The Annals of Probability
1985, Vol. 13, No. 2, 531–557

A STOCHASTIC INTEGRAL REPRESENTATION FOR RANDOM EVOLUTIONS

BY JOSEPH C. WATKINS

University of British Columbia

Previously we established that the martingales

$$M^\theta(t) = \left(\theta,\ Y(t) - Y(0) - \frac{1}{2}\int_0^t \int_\Xi A^2(\xi) Y(s)\mu(d\xi)\, ds\right),$$

with quadratic variation process

$$V^\theta(t) = \int_0^t \int_\Xi (\theta, A(\xi)Y(s))^2 \mu(d\xi)\, ds,$$

characterize the limit process for a sequence of random evolutions. This paper shows the equivalence of this presentation to the questions of existence and uniqueness of the stochastic integral equation

$$Y(t) = Y(0) + \frac{1}{2}\int_0^t \int_\Xi A^2(\xi) Y(s)\mu(d\xi)\, ds + \int_0^t \int_\Xi A(\xi)Y(s)W(d\xi, ds).$$

The paper proceeds in giving strong existence and uniqueness theorems for this integral equation.

1. Introduction. More than a decade has passed since Stroock and Varadhan [6] first used the martingale property to characterize a Markov process. This method has a great advantage because the parametric process of the stochastic integral has been eliminated. The martingale problem formulation focuses on the law of the process at hand. Hence, the tools of weak convergence are available to show that approximating Markov chains converge to a Markov process. These were the advantages that we exploited in giving a central limit theorem for random evolutions. In [7], we began with a sequence $\{Y_n\}$ of random evolutions, stochastic processes which take values in a separable Banach space B. At time t,

$$Y_n(t) = \exp(1/n)A_{[n^2t]} \cdots \exp(1/n)A_2\exp(1/n)A_1 Y_n(0) \tag{1.1}$$

where $Y_n(0)$ is some initial distribution on B, and $A, A_1, A_2, \cdots$ is an independent sequence of identically distributed generators of strongly continuous semigroups. A is defined on a probability space $(\Xi, \mathscr{X}, \mu)$ centered so that $\int_\Xi A(\xi)\mu(d\xi) = 0$. The theorems were a listing of assumptions on $\{A(\xi)\}$, (or more precisely on μ) sufficient for us to conclude that a limit process Y exists and is unique in law on $C_B[0, \infty)$. This limit process was characterized by a martingale problem. The martingale problem formulation that we shall present here is somewhat more

Received August 1982; revised August 1984.
AMS 1980 *subject classifications*. Primary 60H20; secondary 60H05, 60G44.
Key words and phrases. Martingale problem, martingale measures, stochastic integral equations, existence and uniqueness theorems.

cumbersome and less general than the one presented in [7]. In a moment, we shall see that, in addition to previous assumptions, we must hypothesize that $\{A(\xi): \xi \in \Xi\}$ be a set of generators of strongly continuous groups in order to conclude that the martingale problem, as stated below, is satisfied by some process. This seems restrictive, but I know no examples where the A are not group generators but $\int A\, d\mu = EA = 0$.

We shall begin with a short argument which is meant to convince us of the plausibility of the martingale problem as well as to recall the line of reasoning in [7]. Let $r, r_1, r_2, \cdots$ be an independent sequence of Rademacher random variables, i.e.,

$$\mu\{r = 1\} = \mu\{r = -1\} = \tfrac{1}{2}. \tag{1.2}$$

This sequence should be chosen independent of the $\{A_j\}$. Consider the random evolution Y_n that results from an iterated product of $\exp((1/n)r_iA_i)$. Then

$$Y_n(i/n^2) - Y_n((i-1)/n^2) - E[(\exp((1/n)rA) - I)]Y_n((i-1)/n^2) \tag{1.3}$$

is the centering of the ith increment of the process. In order to isolate the contributions to the evolution due to Γ, we make the following definitions:

$$Y_n(\Gamma, 0) = \mu(\Gamma)Y_n(0)$$

$$\begin{aligned} Y_n\left(\Gamma, \frac{i}{n^2}\right) - Y_n\left(\Gamma, \frac{i-1}{n^2}\right) &= \left(\exp\left(\frac{1}{n} I_\Gamma r_i A_i\right) - I\right)Y_n\left(\frac{i-1}{n^2}\right) \\ &= I_\Gamma\left(\exp\left(\frac{1}{n} r_i A_i\right) - I\right)Y_n\left(\frac{i-1}{n^2}\right) \end{aligned} \tag{1.4}$$

In other words, the process is censored outside Γ. Clearly if $\Gamma_1 \cap \Gamma_2 = \varnothing$

$$Y_n(\Gamma_1, t) + Y_n(\Gamma_2, t) = Y_n(\Gamma_1 \cup \Gamma_2, t). \tag{1.5}$$

In addition,

$$Y_n(\Xi, t) = Y_n(t), \tag{1.6}$$

$$Y_n\left(\Gamma, \frac{i}{n^2}\right) - Y_n\left(\Gamma, \frac{i-1}{n^2}\right) - E\left[I_\Gamma\left(\exp\left(\frac{1}{n} rA\right) - I\right)\right]Y_n\left(\frac{i-1}{n^2}\right) \tag{1.7}$$

is the centering of an increment of the process. Therefore, for all $\theta \in \mathscr{D}'$, a certain subspace of B^*,

$$\begin{aligned} &M_n^\theta(\Gamma, t) \\ &= (\theta, Y_n(\Gamma, t) - Y_n(\Gamma, 0) - \textstyle\sum_{i=1}^{[n^2t]} E[I_\Gamma(\exp((1/n)rA) - I)]Y_n((i-1)/n^2)) \end{aligned} \tag{1.8}$$

is a martingale. By Taylor's theorem

$$\begin{aligned} E[I_\Gamma(\exp(1/n)rA - I)] &= (1/2n^2)E[I_\Gamma r^2 A^2] + O(1/n^3) \\ &= (1/2n^2)E[I_\Gamma A^2] + O(1/n^3). \end{aligned} \tag{1.9}$$

Therefore,

$$M^\theta(\Gamma, t) = (\theta, Y_n(\Gamma, t) - Y_n(\Gamma, 0) \qquad (1.10)$$
$$- \tfrac{1}{2} \textstyle\sum_{i=1}^{[n^2t]} E[I_\Gamma A^2] Y_n((i-1)/n^2)(1/n^2)) + O(1/n).$$

By the same type of reasoning, letting $\mathscr{F}_i = \sigma\{r_1, \cdots, r_i, A_1, \cdots, A_i\}$, the quadratic variation process is

$$V_n^\theta(\Gamma, t) = \textstyle\sum_{i=1}^{[n^2t]} E[M_n^\theta(\Gamma, (i/n^2))^2 - M_n^\theta(\Gamma, (i-1)/n^2)^2 \mid \mathscr{F}_{i-1}] \qquad (1.11)$$
$$= \textstyle\sum_{i=1}^{[n^2t]} E[I_\Gamma(\theta, AY_n((i-1)/n^2))^2](1/n^2) + O(1/n).$$

Upon taking limits on the Y_n and changing Riemann sums to integrals, one may easily conjecture the following martingale problem:

Let $\{Y(\Gamma, t): \Gamma \in \mathscr{X}\}$ be a family of processes on $(C_B[0, \infty), P)$ with initial distributions $\mathscr{L}(Y(\Gamma, 0))$ satisfying:

(i) Y is additive in Γ, i.e., if $\Gamma_1 \cap \Gamma_2 = \phi$,

$$Y(\Gamma_1, t) + Y(\Gamma_2, t) = Y(\Gamma_1 \cup \Gamma_2, t) \quad \text{a.s.} \quad [P] \qquad (1.12)$$

(ii) Denote $Y(\Xi, t)$ by $Y(t)$. For a dense subspace of test functionals $\mathscr{D}' \subseteq B^*$,

$$M^\theta(\Gamma, t) = \left(\theta, Y(\Gamma, t) - Y(\Gamma, 0) - \frac{1}{2}\int_0^t \int_\Gamma A^2(\xi) Y(s) \mu(d\xi)\, ds\right) \qquad (1.13)$$

is a continuous orthogonal martingale measure for each $\theta \in \mathscr{D}'$.

(iii) The quadratic variation of $M^\theta(\Gamma, t)$ is

$$V^\theta(\Gamma, t) = \int_0^t \int_\Gamma (\theta, A(\xi) Y(s))^2 \mu(d\xi)\, ds. \qquad (1.14)$$

Some abuses of notation have appeared in this section. In particular, Y_n was used to denote the two quite different processes 1.1 and 1.6. However, if $\{Y_n\}$ in 1.1 give rise to a limit process Y via the theorems in [7], then $\{Y_n\}$ in 1.6 will give the same limit. To see this, all one must realize is that A appears in the martingale problem only with a squaring, and so the Rademacher function disappears.

The second presentation is in terms of the following stochastic integral equation:

$$Y(t) = Y(0) + \frac{1}{2}\int_0^t \int_\Xi A^2(\xi) Y(s) \mu(d\xi)\, ds \qquad (1.15)$$
$$+ \int_0^t \int_\Xi A(\xi) Y(s) W_\mu(d\xi, ds)$$

where W_μ is the unique (in law) orthogonal martingale measure having continuous sample paths and quadratic variation process $\mu(\Gamma)t$. In (1.15), we mean that equality holds for each real-valued process obtained by pairing this equation with an element in $\mathscr{D}'$. This interpretation permits us to view statements 1.13, 1.14 and 1.15 in a mild sense. For example, if $\cap \{\mathscr{D}(A^{*^2}(\xi)) : \xi \in \Xi\} \supseteq \mathscr{D}'$, then we can place the operators $A^*(\xi)$ and $A^{*^2}(\xi)$ on the elements of $\mathscr{D}'$.

The setting we have chosen employs orthogonal martingale measures, a concept developed by Gihman and Skorohod in [5]. Because these objects are uncommon, and because we have an entirely different purpose in mind, we present in Section 2 a synopsis of the relevant properties of orthogonal martingale measures and their associated stochastic integrals.

Theorem 3.1 is the statement of the equivalence of the two formulations for Y given above. Because the linearity plays no essential role, we drop that aspect in Section 3 for a slightly more general and slightly more workable form. The translation from this form to statements 1.13, 1.14, and 1.15 can be easily made.

At the conclusion of Section 3, we have an integral equation with which to work. Many properties of this integral equation can be addressed. However, we shall limit ourselves to questions of the existence and uniqueness of solutions in two cases of particular interest. Section 4 opens with a Stratonovich form, and with variation of parameters form for the stochastic integral equation 1.15. If $\{A(\xi) : \xi \in \Xi\}$ form a bounded set of bounded operators, then by the obvious variant of the stochastic version of the Picard iteration method any of these three forms can be shown to have a unique strong solution. Most of Section 4 is devoted to proving the following uniqueness theorem for the transport problem in $\mathbb{R}^d$:

THEOREM 4.3. *If the coefficient functions a in* $A(\xi) = a(\xi, x)\cdot\nabla$ *have* $a \in C^1(\mathbb{R}^d)$ *and* $\sup\{|(\partial a/\partial x_k)(\xi, x)| : x \in \mathbb{R}^d, k = 1, 2, \cdots, d, \xi \in \Xi\} < \infty$ *and if the closure of* $C = \frac{1}{2}\int_\Xi A^2(\xi)\mu(d\xi)$ *is uniformly elliptic and generates a semigroup, then the stochastic integral equation has the property of pathwise uniqueness.*

On the existence of solutions, we have

THEOREM 5.1. *If the coefficient functions* $a \in C^\infty(\mathbb{R}^d)$, *if for each* ξ *and each multi-index* α, $D^\alpha a(\xi, x)$ *is a Lipschitz function with a common Lipschitz constant, and if C is uniformly parabolic, then the variation of parameters form for the stochastic integral equation has a strong solution.*

2. Stochastic integrals over orthogonal martingale measures.

DEFINITION 2.1. [5]. A family of (locally) square integrable martingales $N(\Gamma, t)$, indexed by $\Gamma \in \mathscr{X}$, σ-algebra on Ξ, and adapted to the filtration $\{\mathscr{F}_t : t \geq 0\}$ is an *orthogonal (local) martingale measure* if the following conditions are satisfied for all Γ, Γ_1, Γ_2 in $\mathscr{X}$ and all $t \geq 0$:

(i) (additivity)

$$N(\Gamma_1, t) + N(\Gamma_2, t) = N(\Gamma_1 \cup \Gamma_2, t) \quad \text{a.s. for} \quad \Gamma_1 \cap \Gamma_2 = \varnothing. \tag{2.1}$$

(ii) (orthogonality)

$$N(\Gamma_1, t)N(\Gamma_2, t) \quad \text{is a (local) martingale for} \quad \Gamma_1 \cap \Gamma_2 = \varnothing. \tag{2.2}$$

(2.3) (iii) $$\langle N(\Gamma, t), N(\Gamma, t)\rangle = \pi(\Gamma, t)$$

where $\pi(\Gamma, t)$ is a random function which, for fixed t, is a measure on $\mathscr{X}$ with probability one, and for fixed Γ, is a continuous monotonically increasing function of t.

REMARK 2.2. (i) $\pi(\Gamma, t)$ is called the *quadratic variation* of the martingale measure $N(\Gamma, t)$. If we additionally require that $\pi(\Gamma, 0) = 0$ for all $\Gamma \in \mathscr{X}$, then π is unique.

(ii) For $\Gamma_1, \Gamma_2 \in \mathscr{X}$,

(2.4) $$\langle N(\Gamma_1, t), N(\Gamma_2, t)\rangle = \pi(\Gamma_1 \cap \Gamma_2, t)$$

(iii) If $\Gamma_1 \cap \Gamma_2 = \varnothing$, then

(2.5) $$\pi(\Gamma_1, t) + \pi(\Gamma_2, t) = \pi(\Gamma_1 \cup \Gamma_2, t)$$

Therefore π is additive and the definition of π is consistent in Γ.

EXAMPLE 2.3. (i) If $\mathscr{X}$ has atoms $\{\Gamma_1, \Gamma_2, \cdots, \Gamma_n\}$, then $\{N(\Gamma_1), N(\Gamma_2), \cdots, N(\Gamma_n)\}$ form a finite family of orthogonal martingales. Conversely, any finite family of orthogonal martingales can be viewed as a martingale measure.

(ii) Let X be Gaussian white noise measure on the positive quadrant in R^2. Thus for a Borel set Σ, $X(\Sigma)$ is a normal random variable with mean zero and variance equal to the area of Σ. In addition, if $\Sigma_1 \cap \Sigma_2 = \varnothing$, $X(\Sigma_1)$ and $X(\Sigma_2)$ are independent. Define $N(\Gamma, t) = X(\Gamma \times [0, t])$ where Γ is a Borel set in $[0, T]$. If $\{\mathscr{F}_{st}: 0 \le s \le T, t \ge 0\}$ is the filtration for X, then N is a martingale measure with respect to the filtration $\{\mathscr{F}_t: t \ge 0\}$ where $\mathscr{F}_t = \sigma\{\mathscr{F}_{st}: 0 \le s \le T\} = \mathscr{F}_{Tt}$. In this situation the quadratic variation process is $tm(\Gamma)$ where m is Lebesgue measure.

(iii) Let μ be a probability measure on $\mathscr{X}$, and set $\pi(\Gamma, t) = t\mu(\Gamma)$. Upon appealing to Lévy's characterization of Brownian motion, one sees that the associated martingale measure $N(\Gamma, t)$ having continuous sample paths is Brownian motion with variance $\mu(\Gamma)$. For the class of random evolutions presented in this paper, the limiting process will be represented as the solution to an integral equation with respect to this martingale measure. Because of its particular interest, we shall denote this martingale measure $W(\Gamma, t)$ or $W_\mu(\Gamma, t)$ whenever the measure μ needs to be emphasized.

Once a martingale on a probability space $(\Omega, \mathscr{F}, P)$ with a filtration $\{\mathscr{F}_t: t \ge 0\}$ has been specified, the construction of a stochastic integral follows three major stages of development. First, one must find some suitable class $\mathscr{E}$ of elementary functions. For these functions, there is, generally speaking, only one way to make a definition worthy of the term "integral". Then one extends by linearity to a class of simple functions $\mathscr{S}$. At this stage, one defines the significant properties of the stochastic integral. For simple functions, these properties follow routinely from the definition. In the last stage, the quadratic variation π of the martingale comes to the front. At the second stage, one of the properties is an L^2-isometry between squares of functions on the probability space and $L^2(\pi)$. This isometry is exploited and the stochastic integral is defined on the completion of $\mathscr{S}$ in the metric induced by the isometry. Via a continuity argument, the properties verified in the middle stage hold equally well in the final stage. The construction for integrals over martingale measures follows the lines of many others (cf. [1], [4], or [3]).

With this outline in mind, we can begin. Let N be a continuous square integrable martingale measure with quadratic variation π, and let $\mathscr{E}$ denote the class of functions of the form $xI_{\Delta\times\Gamma}$, where $\Gamma \in \mathscr{X}$, $\Delta = (t, t']$, and x is a bounded $\mathscr{F}_t$-measurable random variable. Set

$$\int\int xI_{\Delta\times\Gamma}N(d\xi, ds) = x(N(\Gamma, t') - N(\Gamma, t)). \tag{2.6}$$

The simple functions $\mathscr{S}$ are finite linear combinations of elements of $\mathscr{E}$. Each function in $\mathscr{S}$ may be written

$$X = \textstyle\sum_{k=1}^{N} \sum_{j=1}^{M} x_{jk} I_{\Gamma_{jk}\times\Delta_k} \tag{2.7}$$

where $0 \leq t_1 < t_2 < \cdots < t_N$, $\Delta_k = (t_{k-1}, t_k]$, $\Gamma_{jk} \in \mathscr{X}$, and x_{jk} is a bounded $\mathscr{F}_{t_{k-1}}$-measurable random variable. Because we are interested in using stochastic integrals to represent new stochastic processes, we invoke the following notation:

$$\begin{aligned}\int_t^{t'}\int_\Gamma X dN &= \int_t^{t'}\int_\Gamma X(\xi, s)N(d\xi, ds) \\ &= \int\int I_{[t,t')\times\Gamma}\, X(\xi, s)N(d\xi, ds).\end{aligned} \tag{2.8}$$

Proposition 2.4. *If $X \in \mathscr{S}$, and $\Gamma \in \mathscr{X}$, then*

$$M(\Gamma, t) = \int_0^t \int_\Gamma X(\xi, s)N(d\xi, ds) \tag{2.9}$$

defines a continuous square integrable martingale measure. In addition,

(i) (*linearity*) *if $X_1, X_2 \in \mathscr{S}$ and $c_1, c_2 \in \mathbb{R}$, then*

$$\begin{aligned}&\int\int (c_1X_1(\xi, s) + c_2X_2(\xi, s))N(d\xi, ds) \\ &\quad = c_1\int\int X_1(\xi, s)N(d\xi, ds) + c_2\int\int X_2(\xi, s)N(d\xi, ds).\end{aligned} \tag{2.10}$$

(ii)

$$\begin{aligned}&E\left[\left(\int\int X_1(\xi, s)N(d\xi, ds)\right)\left(\int\int X_2(\xi, s)N(d\xi, ds)\right)\right] \\ &\quad = E\left[\int\int (X_1(\xi, s) - X_2(\xi, s))^2\pi(d\xi, ds)\right].\end{aligned} \tag{2.11}$$

(iii) *In particular,*

$$E\left[\left(\int\int X(\xi, s)N(d\xi, ds)\right)^2\right] = E\left[\int\int X^2(\xi, s)\pi(d\xi, ds)\right]. \tag{2.12}$$

This is the promised L^2-isometry.

(iv)
$$E\left[\left(\int\int X_1(\xi, s)N(d\xi, ds) - \int\int X_2(\xi, s)N(d\xi, ds)\right)^2\right] = E\left[\int\int (X_1(\xi, s) - X_2(\xi, s))^2\pi(d\xi, ds)\right]. \tag{2.13}$$

(v) *For $\Gamma_1, \Gamma_2 \in X$, let*

$$I_j(\Gamma, t) = \int_0^t \int_\Gamma X_j(\xi, s)N(d\xi, ds), \quad j = 1, 2$$

then

$$\langle I_1(\Gamma_1, t), I_2(\Gamma_2, t)\rangle = \int_0^t \int_{\Gamma_1\cap\Gamma_2} X_1(\xi, s)X_2(\xi, s)\pi(d\xi, ds). \tag{2.14}$$

(vi) *For an $\mathscr{F}_t$-stopping time τ,*

$$\int_0^{t\wedge\tau} \int_\Gamma X(\xi, s)N(d\xi, ds) = \int_0^t \int_\Gamma X(\xi, s)N^\tau(d\xi, ds) \tag{2.15}$$

where

$$N^\tau(\Gamma, t) = N(\Gamma, t \wedge \tau). \tag{2.16}$$

PROOF. Use the definitions. □

We are now ready to extend the definition to a still wider class of random variables. Let

$$\mathscr{H} = \left\{X = \{X(\xi, t): t \geq 0\}: X \text{ is progressively measurable and } E\left[\int_0^t \int_\Xi X^2(\xi, t)\pi(d\xi, ds)\right] < \infty \text{ for all } t \geq 0\right\}.$$

The proposition above implies that $\mathscr{S} \subseteq \mathscr{H}$. Let $\bar{\mathscr{S}}$ denote the completion of $\mathscr{S}$ in the isometry stated in equation 2.12. Since $\mathscr{H}$ is a closed subspace, $\bar{\mathscr{S}} \subseteq \mathscr{H}$, but we would like $\bar{\mathscr{S}} = \mathscr{H}$. This is the essence of the following lemma.

LEMMA 2.5. *Let $X \in \mathscr{H}$, then there exists a sequence $\{X_n\} \subseteq \mathscr{S}$ such that for all $t \geq 0$*

$$\lim_{n\to\infty} E\left[\int\int (X_n(\xi, s) - X(\xi, s))^2\pi(d\xi, ds)\right] = 0. \tag{2.17}$$

Let $X \in \mathscr{H}$, and choose a sequence $\{X_n\} \subseteq \mathscr{S}$ that fulfills Lemma 2.5, then by equation 2.13,

$$\left\{\int_0^t \int_\Xi X_n(\xi, s) N(d\xi, ds)\right\} \tag{2.18}$$

is a Cauchy sequence in $L^2(P)$ for each t. Thus we specify the limiting process, denoted $\int_0^t \int_\Xi X(\xi, s)N(d\xi, ds)$ for each $t > 0$ as an equivalence class in $L^2(P)$. However we can do better.

THEOREM 2.6. *$\int_0^t \int_\Gamma X(\xi, s)N(d\xi, ds)$ has a continuous version.*

In the future, the term stochastic integral will mean a continuous version of this process.

THEOREM 2.7. *Proposition 2.4 holds with $\mathscr{S}$ replaced by $\mathscr{H}$.*

One further generalization is possible. Now, let $N(\Gamma, t)$ be a local martingale measure. We shall be able to integrate over N, the processes

$\mathscr{L} = \{X(\xi, s)$: X is progressively measurable and

$\int_0^t \int_\Xi X^2(\xi, s)\pi(d\xi, ds) < \infty$ a.s. for all $t > 0\}$.

Let

$$\sigma_n = \inf\left\{t \geq 0: \int_0^t \int_\Xi X^2(\xi, s)\pi(d\xi, ds) \geq n\right\} \tag{2.19}$$

where $\inf \varnothing = \infty$. Then $\lim_{n\to\infty}\sigma_n = \infty$ a.s. If τ_n' is a sequence that arises in defining the local martingale measure N, then $\tau_n = \sigma_n \wedge \tau_n'$ will serve equally well. $N^{\tau_n}(\Gamma, t) = N(\Gamma, t \wedge \tau_n)$ is a continuous square integrable martingale on the filtration $\{\mathscr{F}_{t\wedge\tau_n}: t \geq 0\}$, with quadratic variation $\pi^{\tau_n}(\Gamma, t) = \pi(\Gamma, t \wedge \tau_n)$. Set, for $X \in \mathscr{L}$

$$M_n(\Gamma, t) = \int_0^t \int_\Gamma X(\xi, s)N^{\tau_n}(d\xi, ds) = \int_0^{t\wedge\tau_n} \int_\Gamma X(\xi, s)N(d\xi, ds). \tag{2.20}$$

$X(\xi, s \wedge \tau_n) \in \mathscr{H}$ for N^{τ_n} by equation 2.19 and the fact that $\tau_n \leq \sigma_n$. So, for each n, $M_n(\Gamma, t)$ has been defined. On the set $\{\tau_n > t\}$, $M_{n'}(\Gamma, t) - M_n(\Gamma, t) = 0$ a.s. whenever $n' > n$. In other words, for almost every $\omega \in \Omega$, there exists n so that $M_{n'}(\Gamma, t) = M_n(\Gamma, t)$ for $n' > n$. Thus we can define $M(\Gamma, t) = \lim_{n\to\infty} M_n(\Gamma, t)$. By a standard martingale inequality, $M_n(\Gamma, t) \to M_n(\Gamma, t)$ in probability uniformly in finite intervals. In summary we have:

THEOREM 2.8. *Let $N(\Gamma, t)$ be a local martingale measure with quadratic variation $\pi(\Gamma, t)$ and $X \in \mathscr{L}$. Then*

(i) *$M(\Gamma, t) = \int_0^t \int_\Gamma X(\xi, s)N(d\xi, ds)$ is a continuous locally square integrable martingale measure with continuous quadratic variation*

$$\langle M(\Gamma, t), M(\Gamma, t)\rangle = \int_0^t \int_\Gamma X^2(\xi, s)\pi(d\xi, ds). \tag{2.21}$$

(ii) *If $X_1, X_2 \in \mathscr{L}$ and $c_1, c_3 \in \mathbb{R}$, then*

$$\int\int (c_1X_1(\xi, s) + c_2X_2(\xi, s))N(d\xi, ds) = c_1 \int\int X_1(\xi, s)N(d\xi, ds) + c_2 \int\int X_2(\xi, s)N(d\xi, ds). \tag{2.22}$$

(2.23) (iii) $$\int_0^{t\wedge\tau} \int_\Gamma X_1(\xi, s)N(d\xi, ds) = \int_0^t \int_\Gamma X_1(\xi, s)N^\tau(d\xi, ds)$$

where $N^\tau(\Gamma, t) = N(\Gamma, t \wedge \tau)$ and τ is an $\mathscr{F}_t$-stopping time.

DEFINITION 2.9. Let N_1 and N_2 be two local martingale measures. Then the random process

$$\begin{aligned}\langle N_1(\Gamma, t), N_2(\Gamma, t)\rangle &= \tfrac{1}{2}\{\langle N_1(\Gamma, t) + N_2(\Gamma, t), N_1(\Gamma, t) + N_2(\Gamma, t)\rangle \\ &\quad - \langle N_1(\Gamma, t), N_1(\Gamma, t)\rangle - \langle N_2(\Gamma, t), N_2(\Gamma, t)\rangle\}\end{aligned} \tag{2.24}$$

is called the *covariation* of N_1 and N_2.

We shall refer to equation 2.24 as the polarization identity. The usefulness of this notation is due to the fact that

$$N_1(\Gamma, t)N_2(\Gamma, t) - \langle N_1(\Gamma, t), N_2(\Gamma, t)\rangle \tag{2.25}$$

is a martingale for every Γ. In fact, $\langle N_1(\Gamma, t), N_2(\Gamma, t)\rangle$ is the unique process of bounded variation that is zero a.s. at time zero and makes equation 2.25 a martingale. The major difference between this set up and the usual set up with martingales is that

$$\langle N_1(\cdot, t), N_2(\cdot, t)\rangle \tag{2.26}$$

admits a representation as the difference of two measures on $\mathscr{X} \times \mathscr{X}$.

THEOREM 2.10. *Let N_1 and N_2 be local martingale measures with quadratic variation π_1 and π_2 respectively. Let X_i be in $\mathscr{L}$ for π_i, $i = 1, 2$. Then*

$$\left(\int_0^t \int_\Gamma |X_1(\xi, s)X_2(\xi, s)| \langle N_1(d\xi, ds), N_2(d\xi, ds)\rangle\right)^2 \le \left(\int_0^t \int_\Gamma X_1^2(\xi, s)\pi_1(d\xi, ds)\right)\left(\int_0^t \int_\Gamma X_2^2(\xi, s)\pi_2(d\xi, ds)\right) \tag{2.27}$$

and

$$\left\langle \int_0^t \int_\Gamma X_1(\xi, s)N_1(d\xi, ds), \int_0^t \int_\Gamma X_2(\xi, s)N_2(d\xi, ds)\right\rangle = \int_0^t \int_\Gamma X_1(\xi, s)X_2(\xi, s)\langle N_1(d\xi, ds), N_2(d\xi, ds)\rangle \tag{2.28}$$

THEOREM 2.11. *Let X_1, X_2, and their product be elements of $\mathscr{L}$. Set*

$$M(\Gamma, t) = \int_0^t \int_\Gamma X_2(\xi, s)N(d\xi, ds) \tag{2.29}$$

then

$$\int_0^t \int_\Gamma X_1(\xi, s)M(d\xi, ds) = \int_0^t \int_\Gamma X_1(\xi, s)X_2(\xi, s)N(d\xi, ds). \tag{2.30}$$

THEOREM 2.12. *Let $\pi(d\xi, ds)$ be a positive measure on $\Xi \times [0, \infty)$, and denote $\pi(\Gamma, t) = \int_0^t \int_\Gamma \pi(d\xi, ds)$. Then there exists a process Z, unique in distribution, such that for each $\Gamma \in \mathscr{X}$*

(i) *$Z(\Gamma, \cdot)$ has sample paths in $C([0, \infty), \mathbb{R})$,*
(ii) *$Z(\Gamma, \cdot)$ is a martingale, and*
(iii) *$Z^2(\Gamma, \cdot) - \pi(\Gamma, \cdot)$ is a martingale.*

In addition, Z is a martingale measure.

PROOF. A white noise based on π satisfies (i)–(iii), and therefore such a process Z exists. As for uniqueness, let $p \in \mathbb{R}$ and define $f:[0, \infty) \times \mathbb{R} \to \mathbb{C}$ by

$$f(t, x) = \exp(ipx + \tfrac{1}{2}p^2\pi(\Gamma, t)) \tag{2.31}$$

Then by the Itô formula $f(t, Z(\Gamma, t))$ is a martingale. In particular,

$$E[\exp(ip(Z(\Gamma, t) - Z(\Gamma, s)))\,|\,\mathscr{F}_s] = \exp((-\tfrac{1}{2})p^2(\pi(\Gamma, t) - \pi(\Gamma, s))) \tag{2.32}$$

Therefore $Z(\Gamma, \cdot)$ has independent Gaussian increments. This determines the finite dimensional distributions of $Z(\Gamma, \cdot)$, and guarantees us that $Z(\Gamma, \cdot)$ is unique in distribution.

Let's check that Z is a martingale measure. For Γ_1 and Γ_2 disjoint, both $Z(\Gamma_1 \cup T_2, \cdot)$ and $Z(\Gamma_1, \cdot) + Z(\Gamma_2, \cdot)$ satisfy properties (i), (ii), and (iii) of the theorem, and hence must be equal in distribution. Orthogonality is a consequence of equation 2.4.

$$\langle Z(\Gamma, t), Z(\Gamma, t)\rangle = \pi(\Gamma, t)$$

is just a restatement of property (iii). □

To this point, we have developed the stochastic integral with an eye to the Itô integral. In a similar fashion, we may begin with a symmetric approximating sum and follow the Stratonovich development of the integral. Let N be an orthogonal martingale measure on Ξ. If X is a progressively measurable function and $X(\xi, \cdot)$ is a fixed function on the disjoint sets $\Gamma_1, \cdots, \Gamma_n$, then we may operate by direct analogy.

DEFINITION 2.13

$$\int_0^t \int_\Xi X(\xi, s) \circ N(d\xi, ds) = \textstyle\sum_{j=1}^n \displaystyle\int_0^t X(\xi_j, s) \circ N(\Gamma_j, ds) \tag{2.33}$$

where $\xi_j \in \Gamma_j$.

The left side of the equation is thus defined by the right. The raised small circle on the right denotes the usual Stratonovich integral with respect to martingales. This definition now serves as a basis for an approximation scheme for more general X. For example, from this definition we can pass to

$$X(\xi, \cdot) \quad \text{is a fixed function on the disjoint sets} \quad \Gamma_{1,k}, \cdots \Gamma_{n_k,k}$$

whenever $t_k \leq s < t_{k+1}$, and then on to more general X. Any relevant formula which show the relationship of the two types of integral can be verified using 2.33 and shown to hold in the passage to the limit.

3. The martingale problem and the stochastic integral equation. Let $\pi(\Gamma, t)$ be a positive measure on $(\Xi \times [0, \infty), \mathscr{X} \times \mathscr{B}[0, \infty))$, and let N be the continuous process for which $\langle N(\Gamma, t), N(\Gamma, t)\rangle = \pi(\Gamma, t)$. Let $a, b: \Xi \times B \rightarrow B$ be measurable functions.

The Martingale Problem.

Let $\{Y(\Gamma, t): \Gamma \in \mathscr{X}\}$ be a family of processes on $(C_B[0, \infty), P)$ with initial distributions $\mathscr{L}(Y(\Gamma, 0))$ satisfying

(i) Y is additive in Γ, i.e., if $\Gamma_1 \cap \Gamma_2 = \varnothing$,

$$Y(\Gamma_1, t) + Y(\Gamma_2, t) = Y(\Gamma_1 \cup \Gamma_2, t) \text{ a.s. } [P].$$

(ii) Denote $Y(\Xi, t)$ by $Y(t)$. For a dense subspace of test functionals $\mathscr{D}' \subseteq B^*$

$$M^\theta(\Gamma, t) = \left(\theta, Y(\Gamma, t) - Y(\Gamma, 0) - \frac{1}{2}\int_0^t \int_\Gamma a(\xi, Y(s))\mu(d\xi)\, ds\right) \tag{3.1}$$

is a continuous orthogonal martingale measure for each $\theta \in \mathscr{D}'$.

(iii) The quadratic variation of $M^\theta(\Gamma, t)$ is

$$V^\theta(\Gamma, t) = \int_0^t \int_\Gamma (\theta, b(\xi, Y(s)))^2 \pi(d\xi, ds). \tag{3.2}$$

The Stochastic Integral Equation.

$$Y(t) = Y(0) + \int_0^t \int_\Xi a(\xi, Y(s))\pi(d\xi, ds) + \int_0^t \int_\Gamma b(\xi, Y(s))N(d\xi, ds) \tag{3.3}$$

where 3.3 is read after pairing both sides of the equation with an element from $\mathscr{D}'$. Again, we can view 3.1, 3.2, and 3.3 in any sort of mild sense as long as we are consistent.

Assume the existence of the process in equation 3.3 and define

$$Y(\Gamma, 0) = \nu(\Gamma)Y(0), \quad \text{where } \nu \text{ is some probability measure,}$$

$$Y(\Gamma, t) = Y(\Gamma, 0) + \int_0^t \int_\Gamma a(\xi, Y(s))\pi(d\xi, ds) + \int_0^t \int_\Gamma b(\xi, Y(s))N(d\xi, ds). \tag{3.4}$$

By the basic properties of stochastic integrals, $M^\theta(\Gamma, t)$ is a martingale with quadratic variation $V^\theta(\Gamma, t)$. A progressively measurable process $Y(\Gamma, t)$ with initial distribution $\mathscr{L}(Y(\Gamma, 0))$ is said to be a solution to the martingale problem if 3.1 and 3.2 hold with respect to the measure P and the filtration $\mathscr{F}_t = \sigma\{Y(s, \Gamma): \Gamma \in \mathscr{X}, s \le t\}$. A martingale problem is said to be well posed if there exists a solution and every solution has the same finite dimensional distributions. Thus one easily sees that existence of a solution to the stochastic integral equation in 3.3 gives a solution to the martingale problem. In addition, if the martingale problem is well posed, then the solution to the integral equation is unique. This gives us one direction in each of the following two statements.

THEOREM 3.1. (i) *The stochastic integral equation has a solution if and only if the martingale problem has a solution.*

(ii) *The solution to the stochastic integral equation is unique if and only if the martingale problem is well posed.*

The interest comes in the converse. As stated earlier, the stochastic integral equation involves a process N. If we wish to construct a process that behaves as N, then we must build it from the processes M and V, for they are all we have. The construction follows a familiar line. First, a process $Z(\Gamma, t)$ is defined as a stochastic integral over M. By the definition, Z will be a martingale measure. Second, we show that $\langle Z(\Gamma, t), Z(\Gamma, t)\rangle = \pi(\Gamma, t)$. By Theorem 2.12, this guarantees us that Z and N have the same distribution. The finishing touch is to show that $Y(t)$ solves the stochastic integral equation.

Before we begin, we must be precise in defining stochastic integrals over M. M is not a real valued process, and so it does not fall in line with the theorems in Section 2. However for fixed $\theta \in \mathscr{D}'$, $M^\theta(\Gamma, t) = (\theta, M(\Gamma, t))$ is a real valued martingale measure with quadratic variation process $V^\theta(\Gamma, t)$, and we can apply the results of Section 2 to this situation. If $\eta\colon \Xi \times B \to B^*$ is a simple function, then we may define the stochastic integral $\int \int (\eta(\xi, y), M(d\xi, ds))$ by using the corresponding L^2-isometries piece by piece on the sets where η is constant. Next we can define the integral for any measurable η by approximation to simple functions. Finally we may replace y by an adapted process $Y(s)$. In each instance we use the martingales

$$\left[\int_0^t \int_\Gamma (\eta(\xi, Y(s)), M(d\xi, ds))\right]^2 - \int_0^t \int_\Gamma (\eta(\xi, Y(s)), b(\xi, Y(s)))^2 \pi(d\xi, ds)$$

as the key tool in the extension. Now we go on to construct an η that will be particularly useful.

PROOF OF THEOREM 3.1. If b in equation 3.2 is never zero, then the $M(\Gamma, t)$ and $V(\Gamma, t)$ provide us with enough "randomness" to recover N. To this end, assume that Y is a continuous solution to the martingale problem, and define $\eta\colon \Xi \times B \to B^*$ in the following way:

First, let $(\eta(\xi, y), b(\xi, y)) = 1$. This is always possible since $b(\xi, y)$ is never zero. Then extend $\eta(\xi, y)$ linearly on the subspace generated by $b(\xi, y)$. Further extend η to be defined as a linear functional on all of B in a such a way that the norm is not increased. The Hahn-Banach theorem promises us that such a linear functional exists for each (ξ, y), but it is only an existence theorem, and as such, it does not guarantee us that η is measurable. This is a technical point. Nevertheless, let's satisfy ourselves that a measurable η exists.

If b were a simple function, then η could be chosen to be a constant on the sets where b is constant. In this case, η, being a simple function, is certainly measurable. (Here a simple function can be a countable sum.)

Because B is separable, we can choose a sequence of simple functions $\{b_n\}$ converging to b in norm so that

$$\| b_n(\xi, y) \| \geq \| b(\xi, y) \|$$

a.s. and

$$\| b_n(\xi, y) - b_{n+1}(\xi, y) \| \leq 4^{-n} \| b(\xi, y) \|. \tag{3.5}$$

Define a measurable mapping $\eta_1 \in \mathscr{D}'$ so that

$$(\eta_1(\xi, y), b_1(\xi, y)) = 1 \tag{3.6}$$

and extend η_1 without increasing norm. Having defined $\eta_1, \eta_2, \cdots, \eta_n$, define $\eta_{n+1} \in \mathscr{D}'$ by

$$(\eta_{n+1}(\xi, y), b_{n+1}(\xi, y)) = 1 - \textstyle\sum_{k=1}^{n} (\eta_k(\xi, y), b_{n+1}(\xi, y)) \tag{3.7}$$

and extend η_{n+1} without increasing norm. Observe the following:

$$\begin{aligned} &\| \eta_{n+1}(\xi, y) \| \; \| b_{n+1}(\xi, y) \| \\ &\qquad = (\eta_{n+1}(\xi, y), b_{n+1}(\xi, y)) = 1 - \textstyle\sum_{k=1}^{n} (\eta_k(\xi, y), b_{n+1}(\xi, y)) \\ &\qquad = \textstyle\sum_{k=1}^{n} (\eta_k(\xi, y), b_n(\xi, y)) - \textstyle\sum_{k=1}^{n} (\eta_k(\xi, y), b_{n+1}(\xi, y)) \\ &\qquad = \textstyle\sum_{k=1}^{n} (\eta_k(\xi, y), b_n(\xi, y) - b_{n+1}(\xi, y)) \\ &\qquad \leq \textstyle\sum_{k=1}^{n} \| \eta_k(\xi, y) \| \; \| b_n(\xi, y) - b_{n+1}(\xi, y) \|. \end{aligned} \tag{3.8}$$

Therefore since $\| b_n(\xi, y) - b_{n+1}(\xi, y) \| \leq 4^{-n} \| b_{n+1}(\xi, y) \|$

$$\| \eta_{n+1}(\xi, y) \| \leq (\textstyle\sum_{k=1}^{n} \| \eta_k(\xi, y) \|) 4^{-n}. \tag{3.9}$$

By an easy induction argument $\| \eta_{n+1}(\xi, y) \| \leq 3^{-n} \| \eta_1(\xi, y) \|$. This shows that $\eta = \sum_{k=1}^{\infty} \eta_k$ is a measurable function from $\Xi \times B$ to B^*. Moreover

$$\begin{aligned} &(\eta(\xi, y), b(\xi, y)) \\ &\qquad = \lim_{n\to\infty} (\eta(\xi, y), b_n(\xi, y)) \\ &\qquad = \lim_{n\to\infty} \textstyle\sum_{k=1}^{n} (\eta_k(\xi, y), b_n(\xi, y)) + \textstyle\sum_{k=n+1}^{\infty} (\eta_k(\xi, y), b_n(\xi, y)) \\ &\qquad = \lim_{n\to\infty} 1 + \textstyle\sum_{k=n+1}^{\infty} (\eta_k(\xi, y), b_n(\xi, y)) = 1. \end{aligned} \tag{3.10}$$

With η in hand, we can define Z by

$$Z(\Gamma, t) = \int_0^t \int_\Gamma (\eta(\xi, Y(s)), M(d\xi, ds)). \tag{3.11}$$

Let's check that Z satisfies the axioms for a martingale measure with respect to the filtration $\mathscr{F}_t = \sigma\{Y(s, \Gamma), \Gamma \in \mathscr{X}, s \le t\}$. First of all, $Z(\Gamma, \cdot)$ is a martingale. Note that it is additive. Also,

$$\begin{aligned}\langle Z(\Gamma_1, t), Z(\Gamma_2, t)\rangle &= \int_0^t \int_{\Gamma_1\cap\Gamma_2} (\eta(\xi, Y(s)), b(\xi, Y(s)))^2 \pi(d\xi, ds) \\ &= \pi(\Gamma_1 \cap \Gamma_2, t).\end{aligned} \tag{3.12}$$

Therefore $\langle Z(\Gamma, t), Z(\Gamma, t)\rangle = \pi(\Gamma, t)$, and $Z(\Gamma_1, t)Z(\Gamma_2, t)$ is a martingale if Γ_1 and Γ_2 are disjoint. Since the form for π is sufficiently restricted to determine the distribution, we have $\mathscr{L}(N) = \mathscr{L}(Z)$. Consequently by the change of variables formula

$$\int_0^t \int_\Xi b(\xi, Y(s))Z(d\xi, ds) = \int_0^t b(\xi, Y(s))(\eta(\xi, Y(s)), M(d\xi, ds)).$$

If

$$\begin{aligned}&\int_0^t b(\xi, Y(s))(\eta(\xi, Y(s)), M(d\xi, ds)) \\ &\qquad = \int_0^t (\eta(\xi, Y(s)), b(\xi, Y(s)))M(d\xi, ds)\end{aligned} \tag{3.13}$$

then

$$\begin{aligned}\int_0^t \int_\Xi b(\xi, Y(s))Z(d\xi, ds) &= \int_0^t \int_\Xi M(d\xi, ds) = M(\Xi, t) \\ &= Y(t) - Y(0) - \int_0^t \int_\Xi a(\xi, Y(s))\pi(d\xi, ds)\end{aligned}$$

and the stochastic integral equation holds. Let's check equation 3.13. It is sufficient to show that the square of the difference of the two sides has zero expectation. This is the plan of attack. Let $\theta \in \mathscr{D}'$, then

$$\begin{aligned}&E\left[\int_0^t \int_\Xi (\theta, b(\xi, Y(s)))(\eta(\xi, Y(s)), M(d\xi, ds)) \right.\\ &\qquad\qquad \left. - \int_0^t \int_\Xi (\eta(\xi, Y(s)), b(\xi, Y(s)))(\theta, M(d\xi, ds))^2\right] \\ &\quad = E\left[\int_0^t \int_\Xi (\theta, b(\xi, Y(s)))(\eta(\xi, Y(s)), M(d\xi, ds))^2\right] \\ &\qquad - 2E\left[\int_0^t \int_\Xi (\theta, b(\xi, Y(s)))(\eta(\xi, Y(s)), M(d\xi, ds)) \right.\\ &\qquad\qquad \left. \times \int_0^t \int_\Xi (\eta(\xi, Y(s)), b(\xi, Y(s)))(\theta, M(d\xi, ds))\right] \\ &\qquad + E\left[\int_0^t \int_\Xi (\eta(\xi, Y(s)), b(\xi, Y(s)))(\theta, M(d\xi, ds))^2\right].\end{aligned} \tag{3.14}$$

The first and third terms both equal

$$E \int_0^t \int_\Xi (\eta(\xi, Y(s)), b(\xi, Y(s)))(\theta, b(\xi, Y(s)))\pi(d\xi, ds). \tag{3.15}$$

To finish we show that the expectation in the second term also equals 3.15. This term may be written

$$E\left[\int_0^t \int_\Xi (\theta, b(\xi, Y(s)))(\eta(\xi, Y(s)), b(\xi, Y(s))) \times \langle(\eta(\xi, Y(s)), M(d\xi, ds)), (\theta, M(d\xi, ds))\rangle\right] \tag{3.16}$$

By the polarization identity for the covariance process

$$\begin{aligned}
&\langle(\eta(\xi, Y(s)), M(d\xi, ds)), (\theta, M(d\xi, ds))\rangle \\
&\quad = \tfrac{1}{2}\{\langle(\eta(\xi, Y(s)) + \theta, M(d\xi, ds)), (\eta(\xi, Y(s)) + \theta, M(d\xi, ds))\rangle \\
&\qquad - \langle(\eta(\xi, Y(s)), M(d\xi, ds)),(\eta(\xi, Y(s)), M(d\xi, ds))\rangle \\
&\qquad - \langle(\theta, M(d\xi, ds)), (\theta, M(d\xi, ds))\rangle\} \\
&\quad = \tfrac{1}{2}\{(\eta(\xi, Y(s)) + \theta, b(\xi, Y(s)))^2\pi(d\xi, ds) \\
&\qquad - (\eta(\xi, Y(s)), b(\xi, Y(s)))^2\pi(d\xi, ds) - (\theta, b(\xi, Y(s)))^2\pi(d\xi, ds)\} \\
&\quad = (\eta(\xi, Y(s)), b(\xi, Y(s)))(\theta, b(\xi, Y(s)))\pi(d\xi, ds).
\end{aligned} \tag{3.17}$$

Upon substituting the result in 3.17 into 3.16, one sees that, indeed, the second term is equal to 3.15.

If b is sometimes zero, the processes M and V may lack the randomness necessary to recover N. In this case we shall need more. To this end, let N' be a continuous martingale measure on a space $(\Omega', \mathscr{F}', P')$ with respect to the filtration $\mathscr{F}_t'$. N' has quadratic variation π. Define $Y'(\Gamma, t, \omega, \omega') = Y(\Gamma, t, \omega)$ and $N'(\Gamma, t, \omega, \omega') = N'(\Gamma, t, \omega')$. Thus Y' is additive in Γ. On the augmented probability space $(\Omega \times \Omega', \mathscr{F} \times \mathscr{F}', P \times P')$ with filtration $\mathscr{F}_t \times \mathscr{F}_t$

$$M'^{\theta}(\Gamma, t) = \left(\theta, Y'(\Gamma, t) - Y'(\Gamma, 0) - \int_0^t \int_\Gamma a(\xi, Y'(s))\pi(d\xi, ds)\right) \tag{3.18}$$

is a martingale measure for each $\theta \in \mathscr{D}'$. The quadratic variation of $M'^{\theta}(\Gamma, t)$ is

$$V'^{\theta}(\Gamma, t) = \int_0^t \int_\Gamma (\theta, b(\xi, Y'(s)))^2\pi(d\xi, ds). \tag{3.19}$$

Define $\eta'\colon \Xi \times B \to B^*$ in the following way:

If $b(\xi, y) = 0$, set $\eta'(\xi, y)$ be the zero linear functional. Otherwise, $b(\xi, y) \neq 0$ and η' may be defined in exactly the same manner that η is defined. Let $\rho(\xi, y) = I_{\{b(\xi,y)=0\}}$. As before, we check that,

$$Z'(\Gamma, t) = \int_0^t \int_\Gamma (\eta(\xi, Y'(s), M'(d\xi, ds)) + \int_0^t \int_\Gamma \rho(\xi, Y'(s))N'(d\xi, ds) \tag{3.20}$$

is a martingale measure. Z' is a martingale, additive in Γ,

$$(3.21)\quad \begin{aligned}\langle Z'(\Gamma_1, t), Z'(\Gamma_2, t)\rangle &= \int_0^t \int_{\Gamma_1\cap\Gamma_2} (\eta'(\xi, Y'(s)), b(\xi, Y'(s)))^2 \pi(d\xi, ds)\\ &\quad + \int_0^t \int_{\Gamma_1\cap\Gamma_2} \rho^2(\xi, Y'(s))\pi(d\xi, ds)\\ &= \int_0^t \int_{\Gamma_1\cap\Gamma_2} I_{\{b(\xi, Y(s))\neq 0\}}\pi(d\xi, ds) + \int_0^t \int_{\Gamma_1\cap\Gamma_2} I_{\{b(\xi, Y(s))=0\}}\pi(d\xi, ds)\\ &= \pi(\Gamma_1 \cap \Gamma_2, t).\end{aligned}$$

Therefore $\langle Z'(\Gamma, t), Z'(\Gamma, t)\rangle = \pi(\Gamma, t)$, and $Z'(\Gamma_1, t)Z'(\Gamma_2, t)$ is a martingale if Γ_1 and Γ_2 are disjoint. Again, we have recovered N.

$$(3.22)\quad \begin{aligned}\int_0^t \int_\Xi b(\xi, Y'(s))Z'(d\xi, ds) &= \int_0^t \int_\Xi I_{\{b(\xi, Y'(s))\neq 0\}} b(\xi, Y'(s))Z'(d\xi, ds)\\ &= \int_0^t \int_\Xi I_{\{b(\xi, Y'(s))\neq 0\}} b(\xi, Y'(s))(\eta'(\xi, Y'(s)), M'(d\xi, ds))\\ &\quad + \int_0^t \int_\Xi I_{\{b(\xi, Y'(s))\neq 0\}} \rho(\xi, Y'(s))N'(d\xi, ds)\\ &= \int_0^t \int_\Xi I_{\{b(\xi, Y'(s))\neq 0\}} M'(d\xi, ds) + 0\\ &= M'(\Xi, t) - \int_0^t \int_\Xi I_{\{b(\xi, Y'(s))=0\}} M'(d\xi, ds).\end{aligned}$$

$\int_0^t \int_\Xi I_{\{b(\xi, Y'(s))=0\}} M'(d\xi, ds) = 0$ because the quadratic variation for this martingale is

$$(3.23)\quad \int_0^t \int_\Xi I_{\{b(\xi, Y'(s)))=0\}} (\theta, b(\xi, Y'(s)))^2 \pi(d\xi, ds) = 0.$$

Therefore

$$\begin{aligned}\int_0^t \int_\Xi b(\xi, Y'(s))Z'(d\xi, ds) &= \int_0^t M'(d\xi, ds) = M'(\Xi, t)\\ &= Y'(t) - Y'(0) - \int_0^t \int_\Xi a(\xi, Y'(s))\pi(d\xi, ds)\end{aligned}$$

i.e., the stochastic integral equation holds. □

4. Uniqueness of solution to the stochastic integral equation. The focus of attention now moves to the stochastic integral equation for the limiting random evolutions. Thus we are using Theorem 3.1 in the case

$$a(\xi, y) = A^2(\xi)y/2 \quad b(\xi, y) = A(\xi)y$$

$$\pi(d\xi, ds) = \mu(d\xi)\, ds \quad N(d\xi, ds) = W_\mu(d\xi, ds)$$

and $\mathscr{D}' = \cap \{\mathscr{D}(A^{*2}(\xi)): \xi \in \Xi\}$, which is assumed to be dense in B^*. First of all we present two reformulations of the integral equation.

LEMMA 4.1. *If* Y *solves* 1.15 *then*

$$Y(t) = Y(0) + \int_0^t \int_\Xi A(\xi)Y(s) \circ W(d\xi, ds). \tag{4.1}$$

PROOF. Use the definition to make the transformation. □

This form suggests that Y solves an evolution equation and that Y preserves some sort of exponential behavior from the random evolutions Y_n in 1.1. The noncommutivity manifests itself in the fact that the integral in 4.1 cannot be written as an iterated integral. Although we shall not use this form, this representation shows the connection between central limit theorems for random evolutions and stochastic flows (see [8]).

LEMMA 4.2. *If* Y *solves* 1.15 *and if the closure of* $C = \frac{1}{2}\int_\Xi A^2(\xi)\mu(d\xi)$ *generates a semigroup* $S(t)$ *then*

$$Y(t) = S(t)Y(0) + \int_0^t \int_\Xi S(t-s)A(\xi)Y(s)W(d\xi, ds). \tag{4.2}$$

PROOF. Let $\theta \in \mathscr{D}'$, then

$$\begin{aligned}
&(\theta, Y(t) - S(t)Y(0)) \\
&\quad = \int_0^t (\theta, dS(t-s)Y(s)) \\
&\quad = -\int_0^t (\theta, CS(t-s)Y(s))\, ds + \int_0^t (\theta, S(t-s)\, dY(s)) \\
&\quad = -\int_0^t (\theta, CS(t-s)Y(s))\, ds + \int_0^t (\theta, CS(t-s)Y(s))\, ds \qquad (4.3) \\
&\qquad + \int_0^t \int_\Xi (\theta, S(t-s)A(\xi)Y(s))W(d\xi, ds) \\
&\quad = \int_0^t \int_\Xi (\theta, S(t-s)A(\xi)Y(s))W(d\xi, ds). \quad \square
\end{aligned}$$

Under the variation of parameters form, the average behavior, $S(t)Y(0)$, is separated from the random behavior of the integral term. This presentation gives us a chance to handle the unbounded operators $A(\xi)$. We shall prove pathwise uniqueness for equation 4.2, which in turn, proves pathwise uniqueness for equation 1.15. After a few preparatory lemmas and propositions we shall be set to prove the following theorem.

THEOREM 4.3. *If the coefficient functions a in* $A(\xi) = a(\xi, x)\cdot\nabla$ *are elements in* $C^1(\mathbb{R}^d)$ *for each* $\xi \in \Xi$ *and* $\sup\{|(\partial a/\partial x_k)(\xi, x)| : x \in \mathbb{R}^d, k = 1, \cdots, d, \xi \in \Xi\} < \infty$, *and if the closure of* $C = \frac{1}{2}\int_\Xi A^2(\xi)\mu(d\xi)$ *is uniformly parabolic, and generates a semigroup* $S(t)$, *then the stochastic integral equation has the property of pathwise uniqueness.*

PROPOSITION 4.4. (Gronwall's Inequality) *Let* $f, g:[0, b) \to \mathbb{R}$ *be continuous and nonnegative. Suppose,*

$$f(t) \le A + \int_0^t f(s)g(s)\,ds, \quad A \ge 0. \tag{4.4}$$

Then

$$f(t) \le A \exp \int_0^t g(s)\,ds \quad \text{for} \quad t \in [0, b). \tag{4.5}$$

PROPOSITION 4.5. *Let* $G(x, t; y, u)$ *be the fundamental solution to a uniformly parabolic differential equation in* $\mathbb{R}^d \times [0, T]$. *If the coefficients have continuous and bounded first derivatives, then*

$$\left| \frac{\partial}{\partial y_i} G(x, t; y, u) \right| \le C_1(t-u)^{-(d-1)/2}\exp\left[-C_2\frac{|x-y|^2}{(t-u)}\right] \tag{4.6}$$

for some $C_1, C_2 > 0$, *and all* $i = 1, 2, \cdots, d$.

PROOF. See [2].

LEMMA 4.6. *Under the conditions stated in Theorem* 4.3 *for t sufficiently small*

$$\| S(t)A(\xi) \|^2 \le C/t \tag{4.7}$$

for all $\xi \in \Xi$ *and some* $C > 0$. *C depends only upon the Lipschitz constant.*

PROOF. Let $f \in \cap\, \mathscr{D}(A(\xi))$, and let G be the Green's function associated with S.

$$
\begin{aligned}
\| S(t)A(\xi)f \| &= \left\| \int G(x, t; y, 0)a(\xi, y)\cdot\nabla_y f(y)\, dy \right\| \\
&= \left\| \int \nabla_y \cdot [G(x, t; y, 0)a(\xi, y)] f(y)\, dy \right\| \\
&\le \left\| \int [\nabla_y G(x, t; y, 0)\cdot a(\xi, y)] f(y)\, dy \right\| \\
&\quad + \left\| \int [G(x, t; y, 0)\nabla_y \cdot a(\xi, y)] f(y)\, dy \right\| \\
&\le K_1 \|f\| \left\| \int |\nabla_y G(x, t; y, 0)|\, dy \right\| \\
&\quad + K_1 \|f\| \left\| \int |G(x, t; y, 0)|\, dy \right\| \\
&\le K_1 \|f\|\, dC_1 t^{-(d-1)/2} \int \exp\left[-C_2 \frac{|x - y|^2}{t}\right] dy + K_1 \|f\| \\
&\le K_1 (dC_1 K_2 t^{-1/2} + 1)\, \|f\| \le C t^{-1/2} \|f\|
\end{aligned} \tag{4.8}
$$

for $C = 2dC_1K_2$ and $t < (dC_1)^2$. K_1 is the constant arising from the Lipschitz condition on a and $\partial a/\partial y_i$. □

Now the proof of Theorem 4.3.

PROOF. If Y_1 and Y_2 are both continuous solutions to equation 4.2 with identical initial condition, then the difference

$$
\begin{aligned}
Y_1(t) - Y_2(t) &= S(t)(Y_1(0) - Y_2(0)) \\
&\quad + \int_0^t \int_\Xi S(t - s)A(\xi)(Y_1(s) - Y_2(s))W(d\xi, ds) \\
&= \int_0^t \int_\Xi S(t - s)A(\xi)(Y_1(s) - Y_2(s))W(d\xi, ds).
\end{aligned} \tag{4.9}
$$

Therefore one way to establish uniqueness is to prove that if $Y(t)$ is a continuous solution with $Y(0) = 0$ a.s., then $Y(t)$ must be zero.

In order to establish uniqueness, we shall use the norm

$$
[\sup_{\|\theta\|=1} E(\theta, \cdot)^2]^{1/2}. \tag{4.10}
$$

This will allow us to take advantage of the martingale

$$
\left[\int_0^t \int_\Xi (\theta, X(\xi, s))W(d\xi, ds)\right]^2 - \int_0^t \int_\Xi (\theta, X(\xi, s))^2 \mu(d\xi)\, ds \tag{4.11}
$$

for nonanticipating X. Fix $t > 0$ and choose $\theta \in \mathscr{D}'$ so that $\|\theta\| = 1$. $S(\varepsilon)$

converges strongly to the identity as $\varepsilon \to 0$. As a consequence, for each $\omega \in \Omega$, and each $s \le t$

$$S(\varepsilon)Y(s, \omega) \to Y(s, \omega) \quad \text{as} \quad \varepsilon \to 0 \tag{4.12}$$

and

$$(\theta, S(\varepsilon)Y(s, \omega))^2 \to (\theta, Y(s, \omega))^2. \tag{4.13}$$

Hence we can choose $\varepsilon(\omega)$ so that if $\varepsilon < \varepsilon(\omega)$ and $s \le t$

$$2(\theta, S(\varepsilon)Y(s, \omega))^2 \ge (\theta, Y(s, \omega))^2. \tag{4.14}$$

Since Y has continuous sample paths, the convergence in equation 4.12 is uniform in $[0, t]$. Therefore one $\varepsilon(\omega)$ may be chosen for all $s \le t$. Define

$$G_N(t) = \{\omega\colon 2(\theta, S(1/N)Y(s, \omega))^2 \ge (\theta, Y(s, \omega))^2 \text{ for all } s \le t\} \tag{4.15}$$

Since each ω must eventually be an element of $G_N(t)$,

$$\cup_{N=1}^{\infty} G_N(t) = \Omega.$$

In a similar fashion, for each ω, we may choose $M(\omega)$ so that if $M > M(\omega)$

$$M(\theta, Y(s, \omega))^2 > \| Y(s, \omega) \|^2 \tag{4.16}$$

provided that $(\theta, Y(s, \omega)) \ne 0$. Define

$$H_N(t) = \{\omega\colon N(\theta, Y(s, \omega))^2 > \| Y(s, \omega) \|^2 \text{ for all } s \in [(t - (1/N)) \wedge 0, t]\}.$$

$$F^c(t) = \cup_{N=1}^{\infty} H_N(t), \tag{4.17}$$

where

$$F(t) = \{\omega\colon (\theta, Y(t, \omega)) = 0\}. \tag{4.18}$$

As a final preparation, let us introduce two sequences of stopping times:
For natural numbers L and N,

$$\sigma_L(\omega) = \inf\{s > 0\colon \| Y(s, \omega) \| \ge L\}, \tag{4.19}$$

and

$$\tau_N(\omega) = \inf\{s \ge t - 1/N\colon \omega \in G_N(s) \cap H_N(s)\}. \tag{4.20}$$

Note that on the set $F^c(t)$, $\tau_N \ge t$ if N is large enough. In any case $\tau_N \ge t - (1/N)$. The stopping times σ_L are introduced since the second moment of Y may not be finite. By the continuity, and along with it, the boundedness of $Y(t)$, we may conclude that

$$\sigma_L \to \infty \quad \text{a.s. as} \quad L \to \infty. \tag{4.21}$$

With everything in order, we begin. Multiply equation 4.2 by $S(1/N)$ and sample it via the linear functional θ:

$$\begin{aligned}&(\theta, S(1/N)Y(t \wedge \sigma_L \wedge \tau_N))\\ &= \int_0^{t\wedge\sigma_L\wedge\tau_N} \int_\Xi \left(\theta, S\left(t \wedge \sigma_L \wedge \tau_N - s + \frac{1}{N}\right)A(\xi)Y(s)\right)W(d\xi, ds).\end{aligned} \tag{4.22}$$

If we square both sides of this equation, then we may use the martingale mentioned at the start of the proof, and the strong Markov property to justify

$$
\begin{aligned}
&E[\theta, S(1/N)Y(t \wedge \sigma_L \wedge \tau_N)]^2 \\
&= E\left[\int^{t\wedge\sigma_L\wedge\tau_N} \int_{\Xi} \left(\theta, S\left(t \wedge \sigma_L \wedge \tau_N - s + \frac{1}{N}\right)A(\xi)Y(s)\right)W(d\xi, ds)\right]^2 \\
&= E\int^{t\wedge\sigma_L\wedge\tau_N} \int_{\Xi} \left(\theta, S\left(t \wedge \sigma_L \wedge \tau_N - s + \frac{1}{N}\right)A(\xi)Y(s)\right)^2 \mu(d\xi)\, ds \\
&\leq E\int_0^{t\wedge\sigma_L\wedge\tau_N} \int_{\Xi} \left\| S\left(t \wedge \sigma_L \wedge \tau_N - s + \frac{1}{N}\right)A(\xi)Y(s) \right\|^2 \mu(d\xi)\, ds
\end{aligned}
$$

$$
(4.23) \quad \leq E\int_0^{t\wedge\sigma_L\wedge\tau_N} \int_{\Xi} C\left(t \wedge \sigma_L \wedge \tau_N - s + \frac{1}{N}\right)^{-1} \| Y(s) \|^2 \mu(d\xi)\, ds
$$

by Lemma 4.5

$$
\begin{aligned}
&\leq E\int_0^{t\wedge\sigma_L\wedge\tau_N} CN \| Y(s) \|^2\, ds \\
&= E\int_0^{t\wedge\sigma_L\wedge\tau_N} CN \| Y(s \wedge \sigma_L \wedge \tau_N) \|^2\, ds \\
&\leq \int_0^t CNE \| Y(s \wedge \sigma_L \wedge \tau_N) \|^2\, ds.
\end{aligned}
$$

However for the stopped process

$$
\| Y(s \wedge \sigma_L \wedge \tau_N) \|^2 \leq N(\theta, Y(s \wedge \sigma_L \wedge \tau_N))^2 \leq 2N(\theta, S(1/N)Y(s \wedge \sigma_L \wedge \tau_N))^2.
$$

Therefore,

$$
(4.24) \quad E \| Y(s \wedge \sigma_L \wedge \tau_N) \|^2 \leq \int_0^t C2N^2E \| Y(s \wedge \sigma_L \wedge \tau_N) \|^2\, ds.
$$

Gronwall's inequality applies to yield

$$
(4.25) \quad E \| Y(t \wedge \sigma_L \wedge \tau_N) \|^2 = 0.
$$

Now let $L \to \infty$.

$$
(4.26) \quad E \| Y(t \wedge \tau_N) \|^2 = 0.
$$

$$
(4.27) \quad EI_{F^c(t)}(\theta, Y(t \wedge \tau_N))^2 = 0.
$$

Now let $N \to \infty$

$$
(4.28) \quad EI_{F^c(t)}(\theta, Y(t))^2 = 0.
$$

On $F^c(t)$, $(\theta, Y(t))^2 > 0$. Therefore, $P(F^c(t)) = 0$ or $P(F(t)) = 1$. In other words, $(\theta, Y(t)) = 0$ with probability 1, i.e., $E(\theta, Y(t))^2 = 0$. Since we can do this for all

$\theta \in \mathscr{D}'$, $\|\theta\| = 1$,

$$\sup_{\|\theta\|=1} E(\theta, Y(t))^2 = 0. \tag{4.29}$$

or $Y(t) = 0$ a.s. for each t. Let D by any countable dense set in $[0, T]$, then

$$P\{Y(t) = 0 \text{ for all } t \in D\} = 1 \tag{4.30}$$

but Y must be almost surely continuous. Therefore, equation 4.30 provides us with sufficient justification to conclude that

$$Y(t) = 0 \quad \text{a.s. for all} \quad t \in [0, T]. \tag{4.31}$$

This assures us the uniqueness of solutions. □

5. Existence of solutions to the stochastic integral equation. Recall that for any B-valued nonanticipating X with $E \int_0^t \int_\Xi \|X(\xi, s)\|^2 \mu(d\xi)\, ds < \infty$ for all t we have the following martingale:

$$\int_0^t \int_\Xi (\theta, X(\xi, s))^2 \mu(d\xi)\, ds - \left(\int_0^t \int_\Xi (\theta, X(\xi, s)) W(d\xi, ds)\right)^2 \tag{5.1}$$

for any $\theta \in B^*$. We wish to establish that

$$\int_0^t \int_\Xi \|X(\xi, s)\|^2 \mu(d\xi)\, ds - \left\| \int_0^t \int_\Xi X(\xi, s) W(d\xi, ds) \right\|^2 \tag{5.2}$$

is a martingale. Returning to the language of Section 2, if X is an elementary function, then there exists a random variable θ, so that $(\theta, X(\xi, s))$ is nonanticipating, $\|\theta\| = 1$ and

$$\|X(\xi, s)\| = (\theta, X(\xi, s)). \tag{5.3}$$

In this case, 5.1 and 5.2 are identical. Finite sums of martingales are martingales; therefore equation 5.3 is a martingale for any simple function. Now take limits in order to conclude that 5.2 is a martingale for all nonanticipating X with $E \int_0^t \int_\Xi \|X(\xi, s)\|^2 \mu(d\xi)\, ds < \infty$ for all t. Thus, the martingale 5.2 gives the isometry we use to build Banach valued processes.

THEOREM 5.1. *If the coefficient functions $a \in C^\infty(\mathbb{R}^d)$ and if for each ξ and each multi-index α, $D^\alpha a(\xi, x)$ is a Lipschitz function with a common Lipschitz constant, and if C is uniformly parabolic, then the variation of parameters form (Equation 4.2) has a strong solution.*

The estimates we use in the proof are based upon the following theorem:

THEOREM 5.2. [2] *If the coefficients of a uniformly parabolic system in*

$\mathbb{R}^d \times [0, T]$ *are infinitely differentiable, and if all the derivatives are bounded functions, then the fundamental solution* G *is infinitely differentiable,*

$$(5.4)\quad |D_y^\alpha G(x, t; y, u)| \leq \left| C_1\left(\frac{C_2}{2\pi(t-u)}\right)^{d/2} D^\alpha \exp\left(-C_2\left(\frac{|x-y|^2}{2(t-u)}\right)\right) \right|$$

and

$$(5.5)\quad |D_x^\alpha G(x, t; y, u)| \leq \left| C_1\left(\frac{C_2}{2\pi(t-u)}\right)^{d/2} D^\alpha \exp\left(-C_2\left(\frac{|x-y|^2}{2(t-u)}\right)\right) \right|$$

for any multi-index α.

PROOF OF THEOREM 5.1. Let us attempt to construct a solution to

$$(5.6)\qquad Y(t) = S(t)Y(0) + \int_0^t \int_\Xi S(t-s)A(\xi)Y(s)W(d\xi, ds)$$

for these two cases. First, choose $U_0(t) = S(t)Y(0)$. If a solution exists, U_0 captures the average behavior, since $EU_0(t) = EY(t)$. Then let

$$(5.7)\qquad U_1(t) = \int_0^t \int_\Xi S(t-s)A(\xi)U_0(s)W(d\xi, ds).$$

Iterating this procedure, we have that

$$(5.8)\qquad U_k(t) = \int_0^t \int_\Xi S(t-s)A(\xi)U_{k-1}(s)W(d\xi, ds).$$

We shall define these iterates one at a time. To be precise, if U_{k-1} is a continuous nonanticipating process in the domain of $A(\xi)$, then because $S(t-s)A(\xi)$ is a ξ-uniformly bounded family of operators, with $\cup\mathscr{R}(S(t-s)A(\xi)) \subseteq \cap\, \mathscr{D}(A(\xi))$, then the integrand in equation 5.8 is defined. Once we have established that U_k is bounded, then we can conclude that U_k is a nonanticipating continuous process in the domain of the $A(\xi)$. The $k = 0$ step of the induction is easy to verify. Therefore, owing the proof of the boundedness of the U_k, we are allowed to perform the iteration. By the linearity of the integral

$$(5.9)\quad \sum_{k=0}^n U_k(t) = S(t)Y(0) + \int_0^t \int_\Xi S(t-s)A(\xi) \sum_{k=0}^{n-1} U_k(s)W(d\xi, ds).$$

We shall demand more than boundedness of U_k, because we wish to make sense of

$$\lim_{n\to\infty} \sum_{k=0}^n U_k(t),$$

and to show that equation 5.6 holds for the limit. A typical argument uses the fact that $\sup_{0\leq t\leq T} E\| U_k(t) \|^2$ is bounded by the kth term in an exponential expansion. This fact, along with a martingale inequality and the Borel-Cantelli lemma secures the limit. This is our plan here.

$$
\begin{aligned}
U_k(t) &= \int_0^t \int_\Xi S(t-t_1)A(\xi_1)U_{k-1}(t_1)W(d\xi_1, dt_1) \\
&= \int_0^t \int_\Xi S(t-t_1)A(\xi_1) \int_0^{t_1} \int_\Xi S(t_1-t_2)A(\xi_2)U_{k-2}(t_2) \\
&\qquad \cdot W(d\xi_2, dt_2)W(d\xi_1, dt_1) \\
&= \int_0^t \int_\Xi \int_0^{t_1} \int_\Xi S(t-t_1)A(\xi_1)S(t_1-t_2)A(\xi_2)U_{k-2}(t_2) \\
&\qquad \cdot W(d\xi_2, dt_2)W(d\xi_1, dt_1) \\
&= \int_0^t \int_\Xi \int_0^{t_1} \int_\Xi \cdots \int_0^{t_{k-1}} \int_\Xi S(t-t_1)A(\xi_1)S(t_1-t_2)A(\xi_2) \\
&\qquad \cdots A(\xi_k)S(t_k)Y(0)W(d\xi_k, dt_k) \cdots W(d\xi_2, dt_2)W(d\xi_1, dt_1).
\end{aligned}
\tag{5.10}
$$

Upon k applications of the martingale property we have

$$
\begin{aligned}
&E\| U_k(t) \|^2 \\
&= \int_0^t \int_\Xi \int_0^{t_1} \int_\Xi \cdots \int_0^{t_{k-1}} \int_\Xi E\| S(t-t_1)A(\xi_1)S(t_1-t_2)A(\xi_2) \\
&\qquad \cdots A(\xi_k)S(t_k)Y(0)\|^2 \mu(d\xi_k)\, dt_k \cdots \mu\,(d\xi_2)\, dt_2 \mu(d\xi_1)\, dt_1.
\end{aligned}
\tag{5.11}
$$

The order of business is now to estimate the norm of the product

$$
J_k(t)f = S(t-t_1)A(\xi_1)S(t_1-t_2)A(\xi_2) \cdots A(\xi_k)S(t_k)f, \tag{5.12}
$$

for $f \in B$. Unfortunately giving a detailed proof and avoiding both notational nightmares and expressions that become unmanageably long seems impossible. Luckily, the ideas are not quite as hard to grasp. In this segment, $J_k(t)$ will signify any term like equation 5.12 with A appearing k times and t being the sum of the lengths of time that the semigroup S acts. If we commute the second and third factors in 5.1, we find that

$$
\begin{aligned}
J_k(t)f &= S(t-t_1)S(t_1-t_2)A(\xi_1)A(\xi_2)S(t_2-t_3)A(\xi_3) \cdots A(\xi_k)S(t_k)f \\
&\quad + S(t-t_1)[A(\xi_1), S(t_1-t_2)]A(\xi_2) \cdots A(\xi_k)S(t_k)f \\
&= S(t-t_2)A(\xi_1)A(\xi_2)S(t_2-t_3)A(\xi_3) \cdots A(\xi_k)S(t_k)f \\
&\quad + S(t-t_1)[A(\xi_1), S(t_1-t_2)]A(\xi_2) \cdots A(\xi_k)S(t_k)f.
\end{aligned}
\tag{5.13}
$$

The symbol $[\cdot, \cdot]$ denotes the commutator of the two operators. We shall show that the commutator term acts like $S(t_1 - t_2)$. Hence, we may estimate this term as we would $J_{k-1}(t)f$. This process continues to commute terms, moving the A term to the right and spinning off terms that may be estimated as one would deal with $J_m(t)$ for some $m < k$. The final remaining term is

$$
\begin{aligned}
&| S(t)A(\xi_1)A(\xi_2) \cdots A(\xi_k)f | \\
&\qquad = \left| \int G(x, t; y, 0)A(\xi_1)A(\xi_2) \cdots A(\xi_k)f(y)\, dy \right| \\
&\qquad = \left| \int A^*(\xi_1)A^*(\xi_2) \cdots A^*(\xi_k)G(x, t; y, 0)f(y)\, dy \right| \qquad (5.14)\\
&\qquad \le \sup_{|\alpha|=k} L^k \int | D^\alpha G(x, t; y, 0) |\, | f(y) |\, dy < CL^k t^{-k} \| f \|.
\end{aligned}
$$

The procedure produces $k(k - 1)/2$ terms, each bounded by $CL^m t^{-m}$, with $m < k$. For t bounded, we need worry only about the behavior for t near zero. Thus the final remaining term dominates, and

$$
\| J_k(t) \| \le (k^2/2)\, CL^k t^{-k}. \tag{5.15}
$$

$A(\xi_i)$ is now written $a_i \cdot \nabla$. The subscript i on the gradient operator is meant to show that the derivatives are taken with respect to $x_i = (x_i^1, x_i^2, \cdots, x_i^d)$. For brevity write

$$
e(x, t) = C_1\left(\frac{C_2}{2\pi(t - u)}\right)^{d/2} \exp\left(-C_2\left(\frac{| x - y |^2}{2(t - u)}\right)\right). \tag{5.16}
$$

Now we check to see that the commutator term acts appropriately.

$$
\begin{aligned}
&| J_k(t)f | \\
&= \left| \int\int G(x, t; x_1, t_1)a_1(x)\cdot\nabla G(x_1, t_1; x_2, t_2)a_2(x_2)\nabla_2 J_{k-2}(t_2)f\, dx_2\, dx_1 \right| \\
&\le \int\int | G(x, t; x_1, t_1) |\, | a_1(x) |\, | \nabla G(x_1, t_1; x_2, t_2) | \\
&\qquad\qquad \cdot | a_2(x_2)\cdot\nabla_2 J_{k-2}(t_2)f |\, dx_2\, dx_1 \\
&\le \int\int e(x - x_1, t - t_1)Kd\, | D_{x_2^1} e(x_1 - x_2, t_1 - t_2) | \\
&\qquad\qquad \cdot | a_2(x_2)\cdot\nabla_2 J_{k-2}(t_2)f |\, dx_2\, dx_1 \\
&\le Kd \int\int e(x - x_1, t - t_1)D_{x_2^1} e(x_1 - x_2, t_1 - t_2) \\
&\qquad\qquad \cdot \operatorname{sgn}(x_1^1 - x_2^1)\, | a_2(x_2)\cdot\nabla_2 J_{k-2}(t_2)f |\, dx_2\, dx_1 \\
&= -Kd \int\int e(x - x_1, t - t_1)e(x_1 - x_2, t_1 - t_2) \\
&\qquad\qquad \times D_{x_2^1}[\operatorname{sgn}(x_1^1 - x_2^1)\, | a_2(x_2)\cdot\nabla_2 J_{k-2}(t_2)f |]\, dx_2\, dx_1
\end{aligned}
$$

after integrating by parts,

$$\le -KC_2 d \int e(x - x_2, t - t_2) D_{x_2^1}[\operatorname{sgn}(x_1^1 - x_2^1) \,|\, a_2(x_2)\cdot\nabla_2 J_{k-2}(t_2) f \,|\,]\, dx_2$$

by the semigroup property,

$$\le KC_2 d \int e(x - x_2, t - t_2) 2\delta(x_1^1 - x_2^1) \,|\, a_2(x_2)\cdot\nabla_2 J_{k-2}(t_2) f \,|\, dx_2$$

$$+ KC_2 d \int e(x - x_2, t - t_2)\operatorname{sgn}(x_1^1 - x_2^1) D_{x_2^1} |\, a_2(x_2)\cdot\nabla_2 J_{k-2}(t_2) f \,|\, dx_2$$

$$\le 2KC_2 d \int e(x - x_2, t - t_2) \,|\, a_2(x_2)\cdot\nabla_2 J_{k-2}(t_2) f \,|\, dx_2 \tag{5.17}$$

$$+ KC_2 d \int\int e(x - x_2, t - t_2)\operatorname{sgn}(x_2^1 - x_1^1)$$

$$\times D_{x_2^1} |\, a_2(x_2)\cdot\nabla G(x_2, t_2; x_3, t_3) a_3(x_3)\cdot\nabla_3 J_{k-3}(t_3) f \,|\, dx_3\, dx_2.$$

We can identify the first term as the commutator term. As promised, a differential operator disappeared, and the semigroup terms combined. So this term may be regarded as $J_{k-1}(t)$. The next term is set for the next commutation. The balance of the details is omitted, since the terms become exceedingly long. Guided by the outline above, the reader should be able to produce the course of action.

Assume first that $E\|Y(0)\|^2 < \infty$. Then the iteration procedure in equation 5.8 that generated the U_k is valid. Also by Doob's martingale inequality.

$$E[\sup_{0\le t\le T}\|U_k(t)\|^2] \le 4E\|U_k(T)\|^2$$

$$= 4\int_0^T \int_\Xi \int_0^{t_1} \int_\Xi \cdots \int_0^{t_{k-1}} \int_\Xi E\|J_k(T)Y(0)\|^2$$

$$\cdot\, \mu(d\xi_k)\, dt_k \cdots \mu(d\xi_2)\, dt_2 \mu(d\xi_1)\, dt_1$$

by equations 5.11 and 5.12. Now use the inequality in 5.14.

$$\le 4\int_0^T \int_\Xi \int_0^{t_1} \int_\Xi \cdots \int_0^{t_{k-1}} \int_\Xi \frac{k^2}{2} CL^k T^{-k} E\|Y(0)\|^2$$

$$\cdot\, \mu(d\xi_k)\, dt_k \cdots \mu(d\xi_2)\, dt_2 \mu(d\xi_1)\, dt_1$$

$$= 2k^2 CL^k T^{-k} E\|Y(0)\|^2 \int_0^T \int_0^{t_1} \cdots \int_0^{t_{k-1}} dt_k \cdots dt_2\, dt_1 \tag{5.18}$$

$$= 2k^2 CL^k T^{-k} E\|Y(0)\|^2 \left(\frac{T^k}{k!}\right) = \frac{2k^2}{k!} CL^k E\|Y(0)\|^2.$$

By virtue of the Borel-Cantelli lemma and the Weierstrass-M test,

$$\sum_{k=0}^{\infty} P\{\sup_{0\le t\le T}\|U_k(t)\|^2 > k^{-2}\} \le E\|Y(0)\|^2 2C \sum_{k=0}^{\infty} (k^2/k!)L^k < \infty. \tag{5.19}$$

This implies that $\sum_{k=0}^{n} U_k(t)$ converges almost surely uniformly on $[0,T]$. Therefore, the limit $Y(t)$ exists as a continuous nonanticipating process. Also

$$S(t-s)A(\xi)\sum_{k=0}^{n-1}U_k(t) \to S(t-s)A(\xi)Y(s) \quad (5.20)$$

almost surely $[P]$ for all ξ, and Y solves the integral equation.

A truncation argument allows us to release the restriction $E\|Y(0)\|^2 < \infty$. □

REFERENCES

[1] ARNOLD, L. (1973). *Stochastiche Differentalgleichungen, Theorie und Andwendung*, R. Olderburg Verlag, München.

[2] FRIEDMAN, A. (1964). *Partial Differential Equations of Parabolic Type*. Englewood Cliffs, New Jersey.

[3] FRIEDMAN, A. (1975). *Stochastic Differential Equations and Applications*. Academic, New York.

[4] GIHMAN, I. and SKOROHOD, A. (1972). *Stochastic Differential Equations*. Springer Verlag, Berlin.

[5] GIHMAN, I. and SKOROHOD, A. (1979). *The Theory of Stochastic Processes, III*. Springer Verlag, Berlin.

[6] STROOCK, D. and VARADHAN, S. (1969). Diffusion Processes With Continuous Coefficients. *Comm. Pure Appl. Math.* **XXII** 345–400, 479–530.

[7] WATKINS, J. (1984). A Central Limit Problem in Random Evolution. *Ann. Probab.* **12** 480–514.

[8] ELWORTHY, K. (1982). *Stochastic Differential Equations on Manifolds*. London Mathematical Society Lecture Note Series **70**. Cambridge University Press.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, BRITISH COLUMBIA
V6T 1Y4 CANADA