete separable metric space with metric $\rho_p(f,g) = \|f-g\|_p$. Then it is obvious that

$\mathcal{D} \subset C \subset D \subset L^p \subset L^q \subset M \subset \mathcal{D}' \quad (p > q)$

in the ordinary sense; see Section 1.7 for the definitions of the spaces $\mathcal{D}$, M and $\mathcal{D}'$. If we denote these spaces by T_n, $n=1,2,\ldots,7$, then

$T_n \in B(T_m)$ and $B(T_n) = B(T_m) \cap T_n$, so $T_n \in B(T_m)$

whenever $n < m$; see Section 1.7.

The advantage of L^p is that we can define the evaluation maps e_t and e on L^p:

$e_t : L^p \to \mathbb{R}, \quad f \mapsto f(t)$ (the evaluation map at t),

$e : [0,1] \times L^p \to \mathbb{R}, \quad (t,f) \mapsto f(t)$ (the global evaluation map).

These maps may be defined on $\mathbb{L}^p$ as well, but $\rho(f,g) = 0$ (i.e. $f \sim g$) implies neither $e_t(f) = e_t(g)$ nor $e(t,f) = e(t,g)$, so such maps are not useful on $\mathbb{L}^p$. The evaluation maps on $\mathcal{D}$, C and D are defined in the same way as above.

Theorem 2. The evaluation maps are Borel for $F = \mathcal{D}$, C, D and L^p ($p=0$ or $1 \le p < \infty$).

Remark. $e : [0,1] \times F \to \mathbb{R}$ is Borel (i.e. measurable $\mathcal{B}([0,1] \times \mathbb{R})$) if and only if e is measurable $\mathcal{B}[0,1] \times \mathcal{B}(F)$ (Theorem 1.2.3).

Proof of the theorem. First we remark that if $\Phi : T \times F \to R$ is measurable $\mathcal{J} \times \mathcal{F}$, then the section map of Φ_t at t:

$$\Phi_t : F \to \mathbb{R}, \quad f \longmapsto \Phi(t,f)$$

is measurable $\mathcal{F}$. To prove this we can use the Dynkin class theorem observing that $\mathcal{J} \times \mathcal{F}$ is generated by $A \times B$, $A \in \mathcal{J}$, $B \in \mathcal{F}$. Since e_t is the section map of e at t, we need only prove that e is Borel. Since

$$F \in \mathcal{B}(L^0) \text{ and } \mathcal{B}(F) = \mathcal{B}(L^0) \cap F \text{ for } F = \mathcal{D}, C, D \text{ or } L^p,$$

it is enough to check that e is Borel for L^0.

Let $\bar{e}$ (resp. $\underline{e}$) denote the map $(t,f) \mapsto \bar{R}f(t)$ (resp. $\underline{R}f(t)$) from $[0,1] \times L^0$ into $\bar{\mathbb{R}} = [-\infty, \infty]$. Then

$$e(t,f) = \begin{cases} \bar{e}(t,f) & \text{if } \bar{e}(t,f) = \underline{e}(t,f) \in (-\infty,\infty) \\ 0 & \text{otherwise.} \end{cases}$$

Hence it is enough to prove that both $\bar{e}$ and $\underline{e}$ are Borel, we will prove this only for $\bar{e}$; the proof for $\underline{e}$ is similar.

Since

$$\{(t,f) : \bar{e}(t,f) \le a\} = \underbrace{\{(t,f) : \bar{R}f(t) \le a\}}_{f}$$
$$= \bigcap_n \{(t,f) : t \in \{f \le a+\tfrac{1}{n}\}\} \cup \bigcap_n \{(1,f) : 1 \in \{f \le a+\tfrac{1}{n}\}^{\ell}\},$$

it is enough to prove that each (t,f)-set in the above expression belongs to $\mathcal{B}[0,1] \times \mathcal{B}(L^0)$.

Let

$$\delta_\varepsilon(t,f,b) := \frac{\lambda(\{f \le b\} \cap [t,t+\varepsilon))}{\varepsilon} \qquad (0 < \varepsilon < 1)$$

and

$$\delta_{\varepsilon,\eta}(t,f,b) := \frac{1}{\varepsilon}\int_t^{t+\varepsilon} H_\eta(f(s))ds \qquad (0 < \varepsilon, \eta < 1)$$

where

$$H_\eta(x) := \begin{cases} 0, & x > a+\eta \\ 1, & x \le a \\ \text{linear in} & x \in [a, a+\eta]. \end{cases}$$

Since

$$|\delta_{\varepsilon,\eta}(t,f,b) - \delta_{\varepsilon,\eta}(\tilde{t},\tilde{f},b)|$$
$$\le \frac{1}{\varepsilon}\int_t^{t+\varepsilon} |H_\eta(f(s)) - H_\eta(\tilde{f}(s))| \, ds + \frac{2}{\varepsilon}|t-\tilde{t}|$$
$$\le \frac{1}{\varepsilon}\int_0^1 \left(\frac{|f(s)-\tilde{f}(s)|}{\eta} \wedge 1\right) ds + \frac{2}{\varepsilon}|t-\tilde{t}|$$
$$\le \frac{1}{\varepsilon\eta} \rho_0(f,\tilde{f}) + \frac{2}{\varepsilon}|t-\tilde{t}|,$$

$\delta_{\varepsilon,\eta}(t,f,b)$ is continuous in $(t,f) \in [0,1] \times L^0$. Since

$$\delta_{\varepsilon,\eta} \downarrow \delta_\varepsilon \qquad \text{as} \qquad \eta \downarrow 0,$$

$\delta_\varepsilon(t,f,b)$ is Borel in $(t,f) \in [0,1] \times L^0$. Since

$\delta_\varepsilon(t,f,b)$ is continuous in ε, we have

$$\underline{\delta}(t,f,b) := \lim_{\varepsilon \downarrow 0} \delta_\varepsilon(t,f,b) = \lim_{\substack{\varepsilon \downarrow 0 \\ \varepsilon \in \mathbb{Q}}} \delta_\varepsilon(t,f,b),$$

so $\underline{\delta}(t,f,b)$ is Borel in $(t,f) \in [0,1] \times L^0$. Observing that

$$\underline{\delta}(t,f,b) = 1 \Leftrightarrow t \in \{f \leq b\}^r$$

we can conclude that

$$\{(t,f) : t \in \{f \leq b\}^r\} \in \mathcal{B}([0,1] \times L^0)$$

Similarly we obtain

$$\{(1,f) : 1 \in \{f \leq b\}^\ell\} \in \mathcal{B}([0,1] \times L^0)$$

completing the proof of our theorem.

We can make the above discussion for a general real interval T. No essential charge is necessary for T compact. For T non-compact we express T as a countable union of compact intervals T_n, n=1,2,.... The space C = C(T) of all continuous real functions is a complete separable metric space with metric

$$\rho_n(t,g) = \sum_n 2^{-n} \sup_{t \in T_n} [|f(t) - g(t)| \wedge 1].$$

The space D = D(T) of all right continuous real functions with finite left limits (left continuous at the right endpoint of T if T is right closed) is a complete separable metric space with the Skorohod metric

$$\rho_S(f,g) = \inf_{\varphi \in \Phi} \{\rho_u(\varphi, i) + \rho_u(f \circ \varphi, g)\},$$

where Φ is the family of all order-preserving homeomorphism from T to itself. The space $L^p = L^p(T)$ (p=1 or $1 \leq p < \infty$) is defined in the same way as in the case $T = [0,1]$ but the metric is defined by

$$\rho_p(f,g) = \begin{cases} \int_T \left[\, |f(t) - g(t)| \wedge \frac{1}{1+t^2} \right] dt. & , p=0, \\ \left(\int_T |f(t) - g(t)|^p \, dt\right)^{1/p} & , 1 \quad p \quad . \end{cases}$$

Then (L^p, ρ_p) is a complete separable metric space. Then

$$C \subset D \subset L^0 \quad \text{but} \quad D \not\subset L^p \; (1 \leq p < \infty).$$

in case T is non-compact. Hence we consider the space $L^p_{loc} = L^p_{loc}(T)$ of all locally p-th order summable functions in L^0 endowed with metric

$$\rho_p(f,g) = \sum_n 2^{-n} \left[\int_{T_n} |f(t) - g(t)|^p \, dt \right]^{1/p}.$$

Then (L^p_{loc}, ρ_p) is a complete separable metric space. Using the same argument as before, we can prove

<u>Theorem 3</u> $C \subset D \subset L^p_{loc} \subset L^0$. If we denote these spaces by S_1, S_2, S_3 and S_4, then

$$S_m \in \mathcal{B}(S_n), \quad \mathcal{B}(S_m) = \mathcal{B}(S_n) \wedge S_m \quad \text{and} \quad \mathcal{B}(S_m) \subset \mathcal{B}(S_n)$$

whenever $m < n$.

Thus far we have taken the Lebesque measure as the reference measure in defining the spaces L^p, $p \in \{0\} \cup [1, \infty)$. Now we consider the case where the reference measure μ on T is a general locally finite (=finite on compacts) B-regular measure μ on T. Once we define the space $L^0(T, \mu)$ of all regularly μ-measurable functions, we can proceed as above; the only difference is that we do not have

$$C(T) \subset D(T) \subset L^0(T,\mu)$$

unless μ is atomless (= vanishing on every singleton) and strictly positive (= positive on every non-empty open set).

Let $F : (-\infty, \infty) \to \mathbb{R}$ be a non-decreasing right-continuous function such that $\mu[s,t] = F(t) - F(s-)$ for every $(s,t) \in T \times T$ with $s \leq t$; such a function F is called the <u>distribution function</u> of μ and μ is the Lebesque-Stieltjes measure dF.

Let I be an interval with the endpoints

$$\alpha := \inf_{t \in T} F(t-) < \beta := \sup_{t \in T} F(t),$$

where α (or β) belongs to I if the left (or right) endpoint belongs to T. If μ is atomless and strictly positive, the map $F : T \to I$ is an order-preserving homeomorphism and the image measure $F\mu$ is the Lebesque measure λ on I. Hence we define

$$L^0(T,\mu) := \{g \circ F : \quad g \in L^0(I,\lambda)\} \quad .$$

If μ is general, we define

$$L^0(T,\mu) := \{g \circ \tilde{F} : \quad g \in \tilde{L}^0(I,\lambda)\} \quad ,$$

where $\tilde{F}(t) := F(t-)$ and $\tilde{L}^0(I,\lambda)$ is the space of all functions f that are constant in $(F(t-), F(t))$ for every jump point t of F. The details are left to the reader.

4. Stochastic process and random functions.

Let T be a real interval. A family $X_t(\omega)$, $t \in T$, is called a (stochastic) process on T. For simplicity we deal with the case where T is the unit interval $[0,1]$ endowed with the Lebesgue measure. The general case where T is a general interval endowed with a general locally finite B-regular measure can be treated similarly with some obvious modifications.

A stochastic process $\{X_t, t \in T\}$ is called continuous in probability if the map $t \to X_t$ from T into $\mathcal{L} = \mathcal{L}(\Omega,P)$ with respect to the usual topology on T and the ρ_0-topology on $\mathcal{L}$, i.e. if

$$\lim_{s\to t} P(|X_s - X_t| > \varepsilon) = 0 \quad \text{for every } t \in T \text{ and every } \varepsilon > 0.$$

$\{X_t, t \in T\}$ is called measurable in probability if the map $t \to X_t$ from T into $\mathcal{L}$ is measurable, i.e. measurable $\mathcal{B}(\lambda)/\mathcal{B}(\mathcal{L})$ where $\mathcal{B}(\mathcal{L})$ is the topological σ-algebra on $\mathcal{L}$ with respect to the ρ_0-topology. It is obvious that continuity in probability implies measurability in probability. From now on we will abbreviate 'in probability' to 'i.p.'.

A stochastic process $\{X_t, t \in T\}$ is called measurable if $X_t(\omega)$ is measurable $\mathcal{B}(\lambda) \times \mathcal{B}(P)$ as a function of $(t,\omega) \in T \times \Omega$.

Theorem 1. Measurability implies measurability i.p.

Proof. Suppose that $\{X_t, t \in T\}$ is measurable. Then the function of t :

$$\rho_0(X_t,Y) := \int_\Omega [|X_t(\omega) - Y(\omega)| \wedge 1] P(d\omega) \qquad (Y \in \mathcal{L})$$

is measurable $\mathcal{D}(\lambda)$, so

$$\{t : \rho_0(X_t,Y) < r\} \in \mathcal{D}(\lambda).$$

This means that the inverse image of any open ball in $\mathcal{L}$ under the map $t \to X_t$ belongs to $\mathcal{D}(\lambda)$. Hence the map $t \to X_t$ is measurable $\mathcal{D}(\lambda)/\mathcal{B}(\mathcal{L})$, because the space $(\mathcal{L},\rho_0)$ is a separable metric space (Section 1). ∎

Fixing $\omega \in \Omega$ and moving t in a given stochastic process $X_t(\omega)$, $t \in T$, we obtain a function of t, which will be called the sample function of the process corresponding to the sample point ω, denoted by $X.(\omega)$. A process $\{X_t, t \in T\}$ is called

a C process if $X.(\omega) \in C = C(T)$ for every ω,

a D process if $X.(\omega) \in D = D(T)$ for every ω,

and

an L^p process if it is a measurable process and if $X.(\omega) \in L^p = L^p(T)$ for every ω.

Theorem 2. Every C process is a D process and every D process is an L^p process.

Proof. Since $C \subset D \subset L^p$, it is enough to prove that every D process is measurable. Suppose that $\{X_t, t \in T\}$ is a

D process. Let

$$X_t^n(\omega) := \begin{cases} X_{\frac{k}{n}}(\omega), & t \in [\frac{k-1}{n}, \frac{k}{n}) \quad (k=1,2,\ldots,n-1) \\ X_1(\omega), & t \in [\frac{n-1}{n}, 1] \end{cases}$$

Since $X.(\omega) \in D$,

$$X_t(\omega) = \lim_{n\to\infty} X_t^n(\omega) \quad \text{for every} \quad (t,\omega) \in T \times \Omega.$$

Since the set $\{(t,\omega) : X_t^n(\omega) < a\}$ is expressible in the form

$$\sum_{k=1}^{n} I_k \times A_k, \quad I_k : \text{interval}, \quad A_k \in \mathcal{B}(P),$$

$X_t^n(\omega)$ is measurable $\mathcal{B}(T) \times \mathcal{B}(P)$ (as a function of (t,ω)), so $X_t(\omega)$ is also measurable $\mathcal{B}(T) \times \mathcal{B}(P)$. Now note that $\mathcal{B}(T) \subset \mathcal{B}(\lambda)$.

A C-valued random variable, i.e. a map from Ω into C measurable $\mathcal{B}(P)/\mathcal{B}(C)$, is called a <u>random C function</u>. Similarly we define <u>random D functions</u> and <u>random L^p functions</u>. Let $Y(\omega)$ be a random C (or D or L^p) function. Since the evaluation map e_t is Borel, $e_t(Y(\omega))$ is a real random variable. The stochastic process $e_t(Y(\omega))$, $t \in T$, is called the <u>evaluation process</u> of $Y(\omega)$. Then $Y(\omega)$ is the sample function of the evaluation process of $Y(\omega)$.

<u>Theorem 3</u>. Let $\{X_t\} = \{X_t, t \in T\}$ be a stochastic process.

(i) $\{X_t\}$ is a C process $\Leftrightarrow$ $X.(\omega)$ is a random C function,

(ii) $\{X_t\}$ is a D process $\Leftrightarrow$ $X.(\omega)$ is a random D function,

(iii) $\{X_t\}$ is an L^p process $\Leftrightarrow$ $X.(\omega)$ is a random L^p function.

<u>Proof</u>. First we prove (iii) for $p = 0$. Suppose that $\{X_t\}$ is an L^0 process. Then $X_t(\omega)$ is measurable $\mathcal{D}(\lambda) \times \mathcal{D}(P)$ as a function of (t,ω). Hence Fubini's theorem ensures that the function of ω :

$$\rho_0(X.(\omega), f) = \int_T (|X_t(\omega) - f(t)| \wedge 1)dt \qquad (f \in L^0)$$

is P-measurable, so

$$\{\omega : \rho_0(X.(\omega), f) < \varepsilon\} \in \mathcal{D}(P)$$

i.e. $\quad X^{-1}(U(f,\varepsilon)) \in \mathcal{D}(P)$,

$U(f,\varepsilon)$ being the ε-neighborhood of f. Since (L^0,ρ_0) is a separable metric space, this implies that $X. : \Omega \to L^0$ is measurable $\mathcal{D}(P)/\mathcal{B}(L^0)$, proving that $X.(\omega)$ is a random L^0 function.

Suppose conversely that $X.(\omega)$ is a random L^0 function. Then

$$X_t(\omega) = e_t(X.(\omega)) = e(t, X.(\omega)),$$

where $e : T \times L^0 \to \mathbb{R}$ is the global evaluation map. Since e is measurable $\mathcal{B}(T) \times \mathcal{B}(L^0)$ and since $X.: \Omega \to L^0$ is measurable $\mathcal{D}(P)/\mathcal{B}(L^0)$, it is easy to check that $X_t(\omega) = e(t, X.(\omega))$ is measurable $\mathcal{B}(T) \times \mathcal{D}(P)$, proving that $\{X_t\}$ is an L^0 process.

Next we will prove (iii) for $p \in [1,\infty)$; we can prove (i) and (ii) by the same argument. Suppose that $\{X_t\}$ is an L^p process. Since $L^p \subset L^0$, $\{X_t\}$ is regarded as an L^0 process for which $X.(\omega) \in L^p$ for every ω. Since

$$\mathcal{B}(L^p) \subset \mathcal{B}(L^0) \qquad \text{(see the last section),}$$

the assertion (i) ensures that

$$X.^{-1}(B) \in \mathcal{D}(P) \qquad \text{for every} \quad B \in \mathcal{B}(L^p).$$

Hence $X.(\omega)$ is a random L^p function. Suppose conversely that $X.(\omega)$ is a random L^p function. Since

$$X.(\Omega) \subset L^p \quad \text{and} \quad \mathcal{B}(L^p) = \mathcal{B}(L^0) \cap L^p,$$

$X.(\omega)$ is regarded as a random L^0 function. Hence $\{X_t\}$ is a measurable process. Since $X.(\omega) \in L^p$ for every ω, $\{X_t\}$ is an L^p process. ∎

Let $\{X_t, t \in T\}$ be a stochastic process. A C process $\{Y_t, t \in T\}$ is called a <u>C regularization</u> of $\{X_t\}$ if

$$P(X_t = Y_t) = 1 \quad \text{for } \underline{\text{every}} \quad t \in T.$$

Similarly we define a <u>D regularization</u> of $\{X_t\}$. An L^p process is called an <u>L^p regularization</u> of $\{X_t\}$ if

$$P(X_t = Y_t) = 1 \quad \text{for } \underline{\text{almost every}} \quad t \in T.$$

Theorem 4. A stochastic process $\{X_t, t \in T\}$ has an L^0 regularization if and only if it is measurable i.p. If $\{Y_t\}$ and $\{Y'_t\}$ are L^0 regularizations of $\{X_t\}$, then $P\{Y. = Y'.\} = 1$.

Proof. Let us first prove

Lemma 1. If $\{X_t, t \in T\}$ is continuous i.p., then there exists a measurable process $\{Y_t, t \in T\}$ such that

$$P(X_t = Y_t) = 1 \quad \text{for almost every} \quad t \in T.$$

Proof. Let

$$X_t^n(\omega) := \begin{cases} X_{k/n}(\omega), & t \in [\frac{k-1}{n}, \frac{k}{n}), \quad k = 1,2,\ldots,n \\ X_1(\omega), & t = 1 \end{cases}$$

Then $\{X_t^n\}$ is a measurable process for every n, because the set $\{(t,\omega) : X_t^n(\omega) < a\}$ is expressed in the form:

$$\sum_{k=1}^{n} I_k \times A_k, \qquad I_k : \text{interval}, \qquad A_k \in \mathcal{B}(P).$$

Since $\{X_t\}$ is continuous i.p., we have

$$(1) \qquad \lim_{n\to\infty} \rho_0(X_t^n, X_t) = 0, \qquad t \in T.$$

Using Fubini's theorem and the bounded convergence theorem, we obtain

$$\int_{T\times\Omega} [|X_t^n(\omega) - X_t^m(\omega)| \wedge 1]\, dt\, P(d\omega) \to 0 \quad \text{as} \quad n,m \to \infty.$$

Hence we can find a subsequence $\{Y_t^n(\omega),\ n = 1,2,\ldots\}$ of $\{X_t^n(\omega),\ n = 1,2,\ldots\}$ convergent to a measurable process $Y_t(\omega)$ a.e. on $T \times \Omega$. Hence

$$\rho_0(Y_t^n, Y_t) = \int_\Omega [(|Y_t^n(\omega) - Y_t(\omega)| \wedge 1]\, P(d\omega) \to 0 \quad (n \to \infty) \quad \text{a.e. on } T,$$

so

$$P(Y_t = X_t) = 1 \qquad \text{a.e. on } T$$

by (1). This completes the proof of Lemma 1.

Now we prove our theorem. Suppose that $\{X_t\}$ is measurable i.p., namely that the map $t \to X_t$ from T into $\mathcal{L}^0$ is measurable $\mathcal{B}(\lambda)/\mathcal{B}(\mathcal{L}^0)$. Since $\mathcal{L}^0$ is Polish, we can use the generalized Lusin theorem (Section 1.9) to find a compact set $K_n \subset T$ such that (i) $\lambda(T - K_n) < 1/n$ and (ii) the map $t \to X_t$ restricted to $t \in K_n$ is continuous. This map from K_n into $\mathcal{L}^0$ can be extended to a continuous map $t \to X_t^n$ from T into $\mathcal{L}^0$; since $T - K_n$ is a countable union of intervals $I_1, I_2, \ldots$, we can obtain X_t^n by linear interpolation on each interval I_n. Then $\{X_t^n,\ t \in T\}$ is continuous i.p. and $X_t^n = X_t$ for $t \in K_n$. Use Lemma 1 to find a measurable process $\{Y_t^n,\ t \in T\}$ such that $P(Y_t^n = X_t^n) = 1$ for $t \in T - N_n$ where $\lambda(N_n) = 0$, so

$$P(Y_t^n = X_t) = 1 \quad \text{for} \quad t \in K_n - N_n$$

Define

$$Y_t(\omega) := \begin{cases} Y_t^n(\omega) & \text{if } (t,\omega) \in (K_n - \bigcup_{i=1}^{n-1} K_i) \times \Omega \\ 0 & \text{otherwise} \end{cases}$$

It is obvious that $Y_t(\omega)$ is a measurable process and

$$P(Y_t = X_t) = 1 \quad \text{for } t \in T - N'$$

where $N' := (T - \bigcup_n k_n) \cup N$, so $\lambda(N') = 0$. Let

$$\tilde{Y}_t(\omega) := e_t(R[Y.(\omega)]) \quad \text{(See the last section for } R).$$

Then $\tilde{Y}.(\omega) = R[Y.(\omega)]$ and

$$\rho_0(\tilde{Y}.(\omega), f) = \rho_0(Y.(\omega), f) = \int_T [|Y_t(\omega) - f(t)| \wedge 1]dt.$$

Hence $\rho_0(\tilde{Y}.(\omega), f)$ is a P-measurable function of ω. This implies that $\tilde{Y}. : \Omega \to L^0$ is measurable $\mathcal{D}(p)/\mathcal{B}(L^0)$, because (L^0, p_0) is a separable metric space. Hence $\{\tilde{Y}_t\}$ is an L^0 process by Theorem 3. Since $\{Y_t\}$ and $\{\tilde{Y}_t\}$ are measurable processes and since $\tilde{Y}_t(\omega) = Y_t(\omega)$ a.e. on T for every $\omega \in \Omega$, Fubini's theorem ensures that

$$p(\tilde{Y}_t = Y_t) = 1 \quad \text{for almost every } t \in T.$$

Since $P(Y_t = X_t) = 1$ for almost every $t \in T$,
$p(\tilde{Y}_t = X_t) = 1$ for almost every $t \in T$.

This implies that $\{\tilde{Y}_t\}$ is an L^0 regularization of $\{X_t\}$.

Suppose conversely that $\{Y_t\}$ is an L^0 regularization of $\{X_t\}$. Since $\{Y_t\}$ is an L^0 process, it is measurable, so it is measurable i.p. by Theorem 1. Since X_t and Y_t are identical as points of $\mathcal{L}^0$ for almost every t, $\{t : Y_t \in B\}$ and $\{t : X_t \in B\}$ differ from each other only by a null set. But $\{t : Y_t \in B\} \in \mathcal{B}(\lambda)$, so $\{t : X_t \in B\} \in \mathcal{B}(\lambda)$. This implies that $\{X_t\}$ is measurable i.p.

Let $\{Y_t\}$ and $\{Y_t'\}$ be L^0 regularizations of $\{X_t\}$. Then

$$Y_t(\omega) = Y_t'(\omega) = Y_t(\omega) \quad \text{a.s.} \quad \text{for almost every } t,$$

so $Y_t(\omega) = Y_t'(\omega)$ for almost every $t \in T$ a.s. by Fubini's theorem. Since $Y.(\omega), Y.'(\omega) \in L^0$, this implies that

$$Y.(\omega) = Y.'(\omega) \quad \text{a.s.},$$

completing the proof of our theorem. ∎

Two processes $\{X_t, t \in T\}$ and $\{Y_t, t \in T\}$ are called <u>sample equivalent</u> to each other if $P(X. = Y.) = 1$. Theorem 4 claims that every process measurable i.p. has a unique (up to sample equivalence) L^0 regularization.

If $\{X_t\}$ has a C (or D) regularization $\{Y_t\}$, then $\{Y_t\}$ is also an L^0 regularization of $\{X_t\}$ and every L^0 regularization of $\{X_t\}$ is sample equivalent to $\{Y_t\}$. Hence <u>it is enough to consider only L^0 regularizations</u>.

Theorem 5. (A. Kolmogorov) Suppose that $\{X_t, t \in T\}$ satisfies

$$E(|X_t - X_s|^{\alpha}) \leqq \gamma |t - s|^{1+\beta} \quad \text{for every} \quad (t,s) \in T \times T,$$

where α, β and γ are positive constants. Then $\{X_t\}$ has a C regularization.

Proof. Choose a positive number δ so that $\varepsilon := \beta - \alpha\delta > 0$. Then

$$P\{|X_t - X_s| > |t - s|^{\delta}\} \leq \gamma |t - s|^{1+\varepsilon}$$

Hence

$$P\{|X(k/2^n) - X((k-1)/2^n)| > 2^{-n\delta} \quad \text{for some} \quad k = 1,2,\ldots,2^n-1\}$$
$$\leq \gamma 2^n 2^{-n(1+\varepsilon)} = \gamma 2^{-n\varepsilon}$$

where $X(t)$ denotes X_t. By Borel-cantelli's lemma the following event $\Omega_1 \subset \Omega$ has probability 1:

(2) $$|X(\tfrac{k}{2^n}) - X(\tfrac{k-1}{2^n})| \leqq \gamma 2^{-n\varepsilon}$$

for n sufficiently large $(n \geq N(\omega))$ and for $k = 1,2,\ldots,2^n$.

Let $\mathbb{Q}'$ denote the set of all numbers in T of the form $k/2^n$. We claim that

(3) $$|X_{\rho}(\omega) - X_r(\omega)| \leq 2\gamma(1 - 2^{-\varepsilon})(\rho - r)^{\varepsilon},$$
$$\rho, r \in \mathbb{Q}', \quad 0 < \rho - r < 2^{-N(\omega)}, \quad \omega \in \Omega_1.$$

To prove this, determine n by

$$2^{-(n-1)} > \rho - r \geq 2^{-n}$$

and choose k so that

$$r \leq k2^{-n} < \rho.$$

It is obvious that $n \geq N(\omega)$. Observing that

$$k2^{-n} - r < \rho - r < 2^{-(n-1)}$$

and

$$\rho - k2^{-n} < \rho - r < 2^{-(n-1)},$$

we can expand r and ρ as follows:

$$r = \frac{k}{2^n} - \frac{a_0}{2^n} - \frac{a_1}{2^{n+1}} - \cdots - \frac{a_p}{2^{n+p}} \qquad a_i = 0 \text{ or } 1$$

and

$$\rho = \frac{k}{2^n} + \frac{b_0}{2^n} + \frac{b_1}{2^{n+1}} + \cdots + \frac{b_p}{2^{n+q}} \qquad b_i = 0 \text{ or } 1.$$

Let $r_{-1}, r_0, r_1, \ldots, r_p$ be the partial sums of the expansion of r and $\rho_{-1}, \rho_0, \rho_1, \ldots, \rho_q$ those for ρ; $r_{-1} = \rho_{-1} = k/2^n$, $r_p = r$ and $\rho_q = \rho$. Then (2) implies that for $\omega \in \Omega_1$

$$|X(r_i) - X(r_{i-1})| \leq \gamma 2^{-(n+i)\varepsilon}$$

and

$$|X(\rho_j) - X(\rho_{j-1})| \leq \gamma 2^{-(n+i)\varepsilon},$$

so

$$|X(\rho) - X(r)| \leq \sum_{i=0}^{p} |X(r_i) - X(r_{i-1})| + \sum_{j=0}^{q} |X(\rho_j) - X(\rho_{j-1})|$$

$$\leq 2\gamma 2^{-n\varepsilon}(1 - 2^{-\varepsilon})$$

$$\leq 2\gamma(1 - 2^{-\varepsilon})(\rho - r)^{\varepsilon},$$

proving (3).

By virtue of (3) $X_r(\omega)$ is a uniformly continuous function of $r \in Q'$ for $\omega \in \Omega_1$, so it can be extended to a continuous function $Y_t(\omega)$ of $t \in T$ for $\omega \in \Omega_1$. We define

$$Y_t(\omega) \equiv 0 \qquad \text{for} \quad \omega \in \Omega - \Omega_1.$$

It is obvious that $Y_t(\omega) = X_t(\omega)$ for $\omega \in \Omega_1$ and for $t \in Q'$. Hence

$$P(Y_t = X_t) = 1 \qquad \text{for} \quad t \in Q'.$$

For $t \in T - Q'$ we can find a sequence $t_n \in Q'$ converging to t. Since $\{X_t\}$ is continuous i.p. by the assumption, $X_{t_n} \to X_t$ i.p. Hence we can find a subsequence $\{s_n\}$ of $\{t_n\}$ such that $X_{s_n} \to X_t$ a.s. Since $Y_{s_n}(\omega) \to Y_t(\omega)$ for every ω and since $X_{s_n} = Y_{s_n}$ a.s., we have

$$P(Y_t = X_t) = 1 \qquad \text{for} \quad t \in T - Q'.$$

This completes the proof that $\{Y_t\}$ is a C regularization of $\{X_t\}$. ∎