Nr. 214/1996

A Review of White Noise Analysis from a Probabilistic Standpoint

Th. Deck, J. Potthoff, G. Våge

A Review of White Noise Analysis from a Probabilistic Standpoint

TH. DECK†, J. POTTHOFF†, G. VÅGE‡

Lehrstuhl für Mathematik V
Universität Mannheim
D–68131 Mannheim

Abstract The main notions and tools from white noise analysis are set up on the basis of the calculus of Gaussian random variables and the S–transform. A new proof of the formula for the S–transform of Itô integrals is given. Moreover, measurability and the martingale property with respect to the Brownian filtration are characterized in terms of the S–transform. This allows to extend these notions to random variables and processes, respectively, in the space of Hida distributions.

1. Introduction

This paper is not written for specialists in Gaussian white noise analysis and/or Malliavin calculus, but rather as an introduction for probabilists who might want to apply these methods to concrete problems. In particular, the specialists will not find much here which is really new, except for a slightly different perspective, the extension of the notions of measurability and martingales to the space of Hida distributions, and a – as far as we know – new simple proof for the formula of the S–transform of Itô integrals.

We have tried to use as little as possible from the theory of topological vector spaces, and the theory of Gaussian measures on them. In particular, the chaos expansion which is at the basis of most of the literature is not used at all (although mentioned here and there). Rather, we develop the framework solely out of Gaussian random variables on a general probability space and the S–transform. This entails that we cannot prove all statements in this paper, but sometimes have to refer the reader who is interested in more details to the standard literature. Also, we do not attempt to present things in greatest generality.

The approach taken in this paper makes it possible to unify the basic set-ups of the white noise and Malliavin calculi, which differ essentially only by the concrete choice of the underlying probability space. (The more developed aspects and applications of the two calculi are, of course, quite different.) Moreover, we hope that our approach will be helpful in developing numerical methods for white noise analysis and its applications.

†Partially supported by the *DFG*.

‡Supported by the *DFG*.

Except for a number of examples for more pedagogical purposes, there are no real applications of white noise analysis discussed in this paper. Instead, the interested reader is referred to the large existing literature, cf. e.g., [BD 95, DP 95, DP 96, FH 94 Hi 80, HK 93, HO 96, KL 94, LL 93, Po 94, PV 96, Wa 91], and the references given therein.

In this paper we shall work with three choices for the domain of the time parameter: $[0,1]$, $\mathbb{R}_+$, and $\mathbb{R}$. Many of the statements and notions have straightforward generalizations to more general domains, but this is left to the interested reader. In order to define the corresponding spaces of Hida distributions, we construct (nuclear) spaces of smooth functions of the time parameter in an Appendix. This is done with the help of concrete choices of standard systems of functions, namely the trigonometric, Laguerre, and Hermite functions.

In Section 2 we state our basic assumptions and define the S–transform of random variables. In Section 3 the measurability of random variables with respect to the filtration of Brownian motion is characterized in terms of the S–transform, and the S–transform of Itô integrals is computed. Martingales with respect to the Brownian filtration are characterized via the S–transform in Section 4. In Section 5 we construct the space of Hida distributions, state the characterization theorem and two of its corollaries. These results are then used to extend the notions of measurability and martingales to generalized random variables and processes which are Hida distributions. In particular, with our definition of a generalized martingale white noise becomes a martingale. In Section 6 we give a short introduction to differential calculus with Gaussian random variables.

2. Gaussian Random Variables, S–Transform

Let $(\Omega, \mathcal{B}, P)$ be a probability space. We assume that $\mathcal{B}$ and all other σ–algebras in this paper are P–complete. For the time parameter we choose the following sets: $[0,1]$, $\mathbb{R}_+ = [0,\infty)$, or $\mathbb{R}$, and denote them by T in all three cases. T is equipped with the Borel sets $\mathcal{B}(T)$ and Lebesgue measure λ. (For simplicity, integrals of functions, say f, with respect to λ will be denoted by $\int_T f(t)\,dt$ etc.)

We make the following basic assumption:

(H.1) There exists a family $(X_h, h \in C_c^\infty(T))$ of real centered Gaussian random variables on $(\Omega, \mathcal{B}, P)$, with covariance determined by the inner product $(\cdot,\cdot)$ on $L^2(T,\lambda)$: $\mathbb{E}(X_h X_f) = (h, f)$, where $(h,f) = \int_T h(t) f(t)\,dt$.

Note that if h, $f \in C_c^\infty(T)$, $\alpha \in \mathbb{R}$, then $X_{h+\alpha f}$ is P–a.s. equal to $X_h + \alpha X_f$.

Clearly, (H.1) implies that for all $h \in C_c^\infty(T)$, $X_h \in L^p(\Omega, \mathcal{B}, P) \equiv L^p(P)$, for all $p \geq 1$. (We choose all spaces $L^p(P)$ to be real in this paper, except where otherwise noted.) The monotone convergence theorem entails that also $\exp X_h \in L^p(P)$ for all $p \geq 1$, $h \in C_c^\infty(T)$, and that $\exp(\alpha X_h^2) \in L^2(P)$ if $\alpha < (4\,|h|_2^2)^{-1}$, where $|\cdot|_2$ denotes the norm on $L^2(T,\lambda)$.

Since we can approximate $h \in L^2(T,\lambda)$ by a sequence $(h_n \in C_c^\infty(T), n \in \mathbb{N})$ in $L^2(T,\lambda)$, it follows that $(X_{h_n}, n \in \mathbb{N})$ is a Cauchy sequence in $L^2(P)$, and we denote its limit by X_h. It follows that this extension of X_h to $h \in L^2(T,\lambda)$ is also a.s. linear, and

$(X_h, h \in L^2(T,\lambda))$ forms a centered Gaussian family with covariance $\mathbb{E}(X_h X_f) = (h,f)$. Choosing $h = 1_{[0,t)}$ for $t \in T$, we obtain a Brownian motion $(B_t = X_{1_{[0,t)}}, t \geq 0)$ on our probability space. (As usual, we may assume that B_t denotes the version with a.s. continuous sample paths.)

Conversely, if instead of (H.1) we assume in the cases $T = [0,1]$, $\mathbb{R}_+$ that we are given a standard Brownian motion $(B_t, t \in T)$ on $(\Omega, \mathcal{B}, P)$, then we prove that (H.1) holds: for given $h \in C_c^\infty(T)$, let $[t_0, t]$ be a finite interval in T so that $\operatorname{supp} h \subset [t_0, t]$. Consider a refining sequence $(P_n, n \in \mathbb{N})$ of partitions P_n, given by $t_1, ..., t_{N(n)-1} \in [t_0, t]$, $t_0 < t_1 < t_2 < ... < t_{N(n)-1}$, $N(n) \in \mathbb{N}$ $(t_{N(n)} = t)$, so that the mesh of P_n tends to zero as n tends to infinity. Set

$$Y_n := \sum_{i=1}^{N(n)} h(t_{i-1})(B_{t_i} - B_{t_{i-1}}).$$

Then $(Y_n, n \in \mathbb{N})$ converges in $L^2(P)$ to a centered Gaussian random variable, denoted by X_h, with variance $|h|_2^2$. If $h, f \in C_c^\infty(T)$ one obtains (e.g., by polarization) $\mathbb{E}(X_h X_f) = (h,f)$. Hence we can replace (H.1) by:

(H.1′) $(T = [0,1]$ or $\mathbb{R}_+)$ There exists a Brownian motion $(B_t, t \in T)$ on $(\Omega, \mathcal{B}, P)$.

We shall always assume that the random variables X_h, $h \in L^2(T,\lambda)$, and B_t, $t \geq 0$, are related as above. In particular, we have $\mathbb{E}(X_h B_t) = \int_0^t h(u)\,du$. Furthermore, if for $t \in T$, $\mathcal{B}_t$ denotes the sub–σ–algebra generated by $(B_s, s \in [0,t])$, and h has essential support in $[0,t]$ (i.e., $1_{[0,t]^c} \cdot h = 0$, λ–a.e., where $\cdot^c$ denotes complement in T), then X_h is $\mathcal{B}_t$–measurable. If $\operatorname{ess-supp} h \subset [0,t]^c$, then we find that X_h is independent of $\mathcal{B}_t$.

In our next assumption we require that "the random variables in $L^2(P)$ are functions of the Gaussian random variables X_h". However, instead of formulating this in terms of a measurability condition (in a sense the weakest form) we use a (stronger) functional–analytic condition:

(H.2) The algebra $\mathcal{A}$ generated by the Gaussian random variables $(X_h, h \in C_c^\infty(T))$ is dense in $L^2(P)$.

In case $T = [0,1]$ or $T = \mathbb{R}_+$, we have equivalently

(H.2′) $(T = [0,1]$ or $T = \mathbb{R}_+)$ The algebra $\mathcal{A}'$ generated by $(B_t, t \in T)$ is dense in $L^2(P)$.

(To see the equivalence of (H.2) and (H.2′), one has to observe that with $h_n \to h$ in $L^2(T,\lambda)$, one gets $X_{h_n} \to X_h$ in every $L^p(P)$, $p \geq 1$.) We may also replace (H.2) by

(H.2″) The algebra $\mathcal{E}$ generated by $(e^{X_h}, h \in C_c^\infty(T))$ (which coincides with the linear span of $(e^{X_h}, h \in C_c^\infty(T))$) is dense in $L^2(P)$.

(H.2) and (H.2″) are equivalent because on one hand the Taylor series $\sum_n (n!)^{-1} X_h^n$ of $\exp(X_n)$ converges in $L^2(P)$ to $\exp(X_h)$, and on the other hand

$$n(e^{n^{-1}X_h} - 1)$$

converges in $L^2(P)$ to X_h, as n tends to infinity.

In the standard models for $(\Omega, \mathcal{B}, P)$, namely the white noise space (which is often only considered for the choice $T = \mathbb{R}$) and the Wiener space (usually for $T = [0,1]$ or $T = \mathbb{R}_+$) the hypotheses (H.1) and (H.2) are fulfilled.

Next we define our central tool:

Definition 2.1 For $Y \in L^2(P), h \in C_c^\infty(T)$ define

$$SY(h) := \mathbb{E}(Y : e^{X_h} :),$$

where

$$\begin{aligned} : e^{X_h} : &= (\mathbb{E}e^{X_h})^{-1} e^{X_h} \\ &= e^{X_h - \frac{1}{2}|h|_2^2}. \end{aligned}$$

Remarks 1. Since $\mathcal{E}$ is dense in $L^2(P)$ it follows that S is injective: if $SY(h) = 0$ for all $h \in C_c^\infty(T)$, then $\mathbb{E}(Y\varphi) = 0$ for all $\varphi \in \mathcal{E}$, and therefore by density $Y = 0$ P–a.s. (henceforth we suppress "P–a.s.").

2. We have already observed that X_h has a continuous extension to $h \in L^2(T, \lambda)$ in $L^2(P)$. Thus for $Y \in L^2(P)$, also $h \mapsto SY(h)$ has a continuous extension to $h \in L^2(T, \lambda)$ (which we denote by SY, too).

3. It is useful to note that if g, $h \in L^2(T, \lambda)$, then

$$: \exp(X_{g+h}) : = : \exp(X_g) :: \exp(X_h) :$$

if and only if g and are h are orthogonal.

The most important example of the calculation of an S–transform is probably the following.

Example 2.2 Let $f \in L^2(T, \lambda)$. Then

$$(S : e^{X_f} :)(h) = e^{(f,h)}, \quad h \in L^2(T, \lambda). \tag{2.1}$$

For $h = 0$ the relation is evident. For $h \neq 0$, write $f = |h|_2^{-2}(f,h)h + f^\perp$ with $f^\perp$ orthogonal to h in $L^2(P)$. Thus $X_{f^\perp}$ and X_h are independent and therefore

$$\begin{aligned} (S : e^{X_f} :)(h) &= \mathbb{E}(: e^{|h|_2^{-2}(f,h)X_h} :: e^{X_h} :) \\ &= e^{-\frac{1}{2}|h|_2^{-2}(f,h)^2 - \frac{1}{2}|h|_2^2}\, \mathbb{E}\big(e^{X_h(1+|h|_2^{-2}(f,h))}\big) \\ &= e^{-\frac{1}{2}|h|_2^{-2}(f,h)^2 - \frac{1}{2}|h|_2^2}\, e^{\frac{1}{2}|h|_2^2(1+|h|_2^{-2}(f,h))^2} \\ &= e^{(f,h)}. \end{aligned}$$

Note that $x \mapsto e^{\alpha x - \frac{1}{2}\alpha^2\sigma^2}$ is the generating function of the Hermite polynomials $H_{n,\sigma^2}(x)$, $n \in \mathbb{N}_0$ with (variance) parameter σ^2. Hence from (2.1) we get by differentiation the formula

$$(SH_{n,|f|_2^2}(X_f))(h) = (h,f)^n, \quad n \in \mathbb{N}, h \in L^2(T, \lambda). \tag{2.2}$$

For example,

$$S(B_t^2 - t)(h) = (\int_0^t h(s)\,ds)^2.$$

As another example consider *white noise* $\dot{B}(t)$, i.e., the time derivative of Brownian motion. It is well–known that the limit

$$\lim_{\Delta t \to 0} (\Delta t)^{-1}(B_{t+\Delta t} - B_t) \tag{2.3}$$

cannot exist in the sense of a usual random variable. In particular, it is easy to check that this limit cannot exist in $L^2(P)$. On the other hand, if we compute the S–transform of $(\Delta t)^{-1}(B_{t+\Delta t} - B_t)$ at $h \in C_c^\infty(T)$, we get

$$(\Delta t)^{-1} \int_t^{t+\Delta t} h(s)\,ds \xrightarrow[\Delta t \to 0]{} h(t),$$

which is well–defined, and suggests that the limit (2.3) might exist in a weaker topology than the $L^2(P)$–topology. This is one of the motivations to consider generalized random variables and processes, which will be done in Section 5.

3. *S*–Transform of Itô Integrals

In this section we derive the well–known formula for the S–transform of Itô integrals with respect to Brownian motion. However, we give a proof which is solely based on the usual construction of the Itô integral and quite different to the standard proofs in the literature, which use the representation of Itô's integral in terms of the adjoint of the gradient operator (cf. Section 6).

The formula we are aiming at is very simple, and therefore also quite useful. It is the starting point of very many applications of white noise analysis. For simplicity and without loss of generality, we assume throughout this section that $T = [0,1]$ or $T = \mathbb{R}_+$. We begin with a very simple but useful result.

Lemma 3.1 Let $Y, Z \in L^2(P)$, and for $t \geq 0$, let $\mathcal{B}_t = \sigma(B_s, 0 \leq s \leq t)$. Then $Z = \mathbb{E}(Y|\mathcal{B}_t)$ if and only if $SZ(h) = SY(1_{[0,t]} \cdot h)$ for all $h \in C_c^\infty(T)$.

Proof We denote $h_t = 1_{[0,t]}h$, $h_t^\perp = h - h_t$. Note that h_t and $h_t^\perp$ are orthogonal in $L^2(T, \lambda)$. Since $: \exp(X_{h_t^\perp}) :$ is independent of $\mathcal{B}_t$ and has expectation one, we find

$$\begin{aligned}\mathbb{E}(: e^{X_h} : |\mathcal{B}_t) &= \mathbb{E}(: e^{X_{h_t}} :: e^{X_{h_t^\perp}} : |\mathcal{B}_t) \\ &= : e^{X_{h_t}} : .\end{aligned}$$

Hence we obtain from the projection property of the conditional expectation for $h \in C_c^\infty(T)$:

$$\begin{aligned}S\big(\mathbb{E}(Y|\mathcal{B}_t)\big)(h) &= \mathbb{E}\big(\mathbb{E}(Y|\mathcal{B}_t) : e^{X_h} :\big) \\ &= \mathbb{E}\big(Y\,\mathbb{E}(: e^{X_h} : |\mathcal{B}_t)\big) \\ &= SY(h_t).\end{aligned}$$

The injectivity of the S–transform concludes the proof. $\square$

Corollary 3.2 Let Y be in $L^2(P)$. The following are equivalent:

(i) Y is $\mathcal{B}_t$–measurable.

(ii) For all $h \in C_c^\infty(T)$, $SY(h) = SY(1_{[0,t]} \cdot h)$.

(iii) For all $h \in C_c^\infty(T)$ and all $g \in C_c^\infty([0,t]^c)$, $SY(h+g) = SY(h)$.

Remarks In (ii) one can replace $h \in C_c^\infty(T)$ by $h \in L^2(T,\lambda)$, and in (iii) $h \in C_c^\infty(T)$, $g \in C_c^\infty([0,t]^c)$ resp., by $h, g \in L^2(T,\lambda)$ with $\text{ess-supp}\, g \subset [0,t]^c$, without changing the equivalence of the statements. The point of the reformulation (iii) of (ii) is that (ii) does not extend straightfowardly to generalized random variables in the space of Hida distributions (cf. Section 5).

Proof We use the same notation as in the proof of Lemma 3.1, and show first "(i) $\Longleftrightarrow$ (ii)": Y is $\mathcal{B}_t$–measurable if and only if $\mathbb{E}(Y|\mathcal{B}_t) = Y$. By injectivity of the S–transform this is equivalent to $S\big(\mathbb{E}(Y|\mathcal{B}_t)\big)(h) = SY(h)$ for all $h \in C_c^\infty(T)$. By Lemma 3.1 the last statement is equivalent to $SY(h) = SY(h_t)$ for all $h \in C_c^\infty(T)$.

Next we prove "(ii) $\Longleftrightarrow$ (iii)". Actually, the implication "(ii) $\Rightarrow$ (iii)" is trivial. Assume that (iii) holds, and note that by continuity also (iii) extends to all $h, g \in L^2(T,\lambda)$ with $\text{ess-supp}\, g \subset [0,t]^c$. Now let $\widehat{h} \in C_c^\infty(T)$ and set $h = 1_{[0,t]} \cdot \widehat{h}$ and $g = 1_{[0,t]^c} \cdot \widehat{h}$. But then (iii) implies (ii) for $\widehat{h}$. $\square$

Theorem 3.3 Assume that $(Y_t,\, t \in T)$ is a $\mathcal{B}_t$–adapted process in $L^2(P \otimes \lambda)$. Then for all $h \in C_c^\infty(T)$, $t \in T$,

$$\Big(S \int_0^t Y_s\, dB_s\Big)(h) = \int_0^t SY_s(h)\, h(s)\, ds, \tag{3.1}$$

where the stochastic integral on the lefthand side is the Itô integral.

Proof It is well–known (e.g., [Do 52, Øk 85]) that there exists a refining sequence of partitions of $[0,t]$ – say, given by $0 = t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_{N(n)} = t$, $n \in \mathbb{N}$, and for each $n \in \mathbb{N}$ a simple, $\mathcal{B}_t$–adapted process $\widehat{Y}_n$ which is constant on each subinterval $[t_{k-1}, t_k)$, so that

$$\sum_{k=1}^{N(n)} \widehat{Y}_{n,t_{k-1}} (B_{t_k} - B_{t_{k-1}})$$

converges in $L^2(P)$ to $\int_0^t Y_s\, dB_s$, as n tends to infinity (the mesh of the partition tending to zero with $n \to +\infty$, and the convergence is uniform in t ranging over compacts of T). Moreover, we have that $\widehat{Y}_n$ converges to Y in $L^2(P \otimes \lambda)$, as n tends to $+\infty$.

Using the continuity of the inner product of $L^2(P)$ we get

$$\Big(S \int_0^t Y_s\, dB_s\Big)(h) = \lim_{n\to\infty} \sum_{k=1}^{N(n)} \mathbb{E}\big(\widehat{Y}_{n,t_{k-1}} (B_{t_k} - B_{t_{k-1}}) : e^{X_h} : \big).$$

Let $\triangle_k = [t_{k-1}, t_k)$, $Y_{n,k} := \widehat{Y}_{n,t_{k-1}}$, and note that $B_{t_k} - B_{t_{k-1}} = X_{1_{\triangle_k}}$. Write

$$\mathbb{E}\big(Y_{n,k}\, X_{1_{\triangle_k}} : e^{X_h} :\big) = e^{-\frac{1}{2}|h|_2^2} \frac{d}{d\alpha} \mathbb{E}\big(Y_{n,k}\, e^{X_{h+\alpha 1_{\triangle_k}}}\big)\Big|_{\alpha=0}.$$

One can use for example the dominated convegence theorem to justify the interchange of the expectation and the differentiation w.r.t. α in the last equation. We rewrite the last expectation as an S–transform and get

$$\begin{aligned}\mathbb{E}\big(Y_{n,k}\, X_{1_{\triangle_k}} : e^{X_h} :\big) &= e^{-\frac{1}{2}|h|_2^2} \frac{d}{d\alpha} SY_{n,k}(h + \alpha 1_{\triangle_k})\, e^{\frac{1}{2}|h+\alpha 1_{\triangle_k}|_2^2}\Big|_{\alpha=0} \\ &= e^{-\frac{1}{2}|h|_2^2} \frac{d}{d\alpha} SY_{n,k}(h)\, e^{\frac{1}{2}|h+\alpha 1_{\triangle_k}|_2^2}\Big|_{\alpha=0},\end{aligned}$$

where we used Corollary 3.2 and the remark following it. Hence we find that

$$\mathbb{E}\big(Y_{n,k} X_{1_{\triangle_k}} : e^{X_h} :\big) = SY_{n,k}(h) \int_{t_k-1}^{t_k} h(s)\, ds,$$

and

$$\Big(S \int_0^t Y_s\, dB_s\Big)(h) = \lim_{n\to\infty} \sum_{k=1}^{N(n)} SY_{n,k}(h) \int_{t_k-1}^{t_k} h(s)\, ds.$$

The last sum can be written as follows:

$$\mathbb{E}\Big(\int_T \widehat{Y}_{n,s} 1_{[0,t]}(s) h(s)\, ds\ : e^{X_h} :\Big) = (\widehat{Y}_n,\, 1_{[0,t]} h : e^{X_h} :)_{L^2(P\otimes\lambda)}.$$

Since $\widehat{Y}_n$ converges to Y in $L^2(P \otimes \lambda)$ as n tends to infinity, we obtain (3.1). □

As an illustration of equation (3.1) we give a proof of Itô's formula

$$f(B_t) - f(B_s) = \int_s^t f'(B_u)\, dB_u + \frac{1}{2} \int_s^t f''(B_u)\, du, \quad s,t \in T, s \le t, \tag{3.2}$$

where f belongs to the Schwartz space $\mathcal{S}(\mathbb{R})$ of rapidly decreasing C^∞ functions. The proof below is essentially a copy from [HK 93], where this is proved even for tempered distributions f. Such generalizations can easily be proved from (3.2) by taking appropriate limits. On the other hand, the argument below does unfortunately not extend in an obvious manner to the Itô formula for general (Itô) diffusions.

Here we shall deviate from the convention that we take all spaces of random variables real: for a moment it will be convenient to use complex random variables – the S–transform and all proved relations extend linearly to this case.

In order to show (3.2) it is sufficient to prove for all $h \in C_c^\infty(T)$:

$$\int_s^t Sf'(B_u)(h)\, h(u)\, du = Sf(B_t)(h) - Sf(B_s)(h) - \frac{1}{2} \int_s^t Sf''(B_u)(h)\, du, \tag{3.3}$$

because of Theorem 3.3 and the fact that the S–transform is injective. We write f as the Fourier transform of $\widehat{f}$:

$$f(x) = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} e^{i\alpha x} \widehat{f}(\alpha)\, d\alpha,$$

with $\widehat{f} \in \mathcal{S}(\mathbb{R})$. In the following the various interchanges of integrals are all easily jusitified by Fubini's theorem.

$$\begin{aligned}
\int_s^t Sf'(B_u)(h)\, h(u)\, du &= \int_s^t S\Big((2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} i\alpha e^{i\alpha B_u} \widehat{f}(\alpha)\, d\alpha\Big)(h)\, h(u)\, du \\
&= (2\pi)^{-\frac{1}{2}} \int_s^t \Big(\int_{\mathbb{R}} i\alpha (Se^{i\alpha B_u})(h) \widehat{f}(\alpha)\Big)\, h(u)\, du \\
&= (2\pi)^{-\frac{1}{2}} \int_s^t \Big(\int_{\mathbb{R}} i\alpha\, e^{-\frac{1}{2}\alpha^2 u + i\alpha \int_0^u h(v)\, dv} \widehat{f}(\alpha)\, d\alpha\Big)\, h(u)\, du \\
&= (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} i\alpha\, \widehat{f}(\alpha) \Big(\int_s^t e^{-\frac{1}{2}\alpha^2 u + i\alpha \int_0^u h(v)\, dv} h(u)\, du\Big)\, d\alpha \\
&= (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} \widehat{f}(\alpha) \Big(\int_s^t \Big(\frac{d}{du} + \frac{1}{2}\alpha^2\Big) e^{-\frac{1}{2}\alpha^2 u + i\alpha \int_0^u h(v)\, dv}\, du\Big)\, d\alpha \\
&= (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} \widehat{f}(\alpha) \Big(e^{i\alpha \int_0^t h(v)\, dv - \frac{1}{2}\alpha^2 t} - e^{i\alpha \int_0^s h(v)\, dv - \frac{1}{2}\alpha^2 s} \\
&\qquad + \frac{1}{2}\alpha^2 \int_s^t e^{i\alpha \int_0^u h(v)\, dv - \frac{1}{2}\alpha^2 u}\, du\Big)\, d\alpha \\
&= (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} \widehat{f}(\alpha) \Big(S(e^{i\alpha B_t})(h) - S(e^{i\alpha B_s})(h)\Big)\, d\alpha \\
&\qquad + \frac{1}{2}(2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} \alpha^2 \widehat{f}(\alpha) \int_s^t S(e^{i\alpha B_u})\, du\, d\alpha \\
&= S\Big((2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} \widehat{f}(\alpha)(e^{i\alpha B_t} - e^{i\alpha B_s})\, d\alpha\Big)(h) \\
&\qquad - \frac{1}{2} S\Big(\int_s^t ((2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} (i\alpha)^2 \widehat{f}(\alpha) e^{i\alpha B_u}\, d\alpha)\, du\Big)(h) \\
&= Sf(B_t)(h) - Sf(B_s)(h) - \frac{1}{2} \int_s^t Sf''(B_u)(h)\, du.
\end{aligned}$$

In a similar manner one can solve a number of simple stochastic differential equations, such as linear equations, along the above lines. Here is a typical, simple example:

Example 3.4 Let f be a function in $L^2_{\text{loc}}(T, \lambda)$, and consider the initial value problem given by the following Itô equation

$$dY_t = f(t) Y_t\, dB_t,$$

with the initial condition $Y_0 = y_0 \in \mathbb{R}$. Taking the S-transform of the equation we get for every $h \in C_c^\infty(T)$ the initial value problem

$$d\widehat{Y}_t = f(t) \widehat{Y}_t\, h(t)\, dt, \quad Y_0 = y_0,$$

where $\widehat{Y}_t$ denotes the S–transform of Y_t. The last problem is readily solved:

$$\widehat{Y}_t = y_0 \, e^{\int_0^t f(s)h(s)\,ds}.$$

But according to Example 2.2, this is the S–transform of

$$\begin{aligned} Y_t &= y_0 \; : e^{X_{1_{[0,t)}f}} : \\ &= y_0 \, e^{\int_0^t f(s)\,dB_s - \frac{1}{2}\int_0^t f(s)^2\,ds} \end{aligned}$$

which is the solution of the Itô equation.

4. Martingales

In this section we provide a convenient criterium in terms of the S–transform for when a given $\mathcal{B}_t$–adapted process $(Y_t,\, t \in T)$ in $L^2(P)$ is a martingale w.r.t. the filtration $(\mathcal{B}_t,\, t \in T)$. (Recall that in Corollary 3.2 we characterized $(\mathcal{B}_t,\, t \in T)$–adaptedness in terms of the S–transform.) It will be convenient to assume $T = [0,1]$ or $\mathbb{R}_+$ in this section.

Theorem 4.1 Let $(Y_t,\, t \in T)$ be a $\mathcal{B}_t$–adapted process in $L^2(P)$. $(Y_t,\, t \in T)$ is a $(\mathcal{B}_t,\, t \in T)$–martingale if and only if for all $s, t \in T$ with $s \le t$ and all $h \in C_c^\infty([0,s])$ the following holds

$$SY_t(h) = SY_s(h). \tag{4.1}$$

Proof Let Y be a $(\mathcal{B}_t,\, t \in T)$–martingale, $s, t \in T$ with $s \le t$, and $h \in C_c^\infty(T)$. Taking the S-transform of both sides of the equation $\mathbb{E}(Y_t|\mathcal{B}_s) = Y_s$, we get $S\big(\mathbb{E}(Y_t|\mathcal{B}_s)\big)(h) = SY_s(h)$. Applying Lemma 3.1 with $Z = Y_t$, we obtain $SY_t(1_{[0,s]} \cdot h) = SY_s(h)$, which entails (4.1) for $h \in C_c^\infty([0,s])$.

Conversely, assume that (4.1) holds for all $s \le t$ and all $h \in C_c^\infty([0,s])$. By continuity, (4.1) extends to $h \in L^2(T,\lambda)$ with $\operatorname{ess-supp} h \subset [0,s]$. Let $g \in C_c^\infty(T)$ and denote $g_s = 1_{[0,s]} \cdot g$. By Lemma 3.1,

$$\begin{aligned} S\mathbb{E}(Y_t|\mathcal{B}_s)(g) &= SY_t(g_s) \\ &= SY_s(g_s) \\ &= SY_s(g), \end{aligned}$$

where we used Corollary 3.2 in the last step. Injectivity of the S–transform implies $\mathbb{E}(Y_t|\mathcal{B}_s) = Y_s$. □

Remark It would be nice to characterize also the sub– and supermartingale properties by the S–transform. Necessary conditions are easily derived as above, but sufficiency seems to be a difficult problem, since one would have to conclude from the positivity of an S–transform the positivity of its inverse. But in general the positivity of the S–transform may not imply positivity of the random variable: a counterexample in the context of Hida distributions can be found in [HK 93, Example 4.30]. Whether the condition that $Y_t \in L^2(P)$ rules out all such counterexamples is not known to us.

As a trivial application one just needs to look at equation (3.1) to see that every Itô integral defines a martingale. Moreover, if we define $g_t := 1_{[0,t]} \cdot g$ for $g \in L^2_{\text{loc}}(T, \lambda)$, then the above result together with equation (2.1) tell us that $: \exp(X_{g_t}) :$ is a martingale w.r.t. the filtration $(\mathcal{B}_t,\, t \in T)$.

5. Hida Distributions

As already pointed out at the end of Section 2, it is useful to consider limits of sequences of random variables which converge in weaker topologies than the one of $L^2(P)$. This leads to *generalized random variables*, much in the same way as one discusses generalized functions in the finite dimensional setting. There are many choices of spaces of generalized random variables. In this section we shall only discuss the so-called *Hida distributions*, although many results of this section can be generalized to other spaces of generalized random variables (cf., e.g., [Hi 80, HK 93, IK 89, KLS 94, PT 95]) as well. There are many results about the space of Hida distributions and its various generalizations. We refer the interested reader to, e.g., [HK 93] and the literature quoted there for more details, results, and proofs of some of the results below.

As in the finite dimensional case, we shall first construct a space $(\mathcal{S})$ of smooth random variables in $L^2(P)$ with a stronger topology. This can be done along the recipe for the construction of nuclear spaces given in the Appendix. However, for this we would need the chaos expansion of $L^2(P)$, which we want to avoid in this paper. Instead, we work with analogues of the number operators in the Appendix in their form as differential operators, and define these analogues on the dense linear hull of the exponential functions.

Again, we consider the three cases $T = [0,1]$, $T = \mathbb{R}_+$, or $T = \mathbb{R}$. The corresponding self-adjoint differential operators from the Appendix, which we denote by A in all cases, are:

$$\begin{array}{ll} T = [0,1]: & A = \sqrt{-\Delta} + 2 \text{ (periodic boundary conditions)}; \\ T = \mathbb{R}_+: & A = -\frac{d}{dt} t \frac{d}{dt} + \frac{1}{4} t + 1; \\ T = \mathbb{R}: & A = \frac{1}{2}(-\Delta + t^2) + 1. \end{array}$$

Note that we have added constants to the number operators given in the Appendix, to ensure that the spectrum of the operator A is strictly larger than 1. This guarantees that the space $(\mathcal{S})$ we construct is nuclear.

For $h \in C_c^\infty(T)$, we set

$$\Gamma(A) \, : e^{X_h} : = : e^{X_{Ah}} :,$$

and extend this definition linearly to the dense subspace $\mathcal{E}$. Thus $\Gamma(A)$ is a densely defined operator on $L^2(P)$, and it can be shown that it has a unique, self-adjoint extension (e.g., [Co 53, Si 71]), which we denote by the same symbol. Moreover, the bottom of the spectrum of $\Gamma(A)$ is 1 (the eigenfunctions being the constants: choose $h = 0$ above). Thus the following defines for every $p \in \mathbb{N}_0$ a (Hilbertian) norm on $\mathcal{E}$:

$$\|\varphi\|_{2,p} := \|\Gamma(A)^p \varphi\|_2, \quad \varphi \in \mathcal{E},$$

where $\|\cdot\|_2 = \|\cdot\|_{2,0}$ is the norm of $L^2(P)$. Now we define a subspace $(\mathcal{S})_p$ of $L^2(P)$ as the completion of $\mathcal{E}$ with respect to the norm $\|\cdot\|_{2,p}$. One can check that the sequence

$((\mathcal{S})_p, p \in \mathbb{N}_0)$ of Hilbert spaces forms a chain as in the Appendix, and therefore its intersection $(\mathcal{S})$ is a countably Hilbert space. Furthermore, it turns out that $(\mathcal{S})$ is nuclear. Its dual (i.e., the space of continuous linear functions on $(\mathcal{S})$) is called the space of *Hida distributions* and is denoted by $(\mathcal{S})^*$.

Before we come to the next result – which will enable us to work efficiently with $(\mathcal{S})$ and $(\mathcal{S})^*$ – we mention that by construction $\mathcal{E} \subset (\mathcal{S})$.

In order to make the so far relatively abstract space $(\mathcal{S})^*$ more concrete, we extend our central tool – the S–transform – to $(\mathcal{S})^*$: for $\Phi \in (\mathcal{S})^*$ we set

$$S\Phi(h) := \langle\!\langle \Phi, : e^{X_h} : \rangle\!\rangle, \quad h \in C_c^\infty(T), \tag{5.1}$$

where $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ is the dual pairing between $(\mathcal{S})$ and $(\mathcal{S})^*$. That (5.1) is a legitimate definition is due to the before mentioned fact that $\mathcal{E} \subset (\mathcal{S})$. Note that $S\Phi$ is an everywhere defined function on $C_c^\infty(T)$. It is useful to consider a space of functions on $C_c^\infty(T)$ with the following properties:

Definition 5.1 The space of functions F on $C_c^\infty(T)$ which satisfy:

(A) F has an extension which is *ray entire* on $C_c^\infty(T)$, i.e., for all $g, h \in C_c^\infty(T)$, the function $x \mapsto F(g + xh)$ on $\mathbb{R}$ has an entire analytic extension to the complex plane (which we denote by $z \mapsto F(g + zh)$);

(B) F satisfies the following bound: there exists a positive constant K and a continuous seminorm $\|\cdot\|$ on $\mathcal{S}(T)$, so that for all $h \in C_c^\infty(T)$ and all $z \in \mathbb{C}$,

$$|F(zh)| \leq K\, e^{|z|^2 \|h\|^2};$$

is denoted by $\mathcal{U}$.

In the preceding definition we used the Schwartz space $\mathcal{S}(T)$ constructed in the Appendix, and the notions of convergence and continuity in $\mathcal{S}(T)$ defined there.

The next result goes under the name "characterization theorem for Hida distributions". For the detailed proof, refinements, variants and by-products we refer to [HK 93, KL 94, PS 91] and the literature quoted there.

Theorem 5.2 The S–transform is a bijection from $(\mathcal{S})^*$ onto $\mathcal{U}$.

The most interesting part of this statement for applications is that S is onto, i.e., that the S–transform is invertible. (And this is also the only non-trivial part of its proof: that it is injective follows as injectivity on $L^2(P)$, and that S maps $(\mathcal{S})^*$ into $\mathcal{U}$ is simply due to the fact that the exponential function is analytic. Let us give a quick sketch of the proof that S is surjective for the reader who knows the chaos expansion: given $F \in \mathcal{U}$ we use (A) to write the Taylor series of $z \longmapsto F(zh)$. The Taylor coefficients are n–multilinear functions of h, and we would like to identify them with the kernels of the chaos expansion of a $\Phi \in (\mathcal{S})^*$. To this end, one uses (A), (B), and the Cauchy estimate to establish that these kernels indeed satisfy the properties of those of an element in $(\mathcal{S})^*$.)

Example 5.3 Let $t \in T$, $t > 0$, $n \in \mathbb{N}$, and consider the following mappings on $C_c^\infty(T)$: $h \mapsto h(t)$, $h(t)^n$, $\exp(h(t))$. According to a remark in the Appendix, they are all continuous on $\mathcal{S}(T)$, and they are certainly in the class $\mathcal{U}$. Hence by Theorem 5.2 they correspond to certain Hida distributions, which we denote by $\dot{B}(t)$, $:\dot{B}(t)^n:$, and $:\exp(\dot{B}(t)):$, respectively. In order to understand their probabilistic meaning we need the next result, which is a by-product of the proof of Theorem 5.2 ([DP 96], cf. also [HK 93, KL 94, PS 91]):

Theorem 5.4 Assume that (M, d) is a metric space, and that F is a mapping from M into $\mathcal{U}$. Suppose furthermore, that the continuous seminorm $\|\cdot\|$ on $\mathcal{S}(T)$ and the constant K for $F(\alpha)$, $\alpha \in M$ in Definition 5.1 can be chosen uniformly with respect to $\alpha \in M$, and that for each $h \in C_c^\infty(T)$, the mapping $\alpha \mapsto F(\alpha)(h)$ is continuous. Let $\Phi(\alpha) = S^{-1}F(\alpha)$. Then the mapping $\alpha \mapsto \Phi(\alpha)$ is continuous from (M, d) into $(\mathcal{S})^*$, equipped with its strong topology.

One of the uses of the last theorem is to discuss limits: typically, if $\varepsilon \to 0$ in $\mathbb{R}$, and F_ε is an element in $\mathcal{U}$ with the properties stated in Theorem 5.4: F_ε satifies (A) and (B) above uniformly in ε ranging over a neighborhood of zero, and for every $h \in C_c^\infty(T)$, $F_\varepsilon(h)$ converges to $F_0(h)$, then the corresponding Hida distribution Φ_ε converges strongly to Φ_0. (One can discuss differentiability with respect to a parameter in a similar manner: [DP 96, Po 94].) For the Hida distributions in Example 5.3, we consider again the difference quotient

$$(\Delta t)^{-1}\big(B(t + \Delta t) - B(t)\big),$$

whose S–tranform at $h \in C_c^\infty(T)$ is

$$(\Delta t)^{-1}\int_t^{t+\Delta t} h(s)\,ds.$$

It is quite obvious that the requirements of Theorem 5.4 are fulfilled (say, for Δt in an interval $[-\delta, \delta]$ with $\delta > 0$ small enough so that $t - \delta \geq 0$). Therefore, the difference quotient above converges strongly in $(\mathcal{S})^*$ to $\dot{B}(t)$. Hence we have constructed *white noise* $\dot{B}(t)$ *at time* t as a Hida distribution. The other examples are similar: for $:\dot{B}(t)^n:$, one chooses

$$H_{n,\Delta t^{-1}}\Big((\Delta t)^{-1}\big(B(t + \Delta t) - B(t)\big)\Big),$$

where H_{n,σ^2} is the n–th Hermite polynomial with variance parameter $\sigma^2 > 0$ (cf., Section 2). By equation (2.2) (choose $f = \Delta t^{-1} 1_{[t,t+\Delta t]}$), the S–transform of the above expression is equal to $\big(\Delta t^{-1}\int_t^{t+\Delta t} h(s)\,ds\big)^n$. An application of Theorem 5.4 therefore shows that $H_{n,\Delta t^{-1}}\big((\Delta t)^{-1}(B(t + \Delta t) - B(t))\big)$ converges strongly in $(\mathcal{S})^*$ to $:\dot{B}(t)^n:$ – the *n–th renormalized power of white noise*. For the *renormalized exponential* $:\exp(\dot{B}(t)):$ *of white noise* one starts with the expression

$$\exp\Big(\Delta t^{-1}\big(B(t + \Delta t) - B(t)\big) - \frac{1}{2}\Delta t^{-1}\Big).$$

The details are left to the interested reader.

Another useful by-product of the proof of the characterization theorem pertains to integrals of Hida distributions (e.g., [HK 93, KL 94]):

Theorem 5.5 Let $(X, \mathcal{A}, \mu)$ be a measure space, and assume that F is a function on X with values in $\mathcal{U}$. Suppose that

(i) for each $h \in C_c^\infty(T)$, the mapping $x \mapsto F_x(h)$, $x \in X$ is $\mathcal{A}$–$\mathcal{B}(\mathbb{R})$ measurable;
(ii) there exist positive functions $K_1 \in L^1(X, \mu)$, $K_2 \in L^\infty(X, \mu)$, and a continuous semi-norm $\|\cdot\|$ on $\mathcal{S}(T)$, so that for all $z \in \mathbb{C}$, $h \in C_c^\infty(T)$, and μ–a.s. $x \in X$,

$$|F_x(zh)| \leq K_1(x) e^{K_2(x)|z|^2\|h\|^2}.$$

Then $\int_X F_x \, d\mu(x)$ defines a Hida distribution which is given by the Pettis integral

$$\int_X \Phi_x \, d\mu(x),$$

where $\Phi_x = S^{-1} F_x$.

Remark In the above statement one can actually use as well the notion of the Bochner integral on one of the constituent Hilbert spaces $(\mathcal{S})_p^*$ of $(\mathcal{S})^*$.

With the help of Theorems 5.4 and 5.5 one can construct and analyze many interesting examples of Hida distributions, for instance the classical *Donsker delta function*, see below, and, e.g., [HK 93].

Another useful observation is that the space $\mathcal{U}$ is obviously an algebra under pointwise multiplication. Hence Theorem 5.2 implies that there is a bilinear mapping from $(\mathcal{S})^* \times (\mathcal{S})^*$ into $(\mathcal{S})^*$, called the *Wick product*, so that for any two Hida distributions Φ, Ψ, their Wick product $\Phi \diamond \Psi$ has S–transform given by $S\Phi \cdot S\Psi$. For more details and applications we refer to [HK 93, HO 96, KLS 94] and the literature quoted there.

Next we use our framework to generalize stochastic integration with respect to Brownian motion, and the notion of martingales.

Let $(Y_t,\, t \in T)$ be a *generalized process in* $(\mathcal{S})^*$, i.e., we assume that for each $t \in T$, $Y_t \in (\mathcal{S})^*$, and that $t \mapsto Y_t$ is weakly measurable (that is, for every $\varphi \in (\mathcal{S})$, $\langle Y_t, \varphi \rangle$ is a measurable function of t). Suppose furthermore that $F_t(h)$ given by $F_t(h) = SY_t(h) \cdot h(t)$ satisfies the conditions of Theorem 5.5 with respect to the measure space $(T, \mathcal{B}(T), \lambda)$. Then $\int_T SY_t(h)\, h(t)\, dt$, is the S–transform of a Hida distribution, which itself is an integral over T. In view of formula (3.1) we have extended Itô's stochastic integral to Y_t:

Definition 5.6 Assume that $(Y_t,\, t \in T)$ is a generalized stochastic process in $(\mathcal{S})^*$, so that $(t, h) \mapsto SY_t(h) \cdot h(t)$ satisfies the hypotheses of Theorem 5.5 on $(T, \mathcal{B}(T), \lambda)$. Then the inverse S–transform of $\int_T SY_t(h)\, h(t)\, dt$, called the *Hitsuda–Skorokhod integral of* Y, is denoted by

$$\int_T Y_t \, dB(t).$$

Among other applications, this extended stochastic integral is useful when one wants to integrate processes Y which are not $\mathcal{B}_t$–adapted. In particular, it applies to the integration of processes which anticipate the future of the Brownian motion.

In view of Lemma 3.1 it is natural to make the following definition:

Definition 5.7 Let Φ be a Hida distribution, and let $t \in T$, $t \geq 0$. Φ is called $\mathcal{B}_t$–*measurable*, if for all $h \in C_c^\infty(T)$ and all $g \in C_c^\infty([0,t]^c)$, $S\Phi(h+g) = S\Phi(h)$. Assume that $T' = [a,b]$ or $T' = [a,+\infty)$ is a subinterval of T with $a \geq 0$. If Y is a generalized process in $(\mathcal{S})^*$ indexed by T' so that Y_t is $\mathcal{B}_t$–measurable for every $t \in T'$, then Y is called $(\mathcal{B}_t,\, t \in T')$–*adapted*.

Remark Note that by definition a generalized random variable $\Phi \in (\mathcal{S})^*$ which is $\mathcal{B}_t$–measurable is also $\mathcal{B}_{t'}$–measurable for all $t' \in T$ with $t' \geq t$.

Example 5.8 All examples in *Example 5.3* are for fixed $t > 0$ $\mathcal{B}_s$–measurable for $s \geq t$, but not $\mathcal{B}_s$–measurable for $s < t$. In particular, white noise, its renormalized powers, and its renormalized exponential are $\mathcal{B}_t$–adapted for $t > 0$. As another example consider Donsker's delta function: $\delta_x \circ B(t)$, $t > 0$, $x \in \mathbb{R}$. It is well-known, that it is a Hida distribution, and that its S–transform at $h \in C_c^\infty(T)$ is given by: $p_t(x, \int_0^t h(s)\,ds)$, where $p_t(x,y)$ is the usual heat kernel on the real line. (With Theorems 5.4 and 5.5 this is an easy exercise using the representation

$$\langle \delta_x, f \rangle = \lim_{n\to\infty} (2\pi)^{-1} \int_{-n}^{n} f(y) e^{ik(x-y)}\, dy.)$$

Hence we see immediately that $\delta_x \circ B(t)$ is $\mathcal{B}_t$–adapted for all $t > 0$. Consider the Hitsuda–Skorokhod integral $\int_a^t Y_s\, dB_s$ of a $(\mathcal{B}_t,\, t \in T')$–adapted generalized stochastic process Y. Then it follows from the definition that $(\int_a^t Y_s\, dB_s,\, t \in T')$ is also $(\mathcal{B}_t,\, t \in T')$–adapted.

Finally, we extend the notion of a martingale (with respect to the Brownian filtration) to processes in $(\mathcal{S})^*$. To this end, we use the condition in Theorem 4.1 which characterizes $L^2(P)$–martingales in an equivalent way:

Definition 5.9 Let T' be as in Definition 5.7, and assume that Y is a generalized process which is $(\mathcal{B}_t,\, t \in T')$–adapted. Y is called a *martingale with respect to* $(\mathcal{B}_t,\, t \in T')$ if and only if for all $t, s \in T'$ with $s \leq t$ and all $h \in C_c^\infty([0,s])$, $SY_t(h) = SY_s(h)$.

Remark Based on results in [Hi 80], smaller extensions of the notions of $\mathcal{B}_t$–measurability and –martingales have been considered in [BP 95]. There, sub– and supermartingales were also discussed.

Example 5.10 First, one realizes that it follows directly from the definition that the Hitsuda–Skorokhod integral of an adapted process is a martingale with respect to $\mathcal{B}_t$. Moreover, it is obvious that white noise, its renormalized powers and exponential function are $\mathcal{B}_t$–martingales for $t > 0$.

6. Differential Calculus

We conclude this paper with a short introduction to the differential calculus. Although most of the important notions and formulae can be obtained in our general framework, for a detailed study it is of advantage to use the standard set-ups of white noise and/or Malliavin calculus. In particular, Ω should have the structure of a topological vector space etc. However, since we do not intend to give the full theory of differential operators, the general framework is sufficient for our purposes.

First we prepare the following set of random variables as the domain of the differential operators: for given $n \in \mathbb{N}$, we let $C_e^\infty(\mathbb{R}^n)$ denote the infinitely differentiable functions on $\mathbb{R}^n$ which, together with all their derivatives, have at most exponential growth at infinity. We denote by $C_{fi,e}^\infty(\Omega)$ the set of all random variables Y over $(\Omega, \mathcal{B}, P)$ for which there exists $n \in \mathbb{N}$, $h_1, \ldots, h_n \in C_c^\infty(T)$, and $f \in C_e^\infty(\mathbb{R}^n)$, so that

$$Y = f(X_{h_1}, \ldots, X_{h_n}). \tag{6.1}$$

Note that $C_{fi,e}^\infty(\Omega)$ is a dense subspace of $L^2(P)$. For $Y \in C_{fi,e}^\infty(\Omega)$ of the form (6.1), and $h \in C_c^\infty(T)$ we set

$$D_h Y := \sum_{k=1}^n (\frac{\partial}{\partial x_k} f)(X_{h_1}, \ldots, X_{h_n})(h, h_k)_{L^2(T,\lambda)}. \tag{6.2}$$

(We leave it to the interested reader to check that $D_h Y$ is well-defined.) Note that we have $D_h Y \in C_{fi,e}^\infty(\Omega)$. In particular, we get for g, $h \in C_c^\infty(T)$,

$$D_h \; : \exp(X_g) : = (h, g) : \exp(X_g) : .$$

In order to obtain an integration by parts formula, we compute for f, g, $h \in C_c^\infty(T)$,

$$\begin{aligned}(X_h : e^{X_g} :, : e^{X_f} :) &= e^{-\frac{1}{2}(|g|_2^2+|f|_2^2)} \frac{d}{d\alpha} \mathbb{E}\big(e^{X_{g+f+\alpha h}}\big) \Big|_{\alpha=0} \\ &= (D_h : e^{X_g} :, : e^{X_f} :) + (h, f) e^{(g,f)} \\ &= (D_h : e^{X_g} :, : e^{X_f} :) + (: e^{X_g} :, D_h : e^{X_f} :).\end{aligned}$$

Therefore, the adjoint D_h^* of D_h with respect to P is given by the formula

$$D_h^* Y = X_h \cdot Y - D_h Y. \tag{6.3}$$

(We have shown that formula (6.3) holds at least on the dense domain $\mathcal{E}$. From there it extends to the natural domain of D_h^*, which turns out to be the same as the one of D_h. Anyway, we have equality (6.3) at least on $C_{fi,e}^\infty(\Omega)$.) From (6.3) it is plain to derive the commutation relation: $[D_h, D_g^*] = (h, g)_{L^2(T,\lambda)}$.

Note that we can write (6.2) also as

$$D_h Y = \int_T \partial_t Y h(t)\, dt, \tag{6.4}$$

with

$$\partial_t Y = \sum_{k=1}^n (\frac{\partial}{\partial x_k} f)(X_{h_1}, \ldots, X_{h_n}) h_k(t), \quad t \in T. \tag{6.5}$$

(We might also write ∂_t as D_{δ_t}.) Equation (6.4) shows that ∂_t defines a *gradient*.

An important second order differential operator is the so-called *number–* or *Ornstein–Uhlenbeck*–operator N given by

$$N := \sum_{k=0}^{\infty} D^*_{e_k} D_{e_k},$$

where $(e_k,\ k \in \mathbb{N}_0)$ is a CONS of $L^2(T,\lambda)$. The chaos expansion of $L^2(P)$ that we mentioned several times before is then the spectral decomposition of N. There are other interesting second order differential operators which have been investigated in the literature. We refer for example to [HK 93] and the literature cited there.

Now we derive the intertwining formulae with the S–transform. For $Y \in C^\infty_{fi,e}(\Omega)$, we get (as usual $h,\ g \in C^\infty_c(T)$)

$$\big(SD_g Y\big)(h) = D_g SY(h) \tag{6.6}$$

$$\big(SD^*_g Y(h)\big) = (g,h)SY(h), \tag{6.7}$$

where D_g on the right hand side of (6.6) means the usual directional (Gâteaux) derivative. Let us write the right hand side of (6.7) as $\int_T g(t)h(t)SY(h)\,dt$, and observe that if we set for fixed $t \in T$, $F(h) = h(t)SY(h)$, then $F \in \mathcal{U}$. The corresponding Hida distribution is denoted by $\partial^*_t Y$ so that we have

$$\big(S\partial^*_t Y\big)(h) = h(t)SY(h), \quad t \in T. \tag{6.8}$$

But then a comparison with (3.1) and Definition 5.6 shows that the Itô and Hitsuda–Skorokhod stochastic integrals of a process Y (satisfying the appropriate conditions specified in Sections 3 and 5, respectively) with respect to Brownian motion can be written as

$$\int_T \partial^*_t Y_t\,dt. \tag{6.9}$$

Appendix: On Nuclear Spaces

In this appendix we give a construction and some examples of nuclear vector spaces.

Nuclear topological vector spaces have among others two properties, which single them out among general topological vector spaces: On one hand they are not too far away from finite dimnesional vector spaces in the sense that their compacts are characterized as in finite dimensions, namely as bounded and closed. On the other, they are particularly easy to handle with respect to tensor products: one has the "nuclear" or "kernel" theorem.

Here, we shall only be concerned with nuclear countably Hilbert spaces, which is a sufficiently large class for our applications. The reader who wants to learn more about nuclear spaces is referred to the standard literature, e.g., [Pi 69, GV 64, Tr 67].

A *countably Hilbert space* $\mathcal{H}$ is by definition the limit (i.e., as a set the intersection) of a chain

$$\mathcal{H}_0 \supset \mathcal{H}_1 \supset \cdots \supset \mathcal{H}_n \supset \cdots$$

of separable Hilbert spaces $\mathcal{H}_n$ with norm $|\cdot|_n$, $n \in \mathbb{N}_0$, so that $\mathcal{H}_{n+1}$ is injectively, densely and continuously embedded into $\mathcal{H}_n$ for all $n \in \mathbb{N}_0$. One equips $\mathcal{H}$ with the metric given by:

$$d(u,v) = \sum_{n=0}^{\infty} 2^{-n} (1 + |u-v|_n)^{-1} |u-v|_n, \quad u,v \in \mathcal{H}.$$

$\mathcal{H}$ is called *nuclear* if for every $n \in \mathbb{N}_0$ there exists $m \in \mathbb{N}_0$, $m \geq n$, so that the injection ι_n^m from $\mathcal{H}_m$ into $\mathcal{H}_n$ is of *Hilbert–Schmidt type*, i.e., if for every CONS (e_l^m, $l \in \mathbb{N}$) of $\mathcal{H}_m$ the series $\sum_l |\iota_n^m e_l^m|_n^2$ converges.

From a Hilbert space point of view, nuclear countably Hilbert spaces are easy to construct by the following recipe: assume that we are given a separable Hilbert space $\mathcal{H}_0$ and a CONS (e_k^0, $k \in \mathbb{N}_0$) in $\mathcal{H}_0$. Let N denote the *number operator* with respect to (e_k^0, $k \in \mathbb{N}_0$) on $\mathcal{H}_0$: set

$$\mathcal{D}(N) = \{u \in \mathcal{H}_0;\, u = \sum_{k=0}^{\infty} u_k e_k^0, \text{with } \sum_{k=1}^{\infty} k^2 |u_k|_0^2 < +\infty\},$$

and for $u \in \mathcal{D}(N)$ define

$$Nu = \sum_{k=1}^{\infty} k\, u_k\, e_k^0.$$

Then it is easy to see that N is well–defined and self–adjoint. Similarly, one defines for $n \in \mathbb{N}$,

$$\mathcal{D}(N^n) = \{u \in \mathcal{H}_0;\, u = \sum_{k=0}^{\infty} u_k e_k^0, \text{with } \sum_{k=1}^{\infty} k^{2n} |u_k|_0^2 < +\infty\},$$

and for $u \in \mathcal{D}(N^n)$

$$N^n u = \sum_{k=1}^{\infty} k^n u_k e_k^0.$$

It is obvious that $\mathcal{D}(N^n)$ equipped with the inner product

$$(u,v)_n := ((1+N)^n u, (1+N)^n v)_0,$$

$(\cdot,\cdot)_0$ being the inner product of $\mathcal{H}_0$, is a Hilbert space $\mathcal{H}_n$, and by construction these Hilbert spaces form a chain as above. Moreover, for $n \in \mathbb{N}$, (e_k^n, $k \in \mathbb{N}_0$) defined by $e_k^n := (1+k)^{-n} e_k^0$, is a CONS in $\mathcal{H}_n$. It follows immediately that the embedding of $\mathcal{H}_{n+1}$ into $\mathcal{H}_n$ is of Hilbert–Schmidt type for every $n \in \mathbb{N}_0$. Therefore this chain of Hilbert spaces defines a nuclear space $\mathcal{H} = \bigcap_n \mathcal{H}_n$.

Now we apply this recipe to the space $L^2(T, \lambda)$ (real) for the three choices of T that we use in this paper:

Example A.1 $T = \mathbb{R}_+$. Let L_k, $k \in \mathbb{N}_0$, denote the Laguerre polynomial of degree k, and define

$$l_k(t) := e^{-t/2} L_k(t), \quad t \in \mathbb{R}_+,$$

called *Laguerre function of degree k*. It is well-known that $(l_k,\ k \in \mathbb{N}_0)$ is a CONS in $L^2(\mathbb{R}_+, \lambda)$ (e.g., [Sa 77]). Consider the differential operator A defined by

$$A := -\frac{d}{dt} t \frac{d}{dt} + \frac{1}{4} t.$$

Then the Laguerre functions are eigenfunctions of A:

$$A\, l_k = (k + \frac{1}{2})\, l_k, \quad k \in \mathbb{N}_0.$$

Thus the number operator is in this case equal to $A - \frac{1}{2}$.

We claim that in this case the nuclear space $\mathcal{H}$ consists of infinitely differentiable functions which, together with all their derivatives, vanish at $+\infty$, faster than any inverse power. (Strictly speaking, we show that $u \in \mathcal{H}$ has a version with these properties. But in the sequel we shall not distinguish between the class u and its unique continuous representative.) Let us sketch an elementary proof. Assume that $u \in \mathcal{H}_1$, that is, u has the expansion

$$u = \sum_{k=0}^{\infty} u_k\, l_k,$$

in $L^2(\mathbb{R}_+, \lambda)$ with coefficients u_k so that

$$\sum_{k=0}^{\infty} k^2\, |u_k|^2 < +\infty.$$

Consider $t \in \mathbb{R}_+$, and the following estimate

$$\sum_{k=0}^{\infty} |u_k\, l_k(t)| \leq |u_0\, l_0(t)| + \Big(\sum_{k=1}^{\infty} k^2\, |u_k|^2\Big)^{\frac{1}{2}} \Big(\sum_{k=1}^{\infty} k^{-2} |l_k(t)|^2\Big)^{\frac{1}{2}},$$

due to Schwarz' inequality. Now we use the bound

$$\sup_{t \in \mathbb{R}} |l_k(t)| \leq 1,$$

e.g., [AS 72], and conclude that the series $u(t) := \sum_{k=0}^{\infty} u_k\, l_k(t)$ converges absolutely for all $t \in \mathbb{R}_+$, uniformly in $t \in \mathbb{R}_+$. Hence u is a continuous function, vanishing at $+\infty$. Next suppose that $u \in \mathcal{H}_m$, $m \in \mathbb{N}$, $m \geq 2$. Let $L_k^{(m)}$, $k \in \mathbb{N}_0$, $m \geq 0$, denote the modified Laguerre polynomials, and note that $L_k^{(0)} = L_k$. Then we have the relation:

$$(L_k^{(m)})' = -L_{k-1}^{(m+1)}, \quad k \in \mathbb{N},$$

which can be found in [Sa 77]. On the other hand we have the estimate

$$\sup_{t \in \mathbb{R}_+} e^{-t/2} |L_k^{(m)}(t)| \leq \frac{(k+m)!}{k!\, m!}, \quad k, m \in \mathbb{N}_0,$$

([AS 72, 22.14.13]) which gives

$$\sup_{t\in\mathbb{R}_+} e^{-t/2}|L_k^{(m)}(t)| = O_m(k^m).$$

Hence, by a similar estimate as above, we see that $u \in \mathcal{H}_m$ entails that u is $m-1$ times continuously differentiable, and all its derivatives up to order $m-1$ vanish at $+\infty$. Finally, in order to see that if $u \in \mathcal{H}$, u and all its derivatives vanish faster than any inverse power, one can use the following relation (e.g., [Sa 77]):

$$tL_k^{(m)}(t) = -(k+1)L_{k+1}^{(m)}(t) + (2k+m+1)L_k^{(m)}(t) - (k+m)L_{k-1}^{(m)}(t), \quad t \in \mathbb{R}_+,$$

and proceed as before.

Example A.2 $T = \mathbb{R}$. This example is quite classical, cf., e.g., [Hi 80, RS 72, Si 71]. In this case we choose the Hermite functions $(h_k,\ k \in \mathbb{N}_0)$ as a CONS of $L^2(\mathbb{R},\lambda)$. Since they are eigenfunctions of the differential operator

$$H = \frac{1}{2}(-\frac{d^2}{dt^2} + t^2)$$

with eigenvalues $k+\frac{1}{2}$, we have that in this case the number operator equals $H-\frac{1}{2}$. In [Hi 80, RS 72, Si 71] it is proved that the nuclear space $\mathcal{H}$ in this case is the Schwartz space $\mathcal{S}(\mathbb{R})$ of rapidly decreasing C^∞ functions on the real line. A more elementary argument, similar to the one sketched in Example A.1, can be done if one uses the well-known recursion, differentiation, and multiplications formulae for the Hermite polynomials, together with the bound

$$\sup_{t\in\mathbb{R}} |h_k(t)| \leq O(k^{-1/12}),$$

which can be found, e.g., in [Sz 39].

Example A.3 $T = [0,1]$. We choose the CONS of $L^2([0,1],\lambda)$, given by $(1,\ s_k,\ c_k,\ k \in \mathbb{N})$, with $s_k(t) = \sqrt{2}\sin(2\pi k\,t)$, $c_k(t) = \sqrt{2}\cos(2\pi k\,t)$, $t \in [0,1]$. We make a minor modification of the above recipe by allowing the number operator in this case to have multiplicity 2 of its eigenvalues greater or equal than 1, and choose N equal to $(2^{5/4}\pi)^{-1}\sqrt{-\Delta}$, where Δ is the Laplacian with periodic boundary conditions. Thus the norms of the spaces $\mathcal{H}_n$ above are equivalent to usual Sobolev norms on $[0,1]$, the spaces $\mathcal{H}_n$ are Sobolev spaces of perodic functions. The nuclear space $\mathcal{H}$ is in this case (either by an application of the Sobolev embedding theorem or by a straightforward calculation) the space of infinitely differentiable periodic functions on $[0,1]$.

In all three case we shall denote the nuclear space $\mathcal{H}$ by $\mathcal{S}(T)$ and call it *Schwartz space over T*. We remark that it is easy to use the arguments for $u \in \mathcal{S}(T)$ to be C^∞ sketched above to establish that the seminorms $\sup_t |u(t)|$, $\sup_t |u'(t)|$, etc. are continuous on $\mathcal{S}(T)$.

References.

[AS 72] M. Abramowitz and I.A. Stegun: *Handbook of Mathematical Functions.* New York: Dover (1972).

[BD 95] F.E. Benth, Th. Deck and J. Potthoff: A white noise approach to a class of non linear heat equations; *Preprint* (1995).

[BP 95] F.E. Benth and J. Potthoff: On the Martingale Property for Generalized Stochastic Processes; *Preprint* (1995), to appear in *Stochastics and Stochastics Reports.*

[Co 53] J. Cook: The mathematics of second quantization; *Trans. American Math. Soc.* **74** (1953) 222–245.

[DP 95] Th. Deck and J. Potthoff: On a class of stochastic partial differential equations related to turbulent transport; *Preprint* (1995).

[DP 96] Th. Deck, J. Potthoff, G. Våge and H. Watanabe: Stability and viscosity limit for parabolic stochastic differential equations; *Preprint* (1996).

[Do 52] J.L. Doob: *Stochastic Processes.* New Tork: Wiley (1952).

[GV 64] I.M. Gel'fand and N.Ya. Vilenkin: *Generalized Functions, Vol. 4.* New York: Academic Press (1964).

[FH 94] M. de Faria, T. Hida, L. Streit and H. Watanabe: Intersectional local times as generalized white noise functionals; *Preprint* (1994), to appear in *Acta Appl. Math.*.

[Hi 80] T. Hida: *Brownian Motion.* Heidelberg: Springer (1980).

[HK 93] T. Hida, H.-H. Kuo, J. Potthoff and L. Streit: *White Noise - An Infinite Dimensional Calculus.* Dordrecht: Kluwer (1993).

[HO 96] H. Holden, B. Øksendal, J. Ubøe, T.S. Zhang: *Stochastic Partial Differential Equations - White Noise Functional Methods, Models and Applications.* Manuscript for a monograph to appear (1996).

[IK 89] N. Ikeda and S. Watanabe: *Stochastic Differential Equations and Diffusion Processes* (second edition). Tokyo and Amsterdam: Kondansha and North-Holland (1989).

[KL 94] Yu.G. Kondratiev, P. Leukert, J. Potthoff, L. Streit and W. Westerkamp: Generalized functionals in Gaussian spaces – The characterization theorem revisited; *Preprint* (1994), to appear in *J. Funct. Anal.*.

[KLS 94] Yu.G. Kondratiev, P. Leukert and L. Streit: Wick Calculus in Gaussian Analysis; *Preprint* (1994).

[LL 93] A. Laschek, P. Leukert, L. Streit and W. Westerkamp: Quantum mechanical propagators in terms of Hida distributions; *Reports Math. Phys.* **33** (1993) 221–232.

[Pi 69] A. Pietsch: *Nukleare Lokal Konvexe Räume*, Berlin: Akademie Verlag (1969).

[Po 94] J. Potthoff: White noise approach to parabolic stochastic partial differential equations, in: *Stochastic Analysis and Applications in Physics*, A.I. Cardoso, M. de Faria, J. Potthoff, R. Sénéor and L. Streit (ed.s). Dordrecht: Kluwer Academic Publishers (1994).

[PS 91] J. Potthoff and L. Streit: A characterization of Hida distributions; *J. Funct. Anal.* **101** (1991) 212–229.

[PT 95] J. Potthoff and M. Timpel: On a dual pair of smooth and generalized random variables; *Potential Analysis* **4** (1995) 637–654

[PV 96] J. Potthoff, G. Våge and H. Watanabe: Generalized solutions of linear parabolic stochastic partial differential equations; *Preprint* (1996)

[Øk 85] B. Øksendal: *Stochastic Differential Equations.* Heidelberg: Springer (1985).

[RS 72] M. Reed and B. Simon: *Methods of Modern Mathematical Physics, Vol. 1.* New York: Academic Press (1972).

[Sa 77] G. Sansone: *Orthogonal Functions.* Huntington: Robert Krieger Publ. Co. (1977).

[Si 71] B. Simon: Distributions and their Hermite expansions; *J. Math. Phys.* **12** (1971) 140–148.

[Sz 39] G. Szegö: *Orthogonal Polynomials*, Amer. Math. Soc. Coll. Publ., XXIII, New York, 1939.

[Tr 67] F. Trèves: *Topological Vector Spaces, Distributions and Kernels.* New York: Academic Press (1967).

[Wa 91] H. Watanabe: The local time of self-intersections of Brownian motions as generalized Brownian functionals; *Lett. Math. Phys.* **23** (1991) 1–9.