

# NONSTANDARD CONSTRUCTION OF THE STOCHASTIC INTEGRAL AND APPLICATIONS TO STOCHASTIC DIFFERENTIAL EQUATIONS. II

BY

DOUGLAS N. HOOVER<sup>1</sup> AND EDWIN PERKINS

**ABSTRACT.** H. J. Keisler has recently used a nonstandard theory of Itô integration (due to R. M. Anderson) to construct solutions of Itô integral equations by solving an associated internal difference equation. In this paper we use the same general approach to find solutions  $y(t)$  of semimartingale integral equations of the form

$$y(t, \omega) = h(t, \omega) + \int_0^t f(s, \omega, y(\cdot, \omega)) dz(s),$$

where  $z$  is a given semimartingale,  $h$  is a right-continuous process and  $f(s, \omega, \cdot)$  is continuous on the space of right-continuous functions with left limits, with the topology of uniform convergence on compacts. In addition, we generalize Keisler's continuity theorem and give necessary and sufficient conditions for an internal martingale to be  $S$ -continuous.

**Introduction.** This paper is a continuation of Hoover and Perkins [5]. Notation and definitions presented in [5] are used without further introduction, and references to that paper are made by simply citing the number of the theorem or lemma (due to the consecutive numbering of the two papers, no ambiguity will arise). Finally, the reader is referred to §0 of [5] for a brief description of the contents of this work.

**8. Continuous local martingales.** We assume throughout this section that all processes are  $\mathbf{R}$  or  ${}^*\mathbf{R}$ -valued, since the main results (Theorems 8.1, 8.5 and 8.6) then follow immediately for higher dimensions. We will show (Theorem 8.1) that a continuous local martingale has a  $\mathcal{B}_t$ -local martingale lifting for any internal filtration  $\{\mathcal{B}_t\}$  (recall that, in general, Theorem 5.6 asserted only that such a lifting exists for some internal filtration). For this reason, unless otherwise stated, the internal filtration under consideration will be  $\{\mathcal{G}_t | t \in T\}$  (which was used in the original definition of  $\{\mathcal{F}_t\}$ ).

Note that if  $X: T \times \Omega \rightarrow {}^*\mathbf{M}$  is an  $S$ -continuous process (i.e.,  $X(\cdot, \omega)$  is  $S$ -continuous a.s.) then  $\text{st}(X)$  has continuous paths a.s. and  $X$  is a uniform lifting of  $\text{st}(X)$  in the sense of Keisler [6].

The following lifting theorem was first proved in a more specialized setting in Panetta [9]. The proof given here is different.

---

Received by the editors January 22, 1982. The results of this paper were presented at the second Victoria Symposium on Nonstandard Analysis, Victoria, Canada, June 16–20, 1980.

*AMS (MOS) subject classifications* (1970). Primary 60H20, 02H25; Secondary 60G45.

*Key words and phrases.* Stochastic differential equations, semimartingale, nonstandard analysis.

<sup>1</sup>The first author acknowledges the support of a NATO Postdoctoral Fellowship under the administration of the NSERC Canada.

©1983 American Mathematical Society  
0002-9947/81/0000-0106/\$06.00

**THEOREM 8.1.** *If  $x$  is a continuous local martingale,  $p > 0$  and  $x(0)$  is  $\sigma(\mathcal{Q}_0) \vee \mathcal{H}$ -measurable, then there is an  $S$ -continuous  $\mathcal{Q}_t$ -local martingale lifting of  $x$ ,  $X$ , such that  $|X|^p$  is locally  $S$ -integrable.*

**PROOF.** Choose  $p > 0$  and a sequence of stopping times  $\{U_n\}$  that reduces  $x$  and for which  $\sup_{t \leq U_n} |x(t)| \leq n$ . Let  $X_n$  be a lifting of  $x(U_n)$  such that  $|X_n| \leq n$  and define an  $S$ -martingale  $X^n$  by

$$X^n(t) = \bar{E}(X_n | \mathcal{Q}_t).$$

Then by Lemma 3.3 for each  $t$  in  $\text{ns}(T)$ ,

$$\begin{aligned} {}^\circ X^n(t) &= E({}^\circ X_n | \sigma(\mathcal{Q}_t)) \quad \text{a.s.} \\ &= E(E({}^\circ X_n | \mathcal{F}_{\circ t}) | \sigma(\mathcal{Q}_t)) \quad \text{a.s.} \\ &= E(x({}^\circ t \wedge U_n) | \sigma(\mathcal{Q}_t)) \quad \text{a.s.} \end{aligned}$$

If  ${}^\circ t > 0$ , then  $x({}^\circ t \wedge U_n)$  is  $\sigma(\mathcal{Q}_t) \vee \mathcal{H}$ -measurable by left-continuity, and if  ${}^\circ t = 0$ ,  $x({}^\circ t \wedge U_n) = x(0)$  is  $\sigma(\mathcal{Q}_0) \vee \mathcal{H}$ -measurable by assumption. Therefore,

$$(8.1) \quad {}^\circ X^n(t) = x({}^\circ t \wedge U_n) \quad \text{a.s.}$$

By Theorem 5.2,  $X^n$  is SD and hence Proposition 4.8 implies there is a positive infinitesimal  $\Delta_n t$  in  $T$  such that  $X^n \upharpoonright T_n \times \Omega$  is SDJ, where  $T_n = \{k\Delta_n t \mid k \in {}^*\mathbf{N}_0\}$  (since  ${}^\circ X^n(0) = {}^\circ X^n(\Delta_n t)$  a.s. by (8.1), we may include zero in  $T_n$ ). Since  $\text{st}(X^n)$  is a.s. continuous, it follows from Proposition 2.5 that  $X^n \upharpoonright T_n \times \Omega$  is  $S$ -continuous. We claim that, in fact,  $X^n$  is  $S$ -continuous on  $T \times \Omega$ . Let

$$Y(t) = |X^n(t) - X^n([t]^{T_n})| \vee \max\{|X^n(s + \Delta_n t) - X^n(s)| \mid s \in T_n, s + \Delta_n t \leq t\}.$$

It is easy to check that  $Y$  is an internal submartingale and therefore, by the maximal inequality for submartingales, if  $t_N \approx N$  ( $t_N \in T_n$ ), then

$$\bar{P}\left(\max_{\substack{t \leq t_N \\ t \in T}} Y(t) > N^{-1}\right) \leq N \bar{E}(Y(t_N)).$$

Since  $X^n \upharpoonright T_n \times \Omega$  is a.s.  $S$ -continuous and  $Y$  is uniformly bounded by  $2n$ ,  $\bar{E}(Y(t_N)) \approx 0$  and therefore  $Y(t) \approx 0$  for all  $t$  in  $\text{ns}(T)$ . It follows that  $X^n$  is  $S$ -continuous on  $T \times \Omega$ .

The required lifting of  $x$  is now obtained by a saturation argument. If  $m \leq n$  and  $\{V_i\}$  is a nondecreasing sequence of  $*$ -stopping times such that  ${}^\circ V_i = U_i$  a.s., then

$${}^\circ \bar{E}\left(\max_{t \leq V_m} |X^m(t) - X^n(t)|^p\right) = E\left(\sup_{t \leq U_m} |x(t \wedge U_m) - x(t \wedge U_n)|^p\right) = 0.$$

By saturation there is a  $\gamma$  in  ${}^*\mathbf{N} - \mathbf{N}$  such that  $X^\gamma$  is an  $\mathcal{Q}_t$ -martingale and

$$(8.2) \quad \max_{m \leq \gamma} \bar{E}\left(\max_{t \leq V_m} |X^m(t) - X^\gamma(t)|^p\right) \leq \gamma^{-1}.$$

The  $S$ -continuity of  $X^\gamma$  and the  $S$ -integrability of  $|X^\gamma(V_m \wedge t)|^p$  for all  $t \in T \cup \{\infty\}$  now follow from the corresponding properties of  $X^m$  and (8.2). Finally, for almost

all  $\omega$ , if  $\underline{t} \in \text{ns}(T)$ , then  $\underline{t} \leq V_m$  for large enough  $m$  in  $\mathbf{N}$  and, therefore,

$${}^\circ X^\gamma(\underline{t}) = {}^\circ X^m(\underline{t}) = x({}^\circ \underline{t} \wedge U_m) = x({}^\circ \underline{t}) \quad \text{a.s.}$$

Thus  $X^\gamma$  is the required lifting of  $x$ .  $\square$

We now turn to the problem of deciding when a given  $*$ -martingale is  $S$ -continuous. Establishing the  $S$ -continuity of a  $*$ -martingale is the key step in the construction of the solutions to the stochastic differential equations considered in Keisler [6]. Our main result states that under mild integrability conditions, a  $*$ -martingale  $X$  is  $S$ -continuous if and only if  $[X, X]$  is. The key ingredients are the square function inequalities of Burkholder, Davis, and Gundy (Theorem 1.3) which give us the following result.

**LEMMA 8.2.** *Suppose  $\Phi: {}^*[0, \infty) \rightarrow {}^*[0, \infty)$  is an internal, increasing, convex function and  $X(\underline{t})$  is an  $\mathcal{Q}_{\underline{t}}$ -martingale which satisfies the following conditions:*

(8.3) *If  ${}^\circ x > 0$ ,  ${}^\circ \Phi(x) > 0$ , and  $\Phi(0) = 0$ .*

(8.4) *There is a real constant  $k$  such that  $\Phi(2x) \leq k\Phi(x)$  for all  $x$  in  ${}^*[0, \infty)$ .*

(8.5) *There is a positive infinitesimal  $\varepsilon$  such that if  $\varepsilon < \underline{t}_2 - \underline{t}_1 \approx 0$  and  $\underline{t}_1, \underline{t}_2 \in \text{ns}(T)$ , then*

$$\bar{E}\left(\Phi\left([X, X]_{\underline{t}_2} - [X, X]_{\underline{t}_1}\right)^{1/2}\right)(\underline{t}_2 - \underline{t}_1)^{-1} \approx 0.$$

*Then  $X(\underline{t}) - X(0)$  is an  $S$ -continuous process.*

**PROOF.** For  $N$  in  $\mathbf{N}$  and  $\underline{\delta}$  in  $T$ , let

$$G_{N, \underline{\delta}} = \left\{ \omega \mid \max_{\substack{|\underline{t}_2 - \underline{t}_1| \leq \underline{\delta} \\ \underline{t}_1 \vee \underline{t}_2 \leq N}} |X(\underline{t}_2) - X(\underline{t}_1)| > 3N^{-1} \right\}.$$

Then

$$\begin{aligned} \bar{P}(G_{N, \underline{\delta}}) &\leq \bar{P}\left(\max_{0 \leq k \leq [N/\underline{\delta}]} \max_{k\underline{\delta} \leq \underline{s} \leq (k+1)\underline{\delta}} |X(\underline{s}) - X(k\underline{\delta})| > N^{-1}\right) \\ &\leq ([N/\underline{\delta}] + 1)\Phi(N^{-1})^{-1} \max_{k \leq [N/\underline{\delta}]} \bar{E}\left(\Phi\left(\max_{k\underline{\delta} \leq \underline{s} \leq (k+1)\underline{\delta}} |X(\underline{s}) - X(k\underline{\delta})|\right)\right), \end{aligned}$$

where  $[N/\underline{\delta}]$  is the integer part of  $N/\underline{\delta}$ . By the square function inequality for martingales (Theorem 1.3) there is a real constant  $c$  such that

$$\begin{aligned} \bar{P}(G_{N, \underline{\delta}}) &\leq ([N/\underline{\delta}] + 1)\Phi(N^{-1})^{-1} \\ &\quad \times \max_{k \leq [N/\underline{\delta}]} c\bar{E}\left(\Phi\left([X, X]_{(k+1)\underline{\delta}} - [X, X]_{k\underline{\delta}}\right)^{1/2}\right). \end{aligned}$$

By (8.3) and (8.5),  $\bar{P}(G_{N, \underline{\delta}}) \approx 0$  for all positive infinitesimals  $\underline{\delta}$  such that  $\underline{\delta} > \varepsilon$ . Therefore for each  $N$  in  $\mathbf{N}$  there is a  $\underline{\delta}_N$  in  $T$  such that  ${}^\circ \underline{\delta}_N > 0$  and  $\bar{P}(G_{N, \underline{\delta}_N}) < 2^{-N}$ . By the Borel-Cantelli lemma  $P(G_{N, \underline{\delta}_N} \text{ occurs infinitely often}) = 0$ , and since  $X(\cdot, \omega) - X(0, \omega)$  is  $S$ -continuous whenever  $\omega \in G_{N, \underline{\delta}_N}$  for only finitely many  $N$ , the result follows.  $\square$

**Notation 8.3.** Suppose  $X: T \times \Omega \rightarrow {}^*\mathbf{R}^d$  is an internal process and  $\{\mathcal{B}_{\underline{t}} \mid \underline{t} \in T\}$  is an internal filtration. Let

$$X^{\mathcal{B}_{\cdot}}(\underline{t}) = X(0) + \sum_{\underline{s} < \underline{t}} \bar{E}(\Delta X(\underline{s}) \mid \mathcal{B}_{\underline{s}}),$$

whenever the expression on the right is well defined. Let  $\nu^{\mathfrak{B}^\cdot}(X)$  denote the internal random measure on  $(T, \mathcal{C})$  defined by

$$\nu^{\mathfrak{B}^\cdot}(X)(\{\underline{t}\}) = [X, X]_{\underline{t} + \Delta \underline{t}}^{\mathfrak{B}^\cdot} - [X, X]_{\underline{t}}^{\mathfrak{B}^\cdot}.$$

If, in addition,  $X$  is  $\mathfrak{B}_\cdot$ -adapted and  $\varepsilon: T \times \Omega \rightarrow {}^*[0, \infty)$  is  $\mathfrak{B}_\cdot$ -adapted, define a  $\mathfrak{B}_\cdot$ -martingale  $X^{(\varepsilon, \mathfrak{B}^\cdot)}$  by

$$X^{(\varepsilon, \mathfrak{B}^\cdot)}(\underline{t}) = \sum_{\underline{s} < \underline{t}} \Delta X(\underline{s}) I_{\{\|\Delta X(\underline{s})\| \leq \varepsilon(\underline{s})\}} - \bar{E}(\Delta X(\underline{s}) I_{\{\|\Delta X(\underline{s})\| \leq \varepsilon(\underline{s})\}} | \mathfrak{B}_{\underline{s}}),$$

where for each  $\underline{s} \in T$ ,  $\bar{E}(\Delta X(\underline{s}) I_{\{\|\Delta X(\underline{s})\| \leq \varepsilon(\underline{s})\}} | \mathfrak{B}_{\underline{s}})$  denotes an internal version of the conditional expectation whose norm is bounded above by  $\varepsilon(\underline{s}, \omega)$  for all  $\omega$ . If there is no ambiguity we will write  $X^{(\varepsilon)}$  for  $X^{(\varepsilon, \mathfrak{B}^\cdot)}$  and  $\nu(X)$  for  $\nu^{\mathfrak{B}^\cdot}(X)$ .  $\square$

LEMMA 8.4. *Suppose that  $A$  is an  $\mathcal{Q}_\cdot$ -adapted process such that  $A(\cdot, \omega)$  is a nondecreasing  ${}^*\mathbf{R}$ -valued function for each  $\omega$ . Assume also that  $\sup_{\underline{t} \in \text{ns}(T)} {}^\circ \Delta A(\underline{t}) = 0$  a.s. and  $A(\underline{t}) - A(0)$  is locally  $S$ -integrable. Then*

$$P\left(\sup_{\underline{t} \in \text{ns}(T)} {}^\circ |A^{\mathcal{Q}^\cdot}(\underline{t}) - A(\underline{t})| > 0\right) = 0.$$

PROOF. We clearly may assume that  $A(0, \omega) = 0$  for all  $\omega$ . Suppose that  $A$  is uniformly bounded by a real constant. If  $Y(\underline{t}) = A(\underline{t}) - A^{\mathcal{Q}^\cdot}(\underline{t})$ , then  $Y$  is an internal martingale, and for  $\underline{t}$  in  $\text{ns}(T)$  we have

$$\begin{aligned} \bar{E}([Y, Y]_{\underline{t}}) &= \bar{E}\left(\sum_{\underline{s} < \underline{t}} (\Delta A(\underline{s}) - \bar{E}(\Delta A(\underline{s}) | \mathcal{Q}_{\underline{s}}))^2\right) \\ &= \bar{E}\left(\sum_{\underline{s} < \underline{t}} (\Delta A(\underline{s}))^2 - \bar{E}(\Delta A(\underline{s}) | \mathcal{Q}_{\underline{s}})^2\right) \\ &\leq \bar{E}\left(\sum_{\underline{s} < \underline{t}} (\Delta A(\underline{s}))^2\right) \leq \bar{E}\left(\left(\max_{\underline{s} < \underline{t}} \Delta A(\underline{s})\right) A(\underline{t})\right) \approx 0. \end{aligned}$$

Therefore if  $\varepsilon \in (0, \infty)$ ,

$$P\left(\max_{\underline{s} \leq \underline{t}} |Y(\underline{s})| > \varepsilon\right) \leq \varepsilon^{-2} \bar{E}(Y(\underline{t})^2) = \varepsilon^{-2} \bar{E}([Y, Y]_{\underline{t}}) = 0,$$

and we have shown that  $A(\underline{t}) \approx A^{\mathcal{Q}^\cdot}(\underline{t})$  for all  $\underline{t}$  in  $\text{ns}(T)$  a.s.

For the general case, let  $\{V_n\}$  be a sequence of  ${}^*$ -stopping times such that  $A(V_n)$  is  $S$ -integrable and  $\lim_{n \rightarrow \infty} {}^\circ V_n = \infty$  a.s. If  $A_n(\underline{t}) = A(\underline{t} \wedge V_n)$ , then  $A_n^{\mathcal{Q}^\cdot}(\underline{t}) = A^{\mathcal{Q}^\cdot}(\underline{t} \wedge V_n)$ , and therefore by replacing  $A$  with  $A_n$  we may assume that  $A(\underline{t})$  is  $S$ -integrable for all  $\underline{t}$  in  $T$ . Let  $A^n(\underline{t}) = A(\underline{t}) \wedge n$ . The above argument implies that

$$(8.6) \quad \bar{P}\left(\max_{\underline{t} \leq n} |A^n(\underline{t}) - (A^n)^{\mathcal{Q}^\cdot}(\underline{t})| > n^{-1}\right) < n^{-1}$$

for all  $n$  in  $\mathbf{N}$  and hence for some  $\gamma$  in  ${}^*\mathbf{N} - \mathbf{N}$ . If  $\underline{s} < \underline{t}$ , then

$$(8.7) \quad 0 \leq \Delta A(\underline{s}) - \Delta A^\gamma(\underline{s}) \leq \Delta A(\underline{s}) I_{\{A(\underline{t}) > \gamma\}},$$

and therefore if  $\underline{t} \in T$ ,

$$(8.8) \quad \bar{E}\left(\max_{\underline{s} \leq \underline{t}} (A^{\mathcal{Q}^\cdot}(\underline{s}) - (A^\gamma)^{\mathcal{Q}^\cdot}(\underline{s}))\right) = \bar{E}(A^{\mathcal{Q}^\cdot}(\underline{t}) - (A^\gamma)^{\mathcal{Q}^\cdot}(\underline{t})),$$

since the first inequality in (8.7) implies  $A(\underline{s}) - A^\gamma(\underline{s})$  is increasing, whence  $A^{\mathcal{Q}^\cdot}(\underline{s}) - (A^\gamma)^{\mathcal{Q}^\cdot}(\underline{s}) = (A - A^\gamma)^{\mathcal{Q}^\cdot}(\underline{s})$  is also. By the second inequality in (8.7), the right-hand side of (8.8) is bounded above by

$$E\left(\sum_{\underline{s} \leq \underline{t}} \bar{E}(\Delta A(\underline{s}) I_{\{A(\underline{t}) > \gamma\}} | \mathcal{Q}_{\underline{s}}^\cdot)\right) = \bar{E}(A(\underline{t}) I_{\{A(\underline{t}) > \gamma\}}) \approx 0$$

(the last since  $A(\underline{t})$  is  $S$ -integrable). Combining the above results, we get for almost all  $\omega$  and all  $\underline{t}$  in  $\text{ns}(T)$ ,

$$\begin{aligned} A(\underline{t}) &\approx A^\gamma(\underline{t}) \\ &\approx (A^\gamma)^{\mathcal{Q}^\cdot}(\underline{t}) \quad (\text{by (8.6) with } \gamma \text{ in place of } n) \\ &\approx A^{\mathcal{Q}^\cdot}(\underline{t}) \quad (\text{by (8.8) and the above bound}) \end{aligned}$$

and the result follows.  $\square$

We will need the following inequality of D. L. Burkholder: if  $(Z(\underline{t}), \mathcal{Q}_{\underline{t}})$  is an internal martingale,  $G(\underline{t}, \omega)$  is an  $\mathcal{Q}_{\underline{t}}$ -adapted process,  $Y = G \cdot Z$  and  $a, b > 0$ , then

$$(8.9) \quad \bar{P}(Y^*(\underline{t}) \geq a, [Y, Y]_{\underline{t}}^{\mathcal{Q}^\cdot} \leq b) \leq ba^{-2}.$$

This result may be proved by applying an elementary martingale inequality to the martingale  $Y(\underline{t} \wedge V)$ , where  $V = \min\{\underline{t} \mid [Y, Y]_{\underline{t} + \Delta \underline{t}}^{\mathcal{Q}^\cdot} > b\}$ .

We are finally ready for the main result of this section.

**THEOREM 8.5.** *Let  $X$  be an  $\mathcal{Q}_t$ -martingale.*

(a)  *$X$  is  $S$ -continuous and locally  $S$ -integrable if and only if  $[X, X]^{1/2}$  is  $S$ -continuous and locally  $S$ -integrable.*

(b) *If  $X^2$  is locally  $S$ -integrable and  $\sup_{\underline{t} \in \text{ns}(T)} \circ |\Delta X(\underline{t})| = 0$  a.s., then  $[X, X]_{\underline{t}} \approx [X, X]_{\underline{t}}^{\mathcal{Q}^\cdot}$  for all  $\underline{t}$  in  $\text{ns}(T)$  a.s.*

(c) *If  $[X, X]^{\mathcal{Q}^\cdot}$  is  $S$ -continuous and  $\sup_{\underline{t} \in \text{ns}(T)} \circ |\Delta X(\underline{t})| = 0$  a.s., then  $X$  is  $S$ -continuous. The converse holds if  $X^2$  is locally  $S$ -integrable.*

**PROOF.** (a) It follows from Lemma 6.3(b) and Theorem 6.4 that  $[X, X]^{1/2}$  is locally  $S$ -integrable if and only if  $X$  is. Also, Theorem 7.18 implies that if  $X$  is  $S$ -continuous and locally  $S$ -integrable, then  $[X, X]$  is  $S$ -continuous since it is SDJ and  $\Delta[X, X](\underline{t}) \approx 0$  for all  $\underline{t}$  in  $\text{ns}(T)$  a.s.

Suppose now that  $[X, X]^{1/2}$  is  $S$ -continuous and locally  $S$ -integrable. It remains only to show that  $X$  is  $S$ -continuous. Let  $\{W_n\}$  be the sequence of  $*$ -stopping times obtained in Theorem 6.4 (note that  $X$  is locally  $S$ -integrable by Lemma 6.3(b)).

Assume first that  $|\Delta X(\underline{t}, \omega)| \leq M$  for all  $(\underline{t}, \omega)$  and some  $M$  in  $\mathbb{N}$ . Let

$$\tau(\underline{t}) = \min\{\underline{u} \mid [X, X]_{\underline{u}} + \underline{u} \geq \underline{t}\}.$$

By the optional sampling theorem,  $X'(\underline{t}) = X(\tau(\underline{t}))$  is a  $*$ -martingale with respect to  $\{\mathcal{Q}_{\tau(\underline{t})} \mid \underline{t} \in T\}$ . If  $\underline{t}_1 < \underline{t}_2$ , then

$$\begin{aligned} (8.10) \quad [X', X']_{\underline{t}_2} - [X', X']_{\underline{t}_1} &= [X, X]_{\tau(\underline{t}_2)} - [X, X]_{\tau(\underline{t}_1)} \\ &\leq \underline{t}_2 - \underline{t}_1 + (\Delta X(\tau(\underline{t}_2)) - \Delta \underline{t})^2 + \Delta \underline{t} \\ &\leq 2(\underline{t}_2 - \underline{t}_1) + \max_{\underline{s} \leq \underline{t}_2} \Delta X(\underline{s})^2. \end{aligned}$$

Since  $[X, X]$  is  $S$ -continuous and  $\max_{s < t_2} \Delta X(s)^4 \leq M^4$ , we have

$$\bar{E} \left( \max_{s < N} \Delta X(s)^4 \right) < N^{-1}$$

for all  $N$  in  $\mathbf{N}$ , and therefore for some  $N_0$  in  ${}^*\mathbf{N} - \mathbf{N}$ . If  $\varepsilon = N_0^{-1/2} \approx 0$  and  $t_1, t_2 \in \text{ns}(T)$  satisfy  $\varepsilon < t_2 - t_1 \approx 0$ , then

$$\begin{aligned} & \bar{E} \left( ([X', X']_{t_2} - [X', X']_{t_1})^2 \right) (t_2 - t_1)^{-1} \\ & \leq 8(t_2 - t_1) + 2\bar{E} \left( \max_{s < N_0} \Delta X(s)^4 \right) \varepsilon^{-1} \quad (\text{by (8.10)}) \\ & \approx 0. \end{aligned}$$

Thus the conditions of Lemma 8.2 are satisfied with  $\Phi(x) = x^4$ , and therefore  $X'$  is  $S$ -continuous since  ${}^\circ X'(0) = {}^\circ X(0) = [X, X]_0^{1/2} \in \mathbf{R}$  a.s. Recall that for  $t$  in  ${}^*[0, \infty)$ ,  $[t] = \max\{t \in T \mid t \leq t\}$ . The  $S$ -continuity of  $X'$  and  $[X, X]$  implies the  $S$ -continuity of  $X'([X, X]_t + t, \omega)$ . However, by the definition of  $\tau$  and  $X'$  we have

$$X'([X, X]_t + t, \omega) = X(\tau([X, X]_t + t), \omega) = X(t, \omega),$$

whence the  $S$ -continuity of  $X$ .

Consider now the general case. Let

$$Z(t) = \sum_{s < t} |\Delta X(s)| I_{\{|\Delta X(s)| > 1\}}; \quad W'_n = \min\{t \mid Z(t) \geq n\} \quad (\min \emptyset = \infty).$$

Since  $Z(t) \leq [X, X]_t$ , clearly  $\lim_{n \rightarrow \infty} {}^\circ W'_n = \infty$  a.s., and hence if  $V_n = W'_n \wedge W''_n$ , then  $\lim_{n \rightarrow \infty} {}^\circ V_n = \infty$  a.s. Also

$$Z(V_n) \leq Z(W'_n - \Delta t) + 2X^*(W'_n) \leq n + 2X^*(W'_n),$$

which is  $S$ -integrable. Therefore

$$\begin{aligned} (8.11) \quad & {}^\circ \bar{E} \left( \sum_{s < V_n} |\bar{E}(\Delta X(s) I_{\{|\Delta X(s)| \leq 1\}} \mid \mathcal{Q}_s)| \right) \\ & = {}^\circ \bar{E} \left( \sum_s |\bar{E}(\Delta X(s) I_{\{s < V_n, |\Delta X(s)| > 1\}} \mid \mathcal{Q}_s)| \right) \\ & \leq {}^\circ \bar{E}(Z(V_n)) = E({}^\circ Z(V_n)). \end{aligned}$$

(The first equality holds since  $X$  is a martingale.) However, the  $S$ -continuity of  $[X, X]$  implies that  $Z(t) = 0$  for all  $t$  in  $\text{ns}(T)$  a.s. and so (8.11) equals zero. It now follows from the definition of  $X^{(1)} \equiv X^{(1, \mathcal{Q} \cdot)}$  (see Notation 8.3) that

$$[X^{(1)}, X^{(1)}]_t \approx \sum_{s < t} \Delta X(s)^2 I_{\{|\Delta X(s)| \leq 1\}} \approx [X, X]_t$$

for all  $t$  in  $\text{ns}(T)$  a.s. and, in particular,  $[X^{(1)}, X^{(1)}]$  is  $S$ -continuous. By the previous case  $X^{(1)}$  is  $S$ -continuous. Applying the fact that (8.11) equals zero, we get

$$X^{(1)}(t) \approx \sum_{s < t} \Delta X(s) I_{\{|\Delta X(s)| \leq 1\}} \approx X(t)$$

(the last since  $[X, X]$  is  $S$ -continuous) for all  $t$  in  $\text{ns}(T)$  a.s. Therefore the  $S$ -continuity of  $X$  follows from the  $S$ -continuity of  $X^{(1)}$ .

(b) Suppose that  $X^2$  is locally  $S$ -integrable and  $\sup_{s \in \text{ns}(T)} {}^\circ |\Delta X(s)| = 0$  a.s. By Lemma 6.3(b),  $[X, X]$  is locally  $S$ -integrable. Thus the hypotheses of Lemma 8.4 are satisfied with  $A = [X, X]$ , and therefore  $[X, X]_t \approx [X, X]_t^{\mathcal{Q}}$  for all  $t$  in  $\text{ns}(T)$  a.s.

(c) Suppose  $[X, X]^{\mathcal{Q}}$  is  $S$ -continuous and  $\sup_{t \in \text{ns}(T)} {}^\circ |\Delta X(t)| = 0$  a.s. Note that  $[X^{(1)}, X^{(1)}]^{\mathcal{Q}}$  is  $S$ -continuous because  $[X, X]^{\mathcal{Q}}$  is, and  $\Delta[X^{(1)}, X^{(1)}]^{\mathcal{Q}}(t) \leq \Delta[X, X]^{\mathcal{Q}}(t)$ . Since  $\bar{E}(\Delta X(s) | \mathcal{Q}_s) = 0$ , we have

$$(8.12) \quad \left| \bar{E}(\Delta X(s) I_{\{|\Delta X(s)| \leq 1\}} | \mathcal{Q}_s) \right| = \left| \bar{E}(\Delta X(s) I_{\{|\Delta X(s)| > 1\}} | \mathcal{Q}_s) \right| \\ \leq \Delta[X, X]^{\mathcal{Q}}(s)$$

for all  $s \in T$ ,  $\bar{P}$ -a.s. This, together with the fact that  $\Delta X(t) \approx 0$  for all  $t$  in  $\text{ns}(T)$  a.s., implies that  $\Delta X^{(1)}(t) \approx 0$  for all  $t$  in  $\text{ns}(T)$  a.s. If  $W_n = \min\{t \mid |X^{(1)}(t)| \geq n\}$ , then  $|X^{(1)}(W_n)| \leq n + 2$ . The finiteness of  $[X^{(1)}, X^{(1)}]^{\mathcal{Q}}$  and (8.9) with  $Y = X^{(1)}$  imply that  $\lim_{n \rightarrow \infty} {}^\circ W_n = \infty$  a.s. Therefore  $(X^{(1)})^2$  is locally  $S$ -integrable and the  $S$ -continuity of  $[X^{(1)}, X^{(1)}]$  follows from (b) and the  $S$ -continuity of  $[X^{(1)}, X^{(1)}]^{\mathcal{Q}}$ . By (a),  $X^{(1)}$  is  $S$ -continuous, and therefore by (8.12) and the  $S$ -continuity of  $[X, X]^{\mathcal{Q}}$ , the same is true of  $\sum_{s < t} \Delta X(s) I_{\{|\Delta X(s)| \leq 1\}}$ . Since  $\Delta X(t) \approx 0$  for all  $t$  in  $\text{ns}(T)$  a.s. and  ${}^\circ X(0) = ({}^\circ [X, X]_0^{\mathcal{Q}})^{1/2} \in \mathbf{R}$  a.s., this implies the  $S$ -continuity of  $X$ .

The converse is an immediate consequence of (a) and (b).  $\square$

In practice (c) may give a more useful criterion than (a) for checking whether or not an  $\mathcal{Q}_t$ -martingale  $X$  is  $S$ -continuous. This is because  $[X, X]^{\mathcal{Q}}$  is sometimes easier to compute than  $[X, X]$ . Consider, for example, the case when  $\{Z_t \mid t \in T\}$  is a \*-independent sequence of internal, identically distributed random variables with mean zero and variance  $\sigma^2$  ( ${}^\circ \sigma^2 < \infty$ ), and  $X(t) = \sum_{0 < s \leq t} Z_s (\Delta t)^{1/2}$ . Then  $[X, X]_t = \sum_{0 < s \leq t} Z_s^2 \Delta t$ , but  $[X, X]_t^{\mathcal{Q}} = \sigma^2 t$ .

Part (c) of the above result is a generalization of the “continuity theorem” for \*-martingales in Keisler [6, Theorem 3.2]. That result implies that, on a particular type of internal probability space,  $X$  is  $S$ -continuous if  $\Delta X(t, \omega) \approx 0$  for all  $(t, \omega)$  in  $\text{ns}(T) \times \Omega$  and  $[X, X]^{\mathcal{Q}}$  is Lipschitz continuous on bounded subsets of  $T$ , with a Lipschitz constant independent of  $\omega$ . (Keisler also gives an explicit modulus of continuity for  $\text{st}(X)$ .)

Part (b) and the converse of (c) are false in general without the hypothesis that  $X^2$  is locally  $S$ -integrable. If  $X$  is constant except for a single jump at time 0 of size  $-(\Delta t)^{-1/2}$  with probability  $\Delta t/(1 + \Delta t)$  and of size  $(\Delta t)^{1/2}$  with probability  $1/(1 + \Delta t)$ , then  $X(t) \approx [X, X]_t \approx 0$  for all  $t$  a.s. but  $[X, X]_{\Delta t}^{\mathcal{Q}} = 1$  and  $[X, X]_0^{\mathcal{Q}} = 0$ .

Part (a) has been proved independently by T. L. Lindström [7, Theorem 14; 8, Theorem 2] under the slightly stronger assumption that  ${}^\circ \bar{E}(X^2(V_n)) < \infty$  for some sequence of \*-stopping times such that  $\lim_{n \rightarrow \infty} {}^\circ V_n = \infty$  a.s. The proof given here seems somewhat shorter.

It is not true in general that  $X$  is  $S$ -continuous if  $[X, X]^{1/2}$  is, so that (a) would be false if the local  $S$ -integrability condition is dropped. Consider the following example:

Fix  $\gamma \in {}^*\mathbf{N} - \mathbf{N}$  such that  $\gamma \Delta t \approx 0$ . Let  $\{Y_i \mid i = 1, \dots, \gamma\}$  be internal \*-independent random variables such that  $\bar{P}(Y_i = \gamma) = \gamma^{-2}$  and  $\bar{P}(Y_i = 0) = 1 - \gamma^{-2}$ . Since  $\bar{E}(Y_i) = \gamma^{-1}$ ,  $Y(j\Delta t) = (\sum_{i \leq j \wedge \gamma} Y_i) - (j \wedge \gamma)\gamma^{-1}$  is an internal martingale. Note that

$\bar{P}(Y_i = 0 \text{ for } i = 1, \dots, \gamma) \approx 1$ , so that  $Y(j\Delta t) = -(j \wedge \gamma)\gamma^{-1}$  and  $[Y, Y]_{j\Delta t} = (j \wedge \gamma)\gamma^{-2} \approx 0$  for all  $j\Delta t \in \text{ns}(T)$  a.s. Therefore  $[Y, Y]$  is  $S$ -continuous but  $Y$  is not.

The following result is an easy application of the above “continuity theorem”.

**THEOREM 8.6.** *Let  $X$  be an  $S$ -continuous  $S$ -local martingale and let  $G(\underline{t}, \omega)$  be an  $\mathcal{Q}_t$ -adapted process. If either  $G \in L(0; X, \mathbf{R}, \mathcal{Q}_\cdot)$  or (8.13) for each  $m$  in  $\mathbf{N}$ ,  $I_{\{\cdot \leq m\}}G^2(\cdot, \omega)$  is  $S$ -integrable with respect to  $\nu(X)$  a.s., then  $G \cdot X$  is  $S$ -continuous.*

**PROOF.** Assume first that  $|G(\underline{s}, \omega)|$  is uniformly bounded by  $c \in \mathbf{R}$ . By Theorem 8.5(a),  $[X, X]^{1/2}$  is  $S$ -continuous and locally  $S$ -integrable and hence so is  $[G \cdot X, G \cdot X]_t^{1/2} = (\int_{I_{\{\cdot \leq t\}}} G^2(s) d\lambda^{(2)}(X))^{1/2}$ . Another application of Theorem 8.5 shows that  $G \cdot X$  is  $S$ -continuous.

For the more general cases, let  $G_n = (G \wedge n) \vee (-n)$ . If  $G \in L(0; X, \mathbf{R}, \mathcal{Q}_\cdot)$ , the  $S$ -continuity of  $G \cdot X$  follows from Lemma 7.13(c) and the  $S$ -continuity of each  $G_n \cdot X$ . If (8.13) holds, then for each  $m, n$  in  $\mathbf{N}$ ,

$$\begin{aligned}
 (8.14) \quad & \circ \bar{P} \left( \max_{\underline{t} \leq m} |G \cdot X(\underline{t}) - G_n \cdot X(\underline{t})| \geq m^{-1} \right) \\
 & \leq m^{-2} + \circ \bar{P} \left( \int_{I_{\{\cdot \leq m\}}} (G(\underline{t}) - G_n(\underline{t}))^2 d\nu(X) > m^{-4} \right) \\
 & \quad \text{(by essentially the same argument as for (8.9))} \\
 & \leq m^{-2} + P \left( \int_{I_{\{\cdot \leq m\}}} (\circ G(\underline{t}) - \circ G_n(\underline{t}))^2 dL(\nu(X)) \geq m^{-4} \right) \\
 & \quad \text{(by (8.13)).}
 \end{aligned}$$

By dominated convergence there is a subsequence  $\{n_m\}$  such that (8.14) is less than  $2m^{-2}$  when  $n = n_m$ . The Borel-Cantelli lemma implies that

$$\lim_{m \rightarrow \infty} \circ \max_{\underline{t} \leq m} |G \cdot X(\underline{t}) - G_{n_m} \cdot X(\underline{t})| = 0 \quad \text{a.s.}$$

This, together with the  $S$ -continuity of  $G_{n_m} \cdot X$ , establishes the  $S$ -continuity of  $G \cdot X$ .  $\square$

Note that this result does not follow from Lemma 7.13(e) since  $G$  need not be a  $(0; X)$ -lifting of a standard process.

**REMARK 8.7.** Note that the  $d$ -dimensional versions of Theorems 8.1, 8.5 and 8.6 follow as immediate corollaries to those results.  $\square$

**9. A lifting theorem for semimartingales.** We have shown (Theorem 8.6) that if  $X$  is an  $S$ -continuous,  $S$ -local martingale with respect to  $\{\mathcal{Q}_t\}$  and  $G \in L(0; X, \mathbf{R}, \mathcal{Q}_\cdot)$ , then  $G \cdot X$  is  $S$ -continuous. This followed from a continuity theorem (Theorem 8.5) which said that, under mild integrability conditions,  $X$  is  $S$ -continuous if and only if  $[X, X]$  is. It is natural to ask if these results still hold if one replaces “ $S$ -continuous” by “SDJ” (i.e., the notion of being near-standard in  $C$  is replaced by the corresponding notion in  $D$ ). In fact, an affirmative answer to this question would allow one to obtain solutions of a semimartingale stochastic differential equation in the same way



as Keisler used his continuity theorem to construct solutions to Itô integral equations. Unfortunately both results fail as the following example shows. (A similar example was found independently by T. L. Lindström.)

EXAMPLE 9.1. Let  $V$  be an internal geometric random variable on  $(\Omega, \mathcal{A}, \bar{P})$  such that if  $p = (1 + (\Delta t)^{1/2})^{-1}$ , then  $\bar{P}(V = k) = (1 - p)p^k$  for  $k \in {}^*\mathbf{N}_0$ . Assume that  $1 \in T$  and define  $Y(\underline{t}, \omega)$  by

$$Y(\underline{t}) = 0 \quad \text{for } \underline{t} \leq 1,$$

$$\Delta Y(1 + j\Delta t) = \begin{cases} (-1)^{j+1} \Delta t^{1/2} & \text{if } 0 \leq j < V, \\ (-1)^j & \text{if } j = V, \\ 0 & \text{if } j > V. \end{cases}$$

Then  $Y$  is a uniformly bounded SDJ  $\mathcal{A}_t$ -martingale, where  $\mathcal{A}_t$  is the internal  $\sigma$ -algebra generated by  $\{Y(\underline{s}) \mid \underline{s} \leq \underline{t}\}$ . Note that  $Y(\underline{t})$  is zero before  $\underline{t} = 1$ , oscillates between 0 and  $-(\Delta t)^{1/2}$  for  $\underline{t} \leq 1 + V\Delta t$  and at  $1 + V\Delta t$  has a jump of  $\pm 1$ . If  $G(1 + j\Delta t) = (-1)^j$ , then

$$G \cdot Y(1 + j\Delta t) = \begin{cases} -j(\Delta t)^{1/2} & \text{if } 0 \leq j \leq V, \\ 1 - V(\Delta t)^{1/2} & \text{if } j > V. \end{cases}$$

An easy computation shows that  $V\Delta t \approx 0$  a.s. but  ${}^\circ(V(\Delta t)^{1/2}) > 0$  a.s. Therefore  $Z = G \cdot Y$  is not SDJ. Note also that  $[Z, Z] = [Y, Y]$  is SDJ.  $\square$

In the previous example, it is clear that if one restricts  $Y$  to a coarser time set, then it would a.s. consist of a single jump of  $\pm 1$  that occurs infinitesimally close to one, and the internal integral  $G \cdot Y$  would then of course be SDJ. Our goal in this section is to show that this is always the case. More precisely, Theorem 9.7 states that one can always find an internal filtration  $\{\mathcal{B}_t\}$  and a  $\mathcal{B}_t$ -semimartingale lifting,  $(A; X)$ , of  $(a; x)$  ( $(a, x) \in {}^*\mathbb{C}_0^d \times {}^*\mathbb{C}_0^d$ ) such that  $G \cdot (A + X)$  is SDJ for all  $G \in L(A; X, \mathbf{R}^{n \times d}, \mathcal{B}_t)$ .

Notation 9.2. Suppose  $Y: T \times \Omega \rightarrow {}^*\mathbf{R}^d$  is an internal stochastic process.

(a) If  $\delta, t \in {}^*[0, \infty]$ , let

$$w(Y, t, \delta) = \sup\{\|Y(\underline{s}) - Y(\underline{t})\| \mid \underline{s}, \underline{t} \in T; \underline{s}, \underline{t} \leq t; |\underline{s} - \underline{t}| \leq \delta\}$$

and

$$w_0(Y, t) = \sup\{{}^\circ\|Y(\underline{s}) - Y(\underline{t})\| \mid \underline{s}, \underline{t} \in T; \underline{s}, \underline{t} \leq t; \underline{s} \approx \underline{t}\}.$$

The same notation is used if  $Y: T' \times \Omega \rightarrow {}^*\mathbf{R}^d$  for some other internal  $S$ -dense set  $T'$ .

(b) If  $T' = \{t_i \mid i \in {}^*\mathbf{N}\}$  is an internal subset of  $T$  and  $t_i < t_{i+1}$ , let

$$\text{Var}(Y, T')(t_j) = \sum_{i=1}^{j-1} \|Y(t_{i+1}) - Y(t_i)\|.$$

(c) If  $\varepsilon: T \times \Omega \rightarrow {}^*[0, \infty)$  is an internal process let  $\|\varepsilon\| = \sup_{(t, \omega)} \varepsilon(t, \omega)$  and

$$J(X, \varepsilon)(\underline{t}) = \sum_{\underline{s} < \underline{t}} \Delta X(\underline{s}) I_{\{\|\Delta X(\underline{s})\| > \varepsilon(\underline{s})\}}.$$

(d) If  $\{x_n\}$  is a sequence of random variables converging in probability to  $x_\infty$ , we write  $x_n \xrightarrow{P} x_\infty$ .  $\square$

Note that if  $X$  is a  $\mathfrak{B}_t$ -martingale and  $\epsilon$  is a  $^*[0, \infty)$ -valued  $\mathfrak{B}_t$ -adapted process for some internal filtration  $\{\mathfrak{B}_t\}$ , then

$$(9.1) \quad X(\underline{t}) = X^{(\epsilon, \mathfrak{B}^\cdot)}(\underline{t}) + J(X, \epsilon) - J(X, \epsilon)^{\mathfrak{B}^\cdot}(\underline{t}) \quad \text{for all } \underline{t} \text{ in } T, \bar{P}\text{-a.s.}$$

LEMMA 9.3. *Assume  $\{\mathfrak{B}_t\}$  is an internal filtration and  $X$  is a  $^*\mathbf{R}^d$ -valued  $\mathfrak{B}_t$ -martingale such that  $[X, X]$  is SDJ. If  $\epsilon: T \times \Omega \rightarrow ^*[0, \infty)$  is  $\mathfrak{B}_t$ -adapted and satisfies  $^\circ\|\epsilon\| < \infty$  and  $\sup_{\underline{t} \in \text{ns}(T)} ^\circ\epsilon(\underline{t}, \omega) = 0$  a.s., then  $X^{(\epsilon, \mathfrak{B}^\cdot)}$  is  $S$ -continuous.*

PROOF. Since  $[X, X]$  is SDJ,

$$A_\epsilon(\underline{t}) = \sum_{\underline{s} \leq \underline{t}} \|\Delta X(\underline{s})\|^2 I_{\{\|\Delta X(\underline{s})\| \leq \epsilon(\underline{s})\}}$$

is  $S$ -continuous. Also  $A_\epsilon$  is locally  $S$ -integrable. Therefore, by Lemma 8.4,  $A_\epsilon^{\mathfrak{B}^\cdot}$  is  $S$ -continuous. Note that

$$\begin{aligned} \Delta[X^{(\epsilon)}, X^{(\epsilon)}]^{\mathfrak{B}^\cdot}(\underline{t}) &= \bar{E}\left(\|\Delta X(\underline{t})\|^2 I_{\{\|\Delta X(\underline{t})\| \leq \epsilon(\underline{t})\}} \mid \mathfrak{B}_{\underline{t}}\right) \\ &\quad - \|\bar{E}\left(\Delta X(\underline{t}) I_{\{\|\Delta X(\underline{t})\| \leq \epsilon(\underline{t})\}} \mid \mathfrak{B}_{\underline{t}}\right)\|^2 \\ &\leq \Delta A_\epsilon^{\mathfrak{B}^\cdot}(\underline{t}), \end{aligned}$$

and therefore  $[X^{(\epsilon)}, X^{(\epsilon)}]^{\mathfrak{B}^\cdot}$  is  $S$ -continuous. Moreover by the definition of  $X^{(\epsilon)}$  we have

$$\sup_{\underline{t} \in \text{ns}(T)} ^\circ\|\Delta X^{(\epsilon)}(\underline{t})\| \leq \sup_{\underline{t} \in \text{ns}(T)} ^\circ 2\epsilon(\underline{t}, \omega) = 0 \quad \text{a.s.}$$

By Theorem 8.5(c) and Remark 8.7 we may conclude that  $X^{(\epsilon)}$  is  $S$ -continuous.  $\square$

We have shown (Theorem 7.18) that if  $X$  is an SDJ  $S$ -local martingale then  $[X, X]$  is SDJ, but that the converse is false (Example 9.1). The following theorem gives conditions which are necessary and sufficient for an  $S$ -local martingale to be SDJ.

THEOREM 9.4. *Let  $\{\mathfrak{B}_t \mid t \in T\}$  be an internal filtration and let  $X$  be a  $^*\mathbf{R}^d$ -valued  $S$ -local martingale with respect to  $\{\mathfrak{B}_t\}$ . Suppose  $\{\epsilon_n(\underline{t}, \omega)\}$  is a sequence of  $\mathfrak{B}_t$ -adapted processes taking values in  $^*[0, \infty)$  such that  $\lim_{n \rightarrow \infty} ^\circ\|\epsilon_n\| = 0$  and for each  $n \in \mathbf{N}$ ,*

$$(9.2) \quad ^\circ \inf_{(\underline{t}, \omega)} \epsilon_n(\underline{t}, \omega) > 0.$$

*Then  $X$  is SDJ if and only if  $[X, X]$  is SDJ and  $w_0(J(X, \epsilon_n)^{\mathfrak{B}^\cdot}, t) \xrightarrow{P} 0$  as  $n$  approaches  $\infty$  for all  $t \in [0, \infty)$ .*

PROOF. Suppose  $X$  is SDJ. Then  $[X, X]$  is SDJ by Theorem 7.18. By saturation we may extend  $\{\epsilon_n\}$  internally to  $^*\mathbf{N}$  so that  $\epsilon_n$  is  $\mathfrak{B}_t$ -adapted for all  $n \in ^*\mathbf{N}$  and  $\|\epsilon_H\| \approx 0$  for all  $H \in ^*\mathbf{N} - \mathbf{N}$ . By Lemma 9.3, if  $N \in \mathbf{N}$ , then

$$(9.3) \quad \bar{P}\left(w(X^{(\epsilon_H)}, N, \|\epsilon_H\|) > 2^{-N}\right) < 2^{-N}$$

for all  $H$  in  ${}^*\mathbf{N} - \mathbf{N}$ , and hence (by the permanence principle), for all  $H \geq M_N$  ( $H \in {}^*\mathbf{N}$ ), for some  $M_N \in \mathbf{N}$ . By (9.1) we have for all  $N$  in  $\mathbf{N}$ ,

$$\begin{aligned} w_0(J(X, \varepsilon_n)^{\otimes \cdot}, N) &\leq w_0(X^{(\varepsilon_n)}, N) + w_0(X - J(X, \varepsilon_n), N) \\ &\leq {}^\circ w(X^{(\varepsilon_n)}, N, \|\varepsilon_n\|) + {}^\circ \|\varepsilon_n\| \quad \text{a.s.,} \end{aligned}$$

where we have used the fact that  $X$  is SDJ in the last line. Therefore if  $N \in \mathbf{N}$  and  $n \geq M_N$ ,

$$\begin{aligned} P(w_0(J(X, \varepsilon_n)^{\otimes \cdot}, N) > 2^{-N} + {}^\circ \|\varepsilon_n\|) &\leq {}^\circ \bar{P}(w(X^{(\varepsilon_n)}, N, \|\varepsilon_n\|) > 2^{-N}) \\ &\leq 2^{-N} \quad (\text{by (9.3)}), \end{aligned}$$

as required.

For the converse, we may assume that  $\lim_{n \rightarrow \infty} w_0(J(X, \varepsilon_n)^{\otimes \cdot}, n) = 0$  a.s. by considering a subsequence, if necessary. Since  $[X, X]$  is SDJ by assumption, Lemma 9.3 implies that (9.3) holds for  $N \in \mathbf{N}$  and  $H \in {}^*\mathbf{N} - \mathbf{N}$ , and hence by considering a further subsequence we may also assume that  $\lim_{n \rightarrow \infty} {}^\circ w(X^{(\varepsilon_n)}, n, \|\varepsilon_n\|) = 0$  a.s. Fix  $\omega$  such that both of the above limits are zero, (9.1) holds and  $[X, X](\omega)$  is SDJ. If  $\underline{t}_1 \approx \underline{t}_2$  are in  $\text{ns}(T)$  and  $\underline{t}_1 < \underline{t}_2$ , then by (9.1),

$$\begin{aligned} (9.4) \quad {}^\circ \|X(\underline{t}_1) - X(\underline{t}_2)\| &\leq \limsup_{n \rightarrow \infty} ({}^\circ \|J(X, \varepsilon_n)(\underline{t}_1) - J(X, \varepsilon_n)(\underline{t}_2)\| \\ &\quad + w_0(J(X, \varepsilon_n)^{\otimes \cdot}, n) + {}^\circ w(X^{(\varepsilon_n)}, n, \|\varepsilon_n\|)) \\ &= \limsup_{n \rightarrow \infty} {}^\circ \|J(X, \varepsilon_n)(\underline{t}_1) - J(X, \varepsilon_n)(\underline{t}_2)\|. \end{aligned}$$

If  $\Delta X(\underline{t}) \approx 0$  for all  $\underline{t} \approx \underline{t}_1$ , the above result and (9.2) imply that  $X(\underline{t}_1) \approx X(\underline{t}_2)$ . In particular, the  $S$ -continuity of  $[X, X]$  at zero implies the same is true of  $X$ . If  ${}^\circ \|\Delta X(\underline{t})\| > 0$  for some  $\underline{t} \approx \underline{t}_1$  (there is at most one such  $\underline{t}$  since  $[X, X]$  is SDJ) then by (9.2) and by (9.4),  ${}^\circ X(\underline{t}_1) \neq {}^\circ X(\underline{t}_2)$  if and only if  $\underline{t}_1 \leq \underline{t} < \underline{t}_2$ . Since  $X$  is an  $S$ -local martingale, it is SD by Theorem 5.2, and hence is SDJ by the above.  $\square$

Recall that (Notation 6.1)  $|Y|(\underline{t}, \omega) = \sum_{\underline{s} < \underline{t}} \|\Delta Y(\underline{s}, \omega)\|$ .

**LEMMA 9.5.** *Suppose  $Y$  is a  ${}^*\mathbf{R}^d$ -valued  $\mathbb{B}_t$ -BV lifting of  $y$  for some internal filtration  $\{\mathbb{B}_{\underline{t}} | \underline{t} \in T\}$ . Then for a.a.  $\omega$  and all  $t \in [0, \infty)$ ,  $w_0(|Y|, t) = w_0(Y, t)$ .*

**PROOF.** Choose  $\omega$  such that  $|Y|(\cdot, \omega)$  is an SDJ lifting of  $|y|(\cdot, \omega)$ . Suppose  $\underline{s} \approx \underline{t} \in \text{ns}(T)$  satisfy  $\underline{s} < \underline{t}$  and  ${}^\circ(|Y|(\underline{t}) - |Y|(\underline{s})) > 0$ . Then since  $|Y|(\cdot, \omega)$  is SDJ, there is a  $\underline{u}$  such that  $\underline{s} \leq \underline{u} < \underline{t}$  and  ${}^\circ(|Y|(\underline{t}) - |Y|(\underline{s})) = {}^\circ \|\Delta Y(\underline{u})\|$  (see Proposition 2.5). It follows that  $w_0(|Y|, t) \leq w_0(Y, t)$  for all  $t \in [0, \infty]$  a.s.

The converse inequality is obvious and the result is proved.  $\square$

Recall that our aim is to show the existence of a semimartingale lifting  $(A; X)$  of  $(a; x) \in \mathcal{V}_0^d \times \mathcal{L}_0^d$  such that  $G \cdot (A + X)$  is SDJ for all  $G \in L(A; X, \mathbf{R}^{n \times d}, \mathbb{B}_\cdot)$ . For this we would like a condition on  $X$  that guarantees that the conditions described in Theorem 9.4 are satisfied by  $G \cdot X$ . Such a condition is obtained in the following lemma.

LEMMA 9.6. Assume  $X$  is a  ${}^*\mathbf{R}^d$ -valued SDJ  $S$ -local martingale with respect to  $\{\mathfrak{B}_t\}$  and  $\{\varepsilon_n\}$  satisfies the same hypotheses as in Theorem 9.4. There is an  $S$ -dense subset of  $T, T' = \{k\Delta't \mid k \in {}^*\mathbf{N}_0\}$ , such that for each  $t \in [0, \infty)$ ,

$$w_0\left(\text{Var}\left(J(X, \varepsilon_n)^{\mathfrak{B}_\cdot}, T'\right), t\right) \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty.$$

PROOF. Let  $\{W_m\}$  be the sequence of  $*$ -stopping times introduced in Theorem 6.4 and, for  $n \in \mathbf{N}$ , define

$$Y_n(t) = \sum_{s \leq t} \|\Delta X(s)\| I_{\{\|\Delta X(s)\| > \varepsilon_n(s)\}}.$$

Since  $X$  is SDJ, (9.2) implies that  $Y_n$  is SDJ, and therefore if  $V_m^n = \min\{t \mid Y_n(t) \geq m\}$ , then  $\lim_{m \rightarrow \infty} {}^\circ V_m^n = \infty$  a.s. for each  $n$ . Note that  $Y_n(V_m^n \wedge W_m)$  is  $S$ -integrable because  $Y_n(V_m^n \wedge W_m) \leq m + X^*(W_m)$ . Therefore

$${}^\circ \bar{E}\left(Y_n^{\mathfrak{B}_\cdot}(V_m^n \wedge W_m)\right) = {}^\circ \bar{E}\left(Y_n(V_m^n \wedge W_m)\right) < \infty$$

and, in particular, since  $Y_n^{\mathfrak{B}_\cdot}$  is nondecreasing ( $\bar{P}$ -a.s.),  $Y_n^{\mathfrak{B}_\cdot}$  is SD by Remark 2.4 and the fact that  $\lim_{m \rightarrow \infty} {}^\circ(V_m^n \wedge W_m) = \infty$  a.s. Note also that  $Y_n^{\mathfrak{B}_\cdot}(t) \approx 0$  for all  $t \approx 0$  since if  $t \approx 0$ , then

$${}^\circ \bar{E}\left(Y_n^{\mathfrak{B}_\cdot}(t \wedge V_m^n \wedge W_m)\right) = {}^\circ \bar{E}\left(Y_n(t \wedge V_m^n \wedge W_m)\right) = E({}^\circ Y_n(t \wedge V_m^n \wedge W_m)) = 0.$$

It follows that  $J(X, \varepsilon_n)^{\mathfrak{B}_\cdot}$  is also SD and  $j_n = \text{st}(J(X, \varepsilon_n)^{\mathfrak{B}_\cdot}) \in \mathcal{V}_0^d$ . Lemma 7.5 implies there is a decreasing sequence of  $S$ -dense subsets of  $T, \{T_i\}$ , such that  $T_i = \{k\Delta_i t \mid k \in {}^*\mathbf{N}_0\}$  and for all  $n \leq i$ ,  $J(X, \varepsilon_n)^{\mathfrak{B}_\cdot}([\cdot]^{T_i})$  is a  $\mathfrak{B}_{[t]}T_i$ -BV lifting of  $j_n$ . By saturation there is an  $S$ -dense subset of  $T, T' = \{k\Delta't \mid k \in {}^*\mathbf{N}_0\}$ , such that  $J(X, \varepsilon_n)^{\mathfrak{B}_\cdot}([\cdot]^{T'})$  is a  $\mathfrak{B}_{[t]}T'$ -BV lifting of  $j_n$  for all  $n \in \mathbf{N}$ . By Lemma 9.5, for a.a.  $\omega$  and all  $t$  in  $[0, \infty]$ , we have

$$\begin{aligned} w_0\left(\text{Var}\left(J(X, \varepsilon_n)^{\mathfrak{B}_\cdot}, T'\right), t\right) &= w_0\left(\left|J(X, \varepsilon_n)^{\mathfrak{B}_\cdot}([\cdot]^{T'})\right|, t\right) \\ &= w_0\left(J(X, \varepsilon_n)^{\mathfrak{B}_\cdot}([\cdot]^{T'}), t\right) \leq w_0\left(J(X, \varepsilon_n)^{\mathfrak{B}_\cdot}, t\right). \end{aligned}$$

Since the right side of (9.5) converges in probability to zero as  $n$  approaches  $\infty$  (Theorem 9.4), the result follows.  $\square$

We are ready to prove the main result of this section.

THEOREM 9.7. Let  $(a, x) \in \mathcal{V}_0^d \times \mathcal{L}_0^d$  and let  $h: [0, \infty) \times \Omega \rightarrow \mathbf{R}^m$  be  $\mathcal{F}_t$ -adapted with sample paths in  $D$ . There is an internal filtration  $\{\mathfrak{B}_t \mid t \in T\}$ , a  $\mathfrak{B}_t$ -semi-martingale lifting of  $(a; x), (A; X)$ , and a  $\mathfrak{B}_t$ -adapted SDJ lifting of  $h, H$ , such that if  $Z = A + X$ , then  $(H, G \cdot Z)$  is SDJ for all  $G \in L(A; X, \mathbf{R}^{n \times d}, \mathfrak{B}_\cdot)$ .

PROOF. By Theorem 5.6 and Remark 5.7(b) there is an internal filtration  $\{\mathfrak{B}'_t\}$ , a  $\mathfrak{B}'_t$ -local martingale lifting of  $x, X'$ , and a  $\mathfrak{B}'_t$ -adapted SDJ lifting of  $(h, a), (H', A')$ . Lemma 9.6 implies there is an  $S$ -dense subset of  $T, T' = \{k\Delta't \mid k \in {}^*\mathbf{N}_0\}$ , such that

$$(9.6) \quad w_0\left(\text{Var}\left(J(X', n^{-1})^{\mathfrak{B}'_\cdot}, T'\right), t\right) \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty$$

for all  $t \geq 0$ . Let  $\mathcal{B}_t = \mathcal{B}'_{[t]^{T'}}$ . Clearly we can replace  $T'$  by any  $S$ -dense subset of  $T'$  of the form  $\{k\Delta''t \mid k \in \mathbb{N}_0\}$  and (9.6) will still hold. Therefore we may argue just as in the proof of Theorem 7.6 to obtain a  $\mathcal{B}_t$ -semimartingale lifting of  $(a; x), (A; X)$ , and a  $\mathcal{B}_t$ -adapted SDJ lifting of  $h, H$ , such that  $(H, A, X)$  is SDJ and  $X(t) = X'([t]^{T'} \wedge W)$  for some  $*$ -stopping time  $W$  that satisfies  ${}^\circ W = \infty$  a.s. Let  $Z = A + X$ .

Let  $G \in L(A; X, \mathbb{R}^{n \times d}, \mathcal{B}_\cdot)$ . If  $G_m = GI_{\{\|G\| \leq m\}}$ , then by Lemma 7.13(c), in order to show  $(H, G \cdot Z)$  is SDJ, it suffices to show each  $(H, G_m \cdot Z)$  is SDJ. This will follow if each  $G_m \cdot Z$  is SDJ since  $(H, Z)$  is SDJ and  $\|G_m\|$  is bounded. To show  $G_m \cdot Z$  is SDJ we only need show  $G_m \cdot X$  is SDJ because  $G_m \cdot A$  is clearly SDJ and so is  $(A, X)$ . Moreover, in proving that  $G_m \cdot X$  is SDJ we may assume  $n = d = 1$  by considering the matrix multiplication componentwise and noting that  $X$  is SDJ. Finally, by adding  $m + 1$  to  $G_m$  one can assume that  $1 \leq G_m \leq 2m + 1$  (note that if  $G_m \cdot X + (m + 1)X$  is SDJ then so is  $G_m \cdot X$ ).

By the above argument we may assume  $n = d = 1$ , and it suffices to show  $F \cdot X$  is SDJ for  $F \in L(A; X, \mathbb{R}, \mathcal{B}_\cdot)$  such that  $1 \leq F \leq M$  for some  $M \in \mathbb{N}$ . Fix such an  $F$ . Let  $F'(\underline{s}) = F([\underline{s}]^{T'} + \Delta t - \Delta t)$ . Since  $F$  is  $\mathcal{B}_s$ -adapted,  $F'$  is also  $\mathcal{B}_s$ -adapted and hence  $\mathcal{B}'_s$ -adapted. Note that if  $\underline{t} + \Delta t \in T'$  and  $\underline{t} < W$ , then

$$\begin{aligned} F \cdot X(\underline{t}) &= \sum_{\underline{s} < \underline{t}} F(\underline{s}) \Delta X(\underline{s}) = \sum_{\substack{\underline{s} < \underline{t} \\ \underline{s} + \Delta t \in T'}} F'(\underline{s}) (X'(\underline{s} + \Delta t) - X'(\underline{s} + \Delta t - \Delta t)) \\ &= \sum_{\underline{s} < \underline{t}} F'(\underline{s}) \Delta X'(\underline{s}) = F' \cdot X'(\underline{t}). \end{aligned}$$

Therefore it suffices to show  $F' \cdot X'$  is SDJ because  $F \cdot X$  is constant on  $[\underline{t}, \underline{t} + \Delta t)$  for all  $\underline{t}$  in  $T'$ . By Lemma 7.13(a)  $F' \cdot X'(\underline{t})$  is an  $S$ -local martingale with respect to  $\{\mathcal{B}'_t\}$ . We will use Theorem 9.4 to show that  $F' \cdot X'$  is SDJ. Observe that  $[F' \cdot X', F' \cdot X']$  is SDJ because  $[X', X']$  is, and  $|\Delta[F' \cdot X', F' \cdot X'](\underline{t})| \leq M^2 \Delta[X', X'](\underline{t})$ . To complete the proof it suffices to show that for all  $t \in [0, \infty)$ ,

$$(9.7) \quad w_0(J(F' \cdot X', \epsilon_n)^{\mathcal{B}'_\cdot}, t) \xrightarrow{P} 0 \quad \text{as } n \rightarrow \infty,$$

where  $\epsilon_n(\underline{t}, \omega) = n^{-1}F'(\underline{t}, \omega)$  satisfies the hypotheses of Theorem 9.4 because  $n^{-1} \leq \epsilon_n \leq n^{-1}M$ . For a.a.  $\omega$  if  $\underline{u} \approx \underline{v} \in T, t \in [0, \infty)$  and  $\underline{u} \leq \underline{v} \leq t$ , then since  $F'(\underline{s}) = F'([\underline{s}]^{T'})$  we have

$$\begin{aligned} &|J(F' \cdot X', \epsilon_n)^{\mathcal{B}'_\cdot}(\underline{v}) - J(F' \cdot X', \epsilon_n)^{\mathcal{B}'_\cdot}(\underline{u})| \\ &= {}^\circ \left| \sum_{\underline{u} \leq \underline{s} < \underline{v}} F'(\underline{s}) \bar{E}(\Delta X'(\underline{s}) I_{\{|F'(\underline{s}) \Delta X'(\underline{s})| > n^{-1}F'(\underline{s})\}} \mid \mathcal{B}'_{\underline{s}}) \right| \\ &\leq {}^\circ \sum_{\substack{[\underline{u}]^{T'} < \underline{s} < [\underline{v}]^{T'} \\ \underline{s} \in T'}} |F'(\underline{s})| |J(X', n^{-1})^{\mathcal{B}'_\cdot}(\underline{s} + \Delta t) - J(X', n^{-1})^{\mathcal{B}'_\cdot}(\underline{s})| \\ &\quad + {}^\circ |F'(\underline{u})| |J(X', n^{-1})^{\mathcal{B}'_\cdot}([\underline{u}]^{T'} + \Delta t) - J(X', n^{-1})^{\mathcal{B}'_\cdot}(\underline{u})| \\ &\quad + {}^\circ |F'(\underline{v})| |J(X', n^{-1})^{\mathcal{B}'_\cdot}(\underline{v}) - J(X', n^{-1})^{\mathcal{B}'_\cdot}([\underline{v}]^{T'})| \\ &\leq Mw_0(\text{Var}(J(X', n^{-1})^{\mathcal{B}'_\cdot}, T'), t) + 2Mw_0(J(X', n^{-1})^{\mathcal{B}'_\cdot}, t). \end{aligned}$$

Therefore we have for a.a.  $\omega$  and all  $t \in [0, \infty)$ ,

$$(9.8) \quad w_0(J(F' \cdot X', \varepsilon_n)^{\mathfrak{B}'}, t) \\ \leq Mw_0(\text{Var}(J(X', n^{-1})^{\mathfrak{B}'}, T'), t) + 2Mw_0(J(X', n^{-1})^{\mathfrak{B}'}, t).$$

The right side of (9.8) converges to zero in probability as  $n$  approaches  $\infty$ , by Theorem 9.4 and (9.6). Therefore (9.7) follows and the proof is complete.  $\square$

**10. Stochastic differential equations.** In this section our aim is to construct a solution,  $y(t)$ , of the stochastic differential equation

$$(10.1) \quad y(t, \omega) = h(t, \omega) + \int_0^t f(s, \omega, y(\cdot, \omega)) dz(s, \omega)$$

where  $z$  is a  $d$ -dimensional semimartingale,  $h: [0, \infty) \times \Omega \rightarrow \mathbf{R}^n$ , and  $f: [0, \infty) \times \Omega \times D(\mathbf{R}^n) \rightarrow \mathbf{R}^{n \times d}$ . To describe the exact conditions on  $f$ , we need some notation.

*Notation 10.1.* (a) The topology,  $\mathfrak{T}$ , on  $D(\mathbf{R}^n)$  of uniform convergence on compact sets is the topology induced by the metric

$$\rho_r(d_1, d_2) = \sum_{n=1}^{\infty} \left( \sup_{u \leq n} \|d_1(u) - d_2(u)\| \wedge 1 \right) 2^{-n}.$$

The subscripts  $\mathfrak{T}$  and  $J_1$  will be used to distinguish between these two topologies on  $D(\mathbf{R}^n)$  if there is any ambiguity. The notation “st” still denotes the standard part map in the  $J_1$  topology and  $\mathfrak{D}$  denotes the Borel sets of  $D(\mathbf{R}^n)$  for the  $J_1$  topology.

(b) If  $F: T \rightarrow {}^*\mathbf{R}^m$  is internal and  $\underline{t} \in T$ , let  $F^{\underline{t}}(s) = F(\underline{t} \wedge s)$ . It will be convenient to consider such functions  $F$  as elements of  ${}^*D(\mathbf{R}^m)$  by interpolating them as right-continuous step functions on  ${}^*[0, \infty)$ . If  $d \in D(\mathbf{R}^m)$  and  $t \geq 0$ , let

$$d^{t-}(s) = \begin{cases} d(s) & \text{if } s < t, \\ d(t-) & \text{if } s \geq t, \end{cases}$$

where  $d(0-) = d(0)$ .

(c) If  $L$  and  $M$  are topological spaces, let  $C(L, M)$  denote the space of continuous functions from  $L$  to  $M$  with the compact-open topology.  $\square$

We assume that  $z$ ,  $h$  and  $f$  satisfy the following hypotheses:

(H<sub>1</sub>)  $h$  is an  $\mathfrak{T}_t$ -adapted process with sample paths in  $D(\mathbf{R}^n)$ .

(H<sub>2</sub>)  $z$  is a  $d$ -dimensional semimartingale such that  $z(0) = 0$ .

(H<sub>3</sub>)  $f$  is  $\mathfrak{P} \times \mathfrak{D}$ -measurable (recall that  $\mathfrak{P}$  is the  $\sigma$ -algebra of predictable sets) and for all  $(t, \omega, d) \in [0, \infty) \times \Omega \times D(\mathbf{R}^n)$ ,  $f(t, \omega, d) = f(t, \omega, d^{t-})$ .

(H<sub>4</sub>) There is a process  $\tilde{f} \in L(z, \mathbf{R})$  (see Notation 7.2) such that

$$\sup_{d \in D(\mathbf{R}^n)} \|f(t, \omega, d)\| \leq \tilde{f}(t, \omega) \quad \text{for all } (t, \omega).$$

(H<sub>5</sub>) For each  $(t, \omega)$ ,  $f(t, \omega, \cdot)$  is  $\mathfrak{T}$ -continuous on  $D(\mathbf{R}^n)$ .

Conditions (H<sub>3</sub>) and (H<sub>4</sub>) are needed to ensure that the right side of (10.1) makes sense for all  $t \geq 0$ , as the following lemma shows.

**LEMMA 10.2.** *If  $y: [0, \infty) \times \Omega \rightarrow \mathbf{R}^n$  is  $\mathfrak{T}_t$ -adapted and has sample paths in  $D$ , then  $f(s, \omega, y) \in L(z, \mathbf{R}^{n \times d})$ .*

PROOF. Since  $\|f(s, \omega, y)\| \leq \tilde{f}(s, \omega)$  and  $\tilde{f} \in L(z, \mathbf{R})$ , it clearly suffices to show that  $f(s, \omega, y)$  is predictable. Define  $\phi: [0, \infty) \times \Omega \rightarrow D(\mathbf{R}^n)$  by  $\phi(t, \omega) = y^{t-}(\omega)$ . For  $s$  fixed,  $\phi(\cdot, \cdot)(s)$  is predictable since it is  $\mathcal{F}_t$ -adapted and is left-continuous in  $t$ . Therefore  $\phi$  is  $\mathcal{P}/\mathcal{D}$ -measurable, since  $\mathcal{D}$  is generated by the finite-dimensional sets. It follows that  $f(\cdot, \cdot, y)$  is predictable because  $f(t, \omega, y) = f(t, \omega, \phi(t, \omega))$  and  $f$  is  $\mathcal{P} \times \mathcal{D}$ -measurable (by  $H_3$ ).  $\square$

The main result of this section is the following:

**THEOREM 10.3.** *If conditions  $(H_1)$ – $(H_5)$  are satisfied on an adapted Loeb space  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ , then there is an  $\mathcal{F}_t$ -adapted process,  $y$ , with sample paths in  $D(\mathbf{R}^n)$  such that (10.1) holds for all  $t \geq 0$  a.s.*

PROOF. Choose  $(a, x) \in \mathcal{V}_0^d \times \mathcal{L}_0^d$  such that  $z = a + x$  and  $\tilde{f} \in L_{\text{loc}}(x, \mathbf{R}) \cap L_s(a, \mathbf{R})$ . By Theorem 9.7 there is an internal filtration  $\{\mathcal{B}_t\}$ , a  $\mathcal{B}_t$ -adapted SDJ lifting of  $h, H$ , and a  $\mathcal{B}_t$ -semimartingale lifting of  $(a; x)$ ,  $(A; X)$ , such that if  $Z = A + X$ , then  $(H, G \cdot Z)$  is SDJ for all  $G \in L(A; X, \mathbf{R}^{n \times d}, \mathcal{B}_t)$ . By Theorem 7.11,  $\tilde{f}$  has an  $(A; X)$ -lifting,  $\bar{F}$ , which we may assume is nonnegative for all  $(t, \omega)$ . Let

$$B = \{G: T \times \Omega \rightarrow {}^*\mathbf{R}^{n \times d} \mid G \text{ is } \mathcal{B}_t\text{-adapted and } \|G(t, \omega)\| \leq \bar{F}(t, \omega) \text{ for all } (t, \omega)\}.$$

The next step is to find an appropriate lifting of  $f$ . For this we need the following lifting lemma:

**LEMMA 10.4.** *There is an internal  $\mathcal{C} \times \mathcal{Q} \times {}^*\mathcal{D}$ -measurable mapping  $F: T \times \Omega \times {}^*D(\mathbf{R}^n) \rightarrow {}^*\mathbf{R}^{n \times d}$  such that:*

$$(10.2) \text{ For all } (t, \omega) \in T \times \Omega, \sup_{d \in {}^*D(\mathbf{R}^n)} \|F(t, \omega, d)\| \leq \bar{F}(t, \omega).$$

$$(10.3) \text{ For all } t \in T, F(t, \cdot, \cdot) \text{ is } \mathcal{B}_t \times {}^*\mathcal{D}\text{-measurable.}$$

$$(10.4) \text{ If } G \in B, \text{ then}$$

$${}^\circ F(t, \omega, H^t + (G \cdot Z)^t) = f({}^\circ t, \omega, \text{st}(H + G \cdot Z)) \quad L(\mu(A; X))\text{-a.s.}$$

We defer the proof of this lemma and complete the proof of Theorem 10.3.

Let  $F$  be as in Lemma 10.4 and inductively define a  $\mathcal{B}_t$ -adapted process  $Y$  by setting  $Y(0) = H(0)$  and

$$\Delta Y(t) = \Delta H(t) + F(t, \omega, Y^t) \Delta Z(t).$$

Thus we have

$$(10.5) \quad Y(t) = H(t) + \tilde{F} \cdot Z(t),$$

where  $\tilde{F}(s, \omega) = F(s, \omega, Y^s)$ . By (10.2) and (10.3) we have  $\tilde{F} \in B \subset L(A; X, \mathbf{R}^{n \times d}, \mathcal{B}_t)$  (the latter inclusion follows easily from the fact that  $\bar{F} \in L(A; X, \mathbf{R}, \mathcal{B}_t)$ ). It follows from (10.4) that

$${}^\circ \tilde{F}(s, \omega) = {}^\circ F(s, \omega, H^s + (\tilde{F} \cdot Z)^s) = f({}^\circ s, \omega, \text{st}(Y)) \quad L(\mu(A; X))\text{-a.s.}$$

Therefore if  $y = \text{st}(Y)$  (note that  $Y$  is SDJ by (10.5) and the choice of  $H$  and  $Z$ ), then  $\tilde{F}$  is an  $(A; X)$ -lifting of  $f(s, \omega, y)$ . Taking standard parts in (10.5) and using the definition of the stochastic integral (Definition 7.14), one obtains (10.1).  $\square$

It remains to prove Lemma 10.4. Suppose for the moment that, instead of (10.1), we were constructing solutions of

$$(10.6) \quad y(t) = h(t) + \int_0^t f(s, \omega, y(s^-)) dz(s),$$

where  $f(s, \omega, \cdot)$  is a continuous function from  $\mathbf{R}^n$  to  $\mathbf{R}^{n \times d}$ , satisfying the obvious analogues of (H<sub>3</sub>) and (H<sub>4</sub>). (Such equations were first solved by P. Protter [10, 11] and C. Doléans-Dade [2] when  $f(s, \omega, \cdot)$  is Lipschitz continuous.) Then an appropriate lifting,  $F$ , of  $f$  may be found by using Theorem 7.11 to obtain a lifting  $\hat{F}$ :  $T \times \Omega \rightarrow {}^*C(\mathbf{R}^n, \mathbf{R}^{n \times d})$  of  $\hat{f}(t, \omega)(x) = f(t, \omega, x)$  ( $\hat{f}$ :  $[0, \infty) \times \Omega \rightarrow C(\mathbf{R}^n, \mathbf{R}^{n \times d})$ ) and setting  $F(t, \omega, x) = \hat{F}(t, \omega)(x)$ . A slight modification of the previous proof would then allow one to construct solutions of (10.6). This argument breaks down when considering (10.1) since then  $\hat{f}$  would take values in the nonseparable space  $C(D(\mathbf{R}^n), \mathbf{R}^{n \times d})$ , where  $D$  has the  $\mathfrak{T}$  topology, and hence Theorem 7.11 does not apply. An additional problem arises in establishing (10.4) since “st” refers to the standard part map in the  $J_1$  topology while  $f(t, \omega, \cdot)$  is only assumed to be  $\mathfrak{T}$ -continuous. Both these difficulties are overcome by noting that we are only concerned with the behaviour of  $f(t, \omega, \cdot)$  on a particular subset,  $K(\omega)$ , of  $D$ . In fact, this subset of interest will be a countable union of compact sets on which the two topologies coincide.

*Notation 10.5.* Let  $j(t, \omega) = \|\Delta h(t, \omega)\| + \hat{f}(t, \omega)\|\Delta z(t, \omega)\|$ , and for  $(t, \omega) \in [0, \infty) \times \Omega$  define

$$D(t, \omega) = \{d \in D \mid d = d^{t^-}, \|\Delta d(s)\| \leq j(s, \omega) \text{ for all } s < t\}$$

(here  $d^\infty \equiv d$ ).  $\square$

**LEMMA 10.6.** *For all  $(t, \omega)$ ,  $D(t, \omega)$  is closed in the  $J_1$  topology, and the  $J_1$  and  $\mathfrak{T}$  topologies coincide on  $D(t, \omega)$ .*

**PROOF.** For each  $n \in \mathbf{N}$ ,  $\{t \leq n \mid j(t, \omega) \geq n^{-1}\}$  is finite. Therefore

$$(10.7) \quad {}^*j(s, \omega) \approx 0 \text{ for all } s \text{ in } \text{ns}({}^*[0, \infty)) - [0, \infty) \text{ and all } \omega.$$

Fix  $(t, \omega) \in [0, \infty) \times \Omega$ . If  $d_1, d_2 \in {}^*D(t, \omega)$  satisfy  $d_1 \approx_{J_1} d_2$  and  $d_1 \not\approx_\tau d_2$ , then there are  $t_1 \approx t_2$  in  $[0, t]$  such that  $t_1 \neq t_2$ ,  ${}^\circ\Delta d_1(t_1) \neq 0$ , and  ${}^\circ\Delta d_2(t_2) \neq 0$ . Therefore  ${}^*j(t_i, \omega) > 0$  for  $i = 1, 2$ , which contradicts (10.7). Hence  $d_1 \approx_\tau d_2$  whenever  $d_1 \approx_{J_1} d_2$ , and it follows easily that the  $J_1$  and  $\tau$  topologies coincide on  $D(t, \omega)$ .

Suppose  $d \in {}^*D(t, \omega)$  is nearstandard in the  $J_1$  topology. To see that  $D(t, \omega)$  is closed, we must show that  $\text{st}(d) \in D(t, \omega)$ . By (10.7),  ${}^\circ\Delta d(s) = 0$  for all  $s \approx t$  such that  $s < t$  and therefore  $d = d^{t^-}$  implies that  $\text{st}(d) = \text{st}(d)^{t^-}$ . If  $s \in [0, t)$  then for some  $s' \approx s$  we have

$$\|\Delta \text{st}(d)(s)\| = {}^\circ\|\Delta d(s')\| \leq {}^*j(s', \omega) \leq j(s, \omega)$$

(the last by (10.7)). Therefore  $\text{st}(d) \in D(t, \omega)$  and the proof is complete.  $\square$



LEMMA 10.7. *There is a  $\mathcal{P} \times \mathcal{D}$ -measurable mapping  $g: [0, \infty) \times \Omega \times D(\mathbf{R}^n) \rightarrow \mathbf{R}^{n \times d}$  such that for all  $(t, \omega)$ ,*

$$(10.8) \quad g(t, \omega, d) = g(t, \omega, d^{\tau}) \text{ for all } d,$$

$$(10.9) \quad g(t, \omega, \cdot) \text{ is } J_1\text{-continuous on } D(\mathbf{R}^n),$$

$$(10.10) \quad g(t, \omega, d) = f(t, \omega, d) \text{ for all } d \in D(t, \omega).$$

The proof requires the following auxiliary lemma.

LEMMA 10.8. *Assume  $f: [0, \infty) \times \Omega \times D(\mathbf{R}^n) \rightarrow \mathbf{R}$  satisfies  $(H_3)$ – $(H_5)$ . There is a  $\mathcal{P} \times \mathcal{D}$ -measurable mapping  $g: [0, \infty) \times \Omega \times D(\mathbf{R}^n) \rightarrow \mathbf{R}$  satisfying (10.8) and (10.9), such that for all  $(t, \omega)$ ,*

$$(10.11) \quad |g(t, \omega, d)| \leq \bar{f}(t, \omega)/3 \text{ for all } d,$$

$$(10.12) \quad |g(t, \omega, d) - f(t, \omega, d)| \leq \frac{2}{3} \bar{f}(t, \omega) \text{ for all } d \text{ in } D(t, \omega).$$

PROOF. Define

$$A^+(t, \omega) = \{d \in D(t, \omega) \mid f(t, \omega, d) > \bar{f}(t, \omega)/3\}$$

and

$$A^-(t, \omega) = \{d \in D(t, \omega) \mid f(t, \omega, d) < -\bar{f}(t, \omega)/3\}.$$

Then by Lemma 10.6, for each  $(t, \omega)$ , the  $\tau$ -closure of  $A^\pm, \overline{A^\pm}$ , is also  $J_1$ -closed. Moreover,  $\overline{A^+}$  and  $\overline{A^-}$  are clearly disjoint.

Let  $\rho_J$  denote a metric for the  $J_1$ -topology on  $D(\mathbf{R}^n)$  and define

$$g(t, \omega, d) = (\rho_{J_1}(d^{\tau}, A^-(t, \omega)) \wedge \rho_{J_1}(A^-(t, \omega), A^+(t, \omega))) \\ \times \rho_{J_1}(A^-(t, \omega), A^+(t, \omega))^{-1} \frac{2}{3} \bar{f}(t, \omega) - \bar{f}(t, \omega)/3,$$

where  $\rho_{J_1}(B, C) = \inf\{\rho_{J_1}(b, c) \mid b \in B, c \in C\}$ ,  $\inf \emptyset = \infty$  and  $\infty/\infty = 1$ . (Note that  $\rho_{J_1}(A^-, A^+) > 0$  since  $\overline{A^-}$  and  $\overline{A^+}$  are disjoint  $J_1$ -closed sets.) Clearly  $g$  satisfies (10.8), (10.9), (10.11) and (10.12).

It remains only to show that  $g$  is  $\mathcal{P} \times \mathcal{D}$ -measurable. To this end we construct a sequence of  $\mathcal{F}_t$ -adapted processes  $\{a_i \mid i \in \mathbf{N}\}$  that is a dense subset of  $D(\infty, \omega)$  for each  $\omega$ . Let  $\{T_i\}$  be a sequence of stopping times such that

$$\{T_i \mid T_i < \infty\} = \{t \mid (h, z)(t) \neq (h, z)(t^-)\}$$

and  $T_i \neq T_j$  if  $T_i < \infty$  and  $i \neq j$ , for all  $\omega$ . Let  $S$  be a countable dense subset of the unit ball in  $\mathbf{R}^n$ . For each  $m \in \mathbf{N}$  and  $r = (r_1, \dots, r_m) \in S^m$  define an  $\mathcal{F}_t$ -adapted process  $a(m, r)$  by setting

$$a(m, r)(0) = 0,$$

$$\Delta a(m, r)(T_i) = r_i j(T_i) \quad \text{if } T_i < \infty, i = 1, \dots, m,$$

and then interpolating  $a(m, r)(\cdot)$  as a right-continuous step function. Let  $\{x_i(t) \mid i \in \mathbf{N}\}$  be a countable dense subset of  $C([0, \infty), \mathbf{R}^n)$  and let  $\{a_i(\cdot, \omega) \mid i \in \mathbf{N}\}$  be an enumeration of the countable collection  $\{x_i(\cdot) + a(m, r)(\cdot, \omega) \mid i, m \in \mathbf{N}, r \in S^m\}$ . Clearly each  $a_i$  is  $\mathcal{F}_t$ -adapted and  $a_i(\cdot, \omega) \in D(\infty, \omega)$  for all  $\omega$ . Moreover it is easy to see that  $\{a_i(\cdot, \omega) \mid i \in \mathbf{N}\}$  is  $\tau$ -dense in  $D(\infty, \omega)$  for all  $\omega$ . Therefore  $\{a_i^{\tau} \mid i \in \mathbf{N}\}$  is  $\tau$ -dense in  $D(t, \omega)$  and  $\{a_i^{\tau} \mid f(t, \omega, a_i^{\tau}) < -\bar{f}(t, \omega)/3\}$  is  $\tau$ -dense in  $A^-(t, \omega)$  for all

$(t, \omega)$ , since  $A^-$  is a  $\tau$ -open subset of  $D(t, \omega)$ . In particular we have

$$(10.13) \quad \rho_{J_1}(d^{\tau^-}, A^-(t, \omega)) = \inf \left\{ \rho_{J_1}(d^{\tau^-}, a_i(\cdot, \omega)^{\tau^-}) \mid f(t, \omega, a_i^{\tau^-}) < -\bar{f}(t, \omega)/3 \right\}.$$

The mapping  $\phi_i: [0, \infty) \times \Omega \rightarrow (D(\mathbf{R}^n), \tau)$  defined by  $\phi_i(t, \omega) = a_i(\cdot, \omega)^{\tau^-}$  is adapted, left continuous, and hence  $\mathcal{P}/\mathcal{D}$ -measurable. It follows easily that

$$\psi_i(t, \omega, d) = \rho_{J_1}(d^{\tau^-}, a_i(\cdot, \omega)^{\tau^-})$$

is  $\mathcal{P} \times \mathcal{D}$ -measurable and hence so is  $\rho_{J_1}(d^{\tau^-}, A^-(t, \omega))$  by (10.13). Similarly,  $\rho_{J_1}(A^-, A^+)$  is  $\mathcal{P}$ -measurable and therefore  $g$  is  $\mathcal{P} \times \mathcal{D}$ -measurable.  $\square$

**PROOF OF LEMMA 10.7.** By constructing  $g$  componentwise we may assume  $f: [0, \infty) \times \Omega \times D(\mathbf{R}^n) \rightarrow \mathbf{R}$ . The proof of Lemma 10.7 is now identical to the usual proof of the Tietze extension theorem which may be found on p. 150 of Dugundji [3]. The only difference is that Lemma 10.8 is used in place of the lemma at the top of p. 150 in [3].  $\square$

We require one more technical lemma before we can prove Lemma 10.4.

**LEMMA 10.9.** (a) Suppose  $(M, \rho)$  is a separable metric space,  $K$  is an internal subset of  $\text{ns}(*M)$ , and  $C = \text{st}(K)$ . Then there is an internally measurable mapping  $\phi: *M \rightarrow *M$  (i.e., if  $\mathcal{B}$  is the Borel subsets of  $M$ , then  $\phi^{-1}(B) \in *\mathcal{B}$  for all  $B \in *\mathcal{B}$ ) such that  $\phi(K) \subset *C$  and  $\phi(y) \approx y$  for all  $y \in K$ .

(b) There is a sequence of internal mappings  $\{N_m\}$  from  $B$  to  $\mathcal{Q}$  such that for all  $m \in \mathbf{N}$ ,  $G \in B$  and  $\omega \in N_m(G)$ :

$$(10.14) \quad \bar{P}(N_m(G)) > 1 - 2^{-m},$$

$$(10.15) \quad H + G \cdot Z \text{ is SDJ},$$

$$(10.16) \quad \text{st}(H + G \cdot Z) \in D(\infty, \omega).$$

**PROOF.** (a) By Stroyan and Luxemburg [12, Theorem 8.3.11],  $C$  is compact. For each  $n \in \mathbf{N}$  choose  $x_1, \dots, x_{m_n}$  in  $C$  such that  $C \subset \bigcup_{i=1}^{m_n} B(x_i, n^{-1})$  (here  $B(x, \varepsilon) = \{y \in M \mid \rho(y, x) < \varepsilon\}$ ). Define  $\phi_n: *M \rightarrow *M$  by

$$\phi_n(y) = \begin{cases} x_i & \text{if } y \in B(x_i, n^{-1}) - \bigcup_{j < i} B(x_j, n^{-1}) \text{ for some } i \leq m_n, \\ x_{m_n} & \text{if } y \notin \bigcup_{i=1}^{m_n} B(x_i, n^{-1}). \end{cases}$$

Clearly  $\phi_n$  is internally measurable. Moreover if  $y \in K$ , then  $\rho(y, x_i) < n^{-1}$  for some  $i \leq m_n$  and therefore  $\phi_n(y) \in *C$  and  $\rho(y, \phi_n(y)) < n^{-1}$ . By saturation we may obtain  $\gamma \in *\mathbf{N} - \mathbf{N}$  and an internally measurable mapping  $\phi_\gamma: *M \rightarrow *M$  such that  $\phi_\gamma(K) \subset *C$  and  $\rho(y, \phi_\gamma(y)) < \gamma^{-1}$  for all  $y \in K$ . Hence  $\phi_\gamma$  is the required function.

(b) Let  $\{\Lambda_m \mid m \in \mathbf{N}\}$  be a sequence of sets in  $\mathcal{Q}$  such that  $\bar{P}(\Lambda_m) > 1 - 2^{-m-2}$  and

$$\Lambda_m \subset \{\omega \mid (H, \bar{F} \cdot Z) \text{ is SDJ}, \text{st}(H, \bar{F} \cdot Z) = (h, \bar{f} \cdot z),$$

$$\|\Delta(\bar{f} \cdot z)(u)\| = \bar{f}(u)\|\Delta z(u)\| \text{ for all } u \geq 0\}$$

(we abuse the notation slightly and write  $\bar{F} \cdot Z$  for  $(\bar{F}I_{d \times d}) \cdot Z$ , and  $\bar{f} \cdot z$  for  $(\bar{f}I_{d \times d}) \cdot z$ , where  $I_{d \times d}$  is the  $d \times d$  identity matrix). The existence of such a

sequence follows from the choice of  $H$  and  $Z$ , and the well-known fact that  $\Delta(\bar{f} \cdot z)(u) = \bar{f}(u)\Delta z(u)$  for all  $u \geq 0$  a.s. (this is an easy consequence of our definition of  $\bar{f} \cdot z$ ). If  $m, n \in {}^*\mathbf{N}$  and  $G \in B$ , let

$$\begin{aligned} A(m, n, G) = \Lambda_n \cap \left\{ \omega \mid \max \left\{ \|G \cdot Z(\underline{t}_1) - G \cdot Z(\underline{t}_2)\| \right. \right. \\ \left. \times \left( \max_{\underline{t}_1 \leq \underline{u} < \underline{t}_2} \bar{F}(\underline{u}) \|\Delta Z(\underline{u})\| + 2^{-n} \right)^{-1} \mid \right. \\ \left. \left. \underline{t}_1 < \underline{t}_2 \leq n, \underline{t}_2 - \underline{t}_1 < m^{-1} \right\} \leq 1 \right\} \\ \cap \left\{ \max_{\underline{t} \leq n} \|G \cdot Z(\underline{t})\| \leq m \right\}, \end{aligned}$$

and

$$k(n, G) = \min \{ m \mid \bar{P}(A(m, n, G)) > 1 - 2^{-n-1} \} \quad (\min \emptyset = \infty).$$

Note that  $k(n, G) \in \mathbf{N}$  if  $n \in \mathbf{N}$  because  $G \cdot Z$  is SDJ for  $G \in B$  and  $\bar{P}(\Lambda_n) \geq 1 - 2^{-n-2}$ . Let  $k_n = \sup_{G \in B} k(n, G)$ . If  $n \in \mathbf{N}$ , then  $k_n \in \mathbf{N}$  and

$$(10.17) \quad \bar{P}(A(k_n, n, G)) > 1 - 2^{-n-1} \quad \text{for all } G \in B.$$

By the permanence principle (10.17) holds for all  $n \leq \gamma$  for some  $\gamma \in {}^*\mathbf{N} - \mathbf{N}$ . Let  $N_m(G) = \bigcap_{j=m}^\gamma A(k_j, j, G)$ . It is easy to check that (10.14) and (10.15) are satisfied. To show (10.16) note that if  $m \in \mathbf{N}$ ,  $G \in B$ ,  $\omega \in N_m(G)$  and  $\underline{t} \in \text{ns}(T)$ , then

$$\begin{aligned} {}^\circ \|\Delta H(\underline{t}) + \Delta(G \cdot Z)(\underline{t})\| &\leq {}^\circ \|\Delta H(\underline{t})\| + {}^\circ \|\Delta(\bar{F} \cdot Z)(\underline{t})\| \\ &\leq \|\Delta h({}^\circ \underline{t})\| + \|\Delta(\bar{f} \cdot z)({}^\circ \underline{t})\| \\ &= j({}^\circ \underline{t}, \omega) \quad (\text{since } \omega \in \Lambda_m). \end{aligned}$$

Therefore  $\text{st}(H + G \cdot Z) \in D(\infty, \omega)$  because  $H + G \cdot Z$  is SDJ.  $\square$

**PROOF OF LEMMA 10.4.** Let  $K_m = \{(H + G \cdot Z)^\underline{t} \mid G \in B, \underline{t} \in T, \omega \in N_m(G)\}$  and  $C_m = \text{st}(K_m) \subset D(\infty, \omega)$ , where  $\{N_m\}$  is as in the previous lemma. Then  $C_m$  is  $J_1$ -compact because  $K_m$  is an internal set of nearstandard points in the  $J_1$  topology. Let  $g$  be as in Lemma 10.7 and define  $g_m: [0, \infty) \times \Omega \rightarrow C(C_m, \mathbf{R}^{n \times d})$  (give  $C_m$  the  $J_1$  topology) by  $g_m(t, \omega)(d) = g(t, \omega, d)$ . Then  $g_m$  is predictable,  $C(C_m, \mathbf{R}^{n \times d})$  is separable metric and therefore by Theorem 7.11(a),  $g_m$  has a weak  $(A; X)$ -lifting,  $G_m$ . Let  $\phi_m$  be as in Lemma 10.9(a) with  $K_m$  in place of  $K$  and define  $F_m: T \times \Omega \times {}^*D(\mathbf{R}^n) \rightarrow {}^*\mathbf{R}^{n \times d}$  by

$$F_m(\underline{t}, \omega, d) = \begin{cases} G_m(\underline{t}, \omega)(\phi_m(d)) & \text{if } d \in K_m, \\ 0 & \text{otherwise.} \end{cases}$$

Then  $F_m$  is  $\mathcal{C} \times \mathcal{Q} \times {}^*\mathcal{D}$ -measurable and for each  $\underline{t}$  in  $T$ ,  $F_m(\underline{t}, \cdot, \cdot)$  is  $\mathcal{B}_t \times {}^*\mathcal{D}$ -measurable. The definition of the standard part map on  $\text{ns}({}^*C(C_m, \mathbf{R}^{n \times d}))$  (see Stroyan and Luxemburg [12, Theorem 8.4.41]) implies that

$$(10.18) \quad {}^\circ G_m(\underline{t}, \omega)(d) = g_m({}^\circ \underline{t}, \omega)(\text{st}(d)) \quad \text{for all } d \text{ in } {}^*C_m$$

$L(\mu(A; X))$ -a.s. Therefore if  $G \in B$  and  $(t, \omega) \in \text{ns}(T) \times N_m(G)$  satisfies (10.18), then

$$\begin{aligned} {}^\circ F_m(t, \omega, (H + G \cdot Z)^t) &= {}^\circ G_m(t, \omega)(\phi_m((H + G \cdot Z)^t)) \\ &= g_m({}^\circ t, \omega)(\text{st}((H + G \cdot Z)^t)) \\ &= g({}^\circ t, \omega, \text{st}(H + G \cdot Z)^t) \quad (\text{by (10.8)}) \\ &= f({}^\circ t, \omega, \text{st}(H + G \cdot Z)^t) \\ &\quad (\text{because (10.16) implies } \text{st}(H + G \cdot Z)^t \in D({}^\circ t, \omega)) \\ &= f({}^\circ t, \omega, \text{st}(H + G \cdot Z)). \end{aligned}$$

Therefore

(10.19)

$$\sup_{G \in B} L(\mu(A; X))({}^\circ F_m(t, \omega, (H + G \cdot Z)^t) \neq f({}^\circ t, \omega, \text{st}(H + G \cdot Z))) \leq 2^{-m}$$

and for all  $n \in \mathbb{N}$ ,

$$\begin{aligned} (10.20) \quad \sup_{G \in B} \mu(A; X)(&\|F_n(t, \omega, (H + G \cdot Z)^t) - F_m(t, \omega, (H + G \cdot Z)^t)\| \geq 2^{-n}) \\ &\leq 2^{-m+2} \quad \text{for all } m \leq n. \end{aligned}$$

By saturation we may choose  $\gamma$  in  ${}^*\mathbb{N} - \mathbb{N}$  such that  $F_\gamma$  is  $\mathcal{C} \times \mathcal{Q} \times {}^*\mathcal{D}$ -measurable,  $F_\gamma(t, \cdot, \cdot)$  is  $\mathcal{B}_t \times {}^*\mathcal{D}$ -measurable for all  $t$ , and (10.20) holds with  $\gamma$  in place of  $n$ . Therefore, if  $G \in B$ , then for  $L(\mu(A; X))$ -a.s.  $(t, \omega)$  we have

$$\begin{aligned} {}^\circ F_\gamma(t, \omega, (H + G \cdot Z)^t) &= \lim_{m \rightarrow \infty} {}^\circ F_m(t, \omega, (H + G \cdot Z)^t) \quad \text{by (10.20)} \\ &= f({}^\circ t, \omega, \text{st}(H + G \cdot Z)) \quad (\text{by 10.19}). \end{aligned}$$

Let  $F(t, \omega, d) = F_\gamma(t, \omega, d)(\bar{F}(t, \omega) \wedge \|F_\gamma(t, \omega, d)\| \|F_\gamma(t, \omega, d)\|^{-1})$  (let  $0/0 = 0$ ). Then  $F$  satisfies (10.2) and (10.3), and also satisfies (10.4) since  $F_\gamma$  does.  $\square$

**REMARK 10.10.** That one can find solutions to (10.1) without having to enlarge the given adapted Loeb space is an indication of the richness of such a space. It is natural to ask if the setting of an adapted Loeb space seriously restricts one in studying processes of class  $D$  on an arbitrary  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$  satisfying the “usual hypotheses”.

If one is only interested in the distributions of processes then nothing is lost. In fact much more is true, as we now show.

Let  $(X, \mathcal{G}, Q, (\mathcal{G}_t)_{t \geq 0})$  be a complete probability space equipped with a filtration satisfying the “usual hypotheses” and let  $y(t)$  be an  $M$ -valued ( $M$  is a Polish space),  $\mathcal{G}_t$ -adapted stochastic process with sample paths in  $D$ . Define  $(\Omega, \mathcal{Q}, \bar{P}) = (*X, *\mathcal{G}, *Q)$  and  $\mathcal{Q}_t = *\mathcal{G}_t$  for  $t$  in  $T = *[0, \infty)$ . Then  $(\Omega, \mathcal{Q}, \bar{P}, \mathcal{Q}_t)$  satisfies the conditions of §3 and we may define an adapted Loeb space,  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ , in the usual way. Since  $D$  with the  $J_1$  topology is a Polish space, there is an increasing sequence of compact subsets of  $D$ ,  $\{K_n\}$ , such that  $\lim_{n \rightarrow \infty} Q(y(\cdot, \omega) \in K_n) = 1$ .

Therefore  $\lim_{n \rightarrow \infty} {}^\circ Q(*y(\cdot, \omega) \in *K_n) = 1$  and hence  $P(*y(\cdot, \omega) \in \text{ns}(*D)) = 1$ . By Theorem 2.6,  $*y$  is SDJ and hence  $x = \text{st}(*y)$  is an  $\mathcal{F}_t$ -adapted process with sample paths in  $D$ . Most interesting properties of  $y$  will be inherited by  $x$ . For instance, it is easy to check that if  $y$  is a local martingale reduced by  $\{U_n\}$  (respectively, a semimartingale), then  $x$  is a local martingale reduced by  $\{{}^\circ U_n\}$  (respectively, a semimartingale).

To be more specific, we now show that  $(x, \mathcal{F}_\cdot)$  and  $(y, \mathcal{G}_\cdot)$  are synonymous in the following sense (introduced in Aldous [1]): if  $h_i: M \rightarrow \mathbf{R}$  are bounded and continuous, and  $u_i, t_i \geq 0$  ( $i = 1, \dots, k$ ), then  $(E(h_i(x(t_i)) | \mathcal{F}_{u_i}))_{i=1, \dots, k}$  and  $(E(h_i(y(t_i)) | \mathcal{G}_{u_i}))_{i=1, \dots, k}$  have the same distribution. (Note that this implies that  $x$  and  $y$  have the same law on  $D(M)$ .) Let  $\{t_i, u_i | i = 1, \dots, k\} \subset [0, \infty)$ ,  $\{v_i | i = 1, \dots, k\} \subset \mathbf{R}$  and  $h_1, \dots, h_k$  be bounded, continuous, real-valued functions on  $M$ . Then

$$\begin{aligned} & \int \exp \left\{ \sum_{j=1}^k i v_j E(h_j(y(t_j)) | \mathcal{G}_{u_j}) \right\} dQ \\ &= \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} {}^\circ \int \exp \left\{ \sum_{j=1}^k i v_j \bar{E}(*h_j(*y(t_j + n^{-1})) | \mathcal{G}_{u_j + m^{-1}}) \right\} d*Q \\ &= \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \int \exp \left\{ \sum_{j=1}^k i v_j E(h_j({}^\circ y(t_j + n^{-1})) | \sigma(*\mathcal{G}_{u_j + m^{-1}})) \right\} dP \\ &= \int \exp \left\{ \sum_{j=1}^k i v_j E(h_j(x(t_j)) | \mathcal{F}_{u_j}) \right\} dP. \end{aligned}$$

It follows that  $(x, \mathcal{F}_\cdot)$  and  $(y, \mathcal{G}_\cdot)$  are synonymous. In fact it is not hard to see that  $(x, \mathcal{F}_\cdot)$  and  $(y, \mathcal{G}_\cdot)$  have the same “adapted distribution” in the sense of Hoover and Keisler [4]. This means that if  $f(x, \mathcal{F}_\cdot)$  is any random variable defined using random variables  $x(t)$ ,  $t \in [0, \infty)$ , conditional expectations with respect to  $\sigma$ -fields  $\{\mathcal{F}_s | s \in [0, \infty)\}$ , and bounded continuous functions, and  $f(y, \mathcal{G}_\cdot)$  is the corresponding random variable defined from random variables  $y(t)$  and  $\sigma$ -fields  $\{\mathcal{G}_s\}$ , then  $E(f(y, \mathcal{G}_\cdot)) = E(f(x, \mathcal{F}_\cdot))$ . For example, if  $s_1, s_2, s_3, t_1, t_2, t_3 \in [0, \infty)$ , and  $\phi_1, \phi_3: \mathbf{R} \rightarrow \mathbf{R}$ ,  $\phi_2, \phi_4: \mathbf{R}^2 \rightarrow \mathbf{R}$  are bounded and continuous, then

$$\begin{aligned} & E \left[ \phi_1 \left( E \left( \phi_2 \left( E \left( \phi_3(y_{t_1}) | \mathcal{G}_{s_1} \right), E \left( \phi_4(y_{t_2}, y_{t_3}) | \mathcal{G}_{s_2} \right) \right) | \mathcal{G}_{s_3} \right) \right) \right] \\ &= E \left[ \phi \left( E \left( \phi_2 \left( E \left( \phi_3(x_{t_1}) | \mathcal{F}_{s_1} \right), E \left( \phi_4(x_{t_2}, x_{t_3}) | \mathcal{F}_{s_2} \right) \right) | \mathcal{F}_{s_3} \right) \right) \right]. \end{aligned}$$

This notion of adapted distribution is stronger than that of distribution because it takes into account the interactions of the process with the filtration, and it is also stronger than the intermediate notion of synonymity because it takes into account more of them. Many important properties, such as the Markov property and the martingale property, are shared by processes which have the same adapted distribution (in fact, for such concrete properties, synonymity is enough; see Aldous [1] for further examples). Further properties of adapted distribution are explored in the

paper of Hoover and Keisler. In particular it is shown there that non-path-dependent semimartingale stochastic integral equations (i.e., those of form (10.6)) can be transferred onto any adapted Loeb space, which carries an adapted Brownian motion, in a way which preserves the adapted distribution of both the coefficients  $(f, h, z)$  and the solution (if one exists on the given standard space). Something similar is true in the case with path dependent coefficients, but the result is more complicated to state because of the nonseparability of  $C(D, \mathbf{R}^{n \times d})$ .

As we said in the introduction, although it is of theoretical interest to define the precise strength of the Loeb space method, we think that physical processes may be modelled directly by stochastic processes on a Loeb space.  $\square$

ACKNOWLEDGEMENT. We would like to thank Professor Jean Pellaumail for suggesting that we extend our results on solutions to (10.6) to equations of the form (10.1). Our thanks also go to Professor Keith Stroyan for several helpful suggestions.

#### REFERENCES

1. D. J. Aldous, *Weak convergence and the general theory of processes*, École d'Été de Probabilités de Saint-Flour, 1983 (to appear).
2. C. Doléans-Dade, *On the existence and unicity of solutions of stochastic integral equations*, Z. Wahrsch. Verw. Gebiete **34** (1976), 93–101.
3. J. Dugundji, *Topology*, Allyn and Bacon, Boston, Mass. 1966.
4. D. N. Hoover and H. J. Keisler, *Adapted probability distributions* (preprint).
5. D. N. Hoover and E. Perkins, *Nonstandard construction of the stochastic integral and applications to stochastic differential equations*. I, Trans. Amer. Math. Soc. (to appear).
6. H. J. Keisler, *An infinitesimal approach to stochastic analysis*, Mem. Amer. Math. Soc. (to appear).
7. T. L. Lindström, *Hyperfinite stochastic integration*. I: *The nonstandard theory*, Math. Scand. **46** (1980), 265–292.
8. ———, *Hyperfinite stochastic integration*. II: *Comparison with the standard theory*, Math. Scand. **46** (1980), 293–314.
9. L. Panetta, *Hyperreal probability spaces: some applications of the Loeb construction*, Ph. D. thesis, Univ. of Wisconsin, Madison, Wis., 1978.
10. P. E. Protter, *On the existence, uniqueness, convergence and explosions of solutions of systems of stochastic integral equations*, Ann. Prob. **5** (1977), 243–261.
11. ———, *Right-continuous solutions of systems of stochastic integral equations*, J. Multivariate Anal. **7** (1977), 204–214.
12. K. D. Stroyan and W. A. J. Luxemburg, *Introduction to the theory of infinitesimals*, Academic Press, New York, 1976.

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CONNECTICUT 06520

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, BRITISH COLUMBIA, CANADA V6T 1Y4 (Current address of Edwin Perkins).

*Current address* (D. N. Hoover): Department of Mathematics and Statistics, Queen's University, Kingston (K7L 3N6), Ontario, Canada