PROBABILITY THEORY
BANACH CENTER PUBLICATIONS, VOLUME 5
PWN—POLISH SCIENTIFIC PUBLISHERS
WARSAW 1979

STOCHASTIC DIFFERENTIAL EQUATIONS IN HILBERT SPACES

ANNA CHOJNOWSKA-MICHALIK

Institute of Mathematics, Polish Academy of Sciences, Warsaw, Poland

§ 1. Introduction

Many types of stochastic partial differential equations, some delay equations, and certain filtering problems have a natural interpretation through infinite-dimensional stochastic differential systems with an unbounded operator. This is the motivation behind considering stochastic evolution equations in a Hilbert space. Such equations with state-independent noise have been studied in many papers, for example in [2], [3], by the Lions approach [10] and in [1], [5], [6], [12], [16], by a semigroup approach [8].

We develop the second approach to the study of time-invariant stochastic differential equations in a Hilbert space H, with state-dependent noise. The system under consideration can be formally written as

$$(\blacksquare)\qquad \begin{aligned} dX_t &= AX_t dt + d\mathcal{M}(X)_t, \\ X_0 &= \zeta, \quad t \in [0, T]. \end{aligned}$$

(Here A is the infinitesimal generator of a semigroup T_t on H and $\mathcal{M}(\cdot)$ is a transformation from a space of stochastic processes into a space of H-valued martingales.)

Four types of interpretation of the solution of equation ($\blacksquare$) are possible in the semigroup approach context:

(I) strong solution: $X_t = \zeta + \int_0^t AX_s ds + \mathcal{M}(X)_t$;

(II) weakened solution: $X_t = \zeta + A\int_0^t X_s ds + \mathcal{M}(X)_t$;

(III) mild solution: for any $y \in D(A^*)$ (domain of the adjoint of A)

$$\langle X_t, y\rangle = \langle \zeta, y\rangle + \int_0^t \langle X_s, A^*y\rangle ds + \langle \mathcal{M}(X)_t, y\rangle;$$

(IV) mild integral solution: $X_t = T_t\zeta + \int_0^t T_{t-s} d\mathcal{M}(X)_s$.

In studing these equations we use the recent very nice results on integrals w.r.t. Hilbert space-valued martingales due to Metivier and Pistone [12].

Making use of a certain generalization of the Fubini theorem (§ 3.2), we prove in a straightforward manner the main result, namely that the solutions (II), (III) and (IV) are equivalent (§ 3.3).

Then in the case of a constant transformation $\mathcal{M}$ we obtain the following corollaries (§ 3.4.1):

(1) uniqueness and existence theorem for equations (II) and (III) (Let us note that for (III) we get uniqueness in the class of progressively measurable processes with trajectories belonging to $L^1([0, T], H)$. A similar but weaker result (uniqueness in the class of progressively measurable processes with continuous trajectories) is due to Vinter [16].);

(2) a uniqueness theorem for equation (I);

(3) under some assumptions—a theorem on the existence and uniqueness of solutions of the strong equation (I), i.e. some results obtained by Curtain and Falb [5], Metivier and Pistone [12].

In general the equivalence theorem provides motivation for considering the mild integral equation, which in some situation is more convenient than equations (I), (II), (III). Therefore this theorem may also be important in the study of certain stochastic control problems or of the stability of infinite-dimensional stochastic models ([17]). Similar motivation underlies the "integral equations method" in the study of certain stochastic partial differential equations in [15].

In the case of nonconstant $\mathcal{M}(\cdot)$ we turn our attention to stochastic systems with state-dependent Gaussian white noise. Existence and uniqueness results for such systems (even in the time-varying case) are also obtained by Pardoux [14], who uses the Lions aproach. But for time-invariant linear equations the assumptions in [14] are stronger than ours. (They imply that A is the generator of a semigroup arrising in connection with a coercive bilinear form.) In the spirit of the preceding remarks the second main result (§ 3.3) is a theorem about the existence and uniqueness of solution of a suitable mild integral equation. This theorem is an analogue of certain well-known results for $\boldsymbol{R}^n$-valued stochastic equations. As a corollary of the main theorems, the existence and uniqueness of solutions of (II) and (III) as well as the uniqueness theorem for (I) are obtained (§ 3.4.2). Some sufficient conditions for the existence of solutions of (I) are also supplied (§ 3.4.2).

The results of this paper may be applied to linear stochastic delay systems ([17], [4]).

Acknowledgements. I would like to thank Dr J. Zabczyk for helpful discussions and suggestions.

§ 2. Preliminaries

ASSUMPTION 1. Let H, G be real separable Hilbert spaces. We shall denote the scalar product and the norm by $\langle \cdot, \cdot \rangle$ and $|\cdot|$, respectively.

$\mathcal{L}(H, G)$ is, as usual, the space of linear bounded operators: $H \to G$. The operator norm will be denoted by $\|\cdot\|$.

$\mathcal{L}_2(H, G) \subset \mathcal{L}(H, G)$ is the space of Hilbert–Schmidt operators, with the Hilbert–Schmidt norm denoted by $\|\cdot\|_{\mathrm{HS}}$.

Let us recall that every $B \in \mathcal{L}_2(H, G)$ has the representation: $B(h) = \sum \lambda_{ij} \langle h, e_i \rangle f_j$, where $(e_i), (f_j)$ are orthonormal bases in H, G and $\|B\|_{\mathrm{HS}}^2 = \sum_{i,j} |\lambda_{ij}|^2$.

2.1. Operator-valued functions-measurability and integration. Since we shall consider operator-valued functions (or processes), the following two types of measurability will be useful:

DEFINITION 2.1. Let $(S, \mathcal{S})$ be a measure space and $f\colon S \to \mathcal{L}(H, G)$.

(a) f will be called *measurable* (or *strongly measurable*) iff it is measurable w.r.t. the operator topology in $\mathcal{L}(H, G)$.

(b) f will be called *point-measurable* iff for any $h \in H$ $f(h)$ is measurable (strongly $\approx$ weakly as G is separable), equivalently $\forall h \in H$ $f(h)\colon (S, \mathcal{S}) \to (G, \mathfrak{B}_G)$.

Strong measurability is not very convenient to work with. The space $(\mathcal{L}(H, G), \|\cdot\|)$ is, in general, not separable and so strong measurability and weak measurability do not coincide. Moreover, many $\mathcal{L}(H, G)$-valued functions are not measurable but only point-measurable. For example: If $(T_t)_{t \geq 0}$ is a C_0-semigroup which is not continuous in the operator norm on $]0, \infty[$, then (T_t) is not strongly measurable.

Remark 2.2. If $f\colon (S, \mathcal{S}) \to \mathcal{L}_2(H, G)$, then, under Assumption 1, f is measurable w.r.t. HS-topology iff f is point-measurable.

Proof. $\Rightarrow$ is obvious.

$\Leftarrow$: It follows from the very definition of HS-norm that $\|f\|_{\mathrm{HS}}\colon (S, \mathcal{S}) \to (\boldsymbol{R}, \mathfrak{B})$. Since $\mathcal{L}_2(H, G)$ is separable, this implies the measurability: $f\colon (S, \mathcal{S}) \to (\mathcal{L}_2(H, G), \mathfrak{B}_{\|\ \|_{\mathrm{HS}}})$.

This is the reason for using the HS-norm of operators here.

We shall often want to integrate point-measurable functions; then the following definition will be needed:

DEFINITION 2.3. Let $\varphi\colon (\boldsymbol{R}, \mathfrak{B}) \to \mathcal{L}(H, G)$ be point-measurable and such that

$$\sup_{|h| \leq 1} \int_0^t |\varphi_s(h)|\, ds < +\infty.$$

Then the integral in the Bochner sense, $\int_0^t \varphi_s(h)\, ds$ exists $\forall h \in H$. We define $\int_0^t \varphi_s ds$ as follows:

$$\forall h \in H \left(\int_0^t \varphi_s ds\right)(h) = \int_0^t \varphi_s(h)\, ds.$$

Then

$$\int_0^t \varphi_s\,ds \in \mathscr{L}(H, G) \quad \text{and} \quad \Big\|\int_0^t \varphi_s\,ds\Big\| \leqslant \sup_{|h|\leqslant 1}\int_0^t |\varphi_s(h)|\,ds.$$

In connection with Remark 2.2 we have the estimation

$$\Big\|\int_0^t \varphi_s\,ds\Big\|_{\mathrm{HS}} \leqslant \int_0^t \|\varphi_s\|_{\mathrm{HS}}\,ds \leqslant +\infty.$$

It is obvious that if φ is strongly measurable and $\int_0^t \|\varphi_s\|\,ds < \infty$, the integral so defined coincides with the Bochner integral.

2.2. Semigroups of operators. Let $(T_t)_{t\geqslant 0}$ be a C_0-semigroup of operators belonging to $\mathscr{L}(H, H)$ (i.e. $T_{t+s} = T_t\cdot T_s = T_s\cdot T_t$ and $\forall h \in H$ $|T_t h - h| \to 0$ when $t\to 0$), with A as its infinitesimal generator. Certain properties of (T_t) and A are mentioned here. To get acquainted with the semigroup theory the reader is referred to [8].

2.4. If $T < \infty$, then $\sup_{t\in[0,T]} \|T_t\| < +\infty$.

2.5. A is a closed operator.

2.6. For any $n = 1, 2, \ldots$ the domain of $A^n\big(D(A^n)\big)$ is dense in H.

2.7. For $h \in D(A)$ define the graph norm:

$$|h|_{D(A)} = (|h|^2 + |Ah|^2)^{1/2}.$$

Then $|\cdot|_{D(A)}$ is a well-defined norm in $D(A)$, and it follows from Assumption 1, 2.5 and 2.6 that $(D(A), |\cdot|_{D(A)})$ is a separable Hilbert space. This space will be denoted by $\boldsymbol{D}$.

2.8. $\forall t$ $\ T_t(D(A)) \subset D(A)$ and for $h \in D(A)$: $T_t A h = A T_t h$.

2.9. $\forall h \in D(A)$ $\int_0^t A T_s(h)\,ds = T_t h - h$.

2.10. $D(A)$ is a Borel subset of H and A: $\big(D(A), \mathfrak{B}_H|D(A)\big) \to (H, \mathfrak{B}_H)$. Consequently, $\mathfrak{B}_D = \mathfrak{B}_H|D(A)$.

2.11. Let f: $[0, T] \to D(A)$ be measurable (see 2.10) and $\int_0^T |f(s)|_{D(A)}\,ds < \infty$. Then, for any $t \in [0, T]$,

$$\int_0^t f(s)\,ds \in D(A) \quad \text{and} \quad \int_0^t Af(s)\,ds = A\int_0^t f(s)\,ds.$$

2.12 (an extension of 2.9). For any $h \in H$,

$$\int_0^t T_s(h) \in D(A) \quad \text{and} \quad A\int_0^t T_s(h)\,ds = T_t(h) - h.$$

2.13. Under Assumption 1 the adjoint semigroup (T_t^*) is also a C_0-semigroup on H, with A^* (the adjoint of the operator A) as its infinitesimal generator.

2.3. Separable Hilbert space-valued martingales.

ASSUMPTION 2. Let $(\Omega, \mathscr{F}, P)$ be a complete probability space on which is defined an increasing and right-continuous family $(\mathscr{F}_t)_{t\geqslant 0}$ of complete sub-σ-algebras of $\mathscr{F}$.

$\mathscr{R}$ is, as usual, the class of predictable rectangles on $\boldsymbol{R}_+ \times \Omega$, i.e.

$$\mathscr{R} = \begin{cases}]s, t]\times F, & \text{where} \quad F \in \mathscr{F}_s,\ s < t,\\ \{0\}\times F, & \text{where} \quad F\in\mathscr{F}_0. \end{cases}$$

By $\mathscr{P}$ we denote the predictable σ-field. It is known that $\mathscr{P} = \sigma(\mathscr{R})$. Let

$$\mathfrak{M}^2_{[0,T]}(H) = \{M\colon H\text{-valued, right-continuous martingales adapted to } (\mathscr{F}_t) \text{ such that } M_0 \equiv 0;\ E|M_T|_H^2 < \infty\}.$$

We shall recall some facts about H-valued martingales from [12].

THEOREM 2.14 (see [12], th.1). *Let $M \in \mathfrak{M}^2_{[0,T]}(H)$ and let λ be Dolean's measure of $|M|^2$ (i.e. a measure on $\mathscr{P}$ such that on $\mathscr{R}$: $\lambda(]s, t]\times F) = E[\chi_F(|M_t|^2 - |M_s|^2)]$). There exists a unique $\mathscr{L}(H, H)$-valued, point-predictable process Q such that*

(a) *$Q(s, w)$ is a nuclear, positive, symmetric operator with* $\operatorname{Tr} Q(s, w) = 1$ λ a.e.;

(b) $\forall\,]s, t]\times F \in \mathscr{R}$ $E[\chi_F(M_t - M_s)^{\otimes 2}] = \int_{]s,t]\times F} Q(s, w)\,d\lambda$.(¹)

The process Q will be called a *covariance process of the martingale M*.

By the previous theorem and Remark 2.2 the following class of processes is well-defined:

$$\Lambda_t(M; H, G) = \Big\{\Psi(\cdot,\cdot)\colon \Psi \text{ is an } \mathscr{L}(H, G)\text{-valued, point-predictable process such that } \|\Psi\|_{\Lambda_t} \stackrel{\text{def}}{=} \Big(\int_{\Omega\times[0,t]} \|\Psi\cdot Q^{1/2}\|^2_{\mathrm{HS}}\,d\lambda\Big)^{1/2} < \infty\Big\},$$

and $\|\cdot\|_{\Lambda_t}$ is a prenorm.

The following two propositions have also been proved in [12]:

PROPOSITION 2.15. *Let $\mathscr{E}$ be a class of $\mathscr{L}(H, G)$-valued, strongly predictable, step processes, i.e.*

$$\varphi\in\mathscr{E} \quad \textit{iff} \quad \varphi = \sum_{i=1}^N \chi_{]s_i, t_i]\times F_i}(s, \omega)\varphi_i, \quad \textit{where} \quad]s_i, t_i]\times F_i \in \mathscr{R}.$$

Then $\mathscr{E}$ is dense in Λ_t in the prenorm $\|\cdot\|_{\Lambda_t}$.

PROPOSITION 2.16. *For any $\Psi \in \Lambda_t$ the stochastic integral $\int_0^t \Psi(s, \omega)\,dM_s$ is a well-defined G-valued process such that*

(¹) The customary notation is used: let $h\in H$, $g\in G$; then $h\otimes g\in\mathscr{L}(H, G)$ and $(h\otimes g)(x) \stackrel{\text{def}}{=} \langle h, x\rangle_H g$ for $x\in H$.

(a) $\int_0^t \Psi(s,\omega)\,dM_s$ *is a right-continuous martingale* (*continuous if* M_t *is continuous*),

(b) $E\left|\int_0^t \Psi\,dM_s\right|^2 = \|\Psi\|^2_{\Lambda_t}$.

Here is an important example:

Example 2.17. Let $(W_t)_{t\geqslant 0}$ be an H-valued Wiener process (see [2], [5] for the properties of W). In this case $\lambda = l\times P$ (where l is the Lebesgue measure on $\boldsymbol{R}_+$) and $Q(s,\omega)\equiv Q$ is a constant covariance operator. Then Q has the representation $Q = \sum \alpha_i e_i\otimes e_i$ (where α_i are eigenvalues of Q, $\alpha_i\geqslant 0$, $\sum\alpha_i<\infty$ and $\{e_i\}$ is the orthonormal basis in H, composed of eigenvectors of Q) and $W_t = \sum_{i=1}^{\infty}\sqrt{\alpha_i}\,b_t^i e_i$, where b^i are independent real-valued Wiener processes.

Further,

$$\Lambda_t(W;H,G) = \Big\{\Psi\colon \mathscr{L}(H,G)\text{-valued processes, point-progressively measurable, such that } \|\Psi\|^2_{\Lambda_t} = \sum_{i=1}^{\infty}\alpha_i E\int_0^t |\Psi_s(e_i)|^2\,ds < +\infty\Big\},$$

and

$$\int_0^t \Psi_s\,dW_s = \sum_{i=1}^{\infty}\sqrt{\alpha_i}\int_0^t \Psi_s(e_i)\,db_s^i,$$

where this series is convergent in $L^2(\Omega)$. Let us remark that

2.18. If $\Psi\colon [0,T]\to\mathscr{L}(H,G)$ is a point-measurable (deterministic) function, then it is of course point-predictable.

2.19. If Ψ is point-predictable and $\sup\limits_{(t,\omega)\in[0,T]\times\Omega}\|\Psi(t,\omega)\| < +\infty$, then $\Psi\in\Lambda_T$ and $\|\Psi\|^2_{\Lambda_T}\leqslant \sup\limits_{t,\omega}\|\Psi(t,\omega)\|^2\cdot E|M_T|^2$.

2.20. Let (T_t) be a C_0-semigroup of operators on H. It follows from 2.4, 2.18, and 2.19 that $\int_0^t T_s\,dM_s$ is well defined. Moreover, the process $\left(\int_0^t T_{t-s}\,dM_s\right)_{t\geqslant 0}$ has a progressively measurable modification (see 3.13).

The proposition below, which follows from 2.5 and 2.7, is a stochastic analogue of the semigroup property 2.11:

Proposition 2.21. *Let* $\boldsymbol{D}$ *be the space defined in* 2.7 *and let* $\Psi\in\Lambda_t(M;H,\boldsymbol{D})$. *Then*

$$\int_0^t \Psi_s\,dM_s\in D(A),\qquad A\Psi\in\Lambda_t(M;H,H)\quad\textit{and}\quad A\int_0^t\Psi_s\,dM_s = \int_0^t A\Psi_s\,dM_s.$$

§ 3. Stochastic evolution equations with state-dependent noise

3.1. Stochastic equations with noise transformation. Four types of interpretation.

Assumption 3. Let the finite interval $[0,T]\subset\boldsymbol{R}$ be fixed.

Assumption 4. $(T_t)_{t\geqslant 0}$ is a C_0-semigroup of operators on H, with the infinitesimal generator A.

Assumption 5. Let ζ be an H-valued, $\mathscr{F}_0$-measurable random variable.

Definition 3.1. *Noise transformation* $\mathscr{M}(\cdot)$. Let $\mathfrak{B}$ be, as before, a Borel σ-field of subsets of $[0,T]$ and let

$$\mathscr{N} = \{H\text{-valued, progressively measurable w.r.t. } (\mathfrak{B}\times\mathscr{F}_t)_{t\geqslant 0}\text{ processes}\}.$$

By $\mathscr{M}(\cdot)$ we denote a mapping from some subclass $\mathscr{N}_0\subset\mathscr{N}$ into $\mathfrak{M}^2_{[0,T]}(H)$, having the following property: if $\forall t\in[0,T]$ $X_t = X_t'$ w.p. 1, then $\mathscr{M}(X') = \mathscr{M}(X)$.

Under Assumptions 1–5 we shall consider a stochastic system in H which can be formally written as (■). The following four types of interpretation of this system are possible:

Definition I. An H-valued process X is a *strong solution* of (■) iff

(I.1) $X\in\mathscr{N}_0$;

(I.2) w.p. 1 $\begin{cases} X_t(w)\in D(A) & \text{for a.a. } t,\\ X_\cdot(w)\in L^1([0,T];\boldsymbol{D}); \end{cases}$

(I.3) $\forall t\ \exists\Omega_t\ \ P(\Omega_t) = 1,\ \forall w\in\Omega_t\ \ X_t(w) = \zeta(w)+\int_0^t AX_s(w)\,ds+\mathscr{M}(X)_t(w)$.

Definition II. An H-valued process X is a *weakened solution* of (■) iff

(II.1) $X\in\mathscr{N}_0$;

(II.2) $\begin{cases} \text{w.p. 1 } X_\cdot(w)\in L^1([0,T];H),\\ \forall t \text{ w.p. 1 } \int_0^t X_s(w)\,ds\in D(A); \end{cases}$

(II.3) $\forall t\ \exists\Omega_t\ \ P(\Omega_t) = 1,\ \forall w\in\Omega_t\ \ X_t(w) = \zeta(w)+A\int_0^t X_s(w)\,ds+\mathscr{M}(X)_t(w)$.

Definition III. An H-valued process X is a *mild solution* of (■) iff

(III.1) $X\in\mathscr{N}_0$;

(III.2) w.p. 1 $X_\cdot(w)\in L^1([0,T];H)$;

(III.3) $\forall t\ \forall y\in D(A^*)\ \exists\Omega_{t,y}\ \ P(\Omega_{t,y}) = 1,\ \forall w\in\Omega_{t,y}\ \ \langle X_t(w),y\rangle = \langle\zeta(w),y\rangle + \int_0^t\langle X_s(w),A^*y\rangle\,ds+\langle\mathscr{M}(X)_t,y\rangle$.

Definition IV. An H-valued process X is a *mild integral solution* of (■) iff

(IV.1) $X\in\mathscr{N}_0$;

(IV.3) $\forall t\ \exists\Omega_t\ \ P(\Omega_t) = 1,\ \forall w\in\Omega_t\ \ X_t(w) = T_t\zeta(w)+\left(\int_0^t T_{t-s}\,d\mathscr{M}(X)_s\right)(w)$.

We begin the study of these solutions by formulating a few simple but useful propositions:

3.2. *Let X be a mild integral solution of* (■). *If* $E|\zeta|^2 < \infty$, *then* $\sup_{0\leq t\leq T} E|X_t|^2 < +\infty$.

Proof. The proof is obvious by 2.4 and 2.19.

3.3. *A strong solution of* (■) *is also a weakened solution* (*by* 2.7 *and* 2.11).

3.4. *Conversely, we infer only that*

If X is a weakened solution and satisfies condition (I.2), *then it is a strong solution* (*by* 2.11).

3.5. *It is obvious that a weakened solution is also a mild solution.*

3.6. *Let X be a solution of* (■) *of one of the above types and let X' be a progressively measurable modification of X. Then X' is also a solution of this type.*

The proof for strong solutions. (For the other types of solutions the proof is similar.) Since $X_t - X'_t = 0$ $\forall t$ w.p. 1, by the measurability of X and X', and the Fubini theorem we have

$$0 = \int_0^T E|X_t - X'_t|\,dt = E\int_0^T |X_t - X'_t|\,dt.$$

Then $X_t = X'_t$ w.p. 1 for a.a. t, which implies $X' \in L^1([0,T];\boldsymbol{D})$ and $AX_s = AX'_s$ w.p. 1 for a.a. s, and so $\int_0^t AX_s\,ds = \int_0^t AX'_s\,ds$ (w.p.1 $\forall t$). Finally, by Definition 3.1, $\mathcal{M}(X) = \mathcal{M}(X')$, which finishes the proof.

In connection with 3.6 we also introduce another definition:

DEFINITION 3.7. We say that a process X is an *exact solution* iff it is a solution and for condition (III.3) there exists a universal set $\bar{\Omega}$, $P(\bar{\Omega}) = 1$, independent of t (and of y in Definition III).

Consequently, an exact solution is a solution in the sense of trajectories.

PROPOSITION 3.8. *Let X be a strong solution of* (■). *Then there exists an exact strong solution X', and X' is a modification of X given by the right-hand side of the equality in* (I.3).

Proof. Take the solution X and consider the right-hand side in equality (I.3). By 3.6 it is sufficient to verify that it defines a progressively measurable process with well-defined trajectories.

Let $\bar{\Omega} = \{w\colon$ (I.2) holds$\}$; then $P(\bar{\Omega}) = 1$ and, for fixed $w \in \bar{\Omega}$,

$$X'_t(w) \overset{\text{def}}{=} \zeta(w) + \int_0^t AX_s(w)\,ds + \mathcal{M}(X)_t(w)$$

is sensible $\forall t$. By the measurability of A (Proposition 2.10) the process AX is progressively measurable, which implies that $\forall t \int_0^t AX_s\,ds$ (as a function of w) is $\mathcal{F}_t$-adapted (the Fubini theorem). Moreover, $\int_0^t AX_s(w)\,ds$ is continuous in t on $\bar{\Omega}$. Therefore, by the assumptions about ζ and $\mathcal{M}(X)$, X' is $\mathcal{F}_t$-adapted and right-continuous. So it is progressively measurable ([13], p. 70).

Remark 3.9. As a corollary we infer that an exact strong solution of (■) is right-continuous (continuous, if $\mathcal{M}(X)_t$ is a continuous martingale).

Remark 3.10 (about Definition III; see [16]). For condition (III.3) we can choose for any t a set Ω_t, $P(\Omega_t) = 1$, independent of y.

Proof. By 2.13 and 2.7 $(D(A^*), |\cdot|_{D(A^*)})$ is a separable Hilbert space. Denote by S a countable dense subset of this space. Fix t and define $\Omega_t = \bigcap_{y\in S} \Omega_{t,y}$. Then $P(\Omega_t) = 1$. It is easy to show that under fixed $w_0 \in \Omega_t$ equality (III.3) holds for any $y \in D(A^*)$: $\exists y_n \in S$ $y_n \to y$ and $A^*y_n \to A^*y$ (from the definition of the graph norm $|\cdot|_{D(A^*)}$). So $\sup_n |Ay_n| \leq$ const and $|\langle X_s(w_0), A^*y_n\rangle| \leq |X_s(w_0)|\cdot$ const. Then we can use the Lebesgue convergence theorem, take the limit and replace y_n by y in equality (III.3) at the point w_0.

3.2. Main lemma. The following statement will be important for us and will replace, in a certain sense, Ito's lemma. For the integrals w.r.t. the H-valued Wiener process this fact has been formulated in [5] (Definition 2.18 of the stochastic double integral) by analogy with the scalar case (see, for instance, [9], p. 217).

LEMMA 3.11 (Generalized Fubini theorem). *Let $M \in \mathfrak{M}^2_{[0,T]}(H)$ and $\Psi\colon [0,T]\times\times([0,T]\times\Omega) \to \mathcal{L}(H,G)$ be point-measurable w.r.t. $\mathfrak{B}\times\mathcal{P}$ satisfying the condition*

$$\text{(v)} \qquad \sup_{\substack{u,s\in[0,T]\\ w\in\Omega}} \|\Psi(u,s,w)\| \leq K < +\infty.$$

Define

$$y_t^1(\Psi)(w) \overset{\text{or}}{=} y_t^1(w) = \int_0^t \Big(\int_0^t \Psi(u,s)\,dM_s\Big)\,du$$

and

$$y_t^2(\Psi)(w) \overset{\text{or}}{=} y_t^2(w) = \int_0^t \Big(\int_0^t \Psi(u,s)\,du\Big)\,dM_s$$

(*where* $\int_0^t \Psi(u,s)\,du$ *must be understood as an integral in the sense of Definition* 2.3).

Then $\forall t \in [0,T]$ y_t^1 *and* y_t^2 *are well-defined, $\mathcal{F}_t$-measurable random variables and*

$$\forall t\in[0,T] \quad y_t^1 = y_t^2 \text{ w.p. 1}.$$

Proof. We proceed similarly to Metivier and Pistone [12] in the construction of the stochastic integral and prove the lemma: first for step functions Ψ, secondly for certain strongly measurable functions Ψ, and finally for point-measurable bounded Ψ. The estimations (1) and (3) below make possible the limit passage.

Fix $t \in [0, T]$ and let Q and λ be as in Theorem 2.14. By l we denote the Lebesgue measure. Notice that, by Definition 2.3, $\int_0^t \Psi(u, s, w)\,du$ is point-predictable and by condition (v) it belongs to Λ_t. So y_t^2 is G-valued, $\mathscr{F}_t$-measurable random variable. We have the following estimation for the square mean of y_t^2:

$$E\left|\int_0^t\left(\int_0^t \Psi(u, s)\,du\right)dM_s\right|^2 = \int_{\Omega\times[0,t]} \left\|\left(\int_0^t \Psi(u, s)\,du\right)\cdot Q^{1/2}(s, w)\right\|_{\mathrm{HS}}^2 d\lambda$$

$$\leqslant \int_{\Omega\times[0,t]} \left(\int_0^t \|\Psi(u, s)\cdot Q^{1/2}(s, w)\|_{\mathrm{HS}}du\right)^2 d\lambda \qquad \text{(by the Hölder inequality)}$$

$$\leqslant \int_{\Omega\times[0,t]} t\left(\int_0^t \|\Psi(u, s)\cdot Q^{1/2}(s, w)\|_{\mathrm{HS}}^2du\right) d\lambda.$$

So

$$E|y_t^2(\Psi)|^2 \leqslant t\int_0^t \|\Psi(u)\|_{\Lambda_t}^2\,du. \tag{1}$$

Consider y_t^1; the integral $\int_0^t \Psi(u, s)\,dM_s$ has sense. Suppose that it has a modification measurable w.r.t. $\mathfrak{B}\times\mathscr{F}_t$. Then we shall have the equalities

$$E\int_0^t\left|\int_0^t \Psi(u, s)\,dM_s\right|^2 du = \int_0^t E\left|\int_0^t \Psi(u, s)\,dM_s\right|^2 du = \int_0^t \|\Psi(u)\|_{\Lambda_t}^2\,du, \tag{2}$$

and—by the Hölder inequality—the following estimation:

$$E|y_t^1(\Psi)|^2 \leqslant t\int_0^t \|\Psi(u)\|_{\Lambda_t}^2\,du. \tag{3}$$

Step 1. Let φ be of the form

$$\varphi(u, s, w) = \sum_{i=1}^N \varphi_i(u)\chi_{]s_i, s_i']\times F_i}(s, w), \tag{4}$$

where $\forall i$ $]s_i, s_i']\times F_i \in \mathscr{R}$ and φ_i: $[0, T]\to\mathscr{L}(H, G)$ is strongly measurable such that $\sup_{u\in[0,T]}\|\varphi_i(u)\| < \infty$. Then

$$\int_0^t \varphi(u, s, w)\,dM_s = \sum_{i=1}^N \varphi_i(u)\chi_{F_i}(w)\left(M_{s_i'\wedge t}(w) - M_{s_i\wedge t}(w)\right), \tag{5}$$

and this form implies measurability w.r.t. $\mathfrak{B}\times\mathscr{F}_t$. Therefore in this case $y_t^1(\varphi)$ is well-defined and it is obvious that $y_t^1(\varphi) = y_t^2(\varphi)$.

Step 2. Let $\{h_i\}_{i=1}^\infty$, $\{g_i\}_{i=1}^\infty$ be orthonormal bases in H and G, respectively.

Define $H_n = \mathrm{Lin}\{h_1, \ldots, h_n\}$; $G_n = \mathrm{Lin}\{g_1, \ldots, g_n\}$ and let Π_n $(\tilde{\Pi}_n)$ be the orthogonal projection $H \overset{\text{onto}}{\to} H_n$ $(G \overset{\text{onto}}{\to} G_n)$.

We define $\Psi_n = \tilde{\Pi}_n\cdot\Psi\cdot\Pi_n$: $H\to G$, which means

$$\Psi_n(u, s, w)(h) = \sum_{i,j=1}^n \langle h, h_i\rangle\langle\Psi(u, s, w)(h_i), g_j\rangle g_j.$$

It follows from this form that Ψ_n: $([0, T]\times([0, T]\times\Omega), \mathfrak{B}\times\mathscr{P}) \to \mathscr{L}(H, G)$ is strongly measurable. Then there exists a sequence $(\Psi_{n,k})_{k=1}^\infty$ such that

(1) $\Psi_{n,k}$: $[0, T]\times([0, T]\times\Omega)\to\mathscr{L}(H, G)$ and it is of the form (4);

(2) $\sup_{u,s,w}\|\Psi_{nk}(u, s, w)\| \leqslant 2K$;

(3) $\|\Psi_{nk}(u, s, w) - \Psi_n(u, s, w)\| \underset{k\to\infty}{\to} 0$.

The estimation

$$\|(\Psi_{nk}-\Psi_n)\cdot Q^{1/2}\|_{\mathrm{HS}}^2 \leqslant \|\Psi_{nk}-\Psi_n\|^2\cdot\|Q^{1/2}\|_{\mathrm{HS}}^2 = \|\Psi_{nk}-\Psi_n\|^2\cdot\mathrm{Tr}\,Q = \|\Psi_{nk}-\Psi_n\|^2$$

implies the convergence $\|(\Psi_{nk}-\Psi_n)\cdot Q^{1/2}\|_{\mathrm{HS}}^2 \underset{k\to\infty}{\to} 0$. Therefore, using the boundedness assumptions and the Lebesgue convergence theorem, we have

$$\int_0^t \|\Psi_{nk}-\Psi_n\|_{\Lambda_t}^2\,du = \int_0^t\left(\int_{[0,t]\times\Omega} \|(\Psi_{nk}-\Psi_n)\cdot Q^{1/2}\|_{\mathrm{HS}}^2\,d\lambda\right)du \underset{k\to\infty}{\to} 0.$$

So, it follows from Step 1 and (2) that $y_t^1(\Psi_n)$ is well defined, and from estimations (1) and (3) we have $y_t^1(\Psi_n) = y_t^2(\Psi_n)$.

Step 3. If we show that $\int_0^t \|\Psi_n-\Psi\|_{\Lambda_t}^2 \underset{n\to\infty}{\to} 0$, the proof will be finished. It is clear, by the construction of Ψ_n, that

$$\forall\, h\in H\ \ |\Psi_n(h)-\Psi(h)| \underset{n\to\infty}{\to} 0.$$

Hence

$$|(\Psi_n-\Psi)\cdot Q^{1/2}(h_i)| \underset{n\to\infty}{\to} 0\ \ \forall i.$$

Therefore, using the estimation

$$|(\Psi_n-\Psi)\cdot Q^{1/2}(h_i)|^2 \leqslant 9K^2|Q^{1/2}(h_i)|^2$$

and the summability of the series $\sum_{i=1}^\infty |Q^{1/2}(s, w)(h_i)|^2 = \mathrm{Tr}\,Q = 1$ (λ a.e.) we have

$$\|(\Psi_n-\Psi)\cdot Q^{1/2}\|_{\mathrm{HS}}^2 = \sum_{i=1}^\infty |(\Psi_n-\Psi)\cdot Q^{1/2}(h_i)|^2 \underset{n\to\infty}{\to} 0\ \ (\forall\, u\ \lambda \text{ a.e.}),$$

and, as in Step 2, we finally obtain

$$\int_0^t \|\Psi_n-\Psi\|_{\Lambda_t}^2\,du \underset{n\to\infty}{\to} 0.$$

Remark 3.12. It is easy to show that this lemma is also true if condition (v) is replaced by the weaker condition

$$\text{(v')} \qquad \int_0^t \|\Psi(u,\cdot,\cdot)\|^2_{\Lambda_t} du < +\infty.$$

PROPOSITION 3.13. *Let Ψ and M be as in Lemma* 3.11. *Then the process* $Z_u(\Psi) = \int_0^u \Psi(u,s,\omega)dM_s$ *has a progressively measurable modification.*

Proof. It follows from the proof of Lemma 3.11 that there exists a sequence (Ψ_n) such that $\forall n$ Ψ_n is of the form (4), and so $Z(\Psi_n)$ is progressively measurable, see (5), and $\forall u$ $\|\Psi_n(u)-\Psi(u)\|_{\Lambda_u} \xrightarrow[n\to\infty]{} 0$.

By Proposition 2.16 (b): $E|Z_u(\Psi_n)-Z_u(\Psi)|^2 = \|\Psi_n(u)-\Psi(u)\|_{\Lambda_u}$. Therefore $\forall u\, \exists (n_k)$ $Z_u(\Psi_{n_k}) \to Z_u(\Psi)$ w.p. 1, and if we define

$$Z'_u(\Psi)(w) = \begin{cases} \lim Z_u(\Psi_{n_k})(w) & \text{if this limit exists,}\\ 0 & \text{otherwise,}\end{cases}$$

then $Z'(\Psi)$ is a progressively measurable modification of $Z(\Psi)$.

3.3. Main theorems. Now we can formulate the main result for the general model (■):

THEOREM 3.14. *Let X be an H-valued process on* $[0, T]$. *The following conditions are equivalent*:

(II) *X is a weakened solution of* (■),

(III) *X is a mild solution of* (■),

(IV) *X is a mild integral solution of* (■).

Proof. First, using the main Lemma 3.11, we shall prove two technical lemmas.

LEMMA 1. *Let* $M \in \mathfrak{M}^2_{[0,T]}(H)$ *and introduce the notation* $Y_s \overset{\text{def}}{=} \int_0^s T_{s-u}dM_u$. *Then* $\forall t \in [0,T]$ w.p. 1

$$\int_0^t Y_s ds \in D(A) \quad \textit{and} \quad A\int_0^t Y_s ds = \int_0^t T_{t-u}dM_u - M_t.$$

Proof of lemma. By 2.4 and the continuity of the semigroup the function $\Psi(s,u) \overset{\text{def}}{=} \chi_{[0,s]}(u)T_{s-u}$ satisfies all the assumptions of Lemma 3.11. Thus, in particular, w.p. 1 $Y_\cdot(w) \in L^1([0,T];H)$. From Lemma 3.11 we obtain for fixed t

$$\int_0^t Y_s ds = \int_0^t\Big(\int_u^t T_{s-u}ds\Big)dM_u \text{ w.p. } 1,$$

and the following operator-valued function:

$$\Phi(u) \overset{\text{def}}{=} \int_u^t T_{s-u}ds \quad \Big(\equiv \int_0^t \Psi(s,u)ds\Big)$$

is point-measurable in u. For fixed u we have by 2.12

$$\forall h \in H \quad \Phi(u)(h) \in D(A) \quad \text{and} \quad A\Phi_u(h) = T_{t-u}h - h.$$

Therefore Φ: $[0, T] \to \mathscr{L}(H, D)$ is point-measurable (by 2.10) and (see 2.7 and 2.4) $\forall u \in [0, T]$ $\|\Phi_u\|_{\mathscr{L}(H,D)} < +\infty$. Hence $\Phi \in \Lambda_t(M; H, D)$. This means that (see 2.21)

$$\forall t \int_0^t \Phi_u dM_u \in D(A) \text{ w.p. } 1,$$

so

$$\int_0^t Y_s ds \in D(A) \quad \text{and} \quad A\int_0^t \Phi_u dM_u = \int_0^t A\Phi_u dM_u$$

because A is a bounded operator from D into H. Therefore,

$$A\int_0^t Y_s ds = \int_0^t A\Phi_u dM_u = \int_0^t (T_{t-u}-I)dM_u = \int_0^t T_{t-u}dM_u - M_t.$$

The second basic lemma is a kind of integration-by-parts formula, which can be written formally as

$$d(T_{t-s}M_s) = T_{t-s}dM_s + (dT_{t-s})M_s.$$

A similar formula has been proved in [12] by Ito's lemma under the stronger assumptions on M. The basic tool in our proof is the Main Lemma.

LEMMA 2 (Integration-by-parts). *Let* $M \in \mathfrak{M}^2_{[0,T]}(H)$. *If* $0 \leqslant s \leqslant t \leqslant T$, *then* w.p. 1

$$\int_0^s T_{t-u}M_u du \in D(A)$$

and

$$T_{t-s}M_s = \int_0^s T_{t-u}dM_u - A\int_0^s T_{t-u}M_u du.$$

Proof of lemma. Fix s and t. We can write $M_u = \int_0^u I dM_w$. Then by the Main Lemma we have the identity

$$\int_0^s T_{t-u}M_u du = \int_0^s\Big(\int_0^u T_{t-u}dM_w\Big)du = \int_0^s\Big(\int_w^s T_{t-u}du\Big)dM_w; \tag{1}$$

all the expressions in (1) have a well-defined meaning. Using the same method as that used in Lemma 1, we see that the last expression belongs to $D(A)$ w.p. 1 and

$$\begin{aligned} A\int_0^s\Big(\int_w^s T_{t-u}du\Big)dM_w &= \int_0^s A\Big(\int_w^s T_{t-u}du\Big)dM_w \\ &= \int_0^s (T_{t-w}-T_{t-s})dM_w = \int_0^s T_{t-w}dM_w - T_{t-s}M_s. \end{aligned} \tag{2}$$

Then by (1) and (2) we obtain

$$T_{t-s}M_s = \int_0^s T_{t-u}dM_u - A\int_0^s T_{t-u}M_u du.$$

Proof of Theorem. (IV) $\Rightarrow$ (II). Let X be a solution of (IV)-type; then $\mathcal{M}(X) \in \mathfrak{M}^2_{[0,T]}(H)$ and $\forall\, t$ w.p. 1

$$X_t(w) = T_t\zeta(w) + \int_0^t T_{t-s}d\mathcal{M}(X)_s. \tag{3}$$

We must check that X satisfies conditions (II.2) and (II.3). Consider the first term of the right-hand side in equality (3). Fix w; then

$$|T_t\zeta(w)| \leqslant \sup_{t\in[0,T]} \|T_t\| \cdot |\zeta(w)| < +\infty \tag{4}$$

and by 2.12 we have for every t

$$\int_0^t T_s\zeta(w)\,ds \in D(A), \tag{5}$$

$$A\int_0^t T_s\zeta(w)\,ds = T_t(w) - \zeta(w). \tag{6}$$

For the second term we use Lemma 1 and obtain

$$\int_0^T \Big|\int_0^s T_{s-u}d\mathcal{M}(X)_u\Big|\,ds < +\infty \text{ w.p. } 1, \tag{7}$$

$$\forall\, t \text{ w.p. } 1 \int_0^t \Big(\int_0^s T_{s-u}d\mathcal{M}(X)_u\Big)ds \in D(A), \tag{8}$$

$$\forall\, t \text{ w.p. } 1\ A\int_0^t \Big(\int_0^s T_{s-u}d\mathcal{M}(X)_u\Big)ds = \int_0^t T_{t-u}d\mathcal{M}(X)_u - \mathcal{M}(X)_t. \tag{9}$$

Consequently it follows from (4) and (7) that

$$X_{\cdot}(w) \in L^1([0,T];H) \text{ w.p. } 1.$$

Analogously (5) and (8) imply

$$\forall t \text{ w.p. } 1 \int_0^t X_s(w)\,ds \in D(A).$$

Thus (II.2) follows.

We finally infer from identities (6) and (9) that (II.3) holds;

$$A\int_0^t X_s ds \overset{(6)}{=}{}^{(9)} T_t\zeta - \zeta + \int_0^t T_{t-u}d\mathcal{M}(X)_u - \mathcal{M}(X)_t = X_t - \zeta - \mathcal{M}(X)_t.$$

Therefore a mild integral solution of (■) is also a weakened solution.

It was remarked earlier (see 3.5) that the weakened solution is also a mild solution.

It thus remains to prove the implication (III) $\Rightarrow$ (IV). Let X be a mild solution and fix t. According to Remark 3.10 we can choose a subset Ω_t, $P(\Omega_t) = 1$, such that equality (III.3) holds for any $w \in \Omega_t$ and any $y \in D(A^*)$.

On Ω_t we have

$$\langle \mathcal{M}(X)_t, y\rangle = \langle X_t, y\rangle - \langle \zeta, y\rangle - \int_0^t \langle X_s, A^*y\rangle ds. \tag{10}$$

If we take $s = t$ in the formulation of Lemma 2, then

$$\mathcal{M}_t(X) = \int_0^t T_{t-u}d\mathcal{M}(X)_u - A\int_0^t T_{t-u}\mathcal{M}(X)_u du$$

and therefore, for $y \in D(A^*)$,

$$\langle \mathcal{M}(X)_t, y\rangle = \Big\langle \int_0^t T_{t-u}d\mathcal{M}(X)_u, y\Big\rangle - \int_0^t \langle T_{t-u}\mathcal{M}(X)_u, A^*y\rangle du. \tag{11}$$

Take $y \in D(A^{*2})$; then we have from (10)

$$\begin{aligned}\langle T_{t-u}\mathcal{M}(X)_u, A^*y\rangle &= \langle \mathcal{M}(X)_u, T^*_{t-u}A^*y\rangle \\ &= \langle X_u, T^*_{t-u}A^*y\rangle - \langle \zeta, T^*_{t-u}A^*y\rangle - \int_0^u \langle X_s, A^*T^*_{t-u}A^*y\rangle ds.\end{aligned} \tag{12}$$

All the terms (12) are integrable in u (on the interval $[0, t]$) with probability 1 because for a.a. w $X_{\cdot}(w) \in L^1([0,T];H)$ and $\sup_{0\leqslant t\leqslant T} \|T^*_t\| < \infty$. Let us integrate both sides of this equality:

$$\begin{aligned}&\int_0^t \langle T_{t-u}\mathcal{M}(X)_u, A^*y\rangle du \\ &= \int_0^t \langle X_u, T^*_{t-u}A^*y\rangle du - \Big\langle \zeta, \int_0^t T^*_{t-u}A^*y\,du\Big\rangle - \int_0^t\int_0^u \langle X_s, A^*T^*_{t-u}A^*y\rangle ds\,du.\end{aligned} \tag{13}$$

The expression $\langle X_s, A^*T^*_{t-u}A^*y\rangle$ is measurable in (s, u) as a superposition of measurable functions; moreover, we have the estimation

$$|\langle X_s, A^*T^*_{t-u}A^*y\rangle| \leqslant |X_s| \cdot |A^{*2}y| \cdot \sup_{t\in[0,T]} \|T^*_t\| \leqslant \text{const} \cdot |X_s|.$$

So it is integrable in (s, u) and we can use the usual Fubini theorem. Therefore, if we denote by a the last term in (13), we obtain

$$-a = \int_0^t \Big(\int_s^t \langle X_s, A^*T^*_{t-u}A^*y\rangle du\Big)ds = \int_0^t \Big\langle X_s, \int_s^t A^*T^*_{t-u}A^*y\,du\Big\rangle ds.$$

From the semigroup properties 2.13, 2.8 and 2.9 we have

$$\int_s^t A^*T^*_{t-u}A^*y\,du = A^*\int_s^t T^*_{t-u}A^*y\,du = A^*(T^*_{t-s}y-y);$$

then $a = -\int_0^t \langle X_s, A^*T^*_{t-s}y\rangle ds + \int_0^t \langle X_s, A^*y\rangle ds$ and from (13) we obtain

$$\int_0^t \langle T_{t-u}\mathcal{M}(X)_u, A^*y\rangle du = -\Big\langle \zeta, \int_0^t T^*_{t-u}A^*y\,du\Big\rangle + \int_0^t \langle X_s, A^*y\rangle ds. \tag{14}$$

By (10), (11) and (14), for any $y \in D(A^{*2})$,

$$\langle X_t, y\rangle - \langle \zeta, y\rangle - \int_0^t \langle X_s, A^*y\rangle ds$$

$$= \Big\langle \int_0^t T_{t-u}d\mathcal{M}(X)_u, y\Big\rangle + \Big\langle \zeta, \int_0^t T^*_{t-u}A^*y\,du\Big\rangle - \int_0^t \langle X_s, A^*y\rangle ds \quad \text{on} \quad \Omega_t.$$

Since

$$\int_0^t T^*_{t-u}A^*y\,du = T^*_t y - y,$$

we finally obtain

$$\langle X_t, y\rangle = \langle T_t\zeta, y\rangle + \Big\langle \int_0^t T_{t-u}d\mathcal{M}(X)_u, y\Big\rangle.$$

$D(A^{*2})$ is dense in H; moreover, $T_t\zeta$ and $\int_0^t T_{t-u}d\mathcal{M}(X)_u$ are well-defined elements of H. Therefore $X_t = T_t\zeta + \int_0^t T_{t-u}d\mathcal{M}(X)_u$ on Ω_t.

From 3.3 we immediately have

COROLLARY 3.15. *If X is a solution of the strong equation, then X is a solution of the mild integral equation. Consequently, we have the following scheme of implications*:

$$\text{(I)} \Rightarrow \text{(II)} \Leftrightarrow \text{(III)} \Leftrightarrow \text{(IV)}.$$

Remark 3.15. The opposite implication to the first one in the previous scheme may be false, even if the initial condition ζ belongs to the domain of A. An appropriate example is given in [4].

In the spirit of the remark made in the introduction the second main theorem will be an existence and uniqueness result for certain classes of mild integral equations (IV). These equations describe the evolution of systems disturbed by state (or past) dependent white noise, which are important for applications. The Gaussian white noise may be obtained in our framework as a particular case of the transformation $\mathcal{M}(\cdot)$ (see Remark 3.20).

DEFINITION 3.16. Define

$$N^2_H = \mathcal{N} \cap L^2([0, T]\times\Omega, l\times P, H).$$

Therefore N^2_H, as a closed subspace of $L^2([0, T]\times\Omega, H)$, is a Hilbert space with the L^2-norm (denoted further on by $|\cdot|_2$).

Let W be a G-valued Wiener process adapted to $(\mathcal{F}_t)$ with Q as its covariance operator. As in Example 2.17, $\{\alpha_i\}$ are eigenvalues of Q and $\{e_i\}$ is the orthonormal basis in G composed of eigenvectors of Q.

THEOREM 3.17. *Let B: $H\times[0, T] \to \mathcal{L}(G, H)$ satisfy the following conditions*:

(a) $\forall g \in G$ $B(\cdot, \cdot)(g)$: $H\times[0, T] \to H$ *is measurable in* (h, t);

(b$_1$) $\exists k > 0$ $\forall t \in [0, T]$ $\forall x \in H$ $\|B(x, t)\|^2 \leqslant k(1+|x|^2)$;

(b$_2$) $\exists k > 0$ $\forall t \in [0, T]$ $\forall x, y \in H$ $\|B(x, t)-B(y, t)\| \leqslant k|x-y|$.

Moreover, the H-valued $\mathcal{F}_0$- measurable random variable is assumed to be such that

(c) $E|\zeta|^2 < +\infty$,

Then the equation

$$X_t = T_t\zeta + \int_0^t T_{t-s}B(X_s, s)dW_s \tag{$*$}$$

has a unique solution in the space N^2_H. Moreover, $\sup_{0\leqslant t\leqslant T} E|X_t|^2 < +\infty$ and X is continuous in the mean square sense.

Proof. The proof is similar to that in finite-dimensional spaces, but we give it for completeness and in a less standard manner that usual (compare [7], [11]).

Let $K \overset{\text{def}}{=} \sup_{0\leqslant t\leqslant T} \|T_t\|$. We define the following operator U on N^2_H:

$$U(Y)(t) = T_t\zeta + \int_0^t T_{t-s}B(Y_s, s)dW_s.$$

This form implies that the process $U(Y)$ is progressively measurable (see 2.20). Consider $(|U(Y)|_2)^2$:

$$\int_0^T E|U(Y)(t)|^2dt \leqslant 2\int_0^T E|T_t\zeta|^2dt + 2\int_0^T E\Big|\int_0^t T_{t-s}B(Y_s, s)dW_s\Big|^2 dt. \tag{1}$$

By Example 2.17 we have the equality

$$E\Big|\int_0^t T_{t-s}B(Y_s, s)dW_s\Big|^2 = \sum \alpha_i E\int_0^t |T_{t-s}B(Y_s, s)(e_i)|^2 ds.$$

Therefore from assumption (b$_1$) we obtain the estimation

$$(|U(Y)|_2)^2 \leqslant 2T\cdot K^2\cdot E|\zeta|^2 + 2T\cdot K^2\cdot k\cdot \mathrm{Tr}QE\int_0^T (1+|Y_s|^2)ds < +\infty. \tag{2}$$

Hence U: $N^2_H \to N^2_H$.

We shall show that U is a contraction operator in some norm on N_H^2 equivalent to the standard norm. Define

$$|||\varphi||| = \sup_{0\leqslant t\leqslant T}\Big\{(\exp(-at))\cdot\Big(\int_0^t E|\varphi_s|^2 ds\Big)^{1/2}\Big\}$$

(where a is some constant which we shall define later). It is obvious that $|||\cdot||| \sim |\cdot|_2$.

We have the following inequalities for $X, Y\in N_H^2$:

$$E\int_0^t |U(Y)(s)-U(X)(s)|^2 ds = \int_0^t E\Big|\int_0^s T_{s-u}(B(Y_u, u)-B(X_u, u))dW_u\Big|^2 ds$$

$$\leqslant \operatorname{Tr}Q\cdot K^2\cdot k^2\int_0^t\Big(e^{2as}\cdot\Big(e^{-2as}\int_0^s E|Y_u-X_u|^2 du\Big)\Big)ds$$

$$\leqslant \operatorname{Tr}Q\cdot K^2\cdot k^2\cdot\int_0^t e^{2as}|||X-Y|||^2 ds$$

$$\leqslant \frac{\operatorname{Tr}Q\cdot K^2\cdot k^2\cdot e^{2at}}{2a}|||X-Y|||^2.$$

So

$$|||U(Y)-U(X)|||\leqslant \frac{K\cdot k\cdot\sqrt{\operatorname{Tr}Q}}{\sqrt{2a}}|||X-Y|||,$$

and if $a > \dfrac{K^2\cdot k^2\cdot\operatorname{Tr}Q}{2}$, U is a contraction operator in the norm $|||\cdot|||$. Hence in N_H^2 there exists one and only one solution X of equation $(*)$ (X is the fixed point of the operator U).

By the same estimations as (1) and (2) (with $U(X) = X$) we obtain

$$\sup_{0\leqslant t\leqslant T} E|X_t|^2 < +\infty.$$

Finally, for $\Psi\in\Lambda_t(W; G, H)$ and $t > u$; we have the estimation

$$E\Big|\int_0^t T_{t-s}\Psi_s dW_s-\int_0^u T_{u-s}\Psi_s dW_s\Big|^2$$

$$\leqslant 2E\Big|\int_0^u (T_{t-u}-I)T_{u-s}\Psi_s dW_s\Big|^2+2E\Big|\int_u^t T_{t-s}\Psi_s dW_s\Big|^2$$

$$\leqslant 2K^2\sum_{i=1}^{\infty}\alpha_i E\int_0^T |(T_{t-u}-I)\Psi_s(e_i)|^2 ds+2\int_u^t E||T_{t-s}\Psi_s Q^{1/2}||_{\mathrm{HS}}^2.$$

If $(t-u)\to 0$, then it is obvious that the last term converges to 0. The first term converges also to 0—by the continuity of the semigroup and the Lebesgue convergence theorem. Therefore X is continuous in the mean square.

Remark 3.18. As in the real case we can make assumption (b_2) weaker; namely,

(b_2') $\forall n\,\exists k_n\,\forall t\,\forall x, y\in H$ $(|x|\leqslant n$ and $|y|\leqslant n)$

$$\Rightarrow ||B(x, t)-B(y, t)||\leqslant k_n\cdot|x-y|.$$

Then (a), (b_2') are sufficient for the uniqueness, provided any solution of $(*)$ has a modification with trajectories in $L^\infty([0, T], H)$. (For the idea of the proof see [7].)

Theorem 3.17 is a particular case of the theorem below, including the equations with the noise dependent on the whole past. We have formulated Theorem 3.17 separately, because it is more convenient for applications.

THEOREM 3.19. *Assume that* B: $N_H^2\times[0, T]\to\mathscr{L}(G, H)$ *is such that*

(Fa) $\forall t$ $B(X, t) = B(X^t, t)$, *where* X^t *means the process* X *stopped at* t *and* $\forall g\in G$ $B(\cdot,\cdot)(g)$: $N_H^2\times[0, T]\to H$ *is measurable in two variables*;

(Fb_1) $\exists k\geqslant 0\,\forall X\in N_H^2\,\forall i$ $\quad E\int_0^T |B(X, s)(e_i)|^2 ds\leqslant k\cdot\big(E\int_0^T |X_s|^2 ds+1\big)$;

(Fb_2) $\exists k\geqslant 0\,\forall t\in[0, T]\,\forall X, Y\in N_H^2\,\forall i$ $\quad E\int_0^t |B(X, s)(e_i)-B(Y, s)(e_i)|^2 ds$

$\leqslant k\cdot E\int_0^t |X_s-Y_s|^2 ds$.

Let assumption (c) *also be satisfied. Then the equation*

$$(**)\qquad X_t = T_t\zeta+\int_0^t T_{t-s}B(X^s, s)dW_s$$

has a unique solution in the space N_H^2. *Moreover*, X *is continuous in the mean square sense and* $\sup_{0\leqslant t\leqslant T} E|X_t|^2 < +\infty$.

Proof. Note that the proof of Theorem 3.17 applies here.

Remark 3.20. The assumptions (a), (b_1) ((Fa), (Fb_1)) imply that $\mathscr{M}(\cdot)$, defined as $\mathscr{M}(X)_t = \int_0^t B(X, s)dW_s$, maps $N_H^2 = \mathscr{N}_0$ into $\mathfrak{M}_{[0,T]}^2(H)$. Therefore equation $(**)$ is of (IV)-type.

It follows from assumption (c)—by Proposition 3.2—that the unique solution of $(**)$ in N_H^2 is also a unique (up to modification) solution in the sense of Definition IV.

3.4. Applications of the main theorems. The general model $(\blacksquare)$ includes two important cases:

Stochastic evolution equations with state independent martingale noise (when $\mathscr{M}(\cdot)$: $\mathscr{M}(X)\equiv M\in\mathfrak{M}_{[0,T]}^2$);

Stochastic evolution equations with state dependent Gaussian white noise (when $\mathscr{M}(\cdot)$ is such as in Remark 3.20).

3.4.1. *An application to stochastic evolution equations with constant noise transformation*. Suppose that in Definition 3.1 $\mathscr{N}_0 = \mathscr{N}$ and $\forall X\in\mathscr{N}$ $\mathscr{M}(X) = M$, $M\in\mathfrak{M}_{[0,T]}^2(H)$. Then the model $(\blacksquare)$ has the form

$$(\tilde{\blacksquare})\qquad \begin{aligned} dX_t &= AX_t\,dt+dM_t,\\ X_0 &= \zeta.\end{aligned}$$

DEFINITION 3.21. Let

$$N_H^1 = \{H\text{-valued progressively measurable processes with trajectories belonging to } L^1([0, T]; H) \text{ w.p. } 1\}.$$

As immediate corollaries of § 3.1 and Theorem 3.14 we obtain the following existence and uniqueness results:

THEOREM 3.22. *The progressively measurable process* $X_t = T_t\zeta + \int_0^t T_{t-s}dM_s$ *is the unique (up to modification) in the class* $\mathcal{N}_H^1$:

(a) *weakened solution of* $(\tilde{\blacksquare})$;

(b) *mild solution of* $(\tilde{\blacksquare})$ *(compare with* [16]);

(c) *If there exists a strong solution of* $(\tilde{\blacksquare})$, *then it is a modification of the process* $X_t = T_t\zeta + \int_0^t T_{t-s}dM_s$.

THEOREM 3.23 (proved earlier in [12] and for the Gaussian noise case in [5]). *Let* D *be the space defined in* 2.7. *If* $M \in \mathfrak{M}^2_{[0,T]}(D)$ *and* ζ *is* D*-valued random variable, then the strong equation corresponding to* $(\tilde{\blacksquare})$ *has a unique (in the sense of trajectories) exact solution and this solution is a right-continuous version of the process* $X_t = T_t\zeta + \int_0^t T_{t-s}dM_s$.

3.4.2. *An application to evolution equations with state-dependent Gaussian noise.* Consider the stochastic differential system

$$(\tilde{\tilde{\blacksquare}}) \qquad \begin{aligned} dX_t &= AX_t dt + B(X, t)dW_t, \\ X_0 &= \zeta. \end{aligned}$$

The following theorem is an immediate consequence of § 3.3:

THEOREM 3.24. *Under the assumptions of Theorem* 3.19 (*of Theorem* 3.17) *there exists a unique weakened (as well as mild) solution of* $(\tilde{\tilde{\blacksquare}})$, *given by the solution of the mild integral equation* $(**)$.

Let us turn our attention to the strong equation. As a corollary of the previous theorem, Remark 3.18 and Proposition 3.3 we get

THEOREM 3.25 (*Uniqueness*). *Under the assumptions* (a) *and* (b_2') *there may exist only one strong solution of system* $(\tilde{\tilde{\blacksquare}})$.

Making use of § 3.1 and § 3.3, we can also obtain some existence results for such strong equations. The first one is quite general but not convenient:

PROPOSITION 3.26 (*Existence*). *Let* X *be the solution of equation* $(**)$. *Moreover, suppose that*

(d) $\zeta \in D(A)$ w.p. 1;

(e) $T_t B(X, s)(g) \in D(A)$ w.p. 1 $\forall g \in G$ $\forall s \in [0, T]$, $t \in]0, T]$;

(f) $\sup\limits_{0 \leqslant t \leqslant T} E\int_0^t \|AT_{t-s}(X, s)Q^{1/2}\|^2_{\mathrm{HS}} ds < +\infty$.

Then X *is a solution of the equation*

$$(*{*}*) \qquad X_t = \zeta + \int_0^t AX_s ds + \int_0^t B(X, s)dW_s.$$

Proof. The proof is a consequence of Proposition 3.4 and Theorem 3.14.

Remark 3.27. Let B: $H \times [0, T] \to \mathscr{L}(H, G)$. If one of the following conditions is satisfied:

(1) $\forall x \in H, g \in G$ $B(x, t)(g) \in D(A)$ w.p. 1 $\forall t \in [0, T]$;

(2) $T_t(H) \subset D(A)$ for $t \in]0, T]$,

then assumption (e) is fulfilled.

THEOREM 3.28 (*Existence and uniqueness*). *Suppose that* B: $H \times [0, T] \to \mathscr{L}(G, H)$, *satisfying condition* (a), *is such that*

$$\forall t \in [0, T] \forall g \in G \ B(D(A), t)(g) \subset D(A)$$

(i.e. B: $D \times [0, T] \to \mathscr{L}(G, D)$). *Moreover,* B *fulfils* (b_1), (b_2) *w.r.t. the norms* $\|\cdot\|_{L(G,D)}$, $|\cdot|_D$, ζ *is supposed to satisfy* (d) *and* (c) *w.r.t.* $|\cdot|_D$.

Then there exists a unique, in the sense of trajectories, exact solution of equation $(*{*}*)$. *This solution is continuous* w.p. 1.

Proof. By Theorem 3.17, equation $(*)$ has a unique solution in N_D^2. Moreover, $\sup\limits_{0 \leqslant t \leqslant T} E|X_t|_D^2 < +\infty$. Therefore (e) is satisfied.

We have the estimation

$$\begin{aligned} &\|AT_{t-s}B(X_s, s)Q^{1/2}\|^2_{\mathrm{HS}} \\ &\quad = \sum_{i=1}^{\infty} \alpha_i |AT_{t-s}B(X_s, s)(e_i)|_H^2 \leqslant \sum_{i=1}^{\infty} \alpha_i \|AT_{t-s}\|^2_{\mathscr{L}(D,H)} |B(X_s, s)(e_i)|_D^2 \\ &\quad \leqslant \sum_{i=1}^{\infty} \alpha_i \|T_{t-s}\|^2_{\mathscr{L}(H,H)} \cdot k \cdot (1 + |X_s|_D^2). \end{aligned}$$

Consequently (f) is also fulfilled, and by Proposition 3.26 the existence is proved. The remaining part of the theorem follows from Proposition 3.8 and Remark 3.9.

Remark 3.29. Analogously: If the assumptions of Theorem 3.19, with H replaced by the space D, are satisfied, then equation $(*{*}*)$ has a unique continuous solution.

References

[1] A. V. Balakrishnan, *Introduction to optimization theory in a Hilbert space*, Springer-Verlag, Berlin 1971.

[2] A. Bensoussan, *Filtrage optimal des systems lineaires*, Dunod, Paris 1971.

[3] A. Bensoussan, R. Temam, *Equations aux derivees partielles stochastiques*, Israel J. Math. 11 (1972).

[4] A. Chojnowska-Michalik, *On linear stochastic delay equations*, to appear.

[5] R. Curtain, P. Falb, *Stochastic differential equations in Hilbert space*, J. Differential Equations 10 (1971).

[6] R. Curtain, *Stochastic evolution equations with general white noise disturbance*, Control Theory Centre Report 1975.
[7] I. I. Gihman, A. V. Skorohod, *Stochastic differential equations*, Berlin 1972.
[8] E. Hille, R. S. Philips, *Functional analysis and semigroups*, A. M. S., Providence, R. I. 1957.
[9] [R. Lipcer, A. Širyayev] Р. Липцер, А. Ширяев, *Статистика случайных процессов*, Москва 1974.
[10] J. L. Lions, *Quelques methodes de résolution des problémes aux limites non lineaires*, Dunod-Gauthier Villars, Paris 1969.
[11] H. McKean, *Stochastic integrals*, Academic Press, New York 1969.
[12] M. Metivier, G. Pistone, *Une formule d'isométrie pour l'intégrale stochastique hilbertienne et equations d'evolution lineaires stochastique*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 33 (1975).
[13] P. A. Meyer, *Probability and potentials*, 1966.
[14] E. Pardoux, *Non linear stochastic partial differential equations*, C. R. Acad. Sci. Paris, 1972.
[15] B. Rozowskii, *Stochastic differential equations in infinite dimensional spaces and filtering problems*, Proc. Druskininkai, Nov. 1974.
[16] R. B. Vinter, *A representation of solutions to stochastic delay equations*, Imperial College Report, 1975.
[17] J. Zabczyk, *On stability of infinite-dimensional linear stochastic systems*, this volume, pp. 273–281.

Presented to the Semester
Probability Theory
February 11–June 11, 1976

PROBABILITY THEORY
BANACH CENTER PUBLICATIONS, VOLUME 5
PWN—POLISH SCIENTIFIC PUBLISHERS
WARSAW 1979

ON χ^2 TESTS OF COMPOSITE HYPOTHESES

R. M. DUDLEY

Department of Mathematics, Massachusetts Institute of Technology, Cambridge Mass., USA

For a χ^2 test with m cells and a composite hypothesis in an s-dimensional submanifold V, Birch [4] showed that simple differentiability of V suffices to give a limiting χ^2_{m-s-1} distribution. Dzhaparidze and Nikulin [14] introduced modified X^2 statistics for composite hypotheses, which are easier to compute than the classical ones. Here, these results are proved for topologically non-trivial manifolds (such as circles and spheres).

1. Introduction

For background on the χ^2 test, we refer to Cramér [11], Lancaster [18] and C. R. Rao [25]. Let $\mathscr{L}(X)$ denote the probability distribution or *law* of a random variable X. Let χ_d^2 denote a χ^2 variable with d degrees of freedom, i.e. $\mathscr{L}(\chi_d^2) = \mathscr{L}(G_1^2+\ldots+G_d^2)$ where G_i are independent standard normal variables. Let $N(m, C)$ denote a Gaussian (normal) distribution on $\boldsymbol{R}^d$ with mean vector m and covariance matrix C. The characteristic function of a χ^2 variable is given by

$$E\exp(it\chi_d^2) = (1-2it)^{-d/2} = f(t)^{-d}, \tag{1.1}$$

where $f(t) = (1-2it)^{1/2}$, using the continuous branch of the square root with positive real part.

Let S be a finite set with m elements, say $S = \{1, 2, \ldots, m\}$. In the applications, S often results from decomposing a more general space into m cells. Let p and q be probability measures on S, $p\{j\} = p_j$, $q\{j\} = q_j$, $j = 1, \ldots, m$, with $p_j > 0$ for all j.

Let $Y_1, Y_2, \ldots$ be i.i.d. (independent and identically distributed) with distribution q. Given n, let $n_j = n_j(\omega, n)$ be the number of values of $i \leqslant n$ such that $Y_i = j$. Let

$$X^2 := \sum_{1\leqslant j\leqslant m} (n_j-np_j)^2/np_j.$$

If $q = p$, the central limit theorem in $\boldsymbol{R}^m$ implies that $\mathscr{L}(X^2)\to\mathscr{L}(\chi^2_{m-1})$ as $n\to\infty$ for the usual convergence of laws. Thus if q is unknown but Y_j can be observed, the hypothesis $q = p$ can be tested by the χ^2 test using the X^2 statistic.