JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 31, 434–448 (1970)

Ito's Lemma in Infinite Dimensions

RUTH F. CURTAIN* AND PETER L. FALB*

Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

Submitted by J. P. LaSalle

1. INTRODUCTION

We extend Ito's lemma ([5] or [8], for example) to a Hilbert space context in this paper. Our proof is analogous to that given by Gikhman and Skorokhod ([5]) for the real random variable case. Thus, the crucial points in our treatment involve the proper formulation and juxtaposition of concepts such as Wiener process and stochastic differential in a Hilbert space context. For this, we rely on a modification of the ideas in [1] and [4].

We let $(\Omega, \mathscr{P}, \mu)$ be a probability space with $\mathscr{P}$ as Borel field and μ as measure throughout the paper. We assume that the reader is somewhat familiar with the theory of Banach space valued random variables (see, for example, [7]); however, for convenience, we include a brief appendix containing the definitions and results used in the paper.

We introduce the basic notions of a Hilbert space valued Wiener process and the corresponding stochastic integral in Section 2. Then, we state and prove the extension of Ito's lemma in Section 3. In essence, we show that if H, K, and G are Hilbert spaces and if $u(t)$ is a K-valued stochastic process with stochastic differential $du = q(t)\,dt + \Phi(t)\,dw$ where $w(t)$ is an H-valued Wiener process, $q(t)$ is a K-valued stochastic process, and $\Phi(t)$ is a suitable $\mathscr{L}(H, K)$[1] valued stochastic process, then the stochastic differential of the G-valued stochastic process $z(t) = g(t, u(t))$, where g is a sufficiently smooth nonrandom map of $T \times K$ into G ($T = [T_1, T_2]$ a real interval), can be written down in terms of the derivatives of g.

* Division of Applied Mathematics, Brown University, Providence, R.I. This Research was supported by the National Science Foundation, under grants NSF-GP 7471, NSF-GK 2788 and NSF-GP 9024. Dr. Curtain is currently at the Aerospace Eng. Department, Purdue University.

[1] $\mathscr{L}(H, K)$ is the space of bounded linear maps of H into K.

2. Wiener Processes and Stochastic Integrals

We define Hilbert space valued Wiener processes and develop the concomitant stochastic integral in this section.

Definition 2.1. Let $w(t)$ be an H-valued random process on T. Then $w(t)$ is called a Wiener process if

(i) $E\{w(t) - w(s)\} = 0$ for all s, t in T;

(ii) $w(t)$ is continuous in t w.p.l.;[2]

(iii) $E\{[w(t) - w(s)] \circ [w(t) - w(s)]\}$[3] $= (t - s)\, W$ for all s, t in T where W is a compact, positive, bounded, trace class operator mapping H into itself;

(iv) $E\{\| w(t) - w(s)\|^2\} < \infty$ for all s, t in T; and,

(v) $w(t_2) - w(t_1)$ and $w(s_2) - w(s_1)$ are independent for all s_1, s_2, t_1, t_2 in T such that $s_1 < s_2 \leqslant t_1 < t_2$.

We note that the operator W has countably many eigenvalues $\{\lambda_i\}$, that $\lambda_i \geqslant 0$ for all i, and that $\mathrm{tr}(W) = \sum_{i=0}^{\infty} \lambda_i$. If $\{e_i\}$ is an orthonormal set of eigenvectors of W and if $\{e_i, f_\alpha\}$ is an orthonormal basis for H, then $We_i = \lambda_i e_i$ and $Wf_\alpha = 0$.

We also observe that several alternative versions of Definition 2.1 can be obtained by replacing (v) by either of the weaker conditions

(v)′ $\langle w(t_2) - w(t_1), h_1\rangle$ and $\langle w(s_2) - w(s_1), h_2\rangle$ are independent for all s_1, s_2, t_1, t_2 in T such that $s_1 < s_2 \leqslant t_1 < t_2$ and all h_1, h_2 in H; or,

(v)″ $\langle w(t_2) - w(t_1), e_i\rangle$ and $\langle w(s_2) - w(s_1), e_i\rangle$ are independent for all s_1, s_2, t_1, t_2 in T such that $s_1 < s_2 \leqslant t_1 < t_2$.

These alternative versions of the definition produce deintical results (see [2]).

If $w(t)$ is an H-valued Wiener process, then it can be shown that there are complex-valued stochastic processes $\{\beta_i(t)\}$ on T such that

$$w(t) = \sum_{i=0}^{\infty} \beta_i(t)\, e_i \tag{2.2}$$

almost everywhere in (t, ω), where $\{e_i\}$ is an orthonormal set of eigenvectors of W. Moreover, $\mathrm{Re}\{\beta_i(t)\}$ and $\mathrm{Im}\{\beta_i(t)\}$ are real Wiener processes. Thus, there is no essential difference between Wiener processes in separable and

[2] w.p.1. is shorthand for *with probability one.*

[3] If h_1 and h_2 are elements of H, then $h_1 \circ h_2$ is the element of $\mathscr{L}(H, H)$ given by $(h_1 \circ h_2)h = h_1\langle h, h_2\rangle$ (cf. [4]).

nonseparable Hilbert spaces and so, we shall assume from now on that H is separable.

PROPOSITION 2.3. *If $w(t)$ is an H-valued Wiener process, then*

$$E\{\langle w(t) - w(s), w(t) - w(s)\rangle\} = \operatorname{tr}(W) \mid t - s \mid$$

and

$$E\{\| w(t) - w(s)\|^4\} \leqslant 3[\operatorname{tr}(W)]^2 \mid t - s \mid .$$

A proof of this simple proposition is given in [2].

PROPOSITION 2.4. *If $w(t)$ is an H-valued Wiener process, then there is a family $\{\mathscr{F}_t, t \in T\}$ of σ-algebras such that*

(i) $\mathscr{F}_s \subset \mathscr{F}_t \subset \mathscr{P}$ *for* $s < t$;

(ii) $w(t)$ *is measurable relative to* $\mathscr{F}_t$ *for all* t *in* T;

(iii) $w(t) - w(s)$ *is independent of* $\mathscr{F}_s$ *for* $s < t$;

(iv) $[w(t) - w(s)] \circ [w(t) - w(s)]$ *is independent of* $\mathscr{F}_s$ *for* $s < t$.

Proof. Take, for example, $\mathscr{F}_t$ to be the σ-algebra generated by the sets $w(s)^{-1}(\mathscr{O})$, $s \in T$, $s \leqslant t$, $\mathscr{O}$ a Borel set in H. Properties (i), (ii), and (iii) are obvious. As for property (iv), this is an immediate consequence of (iii) and the fact that the mapping ψ of $H \oplus H$ into $\mathscr{L}(H, H)$ given by $\psi(h_1, h_2) = h_1 \circ h_2$ is continuous (see [4], Propositions 2.2 and 2.4).

COROLLARY 2.5. *If h_1 and h_2 are elements of H, then*

$$\langle [w(t) - w(s)] \circ [w(t) - w(s)]\, h_1, h_2\rangle$$

is independent of $\mathscr{F}_s$ for $s < t$.

If $w(t)$ is a Wiener process, then, for convenience, we fix a family $\{\mathscr{F}_t\}$ satisfying the conditions of Proposition 2.4 and associate it with $w(t)$. We then have

DEFINITION 2.6. Let K be a Hilbert space. Then $\mathscr{M}(H, K) = \{\Phi(\cdot, \cdot) :$ Φ is an $\mathscr{L}(H, K)$-valued stochastic process on $T \times \Omega$[4] such that $\Phi(t)$ is measurable relative to $\mathscr{F}_t$ for all t in $T\}$, $\mathscr{M}_0(H, K) = \{\Phi(\cdot, \cdot) \in \mathscr{M}(H, K) :$ Φ is a t-step function on $T\}$, and

$$\mathscr{M}_1(H, K) = \left\{\Phi(\cdot, \cdot) \in \mathscr{M}(H, K) : \int_T E\{\| \Phi(t)\|^2\}\, dt < \infty\right\}.$$

[4] This means that Φ is measurable with respect to the pair (t, ω). Although this is somewhat more restrictive than usual ([3]), it is adequate for our purposes.

If Φ is an element of $\mathscr{M}_1(H, K)$, then the K-valued stochastic integral, $\int_T \Phi(t, \omega)\, dw$, can be defined in an analogous way to that used in the scalar case by Skorokhod ([8]). More precisely, if Φ is an element of $\mathscr{M}_0(H, K) \cap \mathscr{M}_1(H, K)$, then $\int_T \Phi(t, \omega)\, dw$ is given by a finite sum of the form $\sum \Phi(t_j, \omega)\,(w(t_{j+1}) - w(t_j))$. It is easily checked that $E\{\int_T \Phi(t, \omega)\, dw\} = 0$ and we shall soon show that

$$E\left\{\left\| \int_T \Phi(t, \omega)\, dw \right\|^2\right\} = \operatorname{tr}(W) \int_T E\{\|\Phi(t)\|^2\}\, dt \tag{2.7}$$

for Φ in $\mathscr{M}_0(H, K) \cap \mathscr{M}_1(H, K)$. Now if Φ is any element of $\mathscr{M}_1(H, K)$, then there is a sequence $\{\Phi_n\}$ of elements of $\mathscr{M}_0(H, K) \cap \mathscr{M}_1(H, K)$ such that $\Phi_n \to \Phi$ almost everywhere on $T \times \Omega$ and

$$\lim_{n\to\infty} \int_T E\{\|\Phi - \Phi_n\|^2\}\, dt = 0. \tag{2.8}$$

Moreover, $\{\int_T \Phi_n(t, \omega)\, dw\}$ has a unique limit in $L_2(\Omega, K)$. This limit is the *stochastic integral*, $\int_T \Phi(t, \omega)\, dw$.

PROPOSITION 2.9. *If Φ is an element of $\mathscr{M}_1(H, K)$, then $E\{\int_T \Phi(t, \omega)\, dw\} = 0$ and*

$$E\left\{\left\| \int_T \Phi(t, \omega)\, dw \right\|^2\right\} \leqslant \operatorname{tr}(W) \int_T E\{\|\Phi(t)\|^2\}\, dt. \tag{2.10}$$

Proof. A simple limiting argument shows that $E\{\int_T \Phi(t, \omega)\, dw\} = 0$. So let us turn our attention to (2.10).

Let $\{\Phi_n\}$ be an approximating sequence of t-step functions used to define $\int_T \Phi\, dw$. Then

$$E\left\{\left\| \int_T \Phi\, dw \right\|^2\right\} \leqslant E\left\{\left\| \int_T (\Phi - \Phi_n)\, dw \right\| \left\| \int_T \Phi\, dw \right\|\right\}$$
$$+ E\left\{\left\| \int_T \Phi_n\, dw \right\| \left\| \int_T (\Phi - \Phi_n)\, dw \right\|\right\} + E\left\{\left\| \int_T \Phi_n\, dw \right\|^2\right\}.$$

Suppose, for the moment, that (2.7) holds for elements of

$$\mathscr{M}_0(H, K) \cap \mathscr{M}_1(H, K).$$

Then

$$E\left\{\left\| \int_T \Phi\, dw \right\|^2\right\} \leqslant \lim_{n\to\infty} \operatorname{tr}(W) \cdot \int_T E\{\|\Phi_n\|^2\}\, dt \tag{2.11}$$

by virtue of Schwartz's inequality and (2.8). However, (2.8) also implies that

$$\lim_{n\to\infty} \int_T E\{\| \Phi_n \|^2\}\, dt = \int_T E\{\| \Phi \|^2\}\, dt$$

and so, (2.10) follows.

Thus, all that remains is to verify (2.7) for Φ in $\mathscr{M}_0(H, K) \cap \mathscr{M}_1(H, K)$. For such a Φ,

$$\int_T \Phi\, dw = \sum_1^{n-1} \Phi(t_j, \omega)\,(w(t_{j+1}) - w(t_j)), \tag{2.12}$$

where $\{t_1, ..., t_n\}$ is a finite partition of T. Let

$$\Phi_j = \Phi(t_j, \omega), \qquad \Delta w_j = w(t_{j+1}) - w(t_j)$$

and

$$\Delta t_j = t_{j+1} - t_j .$$

Then

$$E\left\{\left\| \int_T \Phi\, dw \right\|^2\right\} = \sum_j \sum_k E\{\langle \Phi_j\, \Delta w_j , \Phi_k\, \Delta w_k \rangle\}. \tag{2.13}$$

But $E\{\langle \Phi_j\, \Delta w_j , \Phi_k\, \Delta w_k\rangle\} = 0$ if $j \neq k$. For, if $j > k$, then

$$E\{\langle \Phi_j\, \Delta w_j , \Phi_k\, \Delta w_k\rangle\} = E\{\langle \Phi_j\, \Delta w_j , \Phi_k\, \Delta w_k\rangle \mid \mathscr{F}_{t_j}\}\} = 0$$

since the conditional expectation $E\{\langle \Phi_j\, \Delta w_j , \Phi_k\, \Delta w_k\rangle \mid \mathscr{F}_{t_j}\}$ vanishes by virtue of the measurability of Φ_j and $\Phi_k\, \Delta w_k$ relative to $\mathscr{F}_{t_j}$ and the independence of Δw_j of $\mathscr{F}_{t_j}$. It follows that

$$E\left\{\left\| \int_T \Phi\, dw \right\|^2\right\} = \sum_j E\{\| \Phi_j\, \Delta w_j \|^2\} = \sum_j E\{\| \Phi_j \|^2\}\, E\{\| \Delta w_j \|^2\}. \tag{2.14}$$

(Note that $\| \Phi_j \|$ and $\| \Delta w_j \|$ are independent since Φ_j is measurable relative to $\mathscr{F}_{t_j}$.) But, $E\{\| \Delta w_j \|^2\} = \operatorname{tr}(W)\, \Delta t_j$ and so,

$$\sum_j E\{\| \Phi_j \|^2\}\, E\{\| \Delta w_j \|^2\} = \operatorname{tr}(W) \sum_j E\{\| \Phi_j \|^2\}\, \Delta t_j = \operatorname{tr}(W) \int_T E\{\| \Phi \|^2\}\, dt.$$

Thus, the proposition is established.

We now prove a useful convergence lemma.

LEMMA 2.15. *Let $\{\Phi_n\}$ be a sequence of elements of $\mathscr{M}_1(H, K)$. Suppose that*

(i) *there is a* Φ *in* $\mathscr{M}_1(H, K)$ *such that* $\Phi_n \to \Phi$ *almost everywhere on* $T \times \Omega$*; and,* (ii) *there is an* $\alpha(t)$ *in* $L_2(T)$ *such that* $\| \Phi_n(t)\| \leqslant \alpha(t)$ w.p.l. *for all n. Then*

$$\lim_{n\to\infty} \int_T \Phi_n(t)\, dw = \int_T \Phi(t)\, dw \tag{2.16}$$

in $L_2(\Omega, K)$.

Proof. Since

$$E\left\{\left\| \int_T (\Phi - \Phi_n)\, dw \right\|^2\right\} \leqslant \mathrm{tr}(W) \int_T E\{\| \Phi - \Phi_n \|^2\}\, dt,$$

the lemma is an immediate consequence of the Lebesgue dominated convergence theorem.

DEFINITION 2.17. Let $u(t)$, $t \in T$, be the K-valued stochastic process given by

$$u(t) - u(T_1) = \int_{T_1}^{t} q(s, \omega)\, ds + \int_{T_1}^{t} \Phi(s, \omega)\, dw \tag{2.18}$$

where $w(t)$ is an H-valued Wiener process, Φ is an element of $\mathscr{M}_1(H, K)$, and $q(s, \omega)$ is a K-valued stochastic process with $\int_T \| q(s, \omega)\|\, ds < \infty$ w.p.l. which is measurable relative to $\mathscr{F}_t$ for all t in T. Then u is said to have the stochastic differential $q dt + \Phi\, dw$ and we write $du = q\, dt + \Phi\, dw$.

We observe that if u has a stochastic differential, then the real stochastic process $\| u(t)\|$ may be viewed as a separable real process since $u(t)$ is continuous in t w.p.l. (this is proved in [2]) and since $u(t)$ is only determined w.p.l. (See [3] for a discussion of the separability of real processes.) This observation will be useful in the sequel and enables us to avoid the question of generalizing the notion of separability for a random process to the Hilbert space context.

3. ITO'S LEMMA

We are now prepared to state and prove Ito's lemma in a Hilbert space context. Since the proof is essentially along the same lines as that given by Gikhman and Skorokhod ([5]) for the finite dimensional case, we omit many of the details. We begin with the following lemma which is an important tool in the proof of the main theorem.

LEMMA 3.1. *Let H, K and G be Hilbert spaces and let* $w(t)$ *be an H-valued Wiener process. Suppose that* (i) $\Theta(t, \omega)$ *is an* $\mathscr{L}(K, \mathscr{L}(K, G))$*-valued stochastic process which is measurable relative to* $\mathscr{F}_t$ *for all t in* T*;* (ii) $E\{\| \Theta(t)\|^2\} < \infty$

for all t in T; (iii) *Φ_0 is an $\mathscr{L}(H, K)$-valued random variable which is measurable relative to $\mathscr{F}_{T_1}$;* (iv) *$E\{\|\Phi_0\|^4\} < \infty$; and,* (v) *$w(t)$ is real, i.e., $w(t) = \overline{w(t)}$ on $T \times \Omega$. Then*

$$E\{\Theta[\Phi_0\, \Delta w, \Phi_0\, \Delta w] \mid \mathscr{F}_s\} = (t - s) \sum_{i=0}^{\infty} \Theta[\Phi_0 \sqrt{\lambda_i} e_i\,, \Phi_0 \sqrt{\lambda_i} e_i] \text{ w.p.l.} \tag{3.2}$$

and

$$E\{\Theta[\Phi_0\, \Delta w, \Phi_0\, \Delta w]\} = (t - s)\, E\left\{\sum_{i=0}^{\infty} \Theta[\Phi_0 \sqrt{\lambda_i} e_i\,, \Phi_0 \sqrt{\lambda_i} e_i]\right\} \tag{3.3}$$

for almost all s, t with $s < t$ where $\Theta = \Theta(s, \omega)$, $\Delta w = w(t) - w(s)$, and the $\{e_i\}$ form an orthonormal basis of H consisting of eigenvectors of W and with the λ_i as corresponding eigenvalues.

Proof. Clearly, (3.3) follows from (3.2).

Now, recall that $w(t) = \sum_{i=0}^{n} \beta_i(t)\, e_i$ and set $\Delta\beta_i = \beta_i(t) - \beta_i(s)$ and $\Delta w_n = w_n(t) - w_n(s)$ for $s < t$ where $w_n(t) = \sum_{i=0}^{n} \beta_i(t)\, e_i$. We note that since $w(t)$ is real, $\Delta\beta_i = \overline{\Delta\beta_i}$, and so, it follows from Schwartz's inequality that

$$E\{|\, \Delta\beta_i\, \Delta\beta_j\, |\} \leqslant E\{|\, \Delta\beta_i\, |^2\}^{1/2}\, E\{|\, \Delta\beta_j\, |^2\}^{1/2} \leqslant (t - s) \sqrt{\lambda_i \lambda_j} \tag{3.4}$$

for all i and j. Since $\|\, \Theta\, \|\, \|\, \Phi_0\, \|^2$ is measurable relative to $\mathscr{F}_s$ and $|\, \Delta\beta_i\, \Delta\beta_j\, |$ is independent of $\mathscr{F}_s$ (as $w(t) - w(s)$ is), we have

$$E\{|\, \Delta\beta_i\, \Delta\beta_j\, |\, \|\, \Theta\, \|\, \|\, \Phi_0\, \|^2\} \leqslant (t - s) \sqrt{\lambda_i \lambda_j} (E\{\|\, \Theta\, \|^2\})^{1/2}\, E\{\|\, \Theta_0\, \|^4\}^{1/2} \tag{3.5}$$

so that $E\{|\, \Delta\beta_i\, \Delta\beta_j\, |\, \|\, \Theta\, \|\, \|\, \Phi_0\, \|^2\}$ is finite. Since $\|\, \Theta[\Phi_0 e_i\,, \Phi_0 e_j]\| \leqslant \|\, \Theta\, \|\, \|\, \Phi_0\, \|^2$ and since $\Theta[\Phi_0 e_i\,, \Phi_0 e_j]$ is measurable relative to $\mathscr{F}_s$, we have

$$E\{H[\Phi_0\, \Delta w_n\,, \Phi_0\, \Delta w_n] \mid \mathscr{F}_s\} = \sum_{i=0}^{n} \sum_{j=0}^{n} E\{\Delta\beta_i\, \Delta\beta_j \mid \mathscr{F}_s\}\, \Theta[\Phi_0 e_i\,, \Phi_0 e_j] \tag{3.6}$$

w.p.l. for all n. But $E\{\Delta\beta_i\, \Delta\beta_j \mid \mathscr{F}_s\} = E\{\Delta\beta_i\, \Delta\beta_j\}$ since $\Delta\beta_i\, \Delta\beta_j$ is independent of $\mathscr{F}_s$. In view of property (iii) of definition 2.1, $E\{\Delta\beta_i\, \Delta\beta_j\} = (t - s)\, \lambda_i \delta_{ij}$ and so,

$$E\{\Theta[\Phi_0\, \Delta w_n\,, \Phi_0\, \Delta w_n]\} = (t - s) \sum_{i=0}^{n} \Theta[\Phi_0 \sqrt{\lambda_i} e_i\,, \Phi_0 \sqrt{\lambda_i} e_i]. \tag{3.7}$$

The result then follows by a simple application of [7], Theorem 2.5.

As a suggestive shorthand, we write $\widetilde{\operatorname{tr}}(\Theta)\,[\Phi_0 \xi_w]$ in place of

$$\sum_{i=0}^{\infty} \Theta[\Phi_0 \sqrt{\lambda_i} e_i\,, \Phi_0 \sqrt{\lambda_i} e_i] \qquad \text{where} \qquad \xi_w = \sum_{i=0}^{\infty} \sqrt{\lambda_i} e_i\,.$$

THEOREM 3.8 (Ito's Lemma). *Let H, K and G be Hilbert spaces and let $w(t)$ be an H-valued Wiener process. Suppose that $g(t, c)$ is a continuous map of $T \times K$ into G and that $u(t)$ is a K-valued stochastic process with stochastic differential $du = q\,dt + \Phi\,dw$ such that*

(i) *$g_t(t, c)$ is continuous on $T \times K$;*

(ii) *$g(t, \cdot)$ is twice differentiable on K for each fixed t in T;*

(iii) *$g_c(t, c)$ and $g_{cc}(t, c)$ are continuous in (t, c) on $T \times K$;*

(iv) *$q(t)$ is a K-valued process which is measurable relative to $\mathscr{F}_t$, $t \in T$, and integrable on T w.p.1. (i.e., $\int_T \| q(s)\|\, ds < \infty$ w.p.1.);*

(v) *Φ is an element of $\mathscr{M}_1(H, K)$ with $\int_T E\{\| \Phi \|^4\}\, dt < \infty$; and,*

(vi) *$w(t)$ is real.*

Then $z(t) = g(t, u(t))$ has the G-valued stochastic differential

$$\begin{aligned} dz = \{g_t(t, u(t)) + g_c(t, u(t))\,[q(t)] + \tfrac{1}{2}\,\widetilde{\operatorname{tr}}(g_{cc}(t, u(t)))\,[\Phi(t)\,\xi_w]\}\,dt \\ + g_c(t, u(t))\,[\Phi(t)]\,dw. \end{aligned} \tag{3.9}$$

Proof (cf. [5]). Let us suppose for the moment that the theorem holds if q and Φ are t-step functions. Then the general case will follow by a straightforward limiting argument. In other words, we consider sequences $\{q_n(t)\}$, $\{\Phi_n(t)\}$ of t-step functions such that

$$\lim_{n\to\infty} \int_T \| q(t) - q_n(t)\|\, dt = 0 \text{ w.p.l.}$$

and

$$\lim_{n\to\infty} \int_T E\{\| \Phi - \Phi_n \|^4\}\, dt = 0 \text{ w.p.l.}$$

where the $q_n(t)$ satisfy (iv) and the $\Phi_n(t)$ satisfy (v). Letting $u_n(t)$ be the K-valued process with stochastic differential $du_n = q_n\,dt + \Phi_n\,dw$, we can show just as in [5] that $u_n(t)$ converges uniformly to $u(t)$ on T w.p.l., i.e., that $\lim_{n\to\infty} \sup_{t\in T}\{\| u_n(t) - u(t)\|\} = 0$ w.p.l. It follows that there is a subsequence $\{u_{n_i}(t)\}$ of $\{u_n(t)\}$ such that

$$g(t, u_{n_i}(t)) \to g(t, u(t)), \qquad g_t(t, u_{n_i}(t)) \to g_t(t, u(t)),$$

$$g_c(t, u_{n_i}(t)) \to g_c(t, u(t)), \qquad \text{and} \qquad g_{cc}(t, u_{n_i}(t)) \to g_{cc}(t, u(t)),$$

all uniformly on T w.p.l. For simplicity, we write $v_i(t) = u_{n_i}(t)$, $r_i(t) = q_{n_i}(t)$ and $\Psi_i(t) = \Phi_{n_i}(t)$. Then, simple inequality computations show that

$$\begin{gathered}\lim_{i\to\infty} \int_{T_1}^{t} g_t(s, v_i(s))\, ds = \int_{T_1}^{t} g_t(s, u(s))\, ds \\ \lim_{i\to\infty} \int_{T_1}^{t} g_c(s, v_i(s))\, [r_i(s)]\, ds = \int_{T_1}^{t} g_c(s, u(s))\, [q(s)]\, ds \\ \lim_{i\to\infty} \int_{T_1}^{t} \widetilde{\operatorname{tr}}(g_{cc}(s, v_i(s)))\, [\Psi_i(s)\, \xi_w]\, ds = \int_{T_1}^{t} \widetilde{\operatorname{tr}}(g_{cc}(s, u(s)))\, [\Phi(s)\, \xi_w]\, ds\end{gathered} \tag{3.10}$$

all w.p.l. Thus, to show that $z(t) = g(t, u(t))$ has the required stochastic differential, it will be enough to prove that

$$\lim_{i\to\infty} \int_{T_1}^{t} g_c(s, v_i(s))\, [\Psi_i(s)]\, dw = \int_{T_1}^{t} g_c(s, u(s))\, [\Phi(s)]\, dw \tag{3.11}$$

where the convergence is in probability.

Now let $\chi^{(N)}(\cdot)$ be the real random variable given by

$$\chi^{(N)}(t) = \begin{cases} 1 & \text{if} \quad \| u(s)\| \leqslant N \quad \text{for} \quad T_1 \leqslant s \leqslant t \\ 0 & \text{otherwise.} \end{cases} \tag{3.12}$$

Then, for sufficiently large i,

$$g_c(s, u(s))\, [\Phi(s)\, \chi^{(N)}(s)] \qquad \text{and} \qquad g_c(s, v_i(s))\, [\Psi_i(s)\, \chi^{(N)}(s)]$$

will be elements of $\mathscr{M}_1(H, G)$. It then follows from the inequality (2.10) that

$$\lim_{i\to\infty} \int_{T_1}^{t} g_c(s, v_i(s))\, [\Psi_i(s)\, \chi^{(N)}(s)]\, dw = \int_{T_1}^{t} g_c(s, u(s))\, [\Phi(s)\, \chi^{(N)}(s)]\, dw \tag{3.13}$$

for all finite N (where the convergence is in probability). But

$$\begin{aligned}\mu \Big\{ \Big\| &\int_{T_1}^{t} \{g_c(s, v_i(s))\, [\Psi_i(s)] - g_c(s, u(s))\, [\Phi(s)]\}\, dw \\ &- \int_{T_1}^{t} \{g_c(s, v_i(s))\, [\Psi_i(s)\, \chi^{(N)}(s)] - g_c(s, u(s))\, [\Phi(s)\, \chi^{(N)}(s)]\}\, dw \Big\| \neq 0 \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{for } i = 1, 2, \ldots \Big\} \\ &\leqslant \mu\{\omega : \sup_{t\in T} \| u(t, \omega)\| > N\}.^{5}\end{aligned}$$

[5] Note that this probability exists since $\| u(t)\|$ may be viewed as a separable real process.

Since $\mu\{\omega : \sup_{t\in T} \| u(t, \omega)\| > N\}$ goes to zero as N approaches infinity, (3.11) is established.

Now it remains to prove the theorem for case of t-step functions q and Φ. To do this, it will be enough to prove the theorem for the special case where q and Φ are constant, i.e., are independent of t.

So let us assume that q and Φ are constant. Let t be a fixed (but arbitrary) element of $(T_1, T_2]$ and let $t_0, t_1, ..., t_n$ be elements of T with

$$T_1 = t_0 < t_1 < t_2 < \cdots < t_n = t \leqslant T_2 .$$

We set $\Delta t_k = t_{k+1} - t_k$, $u_k = u(t_k)$, $g_k = g(t_k, u_k)$, $\Delta u_k = u_{k+1} - u_k$, and $\Delta g_k = g_{k+1} - g_k$ for $k = 0, 1, ..., n-1$. Then

$$\begin{aligned} z(t) - z(T_1) &= g(t, u(t)) - g(T_1, u(T_1)) \\ &= \sum_{k=0}^{n-1} \{(g_{k+1} - g(t_k, u_{k+1})) + (g(t_k, u_{k+1}) - g_k)\} \qquad (3.14) \\ &= \sum_{k=0}^{n-1} \Delta g_k . \end{aligned}$$

In view of the differentiability assumptions,

$$\Delta g_k = g_t(t_k, u_{k+1}) \Delta t_k + g_c(t_k, u_k)[\Delta u_k] + \tfrac{1}{2} g_{cc}(t_k, u_k)[\Delta u_k, \Delta u_k] + \gamma_k + \delta_k \tag{3.15}$$

where

$$\| \gamma_k \| \leqslant \Delta t_k \sup_{0<\theta<1} \| g_t(t_k + \theta \Delta t_k, u_{k+1}) - g_t(t_k, u_{k+1})\|$$

and

$$\| \delta_k \| \leqslant \| \Delta u_k \|^2 \sup_{0<\theta<1} \| g_{cc}(t_k, u_k + \theta \Delta u_k) - g_{cc}(t_k, u_k)\| .$$

Just as in [5], $\sum_{k=0}^{n-1} (\| \gamma_k \| + \| \delta_k \|) \to 0$ w.p.l. as $\max_k \Delta t_k \to 0$. It follows that

$$\begin{aligned} z(t) - z(T_1) = \sum_{0}^{n-1} \{ & g_t(t_k, u_k) \Delta t_k + g_c(t_k, u_k)[\Delta u_k] \\ & + \tfrac{1}{2} g_{cc}(t_k, u_k)[\Delta u_k, \Delta u_k]\} + \theta_n , \end{aligned} \tag{3.16}$$

where $\| \theta_n \| \to 0$ w.p.l. as $\max_k \Delta t_k \to 0$. Substituting $\Delta u_k = q\Delta t_k + \Phi \Delta w_k$ in (3.16), we obtain the relation

$$z(t) - z(T_1) = \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4 + \Sigma_5 + \theta_n \tag{3.17}$$

where

$$\Sigma_1 = \sum_{k=0}^{n-1} (g_t(t_k, u_k) + g_c(t_k, u_k)[q])\, \Delta t_k \tag{3.18}$$

$$\Sigma_2 = \sum_{k=0}^{n-1} g_c(t_k, u_k)[\Phi]\, \Delta w_k \tag{3.19}$$

$$\Sigma_3 = \sum_{k=0}^{n-1} \tfrac{1}{2} g_{cc}(t_k, u_k)[\Phi\, \Delta w_k, \Phi\, \Delta w_k] \tag{3.20}$$

$$\Sigma_4 = \sum_{k=0}^{n-1} \tfrac{1}{2} g_{cc}(t_k, u_k)[q, q]\, (\Delta t_k)^2 \tag{3.21}$$

$$\Sigma_5 = \sum_{k=0}^{n-1} g_{cc}(t_k, u_k)[q\, \Delta t_k, \Phi\, \Delta w_k]. \tag{3.22}$$

In view of the continuity assumptions and the boundedness of $u(t)$ on T, we immediately deduce that, as $\max_k \Delta t_k$ goes to 0,

$$\Sigma_1 \to \int_{T_1}^{t} (g_t(t, u(t)) + g_c(t, u(t))[q])\, dt \text{ w.p.l.} \tag{3.23}$$

$$\Sigma_4 \to 0 \text{ w.p.l.} \tag{3.24}$$

$$\Sigma_5 \to 0 \text{ w.p.l.} \tag{3.25}$$

In other words, the limiting sums converge to the usual Bochner integral.

We now claim that

$$\Sigma_2 \to \int_{T_1}^{t} g_c(t, u(t))[\Phi]\, dw \text{ in probability} \tag{3.26}$$

and

$$\Sigma_3 \to \tfrac{1}{2} \int_{T_1}^{t} \widetilde{\mathrm{tr}}(g_{cc}(t, u(t)))[\Phi \xi_w]\, dt \text{ in probability} \tag{3.27}$$

as $\max_k \Delta t_k \to 0$. Let $\chi^{(N)}(\cdot)$ be the real random variable given by (3.12). Then

$$\int_T E\{\| g_c(t, u(t))[\Phi \chi^{(N)}(t)]\|^2\}\, dt \leqslant \sup_{\|u(t)\| \leqslant N} \{\| g_c(t, u(t))\|^2\} \int_T E\{\| \Phi \|^2\}\, dt < \infty$$

for all N by virtue of assumptions (iii) and (iv). Since $u(t)$ is bounded w.p.l., it follows that $\int_T E\{\| g_c(t, u(t))[\Phi]\|^2\}\, dt < \infty$ and hence, that $g_c(t, u(t))[\Phi]$

is in $\mathscr{M}_1(H, G)$ as $g_c(t, u(t))[\Phi]$ is measurable relative to $\mathscr{F}_t$ for all t in T. Thus, $\int_{T_1}^t g_c(t, u(t))[\Phi]\, dw$ exists. Consider the sequence $\{g_c^n[\Phi]\}$ where

$$g_c^n(t, u(t))[\Phi] = g_c(t_j, u_j)[\Phi]$$

for t in $[t_j, t_{j+1})$, $j = 0,..., n-1$. Then $g_c^n[\Phi]$ is an element of $\mathscr{M}_1(H, G)$ for every n. Moreover, since g_c is continuous, $\| g_c^n[\Phi]\| \leqslant M$ w.p.l. for all n and some constant M. Thus, Lemma 2.15 applies and (3.26) is established.

Finally, we prove that (3.27) holds. Let $\chi_k^{(N)}$ be given by

$$\chi_k^{(N)} = \begin{cases} 1 & \text{if} \quad \| u_i \| \leqslant N \quad \text{for} \quad i \leqslant k \\ 0 & \text{otherwise.} \end{cases} \tag{3.28}$$

Then

$$E\{\| g_{cc}(t_k, u_k)\, \chi_k^{(N)} \|^2\} \leqslant \sup_{\|u(t)\| \leqslant N} \| g_{cc}(t, u(t))\|^2 < \infty \qquad \text{for all} \qquad N.$$

Since $g_{cc}(t_k, u_k)\, \chi_k^{(N)}$ is measurable relative to $\mathscr{F}_{t_k}$, we deduce from (3.3) that

$$E\{g_{cc}(t_k, u_k)[\Phi\, \Delta w_k, \Phi\, \Delta w_k]\, \chi_k^{(N)}\} = \Delta t_k E\{\widetilde{\mathrm{tr}}(g_{cc}(t_k, u_k)) \chi_k^{(N)} [\Phi\, \xi_w]\}.$$

Setting

$$p_k = g_{cc}(t_k, u_k)[\Phi\, \Delta w_k, \Phi\, \Delta w_k]$$

and

$$v_k = p_k - \Delta t_k\, \widetilde{\mathrm{tr}}(g_{cc}(t_k, u_k))[\Phi \xi_w],$$

we have $E\{v_k \chi_k^{(N)}\} = 0$ for all N. Moreover, by virtue of (3.2),

$$E\{p_k \chi_k^{(N)} \mid \mathscr{F}_{t_k}\} = \Delta t_k\, \widetilde{\mathrm{tr}}(g_{cc}(t_k, u_k))[\Phi \xi_w]\, \chi_k^{(N)} \tag{3.29}$$

and so,

$$E\{v_k \chi_k^{(N)} \mid \mathscr{F}_{t_k}\} = 0 \tag{3.30}$$

for all N. [Note that $g_{cc}(t_k, u_k)[\Phi \Delta w_k, \Phi \Delta w_k]$ is measurable relative to $\mathscr{F}_{t_k}$ and that

$$E\{\| g_{cc}(t_k, u_k)[\Phi\, \Delta w_k, \Phi\, \Delta w_k]\, \chi_k^{(N)} \|\}$$

$$\leqslant \mathrm{tr}(W)\, E\{\| \Phi \|^2\} \sup_{\|u(t)\| \leqslant N} \| g_{cc}(t, u(t))\| < \infty \text{ for all } N.]$$

It is clear that both $E\{\| p_k \chi_k^{(N)} \|^2\}$ and $E\{\| v_k \chi_k^{(N)} \|^2\}$ are finite for all N and k.

Thus, if $j > k$, we have (by Proposition A.9 and (3.30))

$$E\{\langle v_j\chi_j^{(N)}, v_k\chi_k^{(N)}\rangle \mid \mathscr{F}_{t_j}\} = 0 \tag{3.31}$$

since $v_k\chi_k^{(N)}$ is measurable relative to $\mathscr{F}_{t_j}$. It follows that

$$E\{\langle v_j\chi_j^{(N)}, v_k\chi_k^{(N)}\rangle\} = 0$$

if $j \neq k$ and hence that

$$E\left\{\left\|\sum_0^{n-1} v_k\chi_k^{(N)}\right\|^2\right\} = \sum E\{\| v_k\chi_k^{(N)} \|^2\}. \tag{3.32}$$

A simple computation using the independence of $\| \Phi \|$ and $\| \Delta w_k \|$ leads to the estimate

$$\sum E\{\| v_k\chi_k^{(N)} \|^2\} \leqslant 12 \sup_{\|u(t)\| \leqslant N} \{\| g_{cc}(t,u(t))\| \}\, E\{\| \Phi \|^4\} \operatorname{tr}(W) \sum (\Delta t_k)^2. \tag{3.33}$$

We immediately conclude that $\sum_{k=0}^{n-1} v_k\chi_k^{(N)} \to 0$ in probability as $\max_k \Delta t_k \to 0$. Since

$$\mu\left\{\left\|\sum_0^{n-1} (v_k - v_k\chi_k^{(N)})\right\| \neq 0\right\} \leqslant \mu\{\sup \| u(t)\| > N\},$$

we also see that $\sum_{k=0}^{n-1} v_k \to 0$ in probability as $\max_k \Delta t_k \to 0$. But

$$\sum_{k=0}^{n-1} \Delta t_k \, \widetilde{\operatorname{tr}}(g_{cc}(t_k, u_k))\, [\Phi \xi_w]$$

is an approximating sum for the integral $\int_{T_1}^t \widetilde{\operatorname{tr}}(g_{cc}(t, u(t))\, [\Phi \xi_w]\, dt$. Thus, (3.27) and with it, the theorem, are established.

COROLLARY 3.34. *Suppose that, in addition to the hypotheses of the theorem, $H = K$ and $G = \mathfrak{C}$ (or R). Then dz can be written in form*

$$\begin{aligned} dz = {}& \{g_t(t, u(t)) + \langle q(t), \nabla_c g(t, u(t))\rangle + \tfrac{1}{2} \operatorname{tr}[\Phi(t)\, W\Phi^*(t)\, \Theta_{cc} g(t, u(t))]\}\, dt \\ & + \langle \Phi^*(t)\, \nabla_c g(t, u(t)), dw\rangle \end{aligned} \tag{3.35}$$

where $\nabla_c g$ and $\Theta_{cc} g$ are the gradient and Hessian, respectively, of g with respect to c.

APPENDIX. Infinite Dimensional Random Variables

We collect some of the standard definitions and results of the theory of Banach space valued random variables in this appendix as a convenience for the reader. The treatment is along the lines of that given by Scalora [7].

Let $(\Omega, \mathscr{P}, \mu)$ be a probability space with $\mathscr{P}$ as Borel field and μ as measure. We assume that μ is complete. Also, let X be a Banach space. We then have

Definition A.1. A strongly measurable mapping $x(\cdot)$ of Ω into X is called a random variable.

A random variable $x(\cdot)$ is *integrable* on Ω if, and only if, there is a sequence $\{x_n(\cdot)\}$ of finitely valued random variables such that (i) $x_n(\cdot)$ converges to $x(\cdot)$ almost everywhere, and (ii) $\lim_{m,n\to\infty} \int_\Omega \| x_n(\omega) - x_m(\omega)\| \, d\mu = 0$.

Definition A.2. If $x(\cdot)$ is integrable on Ω, then the expectation of x, $E\{x\}$, is the element of X given by

$$E\{x\} = \int_\Omega x(\omega)\, d\mu = \lim_{n\to\infty} \int_\Omega x_n(\omega)\, d\mu. \tag{A.3}$$

Definition A.4. Let $\mathscr{F}$ be a Borel field with $\mathscr{F} \subset \mathscr{P}$ and let $x(\cdot)$ be integrable on Ω. The conditional expectation of x relative to $\mathscr{F}$, $E\{x \mid \mathscr{F}\}$, is a random variable such that

$$\int_F x(\omega)\, d\mu = \int_F E\{x \mid \mathscr{F}\}(\omega)\, d\mu \tag{A.5}$$

for all F in $\mathscr{F}$.

We note that $E\{x \mid \mathscr{F}\}$ is unique w.p.1., is integrable on Ω, and is measurable relative to $\mathscr{F}$.

Definition A.6. Let $T = [T_1, T_2]$ be a finite interval. A mapping $x(t, \omega)$ of $T \times \Omega$ into X is called a stochastic process on T if $x(\cdot, \cdot)$ is measurable in the pair (t, ω) (using Lebesgue measure on T).

Definition A.6 is more restrictive than the usual one (cf. Doob [3]) but is adequate for our purposes. Also, we usually write $x(t)$ in place of $x(t, \omega)$ when discussing stochastic processes.

Definition A.7. Two measurable sets F_1 and F_2 in $\mathscr{P}$ are independent if $\mu(F_1 \cap F_2) = \mu(F_1)\, \mu(F_2)$. If $x(\cdot)$ is a random variable mapping Ω into X and $y(\cdot)$ is a random variable mapping Ω into Y, then $x(\cdot)$ and $y(\cdot)$ are independent if the sets $\{\omega : x(\omega) \in A\}$ and $\{\omega : y(\omega) \in B\}$ are independent for all Borel sets A of X and all Borel sets B of Y. Finally, a random variable $x(\cdot)$ is independent of the Borel field $\mathscr{F} \subset \mathscr{P}$ if the sets F and $\{\omega : x(\omega) \in A\}$ are independent for all F in $\mathscr{F}$ and all Borel sets A of X.

The following propositions contain various results needed in the paper. These propositions are easy extensions of similar results for the ordinary case and are proven in detail in [2].

PROPOSITION A.8. *If $x(\cdot)$ and $y(\cdot)$ are independent X and Y valued random variables, respectively, and if f and g are nonrandom Baire functions mapping X and Y, respectively, into the complex numbers $\mathfrak{C}$, then $f(x(\cdot))$ and $g(y(\cdot))$ are independent random variables.*

PROPOSITION A.9. *Let $\mathscr{F}$ be a Borel field with $\mathscr{F} \subset \mathscr{P}$. Let f, x and Φ be random variables on Ω to $\mathfrak{C}$, X and $\mathscr{L}(X, Y)$, respectively. Then*

(i) *if $E\{\| x \|\} < \infty$, then $E\{E\{x \mid \mathscr{F}\}\} = E\{x\}$;*

(ii) *if $E\{\| x \|\} < \infty$ and x is measurable relative to $\mathscr{F}$, then $E\{x \mid \mathscr{F}\} = x$* w.p.l.;

(iii) *if $E\{\| x \|\} < \infty$, $E\{| f | \, \| x \|\} < \infty$, and x is measurable relative to $\mathscr{F}$, then $E\{fx \mid \mathscr{F}\} = E\{f \mid \mathscr{F}\}\, x$* w.p.l.; *and*

(iv) *if $E\{\| x \|\} < \infty$, $E\{\| x \| \, \| \Phi \|\} < \infty$ and Φ is measurable relative to $\mathscr{F}$, then $E\{\Phi x \mid \mathscr{F}\} = \Phi E\{x \mid \mathscr{F}\}$* w.p.l.

REFERENCES

1. E. CABANA, Stochastic Integration in Separable Hilbert Spaces, Publ. Inst. de Matematica y Estadistica, Uruguay, **IV** (1966), 49–79.
2. R. F. CURTAIN, "Stochastic Differential Equations in a Hilbert Space," Thesis, Brown University, 1969.
3. J. L. DOOB, "Stochastic Processes," Wiley, New York, 1953.
4. P. L. FALB, Infinite dimensional filtering: The Kalman–Bucy filter in Hilbert space, *Information and Control* 11 (1967), 102–137.
5. I. I. GIKHMAN AND A. V. SKOROKHOD, "Introduction to the Theory of Random Processes" (Russ.), Izd.-vo "Nauka", Moscow, 1965.
6. T. KATO, "Perturbation Theory for Linear Operators," Springer-Verlag, Berlin, 1966.
7. F. S. SCALORA, Abstract Martingale Convergence Theorems, *Pacific J. Math.* 11 (1961), 347–374.
8. A. V. SKOROKHOD, "Studies in the Theory of Random Processes," Addison-Wesley, Reading, Mass., 1965.