The Annals of Probability
1975, Vol. 3, No. 4, 573–596

A PREDICTIVE VIEW OF CONTINUOUS TIME PROCESSES[1]

BY FRANK B. KNIGHT

University of Illinois at Champaign—Urbana

DEDICATED TO PROFESSOR J. L. DOOB ON HIS SIXTY-FIFTH BIRTHDAY.

Let $X(t)$, $0 \leqq t$, be an $\mathscr{L} \times \mathscr{F}$-measurable process on $(\Omega, \mathscr{F}, P)$ with state space $(E, \mathscr{E})$, where $\mathscr{L}$ is the Lebesgue σ-field and $\mathscr{E}$ is countably generated. Let $\mathscr{F}(t_1, t_2)$, $0 \leqq t_1 < t_2 \leqq \infty$, be the σ-field generated by $\{\int_{t_1}^{t} f(X(s))\, ds, t_1 < t < t_2, 0 \leqq f \in \mathscr{E}\}$. A new process $Z(t)$ is constructed whose values consist of conditional probabilities in the wide sense over $\mathscr{F}(t, \infty)$ given $\mathscr{F}(0, t+)$. It is shown that $Z(t)$ is a homogeneous strong-Markov process on a compact metric space, with right-continuous paths having left limits for $t > 0$. $Z(t)$ determines $X(t)$ wp 1 except for t in a Lebesgue-null set. We call $Z(t)$ the prediction process of $X(t)$. Some general properties of the construction are developed, followed by two applications.

0. Introduction. Let $(E, \mathscr{E})$ be an arbitrary set and a countably generated σ-field of subsets. A continuous time stochastic process with state space $(E, \mathscr{E})$ is, as usual, a collection of E-valued random variables $X(t)$, $t \in I$, on a complete probability space $(\Omega, \mathscr{F}, P)$, where I is an interval (finite or infinite) of the real line R. From an empirical viewpoint, however, it must be allowed that this concept is something of an artifice, if not even selfcontradictory. Thus if $X(t)$ is to have any operational meaning for individual t, without reference to anything else, we must admit the possibility of a continuum of separate and instantaneous observations, and no amount of ingenuity seems likely to suggest a procedure for carrying these out. Something further is clearly needed to prevent the process from disintegrating into an uncountable number of discrete instants.

This difficulty is present, of course, in deterministic as well as in stochastic processes. In the former case, however, it is always apparent that the intended trajectory has some additional regularity property, such as continuity in some topology on E, which reduces the problem of making observations to a sequence of approximations, each member of which requires a non-zero length of time for its accomplishment. In the stochastic case it is usually not obvious from the finite-dimensional distributions alone that the paths can be chosen to have such a property.

A first step out of this difficulty is the well-known idea of separability due to Doob (1953). If $X(t)$ is real-valued, this yields a standard modification of $X(t)$

Received June 29, 1974; revised December 12, 1974.

[1] This work was supported in part by the National Science Foundation under grant GP-28573X.

AMS 1970 *subject classifications.* Primary 6040, 6060; Secondary 6062.

Key words and phrases. Continuous time processes, prediction, strong Markov processes, conditional probabilities in the wide sense.

(a process equal to $X(t)$ with probability 1 at each t) for which determination of lim sup's and lim inf's of $X(t)$ at all t is reduced to determination of $X(t)$ for t in a countable set. But this alone does not seem to solve the problem for two reasons. First, determination of $X(t)$ is still instantaneous, contrary to all usual processes of measurement. Second, the significance of lim sup's and lim inf's hinges largely upon the presence of some additional regularity property to make these values representative of the behavior of the paths at more than a countable dense set of times.

Another aspect of separability, however, is that it leads to the existence of measurable modifications of the process, and this feature is essential to the present approach. If, for example, $(E, \mathscr{E})$ is a locally compact space with countable base, and $X(t)$ is continuous in probability at all t except for a Lebesgue-null set, then there is always $(\mathscr{L} \times \mathscr{F})$-completion-measurable standard modification [9: Chapter 4, 2, Theorem 1 and Remark 1] where $(\mathscr{L}, L)$ denotes the Lebesgue σ-field and measure on I, and the completion is with respect to the measure $L \times P$. For a general state space, we will take the point of view that such measurability is axiomatic. Thus we assume throughout that for every $A \in \mathscr{E}$, $\{(t, w) : X(t, w) \in A\} \in \mathscr{L} \times \mathscr{F}$ up to an $L \times P$-null set. We shall also assume for convenience, except when stated otherwise, that $I = R^+ = [0, \infty)$. Using the notation $f \in \mathscr{L} \times \mathscr{E}$ for $\mathscr{L} \times \mathscr{E}$-measurability, we recall that for $f \in \mathscr{L} \times \mathscr{E}$ either positive or bounded we have $\int_0^t f(s, X(s))\, ds \in \mathscr{F}$ for each t. This is clear for $f = I_B I_C$, $B \in \mathscr{L}$, $C \in \mathscr{E}$, by Fubini's Theorem, and the general case follows by a standard extension procedure. The starting point of the present synthesis is to disregard the process $X(t)$ for individual t, and to consider those sets of $\mathscr{F}$ obtainable in terms of such integrals of the process. Stated differently, this means that we regard two paths of $X(t)$ as equivalent if they differ at most on a Lebesgue-null set of t. Indeed, it is not difficult to show that the atoms of the σ-subfield of $\mathscr{F}$ generated by all such integrals are precisely the atoms generated by this equivalence relation (in particular, the latter are elements of $\mathscr{F}$). The mathematical advantages of this approach, such as the fact that this subfield is separable, will gradually become apparent. It is hoped that the above remarks will suffice to provide also some heuristic justification.

A fundamental question which immediately arises is how to interpret $X(t)$ as a stochastic process without any essential loss of information for individual t, but using only its indefinite integrals. The answer which we give to this is in fact the basis for the present paper. Stated very briefly, it is that we replace $X(t)$ by the conditional distribution over the entire future after time t given the entire past up to time $t+$. At first glance this admittedly looks quite awkward, but with practice it becomes much less so, and there are unexpected benefits. One of these is that sets contained in the intersection of the past up to $t+$ and the future after t have conditional probability 0 or 1. Thus, to the extent that $X(t)$ is measurable over this intersection, nothing is lost in the reformulation. On the other hand, it turns out that the new process may be viewed as having right-continuous paths with left limits in a compact metric space. There is some non-uniqueness involved in the choice of this topology, but essentially none with regard to the process itself. Perhaps even more surprising is the fact that the new process is always a homogeneous strong-Markov process, irrespective of any Markov property for the original process $X(t)$. The theory of such precesses is quite highly developed (Blumenthal and Getoor (1968), Dynkin (1965), Meyer (1967)), and the present result means that this theory is placed at our disposal in studying an arbitrary measurable process $X(t)$.

In the present paper we develop the general method, setting up the correspondence between measurable processes and what will be called their "prediction processes," as outlined above (the term derives, of course, from the fact that the conditional probabilities also define the best least squares non-linear predictors of future based on past). Some theoretical uniqueness and stability properties of the prediction process are given (Section 2) which may assist the reader to overcome some of the conceptual difficulties involved, and also to alleviate a natural degree of skepticism as to the intrinsic merit of the construction.

By way of application (Section 3) we limit ourselves to two general cases. If $X(t)$ is a right-continuous, homogeneous, strong-Markov process to begin with, we show that the prediction process is completely equivalent with $X(t)$ if and only if the resolvent of $X(t)$ takes continuous functions into continuous functions (i.e. $X(t)$ is a Ray process [17], [18]). More generally, we find that if $X(t)$ is Markovian and time homogeneous, its prediction process corresponds to a Ray compactification of the original process, as carried out in [12]. In the second general case it is shown that if $X(t)$ is stationary and Gaussian with mean 0 and continuous covariance then its prediction process has continuous paths. In other words, even the paths of $X(t)$ are highly discontinuous, the predictive aspect of the process varies continuously. We conclude with two counterexamples. One shows that if the stationarity assumption is omitted, then the prediction process may have fixed discontinuities. The other illustrates the possibility of totally inaccessible discontinuities of the prediction process, even if $X(t)$ (non-Gaussian) has continuous paths.

1. Construction of the prediction processes.

1.1. *The sojourn time process.* Let $0 \leqq h_n(x) \leqq 1$, $n = 1, 2, \cdots$, be a fixed sequence of $\mathscr{E}$-measurable functions generating all of $\mathscr{E}$. For example, we can choose the indicator functions of a sequence of sets generating $\mathscr{E}$, but if E has a topology it will sometimes be more convenient to choose $(h_n) \in C(E)$. We introduce

DEFINITION 1.1.1. The process $Y(t) = (Y_n(t)) = (\int_0^t h_n(X(s))\, ds)$ is called the sojourn time process of $X(t)$ relative to (h_n). For $0 \leqq t_1 < t_2 \leqq \infty$, let $\mathscr{F}(t_1, t_2)$ be the σ-subfield of $\mathscr{F}$ generated by $\{Y_n(t) - Y_n(t_1), t_1 < t < t_2, n \geqq 1\}$. We call $\mathscr{F}(0, t+) = \bigcap_{\varepsilon>0} \mathscr{F}(0, t+\varepsilon)$ the past of Y at time t, and $\mathscr{F}(t, \infty)$ the future of Y at t. For each $\mathscr{F}(0, t+)$-stopping time T, let $\mathscr{F}(T, T+t)$, $0 < t \leqq \infty$, denote the σ-field generated by $\{Y_n(T+s) - Y_n(T), 0 < s < t, n \geqq 1\}$.

All of our considerations of $X(t)$ will be through the medium of $Y(t)$, as a process with countably many components. It will be shown in Section 2 that in a quite strong sense the results do not depend on the choice of (h_n). At present, as a first indication of what is involved, we have

THEOREM 1.1.1. *$\mathscr{F}(t_1, t_2)$ is generated by the family of integrals $\{\int_{t_1}^t f(s, X(s))\, ds$, $t_1 < t < t_2$, $0 \leqq f \in \mathscr{L} \times \mathscr{E}\}$, hence it does not depend on the choice of (h_n). For any pair $w_1, w_2 \in \Omega$, $Y(s, w_1) = Y(s, w_2)$ $0 \leqq s \leqq t$, if and only if $X(s, w_1) = X(s, w_2)$ for all $s \leqq t$ except for a set of Lebesgue measure 0.*

PROOF. If we consider (h_n) as a function $E \to \times_{n=1}^{\infty} [0, 1]$, if follows from [1: 0.2, Proposition 2.7] that any $0 \leqq f \in \mathscr{E}$ has the form $f = g(h_1, h_2, \cdots)$ where g is a measurable function on $\times_{n=1}^{\infty} [0, 1]$. Let $W_n(t) = \limsup_{\Delta t \to 0+} (1/\Delta t) \int_t^{t+\Delta t} h_n(X(s))\, ds$, $n \geqq 1$. Then for $t_1 < t < t_2$, $W_n(t) \in \mathscr{F}(t_1, t_2)$, and moreover $W_n(t) = h_n(X(t))$ except on a t-set of Lebesgue measure 0 for each $w \in \Omega$. In fact, since $W_n(t)$ is $\mathscr{L} \times \mathscr{F}(t_1, t_2)$-measurable on (t_1, t_2) we have

$$\int_{t_1}^t f(X(s))\, ds = \int_{t_1}^t g(h_1(W_1(s)), h_2(W_2(s)), \cdots)\, ds \in \mathscr{F}(t_1, t_2)\,.$$

The case of general $f(s, x) \in \mathscr{L} \times \mathscr{E}$ follows easily by approximations, starting with $f(s, x) = I_B, I_C$, $B \in \mathscr{B}(t_1, t_2)$ (the Borel subsets of (t_1, t_2)) and $C \in \mathscr{E}$. The second assertion of the theorem is an obvious consequence of the first.

Still another way of describing the information carried by the process $Y(t)$ is to define it in terms of the *sojourn time measures*

$$\mu(t, A) = \int_0^t I_A(X(s))\, ds\,, \quad A \in \mathscr{E}\,.$$

COROLLARY 1.1.1. *$\mathscr{F}(t_1, t_2)$ is generated by the family $\{\mu(t, A) - \mu(t_1, A)$, $t_1 < t < t_2$, $A \in \mathscr{E}\}$.*

PROOF. Immediate from the preceding.

1.2. *A path space for the sojourn time processes.* The paths of $Y(t)$ are quite regular, having non-decreasing components which satisfy the Lipshitz condition $Y_n(t+s) - Y_n(t) \leqq s$, and the σ-fields $\mathscr{F}(t_1, t_2)$ are obviously separable. ($\mathscr{F}(0, t+)$, however, is in general not separable, which leads to certain difficulties below.) We need to use regular conditional probabilities over $\mathscr{F}(0, \infty)$ or its equivalent, and there is a question of their existence unless the range of $Y(t)$ is suitably measurable. To circumvent this, we shall systematically replace $Y(t)$ by the process (or measure) induced by $Y(t)$ in the space of all paths which share the above properties. This amounts to using what are called by Doob (1953) "conditional probability distributions in the wide sense," for $Y(t)$. It also has the advantage of providing a space independent of $X(t)$, and hence the possibility of a more unified treatment.

DEFINITION 1.2.1. Let Ω' denote the set of all paths $y(t) = (y_n(t), n = 1, 2, \cdots;$ $0 \leqq t)$ such that $y_n(0) = 0$ and $0 \leqq y_n(t+s) - y_n(t) \leqq s$ for all $s > 0$. Let

$\mathscr{F}'(t_1, t_2)$, $0 \leqq t_1 < t_2 \leqq \infty$, denote the σ-field generated by $(y_n(t) - y_n(t_1)$, $n = 1, 2, \cdots; t_1 < t < t_2)$, and let $\mathscr{F}'(0, t+) = \bigcap_{\varepsilon>0} \mathscr{F}'(0, t + \varepsilon)$. For every $\mathscr{F}'(0, t+)$ stopping time $T' < \infty$, let $\mathscr{F}'(T', T' + t)$ denote the σ-field generated by $\{(y_n(T' + s) - y_n(T')), 0 < s < t\}$ for $0 < t \leqq \infty$.

We shall introduce the natural topology in Ω', in such a way that $\mathscr{F}'(0, \infty)$ is the Borel σ-field. Let $(r_{n,j}, 1 \leqq n, 1 \leqq j)$ be an enumeration of the positive rationals for each n. We can consider Ω' as a subset of the compact product space $R_\infty = \bigtimes_{n,j=1}^{\infty} [0, r_{n,j}]$ by identifying $y_n(r_{n,j})$ with the component in $[0, r_{n,j}]$, and it is obvious that this mapping is one-to-one.

THEOREM 1.2.1. *The topology of Ω' induced by R_∞ is that of uniform convergence in finite time intervals for each component $y_n(t)$. In this way, Ω' is a compact metrizable space and $\mathscr{F}'(0, \infty)$ are the Borel subsets.*

PROOF. Immediate from the definitions.

THEOREM 1.2.2. *For any pair of processes $(X(\cdot), Y(\cdot))$ as above and $0 \leqq t_1 < t_2 \leqq \infty$, $\mathscr{F}(t_1, t_2)$ coincides with all sets of the form $\{w \in \Omega : Y(\cdot) \in S'\}$ where S' ranges over $\mathscr{F}'(t_1, t_2)$. Setting $P'(S') = P\{Y(\cdot) \in S\}$, Y induces a probability measure P' on $(\Omega', \mathscr{F}'(0, \infty))$.*

PROOF. Obvious.

1.3. *The general prediction process.* The entire construction of a prediction process can be carried out for any probability P' on $(\Omega', \mathscr{F}'(0, \infty))$, irrespective of the existence of $(X(\cdot), Y(\cdot))$: only the interpretation in terms of $X(\cdot)$ is missing. Again, the more general approach provides a more tractable state space and a more unified treatment. Not only do we obtain a single theory covering at once all separable $(E, \mathscr{E})$ and measurable $X(t)$, but (remarkably enough!) by treating P' as a variable we find that, in the sense of Markov processes, there is really only *one* prediction process, and it is quite independent from any particular $(X(\cdot), Y(\cdot))$. At a later stage (Section 2) we shall specialize to the case when there is an $(X(\cdot), Y(\cdot))$ which induces the measure P'.

DEFINITION 1.3.1. Let $(\mathrm{H}, \mathscr{H})$ denote the set of all probability measures on $(\Omega', \mathscr{F}'(0, \infty))$ with the topology of weak convergence of measures, and its Borel σ-field. Thus $(\mathrm{H}, \mathscr{H})$ is a compact metrizable space with its Borel sets. (Here "weak convergence" as usual means really "weak* convergence," i.e. convergence of the integrals $\int f(w')P'(dw')$ for each $f \in C(\Omega')$).

THEOREM 1.3.1. *For each t, the mapping i_t: $i_t(y_n(s)) = (y_n(s + t) - y_n(t))$ from Ω' onto Ω' is continuous and $\mathscr{F}'(t, \infty)/\mathscr{F}'(0, \infty)$-measurable. For each $P' \in \mathrm{H}$ there exists an H-valued random measure $Z(t)$ on Ω', $\mathscr{F}'(0, t+)/\mathscr{H}$-measurable, such that for all $S' \in \mathscr{F}'(t, \infty)$*

$$P'(S' \mid \mathscr{F}'(0, t+)) = Z(t)(i_t S)\,, \quad P'\text{-a.s.}$$

Here $Z(t)$ is uniquely determined up to a P'-null set.

REMARK. We emphasize the fact that $Z(t)$ depends essentially upon the choice of $P' \in H$.

PROOF. Since $\mathscr{F}'(t, \infty)$ is generated by the random variables $y_n(t+r) - y_n(t)$, $n \geqq 1$, r rational, where the range of these is a Borel set in $\times_{n,r}[0, r]$, it follows from the Existence Lemma of [13, 27.2a, page 360] that regular conditional probabilities exist over $\mathscr{F}'(t, \infty)$ given $\mathscr{F}'(0, t+)$. The continuity and measurability of i_t, and the fact that its range is Ω' are both obvious. Hence, for any $S \in \mathscr{F}'(0, \infty)$, we have $S = i_t S'$ where $S' = i_t^{-1}S \in \mathscr{F}'(t, \infty)$. Thus if $P'(S' \mid \mathscr{F}'(0, t+))$, $S' \in \mathscr{F}(t, \infty)$, is any regular conditional probability as indicated, and if we define $Z(t)(S) = P'(i_t^{-1}S \mid \mathscr{F}'(0, t+))$, $S \in \mathscr{F}'(0, \infty)$, the main assertion follows. Since $\mathscr{F}'(0, \infty)$ is separable, the uniqueness assertion is clear.

We shall next extend the construction to all t.

THEOREM 1.3.2. *For given $P' \in \mathrm{H}$, let $Z(r)$ be defined as in Theorem 1.3.1 for each rational $r > 0$. Then $Z(r)$ has right limits in* H *at all $t \geqq 0$ and left limits at all $t > 0$ except on a P'-null set. We redefine $Z(t) = \lim_{r \downarrow t} Z(r)$ if this exists, and $Z(t) = P'$ elsewhere. Then the process $Z(t)$ is $\mathscr{F}'(0, t+)/\mathscr{H}$-measurable, and right continuous with left limits in* H *outside a fixed P'-null set. Moreover, for each $\mathscr{F}'(0, t+)$-stopping time $T < \infty$ and $S' \in \mathscr{F}'(0, \infty)$,*

$$P'(i_T^{-1}S' \mid \mathscr{F}'(0, T+)) = Z(T)(S')\,, \quad P'\text{-a.s.}\,, \tag{1.3.1}$$

where, as usual, $\mathscr{F}'(0, T+) = \{S' \in \mathscr{F}'(0, \infty) \colon S' \cap \{T \leqq t\} \in \mathscr{F}'(0, t+)$ for all $t \geqq 0\}$. These properties determine $Z(t)$ uniquely to within a fixed P'-null set.

PROOF. The idea of the proof is to define the convergence of $Z(r)$ in H in such a way that it follows from a martingale convergence theorem. Accordingly, for each $n \geqq 1$, let $0 \leqq f_{n,l} \leqq 1$, $l = 1, 2, \cdots$, enumerate a set of continuous functions of n real variable with limit 0 as $x_1^2 + \cdots + x_n^2 \to \infty$ which is uniformly dense in the set of all such functions. Thus, if C_0^n is the corresponding Banach space of continuous functions with limit 0 at ∞, then $(f_{n,l})$ generates a vector space dense in C_0^n. Let (λ_i) be a fixed enumeration of the positive rationals, for each n let $(\mathbf{r}_{n,m}) = (\mathbf{r}_{n,m}(1), \cdots, \mathbf{r}_{n,m}(n))$, $m = 1, 2, \cdots$, be an enumeration of the n-vectors with nonnegative rational components, and let $(j) = (j_1, \cdots, j_n)$ be any n-vector of positive integers. We shall need to use the following

LEMMA 1.3.2. *For each choice of indices, let $L(i, n, l, m, (\mathrm{j}))$ denote $\int_0^\infty \lambda_i \exp -(\lambda_i s) f_{n,l}(y_{\mathrm{j}_1}(\mathbf{r}_{n,m}(1) + s) - y_{\mathrm{j}_1}(s), \cdots, y_{\mathrm{j}_n}(\mathbf{r}_{n,m}(n) + s) - y_{\mathrm{j}_n}(s))\, ds$. Then for each $P' \in \mathrm{H}$, $E_{Z(r)}L(i, n, l, m, (\mathrm{j}))$ is a λ_i-supermartingale for P' as r varies, when $Z(r)$ is defined as in Theorem 1.3.1. Moreover, convergence of $E_{Z(r)}L(i, n, l, m, (\mathrm{j}))$ for all $(i, n, l, m, (\mathrm{j}))$ is equivalent to convergence of $Z(r)$ in* H.

REMARKS. Here it is understood that $L(i, n, l, m, (\mathrm{j}))$ is defined on Ω', and $E_{Z(r)}$ denotes its expectation with respect to the measure $Z(r)$ on Ω'.

PROOF. For simplicity of notation we write only the case $(j) = (1, 2, \cdots, n)$ and we then omit the index (j) in L. The general case is entirely similar. From Theorem 1.3.1 we have

$$E_{Z(r)} L(i, n, l, m) = E_{P'}(\textstyle\int_0^\infty \lambda_i \exp -(\lambda_i s) f_{n,l}(y_1(\mathbf{r}_{n,m}(1) + s + r) - y_1(s + r), \cdots, y_n(\mathbf{r}_{n,m}(n) + s + r) - y_n(s + r))\, ds \,|\, \mathscr{F}'(0, r+)) .$$

For $r' > 0$, the integral on the right is not less than $\exp -(\lambda_i r') \int_0^\infty \lambda_i \exp -(\lambda_i s) \times f_{n,l}(y_1(\mathbf{r}_{n,m}(1) + s + r + r') - y_1(s + r + r'), \cdots, y_n(\mathbf{r}_{n,m}(n) + s + r + r') - y_n(s + r + r'))\, ds$. Consequently, $E_{P'}(E_{Z(r+r')} L(i, n, l, m) \,|\, \mathscr{F}'(0, r+)) \geqq \exp(\lambda_i r') \times E_{Z(r)} L(i, n, l, m)$, proving the λ_i-supermartingale property.

Next, let $z_k \to z$ in the compact metric space H. Since, for each $s \geqq 0$, the quantity

$$I((y), s) = f_{n,l}(y_1(\mathbf{r}_{n,m}(1) + s) - y_1(s), \cdots, y_n(\mathbf{r}_{n,m}(n) + s) - y_n(s))$$

is a continuous function on Ω', we have $\lim_{k\to\infty} E_{z_k} I((y), s) = E_z I((y), s)$. By the dominated convergence theorem, it follows that, for all (i, n, l, m), $\lim_{k\to\infty} E_{z_k} L(i, n, l, m) = E_z L(i, n, l, m)$.

Conversely, if this convergence holds, then for fixed (n, l, m) we have for every $\lambda > 0$,

$$\lim_{k\to\infty} \textstyle\int_0^\infty \lambda \exp -\lambda s E_{z_k}(I((y), s))\, ds = \int_0^\infty \lambda \exp -\lambda s E_z(I((y), s))\, ds ,$$

since clearly $I((y), s)$ is uniformly continuous and bounded in s, uniformly in (y). By the continuity theorem for Laplace transform [8, 13.1, Theorem 2a], the measures $E_{z_k}(I((y), s))\, ds$ converge weakly to $E_z(I((y), s))\, ds$, and this implies that $\lim_{k\to\infty} E_{z_k}(I((y), s)) = E_z(I((y), s))$ for each s. In particular, for $s = 0$ we obtain as l varies with m and n fixed that the joint distributions of $y_1(\mathbf{r}_{n,m}(1)), \cdots, y_n(\mathbf{r}_{n,m}(n))$ on Ω' converge weakly, and as m varies this extends to $y_1(t_1), \cdots, y_n(t_n)$ for all $t_1, \cdots, t_n > 0$. Reintroducing the case of arbitrary (j), one obtains in this way the weak convergence for all finite sets of coordinates on Ω'. The continuous functions on Ω' depending on only finite sets of coordinates are an algebra and separate points. Hence by the Stone–Weierstrass theorem they are uniformly dense in $C(\Omega')$, and it follows that $z_n \to z$ in H. The lemma is proved.

It now follows by the martingale convergence theorem of Doob [5, page 363], extended easily to λ-supermartingales (as in [11, page 326], for example), that $Z(t) = \lim_{r\downarrow t} Z(r)$ for all t, P'-a.s. Since the exceptional set at each t is in $\mathscr{F}'(0, t+)$, we see that $Z(t) \in \mathscr{F}'(0, t+)$, and is right continuous with left limits outside of a fixed P'-null set. Finally, if T is any finite $\mathscr{F}'(0, t+)$-stopping time, then $T = \lim_k T_k$ where, as usual, $T_k = (n + 1)2^{-k}$ on $\{n2^{-k} \leqq T < (n + 1)2^{-k}\}$ for all $n \geqq 0$. As in the proof of Lemma 1.3.2 we have

$$E_{Z(T_k)} L(i, n, l, m) = E_{P'}(\textstyle\int_0^\infty \lambda_i \exp -(\lambda_i s) f_{n,l}(y_1(\mathbf{r}_{n,m}(1) + s + T_k) - y_1(s + T_k)), \cdots, y_n(\mathbf{r}_{n,m}(n) + s + T_k) - y_n(s + T_k))\, ds \,|\, \mathscr{F}'(0, T_k+)) .$$

It follows by Hunt's Lemma (see [3, Lemma 1.2, proof]) that

$$\begin{aligned}\lim_{k\to\infty} E_{Z(T_k)} L(i, n, l, m) &= E_{Z(T)} L(i, n, l, m) \\ &= E_{P'}(\textstyle\int_0^\infty \lambda_i \exp -(\lambda_i s) f_{n,l}(y_1(\mathbf{r}_{n,m}(1) + s + T) - y_1(s + T), \cdots, \\ &\qquad y_n(\mathbf{r}_{n,m}(n) + s + T) - y_n(s + T))\, ds \,|\, \mathscr{F}'(0, T+)) , \quad P'\text{-a.s.}\end{aligned}$$

Bringing the mathematical expectations under integral signs on both sides, and inverting the transforms as before, it follows that for every $f \in C(\Omega')$, $E_{Z(T)}f = E_{P'}(f \circ i_T \,|\, \mathscr{F}'(0, T+))$. Extending this to $0 \leqq f \in \mathscr{F}'(0, \infty)$ by monotone convergence, we obtain equation (1.3.1). The uniqueness assertion is obvious from that of Theorem 1.3.1.

We can also state the following corollary which will be needed in the next item.

COROLLARY 1.3.2. *For each* $t > 0$ *and* $S' \in F'(0, \infty)$, $P'(i_t^{-1}S' \,|\, \mathscr{F}'(0, t)) = Z(t-)(S')$, *where* $Z(t-)$ *denotes the left limit at time* t.

PROOF. Let $\bigvee_n \mathscr{F}_n$ denote, as usual, the σ-field generated by $(\mathscr{F}_n)$, and let $0 < t_n < t$ increase to t. Then $\mathscr{F}'(0, t) = \bigvee_n \mathscr{F}'(0, t_n+)$. For $f \in C(\Omega')$ it is not hard to see that $f \circ i_t$ is continuous on $R^+ \times \Omega'$. It follows by Hunt's Lemma that

$$\begin{aligned}E_{Z(t-)}f &= \lim_{n\to\infty} E_{Z(t_n)}f \\ &= \lim_{n\to\infty} E_{P'}(f \circ i_{t_n} \,|\, \mathscr{F}'(0, t_n+)) \\ &= E_{P'}(f \circ i_t \,|\, \mathscr{F}'(0, t)) , \quad P'\text{-a.s.}\end{aligned}$$

The extension to $0 \leqq f \in \mathscr{F}'(0, \infty)$ completes the proof as before.

REMARK. A discerning reader may observe that these results might also be proved using somewhat stronger topologies on H. In fact, the martingale convergence of Lemma 1.3.2 would hold using any denumerable collections $0 \leqq f_{n,l} \leqq 1$, measurable in $(x_1, \cdots, x_n)$ for each n (although inversion of the transforms would present a further difficulty). However, the connection between H and a given topology on E, as illustrated by Theorems 3.1.1 and 3.1.2 below, seems to break down unless the $f_{n,l}$ are continuous, and in any case the present weak topology seems both natural and typical of the other possibilities.

1.4. *The Markovian character of the prediction processes.* For the purposes of this item and afterwards, y and z will denote variables on H, hence probability measures on Ω'. Our object is to prove that for any $z \in \mathrm{H}$ the corresponding process $Z(t)$ of Theorem 1.3.2 is a homogeneous Markov process, and that all of these processes have a single homogeneous transition function on $(\mathrm{H}, \mathscr{H})$. If this is true, then it is clear that the transition function is given by

DEFINITION 1.4.1. For $0 \leqq t$, $y \in \mathrm{H}$, and $A \in \mathscr{H}$, let $q(t, y, A) = P_y\{Z(t) \in A\}$, where $Z(t)$ is the prediction process defined from y. For $g(y) \in \mathscr{H}$, and positive or bounded, let

$$Q_t g(y) = \textstyle\int_{\mathrm{H}} q(t, y, dz) g(z) = E_y g(Z(t)) .$$

The principle result concerning $q(t, y, A)$ itself is

THEOREM 1.4.1 *If* $g(y) = E_y f$, $f \in C(\Omega')$, *then* $g(y) \in C(\mathrm{H})$ *and* $Q_t g(y) \in C(\mathrm{H})$ *for* $t \geqq 0$. *More generally for* $0 \leqq g \in \mathscr{H}$, $Q_t g(y)$ *is* $\mathscr{B}[0, \infty) \times \mathscr{H}$*-measurable in* (t, y).

REMARK. It is important to observe that $Q_0 g(y)$ is not in general $g(y)$. By Theorem 1.3.2 it is $E_y g(Z(0))$ where $Z(0)(S') = P_y(S' \mid \mathscr{F}'(0, 0+))$, but in general $\{P_y(S' \mid \mathscr{F}'(0, 0+)) \neq P_y(S')\}$ has positive probability. Anticipating the Markov property, it may be said that in this case y has the behavior of a "branching point" as in Ray (1959). However, this should not be taken too literally since we are not actually dealing with a Ray process. An example is indicated below to show that the "resolvent" of Q_t does not take $C(\mathrm{H}) \to C(\mathrm{H})$. In particular, the Feller property $Q_t g \in C(\mathrm{H})$ does not always hold if $g \in C(\mathrm{H})$ is not of the special form $g(y) = E_y f$, $f \in C(\Omega')$ (for example, if $g(y) = g(E_y f_1, E_y f_2)$ where $g(x_1, x_2)$ is continuous on R^2 and $f_i \in C(\Omega')$, $i = 1$ or 2). This necessitates the rather elaborate measurability proof for $Q_t g$ given below.

PROOF. If $g(y) = E_y f$, $f \in C(\Omega')$, then by (1.3.1) we have

$$\begin{aligned} Q_t g(y) &= E_y E_{Z(t)} f \\ &= E_y E_y(f \circ i_t \mid \mathscr{F}'(0, t+)) \\ &= E_y(f \circ i_t) . \end{aligned}$$

Since i_t is continuous, $f \circ i_t \in C(\Omega')$ and the continuity of $E_y(f \circ i_t)$ follows immediately.

To prove the measurability over $\mathscr{B}[0, \infty) \times \mathscr{H}$, since $Q_t g(y) = E_y g(Z(t))$ and $Z(t)$ is right-continuous (P_y-a.s. for each y), it follows by standard arguments (not requiring the Markov property) that is suffices to prove $Q_t g(y) \in \mathscr{H}$ for fixed t. To this end, consider first $g(y) = g(E_y f_1, \cdots, E_y f_n)$ where, without risk of confusion, $g(x_1, \cdots, x_n)$ is continuous in $n \geqq 1$ real variables on the right, and $f_i \in C(\Omega')$, $i = 1, 2, \cdots, n$. For such $g(y)$ it will be enough to prove that $Q_{t-} g \in \mathscr{H}$ for all t > 0, where the limit

$$\begin{aligned} Q_{t-} g(y) &= \lim_{s \uparrow t} E_y(E_{Z(s)} g) \\ &= E_y g(E_{Z(t-)} f_1, \cdots, E_{Z(t-)} f_n) \end{aligned}$$

exists by continuity of the functions involved. On the right, by Corollary 1.3.2, we have $E_{Z(t-)} f_i = E_y(f_i \circ i_t \mid \mathscr{F}'(0, t))$, $1 \leqq i \leqq n$. The advantage of using $Z(t-)$ is that $\mathscr{F}'(0, t)$, unlike $\mathscr{F}'(0, t+)$, is separable. Let $\mathscr{F}_k$, $k = 1, 2, \cdots$, be an increasing family of finite subfields of $\mathscr{F}'(0, t)$ with $\bigvee_k \mathscr{F}_k = \mathscr{F}'(0, t)$. Then $E_y(f_i \circ i_t \mid \mathscr{F}_k)$ is a martingale in k, $1 \leqq i \leqq n$, and by a theorem of P. Lévy

$$\lim_{k \to \infty} E_y(f_i \circ i_t \mid \mathscr{F}_k) = E_{Z(t-)} f_i , \quad P_y\text{-a.s.}$$

Since $E_y g(E_y(f_1 \circ i_t \mid \mathscr{F}_k), \cdots, E_y(f_n \circ i_t \mid \mathscr{F}_k))$ can be written as a finite sum over the atoms $S_{j,k}$ of $\mathscr{F}_k$ in the form

$$\sum_j P_y(S_{j,k})g(E_y(f_1 \circ i_t; S_{j,k})P_y^{-1}(S_{j,k}), \cdots, E_y(f_n \circ i_t; S_{j,k})P_y^{-1}(S_{j,k}))$$

it is clearly $\mathscr{H}$-measurable in y. By continuity of $g(x_1, \cdots, x_n)$ this converges to $Q_{t-}g(y)$ for each y, hence $Q_{t-}g(y) \in \mathscr{H}$.

The class of $g \in C(\mathrm{H})$ having the above form for some $n \geqq 1$ is an algebra and separates points in H. Hence it is uniformly dense in $C(\mathrm{H})$ by the Stone–Weierstrass theorem, and the measurability of $Q_{t-}g$ thus extends to $g \in C(\mathrm{H})$. By right-continuity of $Z(t)$ we obtain $Q_t g \in \mathscr{H}$, and finally this extends to all $g \in \mathscr{H}$ by the usual monotone class theorem. This completes the proof.

EXAMPLE 1.4.1. Consider the measure $z \in \mathrm{H}$ for which $P_z\{(y_n(t)) = t, 0 \leqq t \leqq 1, n \geqq 1$, and $(y_n(t)) = 1 + \frac{1}{2}(t-1), t \geqq 1\} = \frac{1}{2} = P_z\{(y_n)) = t, 0 \leqq t \leqq 1$, and $(y_n(t)) = 1 + \frac{1}{3}(t-1),\ t \geqq 1\}$, i.e. the measure is concentrated on two paths which are identical in $0 \leqq t \leqq 1$, but differ for $t > 1$. If we define z_k by $P_{z_k}\{(y_n(t)) = (1 - 1/k)t, 0 \leqq t < 1$, and $(y_n(t)) = (1 - 1/k) + \frac{1}{2}(t-1), t \geqq 1\} = \frac{1}{2} = P_{z_k}\{(y_n(t)) = (1 - 2/k)t,\ 0 \leqq t < 1$, and $(y_n(t)) = (1 - 2/k) + \frac{1}{3}(t-1)$, $t \geqq 1\}$, then clearly $\lim_{k\to\infty} z_k = z$ in H. However, the prediction processes $Z_k(t)$ of z_k do not converge weakly to $Z(t)$ for $t < 1$, since the former are concentrated on two quite distinct paths, each having probability $\frac{1}{2}$ and which have point measures as values, while the latter is unique in $0 \leqq t \leqq 1$, and its values are measures which are equally divided between 2 points. Thus if we introduce $R_\lambda g(y) = \int_0^\infty \exp -(\lambda t) Q_t g(y)\, dt$ there are $g \in C(\mathrm{H})$ for which $\lim_{k\to\infty} R_\lambda g(z_k) \neq R_\lambda g(z)$. In short, the topology of H is insufficient to insure weak convergence of the conditional distributions $Z_k(t)$ to those of the limit measure.

We turn now to the main theorem.

THEOREM 1.4.2. *The processes* $(\Omega', \mathscr{F}'(0, \infty), P_z, Z(t))$, $z \in \mathrm{H}$, *are Markovian relative to the σ-fields* $\mathscr{F}'(0, t+)$, *with* $(\mathrm{H}, \mathscr{H})$ *as state space and* $q(t, y, A)$ *as transition function. They are right-continuous with left limits* P_z-a.s. *and have the strong Markov property.*

REMARK. In view of this result, one can define an equivalent Markov process in the sense of Dynkin (1965) on the canonical space of right-continuous, H-valued paths with left limits for $t > 0$, but this process is not normal since in general $P_y\{Z(0) = y\} \neq 1$. It will be shown in the next section (Theorem 2.1.4) that the paths $(y_n) \in \Omega'$ can be recovered from those of $Z(t)$, hence there is no real necessity to maintain Ω' as sample space apart from its simplicity.

PROOF. We first prove the simple Markov property of $Z(t)$ for each y, with transition function $q(t, y, A)$. Taking a mildly restrictive case, we show first that for $A_x = \{z : E_z f \leqq x\}$, $f \in C(\Omega')$, we have

$$(1.4.1) \qquad P_y(Z(t_1 + t_2) \in A_x \mid \mathscr{F}'(0, t_1+)) = q(t_2, Z(t_1), A_x)\,, \quad P_y\text{-a.s.}$$

It will be convenient at times to record explicitly the dependence of $Z(t)$ on y

by denoting it $Z_y(t)$ so that the right side of (1.4.1) is $P_{Z(t_1)}\{E_{Z_{Z(t_1)}(t_2)} f \leqq x\}$. To prove (1.4.1) we can assume without loss of generality that $t_1 + t_2$ is a continuity point of $Z_y(t)$, P_y-a.s., since for each t_1 this is true for t_2 in a dense set, while (1.4.1) for all x is equivalent to

$$(1.4.2) \qquad E_y(g(E_{Z(t_1+t_2)} f) \mid \mathscr{F}'(0, t_1+)) = E_{Z(t_1)} g(E_{Z_{Z(t_1)}(t_2)} f)$$

when g varies in a countable dense set of continuous functions on $[-\max |f|, +\max |f|]$, and (1.4.2) is clearly right-continuous in t_2, P_y-a.s. at each t, along any countable dense set of t_2. At such a continuity point, Corollary 1.3.2 yields that $Z(t_1 + t_2) \in \mathscr{F}'(0, t_1 + t_2)$ up to a P_y-null set, and $\mathscr{F}'(0, t_1 + t_2)$ is separable. Accordingly, let $S(x) \in \mathscr{F}'(0, t_1 + t_2)$ differ from $\{Z(t_1 + t_2) \in A_x\}$ by at most a P_y-null set. We shall need

DEFINITION 1.4.2. For any $S' \in \mathscr{F}'(0, \infty)$, the cylindrical section of S' in $\mathscr{F}'(t, \infty)$ at $(y_n) \in \Omega'$ is the set $C(S', (y_n)) = \{(z_n) \in \Omega'$: for some $(w_n) \in S'$ with $w_n(s) = y_n(s)$, $0 < s \leqq t$, one has $z_n(s) = w_n(s)$, $t \leqq s < \infty\}$.

We show that this concept has the anticipated properties in

LEMMA 1.4.2. *For each* $(y_n) \in \Omega'$, $C(S', (y_n)) \in \mathscr{F}'(t, \infty)$, *and* $P_y(S' \mid \mathscr{F}'(0, t+)) = P_{Z(t)} i_t(C(S', (y_n))$ P_y-a.s., *where* (y_n) *is the identity function on* Ω'.

PROOF. Suppose first that $S' = S_1 \cap S_2$ where $S_1 \in \mathscr{F}'(0, t)$ and $S_2 \in \mathscr{F}'(t, \infty)$. Then it is not hard to see that $C(S, (y_n))$ is either empty (if $(y_n) \notin S_1$) or equals S_2 (if $(y_n) \in S_1$), and that $P_y(S' \mid \mathscr{F}'(0, t+)) = 0$ (if $(y_n) \notin S_1$), or $P_y(S' \mid \mathscr{F}'(0, t+)) = P_{Z(t)}(i_t S_2)$ (if $(y_n) \in S_1$, P_y-a.s.). Just as in the case of product σ-fields, a finite union of sets S' of this form may be written as a finite disjoint union, and by additivity the result follows for any $S' \in \mathscr{F}'(0, t) \vee \mathscr{F}'(t, \infty)$. It only remains to note that since $y_n(s) = y_n(t) + (y_n(s) - y_n(t))$ for $s > t$, where $y_n(t) \in \mathscr{F}'(0, t)$ and $y_n(s) - y_n(t) \in \mathscr{F}'(t, \infty)$, we must in fact have $\mathscr{F}'(0, t) \vee \mathscr{F}'(t, \infty) = \mathscr{F}'(0, \infty)$.

Returning, now, to $S(x)$ we note that $S(x) \in \mathscr{F}'(0, t_1 + t_2)$ and for every $S_1 \in \mathscr{F}'(0, t_1^+)$ and $S_2 \in \mathscr{F}'(t_1, t_1 + t_2)$

$$E_y(f \circ i_{t_1+t_2}; S_1 \cap S_2 \cap S(x)) \leqq x P_y(S_1 \cap S_2 \cap S(x))$$

By varying S_1 in this inequality we obtain

$$(1.4.3) \qquad \begin{aligned} &E_y(f \circ i_{t_1+t_2}; S_2 \cap S(x) \mid \mathscr{F}'(0, t_1^+)) \\ &\qquad \leqq x P_y(S_2 \cap S(x) \mid \mathscr{F}'(0, t_1^+)) , \quad P_y\text{-a.s.} \end{aligned}$$

We next apply Lemma 1.4.2 with

$$S' = S_2 \cap S(x) , \qquad C(S', (y_n)) = S_2 \cap C(S(x), (y_n)) ,$$

and $t = t_1$, to express (1.4.3) in the form

$$(1.4.4) \qquad \begin{aligned} &E_{Z(t_1)}(f \circ i_{t_2}; i_{t_1}(S_2 \cap C(S(x), (y_n)))) \\ &\qquad \leqq x P_{Z(t_1)}\{i_{t_1}(S_2 \cap C(S(x), (y_n)))\} , \quad P_y\text{-a.s.} \end{aligned}$$

But since $\mathscr{F}'(t_1, t_1 + t_2)$ is separable (and $P_{Z(t_1)}$ is a regular measure on $\mathscr{F}'(t_1, t_1 + t_2)$ [10, page 228]) it is easy to see that (1.4.4) holds simultaneously in S_2, P_y-a.s. This implies that P_y-a.s. we have up to a $P_{Z(t_1)}$-null set

$$(1.4.5) \qquad i_{t_1}C(S(x), (y_n)) \subset \{E_{Z(t_1)}(f \circ i_{t_2} \mid \mathscr{F}'(0, t_2)) \leqq x\} .$$

The same argument shows that for $x_1 < x_2$,

$$(1.4.6) \qquad i_{t_1}C(S(x_2) - S(x_1), (y_n)) \subset \{x_1 \leqq E_{Z(t_1)}(f \circ i_{t_2} \mid \mathscr{F}'(0, t_2)) \leqq x_2\}$$

in the same sense. Thus by using a dense set x_n we can easily obtain that for all $C_1 < C_2$, P_y-a.s.,

$$(1.4.7) \qquad \int_{C_1}^{C_2} xP_{Z(t_1)}(i_{t_1}C(S(dx), (y_n))) \leqq \int_{C_1}^{C_2} xP_{Z(t_1)}\{E_{Z(t_1)}(f \circ i_{t_2} \mid \mathscr{F}'(0, t_2)) \in dx\} .$$

On the other hand, for $C_1 = -\infty$ and $C_2 = +\infty$ we have

$$\begin{aligned} E_y \int P_{Z(t_1)}(i_{t_1}C(S(dx), (y_n))) &= E_y \int xP_y(S(dx) \mid \mathscr{F}'(0, t_1+)) \\ &= \int xP_y\{E_y(f \circ i_{t_1+t_2} \mid \mathscr{F}'(0, t_1 + t_2)) \in dx\} \\ &= E_y E_y(f \circ i_{t_1+t_2} \mid \mathscr{F}'(0, t_1 + t_2)) \\ &= E_y f \circ i_{t_1+t_2} , \end{aligned}$$

while on the right side of (1.4.7)

$$\begin{aligned} E_y \int xP_{Z(t_1)}\{E_{Z(t_1)}(f \circ i_{t_2} \mid \mathscr{F}'(0, t_2)) \in dx\} &= E_y E_{Z(t_1)} E_{Z(t_1)}(f \circ i_{t_2} \mid \mathscr{F}'(0, t_2)) \\ &= E_y E_{Z(t_1)} f \circ i_{t_2} \\ &= E_y f \circ i_{t_1+t_2} . \end{aligned}$$

It follows that (1.4.7) holds with equality for all $C_1 < C_2$, P_y-a.s., and this implies

$$(1.4.8) \qquad P_{Z(t_1)}(i_{t_1}C(S(x), (y_n))) = P_{Z(t_1)}\{E_{Z(t_1)}(f \circ i_{t_2} \mid \mathscr{F}'(0, t_2)) \leqq x\}$$

for all x, P_y-a.s. In view of Lemma 1.4.2 and Corollary 1.3.2 this is equivalent to

$$(1.4.9) \qquad P_y(Z(t_1 + t_2) \in A_x \mid \mathscr{F}'(0, t_1+)) = P_{Z(t_1)}\{E_{Z_{Z(t_1)}(t_2-)} f \leqq x\} ,$$

and to prove (1.4.1) we need only remove the minus sign on the right. To do this, we let $g(x)$ be continuous and obtain as for (1.4.2)

$$(1.4.10) \qquad E_y(g(E_{Z(t_1+t_2)} f \mid \mathscr{F}'(0, t_1+)) = E_{Z(t_1)} g(E_{Z_{Z(t_1)}(t_2-)} f) .$$

Since $Z_{Z(t_1)}(t_2)$ is right-continuous in t_2 for each $Z(t_1)$, we may now take right limits in t_2 along a countable set to obtain (1.4.2), and hence (1.4.1).

The remainder of the proof is essentially a repetition. Thus if $A = \bigcap_{j=1}^k \{E_z f_j \leqq x_j\}$, $f_j \in C(\Omega')$, then $\bigcap_{j=1}^k S_j(x_j) \in \mathscr{F}'(0, t_1 + t_2)$ and differs from $\{Z(t_1 + t_2) \in A\}$ by a P_y-null set, while (1.4.2) can be expressed using $g(E_{Z(t_1+t_2)} f_1, \cdots, E_{Z(t_1+t_2)} f_k)$ where $g(x_1, \cdots, x_k)$ is continuous. Since, by the preceding

$$\begin{aligned} i_{t_1}C(\textstyle\bigcap_{j=1}^k S_j(x_j), (y_n)) &= \textstyle\bigcap_{j=1}^k i_{t_1}C(S_j(x_j), (y_n)) \\ &= \textstyle\bigcap_{j=1}^k \{E_{Z(t_1)}(f_j \circ i_{t_2} \mid \mathscr{F}'(0, t_2)) \leqq x_j\} \end{aligned}$$

up to a $P_{Z(t_1)}$-null set, P_y-a.s., we have the extension of (1.4.9), and using $g(x_1, \cdots, x_k)$ in a countable uniformly dense set on $\mathsf{X}_{j=1}^k [-\max |f_j|, \max |f_j|]$ we obtain (1.4.2) for all $g(E_z f_1, \cdots, E_z f_k)$. As was noted in the proof of Theorem 1.4.1, such functions are uniformly dense in $C(\mathrm{H})$, hence the general case of (1.4.1) follows by the monotone class theorem.

The proof of the strong Markov property is essentially the same, upon replacing t_1 by a finite $\mathscr{F}'(0, t_1+)$-stopping time T_1. It is necessary to apply Corollary 1.3.2 at the random time $T_1 + t_2$, but since, for $0 < \varepsilon_n < t_2$, $T_1 + t_2 - \varepsilon_n$ is again a stopping time, we may use the same proof (with $\varepsilon_n \to 0$) where as usual $\mathscr{F}'(0, T_1 + t_2)$ is defined like $\mathscr{F}'(0, T_1 + t_2+)$ (see (1.3.1)), but using $\mathscr{F}'(0, t)$ in place of $\mathscr{F}'(0, t+)$. By the continuity of (y_n) we have $\mathscr{F}'(0, T_1 + t_2) = \bigvee_n \mathscr{F}'(0, T_1 + t_2 - \varepsilon_n+)$. Recalling Definition 1.2.1 of $\mathscr{F}'(T_1, \infty)$, and applying it first with $T_1 + \varepsilon$ in place of T_1, both Definition 1.4.2 and Lemma 1.4.2 extend easily to $\mathscr{F}'(0, t)$-stopping times $T_1 + \varepsilon$, $\varepsilon > 0$, since $\mathscr{F}'(0, \infty) = \mathscr{F}'(0, T_1 + \varepsilon) \vee \mathscr{F}'(T_1 + \varepsilon, \infty)$ follows by standard measurability arguments using $T_1 + \varepsilon \in \mathscr{F}'(0, T_1 + \varepsilon)$. This yields the equivalent expression $E_y(f \mid \mathscr{F}'(0, T_1 + \varepsilon+)) = E_{Z(T_1+\varepsilon)} f(\cdot; (y_n))$, $0 \leqq f \in C(\Omega')$, where $f((y_n)'; (y_n)) = f((z_n))$: $z_n(t) = y_n(t)$ if $t \leqq T_1 + \varepsilon$, and $z_n(t) = y_n(T_1 + \varepsilon) + y_n'(t - (T_1 + \varepsilon))$ otherwise. Letting $\varepsilon \to 0$ and using the uniform convergence of (z_n), the uniform continuity of f, and the weak right continuity of $Z(t)$, we obtain Lemma 1.4.2 for T_1. Finally since $P_{Z(T_1)}$ is a regular measure on the separable $\mathscr{F}'(T_1, T_1 + t_2)$, the remainder of the proof goes through unchanged.

2. Reintroduction of $X(t)$.

2.1. *The prediction process of $X(t)$.* In this item we examine the meaning of $Z(t)$ when its probability z is induced as in Theorem 1.2.2, by a pair $(X(\cdot), Y(\cdot))$. The following theorem is almost immediate from Theorem 1.3.2 and its corollary.

THEOREM 2.1.1. *For each $\mathscr{F}(0, t+)$-stopping time $T < \infty$ and $S \in \mathscr{F}(T, \infty)$*

$$P(S \mid \mathscr{F}(0, T+)) = Z(T)(S'), \quad P\text{-a.s.}$$

for any $S' \in \mathscr{F}'(0, \infty)$ such that $S = \{Y(\cdot) \in i_T^{-1}(S')\}$. Moreover if $0 < T < \infty$ is predictable (i.e. if $T = \lim_{n\to\infty} T_n$ where T_n are $\mathscr{F}(0, t+)$-stopping times less than T, [16]) then $P(S \mid \mathscr{F}(0, T)) = Z(t-)(S')$ P-a.s.

PROOF. We first define an $\mathscr{F}'(0, t+)$-stopping time T' such that $T'(Y(\cdot)) = T$ on Ω. For each n, let $S'_{k,n} \in \mathscr{F}'(0, (k+1)2^{-n})$, $k \geqq 0$, be such that $\{k2^{-n} \leqq T < (k+1)2^{-n}\} = \{w \in \Omega : Y(\cdot) \in S'_{k,n}\}$. This is possible in view of Theorem 1.2.2. Replacing $S'_{k,n}$ by $S'_{k,n} - \bigcup_{j<k} S'_{j,n}$ we may assume that for each n $S'_{k,n}$'s are disjoint. We define $T_n' = (k+1)2^{-n}$ on $S'_{k,n}$, $k \geqq 0$, and $T_n' = \infty$ on $\Omega' - \bigcup_k S'_{k,n}$. Then T_n' is an $\mathscr{F}'(0, t+)$-stopping time and $P'(T_n' < \infty) = 1$, where P' is the measure induced by P. Finally, we set $T' = \limsup_{n\to\infty} T_n'$. It is clear that T' is a stopping time and $P'\{T' < \infty\} = 1$. Also $|T'(Y(\cdot)) - T| < 2^{-n}$ for every n, so that $T'(Y(\cdot)) = T$ on Ω.

This being so, it is easy to show that $\mathscr{F}(0, T+) = \{\{Y(\cdot) \in S'\}, S' \in \mathscr{F}'(0, T'+)\}$, and the first assertion of the theorem follows by (1.3.1) with T' in place of T (where the set $\{T' = \infty\}$ can be neglected), and by definition of the measure P' induced by P.

For the last assertion, if T_n' corresponds as above to T_n, then $T_n'' = \max_{k \leqq n} T_k'$ are stopping times which increase to a limit T' with $T'(Y(\cdot)) = T$ on Ω. Since the $y_n(t)$ are continuous it is clear that $\bigvee_n \mathscr{F}'(0, T_n'') = \mathscr{F}'(0, T')$, and the same proof as in Corollary 1.3.2, together with the observation $\mathscr{F}(0, T) = \{\{Y(\cdot) \in S'\}, S' \in \mathscr{F}'(0, T')\}$, leads as above to the desired result.

Although this theorem and Theorem 1.1.1 show that the meaning of $Z(t)$ for $X(t)$ does not really depend on the choice of (h_n) to generate $Y(t)$, it may be of interest to prove an even stronger uniqueness result to the effect that for another choice (h_n') there is a functional dependence for all t between $Z(t)$ and the corresponding $Z'(t)$. By analogy with the proof of Theorem 1.1.1 it will be useful to introduce

DEFINITION 2.1.1. Let $w_n(t) = \limsup_{\Delta t \to 0+} (\Delta t)^{-1}(y_n(t + \Delta t) - y_n(t))$ on Ω', and for the given $(X(\cdot), Y(\cdot))$ let $W_n(t) = \limsup_{\Delta t \to 0+} (\Delta t)^{-1}(Y_n(t + \Delta t) - Y_n(t))$ on Ω.

It is clear that (in the terminology of [14]) $(w_n(t))$ is a $\mathscr{B}(0, t) \times \mathscr{F}'(0, t+)$-progressively meastrable process, and $(W_n(t))$ is a $\mathscr{B}(0, t) \times \mathscr{F}(0, t+)$-progressively measurable process. Moreover, $y_n(t) = \int_0^t w_n(s)\, ds$ and $Y_n(t) = \int_0^t W_n(s)\, ds$ for all n and t. As in the proof of Theorem 1.1.1 we write $h_n' = g_n(h_1, h_2, \cdots,)$ for each n where $g_n(x_1, x_2, \cdots)$ are measurable on $\mathsf{X}_{n=1}^{\infty} [0, 1]$. We now introduce the measurable mappings $\pi : \Omega' \to \Omega'$ and $\Pi : \mathrm{H} \to \mathrm{H}$ by $\pi(y_n(t)) = (\int_0^t g_n(w_1(s), w_2(s), \cdots)\, ds)$ and $\Pi z(S') = P_z\{(y_n) \in \pi^{-1}(S')\}$, $S' \in \mathscr{F}'(0, \infty)$.

THEOREM 2.1.2.

$$\text{(a)} \quad P\{Z'(t) = \Pi(Z(t)) \ \textit{for all} \ t \geqq 0\} = 1 .$$

$$\text{(b)} \quad P\{Z'(t-) = \Pi(Z(t-)) \ \textit{for all} \ t \geqq 0\} = 1 .$$

PROOF. Since $(W_n(t)) = (h_n(X(t)))$ except for t in a Lebesgue—null set, we have $Y'(\cdot) = \pi(Y(\cdot))$ where $Y'(\cdot)$ is defined using (h_n'). Therefore $\{Y'(\cdot) \in i_T^{-1}S'\} = \{Y(\cdot) \in \pi^{-1} i_T^{-1}(S')\} = \{Y(\cdot) \in i_T^{-1}\pi^{-1}(S')\}$, and by Theorem 2.1.1 we have $Z'(T)(S') = Z(T)(\pi^{-1}(S'))$ P-a.s. for each $\mathscr{F}(0, t+)$-stopping time T. Since $\mathscr{F}'(0, \infty)$ is separable, this implies that $Z'(T) = \Pi(Z(T))$, P-a.s.

According to [14, 8, Theorem 16, Remark (a)] both $Z(t)$ and $Z'(t)$ are well-measurable processes relative to $\mathscr{F}(0, t+)$ at least if these are augmented to contain all P-null sets. Since Π is measurable, it follows that $\{(t, w) : Z'(t) = \Pi(Z(t))\}$ is well-measurable. Accordingly, *unless* its projection on Ω is a.s. empty, there is by Meyer's section theorem [14, 8, Theorem 21] a stopping time T with $P\{Z'(T) \neq \Pi(Z(T))\} > 0$. This contradiction completes the proof of part (a).

For part (b) we use the fact that $\{(t, w): Z'(t-) = \Pi(Z(t-))\}$ is previsible [14, 8, Theorem 19]. Hence by [*ibid.*, Theorem 21] there would be a predictable ("accessible") stopping time T with $P\{Z'(T-) \neq \Pi(Z(T-))\} > 0$, contradicting the second part of Theorem 2.1.1, just as the first part was contradicted before. Thus the proof is complete.

A second, rather theoretical, but nonetheless important question concerns the stability of the construction. In short, if we repeat the construction, using $Z(t)$ in place of $X(t)$ to obtain the prediction process of $Z(t)$, does anything new appear? Not surprisingly the answer is negative, even in the general case when $z \in \mathrm{H}$ need not be induced by any $(X(\cdot), Y(\cdot))$. To make this precise, let $0 \leqq h_n \leqq 1$ be any generating sequence defined on H, and consider for any $z \in \mathrm{H}$ the corresponding $(Y_n'(t)) = \int_0^t h_n(Z_z(s))\, ds)$ where, as before, Z_z denotes the prediction process for z. We introduce the measurable mapping $j: \mathrm{H} \to \mathrm{H}$ defined by $j(z)(S') = P_z\{(Y_n') \in S'\}$.

THEOREM 2.1.3. *$P_z\{Z'(t) = j(Z_z(t))$ for all $t \geqq 0\} = 1$, where $Z'(t)$ is the prediction process for the measure $z' = j(z)$ induced by (Y_n').*

PROOF. For any $\mathscr{F}_{Y'}(0, t+)$-stopping time $T < \infty$, by Theorem 2.1.1, $Z'(T)(S') = P_z(S \mid \mathscr{F}_{Y'}(0, T+)) = P_z\{(Y_n'(T + (\cdot)) - Y_n'(T)) \in S' \mid \mathscr{F}_{Y'}(0, T+)\}$, where $\mathscr{F}_{Y'}$ denotes the σ-field generated by (Y_n') on Ω', and $S = \{(Y_n'(\cdot)) \in i_T^{-1}(S')\}$. Now the mapping $z \to (h_n(z))$ from $\mathrm{H} \to \times_1^\infty [0, 1]$ is one-to-one and Borel. Hence its inverse $(h_n)^{-1}$ is also Borel, and in particular the range of (h_n) is a Borel set.[2]

Since $(h_n(Z_z(t))) = (W_n(t))$ for Lebesgue-a.e. t, there is a dense set $\{t_k\}$ with $P_z\{(h_n(Z_z(t_k))) = W_n(t_k)$ for all n and $k\} = 1$. It follows that $Z_z(t_k) = (h_n)^{-1}(W_n(t_k)) \in \mathscr{F}_{Y'}(0, t_k+)$ up to a P_z-null set, and by right-continuity of $Z_z(t)$ the same measurability is true for every t. Then $Z_z(T) \in \mathscr{F}_{Y'}(0, T+)$ up to a P_z-null set, and by the strong Markov property of $Z(t)$ we have

$$\begin{aligned} P_z((Y_n'(T + (\cdot)) - Y_n'(T)) \in S' \mid \mathscr{F}_{Y'}(0, T+)) &= P_{Z_z(T)}\{(Y_n'(\cdot)) \in S'\} \\ &= P_{j(Z_z(t))}(S')\,, \quad P_z\text{-a.s.} \end{aligned}$$

Thus $Z'(T)(S') = j(Z_z(T))(S')$, P_z-a.s., and since $\mathscr{F}'(0, \infty)$ is separable we have $Z'(T) = j(Z_z(T))$, P_z-a.s. Now since $Z'(t)$ and $Z_z(t)$ are right-continuous and hence well-measurable, our assertion follows as in Theorem 2.1.2 by the section theorem of Meyer.

The last question of interest here is the extent to which $Z(t)$ determines $X(t)$. This is significant in determining the sense in which the present approach to processes is a valid replacement for the usual one. The solution depends, as follows, upon

[2] See, for example, [4, page 98], for the case when the h_n are continuous. For the general case one considers the graph of (h_n) in $\mathrm{H} \times \times_1^\infty [0, 1]$, followed by the continuous projection onto $\times_1^\infty [0, 1]$.

DEFINITION 2.1.2. For $z \in \mathrm{H}$ and $n \geqq 1$, let $v_n(z) = \limsup_{k\to\infty} \int E_z k y_n(k^{-1}) \in \mathscr{H}$.

THEOREM 2.1.4. *For $z \in \mathrm{H}$, we have $y_n(t) = \int_0^t v_n(Z(s))\,ds$ for all t, P_z-a.s. In particular when z is induced by $(X(\cdot), Y(\cdot))$ then $Y_n(t) = \int_0^t v_n(Z(s))\,ds$ for all t, P-a.s., and $P\{h_n(X(t)) = v_n(Z(t))$ for all $n \geqq 1\} = 1$ except for t in a Lebesgue-null set.*

REMARK. Since (h_n) is one-to-one on E, this means that $Z(t)$ determines $X(t) = (h_n)^{-1}(v_n(Z(t)))$ P-a.s., except on a Lebesgue—null set of t.

PROOF. For $z \in \mathrm{H}$ we have by (1.3.1)

$$\int_0^t E_{Z(s)} k y_n(k^{-1})\,ds = \int_0^t E_z(k(y_n(s+k^{-1}) - y_n(s)) \mid \mathscr{F}'(0, s+))\,ds\,. \tag{2.1.1}$$

Recalling Definition 2.1.1 for $w_n(t)$ and using Fatou's Lemma we have P_z-a.s.

$$\begin{aligned} &\limsup_{k\to\infty} \int_0^t E_{Z(s)} k y_n(k^{-1})\,ds \\ &\leqq \int_0^t v_n(Z(s))\,ds \\ &\leqq \int_0^t E(w_n(s) \mid \mathscr{F}'(0, s+))\,ds = \int_0^t w_n(s)\,ds = y_n(t)\,. \end{aligned} \tag{2.1.2}$$

On the other hand, taking E_z on the right of (2.1.1),

$$E_z \limsup_{k\to\infty} \int_0^t E_{Z(s)} k y_n(k^{-1})\,ds = \lim_{k\to\infty} \left(\int_t^{t+k^{-1}} k E_z y_n(s)\,ds\right) = E_z y_n(t)$$

by the dominated convergence theorem. Thus (2.1.2) holds with equality P_z-a.s., and by continuity this extends to all t. This proves the first assertion, and the remaining ones follow immediately by Theorem 1.2.2, the joint measurability of $h_n(X(t))$ and $v_n(Z(t))$, and Fubini's Theorem.

2.2. *Reduction of* H *for prescribed E and (h_n).* It is evident that the space H contains many points not generated by any $(X(\cdot), Y(\cdot))$, and that for different spaces E the subsets of H induced by measurable processes on E do not coincide. Thus it becomes of interest to reduce H at least to the subset induced by all $X(t)$ on a prescribed E, whenever this is possible. We shall not enter here into the possibility of further reductions conforming to a given $X(t)$ on E, or class thereof. In any case, for a fixed generating sequence (h_n) on E we will need to know that the range of (h_n) in $\times_1^\infty [0, 1]$ is measurable in a suitable sense, and to this effect the weakest hypothesis we will use is

ASSUMPTION 2.2.1. Let $(E, \mathscr{E})$ be a topological space homeomorphic to a universally measurable subset of a complete separable metric space D (i.e. of a Polish space D, in the terminology of Bourbaki [2, Chapter 9, Section 6]) together with the restrictions to E of the Borel sets of D, where we consider E as embedded in D. Assume further that the h_n are restrictions to E of measurable functions on D.

REMARK. The most important case is when $E = D$ and D is compact. In this case we can choose $(h_n) \in C(E)$, generating a vector space dense in $C(E)$. Then convergence in H for measures z_n induced by processes $(X_n(\cdot), Y_n(\cdot))$ is equivalent to joint weak convergence of the random sojourn integrals

$\int \mu_n(t, dx)f(x), f \in C(E)$, $t > 0$, where $\mu_n(t, A)$ are defined as in Corollary 1.1.1. This can be viewed as a weak convergence in the space of sojourn time distributions.

We formally introduce the relevant subset of H.

DEFINITION 2.2.1. Let $E - \mathrm{H}$ consist of all elements of H induced by $\mathscr{L} \times \mathscr{F}$-completion-measurable processes $X(\cdot)$ with state space E, for given (h_n) and arbitrary $(\Omega, \mathscr{F}, P)$.

THEOREM 2.2.1. *Under Assumption* 2.2.1, $E - \mathrm{H}$ *is universally measurable in* H. *If E is Borel in D, then $E - \mathrm{H}$ is Borel in* H.

PROOF. We denote by $\mathscr{I}: E \to \mathsf{X}_1^\infty [0, 1]$ the mapping $\mathscr{I}(e) = (h_n(e))$, and for $A \subset E$ the image of A by $\mathscr{I}(A)$. If E is Borel in D, then $\mathscr{I}(E)$ is Borel in H, as follows from the footnote 2 in 2.1. For any finite measure (Radon measure) μ_2 on $\mathsf{X}_1^\infty [0, 1]$, $(h_n)^{-1}$ induces a measure μ_1 on D. Let $E_1 \subset E \subset E_2$ be Borel sets of D with $\mu_1(E_1) = \mu_1(E_2)$. Then clearly $\mathscr{I}(E_1) \subset \mathscr{I}(E) \subset \mathscr{I}(E_2)$ and $\mu_2(\mathscr{I}(E_1)) = \mu_2(\mathscr{I}(E_2))$, hence $\mathscr{I}(E)$ is universally measurable.

LEMMA 2.2.1. *A necessary and sufficient condition for $y \in \mathrm{H}$ to be in $E - \mathrm{H}$ is*

$$P_y\{t = \textstyle\int_0^t I_{\mathscr{I}(E)}(v_n(Z(s)))\, ds \text{ for all } t \geqq 0\} = 1$$

where $v_n(z)$ are given in Definition 2.1.2 *and $I_{\mathscr{I}(E)}(v_n(z)) = 1$ or 0 according as $(v_n(z)) \in \mathscr{I}(E)$ or $(v_n(z)) \notin \mathscr{I}(E)$.*

PROOF. If $y \in E - \mathrm{H}$ is induced by $X(t)$, it follows from Theorem 2.1.4 that $(v_n(Z(s))) = (h_n(X(s))) \in \mathscr{I}(E)$, P-a.s. except for s in a Lebesgue-null set. This means that $(v_n(Z(s))) \in \mathscr{I}(E)$, P_y-a.s. in the same sense, hence $t = \int_0^t I_{\mathscr{I}(E)}(v_n(Z(s)))\, ds$, where the integral is defined by choosing Borel sets $A_1 \subset \mathscr{I}(E) \subset A_2$ with $0 = E_y \int_0^t I_{A_2 - A_1}(v_n(Z(s)))\, ds$. Conversely, let $x_0 \in E$ be fixed, and define a process $X(t)$ on $(\Omega', \mathscr{F}'(0, \infty), P_y)$ by

$$\begin{aligned} X(t) &= \mathscr{I}^{-1}(v_n(Z(t))) && \text{if } (v_n(Z(t))) \in \mathscr{I}(E) \\ &= x_0 && \text{if } (v_n(Z(t))) \notin \mathscr{I}(E)\,. \end{aligned}$$

Then, since $(v_n(Z(t)))$ is $\mathscr{L} \times \mathscr{F}'(0, t+)$-progressively measurable on $\mathsf{X}_1^\infty [0, 1]$, and $\mathscr{I}$ maps universally measurable (respectively, Borel) sets of E onto those of the same type in $\mathsf{X}_1^\infty [0, 1]$, it follows by a routine check that $X(t)$ is $\mathscr{L} \times \mathscr{F}'(0, \infty)$-completion measurable on E. But since $P_y\{t = \int_0^t I_{\mathscr{I}(E)}(v_n(Z(s)))\, ds\} = 1$ and $\mathscr{I}$ is one-to-one on E, we see that $P_y\{h_n(X(s)) = v_n(Z(s))$ for all n and s outside a Lebesgue-null set$\} = 1$. It follows by Theorem 2.1.4 that

$$P_y\{y_n(t) = \textstyle\int_0^t h_n(X(s))\, ds \text{ for all } t \text{ and } n\} = 1\,.$$

This implies that $X(t)$ induces the measure y on Ω', and completes the proof of the lemma.

To complete the proof of the theorem we write the condition of the lemma in the equivalent form

$$t = \int_0^t \int_{\mathrm{H}} q(s, y, dz) I_{\mathscr{I}(E)}(v_n(z))\, ds \tag{2.2.1}$$

for all $t > 0$, or equivalently for t arbitrarily large. Since $\{z : (v_n(z)) \in \mathscr{I}(E)\}$ is universally measurable (respectively Borel) according to the nature of $\mathscr{I}(E)$, and $q(s, y, A)$ is a Borel transition function, it follows by considering measures of the form $\int_0^t \int_{\mathrm{H}} q(s, y, A)\mu(dy)\, ds$ that the right side of condition (2.2.1) is universally measurable (respectively Borel) in y for each t. The rest of the proof is now immediate.

Mostly for the sake of completeness, we include here the following

COROLLARY 2.2.1. *Define* $X(t, (y_n))$ *on* Ω' *by*

$$X(t) = \mathscr{I}^{-1}(w_n(t)) \qquad \text{if} \quad (w_n(t)) \in \mathscr{I}(E)$$
$$= x_0 \qquad \text{if} \quad (w_n(t)) \notin \mathscr{I}(E)\,,$$

where $(w_n(t))$ *is given by Definition* 2.1.1. *Then for every* $y \in E - \mathrm{H}$, $X(t)$ *is* $\mathscr{L} \times \mathscr{F}'(0, \infty)$*-completion measurable and induces* y *as in Theorem* 1.2.2.

PROOF. Since $(w_n(t))$, like $(v_n(Z(t)))$ in Lemma 2.1.1, is $\mathscr{L} \times \mathscr{F}'(0, t+)$-progressively measurable on $\times_1^\infty [0, 1]$, the same check as before shows that $X(t)$ is always $\mathscr{L} \times \mathscr{F}'(0, \infty)$-completion measurable. Again, since $y_n(t) = \int_0^t w_n(s)\, ds$ holds identically, we have by Lemma 2.1.1, $P_y\{v_n(Z(s)) = w_n(s)$ for all n, and s outside of a Lebesgue—null set$\} = 1$. It follows that $P_y\{h_n(X(s)) = w_n(s)$ for all n, and s outside of a Lebesgue—null set$\} = 1$, hence $P_y\{y_n(t) = \int_0^t h_n(X(s))\, ds$ for all n and $t\} = 1$. This completes the proof.

It is clear from (2.2.1) and the semigroup property of Q that for $y \in E - \mathrm{H}$, $q(t, y, E - \mathrm{H}) = 1$ for every $t > 0$. To use $E - \mathrm{H}$ as a reduced state space, however, one needs somewhat more.

THEOREM 2.2.2. *Under Assumption* 2.2.1, *for every* $\mathscr{F}'(0, t+)$*-stopping time* $T < \infty$, $P_y\{Z(T) \in E - H\} = 1$ *for* $y \in E - \mathrm{H}$. *If in addition, E is Borel, then* $P_y\{Z(t) \in E - \mathrm{H}$ *for all* $t \geqq 0\} = 1$.

PROOF. Letting $T \wedge n = \min(T, n)$, we have by (2.2.1) and the strong Markov property of $Z(t)$

$$E_y(T \wedge n + t) = E_y \int_0^{T\wedge n} \left(\int_{\mathrm{H}} q(s, y, dz) I_{\mathscr{I}(E)}(v_n(z))\right) ds$$
$$+ E_y \int_0^t \left(\int_{\mathrm{H}} q(s, Z(T \wedge n), dz) I_{\mathscr{I}(E)}(v_n(z))\right) ds\,.$$

Since the integrands in parentheses are bounded by 1, it follows that for all $t > 0$

$$P_y\{t = \int_0^t \left(\int_{\mathrm{H}} q(s, Z(T \wedge n), dz) I_{\mathscr{I}(E)}(v_n(z))\right) ds\} = 1\,.$$

Letting first $n \to \infty$ and then $t \to \infty$ establishes the first assertion.

If E is Borel, then by Theorem 2.2.1 $\mathscr{I}_{E-\mathrm{H}}(z)$ is also Borel, and by [14, 8, Theorem 16] $\mathscr{I}_{E-\mathrm{H}}(Z(t))$ is well-measurable relative to $\mathscr{F}'(0, t+)$ with all P_y-null sets adjoined. Consequently, Meyer's section theorem [14, 8, Theorem 21] is applicable, and the second assertion is thus a consequence of the first.

REMARK. Even under the condition of the first assertion, one can use $E - \mathrm{H}$

as state space for the standard modification

$$\begin{aligned} Z'(t) &= Z(t) \quad &\text{if} \quad Z(t) \in E - \mathrm{H} \\ &= z_0 \quad &\text{if} \quad Z(t) \notin E - \mathrm{H}\,, \end{aligned}$$

where $z_0 \in E - \mathrm{H}$ is fixed, and it follows that $Z'(t)$ is a strong Markov-process relative to $\mathscr{F}'(0, t+)$ when the latter is completed for P_y, with transition function $q(t, y, A \cap (E - \mathrm{H}))$ on $E - \mathrm{H}$. The right-continuity is of course lost.

3. Two applications: $X(t)$ Markovian and $X(t)$ stationary Gaussian.

3.1. *$X(t)$ Markovian.* The most immediate question concerning the connection of $Z(t)$ and $X(t)$ if $X(t)$ is Markovian is in the case when $X(t)$ has almost the same degree of regularity as $Z(t)$, namely

ASSUMPTION 3.1.1. $(\Omega, \mathscr{F}, \mathscr{F}(t+), X(t), P^x)$, $x \in E$, is a right-continuous strong Markov process with Borel transition function $p(t, x, A)$ on a compact metric space $(E, \mathscr{E})$. The resolvent $R_\lambda f(x) = \int_0^\infty \int_E \exp-(\lambda t)p(t, x, dy)f(y)\,dt$, $0 \leqq f \in \mathscr{E}$, separates points in E.

REMARKS. If E is locally compact with countable base, it may be compactified by a single point Δ in the usual manner. Note that we do not assume the normality condition $P^x\{X(0) = x\} = 1$ nor the existence of left limits. Under Assumption 3.1.1, R_λ is a Ray resolvent [14, 10, Definition 18] if and only if $R_\lambda f \in C(E)$ for $f \in C(E)$, since a countable set $\{R_\lambda f_n, 0 \leqq f_n \text{ dense in } C^+(E)\}$ separates points. In the following theorem we choose $h_n \in C(E)$ in order to connect the topologies of E and $E - \mathrm{H}$.

THEOREM 3.1.1.[3] *Let $\varphi(x)$, $x \in E$, denote the mapping $E \to E - \mathrm{H}$ given by $\varphi(x)(S') = P^x\{Y(\cdot) \in S'\}$, $S' \in \mathscr{F}'(0, \infty)$. We assume that $h_n \in C(E)$, and that $\{h_n\}$ generates a vector space uniformly dense in $C(E)$. The mapping φ is one-to-one and $\mathscr{E}/\mathscr{H}$-measurable. The topology induced by φ on E coincides with the given topology if and only if R_λ maps $C(E) \to C(E)$ (i.e. R_λ is a Ray resolvent). In this case $\varphi(E)$ is compact, $p(t, x, B) = q(t, \varphi(x), \varphi(B))$, $B \in \mathscr{E}$, and $P^x\{Z(t) = \varphi(X(t))$ for all $t \geqq 0\} = 1$.*

PROOF. If $\varphi(x) = \varphi(y)$, then the processes $v_n(Z(t))$ have the same P^x and P^y distributions. Therefore if $f(x_1, \cdots, x_n)$ is continuous on R^n, $R_\lambda f(h_1(x), \cdots, h_n(x)) = E^x \int_0^\infty \exp-(\lambda t)f(v_1(Z(t)), \cdots, v_n(Z(t)))\,dt = R_\lambda f(h_1(y), \cdots, h_n(y))$. As n varies we obtain $R_\lambda f(x) = R_\lambda f(y)$, $f \in C(E)$, and hence $x = y$ since R_λ separates points. Similarly, since $E^x f(Y_1(t_1), \cdots, Y_n(t_n))$ is measurable in x, $\varphi(x)$ is $\mathscr{E}/\mathscr{H}$ measurable.

Now let E_φ denote E with the φ-induced topology. In general, we have

[3] It is not difficult to see from footnote 2 and the Section Theorem of P. A. Meyer that all of the assertions not explicitly involving topology remain true in the general case. Indeed, Theorem 2.1.3 is an instance of this fact. It does not generally follow, however, that $X(t-)$ exists and equals $\varphi^{-1}(Z(t-))$ for all $t > 0$, as is true when the topologies coincide.

$R_\lambda\colon C(E) \to C(E_\varphi)$. To see this, note that

$$E^x \int_{t_1}^{t_2} h_n(X(s))\,ds = E^x(Y_n(t_2) - Y_n(t_1)) = E^{\varphi(x)}(y_n(t_2) - y_n(t_1))$$

where $y_n(t_2) - y_n(t_1)$ is continuous on Ω'. Hence

$$R_\lambda h_n(x) = \lim_{\Delta\to 0} E^x \sum_{k=1}^{\infty} \exp -(k\Delta\lambda) \int_{(k-1)\Delta}^{k\Delta} h_n(X(s))\,ds$$

is continuous on E_φ, since the convergence is uniform in x. But since $\{h_n\}$ generates $C(E)$ the same holds for $R_\lambda f, f \in C(E)$. In particular this implies that the topology of E_φ is as strong as that of E, for if $x_n \to x$ in E_φ then $R_\lambda f(x_n) \to R_\lambda f(x)$, $f \in C(E)$, while if a subsequence $x_{n_j} \to y \neq x$ in E, then $R_\lambda f(x) = R_\lambda f(y)$ would imply that R_λ does not separate x and y. We also note the consequence that if the topologies coincide then $R_\lambda\colon C(E) \to C(E_\varphi) = C(E)$.

To prove the converse, we now need only show that $\varphi(x)$ is continuous on E.

LEMMA 3.1.1. *If $R_\lambda\colon C(E) \to C(E)$, then for $f_1, \cdots, f_n \in C(E)$ and $\lambda_1, \cdots, \lambda_n > 0$, $E^x(\prod_{k=1}^n \int_0^\infty \exp -(\lambda_k t) f_k(X(t))\,dt) \in C(E)$.*

PROOF. For $n = 1$ this is true by hypothesis. For $n > 1$ the function may be expressed as

$$\begin{aligned} n!\, E^x \int_0^\infty \int_{s_1}^\infty \cdots \int_{s_{n-1}}^\infty &\exp -(\lambda_1 s_1) f_1(X(s_1)) \cdots \exp -(\lambda_n s_n) f_n(X(s_n))\,ds_1 \cdots ds_n \\ &= n!\, E^x \int_0^\infty \exp -(\lambda_1 s_1) f_1(X(s_1)) E^{X(s_1)} \\ &\quad \times \left(\int_0^\infty \int_{t_2}^\infty \cdots \int_{t_{n-1}}^\infty \left(\prod_{k=2}^n \exp -\lambda_k(t_k + s_1) f_k(X(t_k))\right) dt_2 \cdots dt_n\right) ds_1 \\ &= n R_{\lambda_1+\cdots+\lambda_n}\left(f_1(x) E^x \prod_{k=2}^n \int_0^\infty \exp -(\lambda_k t) f_k(X(t))\,dt\right). \end{aligned}$$

By successive application of this reduction the expression becomes

$$n!\, R_{\lambda_1+\cdots+\lambda_n}(f_1(x)(R_{\lambda_2+\cdots+\lambda_n} f_2(x)(\cdots f(x_{n-1})(R_{\lambda_n} f_n(x)) \cdots))),$$

and by repeatedly using the case $n = 1$ it follows that this is in $C(E)$.

Integrating by parts, we have

$$\int_0^\infty \exp -(\lambda t) h_n(X(t))\,dt = \lambda \int_0^\infty \left(\exp -(\lambda t) \int_0^t h_n(X(s))\,ds\right) dt,$$

and since the integral is bounded on Ω it follows by Lemma 3.1.1 and the Weierstrass approximation theorem that the joint distributions of $\int_0^\infty \exp -(\lambda_k t) \int_0^t h_k(X(s))\,ds\,dt$, $1 \leqq k \leqq n$, are weakly continuous in x. The same is then also true for any n expressions of the form $\int_0^\infty (c_1 \exp -(\lambda_1 t) + \cdots + c_j \exp -(\lambda_j t)) \int_0^t h_k(X(s))\,ds\,dt$. By the Stone–Weierstrass theorem, the terms $c_1 \exp -(\lambda_1 t) + \cdots + c_j \exp -(\lambda_j t)$ are uniformly dense in $C_0[0, \infty)$ (continuous functions with limit 0 at ∞), and it follows that, for any $g \in C_0[0, \infty)$ and $\varepsilon > 0$, the joint distribution of $\int_0^\infty g_k(t) \exp -(\varepsilon t)(\int_0^t h_k(X(s))\,ds)\,dt$, $1 \leqq k \leqq n$, are weakly continuous in x. Finally, choosing $0 \leqq g_k(t)$ with support in $t_k - \delta$, $t_k + \delta$ and $\int_0^\infty g_k(t)\,dt = 1$, where δ is small, we obtain the joint weak continuity in x of $\int_0^{t_k} h_k(X(s))\,ds$, $1 \leqq k \leqq n$. As in the proof of Lemma 1.3.2, an application of Stone–Weierstrass now shows that $\varphi(x)$ is continuous.

In this case, the compactness of $\varphi(E)$ is well known, and the identification of p with q is immediate. It is easily seen that since $X(t)$ is right continuous, $\mathscr{F}(0, t+) = \mathscr{F}^0(t+)$ [1, page 31]; hence we have $P\{\varphi(X(t)) = Z(t)\} = 1$ for each t by the Markov property and the uniqueness in Theorem 1.3.1. Since both $Z(t)$ and $\varphi(X(t))$ are right-continuous in H, the last assertion follows by taking right limits in t along the rationals.

In the situation where $X(t)$ has less regularity than in Assumption 3.1.1, we may still regard $Z(t)$ as a kind of regularized version of $X(t)$. A particularly explicit connection is obtained if we choose h_n in the manner of [12]. Thus if we begin with any transition function $p(t, x, A)$ measurable in (t, x) for $x \in E$, where $(E, \mathscr{E})$ is an abstract space with a separable σ-field, and assume that the resolvent $R_\lambda f$ separates points, then starting with any generating sequence h_n as in 1.1 we can consider the family $\{r_i R_{r_i} h_n\}$ where r_i is an enumeration of the positive rationals. If we now form the algebraic closure of this set under the two operations (i) applications of $r_i R_{r_i}$ for any r_i, and (ii) multiplication, the resulting set S is still denumerable, although it need not contain $\{h_n\}$. It determines a uniform structure (in fact, a metric) on E and the compactification of E is denoted by $\bar{E}$. We obtain a resolvent on $C(\bar{E})$ by extending $R_\lambda f(x)$, $f \in S$, $x \in E$, to $\bar{E}$ by continuity, and observing that $C(\bar{E})$ is the uniform linear closure of the extensions of $f \in S$. Since this resolvent separates points it is a Ray resolvent, and hence we obtain by Ray (1959) a transition function $\bar{p}(t, x, B)$ on $\bar{E}$, and a canonical Markov process satisfying Assumption 3.1.1 on $\bar{E}$ together with the added property $\bar{R}_\lambda : C(\bar{E}) \to C(\bar{E})$. Let us state the connection of this with the original space E as

THEOREM 3.1.2. *Let the generating sequence h_n consist of an enumeration of the set S defined above. Then given any $\mathscr{L} \times \mathscr{F}$-completion measurable Markov process $(\Omega, \mathscr{F}, P; X(t))$ on $(E, \mathscr{E})$ with transition function p, there is associated a new process $\bar{X}(t) = \lim_{r \to t+} X(r)$, where the limit exists in $\bar{E}$ along the rationals r at all $t \geqq 0$, P-a.s., and $\bar{X}(t)$ is a right-continuous strong Markov process on $\bar{E}$, with transition function $\bar{p}$. Letting $\bar{\varphi}(\bar{x})$ denote the mapping of Theorem 3.1.1 for $\bar{E}$, and $Z(t)$, $\bar{Z}(t)$ denote the prediction processes of $X(t)$ and $\bar{X}(t)$, we have $P\{Z(t) = \bar{Z}(t) = \bar{\varphi}(\bar{X}(t))$ for all $t \geqq 0\} = 1$. Moreover, the closure of $\bar{\varphi}(E)$ in H is $\bar{\varphi}(\bar{E})$. Thus formation of the prediction processes $Z(t)$ is equivalent to formation of the associated Ray processes $\bar{X}(t)$, and the Ray compactification $\bar{E}$ is (homeomorphic to) the closure of $\bar{\varphi}(E)$.*

REMARK. If it is known that to each $x \in E$ there corresponds a measurable Markov process $X(t)$ with $P^x\{X(t) \in A\} = p(t, x, A)$ and p as transition function, then we can define $\varphi(x)(S') = P^x\{Y(\cdot) \in S'\}$, and it follows that $\varphi(x) = \bar{\varphi}(x)$, $x \in E$. Thus one can proceed directly to the closure of $\varphi(E)$ in this case. In general this seems to require some further hypothesis on E.

PROOF. The existence of $\bar{X}(t)$ follows in [12, page 549] from martingale convergence, and we have $P\{X(t) = \bar{X}(t)\} = 1$ except for t in a countable set (namely, where one of the countably many $f(\bar{X}(t))$, $f \in S$, has a fixed discontinuity as in [5, Theorem 11.2 (ii)]). It follows that $X(\cdot)$ and $\bar{X}(\cdot)$ induce the same measure $y \in \mathrm{H}$, hence $Z(\cdot)$ and $\bar{Z}(\cdot)$ are indistinguishable. Since E is dense in $\bar{E}$ and the topologies $\bar{E}$ and $\bar{E}_{\bar{\varphi}}$ coincide (Theorem 3.1.1) the last assertion is immediate.

3.2. *$X(t)$ Stationary Gaussian.* Let $X(t)$ be a stationary real-valued measurable Gaussian process with mean 0 and continuous covariance $R(t-s) = E(X(t)X(s))$, $-\infty < t < \infty$. The prediction theory of such processes has been highly developed, most recently by H. Dym and H. P. McKean, whose survey (1970) will provide all of the facts and additional references needed here. We first remark that since $X(t)$ is continuous in quadratic mean, the σ-fields $\mathscr{F}(-\infty, t+)$ and $\mathscr{F}^0(t+)$ have the same completion. This follows since, for each t, there is a sequence $\Delta_n \to 0$ with $P\{h_n(X(t)) = \lim_{n\to\infty} \Delta_n^{-1}(Y_n(t+\Delta_n) - Y_n(t))\} = 1$. Moreover, the construction of $Z(t)$ applies without any essential change for $-\infty < t < \infty$. We shall only prove one result, which indicates, at any rate, that the discontinuities which may be present in $X(t)$ do not locally carry any information having nonlocal significance.

THEOREM 3.2.1. *$P\{Z(t)$ is continuous for $t \geqq 0\} = 1$.*

PROOF. Set $R(t-s) = \int \exp i\gamma(t-s)\Gamma(d\gamma)$, where $\Gamma(d\gamma)$ is the spectral measure. It is known (the Szegö alternative) that unless

$$\int (1+\gamma^2)^{-1} \ln (\Gamma'(\gamma))\, d\gamma > -\infty \tag{3.2.1}$$

where Γ' denotes the density of the absolutely continuous component of Γ, $X(t)$ is perfectly predictable given $\lim_{t\to-\infty} \mathscr{F}^0(t)$, and hence $Z(t)$ is clearly continuous.

If (3.2.1) holds, then we can write $\Gamma = \Gamma_1 + \Gamma_2$, where Γ_1 is the singular part, and $X(t) \equiv X_1(t) + X_2(t)$, where X_1 and X_2 are independent, X_1 is perfectly predictable, and X_2 is independent of $\lim_{t\to-\infty} \mathscr{F}^0(t)$. Hence for the induced measures of

$$y_n(s) = \int_0^s h_n(X_1(t+\tau) + X_2(t+\tau))\, d\tau \quad \text{given} \quad \mathscr{F}^0(t+),$$

the function $X_1(t+\tau)$ is nonrandom, and does not effect the continuity of $Z(t)$. Thus we may assume that $X_1(t) \equiv 0$.

For fixed T and $t < T$ we consider the process $E(X(T) \mid \mathscr{F}^0(t))$, $-\infty < t \leqq T$, also called the Kolmogorov–Wiener predictor of $X(T)$. It is clearly a centered Gaussian process with orthogonal, hence independent, increments. Moreover [6, pages 1820–21], we can write $\Gamma'(\gamma) = |h(\gamma)|^2$, where $h \in L^2(d\gamma, R)$. Then if $\hat{h} \in L^2(d\gamma, R)$ denotes the Fourier transform of h, we have [6, loc. cit.]

$$E(E(X(T) \mid \mathscr{F}^0(t))^2) = \int_{T-t}^{\infty} |\hat{h}(s)|^2\, ds\,. \tag{3.2.2}$$

Denoting this expression by $V(T-t)$, is follows that $E(X(T) \mid \mathscr{F}^0(t))$ is stochastically equivalent to $B(V(T-t))$ where $B(s)$ is ordinary Brownian motion. In particular, we may choose $E(X(T) \mid \mathscr{F}^0(t))$ to have continuous paths, and then $E(X(T)\,|\,\mathscr{F}^0(t)) = E(X(T)\,|\,\mathscr{F}^0(t+))$. Taking into account that $E(X(T)\,|\,\mathscr{F}^0(t+)) = X(T)$ for $t \geqq T$, we obtain continuous paths for $-\infty < t < \infty$. The conditional variances $E(X^2(T)\,|\,\mathscr{F}^0(t+)) - E^2(X(T)\,|\,\mathscr{F}^0(t+))$ are known to be non-random functions of t, equalling $\int_0^{T-t} |\hat{h}(s)|^2\,ds$ for $T \geqq t$.

More generally, applying the same reasoning to $X(T_2) - X(T_1)$, we see that for any $T_1 < T_2 < \cdots < T_n$, given $\mathscr{F}^0(t+)$ the random variables $X(T_1), \cdots, X(T_n)$ are jointly Gaussian with expectations and covariance continuous in t. Thus their conditional joint density varies continuously with t. It follows that for every $\Delta > 0$ and $0 \leqq f(x_1, \cdots, x_n) \in C_0^n$, the quantity $S_\Delta(t) = E(f(\Delta \sum_{k=1}^{n_1} h_1(X(k\Delta)), \cdots, \Delta \sum_{k=1}^{n_n} h_n(X(k\Delta)))\,|\,\mathscr{F}^0(t+))$ varies continuously with t, where $n_j = [T_j \Delta^{-1}]$. On the other hand, by the classical L^2 martingale maximal inequality [5, 7, Theorem 3.4] we have for $T > T_n$

$$\begin{aligned} &E(\max_{0<t<T} |S_\Delta(t) - E(f(\textstyle\int_0^{T_1} h_1(X(s))\,ds, \cdots, \\ (3.2.3)\qquad &\textstyle\int_0^{T_n} h_n(X(s))\,ds)|\,|\,\mathscr{F}^0(t+)) \\ &\qquad \leqq 4E(S_\Delta(T) - f(\textstyle\int_0^{T_1} h_1(X(s))\,ds, \cdots, \int_0^{T_n} h_n(X(s))\,ds))^2 \end{aligned}$$

Let us assume temporarily that the h_n are continuous. Then as $\Delta \to 0$ the right side tends to 0 since $h_j(X(s))$ are continuous in quadratic mean. A completely analogous argument shows also that for $\tau_1 < T_1, \cdots, \tau_n < T_n$, the expression

$$E(f(\textstyle\int_{\tau_1}^{T_1} h_1(X(s))\,ds, \cdots, \int_{\tau_n}^{T_n} h_n(X(s))\,ds)\,|\,\mathscr{F}^0(t+))$$

varies continuously in t, and by a simple triangle inequality one sees that the same is true with $\tau_1 = \cdots = \tau_n = t$ and T_j replaced by $t + T_j$, $1 \leqq j \leqq n$. By Theorem 2.1.1 and the uniqueness of $Z(t)$, this means that $Z(t)$ has continuous paths for any continuous (h_n). It follows from Theorem 2.1.2, however, that continuity of path for $Z(t)$ does not depend on the choice of (h_n). Thus the proof is complete.

Let us consider two final examples to indicate possible lines of further investigation. Suppose first that

$$\begin{aligned} X(t) &= B(t) && \text{if } 0 \leqq t \leqq 1 \\ &= B(1) + (t-1)X && \text{if } t > 1\,, \end{aligned}$$

where $B(t)$ is ordinary Brownian motion and X is a fixed normal random variable independent of $B(t)$. Then $X(t)$ is a Gaussian process (non-stationary) with continuous paths. However, it is clear that the corresponding $Z(t)$ has a fixed jump discontinuity at $t = 1$. As a second example, let $P(t)$ be an ordinary Poisson process on the nonnegative integers ($P(0) = 0$), and let $X(t) = \int_0^t P(s)\,ds$. Again $X(t)$ has continuous paths, but now the discontinuities of $P(t)$ become totally inaccessible discontinuities of the prediction process $Z(t)$, in the sense of Meyer (1966). In this case, it appears that $Z(t)$ gives a more realistic topological indication of the probabilistic behavior of the underlying process than does $X(t)$.

REFERENCES

[1] BLUMENTHAL, R. M. and GETOOR, R. K. (1968). *Markov Processes and Potential Theory.* Academic Press, New York.

[2] BOURBAKI, N. (1958). Topologie Général (Chapter 9). Hermann, Paris.

[3] CHUNG, K. L. and WALSH, J. B. (1970). To reverse a Markov process. Acta Math. **123** 225–251.

[4] DELLACHERIE, C. (1971). Les théorèmes de Mazurkiewicz–Sierpinski et de Lusin. Séminaire de probabilités **5**, Université de Strasbourg. *Lecture Notes in Mathematics* **191**. Springer-Verlag, Berlin.

[5] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[6] DYM, H. and MCKEAN, H. P. (1970). Extrapolation and interpolation of stationary Gaussian processes. Ann. Math. Statist. **41** 1817–1843.

[7] DYNKIN, E. B. (1965). *Markov Processes I.* Springer-Verlag, New York.

[8] FELLER, W. (1966). *An Introduction to Probability Theory and Its Applications* **2**. Wiley, New York.

[9] GIKHMAN, I. I. and SKOROKHOD, A. V. (1969). *Introduction to the Theory of Random Processes.* Saunders, Philadelphia.

[10] HALMOS, P. R. (1950). *Measure Theory.* Van Nostrand, Princeton.

[11] KNIGHT, F. B. (1963). Markov processes on a entrance boundary. *Illinois J. Math.* **7** 322–336.

[12] KNIGHT, F. B. (1965). Note on regularization of Markov processes. *Illinois J. Math.* **9** 548–552.

[13] LOÈVE, M. (1963). *Probability Theory*, 3rd ed. Van Nostrand, Princeton.

[14] MEYER, P. A. (1966). *Probability and Potentials.* Blaisdell, Waltham.

[15] MEYER, P. A. (1967). *Processus de Markov. Lecture Notes in Mathematics* **26**. Springer-Verlag, Berlin.

[16] MEYER, P. A. (1968). Guide détaillé de la théorie des processus. Séminaire de probabilités 2, Université de Strasbourg. *Lecture Notes in Mathematics* **51**, Springer-Verlag, Berlin.

[17] MEYER, P. A. and WALSH, J. B. (1971). Quelques applications des résolvantes de Ray. *Invent. Math.* **14** 143–166.

[18] RAY, D. (1959). Resolvents, transition functions, and strongly Markovian processes. *Ann. of Math.* **70** 43–75.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT CHAMPAIGN—URBANA
URBANA, ILLINOIS 61801