

# STOCHASTIC PROCESSES DEPENDING ON A CONTINUOUS PARAMETER\*

BY

J. L. DOOB

Stochastic processes depending on a continuous parameter have been defined in various ways. A definition frequently given is in terms of a physical system or other entity which depends on the parameter  $t$  (time) and whose state is specified by the position of a point  $Q = Q(t)$  varying in some space in accordance with a given probability law. The probability that  $Q$  will be in given point sets at given times is specified, or, if the process is a Markoff process, the probability that  $Q$  will be in a given point set at time  $t + \delta$  ( $\delta > 0$ ) if it was in a given position at time  $t$  is specified. Bachelier (I, II, III)† was the first to study these processes. His work was pioneering, and without any attempt at rigor. Most investigations have studied the particular case of Markoff processes‡ or else what are called below differential processes; that is, those in which the changes of  $Q(t)$  in non-overlapping intervals are independent in the usual probability sense.§

If  $Q$  represents the state of a physical system, the combination of the physical system and the probability relations is to be taken as the stochastic process. This, however, is not a mathematical definition, but a concretization of an unstated mathematical definition. Sometimes  $Q(T)$  is simply described as a function varying in accordance with a given law of probability.|| Khintchine (II) defined a stochastic process as a one-parameter family of chance variables. If  $Q$  varies on the  $x$ -axis, the probability relations of such a process are determined by specifying the probability of any set of inequalities of the form

---

\* Presented to the Society, September 1, 1936; received by the editors August 23, 1936.

† Roman numerals refer to the bibliography at the end of the paper. This bibliography does not pretend to completeness. It refers only to those papers on probability which are fairly closely related to this one.

‡ Cf. Hostinsky (I), Khintchine (III, pp. 24–59), Kolmogoroff (IV). Hostinsky has an extended bibliography, including references to papers on the diffusion problem which leads to a study of Markoff processes. Many papers have also been written studying Markoff processes in which  $Q = Q(t)$  can have only a finite number of positions. These have not studied the specific difficulties of definition, many of which do not arise in this special case, so specific references will not be given.

§ Bachelier (I, II, III); Khintchine (III, pp. 68–75); Kolmogoroff (I, II); Lévy (II, III, and several papers in the Paris Comptes Rendus whose results are given in II); Wiener (I, Chapters 9 and 10, and several earlier papers whose results are given in I).

|| Finetti (I, II).

$$(0.1) \quad x(t_j) < \lambda_j, \quad j = 1, \dots, n,$$

where  $\lambda_1, \dots, \lambda_n$  is any set of numbers and  $t_1, \dots, t_n$  any set of values of the (time) parameter.† This definition is not complete enough for the following reason. In many investigations it is necessary to examine the actual paths  $Q = Q(t)$  as to their continuity, etc.‡ The probability that the paths are continuous cannot be defined, however, in terms of the elementary probabilities of the inequalities (0.1).§ It is necessary, therefore, to define a stochastic process in a way which will make it possible to study such classes of paths as continuous paths. In this paper such a definition is proposed, and proofs are given for some new theorems and for some old ones whose earlier proofs (and formulations) have presupposed such a definition. Proofs which would merely be rephrasings of earlier ones have been omitted. Only the case where  $Q$  varies on the  $x$ -axis and where the parameter  $t$  varies from  $-\infty$  to  $+\infty$  is considered. The method in more general cases would be the same.

The point of view is that in which probability becomes (mathematically) a study of Lebesgue and more general measures on suitable spaces. This point of view was developed fully from an abstract standpoint by Kolmogoroff (V).|| It was applied (not in all cases in work based on Kolmogoroff's) by Doob (I, II, III), Hopf (I), Kac and Steinhaus (II, III), Khintchine (I), Lévy (IV, esp. pp. 84–88), Łomnicki and Ulam (II) to discrete stochastic processes (those in which the parameter  $t$  runs through integral values only) by choosing a suitable space of infinitely many dimensions. Essentially the same space had been used earlier by Steinhaus (II) and Wiener (I), and has been used recently by Jessen (I).

1. Definitions. Let  $\Omega^*$  be the space of all functions  $x(t)$  defined over  $-\infty < t < +\infty$ . If  $\{t_1, \dots, t_n\}$  is any finite set of  $t$ -values, and if  $-\infty \leq a_j < b_j \leq +\infty, j = 1, \dots, n$ , the set of elements of  $\Omega^*$  satisfying

$$(1.1) \quad a_j < x(t_j) < b_j, \quad j = 1, \dots, n,$$

---

† Since a chance variable, mathematically, is merely a measurable function  $\phi(\omega)$  defined on some sort of space  $\Omega$  on which a measure is defined, this definition presumably means that we have given a one-parameter family of measurable functions  $\phi_t(\omega)$  all defined on the same space  $\Omega$ , and that the  $\Omega$ -measure of the  $\omega$ -set, where the inequalities

$$\phi_{t_j}(\omega) < \lambda_j, \quad j = 1, \dots, n$$

are satisfied, is specified. Cf. Doob (I).

‡ Kolmogoroff (II, p. 868); Lévy (II, III); Wiener (I).

§ Cf. Khintchine (III, pp. 68–69). If the space in which  $Q$  varies is the real axis, and if  $I$  is an interval on the  $t$ -axis, not even the probability that  $\text{L.U.B.}_{t \in I} x(t) \leq M$  can be defined, for any number  $M$ . Cf. Kolmogoroff (V, p. 26).

|| Cf. also earlier papers of Cantelli (I), Łomnicki (I), Steinhaus (I).

will be called a neighborhood. If  $x_0(t)$  is an element of such a neighborhood, the neighborhood will be called a neighborhood of  $x_0(t)$ . The topological characteristics of the space  $\Omega^*$ , induced by this definition of neighborhood, have been investigated by Tychonoff. † Kolmogoroff (V, pp. 24–30) has shown that a measure function  $P^*(\Lambda^*)$  which is non-negative and completely additive, and satisfies  $P^*(\Omega^*) = 1$ , can be defined on the sets of the Borel system of sets determined by the neighborhoods. ‡ The set function is determined by its values on the neighborhoods (1.1) and the only requirement is that, for  $t_1, \dots, t_n$  fixed,  $P^*(\Lambda^*)$  should become a non-negative, completely additive function defined on the sets of the Borel system over the neighborhoods (1.1), taking on the value 1 for the whole space. § The sets on which  $P^*$  is defined will be called  $P^*$ -measurable. If  $M^*$  is any subset of  $\Omega^*$ , its exterior  $P^*$ -measure,  $\bar{P}^*(M^*)$ , is defined as the greatest lower bound of  $P^*(\Lambda^*)$  for all  $P^*$ -measurable sets  $\Lambda^*$  containing  $M^*$ . It is readily seen that  $\bar{P}^*(M^*)$  is the greatest lower bound of  $P^*(\Lambda^*)$  for all sets  $\Lambda^* \supset M^*$  which are finite or denumerable sums of neighborhoods.

**THEOREM 1.1.** *Let  $\Omega$  be a subset of  $\Omega^*$ . A necessary and sufficient condition that  $\Lambda_1^* \Omega = \Lambda_2^* \Omega$  imply that  $P^*(\Lambda_1^*) = P^*(\Lambda_2^*)$  for every pair of  $P^*$ -measurable sets  $\Lambda_1^*, \Lambda_2^*$  of  $\Omega^*$  is that  $\bar{P}^*(\Omega) = 1$ .*

The condition is sufficient. For if  $\bar{P}^*(\Omega) = 1$ , and if  $\Lambda_1^* \Omega = \Lambda_2^* \Omega$ ,

$$(1.2) \quad (\Lambda_1^* - \Lambda_1^* \cdot \Lambda_2^*) \Omega = 0.$$

Then

$$(1.3) \quad \Omega^* \supset (\Lambda_1^* - \Lambda_1^* \cdot \Lambda_2^*) + \Omega$$

implies

$$(1.4) \quad 1 \geq P^*(\Lambda_1^* - \Lambda_1^* \cdot \Lambda_2^*) + \bar{P}^*(\Omega) = P^*(\Lambda_1^*) - P^*(\Lambda_1^* \cdot \Lambda_2^*) + 1,$$

or

$$(1.5) \quad P^*(\Lambda_1^*) = P^*(\Lambda_1^* \cdot \Lambda_2^*).$$

Interchanging  $\Lambda_1^*, \Lambda_2^*$ ,

$$(1.5') \quad P^*(\Lambda_2^*) = P^*(\Lambda_1^* \cdot \Lambda_2^*),$$

† Mathematische Annalen, vol. 111 (1935), pp. 762–766.

‡ The Borel system of sets determined by any collection of sets is the smallest collection of sets which includes the given collection and which contains the sets  $\sum_1^\infty A_n, \prod_1^\infty A_n$  when it contains the sets  $\{A_n\}$ . We shall suppose also that  $P^*(\Lambda^*)$  is defined for any set differing from a set  $\Lambda^*$  in the Borel system by a subset of a set in the system for which  $P^*$  vanishes:  $P^*(\Lambda_1^*) = P^*(\Lambda^*)$ .

§ This set function is merely a measure in  $n$ -dimensional Euclidean space with coordinates  $x(t_1), \dots, x(t_n)$ .

so that  $P^*(\Lambda_1^*) = P^*(\Lambda_2^*)$ .

The condition is necessary. For if  $\Lambda_1^* \cdot \Omega = \Lambda_2^* \cdot \Omega$  implies that  $P^*(\Lambda_1^*) = P^*(\Lambda_2^*)$ , let  $\Lambda_1^*$  be a  $P^*$ -measurable set containing  $\Omega$ . It is sufficient to show that then  $P^*(\Lambda_1^*) = 1$  necessarily. If  $\Lambda_2^* = \Omega^*$ ,

$$(1.6) \quad \Lambda_1^* \cdot \Omega^* = \Lambda_2^* \cdot \Omega = \Omega,$$

so that  $P^*(\Lambda_1^*) = P^*(\Lambda_2^*) = 1$ , as was to be proved.

This theorem makes possible the following definition.

**DEFINITION.** Let  $\Omega$  be a subset of  $\Omega^*$  of exterior  $P^*$ -measure 1. If  $\Lambda$  is a subset of  $\Omega$  of the form  $\Lambda = \Lambda^* \cdot \Omega$ , where  $\Lambda^*$  is  $P^*$ -measurable, a set function  $P(\Lambda)$  is defined by  $P(\Lambda) = P^*(\Lambda^*)$ . The topological space  $\Omega$  together with the measure function thus defined, will be called a stochastic process.†

The sets  $\Lambda$  on which  $P(\Lambda)$  is defined will be called  $P$ -measurable. Explicit examples of stochastic processes are given below.

**DEFINITION.**  $\Omega$  will be called quasi-separable if for any open interval  $I$ , and for any number  $k$ , the set of elements in  $\Omega$  at which

$$\text{L.U.B.}_{t \in I} x(t) \leq k \quad (\text{G.L.B.}_{t \in I} x(t) \geq k),$$

is the same as the set of elements at which

$$\text{L.U.B.}_{n \geq 1} x(t_n) \leq k \quad (\text{G.L.B.}_{n \geq 1} x(t_n) \geq k),$$

for some sequence  $\{t_n\}$  of points in  $I$ .

The conditions require that certain non-denumerable products of closed sets  $\prod_{t \in I} \{x(t) \leq k\}$ † or  $\prod_{t \in I} \{x(t) \geq k\}$  be equivalent to denumerable partial products, and thus imply less than separability. A stochastic process will be called quasi-separable whenever the space  $\Omega$  of the process is quasi-separable.

**THEOREM 1.2.** Let  $x^*(t)$ ,  $x_*(t)$  be the upper and lower limiting functions of  $x(t)$ ,§ respectively. Then the process is quasi-separable if and only if there is an everywhere dense sequence of points  $\{r_n\}$  such that if  $y^*(t)$  ( $y_*(t)$ ) is the upper (lower) limiting function of  $x(r_i)$ ,||

$$(1.7) \quad x^*(t) \equiv y^*(t), \quad x_*(t) \equiv y_*(t).$$

† Strictly speaking, the stochastic process should be defined as the physical system or other entity whose changing is represented by the mathematical formulation of the definition, but it seems wiser to use the term stochastic process both for the mathematical formulation and for the concretization it represents, than to introduce more terminology.

‡ The set of elements satisfying conditions  $C$  will be denoted by  $\{C\}$ .

§  $x^*(t) = \lim_{n \rightarrow \infty} \text{L.U.B.}_{|\tau - t| < 1/n} x(\tau)$ ,  $x_*(t) = \lim_{n \rightarrow \infty} \text{G.L.B.}_{|\tau - t| < 1/n} x(\tau)$ .

||  $y^*(t) = \lim_{n \rightarrow \infty} \text{L.U.B.}_{|r_j - t| < 1/n} x(r_j)$ ,  $y_*(t) = \lim_{n \rightarrow \infty} \text{G.L.B.}_{|r_j - t| < 1/n} x(r_j)$ .

The condition is certainly sufficient for quasi-separability. The sequence  $\{t_n\}$  of the definition can be taken as the subsequence of  $\{r_n\}$  consisting of the points of  $\{r_n\}$  in  $I$ . Conversely suppose that  $\Omega$  is quasi-separable.

According to the quasi-separability hypothesis, to each open interval  $I$ , and to each number  $k$  correspond two sequences of points in  $I$ . Let  $\{r_n\}$  be the set of all such points for all rational values of  $k$  and for all open intervals  $I$  with rational endpoints. This sequence  $\{r_n\}$  is evidently everywhere dense and has the property that the least upper bound of  $x(t)$  (an element of  $\Omega$ ) in any interval  $I$  with rational endpoints is the same as that of  $x(t)$  for  $t$  running through the values of  $\{r_n\}$  in  $I$ . This statement is still true even if  $I$  is not restricted to have rational endpoints, since any interval can be expressed as the sum of such intervals. The first identity (1.7) then follows at once from the definitions of the functions concerned, and the second identity is proved in the same way.

THEOREM 1.3. *If for some real number  $\lambda$ ,  $\lambda_n \rightarrow \lambda$  implies*

$$(1.8) \quad P\left\{\lim_{n \rightarrow \infty} x(\lambda_n) = x(\lambda)\right\} = 1,$$

*then if  $\{s_n\}$  is a denumerable set with  $\lambda$  as a limit point,*

$$(1.9) \quad P\left\{\lim_{s_n \rightarrow \lambda} x(s_n) = x(\lambda)\right\} = 1.$$

It is sufficient to prove that the upper and lower limiting functions of  $x(s_n)$  at  $\lambda$  are both  $x(\lambda)$ , on an  $\Omega$ -set of  $P$ -measure 1. The proof will be given for the upper limiting function  $x_s^*(s)$ . Unless

$$(1.10) \quad x_s^*(\lambda) \leq x(\lambda)$$

with probability 1—i.e., on an  $\Omega$ -set of  $P$ -measure 1, there is a number  $r$ , a positive number  $\epsilon$ , and a  $P$ -measurable set  $\Lambda$  of positive measure, such that

$$(1.11) \quad x(\lambda) < r < r + \epsilon < \text{L.U.B.}_{|s_j - \lambda| < 1/n} x(s_j), \quad n = 1, 2, \dots,$$

on  $\Lambda$ . Let  $u_1^{(n)}, u_2^{(n)}, \dots$ , be the points of  $\{s_j\}$  in  $|t - \lambda| < 1/n$ , arranged in any order. Then on  $\Lambda$ ,

$$(1.12) \quad x(\lambda) < r < r + \epsilon < \text{L.U.B.}_{j \geq 1} x(u_j^{(n)}), \quad n = 1, 2, \dots$$

Now

$$(1.13) \quad \text{L.U.B.}_{j \geq 1} x(u_j^{(n)}) = \lim_{\nu \rightarrow \infty} \max [x(u_1^{(n)}), \dots, x(u_\nu^{(n)})],$$

so there is an integer  $\nu_n$  so large that

$$(1.14) \quad \max_{j \leq n} x(u_j^{(n)}) > \text{L.U.B. } x(u_j^{(n)}) - \epsilon > r$$

on a subset  $\Lambda_n$  of  $\Lambda$  of  $P$ -measure at least  $P(\Lambda)(1-2^{-n-1})$ . If  $u_1^{(1)}, \dots, u_{\nu_1}^{(1)}, u_1^{(2)}, \dots, u_{\nu_2}^{(2)}, \dots$  are renamed  $\lambda_1, \lambda_2, \dots$  respectively,  $\lambda_n \rightarrow \lambda$  and

$$(1.15) \quad \limsup_{n \rightarrow \infty} x(\lambda_n) \geq r > x(\lambda)$$

on  $\Lambda_0 = \bigcap_1^\infty \Lambda_n$ , where  $P(\Lambda_0) \geq \frac{1}{2}P(\Lambda) > 0$ , contradicting the hypothesis of the theorem. Then  $x_s^*(\lambda) \leq x(\lambda)$  with probability 1, i.e., on a set of  $P$ -measure 1. The inequality  $x_s^*(\lambda) \geq x(\lambda)$  with probability 1 follows at once from the hypotheses, so  $x_s^*(\lambda) = x(\lambda)$  with probability 1, as was to be proved.

**COROLLARY.** *If for some real number  $\lambda$ ,  $\lambda_n \rightarrow \lambda$  implies (1.8) on a quasi-separable process,  $x(t)$  is continuous at  $\lambda$ , with probability 1.*

Since the process is quasi-separable, there is an everywhere dense sequence  $\{r_n\}$  such that the upper and lower limiting functions of  $x(t)$  at  $\lambda$  are the same as those of  $x(t)$  when we consider  $x(t)$  defined only at the points  $\{r_n\}$ . By the theorem the latter limiting functions are both  $x(\lambda)$ , with probability 1.

**2. Measurability of stochastic processes.** In the applications of the theory of stochastic processes,  $P^*$ -measure is usually prescribed on  $\Omega^*$ , and the first problem becomes that of finding the subspace  $\Omega$ ,  $\bar{P}^*(\Omega) = 1$ , whose elements are as simple as possible, e.g., measurable, monotone, etc. As an introductory example we prove the following theorem.

**THEOREM 2.1.** *Let  $P^*$ -measure be so defined that there is a number  $M$  with the property that for each value of  $t$*

$$(2.1) \quad P^*\{ |x(t)| \leq M \} = 1.$$

*Then if  $\Omega$  is the subspace of  $\Omega^*$  consisting of those elements of  $\Omega^*$  for which  $|x(t)| \leq M$ ,  $-\infty < t < +\infty$ ,  $\bar{P}^*(\Omega) = 1$ , and  $\Omega$  is the space of a stochastic process.*

Let  $\Gamma^* = \sum_1^\infty \Gamma_n^*$  be a sum of neighborhoods, and suppose that  $\Gamma^* \supset \Omega$ . If  $t_1, t_2, \dots$  are the  $t$ -values used in defining  $\Gamma_1^*, \Gamma_2^*, \dots$ , and if  $\Lambda^*$  is the  $P^*$ -measurable set determined by  $|x(t_j)| \leq M, j \geq 1$ ,

$$(2.2) \quad P^*(\Lambda^*) = 1.$$

If  $x_0(t) \in \Lambda^*$ , there is an element  $x_1(t)$  in  $\Omega$  such that  $x_1(t_j) = x_0(t_j), j \geq 1$ . For some  $\nu$ ,  $x_1(t) \in \Gamma_\nu^*$ . Since  $\Gamma_\nu^*$  is defined by inequalities on  $x(t)$  at certain of the numbers  $t_1, t_2, \dots$ ,  $x_1(t) \in \Gamma_\nu^*$  implies that  $x_0(t) \in \Gamma_\nu^*$ . Therefore

$$(2.3) \quad \Lambda^* \subset \Gamma^*,$$

so  $P^*(\Gamma^*) = 1$ . Thus any sum of neighborhoods which includes  $\Omega$  has  $P^*$ -measure 1, implying that  $\bar{P}^*(\Omega) = 1$ .

The space of all pairs  $\{\tau, x(t)\}$ , the direct product of the  $\tau$ -axis with  $\Omega$ , will be denoted by  $T \times \Omega$ . If measure is defined on the  $\tau$ -axis as Lebesgue measure, measure can be defined on the product space in the usual way, being determined by the fact that the measure of the direct product of a  $\tau$ -measurable set and a  $P$ -measurable set is the product of their measures.<sup>†</sup> A set in  $T \times \Omega$  is measurable (i.e., its measure is defined) if and only if it is the intersection of  $T \times \Omega$  with a set measurable in  $T \times \Omega^*$ .

**DEFINITION.** A stochastic process will be called measurable if the  $T \times \Omega$  set of elements  $\{\tau, x(t)\}$  for which  $x(\tau) > k$  is measurable in  $T \times \Omega$  for every real number  $k$ , i.e., if  $x(\tau)$  is a measurable function on  $T \times \Omega$ .

This implies, by Fubini's theorem,<sup>‡</sup> that the element  $x(t)$  of  $\Omega$  is a Lebesgue measurable function, with probability 1.

**THEOREM 2.2.** The space  $\Omega^*$  is never the space of a measurable stochastic process.

If  $\Omega^*$  were the space of some measurable stochastic process,  $x(t)$  would be Lebesgue measurable, with probability 1, in terms of  $P^*$ -measure. It is therefore sufficient to show that if  $\Lambda^*$  is the set of elements of  $\Omega^*$  which are not Lebesgue measurable,  $\bar{P}^*(\Lambda^*) = 1$ , however  $P^*$ -measure is defined. Let  $\Gamma^* = \sum_1^\infty \Gamma_n^*$  be a sum of neighborhoods covering  $\Lambda^*$ . Then we must show that  $P^*(\Gamma^*) = 1$ . Let  $t_1, t_2, \dots$  be the set of  $t$ -values involved in defining  $\Gamma^*$ . If  $x_0(t)$  is any element of  $\Omega^*$ , there is a non-measurable function  $x_1(t)$ , an element of  $\Lambda^*$ , such that  $x_1(t_j) = x_0(t_j)$ ,  $j = 1, 2, \dots$ . We need only take any non-measurable function and redefine it, if necessary, at  $t = t_1, t_2, \dots$ . Since  $\Gamma^* \supset \Lambda^*$  contains the element  $x_1(t)$ , it must also contain  $x_0(t)$ . Therefore  $\Gamma^* = \Omega^*$ :  $P^*(\Gamma^*) = 1$ , as was to be proved.

We shall need the following lemma on measurable functions.

**LEMMA 2.1.** Let  $f(\tau, v)$  be a Lebesgue measurable function defined in the strip  $0 < v < 1$  of the  $(\tau, v)$ -plane. Define  $\phi_n(t)$  by:

$$(2.4) \quad \phi_n(t) = k2^{-n} \quad \text{if} \quad (k-1)2^{-n} < t \leq k2^{-n}.$$

Then there is a number  $c$  and a sequence of integers  $\{a_n\}$  such that  $f[\phi_n(t) + c, v]$  is a measurable function in the strip  $0 < v < 1$  of the  $(t, v)$ -plane and that

<sup>†</sup> A treatment of measure functions in product spaces can be found in S. Saks, *Théorie de l'Intégrale*, Warsaw, 1933, pp. 257-263.

<sup>‡</sup> The importance of Fubini's theorem in work of this kind was stressed by P. Lévy (V). Cf. also the investigations of J. von Neumann on flows, *Annals of Mathematics*, vol. 33 (1932), pp. 587-642; esp. pp. 588-589.

$$(2.5) \quad \lim_{n \rightarrow \infty} f[\phi_{a_n}(t) + c, v] = f(t + c, v)$$

almost everywhere in the strip.

(i) Consider the function  $f[\phi_n(t) + s, v]$ , defined in the 3-dimensional  $(s, t, v)$ -region  $0 < v < 1$ . We show first that this function is Lebesgue measurable in the  $(s, t, v)$ -region. It is sufficient to show that if  $E_{u,v}$  is a measurable plane point set on the  $(u, v)$ -plane, and if

$$[\phi_n(t) + s, v] \in E_{u,v}, \quad (u = \phi_n(t) + s)$$

determines the  $(s, t, v)$ -set  $E_{s,t,v}$ , then  $E_{s,t,v}$  is measurable. If  $E_{u,v}$  is a rectangle,  $E_{s,t,v}$  is certainly measurable, as the direct product of an  $(s, t)$  Borel set with a  $v$ -interval. Then if  $E_{u,v}$  is any Borel set  $E_{s,t,v}$  is measurable. Suppose that  $E_{u,v}$  is a Borel set of measure 0. It has just been shown that  $E_{s,t,v}$  is measurable, and we shall show the additional fact that it is of zero measure. The intersection of  $E_{u,v}$  with a line  $v = \text{const.}$  is almost always of measure 0, and it is sufficient to show that the intersection of  $E_{s,t,v}$  with a line  $v = \text{const.}$  is almost always of measure 0. To show this it is sufficient to show that the  $(s, t)$ -set  $E_{s,t}$  determined by  $\phi_n(t) + s \in E_u$  is of measure 0 if it is measurable and if  $E_u$  is of measure 0. This is certainly true since the intersection of  $E_{s,t}$  with a line  $t = \text{const.}$  is congruent to  $E_u$  and so has the same measure. Thus if  $E_{u,v}$  is a Borel set,  $E_{s,t,v}$  is measurable, and has measure 0 if  $E_{u,v}$  has measure 0. If  $E_{u,v}$  is any measurable set of zero measure, it can be enclosed in a Borel set of zero measure, so in this case  $E_{s,t,v}$  is still measurable. Since any measurable set  $E_{u,v}$  differs from some Borel set by at most a set of measure 0, if  $E_{u,v}$  is measurable,  $E_{s,t,v}$  is also measurable, as was to be proved.

(ii) For  $(t, v)$  fixed,  $f[\phi_n(t) + s, v]$  is a measurable function of  $s$  (except possibly for a  $(t, v)$ -set of measure 0) so by a theorem of Auerbach† if  $\{\alpha_n\}$  is any sequence of integers, there is a subsequence  $\{\beta_n\}$  with the property that

$$(2.6) \quad \lim_{n \rightarrow \infty} f[\phi_{\beta_n}(t) + s, v] = f(t + s, v)$$

for almost all  $s$ . Let  $I_s$  be any finite  $s$ -interval,  $\epsilon$  any positive number, and let  $\xi_n(t, v)$  be the  $s$ -measure of the subset of  $I_s$  where

---

† Fundamenta Mathematicae, vol. XI (1928), pp. 196–197. Auerbach's theorem states that if  $F(x)$  is a Lebesgue measurable function of  $x$  and if  $\{\eta_n\}$  is a sequence of numbers converging to 0, there is a subsequence  $\{\eta_{a_n}\}$  with the property that  $F(x + \eta_{a_n}) \rightarrow F(x)$  almost everywhere on the  $x$ -axis. This can be restated in terms of convergence in measure:  $F(x + h)$  converges in measure to  $F(x)$  on any finite interval as  $h \rightarrow 0$ . (Convergence in measure was defined and discussed by F. Riesz, Paris Comptes Rendus, vol. 148 (1909), pp. 1303–1309.)



$$(2.7) \quad |f[\phi_n(t) + s, v] - f(t + s, v)| \geq \epsilon.$$

Then  $\xi_n(t, v)$  is defined almost everywhere in the  $(t, v)$ -strip  $0 < v < 1$ , is uniformly bounded by the length of  $I_s$ , and converges to 0 with  $1/n$ .† If  $R_{t,v}$  is the  $(t, v)$ -rectangle  $[t_0 < t < t_1, 0 < v < 1]$ , the integral

$$\iint_{R_{t,v}} \xi_n(t, v) dt dv$$

which represents the  $(s, t, v)$ -measure of the  $(s, t, v)$ -set where (2.7) is true, restricting  $(s, t)$  to  $[s \in I_s, t_0 < t < t_1]$ , converges to 0 with  $1/n$ . This implies that  $f[\phi_n(t) + s, v]$  converges in measure to  $f(t + s, v)$  on the  $(s, t, v)$ -set  $I_s \times R_{t,v}$  determined by  $s \in I_s, (t, v) \in R_{t,v}$ . There is then a sequence of integers  $\{\alpha_n\}$  such that  $f[\phi_{\alpha_n}(t) + s, v]$  converges almost everywhere on  $I_s \times R_{t,v}$  to  $f(t + s, v)$ . By a familiar use of the diagonal process, there is a subsequence  $\{a_n\}$  of  $\{\alpha_n\}$  such that

$$\lim_{n \rightarrow \infty} f[\phi_{a_n}(t) + s, v] = f(t + s, v)$$

almost everywhere on the  $(s, t, v)$ -set determined by  $0 < v < 1$ . Then except for an  $s$ -set of zero measure  $f[\phi_{a_n}(t) + s, v]$  converges ( $s$  fixed) to  $f(t + s, v)$  almost everywhere on the  $(t, v)$ -strip  $0 < v < 1$ . The number  $c$  of the theorem can be taken as any number not on this exceptional  $s$ -set.

**THEOREM 2.3.** *If  $P^*$ -measure is so defined that a constant  $c$  and a sequence of integers  $\{a_n\}$  exist for which*

$$(2.8) \quad P^* \left\{ \lim_{n \rightarrow \infty} x[c + \phi_{a_n}(t)] = x(c + t) \right\} = 1 \ddagger$$

*for almost all  $t$ , there is a space  $\Omega \subset \Omega^*$  which is the space of a measurable stochastic process.*

*Conversely, if there is a measurable stochastic process with space  $\Omega \subset \Omega^*$ , the corresponding  $P^*$ -measure must satisfy the above condition.*

Suppose that the given  $P^*$ -measure satisfies the condition of the theorem. It will simplify the notation without restricting generality, if we assume that  $c = 0$ . If  $x(t) \in \Omega^*$ , we define  $x^{(i)}(t), x^{(s)}(t)$  by

$$(2.9) \quad \begin{aligned} x^{(i)}(t) &= \liminf_{n \rightarrow \infty} x[\phi_{a_n}(t)], \\ x^{(s)}(t) &= \limsup_{n \rightarrow \infty} x[\phi_{a_n}(t)]. \end{aligned}$$

† Cf. the preceding note.

‡ The function  $\phi_n(t)$  is defined by (2.4).

The functions  $x(t)$ ,  $x[\phi_n(t)]$ ,  $x^{(i)}(t)$ ,  $x^{(s)}(t)$  determine functions  $x(\tau)$ ,  $x[\phi_n(\tau)]$ ,  $x^{(i)}(\tau)$ ,  $x^{(s)}(\tau)$  on  $T \times \Omega^*$ . The function  $x[\phi_{\alpha_n}(\tau)]$  is readily seen to be measurable on  $T \times \Omega^*$ ,  $n = 1, 2, \dots$ , so the functions  $x^{(i)}(\tau)$ ,  $x^{(s)}(\tau)$  are also measurable on  $T \times \Omega^*$ . If  $t$  is not in an exceptional set  $S$  of measure 0,

$$(2.10) \quad P^* \{ x^{(i)}(t) = x(t) = x^{(s)}(t) \} = 1.$$

Let  $\Omega$  be the set of all elements  $x(t) \in \Omega^*$  satisfying

$$(2.11) \quad x^{(i)}(t) \leq x(t) \leq x^{(s)}(t)$$

for all values of  $t$  not in  $S$  for which  $x^{(i)}(t) < +\infty$ ,  $x^{(s)}(t) > -\infty$ , and otherwise unrestricted. We shall show that  $\bar{P}^*(\Omega) = 1$  by showing that if  $\Gamma^* = \sum_1^\infty \Gamma_n^*$  is a sum of neighborhoods covering  $\Omega$ ,  $P^*(\Gamma^*) = 1$  necessarily. Let  $t_1, t_2, \dots$  be the values of  $t$  used in defining  $\Gamma^*$ , of which  $t_{\alpha_1}, t_{\alpha_2}, \dots$  are not in  $S$ ,† and let  $\Lambda^*$  be the  $\Omega^*$ -set determined by

$$(2.12) \quad x^{(i)}(t_{\alpha_j}) = x(t_{\alpha_j}) = x^{(s)}(t_{\alpha_j}), \quad j = 1, 2, \dots$$

By (2.10),  $P^*(\Lambda^*) = 1$ . If  $x_0(t)$  is an element of  $\Lambda^*$ , let  $x_1(t)$  be an element of  $\Omega^*$  equal to  $x_0(t)$  at  $t_1, t_2, \dots$  and at  $t = k2^{-n}$ , for  $k = 0, \pm 1, \dots$ ,  $n = 1, 2, \dots$ , and otherwise restricted only by

$$(2.13) \quad x_1^{(i)}(t) = x_0^{(i)}(t) \leq x_1(t) \leq x_0^{(s)}(t) = x_1^{(s)}(t)$$

at the values of  $t \notin S$  where  $x_0^{(i)}(t) < +\infty$ ,  $x_0^{(s)}(t) > -\infty$ . Then  $x_1(t) \in \Omega$ : for  $t = k2^{-n}$ ,  $x_1^{(i)}(t) = x_1^{(s)}(t) = x_1(t)$  by the definition of the functions  $x^{(i)}(t)$ ,  $x^{(s)}(t)$  so that (2.11) is satisfied; at  $t = t_{\alpha_1}, t_{\alpha_2}, \dots$ , (2.12) implies (2.11); and for all other values of  $t$ , (2.13) implies (2.11) (the qualifying conditions on the two are identical). Since  $x_1(t) \in \Gamma^*$  there is an integer  $\nu$  such that  $x_1(t) \in \Gamma_\nu^*$  and since  $x_1(t_j) = x_0(t_j)$ ,  $j = 1, 2, \dots$ , this means that  $x_0(t) \in \Gamma_\nu^* : \Lambda^* \subset \Gamma^*$ ,  $P^*(\Gamma^*) = 1$ , as was to be proved.

The space  $\Omega$  is thus of exterior  $P^*$ -measure 1 and so the space of a stochastic process, which we shall now show is measurable. The functions  $x^{(i)}(\tau)$ ,  $x^{(s)}(\tau)$  are measurable on  $T \times \Omega^*$  and therefore on  $T \times \Omega$ . For fixed  $\tau = t$  they are equal, with probability 1, by (2.10), except for a  $t$ -set of measure 0, so they are equal to each other almost everywhere on  $T \times \Omega$ . Since for fixed  $\tau = t \notin S$  they are equal to  $x(t)$  with probability 1, they are finite valued almost everywhere on  $T \times \Omega$ . The function  $x(\tau)$  on  $T \times \Omega$  then satisfies

$$(2.11') \quad x^{(i)}(\tau) \leq x(\tau) \leq x^{(s)}(\tau)$$

almost everywhere on  $T \times \Omega$  and is therefore equal to  $x^{(i)}(\tau) = x^{(s)}(\tau)$  almost everywhere on  $T \times \Omega$ , proving the measurability of the process.

† The set  $\{t_{\alpha_n}\}$  may be empty.

Conversely, if  $P^*$ -measure is so defined that there is a measurable stochastic process with space  $\Omega \subset \Omega^*$ , the function  $x(\tau)$  is a measurable function on  $T \times \Omega$ . We apply Lemma 2.1 to this function  $x(\tau)$  defined on  $\{\tau, x(t)\}$  space  $T \times \Omega$ , finding that there is a number  $c$  and a sequence of integers  $\{a_n\}$  such that

$$P\left\{\lim_{n \rightarrow \infty} x[c + \phi_{a_n}(t)] = x(c + t)\right\} = 1,$$

equivalent to (2.8) because of the definition of  $P$ -measure.<sup>†</sup>

The condition of the theorem is in unwieldy form, but has the advantage of being both necessary and sufficient for the existence of a measurable stochastic process. The following theorem gives a necessary condition and a sufficient condition in more usable form.

**THEOREM 2.4.** (i) *Let  $P^*$ -measure be so defined that there is a space  $\Omega \subset \Omega^*$  which is the space of a measurable stochastic process. Then if  $\epsilon > 0$*

$$(2.14) \quad P^*\{ |x(t+h) - x(t)| > \epsilon \}$$

*is a measurable function of  $t$  ( $h$  fixed) which converges in measure to 0 on every finite  $t$ -interval, as  $h \rightarrow 0$ .*

(ii) *If  $P^*$ -measure is so defined that for every positive  $\epsilon$  (2.14) converges to 0 for almost all values of  $t$ , as  $h \rightarrow 0$ , there is a subspace  $\Omega \subset \Omega^*$  which is the space of a measurable stochastic process.*

If  $P^*$ -measure is so defined that the  $P^*$ -measure of a neighborhood (and so of any  $P^*$ -measurable set) is independent of translations of the  $t$ -axis, the conditions (i) and (ii) become identical and the resulting condition is then both necessary and sufficient.

We first prove (i). By hypothesis  $x(\tau)$  is a measurable function of  $\{\tau, x(t)\}$ . Then  $|x(\tau+h) - x(\tau)|$  is measurable, so by Fubini's theorem

$$(2.14') \quad P\{ |x(t+h) - x(t)| > \epsilon \}$$

is a measurable function of  $t$ . The function of (2.14) is equal to this function for all  $t, h$ , and is therefore also measurable in  $t$ . The element  $x(t)$  of  $\Omega$  is measurable, with probability 1. Applying the theorem of Auerbach as in the proof of Lemma 2.1, if  $I$  is any finite  $t$ -interval,  $x(t+h)$  converges in measure on  $I$  to  $x(t)$  as  $h \rightarrow 0$ ; i.e., if  $\xi[x(t), h]$  is the Lebesgue measure of the  $t$ -subset of  $I$  where  $|x(t+h) - x(t)| > \epsilon$ ,  $\xi$  is bounded by the length of  $I$  and converges to 0 with  $h$ . The measure of the  $T \times \Omega$ -set (restricting  $\tau$  to  $I$ ), where  $|x(\tau+h) - x(\tau)| > \epsilon$ , is

<sup>†</sup> The lemma was stated, for convenience, for a function  $f(\tau, v)$  defined on  $(\tau, v)$ -space, but the proof holds for this case also, with no essential change.

$$(2.15) \quad \int \xi[x(t), h] dP, \dagger$$

which converges to 0 with  $h$ . The measure of this set can also be expressed as

$$(2.16) \quad \int_I P\{|x(t+h) - x(t)| > \epsilon\} dt,$$

and the convergence of this integral to 0 with  $h$ , remembering the identity of (2.14) and (2.14'), is equivalent to the statement of the theorem.

Next we prove (ii). We shall show that  $P^*$ -measure satisfies the condition of Theorem 2.3 with  $c=0$  if (ii) of the present theorem is satisfied.

Consider the function  $x[\phi_n(t)]$ , inducing the function  $x[\phi_n(\tau)]$  on  $T \times \Omega^*$ . The latter function is measurable on  $T \times \Omega^*$ . The hypotheses of (ii) imply that  $x[\phi_n(t)]$  ( $t$  fixed) converges (to  $x(t)$ ) in measure almost everywhere on the  $t$ -axis, and by the device used in the proof of the first part of this theorem, this means that  $x[\phi_n(\tau)]$  converges in measure on  $T \times \Omega^*$  for  $\tau$  restricted to any finite interval  $I$ . There is then a sequence of integers  $\{\alpha_n\}$  such that  $x[\phi_{\alpha_n}(\tau)]$  converges almost everywhere on  $I \times \Omega^*$ , and by the diagonal procedure, a subsequence  $\{a_n\}$  of  $\{\alpha_n\}$  can be obtained for which  $x[\phi_{a_n}(\tau)]$  converges almost everywhere on  $T \times \Omega^*$ . By Fubini's theorem,  $x[\phi_{a_n}(t)]$  ( $t$  fixed) converges with probability 1, necessarily to  $x(t)$ , for almost all values of  $t$ , as was to be proved.

**THEOREM 2.5.** *Let  $P^*$ -measure be so defined that  $t_n \rightarrow t$  implies*

$$(2.17) \quad P^*\left\{\lim_{n \rightarrow \infty} x(t_n) = x(t)\right\} = 1,$$

*for all values of  $t$  except possibly those in a set  $S$  of measure 0. Then*

(i) *there is a measurable stochastic process with space  $\Omega$ , whose elements  $x(t)$  are continuous almost everywhere on the  $t$ -axis, with probability 1, and  $x(t)$  is continuous at  $t=t_0$  with probability 1, for almost all values of  $t_0$ ;*

(ii) *every quasi-separable stochastic process with space  $\Omega \subset \Omega^*$  is measurable, and its elements fulfill the regularity conditions of (i);*

(iii) *if the exceptional  $t$ -set  $S$  is empty, and if there is an everywhere dense denumerable set  $\{r_n\}$  with the property that*

$$(2.18) \quad \lim_{k \rightarrow \infty} P^*\left\{\text{L.U.B.}_{r_i \in I} |x(r_i)| \leq k\right\} = 1,$$

*for every finite interval  $I$ , a quasi-separable process with space  $\Omega \subset \Omega^*$  exists (to which (ii) can then be applied).‡*

† The integral of a  $P$ -measurable function  $f[x(t)]$  over a  $P$ -measurable set  $\Lambda$  will be denoted by  $\int_{\Lambda} f[x(t)] dP$ .

‡ Cf. Lévy (II, p. 345).

First we prove (i). By Theorem 2.4, (ii) there is a measurable process with space  $\Omega$ , whose elements  $x(t)$  satisfy (2.11') almost everywhere on  $T \times \Omega$ . Let  $\bar{x}_n(t)$ ,  $\underline{x}_n(t)$  be defined as L.U.B.  $x(r_j)$  (G.L.B.  $x(r_j)$ ) for all numbers  $r_j$  of the form  $\mu 2^{-\nu}$  ( $\mu = 0, \pm 1, \dots$ ,  $\nu = 1, 2, \dots$ ) in the interval

$$(2.19) \quad \frac{k-1}{2^n} \leq r \leq \frac{k+1}{2^n},$$

where  $k$  is determined by

$$(2.20) \quad \frac{k}{2^n} \leq t < \frac{k+1}{2^n}.$$

Then  $\bar{x}_n(\tau)$ ,  $\underline{x}_n(\tau)$  are measurable on  $T \times \Omega$ . Define  $\bar{x}(t)$ ,  $\underline{x}(t)$  by

$$\bar{x}(t) = \lim_{n \rightarrow \infty} \bar{x}_n(t), \quad \underline{x}(t) = \lim_{n \rightarrow \infty} \underline{x}_n(t).$$

The functions  $\bar{x}(t)$ ,  $\underline{x}(t)$  are respectively the upper limiting function of  $x^{(s)}(t)$  and the lower limiting function of  $x^{(i)}(t)$ . The functions  $\bar{x}(\tau)$ ,  $\underline{x}(\tau)$  are measurable on  $T \times \Omega$  and satisfy

$$(2.11'') \quad \underline{x}(\tau) \leq x^{(i)}(\tau) \leq x(\tau) \leq x^{(s)}(\tau) \leq \bar{x}(\tau)$$

almost everywhere on  $T \times \Omega$ . If  $x^*(t)$ ,  $x_*(t)$  are respectively the upper and lower limiting functions of  $x(t)$ , (2.11'') implies

$$(2.11''') \quad \underline{x}(\tau) \leq x_*(\tau) \leq x(\tau) \leq x^*(\tau) \leq \bar{x}(\tau)$$

almost everywhere on  $T \times \Omega$ . Applying Theorem 1.3, if  $t \in S$ , there is equality of the extremes in (2.11''') for  $\tau = t$ , with probability 1. Then there is equality almost everywhere on  $T \times \Omega$ . The equality of  $x^*(t)$ ,  $x_*(t)$  at  $t_0$  is the condition of continuity at  $t_0$ . Applying Fubini's theorem, we have the statement (i) of Theorem 2.5.

We now prove (ii). If  $P^*$ -measure has the property described, and if  $\Omega$  is the space of a quasi-separable process, let  $\{r_n\}$  be a sequence of numbers with the properties described in Theorem 1.2. Let  $\bar{x}(t)$ ,  $\underline{x}(t)$  be defined as in (i), in terms of the values of  $x(t)$  at the points  $\{r_n\}$ , which no longer have the special form prescribed there. Then if  $x(t) \in \Omega$ ,  $\underline{x}(t)$ ,  $\bar{x}(t)$  are its lower and upper limiting functions respectively, whose equality is the condition of continuity. We have on  $\Omega$

$$(2.21) \quad \underline{x}(t) \leq x(t) \leq \bar{x}(t)$$

for all values of  $t$ . Now for fixed  $\tau$ ,  $\underline{x}(\tau)$ ,  $\bar{x}(\tau)$  are equal with probability 1 (except for  $\tau \in S$ ), applying Theorem 1.3 to the present stochastic process.

Then by Fubini's theorem,  $\underline{x}(\tau) = \bar{x}(\tau)$  almost everywhere on  $T \times \Omega$ , so  $x(\tau)$ , equal to a measurable function almost everywhere on  $T \times \Omega$ , is itself measurable on that space. The process with space  $\Omega$  is thus measurable, and the continuity properties of the elements  $x(t)$  are deduced by applying Fubini's theorem, as in the proof of (i).

Finally we consider (iii). Suppose that the conditions of (iii) are satisfied. Form the functions  $\underline{x}(t)$ ,  $\bar{x}(t)$  as in (i), using the sequence  $\{r_n\}$  with the properties described in the statement of (iii). Let  $\Omega$  be the space of all elements  $x(t)$  satisfying (2.21) for every value of  $t$ . We shall show that  $\bar{P}^*(\Omega) = 1$  by showing that if  $\Gamma^* = \sum_1^\infty \Gamma_n^*$  is a sum of neighborhoods covering  $\Omega$ ,  $P^*(\Gamma^*) = 1$  necessarily. Let  $t_1, t_2, \dots$  be the values of  $t$  used in defining  $\Gamma^*$ , and let  $\Lambda_{k,n}^*$  be the  $\Omega^*$ -set determined by

$$\text{L.U.B. } |x(r_j)| \leq k. \\ |r_j| < n$$

The set

$$\Lambda_0^* = \prod_{n=1}^{\infty} \sum_{k=1}^{\infty} \Lambda_{k,n}^*$$

then has  $P^*$ -measure 1. It consists of the elements  $x(t)$  which are bounded functions (considered defined only at the points  $\{t_j\}$ ) on every finite interval. Let  $\Lambda^*$  be the subset of  $\Lambda_0^*$  for which

$$\left. \begin{aligned} \underline{x}(t_j) &= x(t_j) = \bar{x}(t_j) \\ \underline{x}(r_j) &= x(r_j) = \bar{x}(r_j) \end{aligned} \right\} \quad j = 1, 2, \dots$$

The set  $\Lambda^*$  also has  $P^*$ -measure 1. If  $x_0(t) \in \Lambda^*$ , let  $x_1(t)$  be an element of  $\Omega^*$  equal to  $x_0(t)$  at the points  $\{t_j\}$  and  $\{r_j\}$  and otherwise restricted only by

$$\underline{x}_1(t) = \underline{x}_0(t) \leq x_1(t) \leq \bar{x}_0(t) = \bar{x}_1(t).$$

Then  $x_1(t) \in \Omega \subset \Gamma^*$ . Since  $x_1(t) \in \Gamma^*$  there is an integer  $\nu$  such that  $x_1(t) \in \Gamma_\nu^*$ , and since  $x_1(t_j) = x_0(t_j)$ ,  $j \geq 1$ , this means that  $x_0(t_j) \in \Gamma_\nu^* : \Lambda^* \subset \Gamma^* : P^*(\Gamma^*) = 1$ , as was to be proved. The sequence  $\{r_n\}$  can be used as the sequence of  $t$ -values required by the condition of quasi-separability (cf. Theorem 1.2).

**THEOREM 2.6.** *Suppose that  $P^*$ -measure is so defined that for every number  $t$  and positive number  $\delta$ ,*

$$(2.22) \quad P^*\{x(t + \delta) - x(t) \geq 0\} = 1.$$

*There is then a finite or denumerable sequence of points  $d_1, d_2, \dots$  such that if  $\Omega$  is the set of all elements  $x(t)$  of  $\Omega^*$  which are monotone non-decreasing functions and continuous on the right, except possibly at  $d_1, d_2, \dots$ ,  $\bar{P}^*(\Omega) = 1$ :  $\Omega$  is*

the space of a quasi-separable measurable stochastic process. If  $t_0 \neq d_1, d_2, \dots$ ,  $x(t)$  in  $\Omega$  is continuous at  $t_0$  with probability 1.<sup>†</sup>

(i) Let  $t_1, t_2, \dots$  be any sequence of numbers. If  $t_i > t_k$ ,

$$(2.23) \quad P^*\{x(t_i) \geq x(t_k)\} = 1.$$

Since the set of all pairs  $(t_i, t_k)$  is denumerable,  $x(t)$ , considered as defined only at  $t_1, t_2, \dots$  is monotone non-decreasing, with probability 1 (in terms of  $P^*$ -measure).

(ii) Evidently if  $\epsilon > 0$  and if  $a < b$ ,

$$(2.24) \quad P^*\{x(b) - x(a) \geq \epsilon\}$$

is monotone non-decreasing in  $b$  for fixed  $a$  and in  $-a$  for fixed  $b$ . Then if  $a < t < b$  and if  $a \rightarrow t, b \rightarrow t$  the expression (2.24) converges to a limit which is unaltered if  $a, b$  are restricted to run through rational values. Let  $S(\epsilon, \eta)$  be the  $t$ -set for which this limit is at least  $\eta > 0$ . Let  $\Lambda_r^*$  be the subset of  $\Omega^*$  consisting of the elements  $x(t)$  which are monotone non-decreasing when considered only for rational  $t$ -values. By (i),  $P^*(\Lambda_r^*) = 1$ . If  $t_0 \in S(\epsilon, \eta)$ ,  $x(r)$  ( $r$  rational) has a jump of magnitude  $\epsilon$  at  $t_0$ , if  $x(t)$  is in some set  $M_{t_0}^*$  of  $P^*$ -measure at least  $\eta$ . The set  $S(\epsilon, \eta)$  has only a finite number of points in common with any finite interval  $I$ : otherwise, if  $t_j \in I \cap S(\epsilon, \eta)$ ,  $j = 1, 2, \dots$ , there is a  $P^*$ -measurable set  $M^*$  of  $P^*$ -measure at least  $\eta$  each element of which is in infinitely many of the sets  $M_{t_j}^*$ ,<sup>‡</sup> and if  $x(t) \in M^*$ ,  $x(r)$  has infinitely many jumps in  $I$  of magnitude  $\epsilon$ , an impossibility. The set  $S(\epsilon, \eta)$  is therefore at most denumerably infinite, and the set  $\{d_j\}$  of the theorem is the sum of the sets  $S(1/m, 1/n)$ ,  $m, n = 1, 2, \dots$ .

(iii) We show that  $\bar{P}^*(\Omega) = 1$  by showing that if  $\Gamma^* = \sum_1^\infty \Gamma_n^*$  is any sum of neighborhoods covering  $\Omega$ ,  $P^*(\Gamma^*) = 1$ . Let  $t_1, t_2, \dots$  be any everywhere dense sequence which includes the sequence  $\{d_n\}$  and those values of  $t$  used in defining the neighborhoods of  $\Gamma^*$ , and let  $\Lambda^*$  be the subset of  $\Omega^*$  consisting of those elements  $x(t)$  which, when considered defined only at  $t_1, t_2, \dots$ , are monotone non-decreasing. By (i),  $P^*(\Lambda^*) = 1$ . If  $x(t) \in \Lambda^*$ ,  $x(t)$  (considered defined only at  $t_1, t_2, \dots$ ) is either continuous on the right at  $t_i$  or has a jump there. If  $t_i \neq d_\nu$ ,  $\nu = 1, 2, \dots$ , the first case is true, with probability 1. Then if  $\Lambda_0^*$  is the subset of  $\Omega^*$  consisting of those elements  $x(t)$  which are continuous on the right (considered defined only at  $t_1, t_2, \dots$ ) at each point  $t_1, t_2, \dots$  not in the set  $\{d_n\}$ ,  $P^*(\Lambda_0^*) = P^*(\Lambda^*) = 1$ . If  $x_0(t) \in \Lambda_0^*$  define  $x_1(t)$  by

<sup>†</sup> This does not mean that  $x(t)$  in  $\Omega$  is continuous except at  $d_1, d_2, \dots$ . The probability that  $x(t)$  is continuous at any given point  $t_0$  may be 1 for every value of  $t_0$ , although at the same time the probability that  $x(t)$  is continuous in an interval is less than 1. An example of this is given below. The last statement of Theorem 2.6 implies that the condition on  $P^*$ -measure of Theorem 2.5 is satisfied.

<sup>‡</sup> The set  $M^*$  can be defined as  $\prod_{N=1}^\infty \sum_N M_{t_j}^*$ .

$$(2.25) \quad \begin{aligned} x_1(t) &= \lim_{t_n \downarrow t} x_0(t_n) \dagger \text{ if } t \neq t_j, & j = 1, 2, \dots, \\ x_1(t_j) &= x_0(t_j), & j = 1, 2, \dots. \end{aligned}$$

Then  $x_1(t)$  is monotone non-decreasing and continuous on the right, except perhaps at  $t = d_j, j = 1, 2, \dots$ , so  $x_1(t) \in \Omega$ . For some  $\nu$ ,  $x_1(t) \in \Gamma_\nu^*$ . Since  $\Gamma_\nu^*$  is defined by inequalities on  $x(t)$  at certain of the numbers  $t_1, t_2, \dots$ ,  $x_1(t) \in \Gamma_\nu^*$  implies that  $x_0(t) \in \Gamma_\nu^*$ . Therefore  $\Lambda_0 \subset \Gamma^*: P^*(\Gamma^*) = 1, \bar{P}^*(\Omega) = 1$ .

(iv) The numbers  $\{t_n\}$  required in the definition of quasi-separability (cf. §1) can be taken as the rational numbers in the interval. The process is therefore quasi-separable. Let  $t_0$  be any value of  $t$  not in the sequence  $\{d_n\}$ . Then

$$\lim_{t_n \downarrow t_0} x(t_n) = x(t_0 + 0),$$

if  $x(t) \in \Omega$ , and at the same time  $x(t_n)$  converges in measure on  $\Omega$  to  $x(t_0)$ , so  $x(t_n)$  converges to  $x(t_0) = x(t_0 + 0)$  with probability 1. Similarly

$$\lim_{t_n \uparrow t_0} x(t_n) = x(t_0 - 0) = x(t_0)$$

with probability 1, and we can say in general that if  $t_n \rightarrow t_0$ ,

$$(2.26) \quad P^*\left\{\lim_{n \rightarrow \infty} x(t_n) = x(t_0)\right\} = P\left\{\lim_{n \rightarrow \infty} x(t_n) = x(t_0)\right\} = 1.$$

The  $P^*$ -measure thus satisfies the condition (ii) of Theorem 2.5, and the process is therefore measurable. We have proved that

$$(2.27) \quad x(t_0 + 0) = x(t_0 - 0) = x(t_0)$$

on  $\Omega$  with probability 1, if  $t_0 \neq d_1, d_2, \dots$ , completing the proof of the theorem.

The following example<sup>‡</sup> of a stochastic process of the type just discussed has no point  $d_n$ , and so  $x(t)$  is continuous, with probability 1 at every value of  $t$ . If  $t_1 < t_1 + \delta_1 \leq t_2 < \dots \leq t_n < t_n + \delta_n$ , and if  $\nu_1, \dots, \nu_n$  are any positive integers or 0, we define  $P^*$ -measure by

$$(2.28) \quad \begin{aligned} P^*\{x(0) = 0\} &= 1, \\ P^*\{x(t_j + \delta_j) - x(t_j) = \nu_j, & \quad j = 1, \dots, n\} = \prod_1^n e^{-\delta_j} \frac{\delta_j^{\nu_j}}{\nu_j!}. \end{aligned}$$

The elements of  $\Omega$  are monotone non-decreasing and everywhere continuous on the right. It can be shown that the subspace  $\Omega_0$  of  $\Omega$  whose elements  $x(t)$  all vanish at  $t = 0$ , and are constant except for jumps of magnitude 1, is also

<sup>†</sup> The notation  $t_n \uparrow t$  ( $t_n \downarrow t$ ) will be used to denote approach from below (above).

<sup>‡</sup> Cf. Khintchine (III, pp. 23-24).



of exterior  $P^*$ -measure 1. This space  $\Omega_0$  is the natural space of the time series with Poisson distribution.

Let  $M_\tau^*$  be an  $\Omega^*$ -set in the Borel system over the neighborhoods (1.1) with  $t_i \leq \tau, j=1, \dots, n$ , but otherwise unrestricted. If  $\Omega$  is the space of a stochastic process, and if  $\Lambda$  is  $P$ -measurable,  $P(\Lambda \cdot M_\tau)$  is a completely additive non-negative function of the sets  $M_\tau = M_\tau^* \cdot \Omega$  ( $\tau$  fixed), and  $P(M_\tau) = 0$  implies that  $P(\Lambda \cdot M_\tau) = 0$ . There is therefore a  $P$ -measurable function  $P_{\tau, \Lambda}[x(t)]$  such that

$$(2.29) \quad \int_{M_\tau} P_{\tau, \Lambda}[x(t)] dP = P(\Lambda \cdot M_\tau), \dagger$$

for all  $M_\tau$ . The function  $P_{\tau, \Lambda}$  is uniquely determined up to a set of  $P$ -measure 0 and  $P_{\tau, \Lambda} > k$  defines an  $M_\tau$ -set. This function  $P_{\tau, \Lambda}$  is called the conditional probability function of  $\Lambda$  if  $x(t)$  is known for  $t \leq \tau$ .

**THEOREM 2.7.** *If  $P_{\tau, \Lambda}$  is the conditional probability function of  $\Lambda$  when  $x(t)$  is known for  $t \leq \tau$ , and if  $\phi[x(t)]$  is the characteristic function of the set  $\Lambda$ ,<sup>‡</sup>  $\tau_n \rightarrow \infty$  implies*

$$(2.30) \quad \lim_{n \rightarrow \infty} P_{\tau_n, \Lambda}[x(t)] = \phi[x(t)]$$

with probability 1.

The proof of this is the same as that for the corresponding theorem on discrete stochastic processes, proved by Lévy (IV, pp. 88–89), and will therefore be omitted.

Let  $T_s^*$  be the transformation of  $\Omega^*$  taking the element  $x(t)$  into the element  $x(t+s)$ . This point transformation takes  $P^*$ -measurable sets into  $P^*$ -measurable sets. If  $P^*$ -measure has the property that the transformations  $\{T_s^*\}$  are measure preserving, the  $P^*$ -measure, and any process with space  $\Omega \subset \Omega^*$  with the corresponding  $P$ -measure, will be called temporally homogeneous. In the study of temporally homogeneous processes, a natural tool is the ergodic theorem. The following lemma will make its application possible.

**LEMMA 2.2.** *Let  $f[x(t)]$  be a  $P$ -measurable function defined on the space  $\Omega$  of a temporally homogeneous measurable stochastic process, and suppose that  $T_s^* \Omega = \Omega$  for all values of  $s$ . Then*

$$f[x(t + \tau)] = f[T_\tau^* x(t)]$$

is a measurable function in  $\{\tau, x(t)\}$  space  $T \times \Omega$ .

<sup>†</sup> The existence of such a density function was proved by Nikodym, *Fundamenta Mathematicae*, vol. 15 (1930), pp. 168–179. The discussion here is a specialization of one by Kolmogoroff, (V, pp. 41–43).

<sup>‡</sup> The characteristic function of a point set is defined as 1 on the set and 0 on its complement.

The statement that the process is measurable gives the lemma for  $f[x(t)] = x(0)$ . The proof in the general case consists of showing that the lemma is true for  $f[x(t)]$  the characteristic function of a  $P$ -measurable set (going from neighborhoods to more general sets), and then approximating  $f$  in the general case by linear combinations of such functions.

If  $f[x(t)]$  is as described in the lemma,  $f[T_\tau^*x(t)]$  is Lebesgue measurable in  $\tau$  (for  $x(t)$  a fixed element of  $\Omega$ ), with probability 1. If we suppose that the integral

$$\int_{\Omega} f[x(t)] dP$$

exists, the ergodic theorem states that

$$\lim_{Z \rightarrow \infty} \frac{1}{Z} \int_0^Z f[x(t + \tau)] d\tau$$

exists, with probability 1.†

3. **Differential stochastic processes.** If  $P^*$ -measure is defined in such a way that for  $t_1 < t_2 < \dots < t_n$

$$x(t_2) - x(t_1), \dots, x(t_n) - x(t_{n-1})$$

form an independent set of chance variables,‡ the process will be called a differential process.§ The time series discussed above (p. 122) is an example of a differential process.

An  $\Omega$ -set determined by

$$(3.1) \quad a_j < x(t_j + \delta_j) - x(t_j) < b_j, \quad j = 1, \dots, n, \quad (-\infty \leq a_j < b_j \leq +\infty)$$

will be called a differential neighborhood. An  $\Omega$ -set in the Borel system of sets over the differential neighborhoods, or differing from such a set by a subset of such a set which is of  $P$ -measure 0, will be called a differential set. A function  $f[x(t)]$  will be called measurable on the differential sets if the set defined by  $f > k$  is a differential set for every number  $k$ .

† Cf. Doob (I); Khintchine (II, pp. 613–615); Wiener (I, Chapter X); Khintchine's treatment is open to the objection that he does not justify integrating his chance variables with respect to  $t$ .

Doob (II, pp. 763–765), Hopf (I, p. 95), and Khintchine (I) applied the ergodic theorem, in the discrete case, to derive the law of large numbers.

A simple presentation of the ergodic theorem, first proved by Birkhoff, was given by Khintchine (Mathematische Annalen, vol. 107 (1933), pp. 485–488).

‡ This means that for any real numbers  $a_1, \dots, a_n$

$$P^*\{x(t_j) - x(t_{j-1}) > a_j, j = 2, \dots, n\} = \prod_{j=2}^n P^*\{x(t_j) - x(t_{j-1}) > a_j\}.$$

§ Differential processes have been discussed by Bachelier (I, II, III), Khintchine (III, pp. 68–75), Kolmogoroff (I, II), Lévy (II, III), and Wiener (I, Chapters IX and X).

The definition of temporal homogeneity given in the preceding section is not the most convenient one, when differential processes are under investigation. Throughout this section, we shall understand by the temporal homogeneity of a differential process merely the invariance of the  $P^*$ -measure of differential sets under the transformations  $\{T_s^*\}$ .

It will be convenient to define a transformation  $T_s\Lambda$  on the  $P$ -measurable sets of the space  $\Omega$  of a stochastic process by

$$T_s\Lambda = \Omega \cdot T_s^*\Lambda^*,$$

where  $\Lambda^*$  is a  $P^*$ -measurable set satisfying  $\Lambda^* \cdot \Omega = \Lambda$ . If  $T_s^*$  takes sets of  $P^*$ -measure 0 into sets of  $P^*$ -measure 0,  $T_s\Lambda$  is uniquely determined up to a set of  $P$ -measure 0. If  $T_s^*\Omega = \Omega$ ,

$$T_s\Lambda = \Omega \cdot T_s^*\Lambda^* = T_s^*(\Omega \cdot \Lambda^*) = T_s^*\Lambda,$$

so in this case  $T_s$  becomes  $T_s^*$  on the  $P$ -measurable sets. If the process is a temporally homogeneous differential one,  $T_s\Lambda$  can be defined as above for  $\Lambda$  any differential set, and the definition will be unique, up to differential sets of  $P$ -measure 0.

**THEOREM 3.1.** *Let  $B_\tau$  be the Borel system of sets over the  $\Omega$ -sets determined by the differential neighborhoods (3.1) with  $t_i, t_i + \delta_i > \tau, j = 1, \dots, n$ , where  $\Omega$  is the space of a stochastic differential process. Then if  $\Lambda$  is a  $P$ -measurable set which is in  $B_\tau$  for all values of  $\tau$ ,  $P(\Lambda) = 1$  or  $P(\Lambda) = 0$ .†*

Using the notation of Theorem 2.7, if  $M_\tau$  is  $P$ -measurable,

$$(3.2) \quad P(\Lambda \cdot M_\tau) = P(\Lambda) \cdot P(M_\tau).$$

Then it follows that if  $P(\Lambda) > 0$ ,  $P_{\tau, \Lambda} = P(\Lambda)$  almost everywhere on  $\Omega$ , so that, by Theorem 2.7,  $P(\Lambda) = 1$ .

Using this theorem, if  $t_n \rightarrow \infty$ , the probability of convergence of  $x(t_n)$  is either 0 or 1. Or by a simple transformation of the  $t$ -axis, if  $t_n$  approaches  $t$  from one side, the probability of convergence is either 0 or 1.

**THEOREM 3.2.** *Any differential set, invariant (up to a set of  $P$ -measure 0) under the transformations  $\{T_s^*\}$  on the differential sets of  $\Omega$ , the space of a temporally homogeneous differential stochastic process, has  $P$ -measure 0 or 1.‡*

† Cf. Jessen (I, pp. 270–271), Kolmogoroff (III, pp. 60–61), and Lévy (IV, pp. 88–89), for the analogous theorem on discrete stochastic processes.

‡ Another way of stating this theorem is that the transformations  $\{T_s^*\}$  are metrically transitive over the collection of differential sets. The corresponding theorem was proved for discrete processes by Doob (II, pp. 761–763) and Hopf (I, p. 95). The method used here can also be used in that case.

Let  $\Lambda$  be a differential set of  $P$ -measure  $\rho$  invariant under the transformations  $\{T_s\}$ . If  $0 < \rho < 1$ , there is a number  $\epsilon > 0$  so small that

$$(3.3) \quad (\rho + \epsilon)^2 < \rho - 2\epsilon.$$

Let  $\Gamma$  be a finite sum of differential neighborhoods satisfying

$$(3.4) \quad P(\Gamma) < \rho + \epsilon, \quad P\{\Lambda \cdot c\Gamma\} < \epsilon. \dagger$$

Suppose that the  $t$ -values involved in defining  $\Gamma$  are all in the interval  $a < t < b$ , and let  $\Gamma' = T_{b-a}\Gamma$ . Then

$$(3.5) \quad P(\Gamma') = P(\Gamma) < \rho + \epsilon, \quad P(\Gamma \cdot \Gamma') < (\rho + \epsilon)^2, \quad P(\Lambda \cdot c\Gamma') < \epsilon,$$

so

$$(3.6) \quad \begin{aligned} (\rho + \epsilon)^2 &> P(\Gamma \cdot \Gamma') \geq P(\Lambda \cdot \Gamma \cdot \Gamma') \\ &\geq P(\Lambda) - P(\Lambda \cdot c\Gamma) - P(\Lambda \cdot c\Gamma') > \rho - 2\epsilon, \end{aligned}$$

contradicting (3.3).

LEMMA 3.1. *Let  $\Delta_1, \Delta_2, \dots$  be a sequence of independent chance variables. If there is a chance variable  $\Delta$  such that*

$$\Delta = \sum_1^n \Delta_i$$

*is independent of the set  $\{\Delta_1, \dots, \Delta_n\}$  for  $n=1, 2, \dots$  there is a sequence of constants  $\{m_n\}$  such that*

$$(3.7) \quad \lim_{n \rightarrow \infty} \left[ \sum_1^n \Delta_i - m_n \right]$$

*exists and is finite with probability 1.*

The existence of constants  $\{m_n\}$  with the property described was proved by Lévy (I, p. 132).

THEOREM 3.3. *Choose  $m(t)$  to satisfy*

$$(3.8) \quad \int_{\Omega} \arctan [x(t) - m(t)] dP = 0 \dagger$$

*for each value of  $t$ , and set  $y(t) = x(t) - m(t)$ . Then if the process is a differential one, and if  $t$  is any real number, there is a chance variable  $y(t-0)$  ( $y(t+0)$ ) such that  $t_n \uparrow t$  ( $t_n \downarrow t$ ) implies*

$\dagger$  The notation  $c\Gamma$  will be used to denote  $\Omega - \Gamma$ . The existence of  $\Gamma$  can be proved, for example, by transfinite induction.

$\ddagger$  For  $t$  fixed the integral is a monotone decreasing function of  $m$ , varying from  $\pi/2$  to  $-\pi/2$  as  $m$  varies from  $-\infty$  to  $+\infty$ , so there is a unique value of  $m$  satisfying (3.8).

$$(3.9) \quad \lim_{n \rightarrow \infty} y(t_n) = y(t-0), \quad \lim_{n \rightarrow \infty} y(t_n) = y(t+0)$$

with probability 1. Except at most at a denumerable  $t$ -set,

$$y(t-0) = y(t) = y(t+0)$$

with probability 1.†

Suppose that  $t_n \uparrow t$ . We can assume that  $t_1 < t_2 < \dots$ . Then

$$(3.10) \quad y(t_n) - y(t_1) = \sum_1^n \Delta_i, \quad \Delta_i = y(t_{i+1}) - y(t_i).$$

Taking the variable  $\Delta$  of Lemma 3.1 as  $y(t) - y(t_1)$ , it follows that a sequence of constants  $\{m_n\}$  exists such that

$$(3.11) \quad \lim_{n \rightarrow \infty} [y(t_n) - m_n] = y$$

exists and is finite with probability 1. If  $\limsup_{n \rightarrow \infty} |m_n| = \infty$ , some subsequence of  $\{\arctan y(t_n)\}$  must approach  $-\pi/2$  or  $\pi/2$ , contradicting

$$(3.8') \quad \int_{\Omega} \arctan y(t) dP = 0.‡$$

Unless the sequence  $\{m_n\}$  is convergent, there are two subsequences  $\{m_{a_n}\}$ ,  $\{m_{b_n}\}$  converging to different finite limits  $m'$ ,  $m''$  respectively. Then

$$(3.12) \quad \lim_{n \rightarrow \infty} y(t_{a_n}) = y + m', \quad \lim_{n \rightarrow \infty} y(t_{b_n}) = y + m'',$$

contradicting (3.8) again. Thus  $\lim_{n \rightarrow \infty} m_n$  exists and is finite, implying the same (with probability 1) for  $\lim_{n \rightarrow \infty} y(t_n)$ . The latter limit must be independent of the particular sequence  $\{t_n\}$  chosen since any pair of such sequences can be arranged in a single sequence. The limit  $y$  is thus the  $y(t-0)$  of (3.9). The other half of (3.9) is proved in the same way. We shall need the fact that

$$(3.8'') \quad \int_{\Omega} \arctan y(t-0) dP = \int_{\Omega} \arctan y(t+0) dP = 0.$$

To prove the last part of the theorem, define  $\alpha_t(v)$  by

† This theorem is due to Lévy (II, pp. 342–343). Lévy centers his distributions by a different function of  $t$ .

‡ We shall use repeatedly in this discussion the theorem of Lebesgue that a uniformly bounded convergent sequence of measurable functions defined on a measurable set of finite measure can be integrated term by term.

$$(3.13) \quad \alpha_t(v) = \left| \int e^{iv[v(t)-v(0)]} dP \right|, \quad t > 0.$$

Then  $\alpha_t(v)$  is non-negative and continuous in  $v$ . Since  $\alpha_{t+h}(v)$ ,  $h > 0$ ,  $t > 0$ , is the product of  $\alpha_t(v)$  (which is less than or equal to 1) and a function obtained similarly from the distribution of  $y(t+h) - y(t)$ ,<sup>†</sup>  $\alpha_t(v)$  is monotone non-increasing in  $t$ . It can readily be shown that at any  $t$ -point where  $y(t-0) - y(t)$ ,  $y(t+0) - y(t)$  are not identically constant, with probability 1,  $\alpha_t(v)$  must have a jump for some rational value of  $v$ . Then the set of such points is at most denumerable. Moreover if either of these two differences is identically constant with probability 1, the constant must be 0, because of (3.8''), so the theorem is completely proved, for  $t > 0$ . A similar discussion can be given for  $t < 1$ , and the combination of the two results gives the theorem in full.

**COROLLARY.** *If the process of Theorem 3.3 is measurable,  $m(t)$  is a Lebesgue measurable function, and the process whose elements are  $\{y(t)\}$  is a measurable differential process.*

The inequality  $m(t) < k$  is equivalent to

$$(3.14) \quad \int_{\Omega} \arctan [x(t) - k] dP < 0.$$

Now  $\arctan [x(t) - k]$  is a measurable bounded function in  $\{\tau, x(t)\}$ -space  $T \times \Omega$ , so the integral of (3.14), a Lebesgue measurable function of  $t$  (Fubini's theorem), is negative on a Lebesgue measurable  $t$ -set. The function  $m(t)$  is thus Lebesgue measurable. Let  $\Omega'$  be the space whose elements are the functions  $y(t)$ . We define  $P'$ -measure by

$$(3.15) \quad P'\{y(t_j) < \mu_j, j = 1, \dots, n\} = P\{x(t_j) < \mu_j + m(t_j), j = 1, \dots, n\},$$

for any  $t_1, \dots, t_n, \mu_1, \dots, \mu_n$ . It is readily seen that in this way a measurable differential stochastic process with space  $\Omega'$  is defined.

The choice of the function  $m(t)$  is to a certain extent arbitrary. If  $m_1(t) - m(t)$  is continuous,  $m_1(t)$  would serve as well for the purposes of Theorem 3.3. If, whenever  $t_n \uparrow t$  or  $t_n \downarrow t$ ,

$$\lim_{t_n \uparrow t} x(t_n), \quad \lim_{t_n \downarrow t} x(t_n)$$

---

<sup>†</sup> This follows at once from the fact that if  $f$  and  $g$  are bounded  $P$ -measurable functions, which are independent in the usual sense of probability,

$$\int f \cdot g dP = \int f dP \cdot \int g dP,$$

a fact which in turn is simply the fact that an iterated integral, if the limits of integration are constants, is the product of its component integrals.

exist and are finite with probability 1, and if the differences  $x(t-0) - x(t)$ ,  $x(t+0) - x(t)$  can never be identically constant with probability 1 unless the constant is 0, the process will be said to be centered. If a process is centered, the proof of Theorem 3.3 shows that at most a denumerably infinite set of  $t$ -values are points of "fixed discontinuities" (where  $P\{x(t-0) = x(t)\} < 1$  or  $P\{x(t+0) = x(t)\} < 1$ ). Even these cannot occur if the process is temporally homogeneous. The  $P^*$ -measure of a centered differential process thus satisfies the condition of Theorem 2.5 on  $P^*$ -measure, so a measurable process always exists with the same measure relations. This result will be considerably strengthened by Theorem 3.8. The above corollary can be restated as follows. *If the process of Theorem 3.3 is measurable,  $m(t)$  is a Lebesgue measurable function, and the process whose elements are  $\{y(t)\}$  is a centered measurable process.*

**THEOREM 3.4.** *If  $x(t)$  is the general element of a quasi-separable differential stochastic process with the property that the integrals*

$$(3.16) \quad \int_{\Omega} [x(t) - x(0)] dP, \quad \int_{\Omega} [x(t) - x(0)]^2 dP$$

*exist for all  $t$ , and if the first vanishes identically,  $\sigma_t^2$  is monotone non-decreasing in  $t$  for  $t > 0$ , and if  $T > 0$ ,  $\mu > 0$ ,*

$$(3.17) \quad P\left\{ \text{L.U.B.}_{0 < t < T} |x(t) - x(0)| > \mu \right\} < \sigma_T^2 / \mu^2.$$

Since if  $\delta > 0$

$$(3.18) \quad \sigma_{t+\delta}^2 = \sigma_t^2 + \int_{\Omega} [x(t+\delta) - x(t)]^2 dP,$$

$\sigma_{t+\delta} \geq \sigma_t$ . The  $\Omega$ -set  $\Lambda$  where

$$(3.19) \quad \text{L.U.B.}_{0 < t < T} |x(t) - x(0)| > \mu$$

is the set where

$$(3.20) \quad \text{L.U.B.}_{j \geq 1} |x(\lambda_j) - x(0)| > \mu$$

for some sequence  $\{\lambda_n\}$  in  $0 < t < T$ , since the process is quasi-separable. The set  $\Lambda$  is  $P$ -measurable and

$$(3.21) \quad P(\Lambda) = \lim_{n \rightarrow \infty} P\left\{ \max_{j \leq n} |x(\lambda_j) - x(0)| > \mu \right\}.$$

By Kolmogoroff's generalization of the Tchebycheff inequality (III, p. 310,

Theorem 1), the  $P$ -measure of the  $n$ th set on the right is at most  $\sigma(n)^2/\mu^2$ , where if  $\bar{\lambda}_n = \max_{j \leq n} \lambda_j$ ,  $\sigma(n)^2 = \sigma_{\bar{\lambda}_n}^2$ . Then

$$(3.22) \quad P(\Lambda) \leq \limsup_{n \rightarrow \infty} \sigma(n)^2/\mu^2 \leq \sigma_T^2/\mu^2,$$

as was to be proved.

This theorem could be used to derive the analogue of Theorem 3.3 in the special case being considered. Under the hypotheses, the process is already centered.

**THEOREM 3.5.** *Let  $\bar{B}_\tau$  be the Borel system of sets over the  $\Omega$ -sets determined by inequalities of the form*

$$(3.23) \quad a_j < x(t_j - \delta_j) - x(t_j) < b_j, \quad \delta_j > 0, \quad t_j \leq \tau, \quad j = 1, \dots, n,$$

where  $\Omega$  is the space of a centered temporally homogeneous differential stochastic process. Let  $s[x(t)]$  be a  $P$ -measurable function with the property that the  $\Omega$ -set determined by  $s \geq \tau$  is in  $\bar{B}_\tau$  for every number  $\tau$ . Let

$$y(t) = x(t + s[x(t)]).$$

(i) *If  $s[x(t)]$  takes on at most denumerably many values,*

$$(3.24) \quad \begin{aligned} P\{y(t_j + \delta_j) - y(t_j) < \mu_j, j = 1, \dots, n\} \\ = P\{x(t_j + \delta_j) - x(t_j) < \mu_j, j = 1, \dots, n\} \end{aligned}$$

whenever  $0 < t_1 < t_1 + \delta_1 \leq t_2 < \dots \leq t_n < t_n + \delta_n$ .

(ii) *If the process is quasi-separable, the conclusion of (i) holds without the restriction that  $s$  take on at most denumerably many values.†*

(i) If  $s$  assumes only the values  $a_1, a_2, \dots$  on sets  $\Lambda_1, \Lambda_2, \dots$  respectively, and if  $0 < t_1 < t_1 + \delta_1 \leq t_2 < \dots \leq t_n < t_n + \delta_n$ ,

$$(3.25) \quad \begin{aligned} &P\{y(t_j + \delta_j) - y(t_j) < \mu_j, j = 1, \dots, n\} \\ &= \sum_\nu P\{s = a_\nu, x(t_j + a_\nu + \delta_j) - x(t_j + a_\nu) < \mu_j, j = 1, \dots, n\} \\ &= \sum_\nu P\{s = a_\nu\} \cdot P\{x(t_j + a_\nu + \delta_j) - x(t_j + a_\nu) < \mu_j, j = 1, \dots, n\} \\ &= \sum_\nu P\{s = a_\nu\} \cdot P\{x(t_j + \delta_j) - x(t_j) < \mu_j, j = 1, \dots, n\} \end{aligned}$$

which is equivalent to (3.24), since  $\sum_\nu P\{s = a_\nu\} = 1$ .

(ii) Define  $s_n$  by

$$(3.26) \quad s_n = (k-1)/n \quad \text{if} \quad (k-1)/n \leq s < k/n, \quad k = 0, 1, \dots$$

Then  $|s_n - s| \leq 1/n$ . The hypotheses of the theorem imply that

† Doob (III) proved this theorem, in a somewhat more general form, for discrete processes.



$$\lim_{t' \rightarrow t} x(t') = x(t)$$

with probability 1 ( $t$  fixed). There is therefore a positive function  $\delta(\epsilon, \eta)$ , independent of  $t$ , such that if  $t$  is fixed, and if  $\epsilon > 0$ ,  $\eta > 0$ ,

$$(3.27) \quad P\left\{ \text{L.U.B.}_{|\bar{\delta}| < \delta} |x(t + \bar{\delta}) - x(t)| < \epsilon \right\} > 1 - \eta.$$

By (i) the distribution of

$$x(s_n + t + \bar{\delta}) - x(s_n + t), \quad \bar{\delta} > 0,$$

is the same as that of

$$x(t + \bar{\delta}) - x(t),$$

if  $t > 1/n$ , and evidently the elements  $x(s_n + t)$  ( $n$  fixed),  $t > 1/n$ , can be considered as the elements of a quasi-separable stochastic process, so

$$(3.28) \quad P\left\{ \text{L.U.B.}_{|\bar{\delta}| < \delta} |x(s_n + t + \bar{\delta}) - x(s_n + t)| < \epsilon \right\} > 1 - \eta, \quad (t > 1/n), \\ n = 1, 2, \dots$$

If  $\nu = \nu(\epsilon, \eta)$  is so large that  $1/\nu < \delta$ ,

$$(3.29) \quad |x(s + t) - x(s_\nu + t)| = |x(s_\nu + t + s - s_\nu) - x(s_\nu + t)| < \epsilon, \quad t > 1/\nu$$

with probability at least  $1 - \eta$ : the sequence  $x(s_n + t)$  converges in probability, i.e., in measure on  $\Omega$ , to  $x(s + t)$ , if  $t > 0$ . Then  $y(t) = x(s + t)$  is readily seen to have the same differential distribution as  $x(t)$  ( $t > 0$ ), another way of stating (3.24).

We consider now a differential process for which the integrals (3.16) exist, the first vanishing identically. We shall assume, for greater simplicity, that there are no fixed  $t$ -points of discontinuity, meaning in this case that  $\sigma_t^2$  is a continuous function of  $t$ . Then if  $V(t)$  is defined by

$$(3.30) \quad V(t) = \sigma_t^2, \quad t > 0; \quad V(t) = -\sigma_t^2, \quad t < 0; \quad V(0) = 0,$$

we shall define a chance variable (i.e., a  $P$ -measurable function)

$$(3.31) \quad \int f(t) dx(t)$$

for any Baire function  $f(t)$  for which

$$(3.32) \quad \int f(t)^2 dV(t)$$

exists as a Lebesgue-Stieltjes integral.

Let  $f(t)$  be the characteristic function of an open bounded  $t$ -set, the sum of intervals  $a_j < t < b_j$ ,  $j=1, 2, \dots$ . Then we make the definition

$$(3.33) \quad \int f(t) dx(t) = \sum_1^{\infty} [x(b_j) - x(a_j)],$$

where the series on the right must converge, since the terms are independent chance variables whose means are 0 and whose standard deviations squared form a convergent series:†

$$(3.34) \quad \sum_1^{\infty} \int_{\Omega} [x(b_j) - x(a_j)]^2 dP = \sum_1^{\infty} [V(b_j) - V(a_j)].$$

We note that

$$(3.35) \quad \int_{\Omega} \left\{ \int f(t) dx(t) \right\}^2 dP = \int_O dV(t).$$

If  $B$  is any bounded Borel set, there is a monotone sequence  $O_1 \supset O_2 \supset \dots$  of bounded open sets whose product includes  $B$  and differs from it at most by a set of  $V$ -measure 0; i.e., by a set on which  $\int dV(t) = 0$ . If  $f, f_1, f_2, \dots$  are the characteristic functions of  $B, O_1, O_2, \dots$ , respectively,

$$(3.36) \quad \int_{\Omega} \left\{ \int f_n(t) dx(t) - \int f_m(t) dx(t) \right\}^2 dP = \int_{O_n - O_m} dV(t), \quad m > n,$$

so that  $\int f_n(t) dx(t)$  converges in the mean on  $\Omega$ . It can be shown that the limit is independent of the sequence  $\{O_n\}$  chosen (disregarding  $\Omega$ -sets of  $P$ -measure 0). This limit is denoted by (3.31), and (3.35) remains true, substituting  $B$  for  $O$ .

If  $f(t)$  is a Baire function taking on only a finite number of values, and equal to 0 outside some finite interval,  $f(t)$  is a linear combination of functions of the type just treated, and (3.31) is defined as the corresponding linear combination of integrals. Equation (3.35) becomes

$$(3.35') \quad \int_{\Omega} \left\{ \int f(t) dx(t) \right\}^2 dP = \int_{-\infty}^{\infty} f(t)^2 dV(t).$$

If  $f(t)$  is any Baire function for which (3.32) exists, it can be approximated in the mean by Baire functions  $\{f_n(t)\}$  each of which takes on only a finite number of values:

$$(3.37) \quad \lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} [f_n(t) - f(t)]^2 dV(t) = \lim_{n, m \rightarrow \infty} \int_{-\infty}^{+\infty} [f_n(t) - f_m(t)]^2 dV(t) = 0.$$

† Kolmogoroff (III, p. 314, Theorem 7).

Then

$$(3.38) \quad \int_{\Omega} \left\{ \int f_n(t) dx(t) - \int f_m(t) dx(t) \right\}^2 dP \\ = \int_{\Omega} \left\{ \int [f_n(t) - f_m(t)] dx(t) \right\}^2 dP = \int_{-\infty}^{+\infty} [f_n(t) - f_m(t)]^2 dV(t) \rightarrow 0,$$

so that the sequence  $\{ \int f_n(t) dx(t) \}$  converges in the mean. The chance variable (3.31) is defined as the corresponding limit, which is readily seen to be independent of the approximating sequence, neglecting  $\Omega$ -sets of  $P$ -measure 0.

**THEOREM 3.6.** *If  $f_1(t), f_2(t)$  are Baire functions, and if*

$$(3.39) \quad \int_{-\infty}^{+\infty} f_j(t)^2 dV(t) < \infty, \quad j = 1, 2,$$

*then*

$$(3.40) \quad \int f_j(t) dx(t), \quad \int_{\Omega} \left\{ \int f_j(t) dx(t) \right\}^2 dP$$

*are defined, and*

$$(3.41) \quad \int_{\Omega} \left\{ \int f_1(t) dx(t) \cdot \int f_2(t) dx(t) \right\} dP = \int_{-\infty}^{+\infty} f_1(t) \cdot f_2(t) dV(t). \dagger$$

We have already defined  $\int f_j(t) dx(t)$ . Equation (3.35') is a special case of (3.41) with  $f_1 = f_2 = f$ ,  $f$  taking on only a finite number of values, and (3.35') is true whenever the integral (3.32) is finite, since the corresponding equality is true for the approximating functions. Equation (3.41) is deduced from (3.35') by writing (3.35') for  $f_1 + f_2$ ,  $f_1 - f_2$  and finding the difference between the two equations thus obtained.

**THEOREM 3.7.** *Let  $\phi_1(t), \phi_2(t), \dots$  be a complete normal orthogonal set with respect to the measure function  $V(t)$ , for functions defined on the interval  $a \leq t \leq b$  ( $-\infty \leq a < b \leq +\infty$ ). Then if*

$$(3.42) \quad \Phi_n[x(t)] = \int \phi_n(t) dx(t),$$

*the functions  $\{ \Phi_n \}$  form a complete normal orthogonal set for the space of functions (3.40) on  $\Omega$  with  $f(t) = 0$  if  $t < a$ ,  $t > b$ :*

---

<sup>†</sup> The integral over  $\Omega$  in (3.41) exists, by Schwarz' inequality, since the second pair of integrals in (3.40) exists.

$$(3.43) \quad \int_{\Omega} \Phi_n \cdot \Phi_m dP = \delta_{n,m},$$

and

$$(3.44) \quad \begin{aligned} \int f(t) dx(t) &= \text{l.i.m.}_{N \rightarrow \infty} \sum_1^N c_n \Phi_n, \dagger \\ c_n &= \int \left\{ \Phi_n \cdot \int f(t) dx(t) \right\} dP = \int_a^b \phi_n(t) f(t) dV(t). \end{aligned}$$

In particular

$$(3.45) \quad \begin{aligned} x(\tau) - x(a) &= \text{l.i.m.}_{N \rightarrow \infty} \sum_1^N c_n(\tau) \Phi_n, \\ c_n(\tau) &= \int_{\Omega} \left\{ \Phi_n \cdot [x(\tau) - x(a)] \right\} dP = \int_a^{\tau} \phi_n(t) dV(t). \end{aligned}$$

The functions  $\{\Phi_n\}$  form a normal orthogonal set since by (3.41),

$$(3.46) \quad \int_{\Omega} \Phi_n \cdot \Phi_m dP = \int_a^b \phi_n \cdot \phi_m dV(t) = \delta_{n,m},$$

and (3.44) follows by applying (3.41) to

$$f_1 = f_2 = f - \sum_1^N c_n \phi_n.$$

If  $f(t)$  is taken as the characteristic function of the interval  $a \leq t \leq \tau$ , (3.44) becomes (3.45).

The elements of the process have thus been made to depend upon a denumerably infinite set of parameters  $\Phi_1, \Phi_2, \dots$ . The discussion given here is a generalization of one given by Wiener, who supposes that  $x(t+\delta) - x(t)$  has a Gaussian distribution with zero mean and standard deviation proportional to  $(|\delta|)^{1/2}$ .<sup>‡</sup> Wiener defines the elements  $x(t)$  by (3.45), taking advantage of the fact that in this case there is a subset  $\Omega_0$  of  $\Omega^*$  of  $P^*$ -measure 1 on which the series in (3.45) converges to a continuous function of  $\tau$ .<sup>§</sup>

**THEOREM 3.8.** *Let  $P^*$ -measure be so defined that  $\Omega^*$  is the space of a centered differential stochastic process. Let  $\Omega_0$  be the space of elements  $x(t)$  which are con-*

<sup>†</sup> Limit in the mean will be denoted by l.i.m.

<sup>‡</sup> Cf. Wiener (I, Chapter IX).

<sup>§</sup> Cf. our proof of Theorem 3.9.

tinuous except possibly for discontinuities of the first kind,<sup>†</sup> and everywhere continuous on the right except perhaps at fixed points of discontinuity. Then  $P^*(\Omega_0) = 1$ :  $\Omega_0$  is the space of a quasi-separable measurable differential stochastic process.

If  $\Omega$  is the space of a quasi-separable centered differential stochastic process, the elements  $x(t)$  are everywhere continuous, except possibly for discontinuities of the first kind.<sup>‡</sup>

The first part of the theorem is really an existence theorem, and therefore the second part would have no content unless the first part, or some similar statement, were true.

It will be sufficient to prove the theorem under the assumption that there are no points of fixed discontinuity (the denumerably many jumps at such points in any finite interval can be subtracted out). Suppose then that  $\Omega \subset \Omega^*$  is the space of a centered differential process, and that  $t_n \rightarrow t$  implies that

$$P^* \left\{ \lim_{n \rightarrow \infty} x(t_n) = x(t) \right\} = 1.$$

(i) Let  $R: \{r_n\}$  be a sequence of numbers satisfying the following conditions:  $R$  includes the rational numbers; if  $\rho_j \in R, j=1, \dots, n, (\rho_1 \pm \rho_2 \pm \dots \pm \rho_n) \in R$ . The set  $R$  is everywhere dense. Then by Theorem 1.3,

$$(3.47) \quad \lim_{r \rightarrow t} x(r) = x(t) \quad (r \in R)$$

with probability 1.

The oscillation of  $x(r)$ , ( $r \in R$ ) at  $t$  is defined as

$$(3.48) \quad \lim_{\epsilon \rightarrow 0} \text{L.U.B.}_{|t-r|+|t-r'| < \epsilon} |x(r) - x(r')|.$$

This oscillation is 0 almost everywhere on  $\Omega$ , for each value of  $t$ . Let  $I$  be any closed  $t$ -interval  $a \leq t \leq b$ . The set

$$(3.49) \quad \prod_{\nu=1}^{\infty} \sum_{\substack{|r_j - r_k| < 1/\nu \\ a-1/\nu < r_j, r_k < b+1/\nu}} \{ |x(r_j) - x(r_k)| > \alpha - 1/\nu \}, \quad \alpha > 0,$$

consists of the elements  $x(t)$  for which the oscillation of  $x(r)$  is at least  $\alpha$  at some point of  $I$ . This set is  $P$ -measurable, and its complement is also: if  $S_\alpha[(t)]$  is the  $t$ -set at which the oscillation of  $x(r)$  is at least  $\alpha$ , the  $\Omega$ -set of elements for which  $I \cdot S_\alpha[x(t)] = 0$  is  $P$ -measurable.

<sup>†</sup> A function  $x(t)$  is said to have a discontinuity of the first kind at  $a$  if the limits  $\lim_{t \rightarrow a} x(t)$ ,  $\lim_{t \rightarrow a+0} x(t)$  exist and are unequal.

<sup>‡</sup> Cf. Lévy (II, pp. 359-364; III, pp. 217-218).

(ii) We shall now prove that the points of the set  $S_\alpha[x(t)]$  are points of discontinuity of the first kind, with probability 1. Unless this is true, there is an interval  $I: a \leq t \leq b$  so large that the set of elements  $x(t)$  for which there is a point  $t, t \in I \cdot S_\alpha[x(t)], t < b, t$  a point of discontinuity of  $x(r)$  of the second kind, is not of  $P$ -measure 0. The set  $S_\alpha[x(t)]$  is closed for each element  $x(t)$ . Let  $s_1[x(t)]$  be the first point of  $S_\alpha[x(t)]$  in  $I$ , or  $b$ , if  $I \cdot S_\alpha[x(t)] = 0$ . According to (i), the function  $s_1[x(t)]$  is  $P$ -measurable, and evidently it has the property described in Theorem 3.5. Then if  $s_{1,n}$  is defined by

$$(3.50) \quad s_{1,n} = k/n \text{ if } k/n \leq s_1 < (k+1)/n, \quad k = 0, \pm 1, \dots, n = 1, 2, \dots, \\ s_{1,n} + 1/n \text{ has this same property. According to (i), if } \epsilon > 0, \text{ there is a number} \\ \delta = \delta(\epsilon, t) \text{ such that}$$

$$(3.51) \quad P\left\{ \text{L.U.B.} \left| x(r') - x(r'') \right| \leq \epsilon \right\} \geq 1 - \epsilon, \quad r', r'' \in R. \\ \begin{matrix} |t-r'| < \delta \\ |t-r''| < \delta \end{matrix}$$

An immediate application of the Heine-Borel theorem shows that a function  $\delta(\epsilon)$  exists, independent of  $t$ , such that (3.51) is true for  $\delta = \delta(\epsilon)$  when  $t \in I$ . Then if  $\delta = \delta(\epsilon)$ ,

$$(3.52) \quad P\left\{ \text{L.U.B.} \left| x(s_{1,n} + 1/n + r') - x(s_{1,n} + 1/n + r'') \right| \leq \epsilon \right\} \\ = \sum_k P\{s_{1,n} = (k-1)/n\} \\ \cdot P\left\{ \text{L.U.B.} \left| x(k/n + r') - x(k/n + r'') \right| \leq \epsilon \right\} \geq 1 - \epsilon, \\ \begin{matrix} 0 < r', r'' < \delta \end{matrix}$$

or

$$(3.53) \quad P\left\{ \text{L.U.B.} \left| x(r') - x(r'') \right| \leq \epsilon \right\} \geq 1 - \epsilon,$$

where  $r', r''$  are points of  $R$  satisfying

$$(3.54) \quad s_{1,n} + 1/n < r', r'' < s_{1,n} + 1/n + \delta,$$

which will surely be true (for  $1/n < \delta$ ) if

$$(3.55) \quad s_1 + 1/n < r', r'' < s_1 + \delta.$$

Since  $n$  can be made arbitrarily large,

$$(3.56) \quad P\left\{ \text{L.U.B.} \left| x(r') - x(r'') \right| \leq \epsilon \right\} \geq 1 - \epsilon. \\ s_1 < r', r'' < s_1 + \delta$$

This inequality implies that  $\lim_{r \rightarrow s_1} x(r)$  exists with probability 1. If  $s_1 < b$ ,  $s_1$  is therefore an isolated point of  $S_\alpha[x(t)]$  with probability 1. It is then readily seen (applying the results just obtained to the process with elements  $x(-t)$ )

that  $\lim_{r \rightarrow s} x(r)$  exists with probability 1. Then  $x(r)$  has at worst a discontinuity of the first kind at  $s_1$ , with probability 1. If  $s_2$  is the second point of  $S_\alpha$  in  $I$  (or  $b$  if there is no second point), it can be shown in the same way that  $x(r)$  has at worst a discontinuity of the first kind at  $s_2$ , with probability 1. In this way we obtain a sequence  $s_1, s_2, \dots, s_1 \leq s_2 \leq \dots \leq b$ :  $s_n[x(t)] \in S_\alpha[x(t)]$  if  $s_n < b$ ;  $x(r)$  has a discontinuity of the first kind at  $s_n$ , if  $s_n < b$ , with probability 1. If  $\lim_{n \rightarrow \infty} s_n = s$ ,  $s \in S_\alpha[x(t)]$  if  $s < b$ , and the above argument shows that  $s$  is an isolated point of  $S_\alpha[x(t)]$  in that case, with probability 1, an impossibility, unless

$$(3.57) \quad P\left\{\lim_{n \rightarrow \infty} s_n < b\right\} = 0.$$

This however contradicts the hypothesis that the set of elements  $x(t)$  having a point  $t \in I \cdot S_\alpha[x(t)]$ ,  $t < b$ ,  $t$  a discontinuity of  $x(r)$  of the second kind, is not of zero  $P$ -measure. We have thus proved, since  $\alpha$  is arbitrary, that *the function  $x(r)$  ( $r \in R$ ) has at worst discontinuities of the first kind, with probability 1.*

(iii) If in particular  $\Omega$  is quasi-separable, there is a sequence  $R: \{r_n\}$  of points which has the property that the upper and lower limiting functions of  $x(r)$ ,  $r \in R$  are the same as those of  $x(t)$ . The above proof then shows that  $x(t)$  is continuous except for discontinuities of the first kind, with probability 1. The number of such discontinuities for a given element  $x(t)$  is at most denumerable.† This completes the proof of the second part of the theorem.

(iv) Suppose that  $\Omega = \Omega^*$ , and define  $\Omega_0$  as in the statement of the theorem. To show that  $\bar{P}^*(\Omega_0) = 1$ , it is sufficient to show that if  $\Gamma^* = \sum_1^\infty \Gamma_n^*$  is a sum of neighborhoods covering  $\Omega_0$ , then  $P^*(\Gamma^*) = 1$ . Let  $R: \{r_n\}$  be the set of  $t$ -values used in defining the neighborhoods  $\Gamma_1^*, \Gamma_2^*, \dots$ , enlarged, if necessary, to satisfy the requirements on  $R$  in (i). Let  $\Lambda_1^*$  be the  $\Omega^*$ -set of elements  $x(t)$  for which  $x(r)$  is continuous except for discontinuities of the first kind. In accordance with (ii),  $P^*(\Lambda_1^*) = 1$ . If  $\Lambda^*$  is the subset of  $\Lambda_1^*$  for which  $x(r)$  is continuous at  $r_j, j = 1, 2, \dots$ ,  $P^*(\Lambda^*) = 1$ , by (3.47). Let  $x_0(t)$  be an element of  $\Lambda^*$ . Define  $x_1(t)$  by

$$(3.58) \quad x_1(t) = \lim_{r \rightarrow t} x_0(r), \quad r \in R.$$

Then  $x_1(r_j) = x_0(r_j)$ ,  $j = 1, 2, \dots$ , and  $x_1(t) \in \Omega_0 \subset \Gamma^*$ . There is therefore an integer  $\nu$  such that  $x_1(t) \in \Gamma_\nu^*$ . Since  $\Gamma_\nu^*$  is defined by inequalities on  $x(t)$  at certain of the numbers  $r_1, r_2, \dots$ ,  $x_1(t) \in \Gamma_\nu^*$  implies that  $x_0(t) \in \Gamma_\nu^*$ . Therefore  $\Lambda^* \subset \Gamma^*$ , so

$$P^*(\Gamma^*) \geq P^*(\Lambda^*) = 1 : \bar{P}^*(\Omega_0) = 1,$$

as was to be proved.

† Hobson, *The Theory of Functions of a Real Variable*, vol. 1, 3d edition, 1927, p. 304.

Differential processes for which the distribution of  $x(t+\delta) - x(t)$  is Gaussian have been discussed by many authors,<sup>†</sup> who have shown completely or partially that this is the only case in which the elements  $x(t)$  can all be supposed continuous. Such processes can be made temporally homogeneous by a transformation of the  $t$ -axis.

**THEOREM 3.9.** *Let  $P^*$ -measure be so defined that  $\Omega^*$  is the space of a temporally homogeneous differential process for which*

$$(3.59) \quad P^*\{a < x(t+\delta) - x(t) < b\} = (2\pi\delta)^{-1/2} \int_a^b e^{-x^2/(2\delta)} dx, \quad \delta > 0.$$

*Let  $\Omega_c$  be the space of elements  $x(t)$  which are everywhere continuous. Then  $\bar{P}^*(\Omega_c) = 1$ :  $\Omega_c$  is the space of a quasi-separable measurable process.*

*Conversely if  $\Omega$  is the space of a differential stochastic process all of whose elements  $x(t)$  are everywhere continuous,  $x(t+\delta) - x(t)$  has a Gaussian distribution, for all  $t, \delta$ .*

If  $P^*$ -measure is defined by (3.59), it is readily seen that the chance variable  $\int f(t)dx(t)$  discussed above has a Gaussian distribution with mean 0 and standard deviation

$$(3.60) \quad \left\{ \int_{\Omega} \left[ \int f(t)dx(t) \right]^2 dP \right\}^{1/2},$$

more generally that the set

$$(3.61) \quad \int f_1(t)dx(t), \dots, \int f_n(t)dx(t)$$

has an  $n$ -variate Gaussian distribution, the covariance between the  $j$ th and  $k$ th members being given by

$$(3.62) \quad \int_{\Omega} \left\{ \int f_j(t)dx(t) \cdot \int f_k(t)dx(t) \right\} dP.$$

Then the orthogonal set  $\Phi_1, \Phi_2, \dots$  becomes an independent set of functions, in the usual sense of probability, and in (3.44) there is actual convergence, almost everywhere on  $\Omega^*$ : the series is a series of independent chance variables whose standard deviations squared form a convergent series.<sup>‡</sup> In particular the series in (3.45) is convergent almost everywhere on  $\Omega^*$ . Now if  $a, b$  are finite, Wiener has shown that  $\phi_1, \phi_2, \dots$  can be chosen so that

<sup>†</sup> Kolmogoroff (II, p. 868); Lévy (II, pp. 346–347); Wiener (I, Chapter IX).

<sup>‡</sup> Kolmogoroff (III, p. 314, Theorem 7).



$$\sum_1^{\infty} c_n(\tau) \Phi_n$$

converges almost everywhere on  $\Omega^*$  to a continuous function of  $\tau$ .† Since this continuous function coincides with  $x(\tau) - x(a)$  for any fixed  $\tau$  with probability 1 (in terms of  $P^*$ -measure), if  $R: \{r_n\}$  is an everywhere dense set of numbers,  $x(r)$  ( $r \in R$ ) is everywhere continuous, with probability 1. Then if  $\Omega_0$  is the space whose existence was proved in Theorem 3.7, the discontinuities (jumps) are absent, with probability 1. A slight variation of this discussion would show that  $x(t)$  is everywhere continuous with probability 1 on any quasi-separable subspace of  $\Omega^*$  which is the space of a stochastic process.

Conversely if  $x(t)$  is everywhere continuous, the Laplace-Liapounoff theorem can be invoked to show that the distribution of  $x(t+\delta) - x(t)$  is Gaussian.‡

## BIBLIOGRAPHY

L. BACHELIER.

(I) *Calcul des Probabilités*, Paris, 1912.

(II) *Le Jeu, la Chance, et le Hasard*, Paris, 1914.

(III) *Théorie de la spéculation*, Annales de l'École Normale, series 3, vol. 17 (1900), pp. 21-86.

F. P. CANTELLI.

(I) *Una teoria astratta del calcolo delle probabilità*, Giornale dell'Istituto Italiano degli Attuari, vol. 3 (1932), pp. 257-265.

J. L. DOOB.

(I) *Stochastic processes and statistics*, Proceedings of the National Academy of Sciences, vol. 20 (1934), pp. 376-379.

(II) *Probability and statistics*, these Transactions, vol. 36 (1934), pp. 759-775.

(III) *Note on probability*, Annals of Mathematics, vol. 37 (1936), pp. 363-367.

B. DE FINETTI.

(I) *Sulle funzione a incremento aleatorio*, Rendiconti della Reale Accademia Nazionale dei Lincei, Series 6, vol. 10 (1929), pp. 163-168.

E. HOFF.

(I) *On causality, statistics, and probability*, Journal of Mathematics and Physics of the Massachusetts Institute of Technology, vol. 13 (1934), pp. 51-102.

B. HOSTINSKY.

(I) *Méthodes générales du calcul des probabilités*, Mémorial des Sciences Mathématiques, No. 52.

(II) *Application du calcul des probabilités à la théorie du mouvement Brownien*, Annales de l'Institut Henri Poincaré, vol. 3 (1932), pp. 1-74.

B. JESSEN.

(I) *The theory of integration in a space of an infinite number of dimensions*, Acta Mathematica, vol. 63 (1934), pp. 249-323.

M. KAC.

(I) *Sur les fonctions indépendantes*, I, Studia Mathematica, vol. 6 (1936), pp. 46-58.

(II) (with H. Steinhaus) *Sur les fonctions indépendantes*, II. *La loi exponentielle; la divergence de séries*, *ibid.*, pp. 59-66.

† Wiener (I, pp. 148-151, and earlier papers). In the reference given here the elements in the stochastic process are complex-valued.

‡ Lévy (II, pp. 346-347).

- (III) (with H. Steinhaus) *Sur les fonctions indépendantes*, III. *Le mouvement Brownien; la loi de Maxwell*, *ibid.*, pp. 88–97.
- A. KHINTCHINE.
- (I) *Zur mathematischen Begründung der statistischen Mechanik*, *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 13 (1933), pp. 101–103.
  - (II) *Korrelationstheorie der stationäre stochastischen Prozesse*, *Mathematische Annalen*, vol. 109 (1934), pp. 604–615.
  - (III) *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, *Ergebnisse der Mathematik*, vol. 2, No. 4.
- A. KOLMOGOROFF.
- (I) *Sulla forma generale di un processo stocastico omogeneo (Un problema di Bruno de Finetti)*, *Atti della Reale Accademia Nazionale dei Lincei*, Series 6, vol. 15 (1932), pp. 805–808.
  - (II) *Ancora sulla forma generale di un processo stocastico omogeneo*, *ibid.*, pp. 866–870.
  - (III) *Über die Summen durch den Zufall bestimmter unabhängiger Grössen*, *Mathematische Annalen*, vol. 99 (1928), pp. 309–319.
  - (IV) *Über die analytische Methoden in der Wahrscheinlichkeitsrechnung*, *Mathematische Annalen*, vol. 104 (1931), pp. 415–458.
  - (V) *Grundbegriffe der Wahrscheinlichkeitsrechnung*, *Ergebnisse der Mathematik*, vol. 2, No. 3.
- PAUL LÉVY.
- (I) *Sur les séries dont les termes sont des variables éventuelles indépendantes*, *Studia Mathematica*, vol. 3 (1931), pp. 119–155.
  - (II) *Sur les intégrales dont les éléments sont des variables aléatoires indépendantes*, *Pisa Annali*, Series 2, vol. 3 (1934), pp. 337–366.
  - (III) *Observation sur un précédent mémoire de l'auteur*, *ibid.*, vol. 4 (1935), pp. 217–218.
  - (IV) *Propriétés asymptotiques des sommes de variables aléatoires enchainées*, *Bulletin des Sciences Mathématiques*, Series 2, vol. 59 (1935), pp. 84–96.
  - (V) *Sur l'application du théorème de Fubini au calcul des probabilités*, *L'Enseignement Mathématique*, vol. 33 (1934), pp. 265–270.
- Z. ŁOMNICKI.
- (I) *Nouveaux fondements du calcul des probabilités*, *Fundamenta Mathematicae*, vol. 4 (1923), pp. 34–70.
  - (II) (with S. Ulam) *Sur la théorie de la mesure dans les espaces combinatoires et son application au calcul des probabilités*, I. *Variables indépendantes*, *Fundamenta Mathematicae*, vol. 23 (1934), pp. 237–278.
- R. E. A. C. PALEY (with N. Wiener).
- (I) *Fourier Transforms in the Complex Domain*, *American Mathematical Society Colloquium Publications*, vol. XIX, 1934.
- H. STEINHAUS.
- (I) *Les probabilités dénombrables et leur rapport à la théorie de la mesure*, *Fundamenta Mathematicae*, vol. 4 (1923), pp. 286–310.
  - (II) *Sur la probabilité de la convergence de séries*, *Studia Mathematica*, vol. 2 (1930), pp. 21–39. Cf. M. Kac and H. Steinhaus.
- S. ULAM.
- Cf. Z. Łomnicki and S. Ulam.
- N. WIENER.
- Cf. R. E. A. C. Paley and N. Wiener, and earlier papers by Wiener, whose results are given in that reference.

UNIVERSITY OF ILLINOIS,  
URBANA, ILL.