THEORY OF PROBABILITY AND ITS APPLICATIONS

Volume I Number 3

1956

LIMIT THEOREMS FOR STOCHASTIC PROCESSES

A. V. SKOROKHOD

1. Introduction

1.1. In recent years the limit theorems of probability theory, which previously dealt primarily with the theory of summation of independent random variables, have been extended rather widely to the theory of stochastic processes. Among the work done in this connection, we mention the articles of A. N. Kolmogorov [1, 2], M. Donsker [3], I. I. Gikhman [4], Yu. V. Prokhorov [5, 6], the author [7, 8], and N. N. Chentsov [9].

Kolmogorov, Donsker, Gikhman and the author treat various more or less special important cases of this kind of limit theorems. Yu. V. Prokhorov indicates a general approach to limit theorems for stochastic processes, basing his work on compactness criteria of measures in a complete separable metric space. It seems to me that the restriction to a complete metric space is not very natural, since in every specific case it becomes necessary to find in the trajectory space of the random process a complete metric satisfying definite conditions. This is not always possible, and even when it is, it is not always simple.

The present article suggests a new approach to limit theorems which can be used for many topological spaces in which in general no complete metric exists.

1.2. Limit theorems for stochastic processes contain primarily statements from which one can tell whether or not certain quantities $f[\xi_n(t)]$, which depend on a sequence of random processes $\xi_n(t)$, have the same distribution for $n \to \infty$ as the distribution of $f[\xi(t)]$, where $\xi(t)$ is the process which is the limit of the sequence $\xi_n(t)$ (assuming, usually, that the finite dimensional distributions of $\xi_n(t)$ converge weakly to the finite dimensional distributions of $\xi(t)$). Such limit theorems are useful for the following two reasons.

In the first place, the limit process $\xi(t)$ often has a simpler structure; moreover there often exists a good analytic apparatus for its analysis, as is true, for instance, in the case of convergence to processes with independent increments or to Markov processes, describable by differential and integro-differential equations. For these cases it is often much easier to analyze the distribution of $f[\xi(t)]$ than it is to analyze that of $f[\xi_n(t)]$ and this can be done when we have the proper limit theorems (an example is Kolmogorov's well known statistical criterion).

Secondly, a process $\xi(t)$ can often be approximated by a sequence of processes $\xi_n(t)$ whose structure is simpler. In order to analyze the properties of $\xi(t)$ in this way we must know which properties of $\xi_n(t)$ are maintained in the transition to the limit, and this can also be reduced to studying the limit distributions of $f[\xi_n(t)]$, where f is some function defined on the trajectory of the process. Although such treatments are still very rare and often not entirely rigorous (an example is Feller's [10] attempt to analyze some Markov processes by approximating them by discrete ones), analyses of stochastic processes based on approximating them by simpler ones are quite powerful and may soon be very widely used. It therefore becomes even more important to have limit theorems for stochastic processes.

1.3. We shall obtain the limit theorems in this article in the following way.

Consider a sequence of processes $\xi_n(t)$ whose finite dimensional distributions converge to the finite dimensional distributions of the process $\xi(t)$. We shall assume that with probability 1 the trajectories of both $\xi_n(t)$ and $\xi(t)$ belong to a certain set K of functions $x(t)$, and it is natural to suppose that the functions of K are determined by their values on a certain countable everywhere dense set N of values of t. It then turns out that one can construct processes $\bar{\xi}_n(t)$ and $\bar{\xi}(t)$ so that $\bar{\xi}_n(t)$ converges with probability 1 to $\bar{\xi}(t)$ for all t in N, so that $\bar{\xi}_n(t)$ and $\xi_n(t)$ as well as $\bar{\xi}(t)$ and $\xi(t)$ have the same distributions, and so that the trajectories of $\bar{\xi}_n(t)$ and $\bar{\xi}(t)$ belong to K with probability 1. Now let F be a class of functions defined on K, which are continuous in some topology S defined in K. We will find the conditions which must be satisfied by a sequence $x_n(t)$ in K which converges to $x(t)$ in K for t in N, in order that $x_n(t)$ converges to $x(t)$ in the topology S. If these conditions are fulfilled by the sequence $\bar{\xi}_n(t)$, then $\bar{\xi}_n(t)$ converges with probability 1 to $\bar{\xi}(t)$ in the topology S, and this means that for any $f \in F$ the sequence $f[\bar{\xi}_n(t)]$ converges with probability 1 to $f[\bar{\xi}(t)]$. Furthermore, the distribution of $f[\bar{\xi}_n(t)]$ will converge to the distribution of $f[\bar{\xi}(t)]$, and since the distribution of $f[\bar{\xi}_n(t)]$ coincides with that of $f[\xi_n(t)]$, while that of $f[\bar{\xi}(t)]$ coincides with that of $f[\xi(t)]$, we find that the distribution of $f[\xi_n(t)]$ converges to the distribution of $f[\xi(t)]$.

This method of obtaining the limit theorems is used when K is the set of all functions without second order discontinuities. This is perhaps the simplest class of functions which can be defined by its values on an everywhere dense set of values of t (except, of course, the set of all continuous functions; for them, however, the problem has been solved by Prokhorov [5]). In addition, there is a wide class of processes whose trajectories have, with probability 1, no second order discontinuities.

The basic difficulty which arises here is in the choice of reasonable classes of functions, that is in the choice of the topology. In the space of functions without second order discontinuities the author has already suggested a topology [7] in connection with the study of the convergence of processes with independent increments. It was found later that the applicability of this topology is wider than was thought, and that the topology is a very natural one [1]. In the present article some other topologies, occurring naturally from certain points of view, are also considered. Section 2 is devoted to a study of the convergence conditions in these topologies, convergence conditions which contain

[1] See A. N. Kolmogorov, *Probability Theory and Its Applications*, this journal, Vol. 1, No. 2 (1956).

the above mentioned subsidiary conditions must be fulfilled by a sequence of functions converging on an everywhere dense set in order to have convergence in the desired topology. Section 3 contains the limit theorems for stochastic processes.

I should like to note that although the suggested method is used only for a rather special choice of topology and a rather narrow class of processes, it can be used for other topologies and for more complicated processes such as, for instance, for stochastic functions of many variables. If this has not been done in the present article, it is only because the theory of random functions of many variables is not yet sufficiently developed for one to be able to judge which classes of functions are the most natural to consider (except, of course, continuous functions, for which the problem has been solved by Prokhorov [5]).

I take this opportunity to express my deep gratitude to E. B. Dynkin and A. N. Kolmogorov for valuable consultation during the many discussions of the results here presented.

2. The Space K_X. Topologies in this Space

2.1. Let X be a complete metric separable space with elements $x, y, \cdots$, and let $\rho(x, y)$ denote the distance between x and y. We denote by K_X the space of all functions $x(t)$ which are defined on the interval $[0, 1]$, whose values lie in X, and which at every point have a limit on the left and are continuous on the right (and on the left at $t = 1$).

Let us consider certain properties of the functions which belong to K_X. A function $x(t)$ will be said to have a discontinuity $\rho(x(t_0-0), x(t_0+0))$ at the point t_0.

2.1.1. If $x(t) \in K_X$, then for any positive ε there exists only a finite number of values of t such that the discontinuity of $x(t)$ is greater than ε. (This follows from the fact that if there exists a sequence for which $t_k \to t_0$ such that $\rho(x(t_k+0), x(t_k-0)) > \varepsilon$, then at t_0 the function $x(t)$ would have no limit either on the right or on the left.)

2.1.2. Let $t_1, t_2, \cdots, t_k$ be all the points at which $x(t)$ has discontinuities no less than ε. Then there exists a δ such that if $|t'-t''| < \delta$ and if t' and t'' both belong to the same one of the intervals $(0, t_1), (t_1, t_2), \cdots, (t_k, 1)$, then $\rho(x(t'), x(t'')) < \varepsilon$. To prove this, assume the contrary. Then there would exist sequences t_n' and t_n'' which converge to some point t_0 and belong to the same one of the intervals $(0, t_1), \cdots, (t_k, 1)$, and the sequences would have the property that $\rho(x(t_n'), x(t_n'')) \geqq \varepsilon$. Now the points t_n' and t_n'' lie on opposite sides of t_0 (otherwise $\rho(x(t_n'), x(t_n'')) \geqq \varepsilon$ would be impossible), so that $\rho(x(t_0+0), x(t_0-0)) \geqq \varepsilon$. Therefore t_0 is one of the points $t_1, t_2, \cdots, t_k$, which contradicts the statement that t_n' and t_n'' belong to the same one of the intervals $(0, t_1), (t_1, t_2), \cdots, (t_k, 1)$.

2.1.3. If $x(t) \in K_X$, then for all $\eta > 0$ there is a $\delta > 0$ such that every point $t \in [0, 1]$ satisfies one of the inequalities

$$\sup_{0<t_1-t<\delta} \rho(x(t), x(t_1)) < \eta, \qquad \sup_{0<t-t_1<\delta} \rho(x(t_1), x(t)) < \eta. \tag{2.1}$$

This assertion follows directly from 2.1.2. It may be stated in the following more compact way.

2.1.4. If $x(t) \in K_X$, then

$$\sup_{-\delta<t_1-t<0<t_2-t<\delta} \min\left[\rho(x(t_1), x(t)); \rho(x(t), x(t_2))\right] \to 0 \quad \text{for } \delta \to 0. \tag{2.2}$$

Finally, the more or less converse assertion can also be made.

2.1.5. If $x(t)$ satisfies (2.2) then, there exists a $\bar{x}(t) \in K_X$ which coincides with $x(t)$ at all of its points of continuity.

We shall show that $x(t)$ has, at every point, a limit from the left and from the right. If this were not so, there would exist an $\eta > 0$ and three points $t_1 < t_2 < t_3$ arbitrarily close to each other such that $\rho(x(t_1), x(t_2)) > \eta$ and $\rho(x(t_2), x(t_3)) > \eta$. But this contradicts (2.2). We note also that $x(t)$ must definitely coincide either with $x(t-0)$ or $x(t+0)$. Setting $\bar{x}(1) = \lim_{t\to 1} x(t)$ and $\bar{x}(t) = \lim\limits_{\substack{t'\to t\\ t'>t}} x(t'')$, we obtain the desired function.

2.2. We shall now go on to define the topologies in the space K_X. A general analysis of this problem (that is, of classifying the possible topologies and studying their properties) is probably of independent interest. Here we shall discuss the five topologies which are simplest from our point of view, and to which we are led by considering functionals of trajectories of processes usually studied in probability theory. In addition, these are natural generalizations of the uniform topology on the space of continuous functions, generalizations in harmony with the whole development of mathematics.

2.2.1. Definition. *The sequence of functions $x_n(t)$ converges uniformly to $x(t)$ at the point t_0 if for all $\varepsilon > 0$ there exists a δ such that*

$$\overline{\lim_{n\to\infty}} \sup_{|t-t_0|<\delta} \rho(x_n(t), x(t)) < \varepsilon. \tag{2.3}$$

Obviously if $x_n(t)$ converges uniformly to $x(t)$ at every point of some closed set, then $x_n(t)$ converges uniformly to $x(t)$ on this whole set.

A necessary condition for convergence in all the topologies considered here will be that the sequence $x_n(t)$ converges to $x(t)$ uniformly at every point of continuity of $x(t)$.

Thus for continuous functions, all topologies reduce to ordinary uniform convergence. Our topologies begin to differ when we consider the behavior of $x_n(t)$ in the neighborhood of points of discontinuity of $x(t)$. The simplest behavior we might suggest is that the convergence be uniform also at points of discontinuity, in which case we obtain ordinary uniform convergence (in our notation we shall call this topology U).

Although we shall consider this topology in addition to others, we remark that it is not the most natural one for K_X. This is because the uniform topology in K_X requires that the convergence of $x_n(t)$ to $x(t)$ imply that there exists a number such that for all n greater than or equal to this number the points of discontinuity of $x_n(t)$ coincide with the points of discontinuity of $x(t)$. This means that if t is considered to be the time, we must assume the existence of an instrument which will measure time exactly, and physically this is an impossibility. It is much more natural to suppose that the functions we can obtain from each other by small deformation of the times scale lie close to each other. We thus are led to propose the topology $\mathbf{J}_1$.

2.2.2. DEFINITION. *The sequence $x_n(t)$ is called $\mathbf{J}_1$-convergent to $x(t)$ if there exists a sequence of continuous one-to-one mappings $\lambda_n(t)$ of the interval $[0, 1]$ onto itself, such that*

$$\lim_{n\to\infty} \sup_t \rho(x_n(t), x(\lambda_n(t))) = 0; \quad \lim_{n\to\infty} \sup_t |\lambda_n(t)-t| = 0. \tag{2.4}$$

Let us analyze in more detail how we define the behavior of $x_n(t)$ in the neighborhood of a discontinuity point t_0 if we require that $x_n(t)$ converges to $x(t)$ in one of our topologies. Since $x(t)$ has a limit from the left and the right at every point t, we know that in a sufficiently small neighborhood of t_0 it will be approximately equal to $x(t_0-0)$ for $t < t_0$ and to $x(t_0+0)$ for $t \geqq t_0$. Therefore on the left of t_0 there exist intervals where $x_n(t)$ is approximately equal to $x(t_0-0)$, and on the right there are those in which $x_n(t)$ is approximately equal to $x(t_0+0)$ (since the points of continuity of $x(t)$ are everywhere dense, and at these points the convergence is uniform). Thus the question of how $x_n(t)$ behaves in the neighborhood of t_0 is equivalent to the question how $x_n(t)$ changes from $x(t_0-0)$ to $x(t_0+0)$. In the topologies U and $\mathbf{J}_1$ this transition takes the form of a single jump. In both of these topologies, for values of t close to t_0 the function $x_n(t)$ can take on values which are either close to $x(t_0-0)$ or to $x(t_0+0)$. If we wish to keep this last property, but do not require that the transition be in the form of a single jump, that is that a function $x_n(t)$ may change back and forth between the values $x(t_0-0)$ and $x(t_0+0)$ several times in the neighborhood of a point t_0, then we obtain topology $\mathbf{J}_2$.

2.2.3. DEFINITION. *A sequence $x_n(t)$ is said to be $\mathbf{J}_2$-convergent to $x(t)$ if there exists a sequence of one-to-one mappings $\lambda_n(t)$ of the interval $[0, 1]$ onto itself such that*

$$\sup_t |\lambda_n(t)-t| \to 0 \text{ and } \sup_t \rho(x_n(t), x(\lambda_n(t))) \to 0 \text{ as } n \to \infty. \tag{2.5}$$

We shall define the two following topologies for the case in which X is a complete linear normalized space. In this case, together with every function belonging to K_X we shall consider its graph $\Gamma_{x(t)}$, constructed in the following way.

Let $X \times T$ be the space composed of the pairs (x, t), where $x \in X$ and $t \in [0, 1]$. In this space we define the metric R by

$$R[(x_1, t_1); (x_2, t_2)] = \rho(x_1, x_2)+|t_1-t_2|. \tag{2.6}$$

We then define the graph $\Gamma_{x(t)}$ as the closed set in $X \times T$ which contains all pairs (x, t), such that for all t the point x belongs to the segment joining $x(t-0)$ and $x(x+0)$ (henceforth we shall denote the segment joining x_1 and x_2 by $[x_1, x_2]$). The graph $\Gamma_{x(t)}$ is a continuous curve in $X \times T$, and therefore we can define convergence for graphs in the same way as we usually do for space curves. We then obtain the topology M_1.

2.2.4. DEFINITION. *The pair of functions $(y(s), t(s))$ gives a parametric representation of the graph $\Gamma_{(x,t)}$ if those and only those pairs (x, t) belong to it for which an s can be found such that $x = y(s)$, $t = t(s)$, where $y(s)$ is continuous, and $t(s)$ is continuous and monotonically increasing (the functions $y(s)$ and $t(s)$ are defined on the segment $[0, 1]$).*

We note that if $(y_1(s), t_1(s))$ and $(y_2(s), t_2(s))$ are parametric representations of $\Gamma_{x(t)}$, there exists a monotonically increasing function $\lambda(s)$ such that $y_1(s) = y_2(\lambda(s))$ and $t_1(s) = t_2(\lambda(s))$.

2.2.5. Definition. *The sequence $x_n(t)$ is called M_1-convergent to $x(t)$ if there exist parametric representations $(y(s), t(s))$ of $\Gamma_{x(t)}$ and $(y_n(s), t_n(s))$ of $\Gamma_{x_n(t)}$ such that*

$$\lim_{n\to\infty} \sup_s R[(y_n(s), t_n(s)); (y(s), t(s))] = 0. \tag{2.7}$$

We can characterize the topology M_1 in the following way from the point of view of the behavior at a point of discontinuity t_0 of the function $x(t)$. The transition from $x(t_0-0)$ to $x(t_0+0)$ is such that first $x_n(t)$ is arbitrarily close to the segment $[x(t_0-0), x(t_0)]$ and second that $x_n(t)$ moves from $x(t_0-0)$ to $x(t_0)$ almost always advancing. More rigorously, what this last condition means is that if $x(t_0-0) < x_1 < x_2 < x(t_0)$ (using the natural ordering for the segment), then having reached the neighborhood of x_2, we will not return to the neighborhood of x_1.

If we drop the second requirement and assume only that $x_n(t)$ goes from $x(t_0-0)$ to $x(t_0+0)$ remaining always arbitrarily close to the segment $[x(t_0-0), x(t_0)]$, we obtain the topology M_2.

2.2.6. Definition. *The sequence $x_n(t)$ is called M_2-convergent to $x(t)$ if*

$$\lim_{n\to\infty} \sup_{(y_1,t_1)\in\Gamma_{x(t)}} \inf_{(y_2,t_2)\in\Gamma_{x_n(t)}} R[(y_1, t_1); (y_2, t_2)] = 0. \tag{2.8}$$

Let S be any one of our topologies. We shall denote convergence in the topology S by the symbol $x_n(t) \xrightarrow{S} x(t)$. Let us consider the relation between our topologies. It is clear that U is stronger than $\mathbf{J}_1$, and that this in turn is stronger than $\mathbf{J}_2$. It is also clear that M_1 is stronger than M_2. We recall that a topology S_1 is stronger than S_2 if convergence in S_1 implies convergence in S_2. If X is a linear space, we can use any of our topologies. It is easily seen that convergence in M_2 follows from convergence in any of the other topologies, and that convergence in $\mathbf{J}_1$ implies convergence in any of the other topologies except U. Examples can be used to show that the topologies M_1 and $\mathbf{J}_2$ cannot be compared. (Indeed, let X be the real line. We set

$$x(t) = \begin{cases} 0; & t < \frac{1}{2}, \\ 1; & t \geqq \frac{1}{2}, \end{cases}$$

$$x_n'(t) = \begin{cases} 0; & t < \frac{1}{2} - 1/n, \\ n(t-\frac{1}{2}) + \frac{1}{2}; & \frac{1}{2} - 1/n \leqq t < \frac{1}{2} + 1/n, \\ 1; & t > \frac{1}{2} + 1/n, \end{cases}$$

$$x_n''(t) = \begin{cases} 0; & t < \frac{1}{2} - 1/n, \\ 1; & \frac{1}{2} - 1/n \leqq t < \frac{1}{2}, \\ 0; & \frac{1}{2} \leqq t < \frac{1}{2} + 1/n, \\ 1; & t \geqq \frac{1}{2} + 1/n, \end{cases}$$

so that $x_n'(t) \xrightarrow{M_1} x(t)$ although this does not converge to $x(t)$ in $\mathbf{J}_2$, while $x_n''(t) \xrightarrow{J_1} x(t)$ although this does not converge to $x(t)$ in M_1). Denoting the statement "the topology S_1 is stronger than S_2" by $S_1 \to S_2$ all the above can be summarized by

$$U \to \mathbf{J}_1 \begin{matrix} \nearrow M_1 \searrow \\ \searrow \mathbf{J}_2 \nearrow \end{matrix} M_2. \tag{2.9}$$

Let us also mention the relation between our topologies and functionals ordinarily used in probability theory. We shall consider X to be the real line. Consider the following functionals (§2.2.7—§2.2.9):

2.2.7. The upper and lower bound of a function on some segment $[t_1, t_2]$, namely

$$M_{[t_1,t_2]}[x(t)] = \sup_{t_1 \leqq t \leqq t_2} x(t); \quad m_{[t_1,t_2]}[x(t)] = \inf_{t_1 \leqq t \leqq t_2} x(t).$$

2.2.8. The number of intersections $\nu^{[a,b]}_{[t_1,t_2]}[x(t)]$ with the strip $[a, b]$ on the segment $[t_1, t_2]$ is equal to k if it is possible to find a $(k+1)$-point $t^{(0)} < t^{(1)} < \cdots < t^{(k)}$ on $[t_1, t_2]$ such that either $x(t^{(0)}) \leqq a$, $x(t^{(1)}) \geqq b$, $x(t^{(2)}) \leqq a, \cdots$, or $x(t^{(0)}) \geqq b$, $x(t^{(1)}) \leqq a$, $x(t^{(2)}) \geqq b, \cdots$, and it is impossible to give a $(k+2)$-point with these properties.

2.2.9. The first value $\gamma^+_{[t_1,t_2],a}[x(t)]$ above some level a on the segment $[t_1, t_2]$. Let us first define $\gamma^+_{[0,1],a}[x(t)]$. Let $\tau_a[x(t)] = \inf_{x(t) \geqq a} t$, if $\sup x(t) \geqq a$ and $\tau_a[x(t)] = 1$, if $\sup x(t) < a$. Then

$$\gamma^+_{0,1],a}[x(t)] = \begin{cases} x(\tau_a[x(t)]) - a & \text{for } \tau_a[x(t)] < 1, \\ -1 & \text{for } \tau_a[x(t)] = 1. \end{cases}$$

Further, we set

$$\gamma^+_{[t_1,t_2],a}[x(t)] = \gamma^+_{[0,1],a}[\bar{x}(t)], \quad \text{where } \bar{x}(t) = \begin{cases} x(t_1+0), & t < t_1, \\ x(t), & t_1 \leqq t \leqq t_2, \\ x(t_2-0), & t > t_2. \end{cases}$$

The following assertions are easily proved, but we shall not present the proofs here.

2.2.10. A necessary and sufficient condition for $x_n(t) \xrightarrow{M_2} x(t)$ is that

$$M_{[t_1,t_2]}[x_n(t)] \to M_{[t_1,t_2]}[x(t)] \text{ and } m_{[t_1,t_2]}[x_n(t)] \to m_{[t_1,t_2]}[x(t)],$$

where t_1 and t_2 are points of continuity of $x(t)$.

2.2.11. A necessary and sufficient condition for $x_n(t) \xrightarrow{M_1} x(t)$ is that

$$\nu^{[a,b]}_{[t_1,t_2]}[x_n(t)] \to \nu^{[a,b]}_{[t_1,t_2]}[x(t)]$$

for all t_1 and t_2 which are points of continuity of $x(t)$ and almost all a and b.

2.2.12. A necessary and sufficient condition for $x_n(t) \xrightarrow{\mathbf{J}_1} x(t)$ is that

$$\gamma^+_{[t_1,t_2],a}[x_n(t)] \to \gamma^+_{[t_1,t_2],a}[x(t)]$$

for almost all a and for all t_1 and t_2 which are points of continuity of $x(t)$.

2.2.13. A necessary and sufficient condition for $x_n(t) \xrightarrow{\mathbf{J}_2} x(t)$ is that

$$\gamma^+_{[t_1,t_2],a}[x_n(t)] \to \gamma^+_{[t_1,t_2],a}[x(t)] \text{ and } \nu^{[a,b]}_{[t_1,t_2]}[x_n(t)] \to \nu^{[a,b]}_{[t_1,t_2]}[x(t)]$$

for all t_1 and t_2 which are points of continuity of $x(t)$ and for almost all a and b.

Let us now go on to a more detailed analysis of our topologies, starting with the weaker ones and going on to the stronger.

2.3. The topology M_2.

We shall first find a condition for M_2-convergence which is equivalent to the definition but which is given not in terms of graphs but of the functions $x_n(t)$ and $x(t)$ themselves.

2.3.1. *A necessary and sufficient condition for* $x_n(t) \xrightarrow{M_2} x(t)$ *is that*

$$\lim_{c \to 0} \overline{\lim_{n \to \infty}} \Delta_{M_2}(c; x_n(t), x(t)) = 0, \tag{2.10}$$

where

$$\Delta_{M_2}(c; y(t), x(t)) = \sup H(x(t), y(t_c), x(t_c^*)),$$
$$t_c = \max [0, t-c], \quad t_c^* = \min [1, t+c],$$

and $H(x_1, x_2, x_3)$ *is the distance of* x_2 *from the segment* $[x_1, x_3]$.

PROOF. NECESSITY. Let $x_n(t) \xrightarrow{M_2} x(t)$. We shall show that this implies that by choosing c sufficiently small we can make $\overline{\lim}_{n\to\infty} \Delta_{M_2}(c; x_n(t), x(t))$ arbitrarily small. Let c be such that if $t_1 < t_2 < t_3$, $t_2 - t_1 < 2c$, and $t_3 - t_2 < 2c$, then $\min [\rho(x(t_1), x(t_2)); \rho(x(t_2), x(t_3))] < \varepsilon$. Let us now denote by $t_1, t_2, \cdots, t_k$ all the points at which the discontinuities are greater than or equal to ε, and let c be so small that $\rho(x(t'), x(t'')) < \varepsilon$ so long as $|t'-t''| < c$ and t' and t'' both belong to the same one of the intervals $(0, t_1), (t_1, t_2), \cdots, (t_k, 1)$. We choose n so large that

$$\sup_{(y_1, t_1) \in \Gamma_{x(t)}} \inf_{(y_2, t_2) \in \Gamma_{x_n(t)}} R[(y_1, t_1); (y_2, t_2)] < c/2. \tag{2.11}$$

Now if the points t_c, t, t_c^* all belong to the same one of the intervals $(0, t_1)$, $(t_1, t_2), \cdots, (t_k, 1)$, then

$$H(x(t_c), x_n(t), x(t_c^*)) \leqq c/2 + 2\varepsilon.$$

For, it follows from (2.11) that the graph $\Gamma_{x(t)}$ contains a point $(\bar{x}, \bar{t})$ such that $|\bar{t}-t| < c/2$ and $\rho(\bar{x}, x_n(t)) < c/2$, and hence if $\bar{t} < t$, then

$$H(x(t_0), x_n(t), x(t_c^*)) < \rho(x(t_c), x(\bar{t}-0)) + \rho(x(\bar{t}-0), \bar{x}) + \rho(\bar{x}, x_n(t)),$$

while if $\bar{t} > t$, then

$$H(x(t_c), x_n(t), x(t_c^*)) < \rho(x(t_c^*), x(\bar{t}+0)) + \rho(x(\bar{t}+0), \bar{x}) + \rho(\bar{x}, x_n(t)).$$

Let us now find $H(x(t_c), x_n(t), x(t_c^*))$ when $t_c < t_i < t < t_c^*$ (the case $t_c < t < t_i < t_c^*$ is entirely symmetric). It follows from (2.11) that $\Gamma_{x(t)}$ contains a point $(\tilde{x}, \tilde{t})$ such that $\rho(\tilde{x}, x_n(t)) < c/2$ and $|\tilde{t}-t| < c/2$. If $\tilde{t} \neq t$, then either $t_i > \tilde{t}$ and

$$H(x(t_c), x_n(t), x_n(t_c^*)) \leqq \rho(x(t_c), x(\tilde{t}-0)) + \rho(x(\tilde{t}-0), \tilde{x}) + \rho(\tilde{x}, x_n(t)) \leqq 3\varepsilon + c/2,$$

or $t_i < \tilde{t}$ and

$$H(x(t_c), x_n(t), x(t_c^*)) \leqq \rho(x(t_c^*), x(\tilde{t}+0)) + \rho(x(\tilde{t}+0), \tilde{x}) + \rho(\tilde{x}, x_n(t)) \leqq 3\varepsilon + c/2.$$

If, however, $t_i = \tilde{t}$, then

$$H(x(t_c), x_n(t), x(t_c^*)) \leqq H(x(t_c), \tilde{x}, x(t_c^*)) + \rho(\tilde{x}, x_n(t))$$
$$\leqq \rho(x(t_c), x(t_i-0)) + \rho(x(t_c^*), x(t_i+0)) + H(x(t_i-0), \tilde{x}, x(t_i+0)) + \rho(\tilde{x}, x_n(t))$$
$$\leqq 2\varepsilon + c/2,$$

since $H(x(t_i-0), \tilde{x}, x(t_i+0)) = 0$. This proves the necessity.

SUFFICIENCY. Assume that (2.10) is fulfilled. This means that there exists a c sufficiently small so that for n sufficiently large we have

$$\Delta_{M_2}(c; x_n(t), x(t)) < \varepsilon, \tag{2.12}$$

and in addition that c can be chosen so small that if t'' and t' belong to the same one of the intervals

$$(0, t_1),\ (t_1, t_2),\ \cdots,\ (t_k, 1);\ |t_2''-t'| < 2c, \text{ then } \rho(x(t''), x(t')) < \varepsilon$$

and $2c$ is less than the length of any of these intervals. Let us show that for every point of $\Gamma_{x_n(t)}$ there is a point of $\Gamma_{x(t)}$ such that in the metric R the distance between these points is less than δ, and that this last quantity can be made arbitrarily small by appropriate choice of ε and c.

From the fact that $H(x(t_c), x_n(t), x(t_c^*)) < \varepsilon$, if t_c and t_c^* belong to the same one of the intervals $(0, t_1)$, (t_1, t_2), $\cdots$, $(t_k, 1)$, it follows that $\rho(x_n(t), x(t_c)) < 3\varepsilon$, which means that

$$\rho(x_n(t), x(t)) < 4\varepsilon. \tag{2.13}$$

In this case, therefore, for a point (y, t) of $\Gamma_{x_n(t)}$ we can find a point $(\tilde{x}, t)$ of $\Gamma_{x(t)}$ (with the same t) such that $\rho(\tilde{x}, y) < 4\varepsilon$.

Now assume that $t_c < t_i < t < t_c^*$. Then

$$H(x(t_c), x_n(t), x(t_c^*)) < \varepsilon \text{ and } H(x(t_c), x_n(t-0), x(t_c^*)) < \varepsilon.$$

For every point of the segment $[x_n(t-0), x_n(t)]$, therefore, there is a point x' on $[x(t_c), x(t_c^*)]$ such that $\rho(x', x) < 2\varepsilon$. Now since $\rho(x(t_c), x(t_i-0)) < \varepsilon$ and $\rho(x(t_c^*), x(t_i+0)) < 2\varepsilon$, it follows that on the segment $[x(t_i-0), x(t_i+0)]$ there is a point $\tilde{x}$ such that $\rho(x', \tilde{x}) < 2\varepsilon$.

Thus $R[(y, t), (\tilde{x}, t_i)] < 4\varepsilon+c$.

This proves the sufficiency.

2.3.2. Remark. We may consider, instead of $\bar{H}(x_1, x_2, x_3)$, any other quantity such that

$$\varphi_1(H(x_1, x_2, x_3)) \leqq \bar{H}(x_1, x_2, x_3) \leqq \varphi_2(H(x_1, x_2, x_3)),$$

where $\varphi_i(t)$ are nonnegative continuous functions which vanish only for $t = 0$. If X is a Hilbert space, we may set

$$\bar{H}(x_1, x_2, x_3) = \rho(x_1, x_2)+\rho(x_2, x_3)-\rho(x_2, x_3).$$

2.3.3. General conditions for M_2-convergence. *Necessary and sufficient conditions for $x_n(t) \xrightarrow{M_2} x(t)$ are that*

(a) *$x_n(t)$ converges to $x(t)$ on a set of values of t which contains* 0 *and* 1 *and which is everywhere dense on* $[0, 1]$ *and*

(b) $\lim_{c\to 0} \overline{\lim}_{n\to\infty} \Delta_{M_2}[c; x_n(t)] = 0$, *where*

$$\Delta_{M_2}[c; x_n(t)] = \sup_{\substack{t\in[0,1];\, t_1\in[t_c,\, t_c+c/2] \\ t_2\in[t_c^*-c/2,\, t_c^*]}} H(x(t_1), x(t), x(t_2)).$$

Proof. The necessity of (a) is implied by the fact that if $x_n(t) \xrightarrow{M_2} x(t)$, then

$$x_n(t) \to x(t)$$

at every point of continuity of $x(t)$.

Necessity of (b).

$$H(x_n(t_1), x_n(t), x_n(t_2)) \leqq H(\tilde{x}_1, x_n(t), \tilde{x}_2)+\rho(\tilde{x}_1, x_n(t_1))+\rho(\tilde{x}_2, x_n(t_2)).$$

Since $x_n(t) \xrightarrow{M_2} x(t)$, there exist points t_1' and t_2' such that $|t_1'-t_1|$, $|t_2'-t_2|$, $\rho(\tilde{x}_1, x_n(t_1))$, and $\rho(\tilde{x}_2, x_n(t_2))$ are small, where the points $\tilde{x}_1$ and $\tilde{x}_2$ belong to the segments $[x(t_1'-0), x(t_1')]$ and $[x(t_2'-0), x(t_2')]$, respectively. Now let c be such that $x(t)$ cannot have two discontinuities equal to or greater than ε in a distance $3c$, and such that if $\bar{t}_2-\bar{t}_1 \leqq 2c$ and if there is no discontinuity between these points which is equal to or greater than ε, then $\rho(x(\bar{t}_2), x(\bar{t}_1)) < \varepsilon$. Let us evaluate $H(\tilde{x}_1, x_n(t), \tilde{x}_2)$ assuming that $|t_1'-t_1| < c/4$ and $|t_2'-t_2| < c/4$. If in (t_1', t_2') and (t_c, t_c^*) there are no discontinuities of $x(t)$ equal to or greater than ε, then

$$H(\tilde{x}_1, x_n(t), \tilde{x}_2) \leqq H(x(t_c), x_n(t), x(t_c^*))+4\varepsilon.$$

If, however, one of the intervals (t_1', t_2') or (t_c, t_c^*) has a discontinuity equal to or greater than ε at some point $\bar{t}$, we consider the following two possibilities.

1) $t_1' < \bar{t} < t_2'$; $t_c < \bar{t} < t_c^*$. Then $\rho(x(t_c), \tilde{x}_1) \leqq 2\varepsilon$ and $\rho(x(t_c^*), \tilde{x}_2) \leqq 2\varepsilon$, so that

$$H(x_1, x_n(t), \tilde{x}_2) \leqq 4\varepsilon+H(x(t_c), x_n(t), x(t_c^*)).$$

2) $\bar{t}$ lies between t_1' and t_c (and may coincide with t_1'). Then in the interval $(t_{c/4}, t_{c/4}^*)$ all the discontinuities of $x(t)$ are less than ε, so that $\rho(x(t_{c/4}), x(t_{c/4}^*)) < \varepsilon$. This means that $\rho(x(t_{c/4}^*), x_n(t)) < \varepsilon+H(x(t_{c/4}), x_n(t), x(t_{c/4}^*))$. In addition, since $\rho(x(t_{c/4}^*), x(t_c^*)) < \varepsilon$, and the discontinuity in $x(t)$ at t_2' is also less than ε, we have

$$\rho(x_n(t), \tilde{x}_2) < 3\varepsilon+H(x(t_{c/4}), x_n(t), x(t_{c/4}^*)).$$

Therefore

$$H(\tilde{x}_1, x_n(t), \tilde{x}_2) < 3\varepsilon+H(x(t_{c/4}), x_n(t), x(t_{c/4}^*)).$$

This proves the necessity.

Sufficiency. We must show that if conditions (a) and (b) are fulfilled, then

$$\lim_{c\to 0} \varlimsup_{n\to\infty} \Delta_{M_2}(c; x_n(t), x(t)) = 0.$$

Let us assume, on the contrary, that there exist sequences $c_k \to 0$, $n_j^k \to \infty$ (where the arrow indicates that we are considering the limit as $j \to \infty$) and an η such that

$$\Delta_{M_2}(c_k; x_{n_j^k}(t), x(t)) > \eta.$$

This means that there exist points $t_1^{n_j^k}$, $t_2^{n_j^k}$, $t_3^{n_j^k}$ such that $t_3^{n_j^k}-t_2^{n_j^k} = t_2^{n_j^k}-t_1^{n_j^k} = c_k$ (since the $t_i^{n_j^k}$ lie, for sufficiently small c_k, within the segment $[0, 1]$ because (a) and (b) imply that $x_n(t)$ converges uniformly to $x(t)$ at 0 and 1) and

$$H(x(t_1^{n_j^k}), x_{n_j^k}(t_2^{n_j^k}), x(t_3^{n_j^k})) > \eta.$$

Condition (a) implies that for all ε and δ there is an n_0 such that in any interval of length δ there exists a point t^* for which $\rho(x_n(t^*), x(t^*)) \leqq \varepsilon$ when $n > n_0$. Let δ_1 be so small that

$$\sup_{-\delta_1<t_1-t_2<0<t_3-t_2<\delta_1} \min [\rho(x(t_1), x(t_2)); \rho(x(t_2), x(t_3))] < \varepsilon. \tag{2.14}$$

Further, (b) implies that there exist $\bar{c}_1$, $\bar{c}_2$, and n' such that for $n > n'$ we have

$$\begin{gathered} \Delta_{M_2}(\bar{c}, x_n(t)) < \varepsilon; \quad \Delta_{M_2}(\bar{c}_2, x_n(t)) < \varepsilon; \\ \bar{c}_1 > 2\bar{c}_2; \quad \bar{c}_1 > c_k > \bar{c}_2; \quad 2\bar{c}_1 < \delta_1. \end{gathered} \tag{2.15}$$

We then have either

$$\rho x\big((t_1^{n_j^k}), x(t)\big) < 2\varepsilon, \qquad t \in (t_2^{n_j^k}-\bar c_1,\, t_2^{n_j^k}-\bar c_1/2),$$
$$\rho\big(x(t_3^{n_j^k}), x(t)\big) < 2\varepsilon, \qquad t \in (t_2^{n_j^k}+\bar c_1/2,\, t_2^{n_j^k}+\bar c_1),$$

or

$$\rho\big(x(t_1^{n_j^k}), x(t)\big) < 2\varepsilon, \qquad t \in (t_2^{n_j^k}-\bar c_2,\, t_2^{n_j^k}-\bar c_2/2),$$
$$\rho\big(x(t_3^{n_j^k}), x(t)\big) < 2\varepsilon, \qquad t \in (t_2^{n_j^k}+\bar c_2/2,\, t_2^{n_j^k}+\bar c_2)$$

(this follows from (2.14) and from the fact that the four intervals do not intersect in pairs). If we now choose $\delta < \bar c_2/2$, then for $n_j^k > \max(n_0, n')$ there exist three points $t_1, t_2^{n_j^k}, t_3'$ such that

$$H\big(x_{n_j^k}(t_1'), x_{n_j^k}(t_2^{n_j^k}), x_{n_j^k}(t_3')\big) > \eta - 12\varepsilon,$$

and either $\bar c_1/2 \leqq |t_2^{n_j^k} - t_j'| \leqq \bar c_1$ or $\bar c_2/2 \leqq |t_2^{n_j^k} - t_1'| < \bar c_2$, $i = 1, 3$. Thus either $\Delta_{M_2}(\bar c_1, x_{n_i^k}(t)) > \eta - 12\varepsilon$ or $\Delta_{M_2}(\bar c_2, x_{n_j^k}(t)) > \eta - 12\varepsilon$, which contradicts (2.15) if we take $13\varepsilon < \eta$.

We note that in proving the sufficiency we have proved the following stronger assertion.

2.3.4. *A necessary and sufficient condition for* $x_n(t) \xrightarrow{M_2} x(t)$ *is that* (a) *be fulfilled and that*

$$\text{(b}'\text{)} \quad \lim_{c\to 0} \varlimsup_{n\to\infty} \Delta_{M_2}(c, x_n(t)) = 0.$$

Thus we can obtain various sufficient conditions by making different choices of sequences $c_k \to 0$.

In studying the compactness of sets of functions and the topologies $\mathbf{J}_1$, $\mathbf{J}_2$, M_1, and M_2, the following lemma plays an important role.

2.3.5. *Consider a sequence* $x_n(t)$ *which converges to any function* $x(t)$ *for* $t \in N$ *(some set, everywhere dense on* $[0, 1]$, *containing* 0 *and* 1), *and let*

$$\lim_{c\to 0} \varlimsup_{n\to\infty} \Big(\Delta_{M_2}(c, x_n(t)) + \sup_{0<t<c} \rho(x_n(t), x_n(0)) + \sup_{1-c<t<1} \rho(x_n(1), x_n(t))\Big) = 0.$$

Then there exists $\bar x(t) \in K_X$ *such that* $x_n(t) \xrightarrow{M_2} \bar x(t)$.

Proof. Let t_i be the points of N. Let us prove the existence of

$$\lim_{t_i\to t,\, t_i>t} x(t_i) \quad \text{and} \quad \lim_{t_i\to t,\, t_i<t} x(t_i).$$

It is sufficient to prove that one of these limits, say the limit from the right, exists. First let $t = 0$. Let us assume that $\lim_{t_i\to 0} x(t_i)$ does not exist, which means that there is a sequence t_{i_k} and an ε such that $\rho(x(t_{i_k}), x(t_{i_{k+1}})) > \varepsilon$.

Let us choose c so small that $\varlimsup_{n\to\infty} \Delta_{M_2}(c, x_n(t)) < \varepsilon/4$, and k such that t_{i_k} and $t_{i_{k+1}}$ are both less than $c/2$. Then, if t_r is a point such that

$$t_{i_k}+c/2 < t_r < t_{i_k}+c \quad \text{and} \quad t_{i_{k+1}}+c/2 < t_r < t_{i_{k+1}}+c,$$

we have

$$H\big(x(t_{i_k}), x(t_{i_{k+1}}), x(t_r)\big) < \varepsilon/4 \quad \text{and} \quad H\big(x(t_{i_{k+1}}), x(t_{i_k}), x(t_r)\big) < \varepsilon/4.$$

But this contradicts the fact that $\rho(x(t_{i_k}), x(t_{i_{k+1}})) > \varepsilon$. Let us now assume that the limit $\lim_{t_i\to t,\, t_i>t} x(t_i)$ does not exist, where t is an internal point of $[0, 1]$. We then have the following two possibilities.

1) There exists an ε such that in any neighborhood of a point t we can find three points $t_{r_1} < t_{r_2} < t_{r_3}$ lying to the right of t for which

$$\rho(x(t_{r_i}), x(t_{r_j})) > \varepsilon \qquad (i \neq j;\ i, j = 1,\ 2,\ 3).$$

2) There exist $\bar{x}_1$ and $\bar{x}_2$ such that every sequence $x(t_{n_k})$ $(t_{n_k} > t,\ t_{n_k} \to t)$ can be separated into two sequences $x(t_{n'_k})$ and $x(t_{n''_k})$ (the set $\{n_k\}$ is the union of $\{n'_k\}$ and $\{n''_k\}$) such that

$$\lim_{t_{n'_k} \to t} x(t_{n'_k}) = \bar{x}_1 \text{ and } \lim_{t_{n''_k} \to t} x(t_{n''_k}) = \bar{x}_2.$$

We shall show that both of these possibilities contradict the conditions of the lemma.

The first case. Choose c_0 so that $\overline{\lim}_{n\to\infty} \Delta_{M_2}(c, x_n(t)) < \mu$ if $c < c_0$ (we shall choose μ later). Let t' and t'' be points of N such that if $0 < \bar{t}-t < c_0/2$, then, $\bar{t}-c_0 < t' < \bar{t} < c_0/2$ and $\bar{t}+c_0/2 < t'' < \bar{t}+c_0$. Then, for $t < t_i < \bar{t}$, $x(t_i)$ will be at a distance equal to or greater than μ from the segment $[x(t'), x(t'')]$. Now let A be the set of points x of the segment $[x(t'), x(t'')]$ for which $\lim_{t_i\to t, t_i > t} \rho(x(t), x) \leqq \mu$, and let us write $x_1 = \inf_{x\in A} x$ and $x_2 = \sup_{x\in A} x$ (the order on the segment $[x(t'), x(t'')]$ is defined by the fact that we consider $x(t') < x(t''))$. Now consider a point $\bar{x}$ of A such that $\rho(x_1, \bar{x}) > \varepsilon - 2\mu$ and $\rho(x_2, \bar{x}) > \varepsilon - 2\mu$. If $\bar{t}'$ and $\bar{t}''$ are points which, for some $c < c_0$, satisfy $\tilde{t}-c < \bar{t}' < \tilde{t}-c/2$, $\tilde{t}+c/2 < \bar{t}'' < \tilde{t}+c$ for $0 < \tilde{t}-t < c/4$ and $\rho(\bar{x}, x(\bar{t}'')) < \mu$, we find that if $0 < \tilde{t}-t < c/4$, the point $x(\tilde{t})$ lies at a distance no greater than 2μ from the segment $[x(t'), \bar{x}]$. Therefore x_1 and x_2 must lie at a distance no greater than 3μ from $[x(t'), \bar{x}]$ which is impossible if $\varepsilon > 5\mu$

In the second case we choose, for every ε, a δ such that for $0 < t_1 - t < \delta$ we have either $\rho(x(t_1), \bar{x}_1) < \varepsilon$ or $\rho(x(t_1), \bar{x}_2) < \varepsilon$. If $\varepsilon < \rho(\bar{x}_1, \bar{x}_2)/4$, only one of the possibilities is realized. Choose c_0 such that $\lim_{n\to\infty} \Delta(c, x_n(t)) < \varepsilon$ for $c < c_0$. Then if t_1 is a point for which $0 < t_1 - t < \delta$ and $0 < t_1 - t < c_0$ and $\rho(\bar{x}_1, x(t_1)) < \varepsilon$, we find that $\rho(x(\tilde{t}), \bar{x}_1) < \varepsilon$ for $(t_1-t)/3 < \tilde{t}-t < \frac{2}{3}(t_1-t)$ and $\tilde{t} \in N$, since

$$H(\bar{x}_1, x(\tilde{t}), x(t_1)) < \varepsilon,$$

and therefore

$$H(\bar{x}_1, x(\tilde{t}), \bar{x}_1) < 2\varepsilon,$$

or $\rho(\bar{x}_1, x(\tilde{t})) < 2\varepsilon$, and therefore $\rho(x_1, x(\tilde{t})) < \varepsilon$ (there are only two possibilities, namely $\rho(\bar{x}_1, x(\tilde{t})) < \varepsilon$ or $\rho(x_1, x(\tilde{t})) > 3\varepsilon$). Choosing t_2 from N such that $\frac{1}{2}t_1 < t_2 < \frac{2}{3}t_1$, we find that $\rho(x(\tilde{t}), \bar{x}_1) < \varepsilon$ for $(t_2-t)/3 < \tilde{t}-t < \frac{2}{3}(t_2-t)$. Continuing this process *ad infinitum*, we find that $\rho(x(\tilde{t}), \bar{x}_1) < \varepsilon$ for $t < \tilde{t} < \frac{2}{3}t_1$ and therefore $\lim_{\tilde{t}>t, \tilde{t}\to t} x(\tilde{t}) = \bar{x}_1$. Thus $x(t)$ has limits from the right and the left at every point. Setting

$$\bar{x}(t) = \lim_{\tilde{t}>t, \tilde{t}\to t, \tilde{t}\in N} x(\tilde{t}), \qquad \bar{x}(1) = \lim_{\tilde{t}\in N, \tilde{t}\to 1} x(\tilde{t}),$$

we obtain a function from K_X. We now have to prove that $x_1(t) \xrightarrow{M_2} x(t)$. To do this it is sufficient to show that $x_n(t)$ converges to $x(t)$ at all points of continuity of $\bar{x}(t)$. Let t_0 be a point of continuity of $\bar{x}(t)$. Then if we choose t' and t'' from N such that $t_0 - c < t' < t_0 - c/2$ and $t_0 + c/2 < t'' < t_0 + c$, where c is chosen so that $\overline{\lim}_{n\to\infty} \Delta_{M_2}(c, x_n(t)) < \varepsilon$, $\rho(\bar{x}(t'), \bar{x}(t_0)) < \varepsilon$, and $\rho(\bar{x}(t''), \bar{x}(t_0)) < \varepsilon$, we find that $\overline{\lim}_{n\to\infty} \rho(x_n(t_0), \bar{x}(t_0)) < 4\varepsilon$. This proves the lemma.

2.4. The topology M_1.

2.4.1. CONVERGENCE CONDITIONS. *A necessary and sufficient condition that* $x_n(t) \xrightarrow{M_1} x(t)$ *is that* 2.3.3(a) *be fulfilled and that*

$$\text{(b)} \qquad \lim_{c\to 0} \overline{\lim_{n\to\infty}} \Delta_{M_1}(c; x_n(t)) = 0,$$

where

$$\Delta_{M_1}(c; x_n(t)) = \sup_{t_2-c<t_1<t_2<t_3<t_2+c} H(x(t_1), x(t_2), x(t_3)).$$

PROOF. THE NECESSITY OF (b). Let $x_n(t) \xrightarrow{M_1} x(t)$. This means that there exist parametric representations $(y(s), t(s))$ for $\Gamma_{x(t)}$ and $(y_n(s), t_n(s))$ for $\Gamma_{x_n(t)}$ such that $\lim_{n\to\infty} \sup_s |t_n(s)-t(s)| = 0$ and $\lim_{n\to\infty} \sup_s \rho(y(s), y_n(s)) = 0$. Let $t_1 < t_2 < t_3$ and $s_1^{(n)}, s_2^{(n)}, s_3^{(n)}$ be chosen so that $t_n(s_i^{(n)}) = t_i$, $(i = 1, 2, 3)$. Then

$$\begin{aligned} &H(x_n(t_1), x_n(t_2), x_n(t_3)) \\ &= H(y_n(s_1^{(n)}), y_n(s_2^{(n)}), y_n(s_3^{(n)})) \leqq H(y(s_1^{(n)}), y(s_2^{(n)}), y(s_3^{(n)})) \\ &\quad + 3 \sup_s \rho(y_n(s), y(s)). \end{aligned}$$

It remains to prove that for $s_1<s_2<s_3$ we can make $H(y(s_1), y(s_2), y(s_3))$ arbitrarily small if $t(s_3)-t(s_1)$ is sufficiently small.

Let us consider three cases.

1. If $t(s_1) = t(s_2) = t(s_3)$, so that $H(y(s_1), y(s_2), y(s_3)) = 0$.

2. If $t(s_1) = t(s_2) < t(s_3)$, then either the discontinuity in $x(t)$ at $t(s_1)$ is less than ε, i.e. $\rho(y(s_1), y(s_2)) < \varepsilon$ and therefore $H(y(s_1), y(s_2), y(s_3)) < \varepsilon$, or the discontinuity in $x(t)$ at $t(s_1)$ is greater than ε and then for sufficiently small $t(s_3)-t(s_2)$ we have $\rho(y(s_3), x(t(s_1))) < 2\varepsilon$. This means that

$$H(y(s_1), y(s_2), y(s_3)) < 2\varepsilon.$$

3. If, finally, $t(s_1) < t(s_2) < t(s_3)$, then either there is no discontinuity at any of these points greater than ε and therefore

$$H(y(s_1), y(s_2), y(s_3)) < H(x(t(s_1)), x(t(s_2)), x(t(s_3)))+3\varepsilon \tag{2.16}$$

(it follows from 2.1.4 that (2.16) is small), or at $t(s_1)$ the discontinuity is greater than ε. This means that at the other points the discontinuities are less than ε for sufficiently small $t(s_3)-t(s_1)$ so that $\rho(x(t(s_2)), x(t(s_3)))$ is small and so is $H(y(s_1), y(s_2), y(s_3))$ (if the discontinuity greater than ε is at $t(s_3)$, the considerations are similar), or the discontinuity greater than ε is at $t(s_2)$ in which case $\rho(y(s_1), x(t(s_2-0)))$ and $\rho(x(t(s_2)), y(s_3))$ are small, and therefore so is

$$H(y(s_1), y(s_2), y(s_3)).$$

This proves the necessity.

SUFFICIENCY. First we shall show that if $\Delta_{M_1}(c; x(t)) < h$ and if (x_1, t_1), (x_2, t_2), (x_3, t_3) are points of $\Gamma_{x(t)}$ such that $t_2-c < t_1 < t_2 < t_3 < t_2+c$, then $H(x_1, x_2, x_3) < h$. To do this it is sufficient to prove the following two assertions: 1) if $H(\bar{x}_1', x_2, x_3) < h$, $H(\bar{x}_n'', x_2, x_3) < h$, and $x_1 \in [\bar{x}_1', \bar{x}_1'']$, then $H(x_1, x_2, x_3) < h$; 2) if $H(x_1, \bar{x}_2', x_3) < h$, $H(x_1, \bar{x}_n'', x_3) < h$, and $x_2 \in [\bar{x}_2', x_2'']$, then $H(x_1, x_2, x_3) < h$. Assertion (1) follows from the fact that if $y' \in [\bar{x}_1', x_3]$ and $y'' \in [\bar{x}_1'', x_3]$, where $\rho(y', x_2) < h$ and $\rho(y'', x_2) < h$, then any point y on the intersection of $[y', y'']$ and $[x_1, x_3]$ will also satisfy $\rho(x_2, y) < h$. Assertion (2) follows from the fact that if y' and y'' are points of $[x_1, x_3]$ such that $\rho(y', \bar{x}_2') < h$, $\rho(y'', \bar{x}_2'') < h$, and $x_2 = \alpha\bar{x}_2' + (1-\alpha)\bar{x}_2''$, $(0 < \alpha < 1)$, then we know about $y + \alpha y' + (1-\alpha)y''$ that it satisfies

$$\begin{aligned}\rho(x_2, y) &= \rho(\alpha\bar{x}_2' + (1-\alpha)\bar{x}_2'', \alpha y' + (1-\alpha)y'')\\ &\leqq \rho(\alpha\bar{x}_2', \alpha y') + \rho((1-\alpha)\bar{x}_2'', (1-\alpha)y'')\\ &\leqq \alpha\rho(\bar{x}_2', y') + (1-\alpha)\rho(\bar{x}_2'', y'') \leqq h\end{aligned}$$

(it should be recalled that although in this case X is a Banach space, we are using the distance rather than the norm for uniformity of notation). To prove our assertion, it is sufficient to construct, for every ε and for all sufficiently large n, parametric representations $(y_n^\varepsilon(s), t_n^\varepsilon(s))$ for $\Gamma_{x_n(t)}$ and $(y^\varepsilon(s), r^\varepsilon(s))$ for $\Gamma_{x(t)}$ such that

$$\sup_s |t_n^\varepsilon(s) - t^\varepsilon(s)| < \varepsilon; \quad \sup_s \rho(y_n^\varepsilon(s), y^\varepsilon(s)) < \varepsilon.$$

Let $(y(s), t(s))$ be some parametric representation of $\Gamma_{x(t)}$ and let $0 = s_0 < s_1 < s_2 < \cdots < s_m = 1$ be values of the parameter such that $t(s_i)$ are either points of continuity of $x(t)$ or points where the discontinuities are no less than ε, and further let $t(s_i) - t(s_i - 1) < \eta$ for $i = 1, \cdots, m-1$ and $\min[\rho(y(s_i), y(s)); \rho(y(s), y(s_{i+1}))] < \varepsilon/2$ for $s \in (s_i, s_{i+1})$. If finally $t(s_k) = t(s_{k+1})$, then $\varepsilon/4 < \rho(y(s_k), y(s_{k+1})) < \varepsilon/2$. Let us choose an n_0 sufficiently large and an $\eta < \varepsilon$ such that if $t(s_i)$ is a point of continuity of $x(t)$, then $\rho(x_n(t(s_i)), x(t(s_i))) < \varepsilon/4$ and $\Delta_{M_1}(\eta, x_n(t)) < \varepsilon/8$ for $n > n_0$. Then if the point $t' = t(s_r) = \cdots = t(s_{r+j})$ is a point of discontinuity of $x(t)$ at which the discontinuity is no less than ε, then there exist on $\Gamma_{x_n(t)}$ points $(y_r^{(n)}, t_r^{(n)})$, $(y_{r+1}^{(n)}, t_{r+1}^{(n)}), \cdots, (y_{r+j}^{(n)}, t_{r+j}^{(n)})$ such that $|t_{r+1}^n - t') < \eta$ and $\rho(y_{r+i}^{(n)}, y(s_{r+i})) < \frac{7}{8}\varepsilon$ (because the segment $[y(s_{r-1}), y(s_{r+j+1})]$ lies at a distance no greater than $\varepsilon/4$ from $[x_n(t(s_{r-1})), x_n(t(s_{r+j+1}))]$ and at a distance no greater than $\varepsilon/2$ from $[x(t(s_r)-0), x(t(s_{r+j}))]$, while all the points of $\Gamma_{x(t)}$ with $t(s_{r-1}) < t < t(s_{r+j+1})$ lie no further than $\varepsilon/8$ from $[x_n(t(s_{r-1})), x_n(t(s_{r+j+1}))]$.

Thus if $(y_n(s), t_n(s))$ is some parametric representation of $\Gamma_{x_n(t)}$, then there exist points $0 = s_0^{(n)} < s_1^{(n)} < \cdots < s_m^{(n)} = 1$ such that $|t_n'(s_i^{(n)}) - t(s_i)| < \varepsilon$ and $\rho(y_n(s_i^{(n)}), y(s_i)) < \varepsilon$. Let $\lambda(s)$ be the monotonic continuous function which maps $[0, 1]$ onto itself, such that $\lambda(s_i) = s_i^{(n)}$. Then $(y_n(\lambda(s)), t_n(\lambda(s)))$ will also be a parametric representation of $\Gamma_{x_n(t)}$, and it is easily shown that

$$|t_n(\lambda(s)) - t(s)| < 2\varepsilon, \ \rho(y_n(\lambda(s)), y(s)) < 4\varepsilon.$$

This completes the proof.

2.5. The topology $\mathbf{J}_2$.

2.5.1. *A necessary and sufficient condition that* $x_n(t) \xrightarrow{\mathbf{J}_2} x(t)$ *is that*

$$\lim_{c\to 0} \varlimsup_{n\to\infty} \Delta_{\mathbf{J}_2}[c; x_n(t), x(t)] = 0,$$

where

$$\Delta_{\mathbf{J}_2}(c; y(t), x(t)) = \sup_t [\min\{\rho(x(t_c), y(t)); \rho(y(t), x(t_c^*))\}]$$

(*the points* t_c *and* t_c^* *are defined as previously*).

PROOF. NECESSITY. Let $x_n(t) \xrightarrow{J_2} x(t)$, so that there exists a sequenec $\lambda_n(t)$ such that

$$\lim_{n\to\infty} \sup_t \rho(x_n(t), x(\lambda_n(t))) = 0; \quad \lim_{n\to\infty} \sup_t |t-\lambda_n(t)| = 0. \tag{2.17}$$

Therefore

$$\rho(x(t_c), x_n(t)) \leqq \rho(x(t_c), x(\lambda_n(t))) + \rho(x(\lambda_n(t)), x_n(t)),$$
$$\rho(x_n(t), x(t_c^*)) \leqq \rho(x(\lambda_n(t)), x(t_c^*)) + \rho(x(\lambda_n(t)), x_n(t)).$$

Hence

$$\begin{aligned}\min\{\rho(x(t_c), x_n(t)), \rho(x_n(t)), x(t_c^*))\} &\leqq \rho(x(\lambda_n(t)), x_n(t)) \\ &+ \min\{\rho(x(\lambda_n(t)), x(t_c)); \rho(x(\lambda_n(t)), x(t_c^*))\}.\end{aligned} \tag{2.18}$$

In view of (2.17) we have only to show that

$$\lim_{c\to 0} \overline{\lim_{n\to\infty}} \sup_t \min\{\rho(x(t_c), x(\lambda_n(t))); \rho(x(\lambda_n(t)), x(t_c^*))\}. \tag{2.19}$$

But if $|\lambda_n(t)-t| < c$, we have $t_c \leqq \lambda_n(t) \leqq t_c^*$, so that (2.19) follows from 2.1.4.

SUFFICIENCY. Assume there exists such a sequence $c_k \to 0$ and $n_k \to \infty$ that $\Delta_{J_2}(c_k; x_n(t), x(t)) < \varepsilon_k$ for $n > n_k$ and $\varepsilon_k \to 0$.

Let $t_1, t_2, \cdots. t_k$ be the points at which the discontinuities of $x(t)$ are no less than ε. Let us surround these points by intervals $\Delta_1, \Delta_2, \cdots, \Delta_k$ so small that

$$\begin{aligned}&\rho(x(t), x(t_i-0)) < \varepsilon \quad \text{for } t < t_i,\ t \in \Delta_i, \\ &\rho(x(t), x(t_i+0)) < \varepsilon \quad \text{for } t > t_i,\ t \in \Delta_i.\end{aligned} \tag{2.20}$$

Let us choose c_k so small that if $t \in [0, 1] - \bigcup_{i=1}^{k} \Delta_j$ we have

$$\begin{aligned}&\rho(x_n(t), x(t_{c_k}-0)) < 2\varepsilon, \\ &\rho(x(t), x(t_{c_k}^*)) < 2\varepsilon.\end{aligned} \tag{2.21}$$

Then for $n > n_k$ and $t \in \Delta_i$ we have either

$$\begin{aligned}&\rho(x_n(t), x(t_i-0)) < 3\varepsilon+\varepsilon_k, \\ &\text{or} \\ &\rho(x_n(t), x(t_i+0)) < 3\varepsilon+\varepsilon_k.\end{aligned} \tag{2.22}$$

If $\rho(x(t_i-0), x(t_i+0)) \geqq 6\varepsilon+\varepsilon_k$, then only one of (2.22) can be fulfilled. Let us now construct the function $\lambda_n^{\varepsilon,\varepsilon_k}(t)$, which maps the segment $[0, 1]$ onto itself in the following way: $\lambda_n^{\varepsilon,\varepsilon_k}(t) = t$ everywhere except on those intervals Δ_i where $\rho(x(t_i-0), x(t_i+0)) \geqq 6\varepsilon+2\varepsilon_k$, and on these intervals $\lambda_n^{\varepsilon,\varepsilon_k}(t)$ maps the set of values of t for which $\rho(x_n(t), x(t_i-0)) < 3\varepsilon+\varepsilon_k$ into the set $t < t_i$, $t \in \Delta_i$, and the set of values of t for which $\rho(x_n(t), x(t_i+0)) < 3\varepsilon+\varepsilon_k$ into the set $t \geqq t_i$, $t \in \Delta_i$. Obviously $|\lambda_n^{\varepsilon,\varepsilon_k}(t)-t|$ is equal to or less than the greatest of the Δ_i intervals, and $\rho(x_n(t), x(\lambda_n^{\varepsilon,\varepsilon_k}(t))) \leqq 4\varepsilon+\varepsilon_k$ for $n > n_k$. Now by choosing the sequences $\varepsilon \to 0$, $k \to \infty$, and $\delta = \{\text{max length } \Delta_i\} \to 0$, we obtain the desired result.

It is easily seen that in proving the sufficiency we have proved the following stronger assertion.

2.5.2. *A necessary and sufficient condition for $x_n(t) \xrightarrow{\mathbf{J}_2} x(t)$ is that*

$$\underline{\lim}_{c\to 0} \overline{\lim}_{n\to\infty} \Delta_{\mathbf{J}_2}(c; x_n(t), x(t)) = 0.$$

2.5.3. CONVERGENCE CONDITION IN THE TOPOLOGY $\mathbf{J}_2$. *Necessary and sufficient conditions for the functions $x_n(t)$ to be $\mathbf{J}_2$-convergent to $x(t)$ are that*

(a) *$x_n(t)$ converge to $x(t)$ on an everywhere dense set of* $[0, 1]$ *which contains* 0 *and* 1, *and*

(b) $\lim_{c\to 0} \overline{\lim}_{n\to\infty} \Delta_{\mathbf{J}_2}(c, x_n(t)) = 0$, *where*

$$\Delta_{\mathbf{J}_2}(c, y(t)) = \sup_{t_c \leqq t_1 \leqq t_c + c/2,\ t\in[0,1],\ t_c - c/2 \leqq t_2 \leqq t_c} \min\,[\rho(y(t_1), y(t)); \rho(y(t), y(t_2))].$$

PROOF. NECESSITY OF CONDITION (b). Let $x_n(t) \xrightarrow{\mathbf{J}_2} x(t)$. Then there exists a sequence $\lambda_n(t)$ such that $\sup |\lambda_n(t) - t|$ and $\sup_t \rho(x_n(t), x(\lambda_n(t))) \to 0$ as $n \to \infty$. Therefore

$$\rho(x_n(t_1), x_n(t)) \leqq \rho(x(\lambda_n(t_{i1})), x(\lambda_n(t))) + 2 \sup_t \rho(x(\lambda_n(t)), x_n(t)),$$

$$\rho(x_n(t), x_n(t_2)) \leqq \rho(x(\lambda_n(t), \lambda_n(t_2)) + 2 \sup_t \rho(x(\lambda_n(t)), x_n(t)).$$

Hence

$$\min\,[\rho(x_n(t_1), x_n(t)); \rho(x_n(t), x_n(t_2))] \leqq 2 \sup_t \rho(x_n(t), x(\lambda_n(t)))$$
$$+ \min\,[\rho(x(\lambda_n(t_1)), x(\lambda_n(t))); \rho(x(\lambda_n(t)), x(\lambda_n(t_2)))].$$

By choosing c sufficiently small and n so large that $|\lambda_n(t) - t| < c/2$, we can make the expression

$$\sup_{t_c \leqq t_1 \leqq t_c + c/2;\ t\in[0,1],\ t_c^* - c/2 \leqq t_2 \leqq t_c^*} \min\,[\rho(x(\lambda_n(t_1)), x(\lambda_n(t))); \rho(x(\lambda_n(t)), x(\lambda_n(t_2)))]$$

arbitrarily small, so that if $t \in [c, 1-c]$, then $\lambda_n(t_1) < \lambda_n(t) < \lambda_n(t_2)$ and $\lambda_n(t_2) - \lambda_n(t_1) < 3c$. This means we can apply 2.1.4. Now for $t \leqq c$ and $t \geqq 1-c$, all the quantities $x(\lambda_n(t_1))$, $x(\lambda_n(t))$, $x(\lambda_n(t_2))$ can be made, arbitrarily close to $x(0)$ or $x(1)$ by choosing c sufficiently small. The necessity is proven. To prove the sufficiency, we shall prove the following stronger assertion.

2.5.4. *A sufficient condition for $x_n(t) \xrightarrow{\mathbf{J}_2} x(t)$ is that condition* 2.5.3(a) *be fulfilled, and*

$$\text{(b}'\text{)} \qquad \underline{\lim}_{c\to 0} \overline{\lim}_{n\to\infty} \Delta_{\mathbf{J}_2}(c, x_n(t)) = 0.$$

PROOF. We have

$$\rho(x(t_{c_k}), x_n(t)) \leqq \rho(x(t_{c_k}), x(t')) + \rho(x_n(t), x_n(t')) + \rho(x_n(t'), x(t')),$$
$$\rho(x_n(t), x(t_{c_k}^*)) \leqq \rho(x_n(t), x_n(t'')) + \rho(x_n(t''), x(t'')) + \rho(x(t''), x(t''_{c_k})).$$

Since $x_n(t)$ converges to $x(t)$ on an everywhere dense set, there is for every t a $\bar{t}$ such that $|t - \bar{t}'| < \delta$, and as $n \to \infty$, we have $\rho(x_n(\bar{t}'), x(\bar{t}')) \to 0$; further $\rho(x(t), x(\bar{t}')) < \varepsilon$. Then for $\delta < c_k/4$

$$\overline{\lim}_{n\to\infty} \Delta_{\mathbf{J}_2}(c_k, x_n(t), x(t)) \leqq \overline{\lim}_{n\to\infty} \Delta_{\mathbf{J}_2}(c_k + \delta, x_n(t)) + 2\varepsilon.$$

Therefore if

$$\lim_{c_k\to 0}\ \overline{\lim_{n\to\infty}}\, \Delta_{\mathbf{J}_2}(c, x_n(t)) = 0,$$

we have

$$\lim_{c_k\to 0}\ \overline{\lim_{n\to\infty}}\, \Delta_{\mathbf{J}_2}(c, x_n(t), x(t)) = 0,$$

which, in view of 2.5.2, proves our assertion.

2.6. The topologies $\mathbf{J}_1$ and U.

2.6.1. CONVERGENCE CONDITIONS IN THE TOPOLOGY $\mathbf{J}_1$. *Necessary and sufficient conditions for* $x_n(t) \xrightarrow{\mathbf{J}_1} x(t)$ *are that*

(a) $x_n(t)$ *converges to* $x(t)$ *on an everywhere dense set containing* 0 *and* 1, *and*

(b) $\lim_{c\to 0} \overline{\lim}_{n\to\infty} \Delta_{\mathbf{J}_1}(c, x_n(t)) = 0$, *where*

$$\Delta_{\mathbf{J}_1}(c, y(t)) = \sup_{t-c<t_1<t\le t_2<t+c} \min\,[\rho(y(t_1), y(t));\ \rho(y(t), y(t_2))].$$

PROOF. NECESSITY OF CONDITION (b). If $x_n(t) \xrightarrow{\mathbf{J}_1} x(t)$, then there exists a sequence of monotonic continuous functions $\lambda_n(t)$, which map $[0, 1]$ onto itself such that $\lim_{n\to\infty} \sup_t |\lambda_n(t)-t| = 0$ and $\lim_{n\to\infty} \sup_t \rho(x_n(t), x(\lambda_n(t))) = 0$. Then

$$\min\,[\rho(x_n(t_1), x_n(t));\ \rho(x_n(t), x_n(t_2))] < 4 \sup_t \rho(x_n(t), x(\lambda_n(t)))$$
$$+ \min\,[\rho(x(\lambda_n(t_1)), x(\lambda_n(t)));\ \rho(x(\lambda_n(t)), x(\lambda_n(t_2)))],$$

and since (1) $\lambda_n(t_1) < \lambda_n(t) < \lambda_n(t_2)$, and (2), as $c \to 0$ and $n \to \infty$, we know that $\lambda_n(t_2)-\lambda_n(t_1)$ becomes arbitrarily small, it follows that $\Delta_{\mathbf{J}_1}(c, x_n(t))$ approaches zero as $n\to\infty$ and $c \to 0$.

SUFFICIENCY. Assume that (a) and (b) are fulfilled. We must construct a sequence of monotonic functions $\lambda_n(t)$ mapping $[0, 1]$ onto itself such that $\sup_t \rho(x_n(t), x(\lambda_n(t)))\to 0$ and $|\lambda_n(t)-t| \to 0$ as $n\to\infty$. To do this it is sufficient to construct continuous monotonic functions $\lambda_n^\varepsilon(t)$ mapping $[0, 1]$ onto itself, such that for $n > n_\varepsilon$ we have $\rho(x_n(t), x(\lambda_n^\varepsilon(t))) < \varepsilon$ and $\lim_{\varepsilon\to 0} \overline{\lim}_{n\to\infty} |\lambda_n^\varepsilon(t)-t| = 0$. Let us attempt to construct such functions.

Let $t_1, t_2, \cdots, t_k$ be all the points where the discontinuities of $x(t)$ are no less than μ. Let us choose c and N' so that for $n > N'$ we have

$$\Delta_{\mathbf{J}_1}(c, x_n(t)) < \mu. \tag{2.23}$$

Further, let this c be so small that if t' and t'' both belong to the same one of the intervals $(0, t_1)$, (t_1, t_2), $\cdots$, $(t_k, 1)$, then $\rho(x(t'), x(t''))<\mu$ so long as $|t'-t''|<c$. The fact that $x_n(t)$ converges to $x(t)$ on an everywhere dense set implies that there exist points $0 = \tau_0^{(c)} < \tau_1^{(c)} < \tau_2^{(c)} < \cdots < \tau_m^{(c)} = 1$ such that $\tau_i^{(c)}-\tau_{i-1}^{(c)} < c/2$, $(i = 1, 2, \cdots, m)$, and $\rho(x_n(\tau_i^{(c)}), x(\tau_i^{(i)})) < \mu$ for $n > N''$. Let t_l lie between $\tau_{i_l}^{(c)}$ and $\tau_{i_l+1}^{(c)}$. Then if the discontinuity at t_l is greater than 8μ, it follows that $\rho(x(\tau_{i_l}^{(c)}), x(\tau_{i_l+1}^{(c)})) > 6\mu$. This implies that $\rho(x_n(\tau_{i_l}^{(c)}), x_n(\tau_{i_l+1}^{(c)})) > 4\mu$ for $n > N''$. But if $n > \max\{N', N''\}$ and $t\in(\tau_{i_l}^{(c)}, \tau_{i_l+1}^{(c)})$, we have $\min\{\rho(x_n(\tau_{i_l}^{(c)}), x_n(t)), \rho(x_n(\tau_{i_l+1}^{(c)}), x_n(t))\} < \mu$. Therefore in the interval $(\tau_{i_l}^{(c)}, \tau_{i_l+1}^{(c)})$ there is a point $\bar t_l$ such that $t \in (\tau_{i_l}^{(c)}, \bar t_l)$ implies $\rho(x_n(\tau_{i_l}^{(c)}), x_n(t)) < \mu$ and $t\in(\bar t_l, \tau_{i_l+1}^{(c)})$ implies $\rho(x_n(t), x_n(\tau_{i_l+1}^{(c)})) < \mu$. (If this were not so, there would be points

$t_1' < t_2' < t_3' \in (\tau_{i_l}^{(c)}, \tau_{i_l+1}^{(c)})$ such that

$$\rho(x_n(t_1'), x(\tau_{i_l}^{(c)}) < \mu, \ \rho(x_n(t_1'), x_n(\tau_{i_l+1}^{(c)})) < \mu, \ \rho(x_n(t_3'), x(\tau_{i_l}^{(c)})) < \mu,$$

or $\rho(x_n(t_1'), x_n(t_2')) > 2\mu$ and $\rho(x_n(t_2'), x_n(t_3')) > 2\mu$, which is impossible according to (2.23).)

Now let $t_{k_1}, t_{k_2}, \cdots, t_{k_r}$ be points where the discontinuities are greater than 8μ, and let the $\bar{t}_{k_i}$ be points constructed for the t_{k_i} in exactly the same way as $\bar{t}_l$ was constructed for t_l. Let $\lambda_n^{\varepsilon}(t)$ be a function continuous and monotonic on each of the intervals $(0, t_{k_1}), (t_{k_1}, t_{k_2}), \cdots, (t_{k_r}, 1)$ of the form

$$\lambda_n^{\varepsilon}(t) = a_i(t) + b_i, \quad \lambda_n^{\varepsilon}(t_{k_i}) = \bar{t}_{k_i}.$$

Obviously $|\lambda_n^{\varepsilon}(t) - t| < c/2$. Let us evaluate $\rho(x_n(t), x(\lambda_n(t)))$ for $t \in (\tau_j^{(c)}, \tau_{j+1}^{(c)})$. If in this interval there is not a single point $\bar{t}_{k_i}$ (and therefore no point t_{k_i}), then

$$\rho(x_n(t), x(\lambda_n^{\varepsilon}(t))) \leqq \mu + \min\{\rho(x_n(t), x_n(\tau_j^{(c)})); \rho(x_n(t)), x(\tau_{j+1}^{(c)}))\}$$
$$+ \max\{\rho(x(\lambda_n^{\varepsilon}(t), x(\tau_j^{(c)})); \rho(x(\lambda_n^{\varepsilon}(t)), x(\tau_{j+1}^{(c)}))\} < 12\mu.$$

If $t \in (\tau_j^{(c)}, \tau_{j+1}^{(c)})$, however, and t_{k_ν} and $\bar{t}_{k_\nu}$ also belong to $(\tau_j^{(c)}, \tau_{j+1}^{(c)})$, then either $t \in (\tau_j^{(c)}, t_{k_\nu})$, $\lambda_n^{\varepsilon}(t) < \bar{t}_{k_\nu}$ or $t \in (\bar{t}_{k_\nu}, \tau_{j+1}^{(c)})$, $\lambda_n^{\varepsilon}(t) > \bar{t}_{k_\nu}$. This implies that $\rho(x_n(t), \lambda_n^{\varepsilon}(t)) < 3\mu$. If we set $\varepsilon = 12\mu$, we obtain the desired result.

2.6.2. Convergence conditions in the topology U. *Necessary and sufficient conditions for $x_n(t) \xrightarrow{U} x(t)$ are:*

(a) $x_n(t) \xrightarrow{\mathbf{J}_1} x(t)$ *and*

(b) *if $t_1, t_2, \cdots, t_k$ denote points where the discontinuities of $x(t)$ are greater than ε, where ε is arbitrary except that $x(t)$ has no discontinuities whose magnitude is ε, and if $t_1^{(n)}, t_2^{(n)}, \cdots, t_{k_n}^{(n)}$ denote the points where the discontinuities of $x_n(t)$ are greater than ε, then for n greater than a certain value we have $k_n = k$, $t_i = t_i^{(n)}$.*

The necessity of both conditions is obvious. The sufficiency is seen if we note that by constructing $\lambda_n^{\varepsilon}(t)$ (see the proof of the sufficiency in 2.6.1), we obtain

$$\lambda_n^{\varepsilon}(t) = t.$$

2.7. Compactness conditions in the topologies $\mathbf{J}_1$, $\mathbf{J}_2$, M_1, and M_2. We note the following conditions necessary for compactness.

2.7.1. *If a set of functions $K \subset K_X$ is to be compact in one of the topologies $\mathbf{J}_1, \mathbf{J}_2, M_1, M_2$, it is necessary that for all $t \in [0, 1]$ and $x(t) \in K_X$ the values of $x(t)$ belong to a single compact set A of X.*

Indeed, if we have a sequence of points $x_k(t_k)$, then by choosing a sequence n_k, such that $x_{n_k}(t) \xrightarrow{M_2} x_0(t)$, $t_{n_k} \to t_0$, we find that the distance between $x_{n_k}(t_{n_k})$ and the segment $[x_0(t_0 - 0), x_0(t_0)]$ approaches zero, which means that $x_{n_k}(t_{n_k})$ is compact, so that the segment is compact.

2.7.2. *Sufficient condition for compactness.* The set of functions K is compact in a topology S, where S is $\mathbf{J}_1, \mathbf{J}_2, M_1$, or M_2, if 2.7.1 is fulfilled and if

$$\lim_{c\to 0} \limsup_{x(t)\in K} \left(\Delta_S(c, x(t)) + \sup_{0<t<c} \rho(x(0), x(t)) + \sup_{1-c<t<1} \rho(x(1), x(t))\right) = 0 \tag{2.24}$$

(by lim sup of a certain numerical set we mean its maximum limit point).

Proof. Choosing some everywhere dense set N of values of t which contain 0 and 1, we may take from any sequence $x_n(t)$ a subsequence $x_{n_k}(t)$ such that if $t \in N$ then $\lim_{n_k\to\infty} x_{n_k}(t)$ exists. Since $\Delta_{M_2}(c, x(t)) \leqq \Delta_s(c, x(t))$, we have

$$\lim_{c\to 0} \varlimsup_{n_k\to\infty} \left(\Delta_{M_2}(c, x_{n_k}(t)) + \sup_{0<t<c} \rho(x(0), x(t)) + \sup_{1-c<t<1} \rho(x(1), x(t))\right) = 0.$$

According to 2.3.5, therefore, there exists a $\bar{x}(t) \in K_X$ such that $x_{n_k}(t) \xrightarrow{M_2} \bar{x}(t)$. This means that $x_{n_k}(t)$ converges to $\bar{x}(t)$ on an everywhere dense set containing 0 and 1, and that (2.24) is fulfilled. Hence $x_{n_k}(t) \xrightarrow{S} \bar{x}(t)$.

Bearing in mind that $\Delta_{\mathbf{J}_1}(c, x(t))$ and $\Delta_{M_1}(c, x(t))$ are monotonic functions of c, it is easy to obtain the following:

2.7.3. *Necessary and sufficient condition for compactness in the topologies* $\mathbf{J}_1$, M_1. [2]
Let S_1 denote either $\mathbf{J}_1$ or M_1. Conditions 2.7.1 and

$$\lim_{c\to 0} \sup_{x(t)\in K} (\Delta_{S_1}(c, x(t)) + \sup_{0<t<c} \rho(x(0), x(t)) + \sup_{1-c<t<1} \rho(x(1), x(t))) = 0$$

are necessary and sufficient for K to be compact in S_1. The next example shows that condition (2.24) is not necessary either for $\mathbf{J}_2$ or for M_2.

2.7.4. Example. Consider the set K consisting of the functions $x_s(t)$, $s \in [\frac{1}{4}, \frac{1}{2}]$, given by

$$x_s(t) = \begin{cases} 0, & 0 \leqq t < s, \\ 1, & s \leqq t < \frac{1}{4}+s/2, \\ 0, & \frac{1}{4}+s/2 \leqq t < \frac{1}{2}, \\ 1, & \frac{1}{2} \leqq t \leqq 1. \end{cases}$$

The set is $\mathbf{J}_2$-compact and therefore also M_2-compact, but

$$\limsup_{S} \Delta_{\mathbf{J}_2}(c, x_s(t)) = 1,$$

$$\limsup_{S} \Delta_{M_2}(c, x_s(t)) = 1 \quad \text{for } c < \tfrac{1}{2}.$$

2.8. A certain inequality. The inequality we are about to derive hardly has any independent meaning, but will be needed in the proof of the limit theorems.

Let us write

$$R(\delta, x(t), y(t)) = \sup_{0\leqq t\leqq 1-\delta} \inf_{t\leqq t_1\leqq t+\delta} \rho(x(t), y(t)),$$

$$\bar{R}(\delta, x(t), y(t)) = \max\{\rho(x(0), y(0)); \rho(x(1), y(1)); R(\delta, x(t), y(t))\}.$$

Further, we set

$$k(c, x(t)) = \max\{\sup_{0\leqq t\leqq c} \rho(x(0), x(t)); \sup_{1-c\leqq t\leqq 1} \rho(x(1), x(t))\}.$$

Then for $\bar{c} < c/4$, we have

$$\Delta_S(c, y(t)) \leqq L(\bar{R}(\alpha\bar{c}, x(t), y(t)) + \Delta_{\mathbf{J}_1}(3c, x(t)) + k(2c, x(t)) + \Delta_S(\bar{c}, y(t))),$$

where S denotes one of the topologies M_2, M_1, $\mathbf{J}_2$, or $\mathbf{J}_1$, and L and $\alpha < 1$ are certain absolute constants.

We shall give the proof for M_2. We note two properties of the quantity $\Delta_{\mathbf{J}_1}(c, x(t))$. These are: (1) on a segment of length c the function $x(t)$ cannot have two discontinuities larger than $2\Delta_{\mathbf{J}_1}(c, x(t))$ and (2) on a segment of length c there can exist two points t_1 and t_2 such that $\rho(x(t_1), x(t_2)) > 4\Delta_{\mathbf{J}_1}(c, x(t))$ only if there exists $t_1 < t^* < t_2$ such that $\rho(x(t^*-0), x(t^*+0)) > 2\Delta_{\mathbf{J}_1}(c, x(t))$. Let us now evaluate $H(y(t_1), y(t_2), y(t_3))$, where $t_1 \in [t_{2_c}, t_{2_c} + c/2]$ and $t_3 \in [t^*_{2_c} - c/2, t^*_{2_c}]$. If $t \in [c, 1-c]$, then there exist points t'_i, t''_i, $i = 1, 2, 3$, with $t'_i \in [t_{i\bar{c}}, t_{i\bar{c}} + \bar{c}/2]$, $t''_i \in [t^*_{i\bar{c}} - \bar{c}/2, t^*_{i\bar{c}}]$ such that $\rho(x(t'_i), y(t'_i))$ and $\rho(x(t''_i), y(t''_i))$ are less than $\bar{R}(\frac{1}{2}\bar{c}, x(t), y(t))$. If in the interval (t'_2, t''_2) the function $x(t)$ has a

[2] The compactness conditions for the topology $\mathbf{J}_1$ have been obtained by A. N. Kolmogorov.

discontinuity greater than $2\Delta_{\mathbf{J}_1}(2c, x(t))$, then $\rho(x(t_1'), x(t_2'))$ and $\rho(x(t_2''), x(t_2'))$ $< 4\Delta_{\mathbf{J}_1}(2c, x(t))$, and therefore

$$\rho(y(t_1), y(t_2')) < \Delta_{M_2}(\bar{c}, y(t)) + 2\bar{R}(\bar{c}/2, x(t), y(t)) + 4\Delta_{\mathbf{J}_1}(2c, x(t)).$$

Similarly,

$$\rho(y(t_3), y(t_2'')) < \Delta_{M_2}(\bar{c}, y(t)) + 2\bar{R}(\bar{c}/2, x(t), y(t)) + 4\Delta_{\mathbf{J}_1}(2c, x(t)). \tag{2.25}$$

Therefore

$$H(y(t_1), y(t_2), y(t_3)) < 2\Delta_{M_2}(\bar{c}, y(t)) + 2\bar{R}(c/2, x(t), y(t)) + 4\Delta_{\mathbf{J}_1}(2c, x(t)).$$

If, on the other hand, in the interval (t_2', t_2'') the function $x(t)$ has no discontinuity greater than $2\Delta_{\mathbf{J}_1}(2c, x(t))$, then either

$$\rho(x(t_1'), x(t_2')) < 4\Delta_{\mathbf{J}_1}(2c, x(t)),\ \rho(x(t_1''), x(t_2'')) < 4\Delta_{\mathbf{J}_1}(2c, x(t)),$$

which means

$$\rho(y(t_1), y(t_3)) < 2\Delta_{M_2}(\bar{c}, y(t)) + 2\bar{R}(\bar{c}/2, x(t), y(t)) + 4\Delta_{\mathbf{J}_1}(2c, x(t));$$

or

$$\rho(y(t_3'), x(t_2')) < 4\Delta_{\mathbf{J}_1}(2c, x(t)), \quad \rho(x(t_3''), x(t_2'')) < 4\Delta_{\mathbf{J}_1}(2c, x(t))$$

which means

$$\rho(y(t_2), y(t_3)) < 2\Delta_{M_2}(\bar{c}, y(t)) + \bar{R}(\bar{c}/2, x(t), y(t)) + 4\Delta_{\mathbf{J}_1}(2c, x(t)),$$

so that (2.25) remains valid. Now let $t_2 < c$. Then

$$\rho(y(t_2), x(0)) < \Delta_{M_2}(\bar{c}, y(t)) + \bar{R}(\bar{c}/2, x(t), y(t)) + k(c, x(t)),$$

so that

$$\rho(y(t_1), y(t_2)) < 2\Delta_{M_2}(\bar{c}, y(t)) + 2\bar{R}(c/2, x(t), y(t)) + 2k(c, x(t)).$$

The case in which $1-c < t_2 < 1$ is entirely analogous.

3. Fundamental Limit Theorems

Let us consider a stochastic process $\xi(t)$ given over a time interval $[0, 1]$ and taking on values from a certain complete separable metric space X. This means that there exists a space Ω of elementary outcomes ω with a probability measure P on Ω, and we are investigating functions $\xi(t, \omega)$ defined for $t \in [0, 1]$ and for almost all $\omega \in \Omega$ in the sense of the measure P, and which take their values from X. A function $\xi(t, \omega)$ is measurable in the measure P for every fixed t, which means that for every Borel set $A \subset X$, the set of those ω for which $\xi(t, \omega) \in A$ is a P-measurable set. The probabilities $\mathbf{P}\{\xi(t_1) \in A_1; \xi(t_2) \in A_2, \cdots; \xi(t_k) \in A_k\}$ (the A_i are Borel sets in X) are called the finite dimensional distributions of the process $\xi(t, \omega)$. We shall say that the finite dimensional distributions of the processes $\xi_n(t, \omega)$ converge to the finite dimensional distributions of the process $\xi(t)$ if

$$\mathbf{P}\{\xi_n(t_1) \in A_1; \xi_n(t_2) \in A_2; \cdots; \xi_n(t_k) \in A_k\} \to \mathbf{P}\{\xi(t_1) \in A_1; \xi(t_2) \in A_2; \cdots; \xi(t_k) \in A_k\}$$

for all k, for any collection $t_1, t_2, \cdots, t_k \in [0, 1]$, and for any collection of Borel sets $A_1, A_2, \cdots, A_k$ for which

$$\mathbf{P}\{\xi(t_i) \in \bar{A}_i \cap \overline{X - A_i}\} = 0 \qquad (i = 1, 2, \cdots, k).$$

In what follows we shall assume without further mention that for almost all ω (with probability 1) we have $\xi(t, \omega) \in K_X$.

We shall henceforth write $\xi(t, \omega)$ only when Ω is some specific space. Otherwise we shall write simply $\xi(t)$.

3.1. Consider a sequence of random variables ξ_n whose values belong to X and whose probability converges to some random variable ξ, which is to say that

$$\lim_{n\to\infty} \mathbf{P}\{\rho(\xi_n, \xi) > \varepsilon\} = 0 \qquad \text{for arbitrary } \varepsilon.$$

Then it is easy to see that the distribution of ξ_n converges weakly to the distribution of ξ. The converse statement is, of course, not true, since it is possible to give ξ_n and ξ even on different spaces of elementary outcomes. Nevertheless we shall now prove an assertion similar to the converse.

3.1.1. *Assume that the distribution of the quantities ξ_n converges weakly to the distribution of ξ_0. Then there exist random variables $x_n(\omega)$ with $n = 0, 1, \cdots$, such that*

$$P\{\xi_n \in A\} = P\{x_n(\omega) \in A\} \qquad (n = 0, 1, \cdots)$$

and

$$P\{x_n(\omega) \to x_0(\omega), n \to \infty\} = 1.$$

(*Here Ω is the segment* $[0, 1]$, *and P is the ordinary Lebesgue measure*).

PROOF. Let us choose a Borel set $S_{i_1, i_2, \cdots, i_k} \subset X$, where $k, i_1, i_2, \cdots, i_k$ take on all integral values, such that the following properties are fulfilled:

1) $S_{i_1, i_2, \cdots, i_k}$ and $S_{i'_1, i'_2, \cdots, i'_k}$ have no intersection if $i_k \neq i'_k$;

2) $\bigcup_{i_k=1}^{\infty} S_{i_1, i_2, \cdots, i_{k-1}, i_k} = S_{i_1, i_2, \cdots, i_{k-1}}$; $\bigcup_{i_1=1}^{\infty} S_{i_1} = X$;

3) The diameter of $S_{i_1, i_2, \cdots, i_k}$ is less than or equal to $(\frac{1}{2})^k$;

4) $\mathbf{P}\{\xi_0 \in \overline{S_{i_1, i_2, \cdots, i_k}} \cap \overline{X - S_{i_1, i_2, \cdots, i_k}}\} = 0$.

The set $S_{i_1, i_2, \cdots, i_k}$ can be constructed in the following way. Let $x_1^{(k)}, x_2^{(k)}, \cdots$ be a sequence of points such that every point of X lies at a distance no greater than $(\frac{1}{2})^{k+1}$ from at least one point of the sequence. We denote by $S'_r(x_i^{(k)})$ the set of points x such that $\rho(x, x_i^{(k)}) < r$, and we can choose an r_k, $(\frac{1}{2})^{k+1} < r_k < (\frac{1}{2})^k$, such that

$$\mathbf{P}\{\xi_0 \in \overline{S'_{r_k}(x_i^{(k)})} \cap \overline{X - S'_{r_k}(x_i^{(k)})}\} = 0 \qquad \text{for all } i, \tag{3.1}$$

since there is at most a denumerable number of values of r where (3.1) is positive for at least one value of i. Let us write

$$D_i^k = S'_{r_k}(x_i^{(k)}) - \bigcup_{j=1}^{i-1} S'_{r_k}(x_i^k),$$

$$S_{i_1, i_2, \cdots, i_k} = D_{i_1}^1 \cap D_{i_2}^2 \cap \cdots \cap D_{i_k}^k.$$

We denote by $\Delta_{i_1, i_2, \cdots, i_k}^{(n)}$ intervals of the segment $[0, 1]$ defined as follows. The intervals $\Delta_{i_1, \cdots, i_k}^{(n)}$ and $\Delta_{i'_1, \cdots, i'_k}^{(n)}$ have no intersection if $i_k \neq i'_k$, and

$\Delta^{(n)}_{i_1, i_2, \cdots, i_k}$ lies to the left of $\Delta^{(n)}_{i'_1, i'_2, \cdots, i'_k}$ if there exists an r such that $i_j = i'_j$ for $j = 1, 2, \cdots, r-1$ and $i_r < i'_r$. Finally the length of $\Delta^{(n)}_{i_1, i_2, \cdots, i_k}$ is $\mathbf{P}\{\xi_n \in S_{i_1, i_2, \cdots, i_k}\}$. From each set $S_{i_1, i_2, \cdots, i_k}$ we now choose one point $\bar{x}_{i_1, i_2, \cdots, i_k}$. We define the functions $x_n^m(\omega)$ with $\omega \in [0, 1]$ by

$$x_n^m(\omega) = \bar{x}_{i_1, i_2, \cdots, i_m} \text{ for } \omega \in \Delta^{(n)}_{i_1, i_2, \cdots, i_m}.$$

It is easily seen that $\rho(x_n^m(\omega), x_n^{m+p}(\psi)) \leqq (\frac{1}{2})^m$, so that, since X is a complete space, the limit $x_n(\omega) = \lim_{m\to\infty} x_n^m(\omega)$ exists. Since the length of $\Delta^{(n)}_{i_1, i_2, \cdots, i_m}$ tends to that of $\Delta^{(0)}_{i_1, i_2, \cdots, i_m}$, for all internal points of the intervals $\Delta^{(0)}_{i_1, i_2, \cdots, i_m}$ we have

$$\lim_{n\to\infty} \rho(x_0(\omega), x_n(\omega)) \leqq (\tfrac{1}{2})^{m-1},$$

and therefore for all ω except possibly a denumerable set, we have $x_n(\omega)\to x_0(\omega)$.

We have yet to verify that the $x_n(\omega)$ and ξ_n have the same distributions. For this purpose it is sufficient, for instance, to show that $\mathbf{P}\{x_n(\omega) \in A\} = \mathbf{P}\{\xi_n \in A\}$ for A such that

$$\mathbf{P}\{\xi_n \in \overline{A} \cap \overline{X-A}\} = 0.$$

Assume, therefore, that A is such a set. By $A^{(m)}$ we denote the sum of the $S_{i_1, i_2, \cdots, i_m}$ belonging to A, and by $A'^{(m)}$ the sum of the $S_{i_1, i_2, \cdots, i_m}$ which do not belong to $X-A$. It is obvious that $A^{(m)} \subset A \subset A'^{(m)}$ and

(3.2) $\quad \mathbf{P}\{x_n(\omega) \in A^{(m)}\} = P\{\xi_n \in A^{(m)}\}; \ \mathbf{P}\{x_n(\omega) \in A'^{(m)}\} = \mathbf{P}\{\xi_n \in A'^{(m)}\}.$

Further, if $C^{(m)}$ is the set of points whose distance from $\overline{A} \cap \overline{X-A}$ is equal to or less than $(\frac{1}{2})^m$, then $A'^{(m)}-A^{(m)} \subset C^{(m)}$, and since $\mathbf{P}\{\xi_n \in \overline{A} \cap \overline{X-A}\} = 0$ we find that $\mathbf{P}\{\xi_n \in A'^{(m)}-A^{(m)}\} \to 0$ as $m \to \infty$. Together with (3.2) this proves our assertion.

What we have just proved leads to the following corollary for sequences of processes.

3.1.2. *Consider a sequence of processes $\xi_n(t)$ whose trajectories belong to K_X with probability* 1, *and let their finite dimensional distributions tend to the finite dimensional distributions of $\xi_0(t)$. Then it is possible to construct processes $x_n(t, \omega)$ with $n = 0, 1, \cdots$, (Ω is the segment* $[0, 1]$ *and P is the Lebesgue measure on this segment), such that the finite dimensional distributions of $x_n(t, \omega)$ and $\xi_n(t)$ are the same, $x_n(t, \omega) \in K_X$ for almost all $\omega \in [0, 1]$, and $x_n(t, \omega) \to x(t, \omega)$ for almost all ω and for all t belonging to some denumerable set N which is everywhere dense on* $[0, 1]$, *but whose choice is otherwise arbitrary.*

To prove this, let $N = \{t^1, t^2, \cdots\}$ be some everywhere dense set. Denote by $X^{(\infty)}$ the space of the sequences $(x_1, x_2, \cdots)$, where $x_i \in X$. In $X^{(\infty)}$ we introduce the metric $\rho^{(\infty)}$, defined by

$$\rho^{(\infty)}[(x_1, x_2, \cdots), (y_1, y_2, \cdots)] = \sum_{k=1}^{\infty} (1-e^{-\rho(x_k, y_k)}) \frac{1}{k!}.$$

With this metric, $X^{(\infty)}$ is a complete separable metric space. Let us now consider in $X^{(\infty)}$ the random variable

$$\xi_n = \{\xi_n(t^1), \xi_n(t^2), \cdots\}.$$

It is easily shown that the distributions of the ξ_n converge weakly to the distribution of ξ_0. It is therefore possible to construct functions $x_n^{(\infty)}(\omega)$ whose values lie in $X^{(\infty)}$ and which are defined for $\omega \in [0, 1]$, such that $x_n^{(\infty)}(\omega) \to x_0^{(\infty)}(\omega)$ with probability 1 and such that the distributions of the $x_n^{(\infty)}(\omega)$ and ξ_n are the same.

Let $x_n^{(\infty)}(\omega) = (x_1^{(n)}(\omega), x_2^{(n)}(\omega), \cdots)$. We write

$$x_n(t^i, \omega) = x_i^{(n)}(\omega).$$

Since the joint distributions of $x_n(t^i, \omega)$ and $\xi_n(t^i)$ are the same, there almost certainly exists the limit

$$x_n(t, \omega) = \lim_{\substack{t^j \to t \\ t^j > t}} x_n(t^j, \omega). \tag{3.3}$$

We wish to find the functions $x_n(t, \omega)$. That they belong to K_X with probability 1 follows from the fact that the joint distributions of $x_n(t^j, \omega)$ and $\xi_n(t^j)$ coincide and from (3.3). By construction we have $x_n(t^j, \omega) \to x_0(t^j, \omega)$ with probability 1.

Clearly, for our purposes it was not necessary that the finite dimensional distributions of $\xi_n(t)$ converge to those of $\xi(t)$ for all t. It would have been sufficient that this be true for all t in N.

3.2. Let us now recall the convergence conditions in the topologies $\mathbf{J}_1$, $\mathbf{J}_2$, M_1, and M_2, and let us relate these conditions to the results just obtained. The way in which we will obtain the desired limit theorems then becomes immediately evident.

Let S be one of the topologies $\mathbf{J}_1$, $\mathbf{J}_2$, M_1, M_2. Let F be some metric separable space, and let F_S be the set of all functions $f[x(t)]$ defined on K_X whose values lie in F and which are continuous in S.

3.2.1. Theorem. *If for any $f \in F_S$ the distribution of $f[\xi_n(t)]$ is to converge, as $n \to \infty$, to the distribution of $f[\xi_0(t)]$, it is sufficient that*

(a) *the finite dimensional distributions of $\xi_n(t)$ converge to those of ξ_0 for t in some set N everywhere dense on* $[0, 1]$ *and containing* 0 *and* 1, *and*

(b) $\lim_{c \to 0} \overline{\lim_{n \to \infty}} \mathbf{P}\{\Delta_S(c, \xi_n(t)) > \varepsilon\} = 0$ *for all* ε.

PROOF. If condition (a) is fulfilled, we can construct a sequence $x_n(t, \omega)$ which converges with probability 1 to $x_0(t, \omega)$ as $n \to \infty$, such that $x_n(t, \omega)$ has the same finite dimensional distributions as $\xi_n(t)$. This means that the distributions of $f[\xi_n(t)]$ and $f[x_n(t, \omega)]$ are the same. To prove this it is sufficient to show that in terms of their probabilities $f[x_n(t, \omega)] \to f[x_0(t, \omega)]$. According to condition (b) we can choose sequences $c_k \to 0$, $n_k \to \infty$, and $\varepsilon_k \to 0$, such that for $n > n_k$ we have $\mathbf{P}\{\Delta_S(c_k, \xi_n(t)) > \varepsilon_k\} < 1/k^2$. This means that

$$\mathbf{P}\{\Delta_S(c_k, x_n(t, \omega)) > \varepsilon_k\} < 1/k^2. \tag{3.4}$$

Consider a sequence m_k such that $m_k > n_k$ and

$$\mathbf{P}\{\bar{R}(\alpha c_k, x_{m_k}(t, \omega), x_0(t, \omega)) > \varepsilon_k\} < 1/k^2. \tag{3.5}$$

From (2.8) we obtain (see 2.8.1)

$$\Delta_S(c, x_{m_k}(t, \omega)) \leqq L[\Delta_S(c_k; x_{m_k}(t, \omega)) + \bar{R}(\alpha c_k, x_{m_k}(t, \omega), x_0(t, \omega)) \\ + k(2c, x_0(t, \omega)) + \Delta_{\mathbf{J}_1}(2c, x(t, \omega))].$$

Since $\Delta_S(c_k, x_{m_k}(t, \omega))$ and $\bar{R}(\alpha c_k, x_{m_k}(t, \omega))$ approach zero with probability 1 (as follows from (3.4) and (3.5)), we have

$$(3.6)\quad \mathbf{P}\{\overline{\lim_{n\to\infty}} \Delta_S(c, x_{m_k}(t, \omega)) > \varepsilon\} \leqq \mathbf{P}\left\{k(2c, x_0(t, \omega)) + \Delta_{\mathbf{J}_1}(2c, x_0(t, \omega)) > \frac{\varepsilon}{2L}\right\}.$$

On the other hand, from the fact that $x_0(t, \omega) \in K_X$ with probability 1, it follows that

$$\mathbf{P}\{\lim_{c\to 0} k(2c, x_0(t, \omega)) = 0\} = 1$$

and

$$\mathbf{P}\{\lim_{c\to 0} \Delta_{\mathbf{J}_1}(2c, x_0(t, \omega)) = 0\} = 1.$$

Therefore

$$\mathbf{P}\{\lim_{c\to 0} \overline{\lim_{n\to\infty}} \Delta_{\mathbf{J}_1}(2c, x_{m_k}(t, \omega)) = 0\} = 1, \quad \mathbf{P}\{x_{m_k}(t, \omega) \xrightarrow{S} x_0(t, \omega)\} = 1,$$

which means that

$$\mathbf{P}\{\lim_{m_k\to\infty} x_{m_k}(t, \omega) = x_0(t, \omega)\} = 1.$$

In other words, no matter what the sequence $f[x_{n_k}(t, \omega)]$, it always contains a subsequence $f[x_{m_k}(t, \omega)]$ such that $\mathbf{P}\{f[x_{m_k}(t, \omega)] \to f[x_0(t, \omega)]\} = 1$. This, on the other hand, means that the probability of $f[x_n(t, \omega)]$ converges to that of $f[x_0(t, \omega)]$. This proves the theorem.

It is easily seen that conditions (a) and (b) of Theorem 3.2.1 are not necessary for all spaces F. (If F consists, for instance, of only a single point, the distributions of all the $f[\xi_n(t)]$ coincide.) If F is a discrete space, conditions (a) and (b) are also not necessary.

It turns out that these conditions are necessary if F contains a subset which can be mapped continuously onto a line segment. This assertion follows from the following proposition.

3.2.2. Theorem. *If for any continuous bounded functional $f[x(t)]$ (by functional we mean numerical function, i.e., in the present case X is the real line) continuous in the topology S, the distribution of $f[\xi_n(t)]$ converges weakly to the distribution of $f_0[\xi_0(t)]$, then*

(a) *the finite dimensional distributions of $\xi_n(t)$ converge to those of $\xi_0(t)$ for t in some everywhere dense set N of $[0, 1]$ containing 0 and 1, and*

(b) $\lim_{c\to 0} \overline{\lim_{n\to\infty}} \mathbf{P}\{\Delta_S(c, \xi_n(t)) > \varepsilon\} = 0$ *for all* ε.

To prove (a) we show that if $t_1, t_2, \cdots, t_k$ are stochastic continuity points of $\xi_0(t)$ (which means that the probability of $\xi_0(t)$ converges to that of $\xi_0(t_i)$ as $t \to t_i$, $i = 1, \cdots, k$), and if $A_1, A_2, \cdots, A_k \subset X$ are sets such that $\mathbf{P}\{\xi_0(t_i) \in \bar{A}_i \cap \overline{X - A_i}\} = 0$, $i = 1, 2, \cdots, k$, then

$$(3.7)\quad \lim_{n\to\infty} \mathbf{P}\{\xi_n(t_i) \in A_i;\ i = 1, 2, \cdots, k\} = \mathbf{P}\{\xi_0(t_i) \in A_i;\ i = 1, 2, \cdots, k\}.$$

To do this, consider the functional

$$(3.8)\qquad f_\varepsilon^*[x(t)] = \sum_{i=1}^{k} \sup_t g_{t_i,\delta}(t)\, [\lambda(A_i^\varepsilon, x(t)) + 1] / [\lambda(A_i^\varepsilon, x(t)) + 2],$$

where $g_{t_i,\delta}(t)$ is a continuous function never greater than 1, equal to 1 on $(t_i-\delta, t_i+\delta)$, and vanishing outside of $(t_i-2\delta, t_i+2\delta)$; the A_i^ε are the sets of internal points of the A_i whose distance from the boundaries of the A_i is greater than ε, and $\lambda(A, x(t)) = \inf_{y\in A} \rho(x(t), y)$. The functional of (3.8) is continuous in all of the topologies $\mathbf{J}_1, \mathbf{J}_2, M_1, M_2$, since $\lambda(A, x(t))$ gives a continuous mapping of K_X onto K_{R_1} (where R_1 is the real line) in all the topologies. There therefore exist ε_1 arbitrarily small such that

$$\mathbf{P}\left\{f_\varepsilon^*[\xi_n(t)] < \frac{k}{2} + \varepsilon_1\right\} \to \mathbf{P}\left\{f_\varepsilon^*[\xi_0(t)] < \frac{k}{2} + \delta_1\right\}.$$

Since $\sup_t g_{t_i}(t)\,(\lambda(A_i^\varepsilon, x(t)) + 1)/(\lambda(A_i^\varepsilon, x(t)) + 2) \leqq \frac{1}{2}$, it follows that $f_\varepsilon^*[\xi_n(t)] < k/2 + \varepsilon_1$ implies

$$g_{t_i,\delta}(t)\,(\lambda(A_i^\varepsilon, \xi_n(t))+1)/(\lambda(A_i^\varepsilon, \xi_n(t))+2) < \tfrac{1}{2} + \varepsilon_1$$

for all i, and therefore $\lambda(A_i^\varepsilon, \xi_n(t_i)) < \varepsilon_1$ for all i. If $\varepsilon_1 < \varepsilon$, therefore, we have

$$\varliminf_{n\to\infty} \mathbf{P}\{\xi_n(t_i) \in A_i;\ i = 1, 2, \cdots, k\} \geqq \mathbf{P}\left\{f_\varepsilon^*(\xi_0(t)) < \frac{k}{2} + \varepsilon_1\right\}.$$

On the other hand, from the fact that the t_i are stochastic continuity points of $\xi_0(t)$, it follows that the event $\{\lambda(A_i^\varepsilon, \xi_0(t_i)) < \varepsilon_1/k;\ i = 1, 2, \cdots, k\}$ implies the event $\{f_\varepsilon^*[\xi_0(t)] < k/2+2\varepsilon_1\}$ if one neglects a set whose measure can be made arbitrarily small by choosing δ sufficiently small. Since for sufficiently small ε and ε_1, $\mathbf{P}\{\lambda(A_i^\varepsilon, \xi_0(t_i)) < \varepsilon_1;\ i = 1, \cdots, k\}$ is arbitrarily close to $\mathbf{P}\{\xi_0(t_i) \in A_i;\ i = 1, \cdots, k\}$, we have

$$\mathbf{P}\left\{f_\varepsilon^*(\xi_0(t)) < \frac{k}{2} + \varepsilon_1\right\} \geqq \mathbf{P}\{\xi_0(t_i) \in A_i;\ i = 1, 2, \cdots, k\} - \mu,$$

where μ is an arbitrarily small number. Therefore

$$\varliminf_{n\to\infty} \mathbf{P}\{\xi_n(t_i) \in A_i;\ i = 1, 2, \cdots, k\} \geqq \mathbf{P}\{\xi_0(t_i) \in A_i;\ i = 1, 2, \cdots, k\}. \tag{3.9}$$

Since $X = A_i \cup CA_i$ (where CA_i is the complement of A_i) and the sets CA_i have the same properties as $A_i(\overline{A}_i \cap \overline{X-A_i} = \overline{CA}_i \cap \overline{X-CA_i})$, it is not difficult to obtain (3.7) from (3.9).

Let us demonstrate this for $k = 2$. We have the inequalities

$$\begin{aligned}
&\varliminf_{n\to\infty} \mathbf{P}\{\xi_n(t_1) \in A_1, \xi_n(t_2) \in A_2\} \geqq \mathbf{P}\{\xi_0(t_1) \in A_1, \xi_0(t_2) \in A_2\},\\
&\varliminf_{n\to\infty} \mathbf{P}\{\xi_n(t_1) \in A_1, \xi_n(t_2) \in CA_2\} \geqq \mathbf{P}\{\xi_0(t_1) \in A_1, \xi_0(t_2) \in CA_2\},\\
&\varliminf_{n\to\infty} \mathbf{P}\{\xi_n(t_1) \in CA_1, \xi_n(t_2) \in A_2\} \geqq \mathbf{P}\{\xi_0(t_1) \in CA_1, \xi_0(t_2) \in A_2\},\\
&\varliminf_{n\to\infty} \mathbf{P}\{\xi_n(t_1) \in CA_1, \xi_n(t_2) \in CA_2\} \geqq \mathbf{P}\{\xi_0(t_1) \in CA_1, \xi_0(t_2) \in CA_2\}.
\end{aligned} \tag{3.10}$$

If (3.7) were not true, there would be a sequence n_k such that

$$\lim_{n_k} \mathbf{P}\{\xi_{n_k}(t_1) \in A_1, \xi_{n_k}(t_2) \in A_2\} > \mathbf{P}\{\xi_0(t_1) \in A_1, \xi_0(t_2) \in A_2\}.$$

If we were to add to this inequality the last three of (3.10), we would obtain $1 > 1$. Thus (a) is proved.

The proof of (b) is almost the same for the different topologies. Let us go through it for the case of $S = \mathbf{J}_1$. Let

$$\Delta^{(m)}_{\mathbf{J}_1}(c, x(t)) = \max_{t_1<t<t_2} (g_c(t_1, t, t_2))^m \{\min [\rho(x(t_1), x(t)); \rho(x(t), x(t_2))]\},$$

and let $g_c(t_1, t, t_2)$ be a continuous function of the arguments t_1, t, t_2 such that

$$g_c(t_1, t, t_2)\begin{cases} = 1 & \text{for } t-c < t_1 < t < t_2 < t+c, \\ < 1 & \text{for the other } t_1, t, t_2. \end{cases}$$

For every m and c the functional $\Delta^{(m)}_{\mathbf{J}_1}(c, x(t))$ is $\mathbf{J}_1$-continuous, and $\Delta^{(m)}_{\mathbf{J}_1}(c, x(t)) > \Delta_{\mathbf{J}_1}(c, x(t))$. Therefore

$$\overline{\lim_{n\to x}}\, \mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_n(t)) > \varepsilon\} \leqq \mathbf{P}\{\Delta^{(m)}_{\mathbf{J}_1}(c, \xi_0(t)) > \varepsilon\}.$$

Since $\lim_{m\to\infty} \Delta^{(m)}_{\mathbf{J}_1}(c, x(t)) = \Delta_{\mathbf{J}_1}(c, x(t))$, we have

$$\overline{\lim_{n\to\infty}}\, \mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_n(t)) > \varepsilon\} \leqq \mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_0(t)) > \varepsilon/2\}.$$

It remains to be shown that

$$\lim_{c\to 0} \mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_0(t)) > \varepsilon\} = 0.$$

This, however, follows from the fact that for all $x(t)$ in K_X we have

$$\lim_{c\to 0} \Delta_{\mathbf{J}_1}(c, x(t)) = 0.$$

This proves the theorem.

If we review the proof of Theorem 3.2.1, we see that it is not necessary to require that $f[x(t)]$ be continuous. It would be sufficient to use $f[x(t)]$ for which $\xi_n(t) \xrightarrow{S} \xi_0(t)$ with probability 1 implies $f[\xi_n(t)] \to f[\xi_0(t)]$ with probability 1. Let μ_{ξ_0} be the measure on K_X of the process $\xi_0(t)$. Then such functions $f[x, t]$ will be functions continuous in the topology S almost everywhere in terms of the measure μ_{ξ_0}. We then arrive at the following theorem.

3.2.3. Theorem. *Conditions* (*a*) *and* (*b*) *of* 3.2.1 *imply that for every function f defined on K_X whose values lie in F, which is continuous in the topology S almost everywhere in terms of the measure μ_{ξ_0}, the distribution of $f[\xi_n(t)]$ converges weakly to the distribution of $f[\xi_0(t)]$ as $n \to \infty$.*

3.2.4. Corollary. *Let K be a set of K_X open in the topology S, such that* $\mathbf{P}\{\xi_0(t) \in K'_S\} = 0$ (*where K'_S is the boundary of K in the topology S*). *Then condition* (a) *and* (b) *of* 3.2.1 *imply that*

$$\lim_{n\to\infty} \mathbf{P}\{\xi_n(t) \in K\} = \mathbf{P}\{\xi_0(t) \in K\}.$$

3.2.5. Remark. If the space F contains an inverse image of the line segment, the convergence of the distribution of $f[\xi_n(t)]$ to that of $f[\xi_0(t)]$ for all f in F_S implies the convergence of the distribution of $f[\xi_n(t)]$ to that of $f[\xi_0(t)]$ for all f which are S-continuous almost everywhere in terms of the measure μ_{ξ_0}.

3.2.6. Remark. The extension of Theorem 3.2.1 to an almost everywhere continuous functional increases significantly the region of its applicability. Thus, of all the examples we have considered in 2.2.6—2.2.8, only $m_{[t_1, t_2]}[x(t)]$

and $M_{[t_1, t_2]}[x(t)]$ are continuous in U. Nevertheless all these functionals are almost everywhere continuous in the appropriate topologies (see 2.2.9—2.2.12) so long as t_1 and t_2 are stochastic continuity points of $\xi_0(t)$, and a and b are such that at a local extremum $\xi_0(t)$ differs from a and b with probability 1.

3.3. Let us now go on to a consideration of the convergence conditions for functionals continuous in the uniform topology. Let F_U be the set of functions defined on K_X, which take their values from F (some complete metric separable space) and are measurable with respect to the Borel closure of all cylindrical sets. (This last requirement does not in any sense follow from continuity in the topology U, although it would for other topologies considered in this article.)

3.3.1. *Assume the following two conditions to be true.*

(a) *Let ε be such that the probability is 1 that $\xi_0(t)$ has no discontinuities equal to ε; let $\tau_{n,1}^{(\varepsilon)}$, $\tau_{n,2}^{(\varepsilon)}, \cdots, \tau_{n,\nu_n^\varepsilon}^{(\varepsilon)}$ be points where the discontinuities of $\xi_n(t)$ are greater than ε, and let*

$$\mathbf{P}\{\xi_n(t) \in A_1, \cdots, \xi_n(t_k) \in A_k;\ \tau_{n,1}^{(\varepsilon)} \in B_1, \cdots, \tau_{n,m}^{(\varepsilon)} \in B_m;\ \nu_n^\varepsilon = m\} \tag{3.11}$$

approach, as $n \to \infty$,

$$\mathbf{P}\{\xi_0(t_1) \in A_1, \cdots, \xi_0(t_k) \in A_k;\ \tau_{0,1}^{(\varepsilon)} \in B_m, \cdots, \tau_{0,m}^{(\varepsilon)} \in B_m;\ \nu_0^\varepsilon = m\} \tag{3.12}$$

for all $t_1, \cdots, t_k \in [0, 1]$, all Borel sets A_i in X such that

$$\mathbf{P}\{\xi_0(t_i) \in \bar{A}_i \cap X - A_i\} = 0,$$

and all Borel sets $B_1, B_2, \cdots, B_m$ on $[0, 1]$ so that the absolute difference between (3.11) *and* (3.12) *is equal to or less than some function $\lambda_n(t_1, \cdots, t_k, A_1, \cdots, A_k)$ which approaches zero as $n \to \infty$ with fixed $t_1, \cdots, t_k, A_1, \cdots, A_k$;*

(b) $$\lim_{c\to 0} \lim_{n\to\infty} \mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_n(t)) > h\} = 0 \text{ for } all\ h < 0.$$

If both of these conditions are fulfilled then for any function $f \in F_U$ the distribution of $f(\xi_n(t))$ converges to that of $f(\xi_0(t))$.

Proof. Let us first assume that f is uniformly continuous in U, that is, for every ε there exists a δ such that

$$r_F(f(x(t)), f(y(t))) < \varepsilon,$$

if $\sup_t \rho(x(t), y(t)) < \delta$.

If the theorem is not true for f, there exists a $G \subset F$ such that

$$\mathbf{P}\{f(\xi_0(t)) \in \bar{G} \cap \overline{F-G}\} = 0, \tag{3.13}$$

but $\mathbf{P}\{f(\xi_n(t)) \in G\}$ does not approach $\mathbf{P}\{f(\xi_0(t)) \in G\}$. One can therefore find $a > 0$ and subsequences n_k such that

$$|\mathbf{P}\{f(\xi_{n_k}(t)) \in G\} - \mathbf{P}\{f(\xi_0(t)) \in G)\}| > a.$$

Let, for instance,

$$\mathbf{P}\{f(\xi_{n_k}(t)) \in G\} > \mathbf{P}\{f(\xi_0(t)) \in G\} + a.$$

We define the random processes $\xi_n^*(t)$ as follows. Choose points $0 = t_0 < t_1 < \cdots < t_N = 1$, and set $\xi_n^*(t_i) = \xi_n(t_i)$ if $t_i \leqq t < t_{i+1}$ and if between t_i and t_{i+1} there are either no points $\tau_{n,k}^{(\varepsilon)}$ or no less than two points; if, on the other hand, in (t_i, t_{i+1}) there lies one point $\tau_{n,j}^{(\varepsilon)}$, we set

$$\xi_n^*(t) = \begin{cases} \xi_n(t_i) & \text{for } t_i \leqq t < \tau_{n,j}^{(\varepsilon)}, \\ \xi_n(t_{i+1}) & \text{for } \tau_{n,j}^{(\varepsilon)} \leqq t < t_{i+1}. \end{cases}$$

It is not difficult to show that

$$\mathbf{P}\{\sup_t \rho(\xi_n^*(t), \xi_n(t)) > 8\varepsilon\} \leqq \mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_n(t)) > \varepsilon\}, \tag{3.14}$$

where $c = \max_{i=1,\cdots,N-1} (t_{i+1} - t_i)$.

Therefore if $f(\xi_n(t)) \in G$, we have $f(\xi_n^* \in G_\delta^+)$, where G_δ^+ is the set of all points whose distance from G is no greater than δ, and where δ is such that

$$r_F(f(y(t)), f(x(t))) < \delta \quad \text{for } \sup_t \rho(x(t), y(t)) < 8\varepsilon$$

if $\Delta_{\mathbf{J}_1}(c, \xi_n(t)) \leqq \varepsilon$.

On the other hand, if $f(\xi_0^*(t)) \in G_\delta^-$, we have $f(\xi_0(t)) \in G$, where G_δ^- is the set of all points whose distance from $F-G$ is no less than δ, so long as $\Delta_{\mathbf{J}_1}(c, \xi_0(t)) \leqq \varepsilon$.

Therefore

$$\mathbf{P}\{f(\xi_{n_k}^*(t)) \in G_\delta^+\} > a+\mathbf{P}\{f(\xi_0^*(t)) \in G_\delta^-\} -\mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_{n_k}(t)) > \varepsilon\}-\mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_0(t)) > \varepsilon\}.$$

Further,

$$\mathbf{P}\{f(\xi_{n_k}^*(t)) \in G_\delta^+\} > a+\mathbf{P}\{f(\xi_0^*(t)) \in G_{2\delta}^+\}-\mathbf{P}\{f(\xi_0^*(t)) \in G_{2\delta}^+-G_\delta^-\} -\mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_0(t)) > \varepsilon\}-\mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_{n_k}(t)) > \varepsilon\}. \tag{3.15}$$

Since $G_{2\delta}^+-G_\delta^-$ describes, as $\delta\to 0$, a monotonically decreasing sequence of sets converging to $\bar{G} \cap \overline{F-G}$, it follows from (3.14) that

$$\varlimsup_{\delta\to 0} \mathbf{P}\{f(\xi_0^*(t)) \in G_{2\delta}^+-G_\delta^-\} \leqq \mathbf{P}\{\Delta_{\mathbf{J}_1}(c, \xi_n(t)) > \varepsilon\}.$$

Therefore by choosing ε, δ, and c sufficiently small, all the negative terms in (3.15) can be made so small that their sum is less than $a/2$. Then

$$\mathbf{P}\{f(\xi_{n_k}^*(t)) \in G_\delta^-\} > a/2+\mathbf{P}\{f(\xi_0^*(t)) \in G_{2\delta}^+\}. \tag{3.16}$$

Obviously the distribution of $\xi_n^*(t)$ is uniquely defined by the probabilities

$$\mathbf{P}\{\xi_n(t_i) \in A_i;\ i = 0, 1, \cdots, N;\ \tau_{n,j}^{(\varepsilon)} \in B_j;\ j = 1, 2, \cdots, m;\ \nu_n^\varepsilon = m\}.$$

Consider sets A_i^r, $i = 0, 1, \cdots, N$; $r = 1, 2, \cdots, R$, such that

$$\begin{gathered}\mathbf{P}\{\xi_0(t_i) \in \overline{A_i^r} \cap \overline{x-A_i^r}\} = 0,\\ \sum \mathbf{P}\{\xi_0(t_i) \in A_i^r;\ i = 1, \cdots, R\} > 1-a/4,\end{gathered} \tag{3.17}$$

where the sum is taken over all possible different collections $r_1, r_2, \cdots, r_N$, in which each r_i takes on one of the values $1, \cdots, R$ independent of the others. In addition, assume that the diameter of each of the A_i^r is no greater than μ, where μ is a number such that

$$r_F(f(x(t)), f(y(t))) < \delta/4 \quad \text{if} \quad \sup_t \rho(x(t), y(t)) < \mu.$$

Let us now construct new processes $\bar{\xi}_n^*(t)$ in the following way. From each of the A_i^r we choose one point $x_i^{(r)}$ and let y be any point different from all of these $x_i^{(r)}$. We set

$$\bar{\xi}_n^*(t_i) = \begin{cases} x_i^{(r)}, & \text{if} \quad \xi_n^*(t_i) \in A_i^r,\\ y, & \text{if} \quad \xi_n^*(t_i) \notin \bigcup_r A_i^r,\end{cases}$$

and the discontinuity points of $\bar{\xi}_n^*(t)$ coincide with those of $\xi_n^*(t)$, and we set $\bar{\xi}_n^*(t)$ equal to a constant where $\xi_n^*(t)$ is constant. Then it follows from the choice of μ, (3.16), and (3.17) that

$$\mathbf{P}\{f(\bar{\xi}_{n_k}(t)) \in G_{5\delta/4}^+\} > \frac{a}{4} + \mathbf{P}\{f(\bar{\xi}_0^*(t) \in G_{7\delta/4}^+\}. \tag{3.18}$$

There exists therefore at least one m and a collection of points $x_0^{(r_0)}, x_1^{(r_1)}, \cdots, x_N^{(r_N)}$ for which

$$\begin{aligned}\varlimsup_{n\to\infty} \mathbf{P}\{f(\bar{\xi}_n^*(t)) \in G_{5\delta/4}^+,\ \ \nu_n^\varepsilon = m,\ \ \bar{\xi}_n^*(t_i) = x_i^{(r_i)};\ \ i = 1, \cdots, N\}\\ > \mathbf{P}\{f(\bar{\xi}_0^*(t)) \in G_{7\delta/4}^+,\ \ \nu_0^\varepsilon = m,\ \ \bar{\xi}_0^*(t_i) = x_i^{(r_i)};\ \ i = 1, \cdots, N\},\end{aligned}$$

which contradicts (a), since it implies that the variation of

$$\mathbf{P}\{\nu_n^\varepsilon = m,\ \tau_{n,1}^{(\varepsilon)} \in B_1, \cdots, \tau_{n,m}^{(\varepsilon)} \in B_m,\ \xi_n(t_0) \in A_0^{r_0}, \cdots, \xi_n(t_N) \in A_N^{r_N}\}$$

converges, for fixed $m, A_1^{r_1}, \cdots, A_N^{r_N}$ to

$$\mathbf{P}\{\nu_0^\varepsilon = m,\ \tau_{0,1}^{(\varepsilon)} \in B_1, \cdots, \tau_{0,m}^{(\varepsilon)} \in B_m,\ \xi_0(t_0) \in A_0^{r_0}, \cdots, \xi_0(t_N) \in A_N^{r_N}\}.$$

This completes the proof of the theorem for uniformly continuous functions. The transition to arbitrary continuous functions, and even to almost everywhere continuous functions in the measure generated by $\xi_0(t)$ can be performed in exactly the same way as in one of the author's previous works [7] if we note that the distance in the uniform metric from a measurable set is a measurable uniformly continuous function.

3.3.2. REMARK. Condition (a) can be replaced by the following condition: (a′) The variation of the distribution of the $\tau^{(\varepsilon)}_{n,1}, \cdots, \tau^{(\varepsilon)}_{n,m}$ for $\nu^{\varepsilon}_n = m$ converges to that of the $\tau^{(\varepsilon)}_{0,1}, \cdots, \tau^{(\varepsilon)}_{0,m}$ for $\nu^{\varepsilon}_0 = m$; the joint distribution of the $\tau^{(\varepsilon)}_{n_N}, \nu^{(\varepsilon)}_n, \xi_n(t_1), \cdots, \xi_n(t_k)$ converges to the joint distribution of the $\tau^{(\varepsilon)}_{0,j}, \nu^{\varepsilon}_0, \xi_0(t_1), \cdots, \xi_0(t_k)$, if $t_1, \cdots, t_k$ belong to some set everywhere dense on [0,1] which contains 0 and 1.

3.3.3. REMARK. Condition (a′) in 3.3.2 and condition (b) in 3.3.1 are necessary and sufficient for Theorem 3.3.1 if F is a segment of the line. The necessity of (b) follows from 3.2.2. The necessity of (a′) can be shown by considering functionals of the form

$$f(\xi_n(t)) = g(\xi_n(t_1), \xi_n(t_2), \cdots, \xi_n(t_k); \tau^{(\varepsilon)}_{n,1}, \tau^{(\varepsilon)}_{n,2}, \cdots, \tau^{(\varepsilon)}_{n,m}),$$

where g is continuous in the first set of variables and Borel measurable in the second set.

All the limit theorems obtained can be applied to various concrete cases. The author hopes in the near future to publish applications of these theorems to processes with independent increments and to Markov processes.

Received by the editorial office
October 10, 1956.

REFERENCES

1. A. N. KOLMOGOROV, Izv. Akad. Nauk SSSR. Ser. Mat. Estestv. Nauk, (1931), p. 959.
2. A. N. KOLMOGOROV, Izv. Akad. Nauk SSSR. Ser. Mat. Estestv. Nauk, (1933), p. 363.
3. M. DONSKER, Mem. Amer. Math. Soc., 6 (1951).
4. I. I. GIKHMAN, Kiiv. Derzh. Univ. Mat. Sb., (1953), pp. 7, 75.
5. YU. V. PROKHOROV, Uspekhi Mat. Nauk, (1953), pp. 8, 165.
6. YU. V. PROKHOROV, Teor. Veroyatnost. i Primenen., 1: 2 (1956).
7. A. V. SKOROKHOD, Dokl. Akad. Nauk SSSR., 104: 3 (1955).
8. A. V. SKOROKHOD, Dokl. Akad. Nauk SSSR., 106: 5 (1956).
9. N. N. CHENTSOV, Teor. Veroyatnost. i Primenen., 1: 1 (1956).
10. W. FELLER, Trans. Amer. Math. Soc., 77: 1 (1954).

LIMIT THEOREMS FOR STOCHASTIC PROCESSES

A. V. SKOROKHOD (*MOSCOW*)

(*Summary*)

Let us consider a sequence of processes $\xi_n(t)$ such that the multivariate distribution of $\xi_n(t_1), \xi_n(t_2), \cdots, \xi_n(t_k)$ tends to the multivariate distribution of $\xi_0(t_1), \xi_0(t_2), \cdots, \xi_0(t_k)$ for all k and $t_1, t_2, \cdots, t_k$.

Let f be the functional for which $f(\xi_n(t))$ are determined with a probability of 1, the latter being random variables (i.e. those that have probability distributions).

This paper contains several sufficient conditions, for which the distributions of $f(\xi_n(t))$ tend to the distribution of $f(\xi_0(t))$ as $n \to \infty$.

Let K be the space of all functions not having discontinuities higher than simple jumps, and let us assume that $\xi_n(t)$ with a probability of 1 is in K.

Several topologies in K are defined. The necessary and sufficient conditions are found for all functionals f that are continuous in these topologies for which the distribution of $f(\xi_n(t))$ tends to the distribution of $f(\xi_0(t))$.

The results are demonstrated in the example of topology $\mathbf{J}_1$ which is defined as follows.

The sequence $x_n(t)$ tends to $x_0(t)$ in topology $\mathbf{J}_1$ if there exists a sequence of monotonic continuous functions $\lambda_n(t)$ for which

$$\lambda_n(0) = 0,\ \lambda_n(1) = 1,\ \lim_{n\to\infty} \sup_t |\lambda_n(t) - t| = 0,$$

$$\lim_{n\to\infty} \sup_t |x_n(\lambda_n(t)) - t| = 0.$$

Theorem. *The distribution of* $f(\xi_n(t))$ *tends to the distribution of* $f(\xi_0(t))$ *for all* f *that are continuous in topology* $\mathbf{J}_1$, *if and only if*

a) *the multivariate distribution of* $\xi_n(t_1), \cdots, \xi_n(t_k)$ *tends to the multivariate distribution of* $\xi_0(t_1), \cdots, \xi_0(t_k)$ *for all* k, *and* $t_1, t_2, \cdots, t_k$ *from some set* N *that is dense on* $[0, 1]$.

b) *for all* $\varepsilon > 0$

$$\lim_{c\to 0} \overline{\lim_{n\to\infty}} \mathrm{P}\{\sup_{t-c<t_1<t<t_2<t+c} \min [|\xi_n(t_1)-\xi_n(t)|;\ |\xi_n(t)-\xi_n(t_0)|] > \varepsilon\} = 0.$$