J. Math. Soc. Japan
Vol. 20, Nos. 1-2, 1968

The canonical modification of stochastic processes

Dedicated to Professor S. Iyanaga for his sixtieth birthday

By Kiyosi ITÔ

(Received Sept. 11, 1967)

§1. Introduction and summary.

In his book [1], J. L. Doob proved that every stochastic process continuous in probability has a standard separable measurable modification. This theorem plays a fundamental role in the sample path approach of stochastic processes. The aim of our paper is to give a more concrete formulation to this important fact to make it easier to visualize the probability law of the sample path.

In Section 2 we shall introduce the space $M \equiv M(T)$ of *canonical measurable functions* on the time interval T. The space $\tilde{M} \equiv \tilde{M}(T)$ contains bad functions such as the Dirichlet function that takes 1 on rationals and 0 elsewhere. Since we have a good function $f \equiv 0$ equivalent to the Dirichlet function, this can be discarded from $\tilde{M}$ without any essential loss. We shall pick up at least one good function, called canonical measurable function here, from among each equivalent class in $\tilde{M}$ and consider the space $M \equiv M(T)$ of all canonical functions in behalf of $\tilde{M}$. By definition a canonical function takes one of its general approximate limits at each point. All continuous functions are in M and if a function in M is equal to a continuous function almost everywhere on T, they are equal everywhere on T. A similar fact holds for functions with no discontinuities of the second kind. These facts suggest that M is suitable for the function space in which the path of a reasonable stochastic process is ranging.

In Section 3 we shall define a σ-algebra $\mathcal{B} = \mathcal{B}(M)$ of subsets of M which will determine a measurable structure in M. $\mathcal{B}$ is generated by all sets of the following types

$$\text{(i)} \qquad \{f \in M : f(t) < a\}\,,$$

$$\text{(iii)} \qquad \{f \in M : \int_I \arctan f(t)dt < a\}\,.$$

where a ranges over reals, t over T and I over all compact intervals in T. The scaling "arctan" was used in the integral to make it converge. We shall write $\mathcal{B}_K \equiv \mathcal{B}_K(M)$ and $\mathcal{B}_\rho \equiv \mathcal{B}_\rho(M)$, respectively, for the σ-algebra generated by the sets (i) only and that generated by the sets (ii) only. The space $C \equiv C(T)$ of all continuous functions, the space $D \equiv D(T)$ of all functions with no discontinuities of the second kind and the space $L^p \equiv L^p(T)$ of all canonical functions with finite p-th norm are $\mathscr{B}_\rho$-measurable and so $\mathscr{B}$-measurable.

In Section 4 we shall discuss *canonical stochastic processes*. Let $\{X_t(\omega), t \in T\}$ be a stochastic process defined on $(\Omega, \mathscr{F}, P)$. The sample path $X.(\omega)$ is a function of ω ranging in the function space $\bar{R}^T$ in general. $\{X_t\}$ is called *canonical* if $X.(\omega) \in M$ for every ω and if the map: $\omega \to X.(\omega)$ from Ω into M is measurable $(\mathscr{F}, \mathscr{B})$. In other words, the sample path $X.(\omega)$ is an $(M, \mathscr{B})$-valued random variable. We shall prove that every canonical process continuous in probability is *measurable* in the pair (t, ω) with respect to the product measure $dt\, dP$ and also *separable* relative to closed sets with respect to every countable dense subset of T. The (*standard*) *canonical modification* is defined in the same way as Doob's separable measurable modification, but our meaning of "standard" is more strict than Doob's. It will be proved that every stochastic process continuous in probability has one and *only one* (in a reasonable sense) canonical modification.

In Section 5 we shall discuss probability measures on M. Let $\{X_t(\omega)\}$ be a canonical process continuous in probability. Then the sample path $X.(\omega)$ is an $(M, \mathscr{B})$-valued random variable on $(\Omega, \mathscr{F}, P)$. The probability law of $X.(\omega)$ defined as usual will be a complete $\mathscr{B}$-regular probability measure μ on M which satisfies

$$(\Gamma.\ 1) \qquad \mu\{f \in M : |f(t)| < \infty\} = 1 \qquad t \in T,$$

$$(\Gamma.\ 2) \qquad \lim_{s \to t} \mu\{f \in M : |f(s) - f(t)| > \varepsilon\} = 0 \qquad \varepsilon > 0, \qquad t \in T.$$

The finite-dimensional marginal distribution $m_{t_1, \dots, t_n}$ of over $\{t_1, \cdots, t_n\}$ is defined by

$$m_{t_1, \dots, t_n}(E) = \mu\{f \in M : (f(t_1), \cdots, f(t_n)) \in E\}.$$

The system of all such marginal distributions will satisfy

$$(m.\ 1) \qquad m_t\{(-\infty, \infty)\} = 1 \qquad t \in T$$

$$(m.\ 2) \qquad \lim_{s \to t} m_{st}\{(x, y) : |x - y| > \varepsilon\} = 0 \qquad \varepsilon > 0, \qquad t \in T$$

in addition to Kolmogorov's consistency condition. Conversely, if we are given such a system of finite dimensional probability measures $\{m_{t_1, \dots, t_n}\}$, we can construct *one and only one* μ with $(\Gamma.\ 1)$ and $(\Gamma.\ 2)$ whose marginal distributions are the given $\{m_{t_1, \dots, t_n}\}$. It is to be noted that the finite-dimensional marginal distributions determine the probability measure μ not only on $\mathscr{B}_K$ out also on $\mathscr{B}$, provided the conditions $(m.\ 1)$ and $(m.\ 2)$ are satisfied.

REMARKS. We shall list some notations and definitions which will be used repeatedly in this paper.

T stands for a real interval, bounded or unbounded and open, closed or semi-closed. It indicates the time interval. The σ-algebra of all Borel subsets of T is denoted by $\mathscr{B}(T)$.

An open subset of T means a subset of T open in T, not in $(-\infty, \infty)$. Similarly for a closed subset.

R and $\bar{R}$ stand respectively for the open real line $(-\infty, \infty)$ and the closed one $[-\infty, \infty]$. $\mathscr{B}(R)$ denotes the system of all Borel subsets of R. Similarly for $\mathscr{B}(\bar{R})$ and $\mathscr{B}(R^n)$.

A subset E of T is always assumed to be measurable and $|E|$ denotes the Lebesgue measure of E.

I, I_n etc. stand for bounded subintervals of T.

$\mathscr{I}_{\text{rat}}$ stand for the system of all compact subintervals of T expressible as the intersection of T with a rational interval $\subset R$, and $\mathscr{F}_{\text{rat}}$ for the system of all non-empty sets expressible as a finite sum of compact intervals in $\mathscr{I}_{\text{rat}}$. Both are countable systems. Every E with $|E|<\infty$ can be approximated in measure by sets in $\mathscr{F}_{\text{rat}}$.

A function f is always assumed to be a real measurable function defined on T. f may take $\pm\infty$ on a null set.

$\{f<c\}$ denotes the set $\{t\in T: f(t)<c\}$.

"arctan" is abbreviated as "atn". This gives a homeomorphism from $\bar{R}$ onto $[-\pi/2, \pi/2]$.

$(\boldsymbol{X}, \mathscr{B})$ stands for a *measurable space*.

Let $(\boldsymbol{X}, \mathscr{B})$ and $(\boldsymbol{Y}, \mathcal{C})$ be measurable spaces. A map $f: \boldsymbol{X}\to \boldsymbol{Y}$ (into) is said to be *measurable* $(\mathscr{B}, \mathcal{C})$ if $f^{-1}(\mathcal{C})\supset \mathscr{B}$ i.e. if $f^{-1}(C)\in\mathscr{B}$ for every $C\in\mathcal{C}$. f is said to be *measurable* $(\mathscr{B})$ or *$\mathscr{B}$-measurable* in case $\boldsymbol{Y}=R$ or $\bar{R}$ and $\mathcal{C}=\mathscr{B}(R)$ or $\mathscr{B}(\bar{R})$.

The *basic probability space* is denoted by $(\Omega, \mathscr{F}, P)$. $\mathscr{F}$ is always assumed to be P-complete. In other words, every subset of a set with P-measure 0 belongs to $\mathscr{F}$. The *generic element* in Ω is denoted by ω.

Let $\mathscr{F}_1$ be a σ-subalgebra of $\mathscr{F}$. The *P-completion* of $\mathscr{F}_1$, denoted by $\bar{\mathscr{F}}_1^P$ is defined to be the σ-algebra that consists of all $A\in\mathscr{F}$ with the property:

$$\exists A_1, A_2\in\mathscr{F}_1 \quad \text{such that} \quad A_1\subset A\subset A_2,\ P(A_2-A_1)=0\,.$$

Notice that not every P-null set belongs to $\bar{\mathscr{F}}_1^P$.

An $(\boldsymbol{X}, \mathscr{B})$-valued *random variable* X is a function of ω such that the map $\omega\to X(\omega)$ is $(\mathscr{F}, \mathscr{B})$-measurable. The *probability law* of X, denoted by P_X, is defined by

$$P_X(B)=P\{\omega: X(\omega)\in B\}=P(X^{-1}(B)) \quad \text{if } X^{-1}(B)\in\overline{X^{-1}(\mathscr{B})}^P\,.$$

P_X is a *complete $\mathscr{B}$-regular probability measure* on $\boldsymbol{X}$. The P_X-completion of $\mathscr{B}$

is denoted by $\overline{\mathscr{B}}^{P_X}$. $P\{\omega: X\in B\}$ may be meaningful for a set $B\in\overline{\mathscr{B}}^{P_X}$, but we do *not* define P_X for such B.

An $(\bar{R}, \mathscr{B}(\bar{R}))$-valued random variable $X(\omega)$ is called a *real random variable* if $P(X\in R)=1$. $X_t(\omega)$ is written as $X(t, \omega)$ without explanation. The σ-algebra generated by the sets $\{\omega: X_t(\omega)<a,\ t\in T,\ a\in R\}$ is denoted by $\mathscr{B}[X_t, t\in T]$ or $\mathscr{B}[X]$.

§ 2. Canonical measurable function.

Let $\tilde{M}\equiv\tilde{M}(T)$ stand for the space of all real measurable functions defined on T. We shall introduce some notions following Saks [2]. The parameter of regularity of E, denoted by $\alpha(E)$, is defined to be the supremum of $|E|/|I|$, I ranging over all intervals $\supset E$. We shall write $E_n \underset{r}{\rightarrow} t$ (E_n *tends to* t *regularly*), if $t\in E_n$ for every n, if the diameter of E tends to 0 and if $\inf_n \alpha(E_n)>0$. $a\in\bar{R}$ is called a *general approximate limit* of $f\in\tilde{M}$ at t if for every neighborhood U of a, we have $E_n \underset{r}{\rightarrow} t$ such that

$$\lim_n \frac{|\{f\in U\}\cap E_n|}{|E_n|}>0\,.$$

The set of all general approximate limits of f at t is denoted by $L(f, t)$. The *approximate upper limit* of f at t, denoted here by $\bar{f}(t)$, is defined to be the infimum of $b\in\bar{R}$ such that for every $E_n \underset{r}{\rightarrow} t$ we have

$$\overline{\lim_n} \frac{|\{f>b\}\cap E_n|}{|E_n|}=0\,;$$

we set $\bar{f}(t)=-\infty$, if there is no such b. The *approximate lower limit* $\underline{f}(t)$ is defined similarly. It is a well-known important fact [2] that $\bar{f}(t)=\underline{f}(t)$ *a. e.* on T. It is easy to see

PROPOSITION (2.1). *$\bar{f}(t)$ ($\underline{f}(t)$) is the largest (least) element in $L(f, t)$ and so $L(f, t)$ is non-empty. $L(f, t)$ consists of a single point for almost every $t\in T$.*

Let $\mathscr{E}_{mn}(t)$ denote the class of all sets $E\subset T$ such that $E\subset(t-1/m, t+1/m)$ and that $\alpha(E)>1/n$ and $\mathscr{F}_{mn}(t)$ the class of all sets in $\mathscr{F}_{\mathrm{rat}}$ with the same property. The following proposition that will be useful later can be proved by a routine.

PROPOSITION (2.2).

(i) *$a\in L(f, t)$ if and only if for every $\varepsilon>0$, we can find $E_n \underset{r}{\rightarrow} t$ such that*

$$\frac{1}{|E_n|}\int_{E_n}|\operatorname{atn} f(s)-\operatorname{atn} a|\,ds<\varepsilon, \qquad \mathrm{n}=1, 2, \cdots$$

(ii) $$\operatorname{atn}\bar{f}(t)=\sup_n \inf_m \sup_{E\in\mathscr{E}_{mn}(t)} \frac{1}{|E|}\int_E \operatorname{atn} f(s)ds$$

$\mathcal{E}_{mn}(t)$ can be replaced by $\mathcal{F}_{mn}(t)$. $\mathrm{atn}\,\underline{f}(t)$ *is also expressed similarly.*

DEFINITION. $f \in \tilde{M}$ is called a canonical measurable function, if $f(t) \in L(f, t)$ for every $t \in T$. $M \equiv M(T)$ denotes the space of all canonical functions.

We shall use M in behalf of $\tilde{M}$. This is justified by the following

PROPOSITION (2.3). *For every $f \in M$ we have at least one $g \in M$ equal to f a.e. on T. All such g's can be obtained by taking any point in $L(f, t)$ for the value of g at each t.*

"$f = g$ a.e." is a strong condition in M as is seen in the following propositions (2.4), (2.5) and (2.6).

PROPOSITION (2.4). *The following conditions on $f, g \in M$ are equivalent, where U is an open subset of T.*

(i) *$f = g$ a.e. on U,*

(ii) *$L(f, t) = L(g, t)$ for every $t \in U$,*

(iii) *$\bar{f} = \bar{g}$ on U,*

(iv) *$\underline{f} = \underline{g}$ on U.*

Let $C = C(T)$ denote the space of all continuous functions. Then it is obvious that $C \subset M$.

PROPOSITION (2.5). *Suppose $f \in C$, $g \in M$ and U is open in T. Then we have*

(i) *$\bar{f} = \underline{f} = f$: $L(f, t)$ is a single point for every $t \in U$,*

(ii) *$g \leqq (\geqq) f$ a.e. on $U \Rightarrow g \leqq (\geqq) f$ everywhere on U,*

(iii) *$g = f$ a.e. on $U \Rightarrow g = f$ everywhere on U.*

The case $f \equiv$ const in (ii) will be useful later.

Let $D \equiv D(T)$ denote the space of all functions having no discontinuities of the second kind. We assume $f \in D$ to be right or left continuous at every jump point and continuous at the end points (if any). It is clear that $D \subset M$.

PROPOSITION (2.6). *Suppose $f \in D$ and $g \in M$. If $f = g$ a.e. on T, then*

(i) *$g \in D$,*

(ii) *The set of continuity points of g is the same as that for f: $g = f$ on that set and $g(t) = f(t+)$ or $f(t-)$ elsewhere.*

§ 3. The σ-algebras $\mathcal{B}_\rho$, $\mathcal{B}_K$ and $\mathcal{B}$.

The space M of all canonical functions on T is topologized by the following pseudo-metric:

$$\rho(f, g) = \int_T |\mathrm{atn}\, f(t) - \mathrm{atn}\, g(t)| \frac{dt}{1+t^2} \qquad (\mathrm{atn} = \mathrm{arc\,tan});$$

$\rho(f, g) = 0$ does not always imply $f = g$ but only $f = g$ a.e. (see Propositions (2.4), (2.5) and (2.6)). The *ρ-convergence* is equivalent to the convergence in measure on every compact subset of T. The space C of all continuous functions is ρ-dense in M. It is easy to see that the pseudo-metric space $M(\rho)$ is complete and separable. Referring to the *ρ-topology*, we can define *ρ-open sets*, *ρ-Borel sets*, *ρ-continuous functions*, *lower semi-continuous* (ρ) functions *etc.*

Let $\mathscr{B}_\rho \equiv \mathscr{B}_\rho(M)$ denote the σ-algebra of all ρ-Borel sets in M. Another characterization for $\mathscr{B}_\rho$ will be given in Theorem (3.11).

A second σ-algebra $\mathscr{B}_K \equiv \mathscr{B}_K(M)$ is defined to be the σ-algebra generated by

$$\{f \in M : f(t) < a\}, \qquad t \in T, \qquad a \in R.$$

There is no inclusion relation between these two σ-algebras, as we can see in Theorems (3.12) and (3.13). Therefore we shall introduce a third σ-algebra $\mathscr{B} \equiv \mathscr{B}(M)$. It is the join $\mathscr{B}_\rho \vee \mathscr{B}_K$, the least σ-algebra containing both, so that every $\mathscr{B}_\rho$- or $\mathscr{B}_K$-measurable set or function is $\mathscr{B}$-measurable. $\mathscr{B}$ is the σ-algebra we refer to in discussing probability measures on M.

The following notations will be used in this section.

$$A(E,f) = \int_E \operatorname{atn} f(s)ds \qquad E \text{ bounded}$$

$$S(E,f) = \sup f(s), \qquad s \in E$$

$$L(E,f) = \inf f(s), \qquad s \in E$$

$$a(E) = \sup(1+s^2), \qquad s \in E \qquad E \text{ bounded.}$$

PROPOSITION (3.1). *$A(E,f)$ is ρ-continuous in f for E fixed.*

PROOF. $|A(E,f) - A(E,g)| \leqq a(E)\rho(f,g)$.

THEOREM (3.2). *Let U be a non-empty open subset of T, Γ the class of all $E \subset U$ with $|E| < \infty$ and $\Gamma' = \Gamma \cap \mathscr{I}_{\text{rat}}$. Then we have*

$$\operatorname{atn} S(U,f) = \sup_{E \in \Gamma} \frac{A(E,f)}{|E|} = \sup_{I \in \Gamma'} \frac{A(I,f)}{|I|}$$

and a similar identity for $L(U,f)$. $S(U,f)$ ($L(U,f)$) is lower (upper) semi-continuous (ρ) and so $\mathscr{B}$-measurable in f for U fixed.

PROOF. We shall discuss $S(U,f)$ only. First we shall prove that the following conditions are equivalent, where c is a constant.

(a) $\operatorname{atn} f \leqq c$ *on* U,

(b) $A(E,f) \leqq c|E|$ *for every* $E \in \Gamma$,

(c) $A(I,f) \leqq c|I|$ *for every* $I \in \Gamma'$.

(a)$\Rightarrow$(b)$\Rightarrow$(c) is obvious. Assuming (c), we have $f \leqq c$ a.e. on U by the density theorem and so $f \leqq c$ on U by Proposition (2.5) (ii). Hence (c) implies (a). Thus (a), (b) and (c) are equivalent. This proves the identity in the theorem. As $A(I,f)$ is ρ-continuous in f by (3.1), $S(U,f)$ is lower semi-continuous (ρ).

COROLLARY (3.3). *Let U be a non-empty open subset of T. Then $\{f \in M : f \leqq c$ on $U\}$ and $\{f \in M : f \geqq c$ on $U\}$ are ρ-closed and so $\mathscr{B}$-measurable.*

THEOREM (3.4). $C \in \mathscr{B}_\rho$.

PROOF. We shall discuss the case $T=[0,1]$ only. Let U_{ni} be the intersection of the interval $((i-1)/n, (i+1)/n)$ with T, where i ranges over $-1<i<n+1$ so that U_{ni} is non-empty. As each $f \in C$ is bounded and uniformly continuous on T, we have

$$C=\bigcap_m \bigcup_n \bigcap_i \bigcup_{k=-m^2}^{m^2} \left\{f \in M: S(U_{ni}, f) \leqq \frac{k+1}{m},\ L(U_{ni} f) \geqq \frac{k-1}{m}\right\}.$$

The set in the bracket is ρ-closed by (3.3) and so $C \in \mathscr{B}_\rho$.

THEOREM (3.5). $D \in \mathscr{B}_\rho$.

PROOF. We shall discuss the case $T=[0,1]$ only. As every $f \in D$ is bounded on T, we have

$$D \subset \bigcup_n B_n, \qquad B_n=\{f \in M: |f| \leqq n \quad \text{on} \quad T\}.$$

B_n is ρ-closed by (3.3). If we can prove that the set $\Gamma_k(r, r')$ defined by the condition on $f \in M$:

$$\text{(c)} \qquad \exists t_1<t_2<\cdots<t_{2k} \quad \forall i \quad f(t_{2i-1})<r, \quad f(t_{2i})>r'$$

is ρ-open for $r<r'$, then we have

$$B_n-D=B_n \cap \bigcup_{r<r'} \bigcap_k \Gamma_k(r, r') \qquad r, r' \text{ rational}$$

and so

$$D=\bigcup_n B_n \cap D=\bigcup_n [B_n-(B_n-D)] \in \mathscr{B}_\rho.$$

It remains only to prove that $\Gamma_k(r, r')$ is ρ-open. Consider the following condition on $f \in M$:

(c′) $\exists E_1, E_2, \cdots, E_{2k} \subset T$ such that

(i) $|E_i|>0 \qquad i=1,2, \cdots, 2k$

(ii) $A(E_{2i-1}, f)<\operatorname{atn} r \cdot |E_{2i-1}|, \quad A(E_{2i}, f)>\operatorname{atn} r \cdot |E_{2i}| \qquad i=1,2, \cdots, k$

(iii) $E_1<E_2<\cdots<E_{2k}$, namely

$$a_1<a_2<\cdots<a_{2k} \text{ for every } a_i \in E_i,\ 1<i<2k.$$

As $A(E, f)$ is continuous in $f \in M$ by (3.1), (c′) determines a ρ-open set as a union of ρ-open sets. To prove that $\Gamma_k(r, r')$ is open, it suffices to show the equivalence of (c) and (c′). As $A(f, E) \gtrless \operatorname{atn} r \cdot |E|$ implies the existence of $t \in E$ with $f(t) \gtrless r$, (c′) implies (c). Suppose (c) holds. Recalling $f(t_i) \in L(f, t_i)$ for $f \in M$, we have a neighborhood U_i of t_i for each i such that

$$|\{f<r\} \cap U_{2i-1}|>0,\ |\{f>r'\} \cap U_{2i}|>0 \qquad i=1,2, \cdots, k.$$

Write E_{2i-1} and E_{2i} for these t-sets. Then the sets $\{E_i\}$ satisfy. (i) and (ii) in (c′). By taking all U_i small enough, we can assume $U_1 < U_2 < \cdots < U_{2k}$. Therefore E_i, $i=1, \cdots, 2k$ satisfy (iii) in (c′). Thus (c) implies (c′). This completes the proof.

THEOREM (3.6). *The p-th norm* $\|f\|_p$ $(1 \leqq p \leqq \infty)$ *is lower semi-continuous* (ρ) *in* $f \in M$.

PROOF. The set $\{f \in M : \|f\|_\infty \leqq c\}$ is $\{f \in M : -c \leqq f \leqq c$ on $T\}$ and so ρ-closed by (3.3). Hence $\|f\|_\infty$ is lower-semi-continuous (ρ) in f. Suppose $p < \infty$. Consider first the functions

$$F_n(f) = \int_{T \wedge [-n, n]} \min(|f(t)|^p, n) dt \qquad n = 1, 2, \cdots.$$

Since the ρ-convergence is equivalent to the convergence in measure on every compact subset of T, $F_n(f)$ is ρ-continuous in f for each n. But $F_n(f) \uparrow \|f\|_p$ as $n \to \infty$ and so $\|f\|_p$ is lower semi-continuous (ρ).

COROLLARY (3.7). $L^p \equiv \{f \in M : \|f\|_p < \infty\} \in \mathscr{B}_\rho$.

THEOREM (3.8). *For t fixed,* $\bar{f}(t)$ *and* $\underline{f}(t)$ *are* $\mathscr{B}_\rho$*-measurable in* $f \in M$.

PROOF. This follows at once from Propositions (2.2) (ii) and (3.1).

Now we shall discuss the measurability of $\bar{f}(t)$ and $\underline{f}(t)$ in the pair $(t, f) \in T \times M$. Let $\mathscr{B}(T) \otimes \mathscr{B}_\rho$ denote the product σ-algebra. It is the same as the σ-algebra $\mathscr{B}(T \times M)$ of all Borel subsets of $T \times M$ with respect to the product topology.

PROPOSITION (3.9). *Suppose* $I \equiv [u-\varepsilon, v+\varepsilon] \subset T$ *and* $E \subset (-\varepsilon, \varepsilon)$. *Then* $E_t \equiv \{s+t : s \in E\} \subset I$ *for* $t \in [u, v]$ *and* $A(E_t, f)$ *is continuous in* $(t, f) \in [u, v] \times M$.

PROOF.

$$\begin{aligned}
&|A(E_t, f) - A(E_s, g)| \\
&\leqq |A(E_t, f) - A(E_s, f)| + |A(E_s, f) - A(E_s, g)| \\
&\leqq \int_{E_t \triangle E_z} |\operatorname{atn} f(\theta)| d\theta + \int_I |\operatorname{atn} f(\theta) - \operatorname{atn} g(\theta)| d\theta \\
&\leqq \frac{\pi}{2} |E_t \triangle E_s| + a(I)\rho(f, g) \\
&= \frac{\pi}{2} |E_{t-s} \triangle E| + a(I)\rho(f, g) \\
&\to 0, \text{ as } |t-s| + \rho(f, g) \to 0.
\end{aligned}$$

THEOREM (3.10). *Both* $\bar{f}(t)$ *and* $\underline{f}(t)$ *are measurable* $(\mathscr{B}(T) \otimes \mathscr{B}_\rho)$ *in the pair* $(t, f) \in T \times M$.

PROOF. We shall discuss the case $T = [u, v)$ only. Let $\{I_p \equiv [u_p, v_p]\}$ be a sequence of intervals such that

$$u_1 > u_2 > \cdots \to u, \qquad v_1 < v_2 < \cdots \to v.$$

It suffices to show that

$$\{(t, f) \in T \times M : \bar{f}(t) < c\} \in \mathscr{B}(T) \otimes \mathscr{B}_\rho .$$

This set is the union of the following sets

$$\{u\} \times \{f \in M : \bar{f}(u) < c\}, \quad \{(t, f) \in I_p \times M : \bar{f}(t) < c\}, \qquad p = 1, 2, \cdots .$$

The first set is measurable ($\mathscr{B}(T) \otimes \mathscr{B}_p$) by (3.8). It is therefore enough to prove that $\bar{f}(t)$ is measurable ($\mathscr{B}(I_p) \otimes \mathscr{B}_p$) in $(t, f) \in I_p \times M$ for each p. Noticing that $[u_p - 1/m_0, v_p + 1/m_0] \subset T$ for some big m_0, we have

$$\operatorname{atn} \bar{f}(t) = \sup_n \inf_{m < m_0} \sup_{E \in \mathscr{F}_{nm}(0)} A(E_t, f)/|E| \qquad t \in I_p, f \in M$$

by Proposition (2.2) (ii). As $A(E_t, f)$ is continuous in $(t, f) \in I_p \times M$ by (3.9), $\bar{f}(t)$ is measurable ($\mathscr{B}(I_p) \otimes \mathscr{B}_p$) in $(t, f) \in I_p \times M$.

Now we shall prove some facts that will show the difference between $\mathscr{B}_\rho$ and $\mathscr{B}_K$. First we shall prove

THEOREM (3.11). *$\mathscr{B}_p$ is generated by the sets*

$$\{f \in M : \int_I \operatorname{atn} f(t) dt < a\}$$

where I ranges over all compact intervals $\subset T$ and a ranges over R.

PROOF. Let $\mathscr{B}'$ denote the σ-algebra generated by the sets above. It is obvious by (3.1) that $\mathscr{B}' \subset \mathscr{B}_\rho$. To prove the opposite inclusion relation, it is enough to prove that $\rho(f, g)$ is $\mathscr{B}'$-measurable in f for g fixed. For this purpose it suffices to show that

$$F(f) \equiv \int_I |\operatorname{atn} f(t) - \operatorname{atn} g(t)| \frac{dt}{1+t^2}$$

is $\mathscr{B}'$-measurable for every compact interval $I \subset T$, because $\rho(f, g)$ is the limit of a sequence of functions of this form. Let $I = I_{n1} \cup I_{n2} \cup \cdots \cup I_{nn}$ be a non-overlapping decomposition of I into n compact intervals with equal length and set

$$\varphi_n(t) = \frac{1}{|I_{ni}|} \int_{I_{ni}} \operatorname{atn} f(s) ds ,$$

$$\psi_n(t) = \frac{1}{|I_{ni}|} \int_{I_{ni}} \operatorname{atn} g(s) ds , \qquad t \in I_{ni}, \quad i = 1, 2, \cdots, n .$$

By the density theorem we have, as $n \to \infty$,

$$\varphi_n(t) \to \operatorname{atn} f(t), \qquad \psi_n(t) \to \operatorname{atn} g(t) \qquad \text{a.e. on } I .$$

$$F_n(f) \equiv \int_I |\varphi_n(s) - \psi_n(s)| \frac{ds}{1+s^2} \to F(f) \quad \text{as} \quad n \to \infty .$$

Observing

$$F_n(f)=\sum_i\left|\frac{1}{|I_{ni}|}\int_{I_{ni}}\operatorname{atn} f(s)ds - \frac{1}{|I_{ni}|}\int_{I_{ni}}\operatorname{atn} g(s)ds\right|\int_{I_{ni}}\frac{dt}{1+t^2},$$

we can see that $F_n(f)$ is $\mathscr{B}'$-measurable, so that $F(f)$ is also $\mathscr{B}'$-measurable.

THEOREM (3.12). *$\mathscr{B}_K-\mathscr{B}_\rho$ is not empty.*

PROOF. Take a point t_0 strictly inside of T and consider the set

$$A=\{f\in M: f(t_0)=1\}.$$

$A\in\mathscr{B}_K$ is clear. We shall prove $A\notin\mathscr{B}_\rho$. Consider two functions $f_1, f_2\in M$

$$f_1(t)=1 \quad\text{for}\quad t\leqq t_0, \quad =0 \quad\text{for}\quad t>t_0$$

$$f_2(t)=1 \quad\text{for}\quad t<t_0, \quad =0 \quad\text{for}\quad t\geqq t_0.$$

As $\rho(f_1, f_2)=0$, *either* both of f_1, f_2 are in A *or* none of f_1, f_2 is in A, if $A\in\mathscr{B}_\rho$. But $f_1\in A$ and $f_2\notin A$. Therefore $A\notin\mathscr{B}_\rho$.

THEOREM (3.13). *$\mathscr{B}_\rho-\mathscr{B}_K$ is not empty.*

PROOF. Take a compact interval $I\subset T$ and consider the set

$$A=\{f\in M: \int_I \operatorname{atn} f(s)ds=|I|\}.$$

It is obvious by (3.11) that $A\in\mathscr{B}_\rho$. We shall prove $A\notin\mathscr{B}_K$. Suppose $A\in\mathscr{B}_K$. Then we have a countable subset Q of T with the following property

$$g\in M,\ f\in A,\ g=f \ \text{ on }\ Q \Rightarrow g\in A. \tag{3.14}$$

Since $|Q|=0$, we have an open neighborhood U of Q with $|U|<|I|/2$. Consider two functions $f_1, f_2\in M$

$$f_1(t)=\alpha \quad\text{on}\quad T \qquad\qquad (\alpha=\tan 1)$$

$$f_3(t)=\alpha \quad\text{on}\quad U, \quad =0 \qquad \text{elsewhere}.$$

$f_1\in A$ is clear. Since $f_2\in\tilde{M}$, we have $f_3\in M$ such that $f_3=f_1$ a. e. Since $f_3=f_1=\alpha$ a. e. on U, $f_3=\alpha$ on U by Proposition (2.5) (iii). Therefore $f_1=f_3=\alpha$ on Q. But $f_1\in A$. Therefore $f_2\in A$ by (3.14). This is a contradiction, because

$$\int_I \operatorname{atn} f_3(s)ds=\int_I \operatorname{atn} f_2(s)ds=|U\cap I|\leqq|U|<|I|/2.$$

§ 4. Canonical stochastic process.

A stochastic process $\{X_t(\omega), t\in T, \omega\in\Omega, \mathscr{F}, P)\}$ is defined to be a family of real random variables indexed by the time parameter t ranging over T. Fixing ω and changing t in $X_t(\omega)$, we have a function of $t\in T$ which is an element of $\bar{R}^T$. This is denoted by $X.(\omega)$ and is called the *sample path* of the process corresponding to ω. $X.(\omega)$ is considered a function of ω ranging in $\bar{R}^T$ in general. Let $\mathcal{B}[X]$ be the σ-algebra (of subsets of Ω) that is generated by the sets

$$\{\omega : X_t\omega) < a\}, \qquad t \in T, \qquad a \in R\,.$$

$\mathcal{B}[X] \subset \mathcal{F}$ is obvious.

$\{X_t\}$ is said to be *continuous in probability* if

$$\lim_{s\to t} P\{|X_s - X_t| > \varepsilon\} = 0 \qquad \varepsilon > 0,\ t \in T\,,$$

or equivalently if

$$\lim_{r\to t} \boldsymbol{E}\{|\operatorname{atn} X_s - \operatorname{atn} X_t|\} = 0 \qquad t \in T\,,$$

where $\boldsymbol{E}(Y) = \int_\Omega Y(\omega)P(d\omega)$. This can also be stated as follows:

atn X_t is continuous in t with respect to the norm in $L^1(\Omega, \mathcal{F}, P)$.

If this condition is satisfied, then atn X_t is uniformly continuous on every compact $E \subset T$ with respect to this norm. Therefore we have $\delta = \delta(E, \varepsilon)$ for $\varepsilon > 0$ such that

$$|t-s| < \delta, \qquad t, s \in E \Rightarrow \boldsymbol{E}\{|\operatorname{atn} X_s - \operatorname{atn} X_t|\} < \varepsilon\,. \tag{4.1}$$

Definition (4.2). A stochastic process $\{X_t(\omega), t \in T\}$ is called *canonical*, if the following two conditions are satisfied:

(C. 1) $X.(\omega) \in M$ for every $\omega \in \Omega$

(C. 2) The map $\omega \to X.(\omega)$ from Ω into M is measurable $(\mathcal{F}, \mathcal{B})$, i. e. $X.^{-1}(\mathcal{B}) \subset \mathcal{F}$. ($M$ and $\mathcal{B}$ were defined in the previous sections.) In other words, $X.(\omega)$ is an $(M, \mathcal{B})$-valued random variable, so that we can define the probability law of $X.(\omega)$ which is a complete $\mathcal{B}$-regular probability measure on M (see Remarks in Section 1). As it is easy to see

$$X.^{-1}(\mathcal{B}_K) = \mathcal{B}[X_t, t \in T]\,,$$

(C. 2) can be replaced by a weaker condition:

(C. 2′) $$X.^{-1}(\mathcal{B}_\rho) \subset \mathcal{F}\,.$$

Theorem (4.3). (*Measurability of canonical processes*). *Let* $\{X_t(\omega)\}$ *be a canonical process and m the product* (*complete*) *measure of the Lebesgue measure on T and the measure P on* Ω. *Then* $X_t(\omega)$ *is measurable* $(\overline{\mathcal{B}(T) \otimes \mathcal{F}^m})$ *in the pair* $(t, \omega) \in T \times \Omega$.

Proof. First we shall prove the measurability of $\overline{X.(\omega)}(t)$, where the top bar means the approximate upper limit introduced in Section 2. As the map $\omega \to X.(\omega)$ is measurable $(\mathcal{F}, \mathcal{B})$ by (C. 2), the map $(t, \omega) \to (t, X.(\omega))$ is measurable $(\mathcal{B}(T) \otimes \mathcal{F}, \mathcal{B}(T) \times \mathcal{B})$. By Theorem (3.10), the map $(t, f) \to \bar{f}(t)$ is measurable $(\mathscr{B}(T)\otimes\mathscr{B}_\rho)$ and so measurable $(\mathscr{B}(T)\otimes\mathscr{B})$. Composing these two maps, we can see that $(t,\omega)\to\overline{X.(\omega)}(t)$ is measurable $(\mathscr{B}(T)\otimes\mathscr{F})$ and so measurable $(\overline{\mathscr{B}(T)\otimes\mathscr{F}^m})$, namely that $\overline{X.(\omega)}(t)$ is measurable $(\overline{\mathscr{B}(T)\otimes\mathscr{F}^m})$ in the pair (t,ω). Similarly for $\underline{X.(\omega)}(t)$. As it is obvious that $\underline{X.(\omega)}(t)\leqq X_t(\omega)\leqq\overline{X.(\omega)}(t)$ for every pair (t,ω), we get the measurability $(\overline{\mathscr{B}(T)\otimes\mathscr{F}^m})$ of $X_t(\omega)$ in the pair (t,ω), observing

$$m\{(t,\omega):\underline{X.(\omega)}(t)<\overline{X.(\omega)}(t)\}$$

$$=\int_\Omega|\{t\in T:\underline{X.(\omega)}(t)<\overline{X.(\omega)}(t)\}|P(d\omega)\qquad\text{by Fubini's theorem}$$

$$=0\qquad\text{by } X.(\omega)\in M.$$

THEOREM (4.4). (*Separability of canonical processes continuous in probability*). *Every canonical process $\{X_t(\omega)\}$ continuous in probability is separable relative to closed sets with respect to every countable dense subset Q of T.*

PROOF. Let us write $Y_t(\omega)$ for $\operatorname{atn} X_t(\omega)$. Let $\mathscr{F}_{\text{rat}}$ be the set system introduced in Remarks in Section 1 and Θ a countable dense subset of $[-\pi/2, \pi/2]$.

First we shall find, for each pair $(E,\theta)\in\mathscr{F}_{\text{rat}}\times\Theta$, a sequence of non-overlapping decompositions of E into intervals

$$E=I_{n1}\cup I_{n2}\cup\cdots\cup I_{np(n)},\qquad n=1,2,\cdots,$$

and a system of points

$$r_{ni}\in I_{ni}\cap Q\,,\qquad i=1,2,\cdots,p(n)\,,$$

with the following property:

$$\int_E|Y_s(\omega)-\theta|\,dt=\lim_n\sum_i|Y(r_{ni},\omega)-\theta|\,|I_{ni}|\qquad\text{a.e. on }\Omega\,. \tag{4.5}$$

Since E is compact, we can take $\delta(\varepsilon)=\delta(E,\varepsilon)$ in (4.1) to get

$$|t-s|<\delta(\varepsilon),\ t,s\in E\Rightarrow \boldsymbol{E}(|Y_t-Y_s|)<\varepsilon\,. \tag{4.1$'$}$$

As E is a finite sum of compact intervals, E can be decomposed into non-overlapping intervals with length $<\delta(\varepsilon)$, say

$$E=I_1\cup I_2\cup\cdots\cup I_p\,.$$

Since Q is dense in T, $Q\cap I_i$ is non-empty. Take a point r_i from $Q\cap I_i$. Then we have

$$\boldsymbol{E}\Big[\Big|\int_E |Y_s-\theta|\,ds-\sum_i |Y(r_i)-\theta|\,|I_i|\Big|\Big]$$
$$\leqq \sum_i \boldsymbol{E}\Big[\int_{I_i}\Big||Y_s-\theta|-|Y(r_i)-\theta|\Big|ds\Big]$$
$$\leqq \sum_i \boldsymbol{E}\Big[\int_{I_i}|Y_s-X(r_i)|\,ds\Big].$$

As $Y_s(\omega)$ is measurable $(\overline{\mathcal{B}(T)\otimes\mathcal{F}^m})$ in (s, ω) by Theorem (4.3), we can exchange the order of $\boldsymbol{E}$ and $\int$. Using (4.1′), we get

$$\boldsymbol{E}\Big[\Big|\int_E |Y_s-\theta|\,ds-\sum_i |Y(r_i)-\theta|\,|I_i|\Big|\Big]<\varepsilon|E|.$$

Writing I_i and r_i for $\varepsilon=2^{-n}$ as I_{ni} and r_{ni} respectively, we have

$$\boldsymbol{E}\Big[\sum_n\Big|\int_E |Y_s-\theta|\,ds-\sum_i |Y(r_{ni})-\theta|\,|I_{ni}|\Big|\Big]<\infty,$$

which implies (4.5).

Writing $\Omega(E, \theta)$ for the set of all ω for which (4.5) holds and setting

$$\Omega_1=\cap\,\Omega(E, \theta), \qquad E\in\mathcal{F}_{\mathrm{rat}},\ \theta\in\Theta$$

we have $P(\Omega_1)=1$, because $\mathcal{F}_{\mathrm{rat}}\times\Theta$ is countable. To prove our theorem, it suffices to show that for every $\omega\in\Omega_1$, every closed $F\subset[-\pi/2, \pi/2]$, every open $U\subset T$ and every $t\in U$, if $Y_r(\omega)\in F$ for $r\in Q\cap U$ then $Y_t(\omega)\in F$; notice here that atn gives a homeomorphism from $[-\infty, \infty]$ onto $[-\pi/2, \pi/2]$. For this purpose it is enough to find, for every $\varepsilon>0$, $r\in Q\cap U$ such that

$$|Y_r(\omega)-Y_t(\omega)|<\varepsilon,$$

because F is closed. As $X.(\omega)\in M$, we have $X_t(\omega)\in L(X., t)$ and so we can find $E\in\mathcal{F}_{\mathrm{rat}}$ such that $t\in E\subset U$ and that

$$\frac{1}{|E|}\int_E |Y_s(\omega)-Y_t(\omega)|\,ds<\varepsilon/3$$

by virtue of Proposition (2.2) (i). Since $Y_t(\omega)\in[-\pi/2, \pi/2]$ and Θ is dense in this interval, we can find $\theta\in\Theta$ such that

$$|Y_t(\omega)-\theta|<\varepsilon/3. \tag{4.6}$$

These two inequalities will imply

$$\frac{1}{|E|}\int_E |Y_s(\omega)-\theta|\,ds<2\varepsilon/3.$$

As $\omega\in\Omega_1$ and $E\in\mathcal{F}_{\mathrm{rat}}$, we can use (4.5) to get n such that

$$\frac{1}{|E|}\sum_i |Y(r_{ni}, \omega)-\theta|\,|I_{ni}|<2\varepsilon/3.$$

As $|E|=\sum_i |I_{ni}|$, we can find i such that

$$|Y(r_{ni}, \omega)-\theta|<2\varepsilon/3,$$

which, combined with (4.6), implies

$$|Y(r_{ni}, \omega)-Y_t(\omega)|<\varepsilon.$$

As $r_{ni}\in Q\cap I_{ni}\subset Q\cap E\subset Q\cap U$, $r=r_{ni}$ is what we wanted to find out. This completes the proof.

Let $\{X_t(\omega)\}$ be a stochastic process. A canonical stochastic process $\{S_t(\omega)\}$ (defined on the some probability space) is called a *canonical modification* of $\{X_t(\omega)\}$ if

$$P\{X_t=S_t\}=1 \qquad \text{for every } t\in T.$$

CANONICAL MODIFICATION THEOREM (4.7). *Every stochastic process $\{X_t(\omega)\}$ continuous in probability has a canonical modification. If we have two such modifications $\{S_t\}$ and $\{S_t'\}$ for the same process, then we have*

$$P\{S_t=S_t' \qquad \textit{for almost every } t\}=1$$

in addition to the automatic property:

$$P\{S_t=S_t'\}=1 \qquad \textit{for every } t\in T.$$

PROOF. We shall discuss the case $T=[0, 1]$ only, because a small change in the proof will take care of the other cases. The metric ρ in $M=M(T)$ can now be replaced by a simpler one

$$\rho(f, g)=\int_T |\operatorname{atn} f(t)-\operatorname{atn} g(t)|\,dt,$$

which induces the same topology and therefore the same σ-algebras $\mathcal{B}_\rho$ and $\mathcal{B}(\equiv \mathcal{B}_\rho\vee\mathcal{B}_K)$. Let us write Y_t, Y_t^n and V_t respectively for $\operatorname{atn} X_t$, $\operatorname{atn} X_t^n$ and $\operatorname{atn} U_t$. By (4.1) we can take $\delta=\delta(\varepsilon)$ for $\varepsilon>0$ such that

$$(4.1'') \qquad |s-t|<\delta, \quad s, t\in T\Rightarrow \boldsymbol{E}(|Y_s-Y_t|)<\varepsilon.$$

It is harmless to assume that $\delta(\varepsilon)\downarrow 0$ as $\varepsilon\downarrow 0$.

Let us consider for each $n=1, 2, \cdots$, a finite decomposition of I into non-overlapping compact intervals:

$$I=I_{n1}\vee I_{n2}\vee\cdots\vee I_{np(n)}, \quad |I_{ni}|<\delta(2^{-n})$$

and a corresponding step processes

$$X_t^n(\omega)=X(t_{ni}, \omega) \qquad \text{for } I_{ni},\ i=1, 2, \cdots, p(n),$$

where t_{ni} is a point in I_{ni}, say the left end point.

$\rho(X_\cdot^n(\omega), X_\cdot^{n+1}(\omega)$ is a Borel measurable function of $X(t_{ni}, \omega)$, $i=1, 2, \cdots, p(n)$ and $X(t_{nj}, \omega)$, $j=1, 2, \cdots, p(n+1)$ and so measurable $(\mathcal{F})$. $X_t^n(\omega)$ is also measurable $(\mathscr{B}(T)\otimes\mathscr{B}[X])$ in (t, ω) and

$$\boldsymbol{E}[\rho(X_\cdot^n, X_\cdot^{n+1})]$$

$$\leqq \int_E \boldsymbol{E}[\,|Y_t^n - Y_t^{n+1}|\,]dt < 2^{-n}.$$

Therefore we have

$$\boldsymbol{E}[\sum_n \rho(X_\cdot^n, X_\cdot^{n+1})] < \infty,$$

It follows from this that there exists an $\bar{R}$-valued function $W_t(\omega)$ measurable in $(t, \omega)\in T\times\Omega$ such that

(i) $$\int_T |\arctan X_t^n(\omega) - \arctan W_t(\omega)|\,dt \to 0$$

for almost every $\omega\in\Omega$,

(ii) $X_t^n(\omega)\to W_t(\omega)$ for almost every $(t, \omega)\in T\times\Omega$.

Since $X_s(\omega)$ is continuous in probability, it follows from (ii) that

(iii) for almost every t fixed, we have

$$W_t(\omega) = X_t(\omega) \quad \text{for almost every} \quad \omega\in\Omega.$$

By Fubini's theorem we get

$$\int_\Omega |\{t : |W_t(\omega)| = \infty\}|\,P(d\omega)$$

$$= \int_T P\{\omega : |W_t(\omega)| = \infty\}dt$$

$$= \int_T P\{\omega : |X_t(\omega)| = \infty\}dt$$

$$= 0.$$

Therefore $W_\cdot(\omega)\in\tilde{M}$ for almost every ω. Let

$$U_t(\omega) = \overline{W}_t(\omega) \qquad (= \text{the approximate upper limit of } W_\cdot(\omega) \text{ at } t).$$

Then $U_\cdot(\omega)\in M$ for almost every ω and (i) implies that

$$P(\Omega') = 1 \quad \text{for} \quad \Omega' = \{\omega : U_\cdot(\omega)\in M,\ \rho(X_\cdot^n(\omega), U_\cdot(\omega))\to 0\}.$$

As $\rho(X_\cdot^n(\omega), f)$ is equal to a Borel measurable function of $X(t_{ni}, \omega)$, $i = 1, 2, \cdots, p(n)$ for each $f\in M$, it is measurable $(\mathscr{F})$ in ω. Therefore $\rho(U_\cdot(\omega), f)$ is also measurable $(\mathscr{F})$ in ω, because

$$\rho(U_\cdot(\omega), f) = \lim_n \rho(X_\cdot^n(\omega), f) \qquad \text{for } \omega\in\Omega'$$

$$\Omega'\in\mathscr{F}, \qquad P(\Omega') = 1.$$

This shows that the set $\{\omega : U.(\omega) \in N_\rho(f, r)\}$ $(N_\rho(f, r) = \{g \in M : \rho(g, f) < r\})$ is measurable $(\mathcal{F})$. As $\mathcal{B}_\rho$ is generated by the sets $N_\rho(f, r)$, $r > 0, f \in M$, $\{\omega : U.(\omega) \in B\}$ is measurable $(\mathcal{F})$ for $B \in \mathcal{B}_\rho$.

We shall modify $\{U_t(\omega)\}$ at each point t to get a canonical modification $\{S_t(\omega)\}$. Fix the time point t for the moment. Observing

$$\int_I | V_s(\omega) - Y_t(\omega) | ds \mathop{\gtreqless} \int_I | Y_s^n(\omega) - Y_t(\omega) | ds \pm \rho(U.(\omega), X.^n(\omega)),$$

we have

$$\int_I | V_s(\omega) - Y_t(\omega) | ds = \lim_n \int_I | Y_s^n(\omega) - Y_t(\omega) | ds \qquad \text{for } \omega \in \Omega',$$

where I is a compact interval $\subset T$. This shows that the left side is measurable $(\mathcal{F})$. Noticing $P(\Omega') = 1$, we have

$$\boldsymbol{E}\Big[\frac{1}{|I|}\int_I | V_s - Y_t | ds\Big] = \lim_n \boldsymbol{E}\Big[\frac{1}{|I|}\int_I | Y_s^n - Y_t | ds\Big].$$

Since $X_s^n(\omega)$ (and so $Y_s^n(\omega)$) is measurable $(\mathcal{B}(T) \otimes \mathcal{F})$ in (s, ω) and $X_t(\omega)$ (and so $Y_t(\omega)$) is measurable $(\mathcal{F})$ in ω and so measurable $(\mathcal{B}(T) \otimes \mathcal{F})$ in (s, ω), we can use Fubini's theorem to get

$$\boldsymbol{E}\Big[\frac{1}{|I|}\int_I | V_s - Y_t | ds\Big] = \frac{1}{|I|}\int_I \boldsymbol{E}[| V_s - V_t |] ds < \varepsilon, \qquad \text{if } |I| < \delta(\varepsilon).$$

Taking a sequence $\{I_n\}$ with $t \in I_n$ and $|I_n| < \delta(2^{-n})$, we have

$$\boldsymbol{E}\Big[\sum_n \frac{1}{|I_n|}\int_{I_n} | V_s - Y_t | ds\Big] < \sum_n 2^{-n} < \infty.$$

Write Ω_t for the set of all ω for which the infinite series converges. Since each term is measurable $(\mathcal{F})$ as we mentioned above, Ω_t is measurable $(\mathcal{F})$ and $P(\Omega_t) = 1$. It is easy to see by Proposition (2.2) (i) that

$$X_t(\omega) \in L(U., t) \qquad \text{for } \omega \in \Omega_t.$$

Set

$$S_t(\omega) = \begin{array}{ll} X_t(\omega) & \text{for } \omega \in \Omega_t \\ U_t(\omega) & \text{for } \omega \in \Omega - \Omega_t. \end{array}$$

Define $S_t(\omega)$ at each point $t \in T$ in such a way.

Now fix ω. By our construction we have

$$U.(\omega) \in M \quad \text{and} \quad S_t(\omega) \in L(U., t) \qquad \text{for every } t \in T.$$

By Proposition (2.3) we have

$$S.(\omega) \in M \quad \text{and} \quad S_t(\omega) = U_t(\omega) \qquad \text{for almost all } t.$$

This implies $\rho(S.(\omega), U.(\omega)) = 0$. Since that is true for every ω, we have

$$\{\omega : S.(\omega) \in B\} = \{\omega : U.(\omega) \in B\} \qquad \text{for } B \in \mathscr{B}_\rho .$$

But the right side belongs to $\mathscr{F}$ as we mentioned above. Therefore we have

$$S.^{-1}(\mathscr{B}_\rho) \subset \mathscr{F} .$$

Since $\Omega_t \in \mathscr{F}$, $S_t(\omega)$ is measurable ($\mathscr{F}$) for each t fixed by the definition. This shows

$$\{\omega : S. \in B_{t,a}\} \in \mathscr{F} \qquad \text{for } B_{t,a} = \{f : f(t) < a\} .$$

As $\mathscr{B}_K$ is generated by $B_{t,a}$, $t \in T$, $a \in R$, we have

$$\{\omega : S. \in B\} \in \mathscr{F}, \qquad B \in \mathscr{B}_K$$

namely

$$S.^{-1}(\mathscr{B}_K) \subset \mathscr{F} .$$

Therefore

$$S.^{-1}(\mathscr{B}) = S.^{-1}(\mathscr{B}_\rho) \vee S.^{-1}(\mathscr{B}_K) \subset \mathscr{F} .$$

Thus we have proved that $\{S_t(\omega)\}$ is a canonical process. It is obvious by the construction that $P(S_t = X_t) = 1$ for each t. Therefore $\{S_t\}$ is a canonical modification of $\{X_t\}$.

Now we shall prove the uniqueness. Take any arbitrary canonical modification $\{\tilde{S}_t(\omega)\}$. $\{X_t^n(\omega)\}$ is measurable ($\mathscr{B}(T) \otimes \mathscr{F}$) in (t, ω) by the definition. $\{\tilde{S}_t(\omega)\}$ is measurable ($\mathscr{B}(T) \otimes \mathscr{F}$) in (t, ω) by Theorem (4.3). Therefore we can use Fubini's theorem to get

$$\begin{aligned} \boldsymbol{E}[\rho(S., X.^n)] &= \int_T \boldsymbol{E}[|\operatorname{atn} S_t - \operatorname{atn} X_t^n|]dt \\ &= \int_T \boldsymbol{E}[|\operatorname{atn} X_t - \operatorname{atn} X_t^n|]dt \\ &< 2^{-n} \to 0 \quad \text{as} \quad n \to \infty . \end{aligned}$$

Since $\rho(X.^n(\omega), U.(\omega)) \to 0$ a.e. on Ω, we can use the bounded convergence theorem to get

$$\boldsymbol{E}[\rho(X.^n, U.(\omega))] \to 0 \qquad \text{as } n \to \infty .$$

Therefore

$$\boldsymbol{E}[\rho(\tilde{S}., U.)] = 0 .$$

But $\rho(S., U.) = 0$ for every ω. Thus we get

$$\boldsymbol{E}[\rho(\tilde{S}., S.)] = 0$$

i.e. $P[S_t(\omega) = S_t(\omega)$ for almost every $t] = 1$. This completes the proof.

A canonical modification $\{S_t\}$ of $\{X_t\}$ is called standard if

$$S.^{-1}(\mathscr{B}) \subset \overline{\mathscr{B}[X_t\, t \in T]}^P ,$$

where $\mathscr{B}[X_t, t \in T]$ is the σ-algebra of subsets of Ω generated by the sets

$\{\omega : X_t < a\},\ a \in R,\ t \in T.$

STANDARD CANONICAL MODIFICATION THEOREM (4.8). *Every stochastic process $X_t(\omega)$ continuous in probability has a standard canonical modification. It is unique in the same sense as in Theorem* (4.7).

PROOF. Let $\mathcal{F}_1$ denote $\overline{\mathcal{B}[X_t, t \in T]}^P$. Then $\{X_t\}$ is considered a stochastic process continuous in probability on $(\Omega, \mathcal{F}_1, P)$. Therefore we have a canonical process $\{S_t(\omega)\}$ on $(\Omega, \mathcal{F}_1, P)$ such that $P(S_t = X_t) = 1$ for each t. $\{S_t(\omega)\}$ can be considered a canonical process on $(\Omega, \mathcal{F}, P)$. $S.^{-1}(\mathcal{B}) \subset \mathcal{F}_1 = \overline{\mathcal{B}[X_t, t \in T]}^P$ is now obvious, namely $\{S_t\}$ is a standard canonical modification of $\{X_t\}$. The uniqueness part is contained in theorem (4.7).

REMARK. Doob's definition of the standard property is

$$\{\omega : S_t \neq X_t\} \in \overline{\mathcal{B}[X_t, t \in T]}^P \qquad \text{for every } t.$$

This means

$$S.^{-1}(\mathcal{B}_R) \subset \overline{\mathcal{B}[X_t, t \in T]}^P$$

in our case. We required more than this.

§ 5. Probability measure on M.

Let $\{X_t(\omega), t \in T, \omega \in \Omega, \mathcal{F}, P)\}$ be a canonical process. Then the probability law of the sample path $X.(\omega)$ viewed as an $(M, \mathcal{B})$-valued random variable (see Remarks in Section 1), is a complete $\mathcal{B}$-regular probability measure on M satisfying

$$(\Gamma.1) \qquad \mu\{f : f(t) \in R\} = 1, \qquad t \in T.$$

Conversely, if we have such a probability measure μ on M, then $\{\xi_t(f) \equiv f(t), t \in T, f \in M, \mathcal{B}^\mu, \mu)\}$ is a canonical process for which the probability law of the sample path is the measure μ.

If $\{X_t\}$ is *continuous in probability*, then the probability law of the sample path satisfies

$$(\Gamma.2) \qquad \lim_{s \to t} \mu\{f : |f(s) - f(t)| > \varepsilon\} = 0, \qquad \varepsilon > 0,\ t \in T.$$

It is obvious that $\{\xi_t(f)\}$ is also continuous in probability under the condition $(\Gamma.2)$.

Let $m_{t_1 \cdots t_n}$ be the marginal distribution of μ over the time points $t_1 \cdots t_n$, namely

$$(5.1) \qquad m_{t_1 \cdots t_n}(E) = \mu\{f : (f(t_1), \cdots, f(t_n)) \in E\}, \qquad E \in \mathcal{B}[\bar{R}^n].$$

This is the joint distribution of the process $\{\xi_t\}$ (or $\{X_t\}$) over $(t_1 \cdots t_n)$. Let $\mathcal{M}$ denote the system of all marginal distributions. It is obvious that

$(m.0)$ $\mathcal{M}$ satisfies Kolmogorov's consistency condition.

$(\Gamma.1)$ and $(\Gamma.2)$ are formulated in terms of $\mathcal{M}$ as follows.

$(m.1)$ $$m_t(R)=1\,, \qquad t\in T\,,$$

$(m.2)$ $$\lim_{s\to t} m_{st}\{(x,y):|x-y|>\varepsilon\}=0\,, \qquad \varepsilon>0\,, \qquad t\in T\,.$$

Our problem is to determine μ from $\mathcal{M}$.

THEOREM (5.2). *For every $\mathcal{M}$ with $(m.0)$, $(m.1)$ and $(m.2)$, there exists one and only one complete $\mathcal{B}$-regular probability measure μ with* (5.1).

PROOF OF EXISTENCE. Let Ω be $\bar{R}^T$ and $\mathcal{B}(\bar{R}^T)$ the σ-algebra generated by the sets $\{\omega\in\bar{R}^T:\omega(t)<a\}$, $t\in T$, $a\in R$. Since $(m.0)$ is assumed, we can construct a complete regular $(\mathcal{B}(\bar{R}^T))$ probability measure P on Ω such that

$$P\{\omega:(\omega(t_1)\cdots\omega(t_n)\}\in E\}=m_{t_1\cdots t_n}(E)$$

by Kolmogorov's theorem. Then $\{X_t(\omega)\equiv\omega(t),\, t\in T,\, \omega\in(\Omega,\mathcal{F},P)\}$ $(\mathcal{F}=\overline{\mathcal{B}(\bar{R}^T)}^P)$, is a stochastic process continuous in probability by $(m.1)$ and $(m.2)$. By the canonical modification theorem (4.7) we have a canonical process $\{S_t(\omega)\}$ such that $P\{S_t=X_t\}=1$, $t\in T$.

Then the probability law $P_{S.}$ of the sample path $S.(\omega)$ is a complete $\mathcal{B}$-regular probability measure on M such that

$$\begin{aligned}
&P_{S.}\{f:(f(t_1),\cdots f(t_n))\in E\}\\
&=P\{\omega:(S(t_1,\omega)\cdots S(t_n,\omega))\in E\}\\
&=P\{\omega:(X(t_1,\omega)\cdots X(t_n,\omega))\in E\}\\
&=m_{t_1\cdots t_n}(E)\,.
\end{aligned}$$

PROOF OF UNIQUENESS. We shall consider the case $T=[0,1]$ only; the other cases can be treated similarly. Suppose that two complete $\mathcal{B}$-regular measures on M (say μ and ν) satisfy

(5.3) $$\begin{aligned}\mu\{f:(f(t_1),\cdots,f(t_n))\in E\}&=\nu\{f:(f(t_1),\cdots,f(t_n))\in E\}\\&=m_{t_1\cdots t_n}(E)\,.\end{aligned}$$

This is equivalent to

(5.3′) $$\int_M F(f(t_1),\cdots,f(t_n))\mu(df)=\int_M F(f(t_1),\cdots,f(t_n))\nu(df)$$

where F ranges over all continuous functions on $\bar{R}^n$.

Let us write $\hat{f}(I)$ for $\int_I \operatorname{atn} f(t)dt$. Then $\mathcal{B}_\rho$ is generated by $\{f:\hat{f}(I)<a\}$, $a\in R$, I being a compact interval, by virtue of theorem (3.11). $\mathcal{B}_K$ is generated by $\{f:f(t)<a\}$, $a\in R$, $t\in T$. Therefore $\mathcal{B}(=\mathcal{B}_\rho\vee\mathcal{B}_K)$ is generated by

all such sets. (5.3) or (5.3′) shows that μ and ν are equal on $\mathscr{B}_K$. To prove that they are equal on $\mathscr{B}$, it is enough to prove that

$$\int_M G[\hat{f}(I_1), \cdots, \hat{f}(I_p), \quad f(t_1), \cdots, f(t_q)]\mu(df) \tag{5.4}$$

$$= \int_M G[\hat{f}(I_1), \cdots, \hat{f}(I_p), \quad f(t_1), \cdots, f(t_q)]\nu(df)$$

where $p, q = 1, 2, \cdots$, $\{I_i\}$ range over all compact intervals, $\{t_i\}$ range over T and F ranges over all continuous functions on $\bar{R}^{p+q}$. The stochastic processes

$$X_t(f) = f(t), \qquad f \in (M, \overline{\mathscr{B}}^{\mu}, \mu)$$

$$Y_t(f) = f(t), \qquad f \in (M, \overline{\mathscr{B}}^{\nu}, \nu)$$

are canonical processes continuous in probability. Therefore we have $\delta(\varepsilon) > 0$ for every $\varepsilon > 0$ such that

$$|t-s| < \delta(\varepsilon) \tag{5.5}$$

$$\Rightarrow \int_M |\operatorname{atn} f(t) - \operatorname{atn} f(s)| \mu(df) < \varepsilon$$

$$\Rightarrow \int_M |\operatorname{atn} f(t) - \operatorname{atn} f(s)| \nu(df) < \varepsilon \qquad \text{by (5.3')}.$$

Fix a compact interval I, consider its decomposition into non-overlapping intervals

$$I = I_{n1} \cup I_{n2} \cup \cdots \cup I_{np(n)} \qquad |I_{ni}| < \delta(2^{-n})$$

for each n and let t_{ni} be the left and points of I_{ni}. Using the measurability of $X_t(f)$ in (t, f) (Theorem (4.3)), we can get

$$\int_M \sum_n \left| |f(I) - \sum_i \operatorname{atn} f(t_{ni}) |I_{ni}| \right| \mu(df) < \sum_n 2^{-n} |I| < \infty\,.$$

Therefore

$$\mu\{f : \hat{f}(I) = \lim_n \sum_i \operatorname{atn} f(t_{ni}) |I_{ni}|\} = 1\,.$$

Applying this to $I = I_1, I_2, \cdots, I_p$ we have

$$\mu\{f : \hat{f}(I_k) = \lim_n \sum_i \operatorname{atn} f(t^k{}_{ni}) |I^k{}_{ni}|\} = 1\,.$$

We have the same result for ν, i. e.

$$\nu\{f : \hat{f}(I_k) = \lim_n \sum_i \operatorname{atn} f(t^k{}_{ni}) |I^k{}_{ni}|\} = 1\,.$$

Notice here that we can take the same $t^k{}_{ni}$ and $I^k{}_{ni}$ for μ and ν by virtue of (5.5). By (5.3′) we have

$$\int_M G[\sum_i \operatorname{atn} f(t^k{}_{ni})|I^k{}_{ni}|,\ f(t_h),\ k=1, \cdots, p,\ h=1, \cdots, q]\mu(df)$$

$$=\int_M G[\sum_i \operatorname{atn} f(t^k{}_{ni})|I^k{}_{ni}|,\ f(t_h),\ k=1, \cdots, p,\ h=1, \cdots, q]\nu(df)$$

Letting $n\to\infty$, we have (5.4). This completes the proof.

Aarhus University, Denmark

Bibliography

[1] J. L. Doob, Stochastic Processes, New York, 1952.
[2] S. Saks, Theory of the Integral, 2nd rev. ed., English translation by L. C. Young, New York, 1933.