

# ON CONVERGENCE OF STOCHASTIC PROCESSES

BY

JOHN LAMPERTI<sup>(1)</sup>

1. **Introduction.** The “invariance principles” of probability theory [1; 2; 5] are mathematically of the following form: a sequence of stochastic processes  $\{x_t^{(n)}(\omega)\}$  induces a sequence of measures  $\{\mu_n\}$  on some suitable topological function space  $S$ , and one proves that the measures converge weakly to a measure corresponding to a limiting process. Weak convergence of measures means here that

$$(1) \quad \lim_{n \rightarrow \infty} \int_S f d\mu_n = \int_S f d\mu$$

for all bounded continuous real-valued functions  $f$  on  $S$ . This is equivalent to the condition that

$$(2) \quad \lim_{n \rightarrow \infty} \mathcal{L}\{f(x_t^{(n)})\} = \mathcal{L}\{f(x_t)\},$$

where  $\mathcal{L}(X)$  is the distribution function of the random variable  $X$ ,  $f(\cdot)$  is a real-valued function  $S$  continuous almost everywhere ( $\mu$ ), and the limit is in the sense of the usual weak convergence of distributions. Equation (2) is usually the real center of interest, for many “limit-distribution theorems” are implicit in it.

It is clear that for given  $\{\mu_n\}$  and  $\mu$ , the better theorem of this kind would be the one in which (2) is proved for the larger class of functions  $f$ . In this paper we shall show that certain known “invariance principles” can under some hypotheses be improved by considerably enlarging the class of functions for which (2) holds. This will be done by considering spaces  $S$  other than the customary ones. For example, in studying convergence to the Wiener process, it is usual to let  $S$  be the space (denoted  $\mathcal{C}$ ) of continuous functions with the uniform topology. However, this choice does not fully exploit the pleasant properties of the Wiener path-functions, which are not only continuous but also Hölder continuous of any order up to  $1/2$ . Therefore we shall attempt to use spaces  $\text{Lip}_\alpha$  in place of  $\mathcal{C}$  as the function-space  $S$ . When weak convergence can be established using such spaces, the class of functionals for which (2) is known to hold becomes much larger than before.

To carry out the idea sketched above it is necessary to have a criterion which guarantees that the sample functions of a stochastic process are a.s.

---

Received by the editors August 12, 1961.

(<sup>1</sup>) This work was supported in part by National Science Foundation Grant Number 9669 at Stanford University.

Hölder continuous. We shall use that given by Loève in [4, p. 519]. It is shown there that if  $\{x_t\}$  is a separable process, then the condition

$$(3) \quad E\{|x_{t+\Delta t} - x_t|^\alpha\} \leq C |\Delta t|^{1+\beta}, \quad \alpha, \beta > 0,$$

which was shown by Kolmogorov to be sufficient for continuity of sample functions, is actually sufficient for Hölder continuity of any order less than  $\beta/\alpha$ . This fact suggests the task of improving a theorem of Prokhorov [5]. In this theorem it is assumed that the finite-dimensional distributions of a sequence of processes  $\{x_t^{(n)}\}$  converge to those of a process  $\{x_t\}$ , and that condition (3) holds for each  $\{x_t^{(n)}\}$  with the constants  $\alpha, \beta$  and  $C$  independent of  $n$ . The processes  $\{x_t^{(n)}\}$  and  $\{x_t\}$  then induce measures  $\mu_n$  and  $\mu$  on  $\mathcal{C}$  such that  $\mu_n \rightarrow \mu$  weakly. Here we prove that  $\mathcal{C}$  may be replaced by  $\text{Lip}_\gamma$  for any  $\gamma < \beta/\alpha$  with a corresponding enlargement of the class of functionals for which (2) holds. In the context of Donsker's original "invariance principle" [2], our result implies that if Donsker's assumptions are strengthened by requiring the existence of even order moments beyond the second, the conclusion can also be strengthened in the manner described above. An application concerning the convergence of Gaussian processes is also given.

2. **Preliminaries.** Let  $\text{Lip}_\alpha$  be the space of all real-valued functions  $x_t$  defined for  $t \in [0, 1]$  with  $x_0 = 0$  and such that

$$(4) \quad \|x_t\|_\alpha = \sup \frac{|x_{t+\Delta t} - x_t|}{|\Delta t|^\alpha} + \max |x_t| < \infty.$$

Let  $D^\alpha$  be the derivative of order  $\alpha$ , defined as  $DI^{1-\alpha}$  where  $I^{1-\alpha}$  is the Liouville fractional integration operator; let  $\mathcal{C}_\alpha$  be the space of functions  $x_t$  (again with  $x_0 = 0$ ) with continuous  $\alpha$ th derivatives and

$$(5) \quad \|x_t\|^\alpha = \max |D^\alpha x_t| + \max |x_t|.$$

The main fact we shall need about these spaces is the following:

**LEMMA 1.** *A set  $Q$  of functions which is bounded in  $\text{Lip}_\alpha$  is precompact in  $\mathcal{C}_\beta$  and in  $\text{Lip}_\beta$  for any  $\beta$  such that  $0 \leq \beta < \alpha$ .*

**Proof.** According to a theorem of Hardy and Littlewood [3], if  $x_t \in \text{Lip}_\alpha$  and  $0 \leq \beta < \alpha$  then  $D^\beta x_t \in \text{Lip}_{\alpha-\beta}$ . Furthermore,  $x_t$  can be recovered by applying the integral  $I^\beta$ , and this integral is continuous under uniform, and hence under  $\text{Lip}_{\alpha-\beta}$  convergence. The operator  $D^\beta$  mapping  $\text{Lip}_\alpha$  into  $\text{Lip}_{\alpha-\beta}$  is thus closed; by the closed graph theorem it is continuous. The set  $D^\beta x_t$ , where  $x_t$  is in the set  $Q$  which is bounded in  $\text{Lip}_\alpha$ , is therefore bounded in  $\text{Lip}_{\alpha-\beta}$ .

It follows from the above that the  $\beta$ th derivatives of the functions in the bounded set  $Q$  form an equicontinuous family, and hence a precompact one in the topology of uniform convergence. It is, however, easy to see that this in turn means that  $Q$  itself is precompact in the topology of  $\mathcal{C}_\beta$ , the first part of the lemma.

The second part follows from the first together with the fact that the identity mapping is continuous from  $\mathcal{C}_\beta$  into  $\text{Lip}_\beta$ . To see this, note that the identity is always closed, so if  $\mathcal{C}_\beta \subset \text{Lip}_\beta$  the closed graph theorem does the rest. But  $x_t \in \mathcal{C}_\beta$  means that  $I^{1-\beta}x_t \in \mathcal{C}_1 \subset \text{Lip}_1$ , and by the theorem in [3] used above, we obtain  $x_t = D^{1-\beta}(I^{1-\beta}x_t) \in \text{Lip}_\beta$ , q.e.d.

We turn next to conditions ensuring that a family of measures on  $\text{Lip}_\alpha$  is precompact. The basic fact here is a theorem due to Prokhorov [5] relating precompactness of measures *on* a space to compactness of sets *in* the space. This result states that a family  $M$  of probability measures on the Borel sets of a complete separable metric space  $S$  is weakly precompact if and only if for every  $\epsilon$  there exists a compact set  $K_\epsilon \subset S$  such that

$$(6) \quad \sup_{\mu \in M} \mu(S - K_\epsilon) \leq \epsilon.$$

Now let  $S$  be  $\text{Lip}_\alpha$  for some  $0 < \alpha < 1$ . Combining Lemma 1 with the theorem described above we have

LEMMA 2. *If for every  $\epsilon > 0$  there exists  $\eta > 0$  and  $B < \infty$  such that*

$$(7) \quad \inf_{\mu \in M} \mu \{ x_t \in \text{Lip}_\alpha : \|x_t\|_{\alpha+\eta} \leq B \} \geq 1 - \epsilon,$$

*then the family  $M$  of measures is precompact with respect to weak convergence.*

Suppose that  $\{x_t^{(n)}\}$  is a sequence of stochastic processes,  $t \in [0, 1]$ , and that the finite-dimensional distribution functions approach limits as  $n \rightarrow \infty$ . Suppose also that each process  $\{x_t^{(n)}\}$  induces a measure  $\mu_n$  on  $\text{Lip}_\alpha^{(2)}$ , and that these measures form a precompact family. Then since a measure on  $\text{Lip}_\alpha$  is uniquely determined by the measures of finite-dimensional cylinder sets, the measures  $\mu_n$  must have a weak limit. These remarks together with Lemma 2 will be applied in the next section.

**3. The strengthened "invariance principle."** The main result of this paper is (as announced in the introduction) an improved version of Theorem 2.1 of [5]:

**THEOREM.** *Let  $\{x_t^{(n)}\}$ ,  $t \in [0, 1]$ ,  $x_0 = 0$  be a sequence of separable stochastic processes satisfying condition (3) with  $\alpha$ ,  $\beta$  and  $C$  independent of  $n$ . Suppose also that the finite-dimensional distributions of  $\{x_t^{(n)}\}$  converge to limits as  $n \rightarrow \infty$ . Then there exists a process  $\{x_t\}$  whose finite-dimensional distribution are these limits, whose path-functions belong a.s. to  $\text{Lip}_\gamma$  for every  $\gamma < \beta/\alpha$ , and such that*

$$(8) \quad \lim_{n \rightarrow \infty} \mathcal{L}\{f(x_t^{(n)})\} = \mathcal{L}\{f(x_t)\}$$

---

(<sup>2</sup>) It is enough to know that the path-functions belong a.s. to  $\text{Lip}_\alpha$ , in view of the fact that the  $\sigma$ -field of sets generated by the cylinder sets plus  $\text{Lip}_\alpha$  itself contains all the Borel sets in  $\text{Lip}_\alpha$ .

for every functional  $f$  which is continuous at almost all points of  $\text{Lip}_\gamma$  for some  $\gamma < \beta/\alpha$  (with respect to the measure induced by the process  $\{x_t\}$ ).

**Proof.** Choose  $\gamma < \beta/\alpha$ ; by the result from [4] quoted in the introduction the sequence  $\{x_t^{(n)}\}$  then induces a sequence of probability measures on  $\text{Lip}_\gamma$ . Moreover, these measures concentrate all their mass on the subspace  $\text{Lip}_{\gamma+\eta}$ , provided that  $\gamma+\eta < \beta/\alpha$  too. We now appeal to the proof rather than merely the conclusion of Loève's theorem which provides a lower bound for

$$\Pr\{\|x_t^{(n)}\|_{\gamma+\eta} \leq B\}.$$

Indeed, the fact that  $x_t^{(n)} \in \text{Lip}_{\gamma+\eta}$  a.s. shows that this probability goes to one as  $B \rightarrow \infty$  for each  $n$ ; the important point now is that the method of proving this contains implicitly an estimate depending only on  $\alpha$ ,  $\beta$  and  $C$  and hence uniform in  $n$ . We can therefore apply Lemma 2 and conclude that the sequence of measures of the processes  $\{x_t^{(n)}\}$  is precompact in the weak topology of  $\text{Lip}_\gamma$ . The remarks following Lemma 2 and the equivalence of (1) and (2) complete the proof.

Suppose that  $\{\xi_i, i=1, 2, \dots\}$  is a sequence of independent random variables with mean 0 and unit variance. Let  $x_t^{(n)}$  be a continuous "random broken line" function on  $[0, 1]$  defined by

$$x_t^{(n)} = \frac{\xi_1 + \dots + \xi_i}{n^{1/2}}$$

when  $t=i/n$ ,  $i=0, 1, \dots, n$ , and by linear interpolation between these points for other  $t$ ; also let  $\{x_t\}$  be Wiener process. Then Donsker's theorem [2] is equivalent to the assertion (under the additional hypothesis that the  $\xi_i$  are identically-distributed) that (8) holds for functionals which are continuous in the uniform topology at almost all (Wiener measure) points of  $\mathcal{C}$ . By contrast, our theorem has the following.

**COROLLARY 1.** *In the above situation, if the random variables  $\xi_i$  have bounded  $2p$ th moments for some integer  $p > 1$ , then (8) holds for any functional which is continuous in the topology of  $\text{Lip}_\gamma$  a.e. (Wiener) on  $\text{Lip}_\gamma$  for some  $\gamma < (p-1)/2p$ .*

**Proof.** To apply the theorem, it is only necessary to verify (3); let us first consider intervals  $[0, i/n]$ . Then

$$E(|x_{i/n}^{(n)} - x_0^{(n)}|^{2p}) = E\left(\frac{|\xi_1 + \dots + \xi_i|^{2p}}{n^p}\right) = \frac{O(i^p)}{n^p} = O(i^p),$$

where  $O(\ )$  depends on the moments of  $\xi_i$  but not on  $n$ . This relation is easily extended to hold also for other intervals, so that (3) is satisfied with  $\alpha=2p$  and  $\beta=p-1$  and the corollary follows. Note that if all moments are finite, any  $\gamma < 1/2$  will do; this is close to the most that can be expected since

Wiener paths are not Hölder continuous of order 1/2.

It may be remarked that the moment assumption in the above result is indispensable. For example, suppose that

$$(9) \quad \Pr(|\xi_i| > x) = x^{-\delta} \text{ for large } x,$$

so that absolute moments of order less than  $\delta$  exist, but not those of greater order. Now for the processes  $\{x_i^{(n)}\}$  we have constructed,

$$(10) \quad \Pr(\|x_i^{(n)}\|_\alpha > B) \geq \Pr\left(\max_{t \leq n} |\xi_i| > Bn^{1/2-\alpha}\right).$$

The right-hand side of (10), however, if (9) holds is

$$1 - \{1 - Bn^{-(1/2-\alpha)\delta}\}_n,$$

which tends to one for all  $B$  as  $n \rightarrow \infty$  unless  $(1/2-\alpha)\delta \geq 1$ . Clearly if the probability in (10) goes to one weak convergence in the  $Lip_\alpha$  topology cannot hold, so we do not always have such convergence unless moments of order at least  $2(1-2\alpha)^{-1}$  are finite.

It is also easy to obtain from our theorem a result about convergence of Gaussian processes. Let  $\{x_i^{(n)}\}$ ,  $t \in [0, 1]$ ,  $n = 1, 2, \dots$ , be a sequence of real, separable Gaussian stochastic processes; suppose that

$$(11) \quad E[x_i^{(n)}] = \mu_n(t), \quad E[\tilde{x}_s^{(n)} \tilde{x}_t^{(n)}] = \rho_n(s, t),$$

where  $\tilde{x}_t^{(n)} = x_t^{(n)} - \mu_n(t)$ . Convergence of finite-dimensional distributions is assured by assuming that

$$(12) \quad \lim_{n \rightarrow \infty} \mu_n(t) = \mu(t), \quad \lim_{n \rightarrow \infty} \rho_n(s, t) = \rho(s, t)$$

exist for each  $s$  and  $t$  in  $[0, 1]$ . We then have

**COROLLARY 2.** *Under these conditions, suppose that there exist constants  $\xi \in (0, 2]$  and  $A, B < \infty$  such that for  $t, t + \Delta t \in [0, 1]$ ,*

$$(13) \quad |\mu_n(t + \Delta t) - \mu_n(t)| \leq A |\Delta t|^{\xi/2};$$

$$(14) \quad |\rho_n(t, t) - 2\rho_n(t, t + \Delta t) + \rho_n(t + \Delta t, t + \Delta t)| \leq B |\Delta t|^\xi.$$

*Then there is a separable Gaussian process  $\{x_t\}$  with mean function  $\mu(t)$ , covariance  $\rho(s, t)$  and paths belonging a.s. to  $Lip_\gamma$  for every  $\gamma < \xi/2$ , and (8) holds for every functional  $f$  which is continuous in the topology of  $Lip_\gamma$  a.e. ( $\{x_t\}$  measure) for some  $\gamma < \xi/2$ .*

**Proof.** To apply our theorem, it is only necessary to verify that the processes  $\{x_i^{(n)}\}$  satisfy (3), uniformly in  $n$ , for numbers  $\alpha$  and  $\beta$  such that  $\beta/\alpha$  is arbitrarily close to  $\xi/2$ . In fact, let  $\alpha = 2p$  where  $p$  is a positive integer. Then

$$(15) \quad E[(x_{t+\Delta t}^{(n)} - x_t^{(n)})^{2p}]^{1/2p} \leq E[(\tilde{x}_{t+\Delta t}^{(n)} - \tilde{x}_t^{(n)})^{2p}]^{1/2p} + [\mu_n(t+\Delta t) - \mu_n(t)]$$

by the Minkowski inequality. Now  $\tilde{x}_{t+\Delta t}^{(n)} - \tilde{x}_t^{(n)}$  is a normal random variable with mean 0 and variance given by the left-hand side of (14). It follows that the  $2p$ th moment equals  $O(|\Delta t|^{\xi p})$  uniformly in  $t$  and  $\Delta t$ . Using this and (13) in (15), we obtain

$$(16) \quad E[(x_{t+\Delta t}^{(n)} - x_t^{(n)})^{2p}] \leq C |\Delta t|^{\xi p},$$

where  $C$  is independent of  $n$ . Therefore (3) holds with  $\beta/\alpha = (\xi p - 1)/2p$ , which is close to  $\xi/2$  if  $p$  is large.

#### REFERENCES

1. P. Billingsley, *The invariance principle for dependent random variables*, Trans. Amer. Math. Soc. **83** (1956), 250–268.
2. M. Donsker, *An invariance principle for certain probability limit theorems*, Mem. Amer. Math. Soc. No. 6 (1951).
3. G. H. Hardy and J. E. Littlewood, *Some properties of fractional integrals. I*, Math. Z. **27** (1928), 565–606.
4. M. Loève, *Probability theory*, 2nd ed., Van Nostrand, Princeton, N. J., 1960.
5. Yu. V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Teor. Veroyatnost. i Primenen **1** (1956), 289–319.

STANFORD UNIVERSITY,  
STANFORD, CALIFORNIA