

ADAPTED PROBABILITY DISTRIBUTIONS

BY

DOUGLAS N. HOOVER¹ AND H. JEROME KEISLER²

ABSTRACT. We introduce a family of notions of equivalence for stochastic processes on spaces with an underlying filtration. These refine the notion of having the same distribution by taking account of the relation of the processes to their underlying filtrations. The weakest of these notions is the same as the notion of synonymity introduced by Aldous. Analysis of the strongest equivalence property leads to spaces with a strong universality property for adapted stochastic processes, which we call saturation. Spaces having this property contain 'strong' solutions to a large class of stochastic integral equations.

Two random variables are alike if they have the same distribution, and two Markov processes are alike if they have the same finite dimensional distribution. However, in general two stochastic processes may have the same finite dimensional distribution but behave quite differently. This is particularly true in the "Strasbourg" setting where a stochastic process $x(t, \omega)$ lives on an adapted space (or stochastic base), i.e. a probability space (Ω, P) endowed with a filtration \mathcal{F}_t , $t \in \mathbf{R}^+$. The finite dimensional distribution of x does not depend at all on the filtration \mathcal{F}_t .

In this paper we introduce the *adapted distribution* of a stochastic process, which plays the same role for stochastic processes that the distribution plays for random variables. The adapted distribution of a process $x(t, \omega)$ on an adapted space is defined as follows. (For simplicity assume that x is bounded.) We consider the family of new stochastic processes in finitely many time variables obtained from x by iterating the following two operations: (1) composition by continuous functions, and (2) conditional expectation with respect to \mathcal{F}_t . An example is the process

$$\sqrt{E[(x(t))^2 | \mathcal{F}_s]}.$$

Another example with two iterations of the expected value operator is

$$E[\exp(x(s) - E[x(t) | \mathcal{F}_s]) | \mathcal{F}_r].$$

Two stochastic processes x and y (on perhaps different adapted spaces) have the same adapted distribution if each pair of new processes obtained from x and y by iterating (1) and (2) have the same finite dimensional distribution.

Received by the editors June 27, 1983.

1980 *Mathematics Subject Classification*. Primary 60E05, 60G05; Secondary 60T10, 60G48, 03B48.

¹The first author is an NSERC Canada University Research Fellow, and his research was supported by an NSERC Operating Grant.

²Research supported by a National Science Foundation Grant.

By restricting to n iterations of the expected value operator we obtain a sequence of stronger and stronger equivalence relations $x \equiv_n y$, $n = 0, 1, 2, \dots$, between stochastic processes. The weakest relation $x \equiv_0 y$, says that x and y have the same finite dimensional distribution. The next relation, $x \equiv_1 y$, says that x and y are synonymous in the sense of Aldous [1981]. x and y have the same adapted distribution, $x \equiv y$, if $x \equiv_n y$ for all n . For each n , \equiv_{n+1} strictly refines \equiv_n .

Aldous [1981] has shown that several important properties of stochastic processes are preserved under the synonymy relation $x \equiv_1 y$, including the property of being adapted, being a martingale, and the Markov property. The first author has shown that local martingale and semimartingale properties are also preserved under synonymy. We show, however, that the most powerful property of \equiv , the existence of spaces with a saturation property, fails for synonymy.

The objective of this paper is to support the thesis that two processes with the same adapted distribution share the same probabilistic properties. Our results show this for stochastic integration, provided the processes live on a sufficiently rich adapted space.

There is a close analogy between the adapted distribution of a stochastic process and the complete theory of a structure in the sense of model theory. The model-theoretic aspects of the adapted distribution are discussed in Rodenhausen [1982] and Keisler [1983].

§1 of this paper is a brief summary of notation. §2 contains the definition and basic properties of the adapted distribution. In addition to the relation $x \equiv y$ of having the same adapted distribution, we introduce the weaker relation $x \equiv y$ a.e. of having the same adapted distribution for almost all $t \in \mathbf{R}^+$. It is shown that if x and y are Markov processes, then $x \equiv_0 y$ implies $x \equiv y$. If x and y are right continuous, or even right continuous in probability, then $x \equiv y$ a.e. implies $x \equiv y$.

§3 contains examples showing that $x \equiv_n y$ is weaker than $x \equiv_{n+1} y$ for each n . In particular, $x \equiv_1 y$ is weaker than $x \equiv y$.

In §4, we begin the study of rich adapted spaces by defining and constructing complete and atomless adapted spaces. The key existence proof uses nonstandard analysis.

§5 explains what we mean by a sufficiently rich adapted space. An adapted space $\underline{\Omega}$ is said to be *universal* if for every process x on any adapted space $\underline{\Lambda}$ there is a process \hat{x} on $\underline{\Omega}$ such that $x \equiv \hat{x}$. $\underline{\Omega}$ is called *saturated* if for every pair of processes (x, y) on any adapted space $\underline{\Lambda}$ and any process \hat{x} on $\underline{\Omega}$ with $x \equiv \hat{x}$, there is a process \hat{y} on $\underline{\Omega}$ with $(x, y) \equiv (\hat{x}, \hat{y})$. Saturated implies universal. Saturated spaces are convenient because given any stochastic process on the space, there is always room in the space for a second stochastic process having a given relation to the first process. Familiar adapted spaces, such as measures on $C[0, \infty)$ or $D[0, \infty)$ with the natural filtrations, are not saturated. The proof of the existence of saturated adapted spaces uses the results from §4. We also prove that any equivalence relation which preserves the martingale property and has a saturated space is at least as strong as \equiv .

§§6 and 7 contain applications of saturated adapted spaces. In §6 we show that on saturated spaces the local martingale and semimartingale properties depend only on the adapted distribution. More precisely, if x is a local martingale on an adapted space $\underline{\Omega}$, and \hat{x} is an r.c.l.l. process on a saturated adapted space $\underline{\Omega}$ with $x \equiv \hat{x}$, then \hat{x} is a local martingale. The result for semimartingales is similar. §7 contains results in the same vein for stochastic integrals and stochastic integral equations with respect to semimartingales.

1. Notation. As usual, \mathbf{R}^+ is the set of nonnegative real numbers and N the set of natural numbers. By an *adapted (probability) space*, or *stochastic base*, we mean a structure

$$\underline{\Omega} = (\Omega, P, \mathcal{F}_t)_{t \in \mathbf{R}^+},$$

where P is a complete probability measure on the sample space Ω , and \mathcal{F}_t is a σ -algebra on Ω for each $t \in \mathbf{R}^+$ with the following properties (called the *usual conditions*).

Right continuity: $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s$.

Compatibility with P : $\mathcal{F}_t \subset \mathcal{F}_\infty =$ the set of P -measurable sets.

Completeness of \mathcal{F}_0 : Each null set of P belongs to \mathcal{F}_0 .

The family $(\mathcal{F}_t)_{t \in \mathbf{R}^+}$ is called a *filtration* on (Ω, P) if it satisfies the usual conditions.

A *Polish space* M is a complete separable metrizable topological space.

Since we wish to compare stochastic processes on different adapted spaces, we want each stochastic process to carry on with it an adapted space.

By a *stochastic process* (x) (on $\underline{\Omega}$ with values in M) we mean a structure $(x) = (\underline{\Omega}, x)$ where $\underline{\Omega} = (\Omega, P, \mathcal{F}_t)_{t \in \mathbf{R}^+}$ is an adapted space, M is a Polish space, and x is a function $x: \Omega \times \mathbf{R}^+ \rightarrow M$ which is measurable with respect to the product of \mathcal{F}_∞ and Borel sets on \mathbf{R}^+ , and Borel sets on M .

We shall usually write x for the process (x) . x and y always stand for stochastic processes.

Suppose x and y are processes on the same adapted space $\underline{\Omega}$. x is *indistinguishable* from y if for almost all $\omega \in \Omega$,

$$x(\cdot, \omega) = y(\cdot, \omega).$$

x is a *version* of y if for all $t \in \mathbf{R}^+$,

$$x(t, \omega) = y(t, \omega) \quad \text{a.s.}$$

x is an *a.e. version* of y if the above holds for almost all $t \in \mathbf{R}^+$. By an r.c.l.l. process we mean a stochastic process $x(\omega, t)$ such that almost every path $x(\omega, \cdot)$ is right continuous with left limits.

By an *n-fold stochastic process* on $\underline{\Omega}$ we mean a measurable function

$$f: \Omega \times (\mathbf{R}^+)^n \rightarrow M.$$

A tuple of reals is denoted by \vec{t} , and the length of \vec{t} is denoted by $|\vec{t}|$. The indicator function of a set S is denoted by $1(S)$.

2. Adapted distributions. In this section we define our basic equivalence relation $x \equiv y$, which states that the stochastic processes x and y have the same adapted distribution. We let M be a Polish space which remains fixed throughout our discussion. In order to define the adapted distribution we introduce a family of functions f , called conditional processes, which associate with each stochastic process x on $\underline{\Omega}$ an n -fold stochastic process fx on $\underline{\Omega}$. The family of conditional processes is defined inductively and the first step in the induction determines the usual finite dimensional distribution of x .

DEFINITION 2.1. For each n , each bounded continuous function $\Phi: M^n \rightarrow \mathbf{R}$ and each stochastic process x with values in M , $\hat{\Phi}x$ is the n -fold stochastic process

$$\hat{\Phi}x(t_1, \dots, t_n) = \Phi(x_{t_1}, \dots, x_{t_n}).$$

REMARK 2.2. Two stochastic processes x and y have the same finite dimensional distribution if and only if

$$E[\hat{\Phi}x(t_1, \dots, t_n)] = E[\hat{\Phi}y(t_1, \dots, t_n)]$$

for all Φ and all t_1, \dots, t_n .

The finite dimensional distribution of x depends only on (Ω, P, x) and not on the filtration \mathcal{F}_t . Our next notion depends strongly on \mathcal{F}_t .

DEFINITION 2.3. The class CP of *conditional processes* (in M) is defined inductively as follows.

(2.3.1) (*Basis*) For each n and bounded continuous $\Phi: M^n \rightarrow \mathbf{R}$, $\hat{\Phi} \in \text{CP}$.

(2.3.2) (*Composition*) If $f_1, \dots, f_m \in \text{CP}$ and $\varphi: \mathbf{R}^m \rightarrow \mathbf{R}$ is a bounded continuous function, then $\varphi(f_1, \dots, f_m) \in \text{CP}$, where $\varphi(f_1, \dots, f_m)x = \varphi(f_1x, \dots, f_mx)$.

(2.3.3) (*Conditional Expectation*) If f is an n -fold conditional process $fx(t_1, \dots, t_n)$, then $E[f|t]$ is an $n+1$ -fold conditional process, where $E[f|t]x(t, t_1, \dots, t_n)$ is a version of $E[fx(t_1, \dots, t_n)|\mathcal{F}_t]$.

Each conditional process is obtained from basic processes by iterating the composition and conditional expectation operations finitely many times. Since only bounded continuous functions are used, the conditional expectations are always finite, and each conditional process is uniformly bounded. Since conditional expectations are only unique up to a null set, the value of a conditional process at x is only unique up to a version. Here is a typical example of a conditional process with two iterations of the conditional expectation operator.

EXAMPLE 2.4. $f = E[(E[\sin(\hat{t}_1)|t_2])^2|t_3]$ is the 3-fold conditional process

$$fx(t_1, t_2, t_3) = E\left[\left(E[\sin(x(t_1))|\mathcal{F}_{t_2}]\right)^2\middle|\mathcal{F}_{t_3}\right].$$

The number of iterations of the unexpected value operation in f is called the rank of f . Formally, the rank is defined by induction.

DEFINITION 2.5.

(2.5.1) For each n and Φ , the conditional process $\hat{\Phi}$ has rank zero.

(2.5.2) The rank of the composition $\varphi(f_1, \dots, f_m)$ is the maximum of the ranks of the conditional processes f_1, \dots, f_m .

(2.5.3) If f is a conditional process of rank r , then $E[f|t]$ is a conditional process of rank $r + 1$.

The conditional process f in Example 2.4 has rank two. We are now ready to introduce our main notion.

DEFINITION 2.6. Two stochastic processes x and y (on perhaps different adapted spaces) have the same *adapted distribution* (or *adapted law*), in symbols, $x \equiv y$, if

$$(2.6.1) \quad E[fx(t_1, \dots, t_n)] = E[fy(t_1, \dots, t_n)]$$

for every n -fold conditional process f and all $t_1, \dots, t_n \in \mathbf{R}^+$.

x and y have *almost the same adapted distribution*, $x \equiv y$ a.e., if for every f (2.6.1) holds for almost all $t_1, \dots, t_n \in (\mathbf{R}^+)^n$.

x and y have the same adapted distribution up to rank r , $x \equiv_r y$, if (2.6.1) holds for every f of rank at most r and all t_1, \dots, t_n . Thus $x \equiv y$ if and only if $x \equiv_r y$ for all r .

Given a subset T of \mathbf{R}^+ , x and y have the same adapted distribution on T , $x \equiv_T y$ on T , if (2.6.1) holds for every f and all $t_1, \dots, t_n \in T^n$.

We state three easy facts.

PROPOSITION 2.7. (i) If x and y are processes on the same adapted space and y is a version of x , i.e. $x(t) = y(t)$ a.s. for all t , then $x \equiv y$.

(ii) For any x and y , $x \equiv_0 y$ if and only if x and y have the same finite dimensional distributions.

(iii) If $\mathbf{R}^+ - T$ is a null set and $x \equiv y$ on T then $x \equiv y$ a.e.

Aldous [1981] studied the relation $x \equiv_1 y$ ($|x$ and y are synonymous). He showed that various properties of stochastic processes are preserved by synonymity. In particular, if x is \mathcal{F}_t -adapted, \mathcal{F}_t -Markov, or an \mathcal{F}_t -martingale and $x \equiv_1 y$, then y is \mathcal{G}_t -adapted, \mathcal{G}_t -Markov, or a \mathcal{G}_t -martingale, respectively. These results suggest at first sight that the synonymity relation $x \equiv_1 y$ should play the role for adapted spaces which equality in distribution plays for probability spaces. We shall show in this paper that $x \equiv_1 y$ is not strong enough for this purpose, but that the stronger relation $x \equiv y$ is. We shall also see that for each r , $x \equiv_{r+1} y$, and hence $x \equiv y$, is strictly stronger than $x \equiv_r y$.

We say that a stochastic process x is *Markov* if x is \mathcal{F}_t -Markov, that is, $x(t)$ is \mathcal{F}_t -measurable and for every bounded continuous $\Phi: M \rightarrow \mathbf{R}$ and $s < t$,

$$E[\Phi(x(t))|\mathcal{F}_s] = E[\Phi(x(t))|x(s)] \quad \text{a.s.}$$

We now show that if x is Markov, then the adapted distribution of x depends only on the finite dimensional distribution of x .

THEOREM 2.8. Suppose x, y are Markov. Then $x \equiv y$ if and only if $x \equiv_0 y$.

PROOF. Let $x \equiv_0 y$. We shall show by induction that for every n -fold conditional process f and every $\vec{t} \in (\mathbf{R}^+)^n$ there is a bounded Borel function $\psi_{f,\vec{t}}: M^n \rightarrow \mathbf{R}$ such that

$$(2.8.1) \quad \begin{aligned} fx(\vec{t}) &= \psi_{f,\vec{t}}(x(t_1), \dots, x(t_n)) \quad \text{a.s.,} \\ fy(\vec{t}) &= \psi_{f,\vec{t}}(y(t_1), \dots, y(t_n)) \quad \text{a.s.} \end{aligned}$$

For each bounded continuous function $\Phi: M^n \rightarrow \mathbf{R}$, the function $\psi_{\Phi, \vec{t}} = \Phi$ has the required property (2.8.1). If $f \in \text{CP}$ has the form $f = \varphi(f_1, \dots, f_m)$ we take

$$\psi_{f, \vec{t}} = \varphi(\psi_{f_1, \vec{t}}, \dots, \psi_{f_m, \vec{t}}).$$

This takes care of the basis step and the composition step in the induction. For the conditional expectation step, let $g = E[f|s]$ where f is an n -fold conditional process and suppose $s \in \mathbf{R}^+$, $\vec{t} \in (\mathbf{R}^+)^n$ and $\psi_{f, \vec{t}}$ satisfies (1). Since x is Markov and $\psi_{f, \vec{t}}$ is bounded,

$$\begin{aligned} E[\psi_{f, \vec{t}}(x(t_1), \dots, x(t_n)) | \mathcal{F}_s] \\ = E[\psi_{f, \vec{t}}(x(t_1), \dots, x(t_n)) | x(s), x(t_1), \dots, x(t_j)] \quad \text{a.s.} \end{aligned}$$

where $t_1 < \dots < t_j < s < t_{j+1}$.

Therefore there is a bounded Borel function $\psi_{g, s\vec{t}}$ on M^{n+1} such that

$$(2.8.2) \quad E[\psi_{f, \vec{t}}(x(t_1), \dots, x(t_n)) | \mathcal{F}_s] = \psi_{g, s\vec{t}}(x(s), x(t_1), \dots, x(t_n)) \quad \text{a.s.}$$

(In fact, $\psi_{g, s\vec{t}}$ depends only on $x(s)$ and those $x(t_i)$ where $t_i \leq s$.) Applying (2.8.1) for fx , we have

$$\begin{aligned} gx(s\vec{t}) &= E[fx(\vec{t}) | \mathcal{F}_s] = E[\psi_{f, \vec{t}}(x(t_1), \dots, x(t_n)) | \mathcal{F}_s] \\ &= \psi_{g, s\vec{t}}(x(s), x(t_1), \dots, x(t_n)) \quad \text{a.s.} \end{aligned}$$

Thus (2.8.1) holds for gx . Since x and y are Markov processes with the same f.d.d., y has the same transition functions as x . Therefore (2.8.2) and hence (2.8.1) hold for y with the same choice of $\psi_{g, s\vec{t}}$. This completes the induction. Finally, since $x \equiv_0 y$ we have

$$\begin{aligned} E[fx(\vec{t})] &= E[\psi_{f, \vec{t}}(x(t_1), \dots, x(t_n))] \\ &= E[\psi_{f, \vec{t}}(y(t_1), \dots, y(t_n))] = E[fy(\vec{t})] \end{aligned}$$

for all $f \in \text{CP}$ and all \vec{t} , whence $x \equiv y$. \square

We next show that if x and y are right continuous in probability then $x \equiv y$ a.e. implies that $x \equiv y$ (Corollary 2.15).

DEFINITION 2.9. Let z be an n -fold stochastic process with values in a complete separable metric space (M, d) . z is *right continuous in probability* if for every $\vec{t} \in (\mathbf{R}^+)^n$ and every $\varepsilon > 0$,

$$\lim_{\vec{u} \downarrow \vec{t}} P[d(z(\vec{u}), z(\vec{t})) \geq \varepsilon] = 0$$

where $\vec{t} \leq \vec{u}$ means $t_1 \leq u_1, \dots, t_n \leq u_n$.

Notice that if $d(z(\vec{u}), z(\vec{t}))$ is uniformly bounded then this is equivalent to the condition

$$\lim_{\vec{u} \downarrow \vec{t}} E[d(z(\vec{u}), z(\vec{t}))] = 0.$$

If almost every path of z is right continuous, then z is right continuous in probability.

THEOREM 2.10. *Suppose x is right continuous in probability. Then for every n -fold conditional process f , fx is right continuous in probability.*

PROOF. The proof is by induction on the construction of f . The only nontrivial step is the conditional expectation operation. Suppose $f = E[g|s]$ where g is an n -fold conditional process and gx is right continuous in probability. To simplify notation let $n = 1$, so g is one-fold and f is two-fold. Fix $s, t \in \mathbf{R}^+$, so

$$(2.10.1) \quad fx(s, t) = E[gx(t)|\mathcal{F}_s].$$

Then

$$(2.10.2) \quad \lim_{(u,v) \downarrow (s,t)} E[|fx(u, v) - fx(s, t)|] \leq \lim_{(u,v) \downarrow (s,t)} E[|fx(u, v) - fx(u, t)|] \\ + \lim_{u \downarrow s} E[|fx(u, t) - fx(s, t)|].$$

We have

$$\begin{aligned} E[|fx(u, v) - fx(u, t)|] &= E[|E[gx(v)|\mathcal{F}_u] - E[gx(t)|\mathcal{F}_u]|] \\ &= E[|E[gx(v) - gx(t)|\mathcal{F}_u]|] \\ &\leq E[E[|gx(v) - gx(t)||\mathcal{F}_u]] \\ &= E[|gx(v) - gx(t)|]. \end{aligned}$$

Therefore

$$(2.10.3) \quad \lim_{(u,v) \downarrow (s,t)} E[|fx(u, v) - fx(u, t)|] \leq \lim_{v \downarrow t} E[|gx(v) - gx(t)|] = 0.$$

On the other hand, since

$$h(u) = fx(u, t) = E[gx(t)|\mathcal{F}_u]$$

is a martingale, it has a right continuous version and hence is right continuous in probability. Thus

$$(2.10.4) \quad \lim_{u \downarrow s} E[|fx(u, t) - fx(s, t)|] = 0.$$

From (2.10.2)–(2.10.4) it follows that

$$\lim_{(u,v) \downarrow (s,t)} E[|f(u, v) - f(s, t)|] = 0. \quad \square$$

If the process x has the stronger property that for each t

$$\lim_{u \downarrow t} x(u) = x(t) \quad \text{a.s.,}$$

it can be shown that fx has this property for every conditional function f . However, if x has the still stronger property that almost every path of x is right continuous, we do not know whether fx must have that property (jointly in n variables).

In the following lemma we will consider a slightly extended notion of conditional process. A *Borel conditional process* is defined as in 2.3 except that in the basis step (2.3.1) of the formation of a Borel conditional process, we will allow any bounded Borel function $M^n \rightarrow \mathbf{R}$ to be used.

LEMMA 2.11. *Let f be an n -fold Borel conditional process. Let (Φ_1, \dots, Φ_m) be the Borel functions used in building f in the basis step (2.3.1), and let $\varphi_1, \dots, \varphi_k$ be the continuous functions used in building f in the composition clause (2.3.2). Suppose that as $j \rightarrow \infty$*

$$\begin{aligned} (\Phi_1^j, \dots, \Phi_m^j) &\rightarrow (\Phi_1, \dots, \Phi_m) \text{ pointwise,} \\ (\varphi_1^j, \dots, \varphi_k^j) &\rightarrow (\varphi_1, \dots, \varphi_k) \text{ uniformly on compact sets,} \end{aligned}$$

where Φ_i^j are Borel, φ_i^j are continuous, and all are uniformly bounded. Let f^j be the Borel conditional process formed in the same way as f but with $\Phi_1^j, \dots, \Phi_m^j, \varphi_1^j, \dots, \varphi_k^j$ in place of $\Phi_1, \dots, \Phi_m, \varphi_1, \dots, \varphi_k$. Then $f^j x(t) \rightarrow f x(t)$ a.s. for every $t \in (\mathbf{R}^+)^n$.

PROOF. We argue by induction on the construction of f . The basis step (2.3.1) is trivial. For the composition step (2.3.2), we will assume, to simplify notation, that $f = \varphi(g)$ for some $g \in \text{CP}$ and bounded continuous $\varphi: \mathbf{R} \rightarrow \mathbf{R}$. By induction hypothesis, $g^j x(t) \rightarrow g x(t)$ a.s. If we take any ω such that $g^j x(t, \omega) \rightarrow g x(t, \omega)$, $\{g^j x(t, \omega), g x(t, \omega) : j \in \mathbf{N}\}$ is a compact set. Since $\varphi^j \rightarrow \varphi$ uniformly on compact sets, it follows easily that $\varphi^j(g^j x(t, \omega)) \rightarrow \varphi(g x(t, \omega))$. For clause (2.3.3), let $f = E[g(\vec{t})|s]$ and suppose $g^j x(\vec{t}) \rightarrow g x(\vec{t})$ a.s. for all \vec{t} . Since $g x$ is uniformly bounded, it follows by dominated convergence that

$$E[g^j x(\vec{t})|\mathcal{F}_s] \rightarrow E[g x(\vec{t})|\mathcal{F}_s] \text{ a.s.}$$

for all s, \vec{t} , that is

$$f^j x(s, \vec{t}) \rightarrow f x(s, \vec{t}) \text{ a.s. } \square$$

DEFINITION 2.12. (i) If S is a set of real-valued functions, the set generated by S by bounded pointwise convergence is the smallest set of functions containing S and closed under pointwise convergence of uniformly bounded sequences of functions.

(ii) For each n , choose a countable set of bounded functions $S_n \subseteq C(M^n, \mathbf{R})$, such that the class of bounded Borel functions $M^n \rightarrow \mathbf{R}$ is the set generated by S_n by bounded pointwise convergence, and a countable set of bounded functions $U_n \subseteq C(\mathbf{R}^n, \mathbf{R})$ such that for each integer m , $U_n \cap C(\mathbf{R}^n, [-m, m])$ is dense in the compact-open topology on $C(\mathbf{R}^n, [-m, m])$. Let CP^0 be the set of conditional processes formed using only functions from the countable family

$$\bigcup_{n=1}^{\infty} S_n \cup \bigcup_{n=1}^{\infty} U_n.$$

Thus CP^0 is a countable subset of CP .

COROLLARY 2.13. *Let $T \subset \mathbf{R}^+$.*

(i) *$x \equiv_r y$ on T if and only if $E[f x(\vec{t})] = E[f y(\vec{t})]$ for all n -fold $f \in \text{CP}^0$ of rank $\leq r$ and all $\vec{t} \in T^n$.*

(ii) *Let x and y be right continuous in probability and let T be dense in \mathbf{R}^+ . Then $x \equiv y$ if and only if $x \equiv y$ on T .*

(iii) *Let T be countable and let $(x_t)_{t \in T}$ and $(y_t)_{t \in T}$ be the M^T -valued random variables*

$$(x_t)_{t \in T}(\omega) = (x(t, \omega))_{t \in T}, \quad (y_t)_{t \in T}(\omega) = (y(t, \omega))_{t \in T}.$$

$(x_t)_{t \in T}$ and $(y_t)_{t \in T}$ may be considered as M^T -valued processes with constant sample paths.

Then $x \equiv y$ on T if and only if $(x_t)_{t \in T} \equiv (y_t)_{t \in T}$ on T .

PROOF. (i) The proof follows from Lemma 2.11.

(ii) The proof follows from Theorem 2.10.

(iii) Observe that every conditional process $fx(\vec{t})$ with $\vec{t} \in T$ can be identified with a conditional process $g(x_s)_{s \in T^*}(t)$. Conversely, any conditional process $g(x_s)_{s \in T^*}(\vec{t})$, $\vec{t} \in T$, formed using only bounded continuous functions $(M^T)^n \rightarrow \mathbf{R}$ which depend on only finitely many coordinates of M^T may be identified with a conditional process $fx(\vec{s}, \vec{t})$, $\vec{s} \in T$. Since the bounded Borel functions $(M^T)^n \rightarrow \mathbf{R}$ are generated by the finite dimensional bounded continuous functions by pointwise convergence, the conclusion follows by Lemma 2.11. \square

PROPOSITION 2.14. Suppose $x \equiv y$ a.e. and let μ be a measure on \mathbf{R}^+ which is absolutely continuous with respect to Lebesgue measure. Then (1) $x \equiv y$ on T and (2) $(x, x_t)_{t \in T} \equiv (y, y_t)_{t \in T}$ a.e. for μ^N -almost every sequence $T = \{t_n: n \in \mathbf{N}\} \subseteq \mathbf{R}^+$. In particular, there is a countable dense set T satisfying (1) and (2). The same holds for \equiv , in place of \equiv .

PROOF. Since $x \equiv y$ a.e., for each n (1) and (2) hold for μ^n -almost every sequence $T_n = \{t_1, \dots, t_n\}$. By the Fubini theorem, for each $n \in \mathbf{N}$ and μ^N -almost every sequence $\{t_k: k \in \mathbf{N}\}$, $T_n = \{t_1, \dots, t_n\}$ satisfies (1) and (2). Thus for μ^N -almost every sequence $T = \{t_k: k \in \mathbf{N}\}$, each $T_n = \{t_1, \dots, t_n\}$ satisfies (1) and (2), and hence T satisfies (1) and (2).

If μ is the measure given by

$$\mu[a, b] = \int_a^b e^{-t} dt,$$

then μ^N -almost every sequence $T = \{t_n: n \in \mathbf{N}\}$ is dense in \mathbf{R}^+ , so (1) and (2) hold for a countable dense T . \square

COROLLARY 2.15. If $x \equiv y$ a.e. and x and y are right continuous in probability then $x \equiv y$.

PROOF. By Theorem 2.10. \square

REMARK. We shall use the fact that certain properties of stochastic processes depend only on the finite dimensional distributions, that is are preserved under the relation $x \equiv_0 y$. For example:

(i) If x is right (left) continuous in probability and $x \equiv_0 y$, then y is right (left) continuous in probability.

(ii) If for each $t \in \mathbf{R}^+$, x is a.s. right (left) continuous at t and $x \equiv_0 y$, then y has a version z such that for each $t \in \mathbf{R}^+$, z is a.s. right (left) continuous at t .

(iii) If every path of x is right continuous with left limits and $x \equiv_0 y$, then y has a version z such that every path of z is right continuous with left limits.

(iv) If every path of x is a right (left) continuous step function and $x \equiv_0 y$, then y has a version z all of whose paths are right (left) continuous step functions.

DEFINITION 2.16. The stochastic processes $m_r x$, $r = 1, 2, \dots$, and mx are defined by

$$\begin{aligned} m_r x(t) &= \langle E[fx(\bar{s})|\mathcal{F}_t] : f \in \text{CP}^0, f \text{ has rank} < r, \bar{s} \text{ is rational} \rangle, \\ mx(t) &= \langle E[fx(\bar{s})|\mathcal{F}_t] : f \in \text{CP}^0, \bar{s} \text{ rational} \rangle. \end{aligned}$$

For each t and ω , the value of $mx(t, \omega)$ is a countable sequence of reals and will be identified with an element of the Polish space $\mathbf{R}^{\mathbf{N}}$ with the product topology.

PROPOSITION 2.17. (i) $mx(t)$ and $m_r x(t)$ have right continuous, left limit, \mathcal{F}_t -adapted versions.

(ii) $mx(t)$ is \mathcal{F}_t -Markov.

PROOF. $mx(t)$ and $m_r x(t)$ are \mathcal{F}_t -martingales in each coordinate, and (i) follows.

To prove (ii) it suffices to show that for each n -tuple h_1, \dots, h_n of coordinates of mx , each bounded continuous function $\psi: \mathbf{R}^n \rightarrow \mathbf{R}$ from a countable dense set, and each $s < t$ in \mathbf{R}^+ ,

$$(2.17.1) \quad E[\psi(h_1(t), \dots, h_n(t))|\mathcal{F}_s] = E[\psi(h_1(t), \dots, h_n(t))|mx(s)] \quad \text{a.s.}$$

From the definition of mx we see that for each $i \leq n$,

$$h_i(t) = E[f_i x(\bar{u}_i)|\mathcal{F}_t]$$

where $f_i \in \text{CP}^0$ and \bar{u}_i is rational. Moreover, if $\psi \in U_n$,

$$\psi(h_1(t), \dots, h_n(t)) = fx(t, \bar{u})$$

for some $f \in \text{CP}^0$ and rational \bar{u} . For each rational v ,

$$E[\psi(h_1(v), \dots, h_n(v))|\mathcal{F}_s] = E[fx(v, \bar{u})|\mathcal{F}_s]$$

is itself a coordinate of mx and is therefore $mx(s)$ -measurable. Hence (2.17.1) holds whenever t is rational. Moreover, since $mx(t)$ has a right continuous version, and ψ is bounded,

$$E[\psi(h_1(t), \dots, h_n(t))|\mathcal{F}_s] = \lim_{\substack{r \downarrow t \\ r \in Q}} E[\psi(h_1(r), \dots, h_n(r))|\mathcal{F}_s] \quad \text{a.s.}$$

by dominated convergence. Hence (2.17.1) holds for all t . \square

The next result shows that if x is right continuous in probability, the adapted distribution of x is determined by the finite dimensional distribution of the Markov process mx .

PROPOSITION 2.18. Let x and y be right continuous in probability.

(i) For each $r \geq 1$ and $k \in \mathbf{N}$,

$$m_r x \equiv_k m_r y$$

if and only if $x \equiv_{r+k} y$.

(ii) $x \equiv y$ if and only if $mx \equiv_0 my$.

PROOF. (i) implies (ii), since if $mx \equiv_0 my$, then $m_r x \equiv_0 m_r y$ for all $r \geq 1$, so $x \equiv y$ by (i).

To prove (i), suppose first that $m_r x \equiv_k m_r y$. By Lemma 2.11, we may further restrict h to be in CP^0 . To prove that $x \equiv_{r+k} y$ it will, by the Stone-Weierstrass Theorem and Corollary 2.13, suffice to show that $E[fz(\vec{t})] = E[fy(\vec{t})]$ for \vec{t} rational and f of rank $\leq r + k$ of the form

$$\varphi(E[g_i|t_i]: i \leq n) \cdot \psi(\Phi_i(t_i): i \leq n),$$

where $\vec{t} = t_1, \dots, t_n$.

Now for any \vec{t} and process z ,

$$E[fz(\vec{t})] = E[E[fz(\vec{t})|s]].$$

If s is chosen to be greater than or equal to any parameter in \vec{t} , then

$$E[fz(\vec{t})|s] = \varphi(E[g_i z(\vec{t})|t_i]: i \leq n) E[\psi(\Phi_i z(t_i): i \leq n)|s].$$

Since $r \geq 1$, $E[fz(\vec{t})|s]$ has rank $\leq r + k$. $E[fz(\vec{t})|s]$ also has the property that any subexpression of the form $\Phi(v)$ occurs only inside some subexpression of the form $E[g|u]$. Thus it suffices to check that

$$E[hx(\vec{s})] = E[hy(\vec{s})]$$

for $\vec{s} \in \mathcal{Q}$ and h having this property. But for such h one may easily show that for any $\vec{t} \in \mathcal{Q}$ there is a conditional process h' on \mathbb{R}^N of rank $\leq k$ such that $hz(\vec{t}) = h'm_r z(\vec{t})$ for any z . Thus

$$E[hx(\vec{t})] = E[h'm_r x(\vec{t})] = E[h'm_r y(\vec{t})] = E[hy(\vec{t})]$$

and it follows that $x \equiv_{r+k} y$.

Now assume $x \equiv_{r+k} y$ and show that $m_r x \equiv_k m_r y$. For each n , the functions with support on only finitely many coordinates generate the Borel functions: $(\mathbb{R}^N)^n \rightarrow \mathbb{R}$ by bounded pointwise convergence. Hence by Lemma 2.11 it suffices to show that $E[fm_r x(\vec{t})] = E[fm_r y(\vec{t})]$ for f of rank $\leq k$ formed using only such functions. But for every such f there are a conditional process g on M of rank $\leq k + r$ and rationals \vec{s} such that $fm_r z(\vec{t}) = gz(\vec{t}, \vec{s})$ for every M -valued process z . Since $x \equiv_{r+k} y$ it follows that $m_r x \equiv_k m_r y$. \square

Notice that by 2.8 and 2.18, $x \equiv y$ if and only if $mx \equiv my$. The process $m_1 x$ is similar to the Knight prediction process [1975] as modified by Aldous [1981], and the processes $m_r x$, mx are closely related to the processes obtained by finite and infinite iterations of the construction of $m_1 x$. Knight confines his attention to the case where \mathcal{F}_t is the filtration generated by $x(t)$, and shows that in that case one does not get anything new by iterating the prediction process. In our notation, he shows that if $x, y \in \mathcal{A}$ are such that the filtrations \mathcal{F}_t of x and \mathcal{G}_t of y are generated by $x(t)$ and $y(t)$, then $x \equiv_1 y$ if and only if $x \equiv_2 y$, and hence if and only if $x \equiv y$. In the next section we show that this fails in general.

LEMMA 2.19. Let (M, ρ) be a Polish space and let x_n, y_n be processes on the same adapted space with values in M such that for all t ,

$$\rho(x_n(t), y_n(t)) \rightarrow 0$$

a.s. or in probability. Then for every $\vec{t} \in (\mathbf{R}^+)^n$ and every conditional process f ,

$$|fx_n(\vec{t}) - fy_n(\vec{t})| \rightarrow 0$$

a.s. or in probability.

PROOF. Similar to Lemma 2.11. \square

PROPOSITION 2.20. Let x_n, x, y_n, y be processes such that:

(i) All the x_n 's and x are on the same adapted space and for all t , $\lim_{n \rightarrow \infty} x_n(t) = x(t)$ in probability.

(ii) All the y_n 's and y are on the same adapted space and for all t , $\lim_{n \rightarrow \infty} y_n(t) = y(t)$ in probability.

(iii) For each $n \in \mathbf{N}$, $x_n \equiv y_n$.

Then $x \equiv y$.

PROOF. By Lemma 2.19, for each conditional process f and each $\vec{t} \in (\mathbf{R}^+)^n$ we have

$$\begin{aligned} fx_n(\vec{t}) &\rightarrow fx(\vec{t}) \quad \text{in probability,} \\ fy_n(\vec{t}) &\rightarrow fy(\vec{t}) \quad \text{in probability.} \end{aligned}$$

Therefore

$$E[fx_n(\vec{t})] \rightarrow E[fx(\vec{t})], \quad E[fy_n(\vec{t})] \rightarrow E[fy(\vec{t})].$$

Since $x_n \equiv y_n$, $E[fx_n(\vec{t})] = E[fy_n(\vec{t})]$, so $E[fx(\vec{t})] = E[fy(\vec{t})]$. \square

The above result also holds with \equiv replaced everywhere by \equiv_k , by the same proof.

The main results concerning adapted distribution in §5 are most conveniently proved for r.c.l.l. processes. We can extend them to both \equiv and \equiv_k a.e. for general processes by showing how \equiv and \equiv_k a.e. can be expressed in terms of \equiv for certain associated r.c.l.l. processes.

DEFINITION 2.21. Let Φ_n , $n \in \mathbf{N}$, be a fixed sequence of bounded continuous functions $M \times \mathbf{R}^+ \rightarrow \mathbf{R}$ which generate the Borel-measurable functions $M \times \mathbf{R}^+ \rightarrow \mathbf{R}$. Define

$$Sx: \Omega \times \mathbf{R}^+ \rightarrow \mathbf{R}^{\mathbf{N}},$$

the sojourn process of x , by

$$Sx(t) = \left(\int_0^t \Phi_n(x(s), s) ds : n \in \mathbf{N} \right). \quad \square$$

Sx is a process on the same adapted space as x , and since each coordinate of Sx is an integral, Sx is continuous.

THEOREM 2.22. *For any processes x and y , the following are equivalent.*

(1) $x \equiv y$ a.e.

(2) $Sx \equiv Sy$.

(3) $(x, Sx) \equiv (y, Sy)$ a.e.

The same holds for \equiv_r , $r \in \mathbf{N}$.

PROOF. There is no difference in the \equiv and \equiv_r cases. It suffices to show (1) \Rightarrow (3) and (2) \Rightarrow (3), since (3) \Rightarrow (1), (2) is trivial.

(1) \Rightarrow (3): We show that there is a measurable function $\varphi: M^{\mathbf{N}} \rightarrow C(\mathbf{R}^+, \mathbf{R})^{\mathbf{N}}$ and a countable set $T \subseteq \mathbf{R}^+$ such that

$$(x, x_t; t \in T) \equiv (y, y_t; t \in T) \quad \text{a.s.}$$

and

$$Sx = \varphi(x_t; t \in T) \text{ a.s.}, \quad Sy = \varphi(y_t; t \in T) \text{ a.s.}$$

Give $[0, u)$ the uniform probability. By the conditional form of the Strong Law of Large Numbers, for each bounded continuous $\Phi: M \times \mathbf{R}^+ \rightarrow \mathbf{R}$, each $u \in \mathbf{R}^+$, and almost all sequences $\{t_n; n \in \mathbf{N}\} \subseteq [0, u)$,

$$\int_0^u \Phi(x(s), s) ds = \lim_{N \rightarrow \infty} \frac{u}{N} \sum_{n=1}^N \Phi(x(t_n), t_n) \quad \text{a.s.}$$

It follows that for almost all $\{t_n; n \in \mathbf{N}\} \subseteq [0, u)$, $Sx(u)$ is a.s. $\sigma(x(t_n), n \in \mathbf{N})$ -measurable. The same goes for y . Let $\{q_k; k \in \mathbf{N}\}$ enumerate the rationals. By the foregoing and Proposition 2.14 we may successively choose for each $k \in \mathbf{N}$ a sequence $\{t_n^k, n \in \mathbf{N}\}$ in $[0, q_k)$ such that $Sx(q_k)$ and $Sy(q_k)$ are a.s. $\sigma(x(t_n^k))$ - and $\sigma(y(t_n^k))$ -measurable, respectively, and

$$(x, x(t_n^j); j \leq k, n \in \mathbf{N}) \equiv (y, y(t_n^j); j \leq k, n \in \mathbf{N}) \quad \text{a.s.}$$

Then $T = \{t_n^k; k, n \in \mathbf{N}\}$ is a countable set such that Sx is a.s. $\sigma(x(t); t \in T)$ -measurable, Sy is $\sigma(y(t); t \in T)$ -measurable, and

$$(4) (x, x_t; t \in T) \equiv (y, y_t; t \in T) \text{ a.s.}$$

It follows that there is a Borel-measurable function φ such that

$$\varphi(x_t; t \in T) = Sx \text{ a.s.}, \quad \varphi(y_t; t \in T) = Sy \text{ a.s.}$$

By (4), it follows that $(x, Sx) \equiv (y, Sy)$ a.e.

(2) \Rightarrow (3): consider x and Sx as measurable functions on the space $\Omega \times \mathbf{R}^+$. For each open set $U \subseteq M$, the increasing, absolutely continuous process

$$z_t = \int_0^t 1(x(s) \in U) ds$$

is Sx -measurable and $\{x \in U\}$ is a.e. the set where the Radon-Nikodym derivative with respect to $P \times ds$ of the measure on $(\Omega \times \mathbf{R}^+, \sigma(Sx) \times \mathcal{B})$ induced by z is 1. It follows that there is a Borel-measurable function $\varphi: C(\mathbf{R}^+, \mathbf{R}^{\mathbf{N}}) \times \mathbf{R}^+ \rightarrow M$ such that $x = \varphi(Sx, \cdot)$ a.e. $P \times ds$. Let

$$\hat{x} = \varphi(Sx, \cdot), \quad \hat{y} = \varphi(Sy, \cdot).$$

Since $\hat{x} = x$ a.e., it follows easily that $Sx = S\hat{x}$. By the proof of (1) \Rightarrow (3), this implies that

$$(\hat{x}, Sx, S\hat{x}) \equiv (\hat{y}, Sy, S\hat{y}),$$

hence $S\hat{y} = Sy$ a.s. But this just says that

$$\int \Phi(y(s), s) ds = \int \Phi(\hat{y}(s), s) ds \quad \text{a.s.}$$

for every measurable function Φ , so $y = \hat{y}$ a.e. $\mathcal{Q} \times ds$. Since $(\hat{x}, Sx) \equiv (\hat{y}, Sy)$ it follows that

$$(x, Sx) \equiv (y, Sy) \quad \text{a.e.} \quad \square$$

COROLLARY 2.23. *If $x \equiv y$ a.e. (or $x \equiv_r y$ a.e.) then there are processes \hat{x}, \hat{y} such that*

$$x = \hat{x} \text{ a.e.}, \quad y = \hat{y} \text{ a.e.}, \quad \hat{x} \equiv \hat{y} \quad (\text{or } \hat{x} \equiv_r \hat{y}).$$

PROOF. Let \hat{x}, \hat{y} be as in the proof of Theorem 2.22. \square

PROPOSITION 2.24. *If x is a stochastic process then there is an r.c.l.l. process*

$$z: \mathbf{R}^+ \times \Omega \rightarrow \{0, 1\}^{\mathbf{N}}$$

and a Borel-measurable function $\varphi: \{0, 1\}^{\mathbf{N}} \rightarrow M$ such that for all $t \in \mathbf{R}^+$

$$\varphi(z(t)) = x(t) \quad \text{a.s.}$$

If $x \equiv y$, we may choose φ and r.c.l.l. processes z, w such that $z \equiv w$ and $x(t) = \varphi(z(t))$ a.s. and $y(t) = \varphi(w(t))$ a.s. for all $t \in \mathbf{R}^+$.

PROOF. First observe that for every $\mathcal{F}_\infty \times \mathcal{B}$ -measurable set S , there is a $\sigma(x(t): t \in \mathbf{R}^+)$ -measurable process e such that for every $t \in \mathbf{R}^+$,

$$e(t) = E[S(t) | \sigma(x_t: t \in \mathbf{R}^+)] \quad \text{a.s.}$$

This is trivial for finite unions of measurable rectangles. Since sections and conditional expectations commute with monotone limits, it follows for all $\mathcal{B} \times \mathcal{F}_\infty$ -measurable sets by the monotone class theorem. Now if S is $\{x \in U\}$ for U an open set of M , $S(t)$ is $\sigma(x(t): t \in \mathbf{R}^+)$ -measurable for each t , so for all t , $S(t) = e(t)$ a.s. It follows that x has a $\mathcal{B} \times \sigma(x_t: t \in \mathbf{R}^+)$ -measurable version \hat{x} . Thus there must be $F_n \in \sigma(x(t_n)), t_n \in \mathbf{R}^+, a_n, b_n \in \mathbf{R}^+$, such that

$$\hat{x} = \varphi(1([a_n, b_n) \times F_n): n \in \mathbf{N})$$

for some measurable function from $\{0, 1\}^{\mathbf{N}}$ to M . Now the process

$$z = (1([a_n, b_n) \times F_n): n \in \mathbf{N})$$

is r.c.l.l. and of the form

$$z(t) = \psi(t, (x(t_n): n \in \mathbf{N}))$$

for some measurable function ψ which is r.c.l.l. in t . Since \hat{x} is a version of x , $x(t) = \varphi(z(t))$ a.s. for all t . Now if w is given by $w = \psi(t, (y(t_n): n \in \mathbb{N}))$ then since $x \equiv_0 y$, for all t

$$y(t) = \varphi(w(t)) \quad \text{a.s.}$$

Clearly z , w and φ satisfy the conclusions of the proposition. \square

We finish this section with a result which shows that CP^0 is a convergence determining class for a notion of convergence in adapted distribution.

DEFINITION 2.25. Let x_n , $n \in \mathbb{N}$, and x be stochastic processes on $(\underline{\Omega} \times \mathbb{R}^+) \rightarrow M$, and let $T \subseteq \mathbb{R}^+$.

(2.25.1) We write

$$x_n \xrightarrow{\text{ad}} x \quad \text{on } T,$$

x_n converges in adapted distribution to x on T , if $E[f x_n(\vec{t})] \rightarrow E[f x(\vec{t})]$ for each conditional process f and \vec{t} in T .

(2.25.2) We write

$$x_n \xrightarrow{\text{ad}} \quad \text{on } T$$

if $E[f x_n(\vec{t})]$ converges for each conditional process f and \vec{t} in T .

(2.25.3) The sequence $x_n(t)$ is said to be *tight* (see Billingsley [1968]) if for each real $\varepsilon > 0$ there is a compact set $K \subseteq M$ such that $P[x_n(t) \in K] \geq \varepsilon$ for each $n \in \mathbb{N}$.

The notion $x_n \xrightarrow{\text{ad}} x$ also makes sense if the x_n and x are on different adapted spaces. If $x_n \xrightarrow{\text{ad}} x$ on T , then $x_n(t)$ converges in distribution to $x(t)$ for each t in T , and hence by Prokhorov's theorem, the sequence $x_n(t)$ is tight.

THEOREM 2.26. Let x , x_n , $n \in \mathbb{N}$, be adapted processes, not necessarily on the same adapted space, and let T be a subset of \mathbb{R}^+ such that, for each $t \in T$, $\{x(t), x_n(t)\}$ is tight, and for every $f \in \text{CP}^0$ and $\vec{t} \in T$

$$E[f x(\vec{t})] \rightarrow E[f x(\vec{t})],$$

then $x_n \xrightarrow{\text{ad}} x$ on T .

PROOF. We show by induction on formation that for each $t \in T$, $f \in \text{CP}$, and $\varepsilon > 0$, there is f^ε in CP^0 such that for sufficiently large n ,

$$(2.26.1) \quad E[|f^\varepsilon x_n(t) - f x_n(t)|] < \varepsilon \quad \text{and} \quad E[|f^\varepsilon x(t) - f x(t)|] < \varepsilon.$$

The basis case is most difficult. If $\vec{t} = t_1, \dots, t_k$, let us write $x(\vec{t}) = (x(t_1), \dots, x(t_k))$. We will first show that, for each $t \in T$, $x_n(t)$ converges to $x(t)$ in distribution. Because products of compact sets are compact, $\{x_n(\vec{t}), x(\vec{t})\}$ is tight. By Prokhorov's Theorem, any subsequence of $x_n(\vec{t})$ has a subsequence which converges in distribution, so it suffices to show that every subsequence of $x_n(\vec{t})$ which converges in distribution converges to $x(\vec{t})$. But if $E[\Phi(x_n(\vec{t}))] \rightarrow E[\Phi(z)]$ for some z , and for every bounded continuous Φ , then

$$E[\Phi(x(\vec{t}))] = E[\Phi(z)]$$

for every $\Phi \in S_k$, hence for every bounded Borel function $\Phi: M^k \rightarrow \mathbf{R}$, since S_k generates these by bounded pointwise convergence. Thus we have in fact $x_{n_i}(\vec{t}) \rightarrow x(\vec{t})$ in distribution. For each bounded continuous $\Phi: M_k \rightarrow \mathbf{R}$, choose now Φ^ϵ in S_k such that

$$E\left[|\Phi^\epsilon(x(\vec{t})) - \Phi(x(\vec{t}))|\right] < \epsilon/2.$$

Then since

$$E\left[|\Phi^\epsilon(x_n(\vec{t})) - \Phi(x_n(\vec{t}))|\right] \rightarrow E\left[|\Phi^\epsilon(x(\vec{t})) - \Phi(x(\vec{t}))|\right]$$

as $n \rightarrow \infty$, (2.26.1) holds for Φ and Φ^ϵ for sufficiently large n . This completes the basis case. Now suppose f is of the form $\varphi(g_1, \dots, g_p)$, where φ is bounded and uniformly continuous. For notational simplicity, take $p = 1$, $g_1 = g$, and assume that φ is bounded by 1. Now g has bounded range. Hence we may choose a bounded continuous φ^ϵ in u_n , also bounded by 1, which uniformly approximates φ within ϵ on a compact set K containing an ϵ -neighbourhood of the range of g . Choose $\delta < \min(1, \epsilon)$ so that $|a - b| < \delta$ implies both $|\varphi(a) - \varphi(b)| < \epsilon$ and $|\varphi^\epsilon(a) - \varphi^\epsilon(b)| < \epsilon$. Then if g^{δ^2} is as in (2.26.1), then for sufficiently large n ,

(2.26.2)

$$P(|g^{\delta^2}x_n(t) - gx_n(t)| < \delta) < 1 - \delta, \quad P(|g^{\delta^2}x(t) - gx(t)| < \delta) < 1 - \delta.$$

If $|g^{\delta^2}x_n(t) - gx_n(t)| < \delta$, then $g^{\delta^2}x_n(t)$ is in K , so

$$\begin{aligned} E\left[|\varphi^\epsilon(g^{\delta^2}x_n(t)) - \varphi(gx_n(t))|\right] &\leq E\left[|\varphi^\epsilon(g^{\delta^2}x_n(t)) - \varphi(g^{\delta^2}x_n(t))|\right] \\ &\quad + E\left[|\varphi(g^{\delta^2}x_n(t)) - \varphi(gx_n(t))|\right] \\ &\leq (\epsilon + 2\delta) + (\epsilon + 2\delta) \quad \text{by (2.26.2)} \\ &\leq 6\epsilon. \end{aligned}$$

(2.26.1) now follows by ϵ -juggling. The case $f = E[g|t]$ follows immediately by Jensen's inequality. \square

3. Some counterexamples. In this section we shall give finite examples showing that the equivalence relation $x \equiv y$ is stronger than any of the relations $x \equiv_n y$ for $n \in \mathbf{N}$. In fact, each $x \equiv_n y$ is stronger than $x \equiv_{n-1} y$. This shows that each additional iteration of the expected value operator $E[f|t]$ distinguishes more properties of a stochastic process. The examples proceed inductively from the following simple example.

EXAMPLE 3.1. Two processes x^1 and y^1 on the same adapted space $\underline{\Omega}^1$ such that $x^1 \equiv_0 y^1$ but not $x^1 \equiv_1 y^1$.

Let $\underline{\Omega}^1 = (\Omega^1, P^1, \mathcal{F}_t^1)$ be the following adapted space:

$$\Omega^1 = \{1, 2\}^{\{1, 2\}} = \{(1, 1), (1, 2), (2, 1), (2, 2)\}.$$

P^1 is the counting probability measure on Ω^1 .

For $0 \leq t < 1$, \mathcal{F}_t^1 is the algebra $\mathcal{F}_t^1 = \{\emptyset, \Omega^1\}$. For $1 \leq t < 2$, \mathcal{F}_t^1 is the algebra

$$\mathcal{F}_t^1 = \{\emptyset, \Omega^1, \{(1, 1), (1, 2)\}, \{(2, 1), (2, 2)\}\}.$$

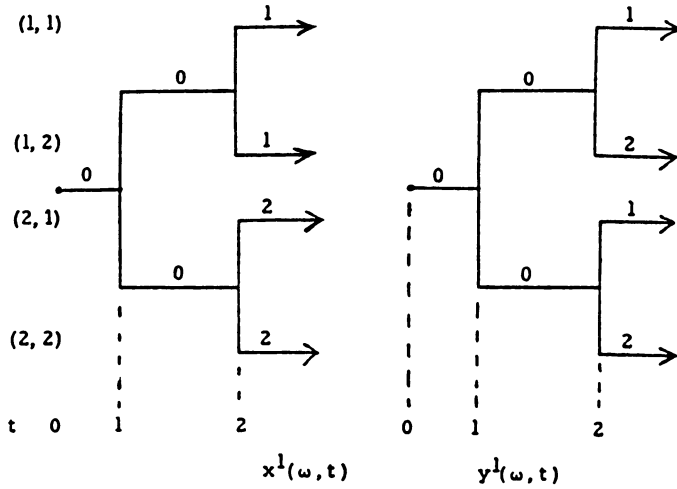
For $2 \leq t$, \mathcal{F}_t^1 is the algebra of all subsets of Ω^1 .

The processes x^1 and y^1 are defined by

$$x^1(\omega, t) = \begin{cases} 0 & \text{if } t < 2, \\ \omega(1) & \text{if } t \geq 2, \end{cases}$$

$$y^1(\omega, t) = \begin{cases} 0 & \text{if } t < 2, \\ \omega(2) & \text{if } t \geq 2. \end{cases}$$

In the following figure we identify the four elements of Ω^1 with branches of a binary tree and indicate the values of $x^1(\omega, t)$ and $y^1(\omega, t)$ along each branch.



Let h be the bijection of Ω^1 defined by $h(\omega) = (\omega(2), \omega(1))$. In the figure, h interchanges the middle two branches. We see at once that

$$(3.1.1) \quad x^1(h\omega, t) = y^1(\omega, t).$$

It follows that x^1 and y^1 have the same f.d.d., that is,

$$(3.1.2) \quad x^1 \equiv_0 y^1.$$

On the other hand we have

$$E[x^1(\cdot, 2) | \mathcal{F}_1^1](\omega) = \begin{cases} 1 & \text{if } \omega(1) = 1, \\ 2 & \text{if } \omega(1) = 2 \end{cases}$$

while

$$E[y^1(\cdot, 2) | \mathcal{F}_1^1](\omega) = \frac{3}{2} \quad \text{for all } \omega.$$

It follows that

$$(3.1.3) \quad x^1 \not\equiv_1 y^1.$$

EXAMPLE 3.2. For each $n \in \mathbb{N}$, two processes x^n and y^n on the same adapted space $\underline{\Omega}^n$ such that $x^n \equiv_{n-1} y^n$ but not $x^n \equiv_n y^n$.

Assume $n > 1$. Let $\Omega^n = \{1, 2\}^{\{1, 2, \dots, 2n\}}$. Let P^n be the counting probability measure on Ω^n . Let \mathcal{F}_t^n be the set of all $A \subset \Omega^n$ such that whether $\omega \in A$ depends only on $\omega(i)$ for $i \leq t$. In particular,

$$\mathcal{F}_t^n = \{\emptyset, \Omega^n\} \quad \text{for } 0 \leq t < 1$$

and

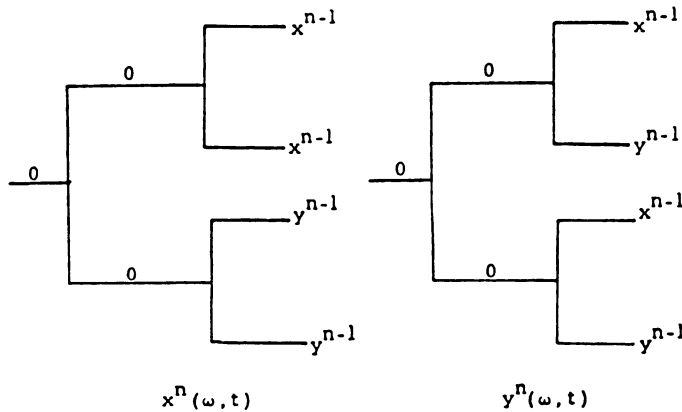
$$\mathcal{F}_t^n = \mathcal{P}(\Omega^n) \quad \text{for } 2n \leq t.$$

Let $\pi: \Omega^n \rightarrow \Omega^{n-1}$ be the mapping $\pi(\omega) = (\omega(3), \omega(4), \dots, \omega(2n))$ obtained by removing the first two terms of ω . The processes x^n and y^n on $\underline{\Omega}^n$ are defined inductively by

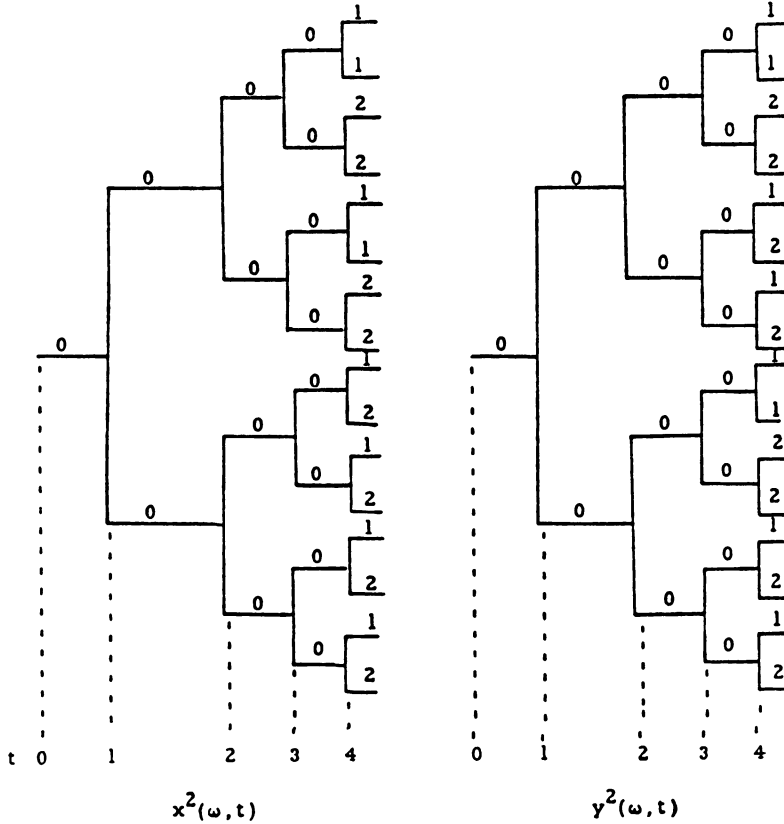
$$x^n(\omega, t) = \begin{cases} 0 & \text{if } t < 2n, \\ x^{n-1}(\pi\omega, t) & \text{if } t \geq 2n \text{ and } \omega(1) = 1, \\ y^{n-1}(\pi\omega, t) & \text{if } t \geq 2n \text{ and } \omega(1) = 2, \end{cases}$$

$$y^n(\omega, t) = \begin{cases} 0 & \text{if } t < 2n, \\ x^{n-1}(\pi\omega, t) & \text{if } t \geq 2n \text{ and } \omega(2) = 1, \\ y^{n-1}(\pi\omega, t) & \text{if } t \geq 2n \text{ and } \omega(2) = 2. \end{cases}$$

The tree representation of x^n is obtained by replacing each 1 in the tree for x^1 by a copy of the tree for x^{n-1} , and each 2 by a copy of the tree for y^{n-1} . The tree for y^n is obtained in analogous manner. The full tree has 2^{2^n} branches.



The figure below shows the trees for x^2 and y^2 .



Let h be the bijection of Ω^n obtained by interchanging $\omega(1)$ and $\omega(2)$, that is

$$h(\omega) = (\omega(2), \omega(1), \omega(3), \omega(4), \dots, \omega(2n)).$$

Then $\pi(h(\omega)) = \pi(\omega)$. In the tree, h interchanges the second and third subtrees beginning at time $t = 2$. We now prove some results about the processes x^n and y^n .

$$(3.2.1) \quad x^n(h\omega, t) = y^n(\omega, t).$$

PROOF. By induction on n . The case $n = 1$ is given in (3.1.1). Assume the result for $n - 1$. $x^n(\cdot, t)$ and $y^n(\cdot, t)$ are zero for $t < 2n$. Let $t \geq 2n$. Suppose first that $\omega(2) = 1$. Then $(h\omega)(1) = 1$, and

$$x^n(h\omega, t) = x^{n-1}(\pi(h\omega), t) = x^{n-1}(\pi\omega, t) = y^n(\omega, t).$$

The case $\omega(2) = 2$ is analogous. \square

We now wish to prove

(3.2.2) For each conditional process f of rank $< n$, $fx^n(h\omega) = fy^n(\omega)$. Hence

$$E[fx^n] = E[fy^n], \quad x^n \equiv_{n-1} y^n.$$

Given a conditional process $f(t_1, \dots, t_m)$, let f^- be the process

$$f^-(t_1, \dots, t_m) = f((t_1 - 2) \vee 0, \dots, (t_m - 2) \vee 0)$$

obtained by shifting the time two units to the left.

We shall prove (3.2.2) and the following condition (3.2.3) simultaneously by induction.

(3.2.3) *Let $n > 1$. For every conditional process f of rank less than n ,*

$$\begin{aligned} fx^n(\omega) &= \begin{cases} f^-x^{n-1}(\pi\omega) & \text{if } \omega(1) = 1, \\ f^-y^{n-1}(\pi\omega) & \text{if } \omega(1) = 2, \end{cases} \\ fy^n(\omega) &= \begin{cases} f^-x^{n-1}(\pi\omega) & \text{if } \omega(2) = 1, \\ f^-y^{n-1}(\pi\omega) & \text{if } \omega(2) = 2. \end{cases} \end{aligned}$$

PROOF OF (3.2.2) AND (3.2.3). We have already shown in Example 3.1 that (3.2.2) holds for $n = 1$. Now assume that (3.2.2) holds for $n - 1$. We shall prove (3.2.3) for n , arguing by induction on the formation of f . From the definition of x^n and y^n we see that (3.2.3) holds for every basic conditional process $\hat{\Phi}$. The set of f for which (3.2.3) holds is obviously closed under composition by bounded continuous functions. Suppose finally that f has the form $f = E[g|t]$ where (3.2.3) holds for g . Since f has rank less than n , g has rank less than $n - 1$. Then g^- has rank less than $n - 1$, and by (3.2.2) for $n - 1$,

$$E[g^-x^{n-1}] = E[g^-y^{n-1}].$$

Suppose first that $0 \leq t \leq 1$. Then by (3.2.3) for g ,

$$\begin{aligned} fx^n(\omega) &= E[gx^n|\mathcal{F}_0^n] = E[gx^n] = \frac{1}{2}E[g^-x^{n-1}] + \frac{1}{2}E[g^-y^{n-1}] \\ &= E[g^-x^{n-1}] = f^-x^{n-1}(\pi\omega) = E[g^-y^{n-1}] = f^-y^{n-1}(\pi\omega). \end{aligned}$$

Similarly,

$$fy^n(\omega) = f^-x^{n-1}(\omega) = f^-y^{n-1}(\omega).$$

Suppose next that $1 \leq t < 2$. Then

$$\begin{aligned} fx^n(\omega) &= E[gx^n|\mathcal{F}_1^n](\omega) \\ &= \begin{cases} E[g^-x^{n-1}] & \text{if } \omega(1) = 1, \\ E[g^-y^{n-1}] & \text{if } \omega(1) = 2, \end{cases} \\ &= E[g^-x^{n-1}] = f^-x^{n-1}(\pi\omega) \\ &= E[g^-y^{n-1}] = f^-y^{n-1}(\pi\omega). \end{aligned}$$

The computation of $fy^n(\omega)$ is the same as in the case $0 \leq t < 1$. Finally, suppose $2 \leq t$. Then $t - 2 \geq 0$, so

$$\begin{aligned} fx^n(\omega) &= E[gx^n|\mathcal{F}_t^n](\omega) \\ &= \begin{cases} E[g^-x^{n-1}|\mathcal{F}_{t-2}^{n-1}](\pi\omega) & \text{if } \omega(1) = 1, \\ E[g^-y^{n-1}|\mathcal{F}_{t-2}^{n-1}](\pi\omega) & \text{if } \omega(1) = 2, \end{cases} \\ &= \begin{cases} f^-x^{n-1}(\pi\omega) & \text{if } \omega(1) = 1, \\ f^-y^{n-1}(\pi\omega) & \text{if } \omega(1) = 2. \end{cases} \end{aligned}$$

The computation of $fy^n(\omega)$ is similar. Thus (3.2.3) holds for all conditional processes f of rank less than n .

Now assuming (3.2.3) for n , we prove (3.2.2) for n . Let f have rank less than n . Then $(h\omega)(1) = \omega(2)$ and $\pi(h\omega) = \pi(\omega)$, so

$$\begin{aligned} fx^n(h\omega) &= \begin{cases} f^{-}x^{n-1}(\pi\omega) & \text{if } \omega(2) = 1, \\ f^{-}y^{n-1}(\pi\omega) & \text{if } \omega(2) = 2, \end{cases} \\ &= fy^n(\omega). \quad \square \end{aligned}$$

To complete the example we show that

$$(3.2.4) \quad x^n \not\equiv_n y^n.$$

PROOF. The case $n = 1$ is shown in Example 3.1. Assume the result for $n - 1$. That is, there is a conditional process g of rank $n - 1$ such that $E[gx^{n-1}] \neq E[gy^{n-1}]$. Form the conditional process g^+ by shifting times two to the right, so that $(g^+)^- = g$. By (3.2.3),

$$g^+x^n(\omega) = \begin{cases} gx^{n-1}(\pi\omega) & \text{if } \omega(1) = 1, \\ gy^{n-1}(\pi\omega) & \text{if } \omega(1) = 2. \end{cases}$$

Let f be the conditional process $f = E[g^+|\cdot]$ of rank n . Then

$$fx^n(\omega) = E[g^+x^n|\mathcal{F}_1^n](\omega) = \begin{cases} E[gx^{n-1}] & \text{if } \omega(1) = 1, \\ E[gy^{n-1}] & \text{if } \omega(1) = 2. \end{cases}$$

By a similar computation,

$$fy^n(\omega) = \frac{1}{2}E[gx^{n-1}] + \frac{1}{2}E[gy^{n-1}].$$

It follows that $x^n \not\equiv_n y^n$.

By conditioning backwards we can show that the relation \equiv_n is stronger than \equiv_{n-1} even for martingales.

EXAMPLE 3.3. Let $n > 1$ and

$$\bar{x}^n(\omega, t) = E[x^n(\cdot, 2n)|\mathcal{F}_t^n], \quad \bar{y}^n(\omega, t) = E[y^n(\cdot, 2n)|\mathcal{F}_t^n]$$

Then \bar{x}^n and \bar{y}^n are martingales on the same space $\underline{\Omega}^n$ such that $\bar{x}^n \equiv_{n-2} \bar{y}^n$ but not $\bar{x}^n \equiv_{n-1} \bar{y}^n$.

The proof is straightforward.

Recall that Proposition 2.17 states that for each stochastic process x , the process m_x of Definition 2.16 is \mathcal{F}_t -Markov. We conclude this section with an example showing that the processes $m_r x$ are not necessarily \mathcal{F}_t -Markov.

EXAMPLE 3.4. A process x on an adapted space $\underline{\Omega}$ such that for each $r \in \mathbb{N}$, $m_r x$ is not \mathcal{F}_t -Markov.

Begin with the adapted spaces $\underline{\Omega}^n$ of the preceding examples. Let $\Omega = \bigcup_{n \in \mathbb{N}} \Omega^n$. Note that the sets Ω^n are disjoint.

Let P be the probability measure on Ω such that for each n , each element of Ω^n has measure 2^{-3n} . That is, $P(\Omega^n) = 2^{-n}$ and all elements of Ω^n have equal weight.

Let \mathcal{F}_t be the σ -algebra on Ω generated by $\bigcup_{n \in \mathbb{N}} \mathcal{F}_t^n$ and let x be the process on $\underline{\Omega}$ defined by

$$x(\omega, t) = x^n(\omega, t) \quad \text{if } \omega \in \Omega^n.$$

PROOF THAT $m_r x$ IS NOT \mathcal{F}_t -MARKOV. By (3.2.2) and (3.2.3), for every conditional process g of rank $< r$, and all $\omega \in \Omega^{r+1}$,

$$E[gx^{r+1}|\mathcal{F}_1^{r+1}](\omega) = E[g^-x^r] = E[g^-y^r]$$

and hence on $\underline{\Omega}^{r+1}$,

$$E[gx|\mathcal{F}_1](\omega) = E[g^-x^r].$$

Each coordinate of $m_r x(t)$ has the form $E[gx|\mathcal{F}_t]$ for some $g \in \text{CP}^0$ of rank $< r$. Therefore $m_r x(\omega, 1)$ has the same value for all $\omega \in \Omega^{r+1}$. By (3.2.4) there is a conditional process f of rank r such that

$$\alpha = E[fx^r] \neq E[fy^r] = \beta.$$

Moreover, by (3.2.3), on $\bar{\Omega}^{r+1}$ we have

$$E[f^+x^r|\mathcal{F}_1^r](\omega) = \begin{cases} \alpha & \text{if } \omega(1) = 1, \\ \beta & \text{if } \omega(1) = 2. \end{cases}$$

Working on the space $\underline{\Omega}$ we have for each $\omega \in \Omega^{r+1}$,

$$E[f^+x|\mathcal{F}_1](\omega) = \begin{cases} \alpha & \text{if } \omega(1) = 1, \\ \beta & \text{if } \omega(1) = 2. \end{cases}$$

By Corollary 2.13, f may be taken to be of the form

$$f = \varphi(E[g_1|t_1], \dots, E[g_n|t_n]),$$

where φ is a bounded continuous function and g_1, \dots, g_n belong to CP_{r-1}^0 . Thus the $E[g_i x|\mathcal{F}_{t_i}]$ are coordinates of $m_r x(t_i)$, and therefore there is a bounded continuous function $\psi: (\mathbb{R}^n)^n \rightarrow \mathbb{R}$ such that

$$f^+x = \psi(m_r x(t_1 + 2), \dots, m_r x(t_n + 2)).$$

Then for $\omega \in \Omega^{r+1}$,

$$E[\psi(m_r x(t_1 + 2), \dots, m_r x(t_n + 2))|\mathcal{F}_1](\omega) = \begin{cases} \alpha & \text{if } \omega(1) = 1, \\ \beta & \text{if } \omega(1) = 2. \end{cases}$$

This shows that $m_r x$ is not \mathcal{F}_t -Markov. \square

4. Complete and atomless spaces. It is easy to use the method of Maharam [1950] to show that there exists a probability space Ω which has the following saturation property (cf. Corollary 4.5): If x_1 is a random variable on Ω , and y_1, y_2 are random variables on a probability space Λ such that x_1 and y_1 have the same distributions, in our notation $x_1 \equiv_0 y_1$, then there is an x_2 on Ω such that

$$(x_1, x_2) \equiv_0 (y_1, y_2).$$

Our aim is to prove that there are adapted spaces with an analogous saturation property for stochastic processes and adapted distribution (Theorem 5.2). In this

section we introduce notions of rich adapted spaces necessary for these saturation results.

DEFINITION 4.1. An adapted space $\underline{\Omega}$ is *complete* if for each countable $T \subseteq \mathbf{R}^+$ and all stochastic processes $x, y_n, n \in \mathbf{N}$ on $\underline{\Omega}$ such that $y_n(t)$ is tight for all $t \in T$ and

$$(x, y_n) \xrightarrow{\text{ad}} \quad \text{on } T,$$

there is a stochastic process y on $\underline{\Omega}$ such that

$$(x, y_n) \xrightarrow{\text{ad}} (x, y) \quad \text{on } T.$$

Every finite adapted space is complete for trivial reasons.

We now show that nontrivial complete adapted spaces exist. The only method we know of obtaining such spaces uses the Loeb measure construction from nonstandard analysis (see Anderson [1976], Keisler [1982], or Hoover and Perkins [1982] for the necessary background).

THEOREM 4.2. *Every adapted Loeb space is complete.*

PROOF. Let

$$\bar{\underline{\Omega}} = (\bar{\Omega}, \bar{\mathcal{G}}_t, \bar{P})_{t \in {}^*\mathbf{R}^+}$$

be an internal adapted space and let $\underline{\Omega} = (\Omega, \mathcal{F}_t, P)$ be the Loeb space of $\bar{\underline{\Omega}}$. That is $P = L(\bar{P})$ is the Loeb measure of \bar{P} and $\mathcal{N}(P)$ is the set of null sets of P and

$$\mathcal{F}_t = \bigcap_{o_t > t} \sigma(\mathcal{G}_t) \vee \mathcal{N}(P), \quad t \in \mathbf{R}^+.$$

Let $T \subseteq \mathbf{R}^+$ be countable and let x, y_n be processes on $\underline{\Omega}$ such that $y_n(t)$ is tight for each $t \in T$ and

$$(x, y_n) \xrightarrow{\text{ad}} \quad \text{on } T.$$

It follows from results in Hoover and Perkins [1982] that

$$(4.2.1) \quad \mathcal{F}_t = \bigcup_{o_t \approx t} \sigma(\mathcal{G}_t) \vee \mathcal{N}(P).$$

(4.2.2) For each bounded random variable $z: \Omega \rightarrow \mathbf{R}$, bounded lifting $Z: \Omega \rightarrow {}^*\mathbf{R}$, and $u \in {}^*\mathbf{R}^+$,

$${}^o\bar{E} [Z | \mathcal{G}_u] = E [z | \sigma(\mathcal{G}_u)] \quad \text{a.s.}$$

(4.2.3) For each bounded random variable $z: \Omega \rightarrow \mathbf{R}$ and $t \in \mathbf{R}^+$, there is a $u \in {}^*\mathbf{R}^+$ such that $u \approx t$ and $E[z | \mathcal{F}_t]$ is $\sigma(\mathcal{G}_u) \vee \mathcal{N}(P)$ -measurable. Thus

$$(4.2.4) \quad E [z | \mathcal{F}_t] = E [z | \sigma(\mathcal{G}_u)] \quad \text{a.s.}$$

for all sufficiently large $u \approx t$.

Since any countable subset of the monad of t has an upper bound, it follows that for each $t \in \mathbf{R}^+$ there is a $t_1 \in {}^*\mathbf{R}^+$ such that $t_1 \approx t$ and for each $f \in \text{CP}^0$, \bar{s} in T , and $n \in \mathbf{N}$,

$$(4.2.5) \quad E [f(x, y_n)(\bar{s}) | \mathcal{F}_t] = E [f(x, y_n)(\bar{s}) | \sigma(\mathcal{G}_{t_1})] \quad \text{a.s.}$$

For each $n \in \mathbf{N}$ and $t \in T$, choose liftings $X(t)$ of $x(t)$ and $Y_n(t)$ of $y_n(t)$. For each $f \in \mathbf{CP}^0$, \vec{s} in T , and \vec{s}_1 in ${}^*\mathbf{R}^+$ with $\vec{s}_1 \approx \vec{s}$, let $F(X, Y_n)(\vec{s}_1): \Omega \rightarrow {}^*\mathbf{R}$ be obtained from $f(x, y_n)(\vec{s})$ by replacing $x(t)$ by $X(t)$, $y_n(t)$ by $Y_n(t)$, Φ by ${}^*\Phi$, φ by ${}^*\varphi$, and \mathcal{F}_t by \mathcal{G}_{t_1} . By (4.2.2) and induction on f , for each $s \in T$ we have

$$(4.2.6) \quad {}^oF(X, Y_n)(\vec{s}_1) = f(x, y_n)(\vec{s}) \quad \text{a.s.}$$

for each $f \in \mathbf{CP}^0$ and $n \in \mathbf{N}$. Extend the double sequence $Y_n(t)$, $n \in \mathbf{N}$, $t \in T$ to an internal function $Y_n(t)$, $n \in {}^*\mathbf{N}$, $t \in {}^*T$.

Let M be the Polish space of values for the processes $y_n(t)$. We shall now use the tightness assumption to show that for all sufficiently small infinite $H \in {}^*\mathbf{N}$ and each $t \in T$, $Y_H(t)$ is almost surely near-standard. Let ρ be a metric on M . Let $t \in T$. For each $m \in \mathbf{N}$, let $K_m \subseteq M$ be a compact set such that

$$P[y_n(t) \in K_m] \geq 1 - 1/m.$$

Then for each $n \in \mathbf{N}$,

$$(4.2.7) \quad P[{}^*\rho(Y_n(t), {}^*K_m) \leq 1/n] \geq 1 - 1/m$$

and thus (4.2.7) holds for all sufficiently small infinite $n \in {}^*\mathbf{N}$. If ${}^*\rho(Y_n(t), {}^*K_m) \approx 0$ then $Y_n(t)$ is near-standard. By saturation, for all sufficiently small infinite $H \in {}^*\mathbf{N}$, $Y_H(t)$ is a.s. near-standard. Since T is countable, for all sufficiently small infinite H , $Y_H(t)$ is a.s. near-standard for all $t \in T$. We may thus define

$$y_H(t) = {}^oY_H(t), \quad t \in T.$$

Since $E[f(x, y_n)(\vec{s})]$ converges, for each $f \in \mathbf{CP}^0$, \vec{s} in T , all sufficiently small infinite $H \in {}^*\mathbf{N}$, and all sufficiently large $\vec{s}_1 \approx \vec{s}$, we have

$$(4.2.8) \quad {}^oE[F(X, Y_H)(\vec{s}_1)] = \lim_{n \rightarrow \infty} E[f(x, y_n)(\vec{s})].$$

It follows that there is an infinite $H \in {}^*\mathbf{N}$ such that $Y_H(t)$ is a.s. near-standard for each $t \in T$ and there is a positive infinitesimal ε such that for all $f \in \mathbf{CP}^0$ and $\vec{s} \in T$, (4.2.8) holds whenever $\vec{s}_1 \geq \vec{s} + \varepsilon$ and $\vec{s}_1 \approx \vec{s}$. ε may be chosen large enough so that whenever $f \in \mathbf{CP}^0$, \vec{s} , $t \in T$, and $t + \varepsilon \leq t_1 \approx t$, (4.2.5) holds for all $n \in \mathbf{N}$ and also for $n = H$. Then (4.2.6) holds whenever $f \in \mathbf{CP}^0$, \vec{s} in T , $\vec{s} + \varepsilon \leq \vec{s}_1 \approx \vec{s}$, and $n \in \mathbf{N}$ or $n = H$. Therefore

$$(4.2.9) \quad E[f(x, y_H)(\vec{s})] = \lim_{n \rightarrow \infty} E[f(x, y_n)(\vec{s})]$$

for each $f \in \mathbf{CP}^0$ and \vec{s} in T . By Theorem 2.26,

$$(x, y_n) \xrightarrow{\text{ad}} (x, y_H) \quad \text{on } T. \quad \square$$

DEFINITION 4.3. Let (Ω, \mathcal{F}, P) be a probability space.

(4.3.1) If \mathcal{A} , \mathcal{D} are σ -algebras on Ω and $\mathcal{A} \subseteq \mathcal{D} \subseteq \mathcal{F}$, we say that \mathcal{D} is *atomless over* \mathcal{A} if for every $D \in \mathcal{D}$ such that $P(D) > 0$ there is $D_0 \in \mathcal{D}$, $D_0 \subseteq D$, such that on some set of positive probability,

$$0 < P[D_0 | \mathcal{A}] < P[D | \mathcal{A}].$$

A σ -algebra is *atomless* if it is atomless over the trivial σ -algebra.

(4.3.2) We say \mathcal{D} is \aleph_1 -atomless over \mathcal{A} if \mathcal{D} is atomless over every $\mathcal{A}' \subseteq \mathcal{D}$ which is countably generated over \mathcal{A} , i.e. of the form $\mathcal{A} \vee \mathcal{C}$ with \mathcal{C} countably generated.

(4.3.3) A probability space $\underline{\Omega} = (\Omega, \mathcal{F}, P)$ is *atomless*, or \aleph_1 -atomless if \mathcal{F} is. For example, given any probability space $\underline{\Omega}$ with at least one set of measure different from 0 and 1, the product space $\underline{\Omega}^{\mathbb{N}}$ is atomless, and the product space $\underline{\Omega}^{\mathbb{R}}$ is \aleph_1 -atomless.

(4.3.4) A filtration \mathcal{F}_t , $t \in \mathbb{R}^+$, on (Ω, \mathcal{F}, P) is said to be *atomless* if \mathcal{F}_0 is atomless, \mathcal{F}_∞ is atomless over each \mathcal{F}_t , and \mathcal{F}_t is atomless over \mathcal{F}_s whenever $s < t$. The notion of an \aleph_1 -atomless filtration is defined analogously. An adapted space $\underline{\Omega}$ is said to be *atomless* or \aleph_1 -atomless if its filtration is atomless or \aleph_1 -atomless.

The necessary techniques for dealing with atomlessness were perfected in Maharam [1942, 1950]. The essential facts are assembled in the following lemma.

LEMMA 4.4. (i) If x is an M -valued random variable, then $\sigma(x)$ is atomless iff the distribution of x has no point masses.

(ii) If \mathcal{D} is atomless over \mathcal{A} then \mathcal{D} is atomless over every σ -algebra $\mathcal{C} \subseteq \mathcal{D}$ which is finitely generated over \mathcal{A} .

(iii) Suppose \mathcal{D} is atomless over \mathcal{A} . Let G be an \mathcal{A} -measurable random distribution on M . That is, for each Borel set $U \subseteq M$, $G(U): \Omega \rightarrow [0, 1]$ is \mathcal{A} -measurable, $G(\emptyset) = 0$, $G(M) = 1$, and if $U = \bigcup_n U_n$ is a disjoint union then $G(U) = \sum_n G(U_n)$ a.s. Then there is a \mathcal{D} -measurable random variable x such that

$$P[x \in U | \mathcal{A}] = G(U) \quad \text{a.s.}$$

for each Borel set $U \subseteq M$.

(iv) If \mathcal{D} is a σ -algebra, \mathcal{A} and \mathcal{C} are sub- σ -algebras of \mathcal{D} , \mathcal{C} is atomless, and \mathcal{A} and \mathcal{C} are independent, then \mathcal{D} is atomless over \mathcal{A} .

PROOF. (i) If $p[x = m] > 0$ for $m \in M$, then $\{x = m\}$ is an atom of $\sigma(x)$. Conversely, if A is an atom of $\sigma(x)$, and U_n , $n \in \mathbb{N}$, is a countable basis for the topology on M , then for each $n \in \mathbb{N}$, either $x(A) \in U_n$ a.s. or $x(A) \notin U_n$ a.s. Since M is Hausdorff,

$$\bigcap_{x(A) \in U_n \text{ a.s.}} U_n$$

must be a singleton $\{m\}$, and $x(A) = m$ a.s., so m is a point mass of the distribution of x .

(ii) Let \mathcal{D} be atomless over \mathcal{A} , and suppose that $\mathcal{C} \subseteq \mathcal{D}$ is finitely generated over \mathcal{A} . We may assume that \mathcal{C} is obtained by adjoining sets $D_1, \dots, D_n \in \mathcal{D}$ to \mathcal{A} , where $\{D_i\}$ forms a partition of Ω . Fix $C \in \mathcal{D}$, and $i \leq n$ such that $P[C \cap D_i] > 0$. We may assume wlog that $C \subseteq D_i$. Since \mathcal{D} is atomless over \mathcal{A} , there is a set $C_0 \subseteq C$ in \mathcal{D} such that

$$(4.4.1) \quad 0 < P[C_0 | \mathcal{A}] < P[C | \mathcal{A}]$$

on a set $A \in \mathcal{A}$ of positive measure. Then

$$P[A \cap C_0] = \int_A P[C_0 | \mathcal{A}] > 0.$$

Now observe that any \mathcal{C} -measurable function must have the form $\sum_{i=1}^n g_i \cdot 1(D_i)$, where the g_i are \mathcal{A} -measurable. Hence if $F \subseteq D_i$, there is an \mathcal{A} -measurable function g such that $P[F|\mathcal{C}] = g \cdot 1(D_i)$. Then

$$P[F|\mathcal{A}] = E[P[F|\mathcal{C}]|\mathcal{A}] = g \cdot P[D_i|\mathcal{A}].$$

Since $P[D_i|\mathcal{A}] > 0$ a.s. on D_i , it follows that for any $F \subseteq D_i$,

$$(4.4.2) \quad P[F|\mathcal{C}] = 1(D_i) \cdot P[F|\mathcal{A}]/P[D_i|\mathcal{A}] \quad \text{a.s.}$$

Then almost surely on D_i ,

$$P[C_0|\mathcal{C}] = P[C_0|\mathcal{A}]/P[D_i|\mathcal{A}], \quad P[C|\mathcal{C}] = P[C|\mathcal{A}]/P[D_i|\mathcal{A}].$$

By (4.4.1), it follows that

$$0 < P[C_0|\mathcal{C}] < P[C|\mathcal{C}]$$

almost surely in $A \cap D_i$. By (4.4.1) and $C \subseteq D_i$, $P[A \cap D_i] = E[P[D_i|\mathcal{A}] \cdot 1(A)] > 0$. This shows that \mathcal{D} is atomless over \mathcal{C} .

(iii) By Maharam [1950, Lemma 3, p. 146], if \mathcal{D} is atomless over \mathcal{A} , then for any \mathcal{A} -measurable $f: \Omega \rightarrow [0, 1]$, there is $D \in \mathcal{D}$ such that $P[D|\mathcal{A}] = f$ a.s. Let $U_n, n \in \mathbb{N}$, be a countable basis for the topology on M . Using (ii) and Maharam's Lemma, inductively define D_1, D_2, D_3, \dots in \mathcal{D} so that

$$P[D_1|\mathcal{A}] = G(U_1) \quad \text{a.s.},$$

$$P[D_2|\mathcal{A}(D_1)] = \begin{cases} G(U_1 \cap U_2)/G(U_1) & \text{on } D_1, \\ G(U_1^c \cap U_2)/G(U_1^c) & \text{on } D_1^c, \end{cases}$$

and in general

$$P[D_{n+1}|\mathcal{A}(D_1, \dots, D_n)] = \frac{G(p(U_1, \dots, U_n) \cap U_{n+1})}{G(p(U_1, \dots, U_n))} \quad \text{on } p(D_1, \dots, D_n)$$

for each Boolean expression $p(X_1, \dots, X_n)$ of the form $Y_1 \cap \dots \cap Y_n$, where each Y_i is either X_i or X_i^c . One may check by induction that for each n and p ,

$$P[p(D_1, \dots, D_n)|\mathcal{A}] = G(p(U_1, \dots, U_n)) \quad \text{a.s.}$$

Let x be an M -valued random variable such that $\{x \in U_n\} = D_n$ a.s. for each n . For each finite m it is easy to find x_m that satisfies this for each of D_1, \dots, D_m ; these converge a.s. to the desired x . Then for each Borel set $U \subseteq M$,

$$P[x \in U|\mathcal{A}] = G(U) \quad \text{a.s.}$$

(iv) Choose any $D \in \mathcal{D}$ such that $P(D) > 0$. By (iii), there are $C_1, \dots, C_n \in \mathcal{C}$ forming a partition of Ω such that

$$0 < P(C_i) \leq \frac{1}{2}P(D) \quad \text{for each } i.$$

Then for each i ,

$$P[C_i \cap D|\mathcal{A}] \leq P[C_i|\mathcal{A}] = P(C_i) \leq \frac{1}{2}P(D) \quad \text{a.s.},$$

the equality holding since \mathcal{C} and \mathcal{A} are independent. Now $\{P[D|\mathcal{A}] > \frac{1}{2}P(D)\}$ has positive measure, and since $\bigcup C_i = \Omega$, for some i , $P[C_i \cap D|\mathcal{A}]$ must be positive on a nonnull subset of this set. Therefore

$$0 < P[C_i \cap D|\mathcal{A}] < P[D|\mathcal{A}]$$

on a set of positive measure. Thus \mathcal{D} is atomless over \mathcal{A} . \square

COROLLARY 4.5. (i) *A necessary and sufficient condition for a probability space $\underline{\Omega} = (\Omega, \mathcal{F}, P)$ to be \aleph_1 -atomless is that for all random variables x_1 on $\underline{\Omega}$ and y_1, y_2 on another space $\underline{\Lambda} = (\Lambda, \mathcal{G}, Q)$ with $x_1 \equiv_0 y_1$ (i.e. x_1 and y_1 have the same distribution), there is a random variable x_2 on $\underline{\Omega}$ with*

$$(x_1, x_2) \equiv_0 (y_1, y_2).$$

(ii) *Let $\underline{\Omega}$ be an adapted space such that \mathcal{F} is \aleph_1 -atomless. Then for all processes right continuous in probability x_1 on $\underline{\Omega}$ and y_1, y_2 on $\underline{\Lambda}$ with $x_1 \equiv_0 y_1$ there is a process right continuous in probability x_2 on $\underline{\Omega}$ with*

$$(x_1, x_2) \equiv_0 (y_1, y_2).$$

PROOF. (i) The sufficiency follows from 4.4(iv). We prove the necessity. Let y_1, y_2 have values in the Polish spaces M_1, M_2 . For each Borel set $U \subseteq M_2$ there is a Borel function

$$g_U: M_1 \rightarrow [0, 1]$$

such that in Δ ,

$$Q[y_2 \in U|y_1] = g_U(y_1) \quad \text{a.s.}$$

Since \mathcal{F} is \aleph_1 -atomless and $\sigma(x_1)$ is countably generated, \mathcal{F} is atomless over $\sigma(x_1)$. By Proposition 4.4(iii) there is an x_2 on $\underline{\Omega}$ such that for each U ,

$$P[x_2 \in U|x_1] = g_U(x_1) \quad \text{a.s.}$$

Then

$$\begin{aligned} P[(x_1, x_2) \in U_1 \times U_2] &= \int P[x_2 \in U_2|x_1] \cdot 1\{x_1 \in U_1\} dP \\ &= \int g_{U_2}(x_1) \cdot 1\{x_1 \in U_1\} dP \\ &= \int g_{U_2}(y_1) \cdot 1\{y_1 \in U_1\} dQ \\ &= Q[(y_1, y_2) \in U_1 \times U_2], \end{aligned}$$

so $(x_1, x_2) \equiv_0 (y_1, y_2)$.

To prove (ii), replace each processes $z(t)$ by the random variable $\langle z(t): t \in \mathbf{Q}^+ \rangle$.

\square

COROLLARY 4.6. *Atomless complete spaces exist.*

PROOF. It follows by 4.4(i) and (iv) that any adapted space $\underline{\Omega}$ which carries an \mathcal{F}_t -Brownian motion is atomless. By Proposition 4.2 any adapted Loeb space which

carries an \mathcal{F}_t -Brownian motion is atomless and complete. Such spaces exist by Anderson [1976]. \square

Completeness has the effect of strengthening atomlessness.

PROPOSITION 4.7. *If an adapted $\underline{\Omega}$ is complete and atomless, then it is \aleph_1 -atomless.*

PROOF. We shall show that whenever $s < t$ in \mathbf{R}^+ , \mathcal{F}_t is \aleph_1 -atomless over \mathcal{F}_s . Suppose $\mathcal{A} \subset \mathcal{F}_t$ is countably generated over \mathcal{F}_s . Let x be an \mathcal{F}_t -measurable random variable such that $\mathcal{A} = \mathcal{F}_s(x)$, and let x_n , $n \in \mathbf{N}$, be real-valued random variables such that each $\sigma(x_n)$ is finite and contained in $\sigma(x)$, and

$$(4.7.1) \quad x_n \rightarrow x \quad \text{a.s.}$$

By 4.4(ii), \mathcal{F}_t is atomless over $\mathcal{F}_{t-1/n}(x_n)$ for each n . Then by 4.4(iii) we may choose \mathcal{F}_t -measurable random variables u_n which are uniformly distributed over $[0, 1]$ and independent of $\mathcal{F}_{t-1/n}(x_n)$. Let \hat{x} , \hat{x}_n , \hat{u}_n be the constant stochastic processes

$$\hat{x}(r) = x, \quad \hat{x}_n(r) = x_n, \quad \hat{u}_n(r) = u_n,$$

and let $T = \mathbf{Q}^+ - \{t\}$. One can show by induction on rank that

$$(4.7.2) \quad (\hat{x}_n, \hat{u}_n) \xrightarrow{\text{ad}} \quad \text{on } T.$$

By (4.7.1) and Lemma 2.19,

$$(4.7.3) \quad (\hat{x}, \hat{u}_n) \xrightarrow{\text{ad}} \quad \text{on } T.$$

The sequence $(\hat{x}, \hat{u}_n)(r)$ is tight for each r .

By completeness of $\underline{\Omega}$ there is a process v on $\underline{\Omega}$ such that

$$(4.7.4) \quad (\hat{x}, \hat{u}_n) \xrightarrow{\text{ad}} (\hat{x}, v) \quad \text{on } T.$$

By (4.7.1) and 2.19,

$$(4.7.5) \quad (\hat{x}_n, \hat{u}_n) \xrightarrow{\text{ad}} (\hat{x}, v) \quad \text{on } T.$$

Since each \hat{u}_n is constant on T , v is a.s. constant on T , so we may take v to be of the form \hat{u} for some random variable u on $\underline{\Omega}$. Since each u_n is \mathcal{F}_t -measurable, u must be \mathcal{F}_t -measurable. Similarly, since u_n is independent of $\mathcal{F}_s(x_n)$ for $1/n \leq t -$, u is independent of $\mathcal{A} = \mathcal{F}_s(x)$. Since each u_n and hence u is uniformly distributed over $[0, 1]$, $\sigma(u)$ is atomless. Then by 4.4(iv), \mathcal{F}_t is atomless over \mathcal{A} , and hence \aleph_1 -atomless over \mathcal{F}_s .

The proofs that \mathcal{F}_0 is \aleph_1 atomless, and that \mathcal{F}_∞ is \aleph_1 -atomless over each \mathcal{F}_s , are similar. \square

5. Saturated adapted spaces. In this section we prove the main theorem of the paper, which says that complete atomless spaces possess a saturation property with respect to adapted distribution. In particular, any system of processes can be duplicated, up to adapted distribution, on any complete atomless space. The saturation property is the most powerful tool in dealing with adapted distributions.

DEFINITION 5.1. An adapted space $\underline{\Omega} = (\Omega, P, \mathcal{F}_t)_{t \in \mathbf{R}^+}$ is *universal* if for every Polish space M and every stochastic process

$$(y) = (\Lambda, Q, \mathcal{G}_t, y)_{t \in \mathbf{R}^+}$$

with values in M there is a stochastic process x on $\underline{\Omega}$ such that $x \equiv y$. $\underline{\Omega}$ is *saturated* if for every stochastic process x_1 , on $\underline{\Omega}$, adapted space $\underline{\Lambda}$, and stochastic process (y_1, y_2) on $\underline{\Lambda}$ such that $x_1 \equiv y_1$, there is a stochastic process x_2 on $\underline{\Omega}$ such that $(x_1, x_2) \equiv (y_1, y_2)$. Thus every saturated adapted space is universal.

THEOREM 5.2. *An adapted space is saturated if and only if it is complete and atomless. Hence there exists a saturated adapted space.*

We shall omit the easy proof that every saturated space is complete and atomless. We prepare the proof of the other direction with a series of lemmas.

DEFINITION 5.3. Let x be r.c.l.l. process on $\underline{\Omega}$ and let T be a finite or countable subset of \mathbf{R}^+ . The stochastic process

$$m^T x: (\mathbf{R}^+ \cup \{\infty\}) \times \Omega \rightarrow \mathbf{R}^N$$

is defined as follows:

$$\begin{aligned} m^T x(\infty) &= \langle fx(\vec{s}): f \in \mathbf{CP}^0, \vec{s} \text{ in } T \rangle, \\ m^T x(t) &= E[m^T x(\infty) | \mathcal{F}_t], \quad t \in \mathbf{R}^+. \end{aligned}$$

LEMMA 5.4. $m^T x$ is a martingale (in each coordinate) and a Markov process with respect to $(\Omega, P, \mathcal{F}_t)_{t \in \mathbf{R}^+ \cup \{\infty\}}$.

PROOF. Argue as in Proposition 2.17. \square

LEMMA 5.5. The following are equivalent for finite or countable $T \subseteq \mathbf{R}^+$.

- (i) $x \equiv y$ on T .
- (ii) $m^T x(\infty)$ and $m^T y(\infty)$ have the same distributions.
- (iii) $m^T x \equiv_0 m^T y$ on $T \cup \{\infty\}$.
- (iv) $m^T x \equiv_0 m^T y$ on T .

PROOF. The implications (i) \rightarrow (ii) \rightarrow (iii) \rightarrow (iv) are obvious. The proof that (iv) \rightarrow (i) is like the proof of Proposition 2.18. \square

LEMMA 5.6. Let $T \subseteq \mathbf{R}^+$ be countable, let $z: (T \cup \{\infty\}) \times \Omega \rightarrow \mathbf{R}^N$ be a martingale, and let y be an M -valued stochastic process.

- (i) If $(x, z) \equiv_0 (y, m^T y)$ on $T \cup \{\infty\}$, then $z = m^T x$ a.s. for all $t \in T \cup \{\infty\}$.
- (ii) Suppose $z \equiv_0 m^T y$ on $T \cup \{\infty\}$. Then there is a stochastic process $x: T \times \Omega \rightarrow M$ such that $z = m^T x$ a.s. for all $t \in T \cup \{\infty\}$.

PROOF. (i) One shows by induction on formation that for $f \in \mathbf{CP}^0$ and $\vec{s}, t \in T \cup \{\infty\}$,

$$z_{f(\vec{s})}(\infty) = fx(\vec{s}) \text{ a.s.}, \quad z_{f(\vec{s})}(t) = E[fx(\vec{s}) | \mathcal{F}_t] \text{ a.s.}$$

- (ii) For each $\varphi \in S_n$, and $s_1, \dots, s_n \in T$, $\varphi(y(s_1), \dots, y(s_n))$ is a coordinate of $m^T y(\infty)$.

Define $\Phi: M \rightarrow \mathbf{R}^N$ by $\Phi(y) = (\varphi(y))_{\varphi \in S_1}$. Then Φ is a Borel isomorphism of M onto a Borel subset of \mathbf{R}^N by the Souslin-Lusin Theorem. (Dellacherie and Meyer [1978, p. 49–III]). Applying Φ^{-1} to the appropriate coordinates of $m^T y(\infty)$ we obtain a Borel function $\psi: \mathbf{R}^N \rightarrow M$ such that

$$(y(t): t \in T) = \psi(m^T y(\infty)),$$

Let $x: T \times \Omega \rightarrow M$ be defined by $(x(t): t \in T) = \psi(z(\infty))$. Then

$$(x, z) \equiv_0 (y, m^T y) \quad \text{on } T \cup \{\infty\},$$

and it follows by (i) that $z = m^T(x)$ a.s. \square

We now show that \aleph_1 -atomless spaces satisfy the saturation property for finite sets T .

LEMMA 5.7. *Let $\underline{\Omega}$ be \aleph_1 -atomless and let $T \subseteq \mathbf{R}^+$ be finite. For all stochastic processes x_1 on $\underline{\Omega}$ and y_1, y_2 on another space $\underline{\Lambda}$ such that $x_1 \equiv y_1$ on T , there is a stochastic process x_2 on $\underline{\Omega}$ with $(x_1, x_2) \equiv (y_1, y_2)$ on T .*

PROOF. Let $T = \{t_1, \dots, t_n\}$ with $t_1 < \dots < t_n$. It suffices to prove the result when x_1, y_1 , and y_2 are random variables, because a stochastic process u may be replaced by the random variable $\langle u(t_1), \dots, u(t_n) \rangle$. By Lemma 5.5, $m^T x_1 \equiv_0 m^T y_1$ on $T \cup \{\infty\}$.

Put $t_{n+1} = \infty$. We prove by induction that for each $k \leq n+1$ there is a martingale $z(t_1), \dots, z(t_k)$ on $\underline{\Omega}$ such that

$$(5.7.1) \quad (m^T x_1, z) \equiv_0 (m^T y_1, m^T(y_1, y_2)) \quad \text{on } \{t_1, \dots, t_k\}.$$

Let us first show that if (5.7.1) holds for $k = n+1$ then the lemma is true. Assume (5.7.1) for $k = n+1$. By Lemma 5.6 there is a random variable u on $\underline{\Omega}$ with $z(\infty) = m^T u(\infty)$ a.s. Using (5.7.1) again we find that there is an x_2 on $\underline{\Omega}$ with $u = (x_1, x_2)$ a.s. By Lemma 5.5 it follows that

$$(x_1, x_2) \equiv (y_1, y_2) \quad \text{on } T.$$

We now prove (5.7.1) when $k = 1$. Since \mathcal{F}_{t_1} is \aleph_1 -atomless and $m^T x_1(t_1)$ is \mathcal{F}_{t_1} -measurable, Corollary 4.5 shows that there is an \mathcal{F}_{t_1} -measurable $z(t_1)$ on $\underline{\Omega}$ such that

$$(m^T x_1, z) \equiv_0 (m^T y_1, m^T(y_1, y_2)) \quad \text{on } \{t_1\}.$$

Now let $1 \leq k < n+1$ and assume there is a martingale $z(t_1), \dots, z(t_k)$ which satisfies (5.7.1) for k . To simplify notation let $s = t_k, t = t_{k+1}, \hat{x}_1 = m^T x_1, \hat{y}_1 = m^T y_1, \hat{y}_2 = m^T(y_1, y_2)$. Let \hat{y}_1, \hat{y}_2 take values in M_1, M_2 .

We show first that

$$(5.7.2) \quad (\hat{x}_1(t_1), \dots, \hat{x}_1(t_{k+1}), z(t_1), \dots, z(t_k)) \equiv_0 (\hat{y}_1(t_1), \dots, \hat{y}_1(t_{k+1}), \hat{y}_2(t_1), \dots, \hat{y}_2(t_k)).$$

Since \hat{x}_1 and \hat{y}_2 are Markov processes, for each $\varphi \in C(M_1, \mathbf{R})$ there is a Borel function $f_\varphi: M_1 \rightarrow \mathbf{R}$ such that

$$E[\varphi(\hat{y}_1(t)) | \mathcal{G}_s] = f_\varphi(\hat{y}_1(s)) = E[\varphi(\hat{y}_1(t)) | (\hat{y}(t_l), \hat{y}_2(t_l))_{l \leq k}] \quad \text{a.s.,}$$

and

$$E[\varphi(\hat{x}_1(t))|\mathcal{F}_s] = f_\varphi(\hat{x}_1(s)) = E[\varphi(\hat{x}_1(t))|(\hat{x}_1(t_l), z(t_l))_{l \leq k}] \quad \text{a.s.}$$

Condition (5.7.2) now follows. For each Borel set $U' \subseteq M_2$, let

$$g_U: M_1 \times M_2 \rightarrow [0, 1]$$

be a Borel function such that

$$P[\hat{y}_2(t) \in U | (\hat{y}_1(t), \hat{y}_2(s))] = g_U(\hat{y}_1(t), \hat{y}_2(s)).$$

Since $\sigma(\hat{x}_1(t))$ is countably generated, \mathcal{F}_t is atomless over $\mathcal{F}_s(\hat{x}_1(t))$. By Proposition 4.4(iii) there is an \mathcal{F}_t -measurable random variable $z(t)$ on $\underline{\Omega}$ such that for each Borel set $U \subseteq M_2$,

$$P[z(t) \in U | \mathcal{F}_s(\hat{x}_1(t))] = g_U(\hat{x}_1(t), z(s)) \quad \text{a.s.}$$

Then

$$P[z(t) \in U | (\hat{x}_1(t_1), \dots, \hat{x}_1(t_{k+1}), z(t_1), \dots, z(t_k))] = g_U(\hat{x}_1(t), z(s)),$$

and

$$P[\hat{y}_2(t) \in U | (\hat{y}_1(t_1), \dots, \hat{y}_1(t_{k+1}), \hat{y}_2(t_1), \dots, \hat{y}_2(t_k))] = g_U(\hat{y}_1(t), \hat{y}_2(s)).$$

It follows that

$$(\hat{x}_1, z) \equiv_0 (\hat{y}_1, \hat{y}_2) \quad \text{on } \{t_1, \dots, t_{k+1}\},$$

that is, (5.7.1) holds for $k + 1$.

It remains to show that z is a martingale on $\{t_1, \dots, t_{k+1}\}$, that is,

$$E[z(t) | \mathcal{F}_s] = z(s) \quad \text{a.s.}$$

In the following computation we use the fact that since \hat{x}_1 is \mathcal{F}_t -Markov and $\sigma(\hat{x}_1(s)) \subseteq \sigma(z(s)) \subseteq \mathcal{F}_s$,

$$E[h(\hat{x}_1(t), z(s)) | \mathcal{F}_s] = E[h(\hat{x}_1(t), z(s)) | z(s)]$$

for any Borel $h: M_1 \times M_2 \rightarrow M_1$. By definition of $z(t)$ there is an h such that

$$E[\hat{y}_2(t) | (\hat{y}_1(t), \hat{y}_2(s))] = h(\hat{y}_1(t), \hat{y}_2(s)) \quad \text{a.s.}$$

and

$$E[z(t) | \mathcal{F}_s(\hat{x}_1(t))] = h(\hat{x}_1(t), z(s)) \quad \text{a.s.}$$

Then

$$E[h(\hat{y}_1(t), \hat{y}_2(s)) | \hat{y}_2(s)] = E[\hat{y}_2(t) | \hat{y}_2(s)] = \hat{y}_2(s) \quad \text{a.s.}$$

By (5.7.2) and the above,

$$E[h(\hat{x}_1(t), z(s)) | z(s)] = z(s) \quad \text{a.s.}$$

Then

$$\begin{aligned} E[z(t) | \mathcal{F}_s] &= E[E[z(t) | \mathcal{F}_s(\hat{x}_1(t))] | \mathcal{F}_s] \\ &= E[h(\hat{x}_1(t), z(s)) | \mathcal{F}_s] \\ &= E[h(\hat{x}_1(t), z(s)) | z(s)] = z(s) \quad \text{a.s.} \quad \square \end{aligned}$$

PROOF OF THEOREM 5.2. Let $\underline{\Omega}$ be a complete atomless adapted space, and let $\underline{\Lambda}$ be some other adapted space. Let x_1 be a stochastic process on $\underline{\Omega}$ and (y_1, y_2) a stochastic process on $\underline{\Lambda}$ such that $x_1 \equiv y_1$. We show that there is a stochastic process x_2 on $\underline{\Omega}$ such that $(x_1, x_2) \equiv (y_1, y_2)$.

Case 1. x_1, y_1 , and y_2 are r.c.l.l. processes. Let $T_1 \subseteq T_2 \subseteq \cdots$ be an increasing chain of finite subsets of \mathbf{R}^+ such that $T = \bigcup_n T_n$ is dense. By Proposition 4.7, $\underline{\Omega}$ is \aleph_1 -atomless. By Lemma 5.7, for each n there is a stochastic process x_2^n on $\underline{\Omega}$ such that

$$(x_1, x_2^n) \equiv (y_1, y_2) \quad \text{on } T_n.$$

Since each finite \bar{t} in T belongs to some T_n , it follows that

$$(x_1, x_2^n) \xrightarrow{\text{ad}} (y_1, y_2) \quad \text{on } T.$$

Then there is a stochastic process u on $\underline{\Omega}$ such that

$$(x_1, x_2^n) \xrightarrow{\text{ad}} (x_1, u) \quad \text{on } T.$$

It follows that $(x_1, u) \equiv (y_1, y_2)$ on T . Since $u \equiv_0 y_2$ on T , we may define x_2 on $\underline{\Omega}$ by

$$x_2(t) = \lim_{\substack{s \downarrow t \\ s \in T}} u(s) \quad \text{a.s.}$$

$x_2 = u$ on T , so $(x_1, x_2) \equiv (y_1, y_2)$ on T . Moreover, x_2 is r.c.l.l., so by Corollary 2.13 we have $(x_1, x_2) \equiv (y_1, y_2)$. This proves the result in Case 1.

We now turn to the general case. By Proposition 2.24 there are r.c.l.l. processes w_1 on $\underline{\Omega}$ and (z_1, z_2) on $\underline{\Lambda}$ and Borel functions φ_1, φ_2 such that $w_1 \equiv z_1$ and for all $t \in \mathbf{R}^+$

$$x_1(t) = \varphi_1(w_1(t)) \text{ a.s.}, \quad y_1(t) = \varphi_1(z_1(t)) \text{ a.s.}, \quad y_2(t) = \varphi_2(z_2(t)) \text{ a.s.}$$

By Case 1 there is a w_2 on $\underline{\Omega}$ such that $(w_1, w_2) \equiv (z_1, z_2)$. Define x_2 on $\underline{\Omega}$ by

$$x_2(t) = \varphi_2(w_2(t)), \quad t \in \mathbf{R}^+.$$

It follows that

$$(\varphi_1(w_1), \varphi_2(w_2)) \equiv (\varphi_1(z_1), \varphi_2(z_2)),$$

so $(x_1, x_2) \equiv (y_1, y_2)$. \square

COROLLARY 5.8. *If $\underline{\Omega}$ is saturated then it has the saturation property for r.c.l.l. processes, that is:*

If x_1 is an r.c.l.l. process on $\underline{\Omega}$, (y_1, y_2) is an r.c.l.l. process on $\underline{\Lambda}$, and $x_1 \equiv y_1$, then there is an r.c.l.l. process x_2 on $\underline{\Omega}$ such that $(x_1, x_2) \equiv (y_1, y_2)$.

PROOF. This was proved in Case 1 of the proof of Theorem 5.2. \square

The next result show that saturation for \equiv implies saturation for \equiv a.e.

COROLLARY 5.9. *If x_1 is a process on a saturated adapted space $\underline{\Omega}$ and y_1, y_2 are processes on an adapted space $\underline{\Lambda}$ such that $x_1 \equiv y_1$ a.e., then there is x_2 on $\underline{\Omega}$ such that*

$$(5.9.1) \quad (x_1, x_2) \equiv (y_1, y_2) \quad \text{a.e.}$$

and

$$(5.9.2) \quad x_2 \equiv y_2.$$

PROOF. By Corollary 2.23 there are x'_1 on $\underline{\Omega}$ and y'_1 on $\underline{\Delta}$ such that

$$x'_1 = x_1 \text{ a.e.}, \quad y'_1 = y_1 \text{ a.e.}, \quad x'_1 \equiv y'_1.$$

Since $\underline{\Omega}$ is saturated there is x_2 on $\underline{\Omega}$ such that $(x'_1, x_2) \equiv (y'_1, y_2)$. Then (5.9.1) and (5.9.2) hold for x_1, x_2, y_1, y_2 . \square

COROLLARY 5.10 (AMALGAMATION THEOREM). *Let x_1, y_1 be stochastic processes on an adapted space $\underline{\Delta}_1$, and let (y_2, z_2) be processes on an adapted space $\underline{\Delta}_2$ such that $y_1 \equiv y_2$. Then there is an adapted space $\underline{\Omega}$ with stochastic processes x, y, z such that $(x, y) \equiv (x_1, y_1)$ and $(y, z) \equiv (y_2, z_2)$. The same holds with \equiv a.e. instead of \equiv . If x_1, y_1, y_2 , and z_2 are r.c.l.l., then x, y may be taken to be r.c.l.l.*

PROOF. Take a saturated space for $\underline{\Omega}$. Choose (x, y) so that $(x, y) \equiv (x_1, y_1)$. Then choose z so that $(y, z) \equiv (y_2, z_2)$. By Corollary 5.9, the same can be done with \equiv a.e. instead of \equiv . \square

We will now prove some results which are in a sense converse to Theorem 5.2 and Corollary 5.10. They assert that \equiv is the weakest reasonable equivalence relation on stochastic processes for which a certain amalgamation theorem holds or for which a saturated space exists. In these results we will allow stochastic processes to have parameters either in \mathbf{R}^+ or $\mathbf{R}^+ \cup \{\infty\}$.

THEOREM 5.11. *Let \sim and \approx be equivalence relations on stochastic processes, and suppose these relations have the following properties:*

- (1) $x \approx y$ implies $x \equiv_0 y$.
- (2) If $(x_1, y_1) \approx (x_2, y_2)$, then $y_1 \approx y_2$.
- (3) \approx preserves the martingale property.
- (4) Given $(x_1, y_1), (x_2, z_2)$ such that $x_1 \sim x_2$, there is a process (x, y, z) such that

$$(x, y) \approx (x_1, y_1) \quad \text{and} \quad (x, z) \approx (x_2, z_2).$$

Then $x_1 \sim x_2$ implies $x_1 \equiv x_2$.

PROOF. Suppose $x_1 \sim x_2$ and, by (4), choose (x, y, z) such that $(x, y) \approx (x_1, m^T x_1)$ and $(x, z) \approx (x_2, m^T x_2)$, where T is any countable subset of $\mathbf{R}^+ \cup \{\infty\}$. By (2) and (4), $x_1 \approx x_2$. By (2) and (3) y and z are both martingales, so by (1) and Lemma 5.6(i), both are versions of $m^T x$. By Lemma 5.5, $x_1 \equiv x_2$ on T . Since T was any countable subset of $\mathbf{R}^+ \cup \{\infty\}$ we must have $x_1 \equiv x_2$. \square

It follows from this theorem that the amalgamation property (4) cannot hold when \sim is \equiv_n and \approx is \equiv_1 , for any finite n .

THEOREM 5.12. *Let \approx be an equivalence relation on stochastic processes having properties (1) and (2) of Theorem 5.11 and also the following property:*

- (3') If $x \approx y$ and x is adapted then y is adapted.

Suppose $\underline{\Omega}_1$ and $\underline{\Omega}_2$ are spaces having the following saturation property:

(4') Whenever (z_1, v_1) are processes on $\underline{\Omega}_1$ and (x_2, w_2) are processes on $\underline{\Omega}_2$ such that $z_1 \approx z_2$, then there are processes v_2 on $\underline{\Omega}_2$ and w_1 on $\underline{\Omega}_1$ such that $(z_1, v_1) \approx (z_2, v_2)$ and $(z_1, w_1) \approx (z_2, w_2)$.

If x is a process on $\underline{\Omega}_1$, y a process on $\underline{\Omega}_2$ such that $x \approx y$, then $x \equiv y$.

PROOF. Let $x \approx y$ be processes as in the hypothesis. Choose z on $\underline{\Omega}_1$ so that $(x, z) \approx (y, m^T y)$, T a countable subset of $\mathbf{R}^+ \cup \{\infty\}$. If we show that z is a martingale, then it will follow as in 5.11 that $x \equiv y$. Suppose z is not a martingale. Then there are $t_1, t_2 \in \mathbf{R}^+ \cup \{\infty\}$, $t_1 < t_2$, and $F \in \mathcal{F}_{t_1}^1$ such that

$$(5.12.1) \quad E[z(t_1) \cdot 1(F)] \neq E[z(t_2) \cdot 1(F)].$$

Let v be the process

$$v(t) = \begin{cases} 0, & t < t_1, \\ 1(F), & t \geq t_1. \end{cases}$$

By (2) and (4'), choose w such that $(z, v) \approx (m^T y, w)$. Then by (1), there is a set G such that

$$w(t) = \begin{cases} 0, & t < t_1, \\ 1(G), & t \geq t_1 \text{ a.s.} \end{cases}$$

By (2) and (3'), $G \in \mathcal{F}_{t_1}^2$. But then by (5.12.1) and (1),

$$E[m^T y(t_1) \cdot 1(G)] = E[z(t_1) \cdot 1(F)] \neq E[z(t_2) \cdot 1(F)] = E[m^T y(t_2) \cdot 1(G)].$$

But this is impossible, since $m^T y$ is a martingale. Hence z must indeed be a martingale. \square

6. Application to semimartingales. Let x be a local martingale on some adapted space $\underline{\Delta}$. In this section we shall show that if \hat{x} is an r.c.l.l. process on a saturated space and \hat{x} has the same adapted law as x (i.e. $x \equiv \hat{x}$) then \hat{x} is a local martingale.

The first author has in fact shown that the local and semimartingale properties are even preserved under synonymy. We include the weaker results in this section as a first example of a general method using saturated spaces. As an application we improve a result of Barlow [1981] and Perkins [1982] on local martingales with prescribed absolute value. The results and methods of this section will be applied to stochastic integral equations in §7.

We first state the definition of local martingale (see Kussmaul [1977]).

DEFINITION 6.1. A stochastic process x with values in \mathbf{R} is a *local martingale* if x is adapted and there is a sequence τ_n , $n \in \mathbf{N}$, of stopping times such that τ_n in increasing $\lim_{n \rightarrow \infty} \tau_n = \infty$ a.s., and for each n the stopped process $x(\tau_n(\omega) \wedge t, \omega) - x(0, \omega)$ is a uniformly integrable martingale. In this case we shall say that the sequence τ_n *reduces* x .

Each local martingale is a right continuous left limit process.

THEOREM 6.2. Let x be a local martingale on some space $\underline{\Delta}$ and \hat{x} a right continuous left limit process on a saturated adapted space $\underline{\Omega}$. If \hat{x} has the same adapted law as x , then \hat{x} is a local martingale.

PROOF. x is a right continuous left limit adapted process. Let τ_n , $n \in \mathbb{N}$, be a sequence of stopping times on $\underline{\Lambda}$ which reduce x , that is,

$$(6.2.1) \quad \tau_n \rightarrow \infty \quad \text{a.s.}$$

$$(6.2.2) \quad E(x(\tau_n \wedge t) | \mathcal{G}_s) = x(\tau_n \wedge s) \quad \text{a.s. when } s \leq t,$$

Uniform integrability: for each $n \in \mathbb{N}$,

$$(6.2.3) \quad \lim_{m \rightarrow \infty} E[|x(\tau_n \wedge t)| \cdot 1_{|x(\tau_n \wedge t)| \geq m}] = 0$$

uniformly in t . (I_U is the indicator function of U .)

For each $n \in \mathbb{N}$, let z_n be the indicator function of $\tau_n \leq t$, that is

$$z_n(t, \omega) = \begin{cases} 1 & \text{if } \tau_n(\omega) < t, \\ 0 & \text{if } \tau_n(\omega) \geq t. \end{cases}$$

The paths of z_n are right continuous step functions.

Since $\underline{\Omega}$ is saturated and $x \equiv \hat{x}$, there are processes \hat{z}_n , $n \in \mathbb{N}$, on $\underline{\Omega}$ such that

$$(6.2.4) \quad (\hat{x}, \hat{z}_n)_{n \in \mathbb{N}} \equiv (x, z_n)_{n \in \mathbb{N}}.$$

By the remark following Corollary 2.15, the \hat{z}_n may be taken to have paths which are indicator functions of intervals of the form $[0, \tau)$. Now using only the synonymity relation

$$(6.2.5) \quad (\hat{x}, \hat{z}_n)_{n \in \mathbb{N}} \equiv_1 (z, z_n)_{n \in \mathbb{N}}$$

we shall show that \hat{x} is a local martingale on $\underline{\Omega}$.

Since x and z_n are adapted, \hat{x} and \hat{z}_n are adapted by (6.2.5). It follows that for each n there is a stopping time $\hat{\tau}_n$ on $\underline{\Omega}$ such that

$$\hat{z}_n = 1_{[0, \hat{\tau}_n)}.$$

By (6.2.5) and $\tau_n \rightarrow \infty$ a.s., we have $\hat{\tau}_n \rightarrow \infty$ a.s. Properties (6.2.2) and (6.2.3) for \hat{x} and $\hat{\tau}_n$ follow easily from the synonymity (6.2.5) and the corresponding properties for x , τ_n . Thus $\hat{\tau}_n$ reduces \hat{x} , and \hat{x} is a local martingale. \square

Using the above result we now prove the analogous theorem for semimartingales.

DEFINITION 6.3. Let x be a stochastic process with values in \mathbb{R}^d . x is a process of *bounded variation* if almost every path of x is r.c.l.l. and of bounded variation on each interval $[0, t]$. x is a *semimartingale* if x can be written as a sum $x = m + v$, where m is a local martingale and v is an adapted process of bounded variation. A sequence of stopping times τ_n is said to *reduce* x if it reduces m . Note that every semimartingale is a right continuous left limit process.

LEMMA 6.4. If x is a process of bounded variation and y is an r.c.l.l. process such that $x \equiv_0 y$, then y is of bounded variation.

PROOF. The variation process $\text{var}(x)$ of the process x is given by

$$\text{var}(x)(t) = \lim_{n \rightarrow \infty} \sum_{0 \leq i < nt} \left| x\left(\frac{i+1}{n}\right) - x\left(\frac{i}{n}\right) \right|.$$

As the same is true of y , and $x \equiv_0 y$, $\text{var}(x) \equiv_0 \text{var}(y)$. Since $\text{var}(x)(t)$ is almost surely finite for all t , so is $\text{var}(y)$. Since $\text{var}(y)$ is increasing, this implies that y is of bounded variation. \square

THEOREM 6.5. *Let x be a semimartingale on some adapted space $\underline{\Delta}$ and \hat{x} an r.c.l.l. process on a saturated adapted space $\underline{\Omega}$. If \hat{x} has the same adapted law as x , then \hat{x} is a semimartingale.*

PROOF. Let $x = m + v$, where m is a local martingale and v is a process of bounded variation. Choose \hat{m} and \hat{v} on $\underline{\Omega}$ such that

$$(\hat{x}, \hat{m}, \hat{v}) \equiv (x, m, v),$$

and \hat{m}, \hat{v} are r.c.l.l. By Lemma 6.4, \hat{v} is of bounded variation, and by Theorem 6.2, \hat{m} is a local martingale. Since $x = m + v$ we have $\hat{x} = \hat{m} + \hat{v}$ a.s., whence \hat{x} is a semimartingale. \square

LEMMA 6.6. *For every semimartingale x there is an increasing sequence of stopping times τ_n such that τ_n reduces x and $|x|$ is bounded by n on $[0, \tau_n)$.*

PROOF. Let σ_n be an increasing sequence of stopping times which reduces x . Let

$$\tau_n = \sigma_n \wedge \inf\{t: |x(t)| \geq n\}.$$

Then τ_n is a stopping time and $|x|$ is bounded by n on $[0, \tau_n)$. Since $\sigma_n \rightarrow \infty$ a.s. and the paths of x are right continuous with left limits, $\tau_n \rightarrow \infty$ a.s. Moreover, $\tau_n \leq \sigma_n$, and hence τ_n reduces x . \square

THEOREM 6.7. *Let x be a semimartingale and τ_n an increasing sequence of stopping times on an adapted space $\underline{\Delta}$, and \hat{x} a semimartingale and $\hat{\tau}_n$ an increasing sequence of stopping times on a saturated adapted space $\underline{\Omega}$, such that*

$$(x, 1_{[0, \tau_n)})_{n \in \mathbb{N}} \equiv (\hat{x}, 1_{[0, \hat{\tau}_n)})_{n \in \mathbb{N}}.$$

Then:

- (i) *If τ_n reduces x , then $\hat{\tau}_n$ reduces \hat{x} .*
- (ii) *If x is bounded by b_n on each interval $[0, \tau_n)$, then \hat{x} is almost surely bounded by b_n on each interval $[0, \hat{\tau}_n)$.*

PROOF. Let $x = m + v$, $\hat{x} = \hat{m} + \hat{v}$, where m, \hat{m} are local martingales and v, \hat{v} are of bounded variation and

$$(x, m, v, (1_{[0, \tau_n)}))_{n \in \mathbb{N}} \equiv (\hat{x}, \hat{m}, \hat{v}, (1_{[0, \hat{\tau}_n)}))_{n \in \mathbb{N}}.$$

Suppose τ_n reduces m . Then by the proof of Theorem 6.2, $\hat{\tau}_n$ reduces \hat{m} . Suppose, finally, that x is bounded by b_n on each interval $[0, \tau_n)$. That is, $x \cdot 1_{[0, \tau_n)} \leq b_n$. Then

$$\hat{x} \cdot 1_{[0, \hat{\tau}_n)} \leq b_n \quad \text{a.s.} \quad \square$$

A *submartingale* is an r.c.l.l. adapted process x such that whenever $s \leq t$, $E[x(t)|\mathcal{F}_s] \geq x(s)$ a.s. The definition of a *local submartingale* is analogous to the definition of a local martingale. By essentially the same proof, we obtain an analogue of Theorem 6.2 for local submartingales.

THEOREM 6.8. *Let x be a local submartingale on some space $\underline{\Delta}$ and \hat{x} be an r.c.l.l. process on a saturated space $\underline{\Omega}$. If \hat{x} has the same adapted law as x then \hat{x} is a local submartingale.*

As an application we improve a result of Perkins [1982] and Barlow [1981] on martingales with a given absolute value. The improvement we obtain is the following. The result holds with or without the word “local”.

THEOREM 6.9. *Let $\underline{\Omega}$ be a saturated adapted space. For every nonnegative (local) submartingale x on $\underline{\Omega}$ there exists a (local) martingale m on $\underline{\Omega}$ such that*

$$|m(\omega, \cdot)| = x(\omega, \cdot) \quad \text{a.s.}$$

COROLLARY 6.10. *Let $\underline{\Omega}$ be a saturated adapted space. For every nonnegative (local) submartingale y on an arbitrary space $\underline{\Delta}$ there exists a (local) martingale m on $\underline{\Omega}$ such that $|m| \equiv y$.*

We now state the result of Perkins [1982] and obtain Theorem 6.9 as a consequence. Perkins constructed a pair of saturated adapted spaces

$$\underline{\Omega}^1 = (\Omega^1, P^1, \mathcal{F}_t^1)_{t \in \mathbf{R}^+}, \quad \underline{\Omega}^2 = (\Omega^2, P^2, \mathcal{F}_t^2)_{t \in \mathbf{R}^+}$$

and a mapping $\pi: \Omega^2 \rightarrow \Omega^1$ with the following properties:

(a) For every nonnegative (local) submartingale \hat{x} on $\underline{\Omega}^1$ there is a (local) martingale \hat{m} on $\underline{\Omega}^2$ such that

$$|\hat{m}(\omega, \cdot)| = \hat{x}(\pi\omega, \cdot) \quad \text{a.s. } (P^2).$$

(b) $\pi^{-1}(\mathcal{F}_\infty^1) \subseteq \mathcal{F}_\infty^2$ and $\pi^{-1}(\mathcal{F}_t^1) \subseteq \mathcal{F}_t^2$ for all $t \in \mathbf{R}^+$.

(c) π^{-1} is measure-preserving.

(d) For every random variable $u: \Omega^1 \rightarrow \mathbf{R}^+$ on $\underline{\Omega}^1$,

$$E[u|\mathcal{F}_t^1](\pi\omega) = E[u \circ \pi|\mathcal{F}_t^2](\omega) \quad \text{a.s. } (P^2).$$

To prove Theorem 6.9 let x be a nonnegative (local) submartingale on $\underline{\Omega}$. Since $\underline{\Omega}^1$ is saturated and x is right continuous there is a process \hat{x} on $\underline{\Omega}^1$ such that $x \equiv \hat{x}$ and \hat{x} is right continuous. Moreover, \hat{x} is a nonnegative (local) submartingale by Theorem 6.8. Let z be the process on $\underline{\Omega}^2$ defined by $z = \hat{x} \circ \pi$, i.e.

$$z(\omega, t) = \hat{x}(\pi\omega, t).$$

Using properties (b), (c), (d), it can be shown that for every conditional process $f(\vec{t})$,

$$fz(\vec{t})(\omega) = f\hat{x}(\vec{t})(\pi\omega) \quad \text{a.s. } (P^2).$$

The main inductive step, for conditional expectations, uses (d). It follows that $E[fz(\vec{t})] = E[f\hat{x}(\vec{t})]$, and therefore $z \equiv \hat{x}$.

Now consider the (local) martingale \hat{m} in condition (a) and form the joint process (\hat{m}, z) . By saturation of $\underline{\Omega}$ there is a right continuous process m on $\underline{\Omega}$ such that $(\hat{m}, z) \equiv (m, x)$. We have $|\hat{m}(t)| = z(t)$ a.s. and hence $|m(t)| = x(t)$ a.s. for all $t \in \mathbf{R}^+$. By Theorem 6.2, m is a (local) martingale on $\underline{\Omega}$, as required. \square

7. Application to stochastic integral equations. In this section we consider stochastic integrals and stochastic integral equations with respect to a semimartingale. Given a solution x to a stochastic integral equation E on some adapted space $\underline{\Delta}$, we shall show that on a saturated space $\underline{\Omega}$, any stochastic integral equation F whose coefficients and semimartingale have the same adapted law as those of E has a solution y with the same adapted law as x . Thus on saturated spaces, the existence of a solution, and the uniqueness of the adapted law of a solution, depend only on the adapted laws of the coefficients and of the semimartingale.

We begin with the relevant definitions.

DEFINITION 7.1. The *predictable* (or *previsible*) σ -field on an adapted space $\underline{\Omega} = (\Omega, P, \tau_t)_{t \in \mathbf{R}^+}$ is the σ -field on $\Omega \times \mathbf{R}^+$ generated by all sets of the form

$$[0, \tau] = \{(\omega, t) : 0 \leq t \leq \tau(\omega)\}$$

where τ is a stopping time. A stochastic process x on $\underline{\Omega}$ is *predictable* if it is measurable with respect to the predictable σ -field.

DEFINITION 7.2. A *basic predictable process* is a finite linear combination of indicator functions

$$a_1 \cdot 1_{[0, \tau_1]} + a_2 \cdot 1_{(\tau_1, \tau_2]} + \cdots + a_n \cdot 1_{(\tau_{n-1}, \tau_n]}$$

where τ_1, \dots, τ_n are stopping times, a_1, \dots, a_n are real numbers, and $(\sigma, \tau] = \{(\omega, t) : \sigma(\omega) \leq t \leq \tau(\omega)\}$.

The basic predictable processes form a vector space closed under the minimum operation $f \wedge g$. It follows from the monotone class theorem that the closure of the class of basic predictable processes under bounded monotone convergence is exactly the class of bounded predictable processes. Thus h is bounded predictable process if and only if there is a sequence h_n of basic predictable processes and a bounded Borel function $\varphi: \mathbf{R}^{\mathbf{N}} \rightarrow \mathbf{R}$ such that for all t and ω ,

$$h(t, \omega) = \varphi((h_n(t, \omega))_{n \in \mathbf{N}}).$$

The stochastic integral of a bounded predictable process with respect to a semimartingale may be defined as follows. We could consider the more general case of integration of a predictable process which is locally integrable with respect to a semimartingale m , but the additional technical complications would obscure our exposition. All our results are proved for one dimension but generalize readily to d dimensions.

DEFINITION 7.3. The stochastic integral is the unique function assigning to each bounded \mathbf{R} -valued predictable process h and \mathbf{R} -valued semimartingale m a right continuous left limit \mathbf{R} -valued process $\int_0^t h \, dm$ such that

(7.3.1) $\int_0^t h \, dm$ is linear in both h and m .

(7.3.2) If τ is a stopping time then $\int_0^t 1_{[0, \tau]} \, dm = m(\tau \wedge t)$.

(7.3.3) If h_k is a bounded monotone sequence of predictable processes converging to h , then

$$\int_0^t h \, dm = \lim_{k \rightarrow \infty} \int_0^t h_k \, dm$$

in probability.

The stochastic integral can be shown to exist and to be unique up to indistinguishability.

For the remainder of this section let $\underline{\Omega} = (\Omega, P, \mathcal{F}_t)_{t \in \mathbb{R}^+}$ be a saturated adapted probability space and let $\underline{\Lambda} = (\Lambda, Q, \mathcal{G}_t)_{t \in \mathbb{R}^+}$ be arbitrary.

LEMMA 7.4. *Let h be a basic predictable process on $\underline{\Lambda}$ and let \hat{h} be a process on $\underline{\Omega}$ such that $h \equiv_1 \hat{h}$. Then \hat{h} has a version \hat{g} which is a basic predictable process.*

PROOF. Let

$$h = a_1 \cdot 1_{[0, \tau_1]} + a_2 \cdot 1_{(\tau_1, \tau_2]} + \cdots + a_n \cdot 1_{(\tau_{n-1}, \tau_n]}.$$

Then the paths of h are left continuous step functions with values at a_1, \dots, a_n . By the remark following Corollary 2.15, since $h \equiv_0 \hat{h}$, \hat{h} has a version \hat{g} whose paths are also left continuous step functions with values at a_1, \dots, a_n . Then $h \equiv_1 \hat{h}$, $h \equiv_1 \hat{g}$. It follows that the steps of \hat{g} occur at stopping times, so \hat{g} is a basic predictable process. \square

THEOREM 7.5. *Let h, z, m be processes on $\underline{\Lambda}$, and $\hat{h}, \hat{z}, \hat{m}$ be processes on $\underline{\Omega}$ such that:*

- (i) $(h, m, z) \equiv (\hat{h}, \hat{m}, \hat{z})$.
- (ii) h and \hat{h} have values in \mathbb{R}^N and all coordinates h_n and \hat{h}_n are basic predictable processes.
- (iii) m and \hat{m} are semimartingales. Then for every bounded Borel function $\varphi: \mathbb{R}^N \rightarrow \mathbb{R}$ we have

$$\left(h, m, z, \int_0^t \varphi(h) dm \right) \equiv \left(\hat{h}, \hat{m}, \hat{z}, \int_0^t \varphi(\hat{h}) d\hat{m} \right).$$

PROOF. If φ depends on finitely many coordinates, then $\varphi(h)$ is basic and the result is clear. A monotone class argument together with (7.3.3) and Proposition 2.20 now completes the proof. \square

It follows from Theorem 7.5 that if h, \hat{h} are real predictable processes such that h is predictable and $(h, m) \equiv (\hat{h}, \hat{m})$, then \hat{h} has a predictable version \tilde{h} such that

$$\int h dm \equiv \int \tilde{h} d\hat{m},$$

but in general we need not have $\int h dm \equiv_0 \int \hat{h} d\hat{m}$ even when \hat{h} is predictable.

We conclude this paper with an application to stochastic integral equations.

If $x(t, \omega)$ is an r.c.l.l. process, the process $x(t-, \omega)$ is defined by

$$x(t-, \omega) = \begin{cases} \lim_{s \uparrow t} x(s, \omega) & \text{if } t > 0, \\ x(0, \omega) & \text{if } t = 0. \end{cases}$$

LEMMA 7.6. *If $x(t, \omega)$ is an adapted r.c.l.l. process, then $x(t-, \omega)$ is predictable. In fact, if S is the Polish space of r.c.l.l. functions from \mathbb{R}^+ into \mathbb{R}^d , there are Borel functions*

$$\theta: S \rightarrow S^N, \quad \psi: (\mathbb{R}^d)^N \rightarrow \mathbb{R}^d$$

such that for each adapted r.c.l.l. process $x(t, \omega)$, each coordinate $\theta_n(x)(t)$ is a basic predictable process and for all t ,

$$x(t-) = \psi(\theta(x)(t)).$$

PROOF. For $u \in \mathbb{R}^d$ let $l_n(u)$ be the greatest $(1/n)$ -lattice point which is $\leq u$ in each coordinate. Given an r.c.l.l. function x and positive integer p , define points t_0^p, t_1^p, \dots , as follows.

$$t_0^p = 0, \quad t_{i+1}^p = \text{least } t \geq t_i^p \text{ such that } \|x(t) - x(t_i^p)\| \geq 1/p.$$

This sequence $\{t_i^p\}$ is obtained as a Borel function of x . Let $\langle \cdot, \cdot \rangle$ be a bijection of \mathbb{N}^2 and \mathbb{N} . Define $\theta_n, n \in \mathbb{N}$, by

$$\theta_{\langle p, q \rangle}(x)(0) = l_q(x(0)).$$

For $t > 0$,

$$\begin{aligned} \theta_{\langle p, q \rangle}(x)(t) &= l_q(x(t_i)) \quad \text{for } t \in [t_i, t_{i+1}], i < p, \\ &= l_q(x(t_p)), \quad t > t_p. \end{aligned}$$

Then each θ_n is a Borel function from D to the space of left-continuous step functions, if x is adapted then $\theta_n(x)$ is basic predictable, and if

$$\psi(y_n: n \in \mathbb{N}) = \lim_{p \rightarrow \infty} \lim_{q \rightarrow \infty} y_{\langle p, q \rangle},$$

then $x(t-) = \psi(\theta(x)(t))$ for all t . \square

DEFINITION 7.7. Let m be a semimartingale on $\underline{\Delta}$, and let f be a bounded measurable function $((\mathbb{R}^+ \times \Omega) \times \mathbb{R}, \mathcal{P} \times \mathcal{B}) \rightarrow (\mathbb{R}, \mathcal{B})$. Let x_0 be a random variable on $\underline{\Delta}$. By a *solution* of the stochastic integral equation

$$(7.7.1) \quad x(t, \omega) = x_0(\omega) + \int_0^t f(s, \omega, x(s-, \omega)) dm(s, \omega)$$

we mean an adapted r.c.l.l. process x on $\underline{\Delta}$ such that the equation (7.7.1) holds almost surely.

The process $f(s, x(s-))$ is predictable for the following reason: By Fubini's Theorem, any f as specified must have the form

$$f(t, \omega, r) = \psi(g_n(t, \omega)h_n(r))$$

where for each $n \in \mathbb{N}$, g_n is basic predictable, h_n is continuous, and $\psi: \mathbb{R}^N \rightarrow \mathbb{R}$ is a bounded Borel function. Since $x(t-)$ is predictable, so is $g_n(t)h_n(x(t-))$. Hence $f(s, x(s-))$ is predictable.

THEOREM 7.8. Let f, m, z, x be processes on $\underline{\Delta}$ and $\hat{f}, \hat{m}, \hat{z}$ processes on $\underline{\Omega}$ such that

(i) $(f, m, z) \equiv (\hat{f}, \hat{m}, \hat{z})$.

(ii) f is a sequence of processes $(f_n: n \in \mathbb{N}), f_n: \mathbb{R}^+ \times \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ where for each n

$$f_n(t, \omega, r) = g_n(t, \omega)h_n(r),$$

$g_n: \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}$ being basic predictable, and $h_n: \mathbb{R} \rightarrow \mathbb{R}$ being continuous. The analogous statement holds for \hat{f} .

(iii) m and \hat{m} are semimartingales.

(iv) $\varphi: \mathbb{R}^N \rightarrow \mathbb{R}$ is a bounded Borel function.

(v) x is a solution on Λ of the stochastic integral equation

$$x(t) = x(0) + \int_0^t (\varphi \circ f)(s, x(s-)) dm(s).$$

Then there is a solution \hat{x} on $\underline{\Omega}$ of

$$\hat{x}(t) = \hat{x}(0) + \int_0^t (\varphi \circ \hat{f})(s, \hat{x}(s-)) d\hat{m}(s)$$

such that $(f, m, z, x) \equiv (\hat{f}, \hat{m}, \hat{z}, \hat{x})$.

PROOF. By saturation there is an r.c.l.l. process \hat{x} on $\underline{\Omega}$ such that $(f, m, z, x) \equiv (\hat{f}, \hat{m}, \hat{z}, \hat{x})$. We will show that \hat{x} is the desired solution. Let θ and ψ be the Borel functions introduced in Lemma 7.6, so that $\theta_n(x)$ and $\theta_n(\hat{x})$ are basic predictable processes and

$$\begin{aligned} x(t-) &= \psi(\theta(x)(t)) = \lim_{n \rightarrow \infty} \theta_n(x)(t), \\ \hat{x}(t-) &= \psi(\theta(\hat{x})(t)) = \lim_{n \rightarrow \infty} \theta_n(\hat{x})(t). \end{aligned}$$

Then $f_k(\theta_n(x))$ and $\hat{f}_k(\theta_n(\hat{x}))$ are basic predictable processes and

$$(f_k(\theta_n(x)), m, z, x)_{k, n \in \mathbb{N}} \equiv (\hat{f}_k(\theta_n(\hat{x})), \hat{m}, \hat{z}, \hat{x})_{k, n \in \mathbb{N}}.$$

By continuity of f, \hat{f} in the last coordinate,

$$\begin{aligned} f_k(x(t-)) &= \lim_{n \rightarrow \infty} f_k(\theta_n(x)(t)) = \psi(f_k(\theta(x)(t))), \\ \hat{f}_k(\hat{x}(t-)) &= \lim_{n \rightarrow \infty} \hat{f}_k(\theta_n(\hat{x})(t)) = \psi(\hat{f}_k(\theta(\hat{x})(t))). \end{aligned}$$

Hence

$$\begin{aligned} (\varphi \circ f)(t, x(t-)) &= (\varphi \circ \psi)(f(t, \theta(x)(t))), \\ (\varphi \circ \hat{f})(t, \hat{x}(t-)) &= (\varphi \circ \psi)(\hat{f}(t, \theta(\hat{x})(t))). \end{aligned}$$

Hence by Theorem 7.5,

$$\begin{aligned} &\left(f, m, z, x, x(0) + \int_0^t (\varphi \circ f)(s, x(s-)) dm(s) \right) \\ &\equiv \left(\hat{f}, \hat{m}, \hat{z}, \hat{x}, \hat{x}(0) + \int_0^t (\varphi \circ \hat{f})(s, \hat{x}(s-)) d\hat{m}(s) \right). \end{aligned}$$

Thus, since x is a solution of the stochastic integral equation and \hat{x} is an adapted r.c.l.l. process, \hat{x} is a solution of the corresponding equation. \square

The following two corollaries adapt Theorem 7.8 in order to show directly how a stochastic integral equation and its solution may be transferred to a saturated space.

Let \mathcal{P} denote the predictable σ -algebra, \mathcal{B} the Borel sets.

COROLLARY 7.9. Let h, m, z , and x be processes on $\underline{\Lambda}$ and \hat{m} and \hat{z} be processes on $\underline{\Omega}$ satisfying:

- (1) $h: \mathbf{R}^+ \times \Lambda \times \mathbf{R} \rightarrow \mathbf{R}$ is bounded and $\mathcal{P} \times \mathcal{B}$ -measurable.
- (2) m is a semimartingale.
- (3) $(m, z) \equiv (\hat{m}, \hat{z})$.
- (4) $x_t = x_0 + \int_0^t h(s, \omega, x_{s-}) dz_s$.

Then there exist \hat{h} and \hat{x} on $\underline{\Omega}$ such that:

(1)' $\hat{h}: \mathbf{R}^+ \times \Omega \times \mathbf{R} \rightarrow \mathbf{R}$ is bounded and $\mathcal{P} \times \mathcal{B}$ -measurable.

(2)' \hat{x} is an r.c.l.l. process and satisfies

$$\hat{x}_t = \hat{x}_0 + \int_0^t \hat{h}(s, \omega, \hat{x}_{s-}) dm_s.$$

(3)' For every sequence $r_k: k \in \mathbf{N}$ of reals,

$$\begin{aligned} (m, z, x, (h(r_k))_{k \in \mathbf{N}}, h(\cdot, \omega, x - (\omega))) \\ \equiv (\hat{m}, \hat{z}, \hat{x}, (\hat{h}(r_k))_{k \in \mathbf{N}}, \hat{h}(\cdot, \omega, \hat{x} - (\omega))). \end{aligned}$$

PROOF. $h(t, \omega, r) = \psi((g_n(t, \omega)h_n(r))_{n \in \mathbf{N}})$ where ψ , g_n and h_n are as in 7.8(ii). By saturation and Lemma 7.4, there exist basic predictable \hat{g}_n , and r.c.l.l. \hat{x} s.t.

$$(m, z, (g_n)_{n \in \mathbf{N}}, x) \equiv (\hat{m}, \hat{z}, (\hat{g}_n)_{n \in \mathbf{N}}, \hat{x}).$$

Let $\hat{h}(t, \omega, r) = \psi((\hat{g}_n(t, \omega)h_n(r))_{n \in \mathbf{N}})$. The result now follows from Theorem 7.5 by the proof of Theorem 7.8. \square

COROLLARY 7.10. Let h, m, z, x and \hat{m}, \hat{z} be as in Corollary 7.9, but assume in addition that $h(t, \omega, \cdot)$ is continuous for each t, ω . Let $\tilde{h}: \mathbf{R}^+ \times \Lambda \rightarrow C(\mathbf{R}, \mathbf{R})$ be the predictable process defined by $\tilde{h}(t, \omega)(x) = h(t, \omega, x)$. Then there exist \hat{h}, \hat{x} on $\underline{\Omega}$ such that:

(1)'' $\hat{h}[0, \infty) \times \Omega \rightarrow C(\mathbf{R}, \mathbf{R})$ is predictable.

(2)'' x is r.c.l.l. and satisfies $\hat{x}(t) = \hat{x}(0) + \int_0^t \hat{h}(s, \omega)(\hat{x}_{s-}) d\hat{m}(s)$.

(3)'' $(\hat{m}, \hat{z}, \hat{h}, \hat{x}) \equiv (m, z, h, x)$.

PROOF. Define basic predictable $C(\mathbf{R}, \mathbf{R})$ -valued processes as in 7.2 except that the values a_i are taken to be elements of $C(\mathbf{R}, \mathbf{R})$. Since $C(\mathbf{R}, \mathbf{R})$ is a Polish space the standard results on measure theory for real functions can be adapted to show that there exist basic predictable $h_n: \mathbf{R}^+ \times \Omega \rightarrow C(\mathbf{R}, \mathbf{R})$ and a measurable Borel function $\varphi: C(\mathbf{R}, \mathbf{R})^{\mathbf{N}} \rightarrow C(\mathbf{R}, \mathbf{R})$ such that $\varphi[C(\mathbf{R}, \mathbf{R})^{\mathbf{N}}]$ is a uniformly bounded set of functions and $\varphi(\tilde{h}_n; n \in \mathbf{N}) = \tilde{h}$. Now if $\theta_k, k \in \mathbf{N}$, are the functions introduced in Lemma 7.6, we see that for each n ,

$$h_n = \lim_{k \rightarrow \infty} h_n \circ \theta_k$$

(limit in the topology of uniform convergence on bounded sets). By saturation and Lemma 7.4, there exist basic predictable $\hat{h}_n, n \in \mathbf{N}$, and r.c.l.l. \hat{x} on $\underline{\Omega}$ such that

$$(m, z, x, \tilde{h}_n)_{n \in \mathbf{N}} \equiv (\hat{m}, \hat{z}, \hat{x}, \hat{h}_n)_{n \in \mathbf{N}}.$$

Since evaluation is a continuous map $C(\mathbf{R}, \mathbf{R}) \times \mathbf{R} \rightarrow \mathbf{R}$, we see that

$$(m, z, x, \tilde{h}_n, \tilde{h}_n(\theta_k(x)))_{n, k \in \mathbf{N}} \equiv (\hat{m}, \hat{z}, \hat{h}_n, \hat{h}_n(\theta_k(\hat{x})))_{n, k \in \mathbf{N}}.$$

$\tilde{h}_n(\theta_k(x))$ and $\hat{h}_n(\theta_k(\hat{x}))$ are both basic predictable real-valued processes, and now the result follows as usual from Theorem 7.5.

REFERENCES

- [1981] D. Aldous, *Weak convergence and the general theory of processes*, preprint.
- [1976] R. M. Anderson, *A non-standard representation of Brownian motion and Itô integration*, Israel J. Math. **25** (1976), 15–46.
- [1981] M. T. Barlow, *Construction of a martingale with a given absolute value*, Ann. Probab. **9** (1981), 314–320.

- [1968] P. Billingsley, *Convergence of probability measure*, Wiley, New York, 1968.
- [1982] D. N. Hoover and E. Perkins, *Nonstandard construction of the stochastic integral and applications to stochastic differential equations*. I, II, Trans. Amer. Math. Soc. **275** (1983), 1–58.
- [1979] H. J. Keisler, *Hyperfinite probability theory and probability logic*, Lecture Notes, Univ. of Wisconsin (unpublished).
- [1982] ———, *An infinitesimal approach to stochastic analysis*, Mem. Amer. Math. Soc. (to appear).
- [1983] ———, *Probability quantifiers*, Abstract Model Theory and Logics of Mathematical Concepts (J. Barwise and S. Feferman, eds.), Springer-Verlag, Berlin and New York (to appear).
- [1975] F. B. Knight, *A predictive view of continuous time processes*, Ann. Probab. **3** (1975), 573–596.
- [1977] A. U. Kussmaul, *Stochastic integration and generalized martingales*, Pitman, New York, 1977.
- [1979] P. A. Loeb, *An introduction to nonstandard analysis and hyperfinite probability theory*, Probabilistic Analysis and Related Topics 2, (Bharucha and Reid, eds.), Academic Press, New York, 1979, pp. 105–142.
- [1942] D. Maharam, *On homogeneous measure algebras*, Proc. Nat. Acad. Sci. U.S.A. **28** (1942), 108–111.
- [1950] ———, *Decompositions of measure algebras and spaces*, Trans. Amer. Math. Soc. **69** (1950), 142–160.
- [1980] M. Metivier and J. Pellaumail, *Stochastic integration*, Academic Press, New York, 1980.
- [1982] E. Perkins, *On the construction and distribution of a local martingale with a given absolute value*, Trans. Amer. Math. Soc. **271** (1982), 261–281.
- [1982] H. Rodenhausen, *The completeness theorem for adapted probability logic*, Ph.D. Thesis, Heidelberg University.
- [1983] K. D. Stroyan and J. M. Bayod, *Foundations of infinitesimal stochastic analysis*, North-Holland, Amsterdam (to appear).
- [1978] P. A. Meyer and C. Dellacherie, *Probabilities and potential*, North-Holland Mathematical Studies No. 29, North-Holland, Amsterdam, 1978.

DEPARTMENT OF MATHEMATICS AND STATISTICS, QUEEN'S UNIVERSITY, KINGSTON, ONTARIO, CANADA K7L 3N6

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN 53706