ALBERT BADRIKIAN

Transformation of gaussian measures


<http://www.numdam.org/item?id=AMBP_1996__S3__13_0>
Transformation of Gaussian measures
Introduction

We shall be, in our lecture, mainly concerned by some particular cases of the following problem :

Let \((X, \mathcal{F}, \mu)\) be a measure space and \(T : X \to X\) measurable. We denote by \(T(\mu)\) or \(\mu \circ T^{-1}\) the image of \(\mu\) by \(T\):

\[
T(\mu) (A) = \mu \circ T^{-1} (A) = \mu (T^{-1} A), \quad \forall A \in \mathcal{F}.
\]

When does \(T(\mu) \ll \mu\) and how to compute the density?

**Example 1** : Let \(X = \mathbb{R}^n\), \(\mu = \lambda_n\) (the Lebesgue measure) and \(T : X \to X\) a diffeomorphism. Then from the formula

\[
\int f(T(x))|\det T'(x)|dx = \int f(y)dy,
\]

we conclude that \(T(\lambda_n)\) is absolutely continuous with respect to \(\lambda_n\) and

\[
T(\lambda_n) (dy) = |\det T'(T^{-1}y)|^{-1}dy = |\det (T^{-1})'(y)|dy.
\]

**Example 2** : Let \((\Omega, \mathcal{F}, P)\) be the classical Wiener space, \(\Omega = C_0([0,1])\), \(\mathcal{F}\) the Borel \(\sigma\)-field, \(P\) the Wiener measure. Let \(u : [0, 1] \times \Omega \to \mathbb{R}\) be a measurable and adapted stochastic process such that \(\int_0^1 u_t^2(\omega)dt < \infty\) almost surely, and let \(T : \Omega \to \Omega\) be defined by :

\[
(T\omega)_t = \omega_t + \int_0^t u_s(\omega) \, ds.
\]

Girsanov has proven that

\[
T(P) \ll P.
\]

On the other hand, let

\[
\xi = \exp\left\{ -\int_0^1 u_t d\omega_t - \frac{1}{2} \int_0^1 u_t^2(\omega) \, dt \right\}
\]
then, if $\mathbb{E}(\xi) = 1$. $(T\omega)_t$ is a Brownian motion with respect to $(\Omega, \mathcal{F}, Q)$, where $\frac{dQ}{dP} = \xi$.

That is $Q \circ T^{-1} = P$.

(This fact was first proven by means of the Itô-calculus, but as we shall see, we can obtain this by analytic methods).

This has an application in Statistical Communication Theory:

Suppose we are receiving a signal corrupted by noise, and we wish to determine if there is indeed a signal or if we are just receiving noise.

If $x(t)$ is the received signal, $\xi(t)$ the noise and $s(t)$ the emitted signal:

$$x(t) = s(t) + \xi(t) \quad (A)$$

In general, we make an hypothesis on the noise: it is a **white noise**.

The "integrated" version of (A) is

$$X(t) = \int_0^t s(u) \, du + W_t = S_t + W_t \quad (A')$$

($W$ is the standard Wiener process, $X(t) = \int_0^t x(s) \, ds$ is the cumulative received signal).

Now we ask the question: is there a signal corrupted by noise, or is there just a noise ($s(t) = 0, \forall t$)?

The hypotheses are:

$$H_0 : X_t = W_t$$

$$H_1 : X_t = \int_0^t s(u) \, du + W_t.$$  

We consider the likelihood ratio

$$\frac{d\mu_w}{d\mu_x} = \exp \left( - \int_0^1 s(t) \, dW_t - \frac{1}{2} \int_0^1 s(t)^2 \, dt \right)$$

and we fix a threshold level for the type I-error:

- if $\frac{d\mu_w}{d\mu_x}(\omega) \leq \lambda$ we reject $(H_0)$
- if $\frac{d\mu_w}{d\mu_x}(\omega) \geq \lambda$ we accept $(H_0)$.  

Some general considerations and examples.

If \( P \ll Q, \) then \( T(P) \ll T(Q). \) (a)

Therefore, we do not lose very much if we suppose that \( P \) and \( Q \) are probabilities.

In the case where \( Q \) is a probability, we can have an expression of \( \frac{dT(P)}{dT(Q)} \) as conditional mathematical expectation.

Remark : From (a) we see that, if there exists a probability \( Q \) such that

\[ P \ll Q \quad \text{and} \quad T(Q) = P, \] then \( T(P) \ll P. \)

The converse is true if moreover \( \frac{dT(P)}{dP} > 0. \) (The measures are equivalent). Therefore the following properties are equivalent :

(i) : \( T(P) \sim P, \)

(ii) : \( \exists Q \sim P \) such that \( T(Q) = P. \)

Let us now consider an example which allows us to guess the situation in infinite dimensional space.

Let \( \Omega = \mathbb{R}^n \) and \( P = \gamma_n \) the canonical Gaussian measure with density :

\[ \frac{1}{(2\pi)^{n/2}} \exp \left( -\frac{\|x\|^2}{2} \right) \]

and let \( T : \mathbb{R}^n \to \mathbb{R}^n \) be a diffeomorphism, then

\[
\int_{\mathbb{R}^n} f(y) T(\gamma_n)(dy) = \int_{\mathbb{R}^n} f(Tx) \gamma_n(dx) \\
= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(Tx) \exp \left( -\frac{\|x\|^2}{2} \right) dx \\
= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(Tx) \exp \left( -\frac{1}{2} \|T^{-1}Tx\|^2 \right) dx \\
= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \frac{f(y)}{|\det T'(T^{-1}y)|} \exp \left( \frac{1}{2} \|y\|^2 - \frac{1}{2} \|T^{-1}y\|^2 \right) \exp \left( -\frac{1}{2} \|y\|^2 \right) dy.
\]
Therefore:

\[
\frac{dT(\gamma_n)}{d\gamma_n}(y) = \frac{1}{|\det T'(T^{-1}y)|} \exp \left( \frac{1}{2} \|y\|^2 - \frac{1}{2} \|T^{-1}y\|^2 \right)
\]

\[
= |\det (T^{-1})'(y)| \exp \left( \frac{1}{2} \|y\|^2 - \frac{1}{2} \|T^{-1}y\|^2 \right).
\]

Now if we write:

\[
T^{-1} = (I + S) \text{ with } S \text{ self adjoint},
\]

then:

\[
(T^{-1})'(y) = I + S'(y)
\]

and we obtain:

\[
\frac{d(I + S)^{-1}(\gamma_n)}{d\gamma_n}(y) = |\det (I + S'(y))| \exp \left\{ -(Sy, y)_{\mathbb{R}^n} - \frac{1}{2} \|S(y)\|^2 \right\}. \quad (B)
\]

This can be written as:

\[
|\det (I + S'(y))| \exp \left( -\text{Trace } S'(y) \right) \exp \left\{ -(Sy, y)_{\mathbb{R}^n} + \text{Trace } S'(y) - \frac{1}{2} \|S(y)\|^2 \right\},
\]

where $|\det (I + S'(y))| \exp \left( -\text{Trace } S'(y) \right)$ is the Carleman determinant.

**General remark:** If $T = Id(\Omega)$, it is clear that $TP = P$ for every $P$. The idea is to perturb the identity operator.

The problem is:

"what does the word *perturbation* mean?"
CHAPTER ONE

Anticipative stochastic integral

1 - Gaussian measures on Banach spaces

Let $E$ be a (real) separable Banach space, $E'$ its dual. A (Borelian) probability $\mu$ on $E$ is said to be "Gaussian centered" if for every $x' \in E'$, $\langle \cdot, x' \rangle_{E,E'} = x'(\cdot)$ is a Gaussian centered (real) variable (eventually degenerated) under $\mu$. All what we shall say is true whatever be the dimension of $E$ (finite or infinite).

If $x' \in E'$ we define $A : E' \to E$ by

$$Ax' = \int_E \langle x, x' \rangle_{E,E'} x \, d\mu(x),$$

(Bochner integral of a vector function). It is the barycenter of the measure $\langle \cdot, x' \rangle d\mu$.

$A$ is injective if $\text{Supp} \, \mu = E$.

Let $x \in A(E')$ so $x = A(u')$ and let $y \in A(E')$ so $y = A(v')$, we shall put on $A(E') \subset E$ the following scalar product:

$$(x, y) \mapsto (x, y)_\mu := \int_E \langle u', z \rangle \langle v', z \rangle \, d\mu(z)$$

(it does not depend on $u'$ and $v'$).

$A : E' \to E$ is continuous. (Since $\int_E \|x\|^2 \, d\mu(x) < \infty$ by Fernique's theorem).

Therefore, if $i$ denotes the canonical injection of $A(E')$ into $E$:

$$i : (A(E'), \|\cdot\|_\mu) \to (E, \|\cdot\|) \text{ is continuous.}$$

Actually:

$$\|Ax'\|_E = \sup_{\|y'\| \leq 1} \left| \int_E \langle x', x \rangle \langle y', x \rangle \, d\mu(x) \right|$$

$$\leq \sup_{\|y'\| \leq 1} \left( \int_E |\langle x', x \rangle|^2 \, d\mu(x) \right)^{\frac{1}{2}} \left( \int_E |\langle y', x \rangle|^2 \, d\mu(x) \right)^{\frac{1}{2}}$$

$$\leq \left( \int |\langle x', x \rangle|^2 \, d\mu(x) \right)^{\frac{1}{2}} \left( \int \|x\|^2 \, d\mu(x) g \right)^{\frac{1}{2}};$$
hence,

\[ \|Ax\|_E \leq C \|Ax\|_\mu \quad (\text{where } C \text{ is a constant}). \]

Let \( H_\mu \) be the completion of \( A(E') \) with respect to \( \|\cdot\|_\mu \). We have \( \hat{i} : H_\mu \to E \). I say that \( \hat{i} \) is injective (it will allow us to consider \( H_\mu \) as a subspace of \( E \)).

\( H_\mu \) is called the "reproducing kernel Hilbert space" (r.k.H.s.) of \( \mu \).

**Example 1 : Finite dimension**

\( E = \mathbb{R}^n \), \( \text{Supp } \mu = \mathbb{R}^n \):

\[
Ax = \int_E \langle x, x' \rangle x \, d\mu(x),
\]
or:

\[
\langle Ax', y' \rangle = \int_E \langle x, x' \rangle \langle x, y' \rangle \, d\mu(x).
\]

\( A \) is the covariance, it is invertible and

\[
(x, y)_\mu = \int_E \langle A^{-1}x, z \rangle \langle A^{-1}y, z \rangle \, d\mu(z) = \langle x, A^{-1}y \rangle,
\]

and therefore:

\( H_\mu = \mathbb{R}^n \).

**Example 2 : Brownian motion, Wiener space.**

Let \( T > 0 \) and \( \Omega = E = C([0,T], \mathbb{R}) \) be the space of real continuous functions on \([0,T]\).

There exists an unique centered measure \( \mu \) such that:

a) the support of \( \mu \) is \( C_0([0,T], \mathbb{R}) \), the space of the continuous functions vanishing at 0,

b) \( \forall t \in [0,T] : \omega \rightsquigarrow \omega_t \) has the variance \( t \),

c) let \( 0 \leq t_1 < t_2 < ... < t_n \leq T \), then: \( \omega_{t_1}, \omega_{t_2} - \omega_{t_1}, ..., \omega_{t_n} - \omega_{t_{n-1}} \) are independent.

We shall call \( \mu \) the Wiener measure on \( C([0,T], \mathbb{R}) \); then \( E' \) is the space of signed measures \( \nu \) on \([0,T]\). We shall also denote:

\[
\omega_t = B(t, \omega)
\]

and call \( t \rightsquigarrow B(t, \cdot) \) : the "Brownian motion" on \([0,T]\).
For $\nu_1, \nu_2 \in E'$ let:

$$B(\nu_1, \nu_2) = E \left[ \langle \nu_1, B \rangle \langle \nu_2, B \rangle \right]$$

$$= \int_{\Omega} \langle \nu_1, \omega \rangle \langle \nu_2, \omega \rangle \, d\mu(\omega).$$

We have for $\nu \in E'$

$$\langle \nu, B \rangle = \int_{[0,T]} B(t, \omega) \, d\nu(t) = \int_0^T \nu([t,T]) dB(t) \text{ (stochastic integral)}.$$ 

This fact can be verified as follows:

- it is true for $\nu = \delta_\omega$ (by definition of Brownian motion),
- by linearity this remains true if $\nu = \sum \alpha_i \delta_{\omega_i}$,
- then we apply a continuity argument.

Therefore

$$B(\nu_1, \nu_2) = \int_{[0,T]} \nu_1([t,T]) \nu_2([t,T]) \, dt.$$ 

Now let $\nu_1$ be a measure on $[0,T]$. We want to find the barycenter $m_{\nu_1}$ of the random variable on $\Omega : \omega \mapsto \langle \omega, \nu_1 \rangle$. ($m_{\nu_1}$ is an element of $\Omega = C([0,T])$). It is defined by

$$\nu \mapsto \langle m_{\nu_1}, \nu \rangle = \int_{[0,T]} m_{\nu_1}(t) \, d\nu(t) = B(\nu, \nu_1) = \int_{[0,T]} \nu_1([t,T]) \, \nu([t,T]) \, dt.$$ 

By the generalized integration by parts this is equal to:

$$\int_{[0,T]} J(\nu_1)(t) \, d\nu(t)$$

where

$$J(\nu_1)(t) = \int_0^t \nu_1([u,T]) \, du.$$ 

$J(\nu_1)$ is then absolutely continuous. On the space

$$\left\{ J(\nu_1), \ \nu_1 \in \mathcal{M}([0,T]) \right\}$$
we put the norm
\[ J(\nu_1) \sim \int_0^T \nu_1([t,T])^2 dt. \]
Its completion is the space of functions from \([0, T]\) into \(\mathbb{R}\) absolutely continuous, null at zero, whose derivative belongs to \(L^2([0, T], dt)\). It is the Cameron-Martin space.

Then the Cameron-Martin space is the reproducing kernel Hilbert space of the Wiener measure.

**Definition:** We call an "abstract Wiener space" a triple \((H, E, \mu)\) where:
- \(E\) is a separable Banach space, and \(\mu\) is a centered Gaussian measure on \(E\), whose topological support is \(E\).
- \(H\) is the r.k.h.s. associated to \(\mu\).

Actually \(H\) is dense in \(E\). This can be proven as follows:
Let \(i : H \rightarrow E\) be the canonical injection and \(i^* : E' \rightarrow H\) its transpose (we identify \(H\) to its dual).
Suppose that \(\langle x', i(x) \rangle_{E,E'} = 0\) for every \(x \in H\). This is equivalent in saying that:
\[
\langle x | i^*(x') \rangle_H = 0, \text{ for every } x \in H.
\]
Therefore
\[
i^*(x') = 0.
\]
This means that
\[
\|i^*(x')\|_H^2 = \int_E |\langle x', y \rangle_{E,E'}|^2 d\mu(y) = 0.
\]
Therefore
\[
\langle x', y \rangle = 0 \text{ almost surely},
\]
so this holds for all \(y \in E\) since \(\text{Supp } \mu = E\) and \(x'\) is continuous.
The transpose \(i^*\) from \(i : H \rightarrow E\) is therefore injective and dense and we have:
\[
E' \xrightarrow{i^*} H \xrightarrow{i} E \quad (\text{\(i\) is the canonical injection}).
\]
Every \(x' \in E',\) defines a Gaussian centered random variable on \(E',\) whose variance is
\[
\|i^*(x')\|_{H'}^2.
\]
Now we give without proof some properties of an abstract Wiener space:

1) $H$ is separable, as a Hilbert space. Therefore it is a borelian subset of $E$,
2) $\mu(H) = 0$ or 1 and $\mu(H) = 0 \Leftrightarrow \dim H = +\infty$ (therefore $\mu(H) = 1 \Leftrightarrow \dim H < \infty$),
3) $H$ is the intersection of the family of measurable subspaces of $E$, whose probability is equal to one,
4) the canonical injection $i : H \to E$ is compact,
5) for every Hilbert space $K$ and $u : E \to K$ linear continuous, $u \circ i : H \to K$ is Hilbert-Schmidt,
6) for every Hilbert space $K$ and $v : K \to E'$ linear continuous, $i^* \circ v : K \to H$ is Hilbert-Schmidt.

As a consequence of 5) and 6) we have:

7) let $K_1, K_2$ two Hilbert spaces ; $u_1 : K_1 \to E'$ and $u_2 : E \to K_2$ linear continuous then

$$K_1 \xrightarrow{u_1} E' \xrightarrow{i^*} H \xrightarrow{i} E \xrightarrow{u_2} K_2,$$

the composition $u_2 \circ i \circ i^* \circ u_1$ is nuclear (i.e. it possesses a trace).

2 - $L^2$-functionals on an abstract Wiener space

Let $(H, E, \mu)$ be an abstract Wiener space.

Suppose $(e_j)_{j \geq 1}$ is a sequence of elements of $E'$ such that $(i^*(e_j))_{j \geq 1}$ is an orthonormal basis in $H$. A function $f : E \to \mathbb{R}$ is said to be a polynomial in the $(e_j)$ if there exists an integer $n$ and a polynomial function $P$ on $\mathbb{R}^n$ such that

$$f(x) = P(e_1(x), ..., e_n(x)), \quad \forall x \in E.$$

We denote $\deg f := \deg P$ ($P$ is not defined uniquely but the degree of $f$ is independent of the choice of $P$).

We denote by $\mathcal{P}((e_j))$ the set of polynomials and by $\mathcal{P}^n((e_j))$ the set of polynomials of degree $\leq n$. It is easy to see that $\mathcal{P}((e_j))$ is contained in each $L^p(E, \mu)$ $0 \leq p < \infty$ (but clearly not in $L^\infty(E, \mu)$). Moreover, $\mathcal{P}((e_j))$ is dense in $L^p(E, \mu)$ for these $p$. Therefore, $\mathcal{P}((e_j))_{L^p}$ is independent of the chosen orthonormal family $(e_j)$. The same is true for each $\mathcal{P}^n((e_j))$. 
Example: If $n = 1$, $\mathcal{P}^1 ((e_j))$ is the family of affine continuous functions: an element of $\mathcal{P}^1 ((e_j))$ is a linear continuous function on $E$ plus a constant.

We have:

$$\overline{\mathcal{P}^1}_{L^2(E, \mu)} = H \oplus \mathbb{R} \quad \text{(see infra).}$$

We call $\overline{\mathcal{P}^n}_{L^2}$ the set of generalized polynomials of degree at most $n$; $\overline{\mathcal{P}^n}_{L^2}$ is a Hilbert space.

Let us now introduce the "Wiener chaos decomposition" (or "Wiener-Itô decomposition"). Let $C_0 = \overline{\mathcal{P}^0}_{L^2}$ the vector space of ($\mu$-equivalence classes of) constants. We define $C_n$ inductively as follows:

$C_n$ is the orthogonal complement in $\overline{\mathcal{P}^n}_{L^2}$ of $(C_0 \oplus \ldots \oplus C_{n-1})$.

(In other words $C_n$ is the set of generalized polynomials of degree $n$, orthogonal to all generalized polynomials of degree less than $n$).

It is clear that for every $n$:

$$\overline{\mathcal{P}^n}_{L^2} = C_0 \oplus \ldots \oplus C_n$$

and moreover

$$L^2(E, \mu) = \bigoplus_{n=0}^{\infty} C_n.$$

The $C_n$ are called the "nth chaos" (or "chaos of order $n$"). $C_1$ is isomorphic to $H$. We have a description of elements of $C_n$ in term of Hermite polynomials.

We recall that the Hermite polynomials in one variable are defined by:

$$H_n(t) = \exp\left\{\frac{t^2}{2}\right\} \frac{d^n}{dt^n} \left(\exp\left\{-\frac{t^2}{2}\right\}\right), \quad n \in \mathbb{N}.$$

Then they satisfy:

- $\sum_{n=0}^{\infty} \lambda^n H_n(t) = \exp\left\{-\frac{\lambda^2}{2} + \lambda t\right\}$

- $\frac{d}{dt} H_n(t) = H_{n-1}(t)$

- $\int_{\mathbb{R}} H_m(t) H_n(t) \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} dt = \frac{1}{n!} \delta_{nm}.$

Let $\alpha = (\alpha_1, \alpha_2, \ldots) \in \mathbb{N}^\mathbb{N}$ such that $|\alpha| := \sum_{i=1}^{\infty} \alpha_i < \infty$. We set $\alpha! := \prod_{i=1}^{\infty} \alpha_i!$. 
Now let $(e_n)_{n \geq 1}$ be a sequence of elements of $E'$ which is an orthonormal basis in $H$. If $\alpha \in \mathbb{N}^\mathbb{N}$ let

$$H_\alpha(x) := \prod_{i=1}^{\infty} H_{\alpha_i}(e_i(x))$$

(This product is well defined). Then:

$$\{\sqrt{\alpha!} H_\alpha(x), \alpha \in \mathbb{N}^\mathbb{N} \text{ and } |\alpha| < +\infty\}$$

is an orthonormal basis in $L^2(E, \mu)$ and:

$$\{\sqrt{\alpha!} H_\alpha(x), |\alpha| = n\}$$

is an orthonormal basis in $C_n$.

In the case of the Wiener measure associated to Brownian motion, we have the following characterization of $C_n$ in terms of multiple stochastic integrals:

$F : C([0, T], \mathbb{R}) \to \mathbb{R}$ belongs to $L^2(P)$ where $P$ is the Wiener measure if and only if for each $n$ there exists $f_n \in L^2(\Delta_n, dt)$ where $\Delta_n = \{t \in \mathbb{R}^n, 0 \leq t_1 \leq t_2 \leq \ldots \leq t_n \leq T\}$ such that

$$F = \sum_n \int_{\Delta_n} f_n(t_1, \ldots, t_n) \, dB(t_1) \ldots dB(t_n) = \sum_n F_n.$$

Here

$$F_0 = \mathbb{E}(F) \in C_0 \text{ and } F_n \in C_n.$$

3 - Measurable linear functionals
and linear measurable operators

Let $(H, E, \mu)$ be an abstract Wiener space. Without loss of generality, we shall identify $H$ as a subspace of $E$ (i.e., $i(x) = x$).

A linear mapping $f : E \to \mathbb{R}$ is said to be a "linear measurable functional" if there exists a sequence of linear continuous functionals on $E$, converging to $f$, $\mu$-almost surely.

If $x \in H$, it defines a linear measurable functional $\tilde{x}(\cdot)$. Actually, if $x_n$ is a sequence of elements of $E' \subset H$ such that $x_n \rightarrow x$ in $H$, then $x_n(\cdot)$ converges to the random variable $\tilde{x}$ defined by $x$, in $L^2(E, \mu)$. Therefore, there exists a subsequence converging almost surely to $\tilde{x}$. Moreover,

$$\int_E |\tilde{x}(x)|^2 \, d\mu(x) < \infty.$$

The converse is true, shown by the following proposition.
If \( h \in H \), the random variable \( \tilde{h} \) on \( E \) will be denoted by
\[
x \sim (x, h)_H.
\]

**Proposition:** Every linear measurable functional, \( f \), has a restriction to \( H \) which is continuous (for the Hilbertian topology). If we denote by \( f_0 \) this restriction we have
\[
\|f\|_{L^2(E, \mu)} = \|f_0\|_H.
\]
The converse is true.

**Proof:**

We have already noticed that the converse is true. Let \( (x_n) \subset E' \subset H \) such that
\[
x_n(x) \to f(x) \quad \forall x \in A, \text{ where } \mu(A) = 1.
\]
Take \( E \) the linear subspace generated by \( A \), we see that the above convergence holds for all \( x \in E \). Since \( \mu(E) = 1 \), then \( H \subset E \) and therefore
\[
x_n(x) \to f(x), \quad \forall x \in H.
\]
Therefore the restriction of \( f \) to \( H \) is uniquely defined.

Now,
\[
\int_E \exp \{i(x_n - x_m)(x)\} \mu(dx) = \exp \{-\frac{1}{2}\|x_n - x_m\|_H^2\} \to 1.
\]
Therefore, \( (x_n) \) converges in \( H \), and
\[
\int_E |x_n(x) - x_m(x)|^2 \mu(dx) = \|x_n - x_m\|_H^2 \to 0 \quad \text{as } m, n \to \infty.
\]
Therefore \( (x_n(\cdot)) \) converges in \( L^2(\mu) \). The limit is equal to \( f \) almost surely, as we can see immediately.

— Q.E.D. —

Now let \( K \) be a Hilbert space. As before we define a linear measurable function from \( E \) to \( K \), as the almost sure limit of a sequence of linear continuous functions from \( E \) to \( K \).

And, as before, if \( A \) is a linear measurable function from \( E \) into \( K \), its restriction to \( H \) is well defined and continuous from \( H \) to \( K \).

Let us remark that if \( A \) is a linear measurable function from \( E \) to \( K \), we can define its transpose as a linear function from \( K \) to \( H \) since, for every \( \varphi \in K \),
\[
\langleAx, \varphi\rangle_K = (A^*\varphi)(x), \quad \text{almost surely}
\]
\[
= (x, A^*\varphi)_H
\]
where \( A^* \) is the conjugate of the restriction of \( A \) to \( H \).
Now we can prove the following result:

**THEOREM**: If $A$ is a linear measurable function from $E$ to $K$ such that 
$$
\int \|Ax\|_K^2 \, d\mu(x) < \infty,
$$
then its restriction to $H$ is a Hilbert-Schmidt mapping $B$ from $H$ to $K$. Conversely if $B$ is a Hilbert-Schmidt mapping from $H$ to $K$, (we shall note $B \in \mathcal{L}^2(H, K)$ or $B \in \mathcal{L}_2(H, K)$), it possesses a linear measurable continuation on $E$, denoted by $A$.

Moreover, we have:

$$
\int_E \|Ax\|_K^2 \, d\mu(x) = \|B\|_{H-S}^2.
$$

**Proof**:

Let $(\varphi_j)$ be an orthonormal basis of $K$.

We have:

$$
\|Ax\|_K^2 = \sum_j (Ax, \varphi_j)_K^2 = \sum_j (x, A^\ast \varphi_j)_H^2.
$$

If we integrate term by term these equalities, we obtain:

$$
\int_E \|Ax\|_K^2 \, d\mu(x) = \sum_j \int_E (x, A^\ast \varphi_j)_H^2 \, d\mu(x)
$$

$$
= \sum_j \|A^\ast \varphi_j\|_H^2 = \sum_j \|B^\ast \varphi_j\|_H^2 = \|B^\ast\|_{H-S}^2.
$$

Conversely let $B \in \mathcal{L}_2(H, K)$. We have for $x \in H$:

$$
Bx = \sum_j (Bx, \varphi_j)_K \varphi_j
$$

$$
= \sum_j (x, B^\ast \varphi_j)_H \varphi_j.
$$

Now each term in the right-hand member possesses a linear measurable continuation to $E$, and the series converges in $\mathcal{L}_2(E, \mu, K)$.

We have then defined a linear measurable extension of $A$ to $E$.

--- Q.E.D. ---
4 - Derivatives of functionals on a Wiener space

Let \((E, H, \mu)\) be an abstract Wiener space and let \(K\) be another Hilbert space. Let \(f : E \to K\) be a function.

We say that \(f\) possesses a Fréchet derivative in the direction of \(H\), at the point \(x_0 \in E\) if there exists an element denoted \(f'(x_0)\) or \(Df(x_0)\) or \(\nabla f(x_0) \in \mathcal{L}(H, K)\) such that
\[
f(x_0 + h) - f(x_0) = f'(x_0) \cdot h + o(\|h\|_H), \quad \forall h \in H.
\]

Inductively we can define derivatives of all orders.

Example: Let \(x_0 \in H \setminus \text{i}^*(E')\) and let \(f\) be a measurable continuation of \(h \mapsto (x_0, h)_H\) to \(E\). \((f\) is not continuous).

Then \(f\) is derivable at every \(x\), and \(f'(x_0) \in H\).

This example shows that a discontinuous function may have Fréchet derivatives in the direction of \(H\).

Definition 1: Let us denote by \(C^{2,1}(E, K)\) the set of functions \(f : E \to K\) possessing the following properties:
- \(f\) possesses \(H\)-derivatives at every point \(x \in E\) and \(f'(x)\) is Hilbert-Schmidt for every \(x\),
- \(f\) and \(f'\) are continuous from \(H\) to \(K\) and to \(\mathcal{L}_2(H, K)\) respectively,
- \(\|f\|_{2,1}^2 := \int_E \left[ \|f(x)\|^2_K + \|f'(x)\|^2_{\mathcal{L}_2(H, K)} \right] \mu(dx) < \infty\).

Then \(C^{2,1}(E, K)\) is a vector space and \(\|\cdot\|_{2,1}\) is a Hilbertian norm on this space.

Definition 2: Let \(D^{2,1}(E, K)\) be the completion of \(C^{2,1}(E, K)\) for the preceding norm; \(D^{2,1}(E, K)\) is then a Hilbert space.

Clearly the elements of \(D^{2,1}(E, K)\) are \(\mu\)-equivalence classes of functions.

Convention: Often we shall write \(D^{2,1}(K)\) instead of \(D^{2,1}(E, K)\). In the same manner we shall write \(D^{2,1}\) instead of \(D^{2,1}(E, \mathbb{R})\) or \(D^{2,1}(\mathbb{R})\).

Now the map \(f \mapsto f'\) from \(C^{2,1}(E, K)\) into \(L^2(E, \mu, \mathcal{L}_2(H, K))\) is clearly continuous; therefore it possesses a unique continuous extension from \(D^{2,1}(H, K)\) into \(L^2(E, \mu, \mathcal{L}_2(H, K))\). This extension is again denoted by \(f'\), or \(Df\), or \(\nabla f\).

Example 1: Let \(f\) be a polynomial function on \(E\), with values in \(\mathbb{R}\):
\[
f(x) = P((f_1, x)_{E', E}, \ldots, (f_n, x)_{E', E}), \quad f_1, \ldots, f_n \in E'.
\]
Then \( f \in \mathcal{C}^{2,1} \) and

\[
f'(x) = \sum_{j=1}^{n} \frac{\partial P}{\partial y_j} \left( (f_1, x)_{E'}, ..., (f_n, x)_{E'} \right) i^*(f_j).
\]

The same result is true if \( P \) is a \( \mathcal{C}^1(\mathbb{R}^n) \)-function such that \( P \) and the partial derivatives \( \frac{\partial P}{\partial y_j} \) have polynomial growth.

In the same manner if \( f \) is defined (\( \mu \)-almost everywhere) as

\[
f(\cdot) = P(\tilde{h}_1(\cdot), \ldots, \tilde{h}_n(\cdot)), \quad h_j \in H
\]

with \( P \) a polynomial function (or a \( \mathcal{C}^1(\mathbb{R}^n) \)-function with polynomial growth together with its derivatives),

\[
\nabla f = \sum_{j=1}^{n} \frac{\partial P}{\partial y_j} (\tilde{h}_1(\cdot), \ldots, \tilde{h}_n(\cdot)) h_j.
\]

**Example 2**: Let \( \mu = \gamma_n \) the canonical Gaussian measure on \( \mathbb{R}^n \), \( \mathcal{D}^{2,1} \) is the Sobolev space \( W^{2,1}(\gamma_n) \) of the distributions in \( \mathbb{R}^n \) such that:

- \( f \in L^2(\mathbb{R}^n, \gamma_n) \),
- the distribution derivatives of \( f \) belong to \( L^2(\mathbb{R}^n, \gamma_n) \). The norm of \( \mathcal{D}^{2,1} \) is the usual Hilbertian norm:

\[
f \sim \left( \int_{\mathbb{R}^n} \left[ |f(x)|^2 + \sum_{j=1}^{n} \left| \frac{\partial f}{\partial x_j} (x) \right|^2 \right] d\gamma_n(x) \right)^{\frac{1}{2}}.
\]

**Example 3**: If \( f \) is a polynomial function with values in \( K \) :

\[
f(x) = \sum_{j=1}^{m} P_j((f_1, x)_{E'}, ..., (f_n, x)_{E'}) k_j
\]

\((k_j \in K, \quad f_1, ..., f_n \in E')\).

\[
\nabla f(x) = \sum_j \sum_i \frac{\partial P_j}{\partial y_i} ((f_1, x)_{E'}, ..., (f_n, x)_{E'}) f_i \otimes k_j.
\]

(Analogous assertion for generalized polynomials, or “moderate” regular functions \( P_j \)).
Example 4: Characterization of the elements of $\mathbb{D}^{2,1}$ in the case of the Wiener measure.

If $E = C_0([0,T], \mathbb{R})$ and $\mu$ is the Wiener measure, we have seen that an element of $L^2(\mu)$ can be written as a series

$$F = \sum_{n=0}^{\infty} \sqrt{n!} \int_{\Delta_n} f_n(t_1, t_2, ..., t_n) \, dB_{t_1}, ..., dB_{t_n}$$

with

$$\sum_{n=0}^{\infty} n! \|f_n\|_{L^2(\Delta_n)}^2 < \infty.$$ 

Then $F$ belongs to $\mathbb{D}^{2,1}$ if and only if

$$\sum_{n=1}^{\infty} nn! \|f_n\|_{L^2(\Delta_n)}^2 < \infty$$

and in this case

$$\nabla F = \sum_{n=1}^{\infty} n J(I_{n-1}(f^t_n))$$

where $f^t_n$ is the function defined on $\Delta^t_{n-1} = \{0 \leq t_1 < t_2 < ... < t_{n-1} < t\}$ by

$$f^t_n(t_1, t_2, ..., t_{n-1}) = f_n^{SYM}(t_1, t_2, ..., t_{n-1}, t),$$

$f_n^{SYM}$ being the symetrisation of $f_n$.

The formula needs an explanation:

In the right member

$$(t, \omega) \rightsquigarrow I_{n-1}(f^t_n)(\omega) = g(t, \omega)$$

belongs to

$$L^2([0,T] \times \Omega, dt \otimes dP),$$

therefore for almost $\omega$,

$$t \rightsquigarrow g(t, \omega)$$

is a $L^2([0,T], dt)$ function.

$J(I_{n-1}(f^t_n)) (\omega)$ is the indefinite integral null at zero of $I_{n-1}(f^t_n)(\omega)$:

$$J(I_{n-1}(f^t_n)) = \int_0^t I_{n-1}(f^s_n) \, ds.$$ 

Therefore $\nabla F(\omega)$ is an element of the Cameron-Martin space.
We now give several useful properties of $\mathbb{D}^{2,1}(E, K)$:

- The set of polynomial functions on $E$, with values in $K$ is dense in $\mathbb{D}^{2,1}(K)$.
- Therefore the algebraic sum of chaos $\sum C_n$ is dense in $\mathbb{D}^{2,1}$.
- The set of smooth functions on $E$ is dense in $C^{2,1}$ (a function is said to be "smooth" if it has the form:

\[
x \sim f((f_1, x)_{E'}, E, ..., (f_n, x)_{E'})
\]

with $f$ belonging to $C_b^\infty(\mathbb{R}^n)$; $f$ and its derivatives are bounded).
- Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a function in $C^1_b(\mathbb{R}^n)$ and let $F^1, ..., F^n \in \mathbb{D}^{2,1}$. Then $\varphi(F^1, ..., F^n)$ is in $\mathbb{D}^{2,1}$ and

\[
\nabla(\varphi(F^1, ..., F^n)) = \sum_{i=1}^n \frac{\partial \varphi}{\partial y_i}(F^1, ..., F^n) \nabla F^i.
\]

This result is false if the above hypothesis is not satisfied. For instance on $\mathbb{R}$,$$
 f = g = e^x \in \mathbb{D}^{2,1}, \text{ but } f \circ g \notin L^2(\mathbb{R}^n, \gamma_n).
$$

**Remark**: The operator $\nabla$, called the "stochastic" gradient, or "stochastic" derivative, is very close to the ordinary gradient as we can see. The usual gradient at the point $x_0$ is an element of $E'$ (if the function takes its values in $\mathbb{R}$). The stochastic gradient is the composite of the ordinary gradient by the application $i^*$ from $E'$ to $H$.

In an analogous manner if $f : E \to K$ has an ordinary gradient, this gradient is a linear mapping of $E$ into $K$; $f' : E \to K$.

The transpose $^t f'$ is a linear continuous mapping from $K$ into $E'$. Then the stochastic gradient is equal to $i^*(^t f') \in \mathcal{L}(K, H)$.

In his lectures at the EIPES in 1989, D. Nualart, in the case of usual Wiener space defined the stochastic derivative of the functional of the form:

\[
F = f(W_{t_1}, ..., W_{t_n}), \quad f \in C^\infty_b(\mathbb{R}^n) \quad (\text{or } f \text{ polynomial})
\]

by

\[
DF = \sum_{j=1}^n \frac{\partial F}{\partial y_j}(W_{t_1}, ..., W_{t_n}) 1_{[0,t_j]}.
\]
This definition is actually equivalent to ours, up to the notations.

Actually, let \( h_j(t) = \int_0^t 1_{[0,t_j]} (s) \, ds \), \( h_j \) belongs to the Cameron-Martin space and

\[
W_{t_j} = \tilde{h}_j = \langle h_j, \bullet \rangle_{\mathcal{C}-M}
\]

The stochastic derivative of \( F \) in our notations is therefore

\[
\sum_{j=1}^n \frac{\partial F}{\partial y_j} (\tilde{h}_1, \ldots, \tilde{h}_n) h_j.
\]

There are actually equivalent since the Cameron-Martin space is isomorphic as Hilbert space to \( L^2([0,T], dt) \). We shall have to consider \( \nabla \) as an operator (densely defined) from \( L^2(E, \mu, K) \) into \( L^2(E, \mu, \mathcal{L}_2(H,K)) \). It is a closed operator, naturally not continuous.

5 - Anticipative stochastic integral

**Definition:** The transpose of the operator \( \nabla \) is called the "Skorokhod integral", or the "divergence operator".

The definition needs an explanation: on \( L^2(E, \mu, K) \) (\( K \) : Hilbert space) we have defined the scalar product

\[
(f, g) \sim \int_E \langle f(x), g(x) \rangle_K \, d\mu(x)
\]

and on \( L^2(E, \mu, \mathcal{L}_2(H,K)) \) we have the pairing:

\[
(F, G) \sim \int_E \langle F(x), G(x) \rangle_{\mathcal{L}_2(H,K)} \, d\mu(x)
\]

\[
\quad = \int_E \text{Trace} \left( G^*(x) \circ F(x) \right) \, d\mu(x).
\]

Then \( G \in L^2(E, \mu, \mathcal{L}_2(H,K)) \) belongs to \( \text{dom}(\delta) \) if and only if the linear form on \( \mathbb{D}^{2,1}(K) \):

\[
F \sim \int_E \langle DF, G \rangle_{\mathcal{L}_2(H,K)}(x) \, d\mu(x)
\]

is continuous for the topology induced by \( L^2(E, \mu, K) \).

We denote \( \delta \) the Skorokhod integral and we have by definition, for every \( F \in \mathbb{D}^{2,1}(K) \),

\[
\int_E \langle F, \delta G \rangle_K \, d\mu = \int_E \langle \nabla F, G \rangle_{\mathcal{L}_2(H,K)} \, d\mu \quad \text{if } \delta(G) \text{ is defined}.
\]
Example 1: Let $a \in H$, and $\varphi \in D^{2,1}(K)$. Then $G := \varphi \otimes a$ is Skorokhod integrable and

$$\delta(a \otimes \varphi) = \tilde{a}(\cdot) \varphi - \langle \nabla \varphi, a \rangle.$$ 

In particular, if $G : E \to H$ is such that $G(x) = a, \forall x :$

$$\delta G = \tilde{a}(\cdot).$$

Example 2: $E = \mathbb{R}^n$, $\mu = \gamma_n$, $G : \mathbb{R}^n \to \mathbb{R}^n$.

Then

$$\delta G(x) = \langle x, G(x) \rangle_{\mathbb{R}^n} - \sum_{j=1}^n \frac{\partial G_j}{\partial x_j}(x)$$

$$= \langle x, G \rangle - \text{div} \ G(x).$$

This formula can be written in another manner:

$$\delta G = \langle *, G \rangle - \text{Trace} (\nabla G).$$

Example 3: If $G \in D^{2,1}(E, \mu, L^2(H, K))$, then it is $\delta$-integrable, and $\delta$ is continuous from $D^{2,1}(L^2(H, K))$ in $L^2(E, \mu, K)$.

Example 4: Let $F \in L^2(E, \mu, H)$ such that for every $h \in H : \nabla \langle (F, h) \rangle_H$ exists. Then for every linear continuous operator $A : H \to H$ with finite rank, $A(F)$ is Skorokhod integrable.

More precisely, if $A = \sum_{j=1}^n \langle *, a_j \rangle_H e_j$ (with $a_j$ and $e_j$ in $H$, $(e_j)$ being orthonormal) we have:

$$A(F) = \sum_{j=1}^n \langle F, a_j \rangle_H e_j$$

$$\delta(A(F)) = \sum_{j=1}^n \left[ \langle F, a_j \rangle \overline{e_j} - \nabla e_j \langle (F, a_j) \rangle \right].$$

(see example 1).
This can be written in another manner:

Let $A^*$ be the transpose of $A$: 

$$A^* = \sum_{j=1}^{n} \langle \cdot, e_j \rangle_H a_j$$

and let $\tilde{A}^*$ defined as:

$$\tilde{A}^* = \sum_{j=1}^{n} a_j \tilde{e}_j .$$

Then

$$\delta(A(F)) = \langle F, \tilde{A}^* \rangle_H - \sum_{j=1}^{n} \nabla_{e_j} \left( \langle F, a_j \rangle \right).$$

If we now suppose that $DF$ exists, we have:

$$\sum_{j=1}^{n} \nabla_{e_j} \left( \langle F, a_j \rangle \right) = \text{Trace} \left( A \circ DF \right).$$

Therefore, we have:

$$\delta(A(F)) = \langle F(\cdot), \tilde{A}^*(\cdot) \rangle_H - \text{Trace} \left( A \circ DF \right).$$

**Example 5:** The Skorokhod integral coincides with the ordinary Itô-Integral for adapted processes (see the above mentioned Nualart's Lecture Notes for a precise statement of this fact).

**Now we give some properties of the Skorokhod integral:**

a) Let $A: K \to K'$ be a linear continuous operator ($K$ and $K'$ Hilbert spaces) and let $F \in L^2(E, \mu, \mathcal{L}_2(H, K))$. If $F$ is Skorokhod-integrable so is $A \circ F$ and we have

$$\delta(A \circ F) = A(\delta F).$$

As a consequence we have:

- Let $F \in L^2(E, \mu, \mathcal{L}_2(H, K))$ such that $\delta(F)$ exists, then for every $k$ in $K$ we have $\langle \delta(F), k \rangle = \delta(F^*(k)).$
Let $F \in L^2\left(E, \mu, \mathcal{L}_2(H, \mathcal{L}_2(H, K))\right)$ such that $\delta(F)$ exists, then

for every $h \in H$, $\delta \left(\dot{F}(\cdot)(h)\right)$ exists

and

$$\delta (F)(h) = \delta \left(\dot{F}(\cdot)(h)\right).$$

If $F \in L^2(H, \mathcal{L}_2(H, K))$, $\dot{F}$ denotes the operator of $L^2\left(H, \mathcal{L}_2(H, K)\right)$ such that :

$$\dot{F}(h)(h') = F(h')(h), \quad h, h' \in H.$$ 

b) Let $\varphi \in \mathcal{D}^{2,1}$, $F \in L^2(E, \mu, H)$ such that $F$ is Skorokhod integrable. Suppose that $\varphi F \in L^2(E, \mu, H)$ and that $\delta(F)\varphi - \langle F, D\varphi \rangle_H$ belongs to $L^2(E, \mu)$, then $\varphi F$ is Skorokhod integrable and

$$\delta(\varphi F) = \delta(F)\varphi - \langle F, D\varphi \rangle_H.$$ 

c) Let $A_n : H \rightarrow H$ a sequence of linear continuous operators such that $A_n \rightarrow I_{d_H}$ in the simple convergence.

Let $F \in \mathcal{D}^{2,1}(\mathcal{L}_2(H, K))$, then $\delta(F \circ A_n) \rightarrow \delta(F)$ in $L^2(E, \mu, K)$. In particular, if $(e_n)$ is an orthonormal basis of $H$, the sequence

$$\left(\sum_{i=1}^{n} \tilde{e}_i F(e_i) - \nabla_{e_i} F(e_i)\right)$$

converges to $\delta(F)$.

d) Let $F, G$ in $\mathcal{D}^{2,1}(H)$ we have :

$$\mathbb{E}(\delta(F)\delta(G)) = \mathbb{E}\{\langle F, G \rangle_H\} + \mathbb{E}\{\langle DF, (DG)^\ast \rangle_{\mathcal{L}_2(H, H)}\}$$

$$= \mathbb{E}\{\langle F, G \rangle_H\} + \mathbb{E}\{\text{Trace } DG(\cdot) \circ DF(\cdot)\}.$$ 

More generally, if $F$ and $G$ belong to $\mathcal{D}^{2,1}(\mathcal{L}_2(H, K))$ we have :

$$\mathbb{E}\{\langle \delta F, \delta G \rangle_K\} = \mathbb{E}\{\langle F, G \rangle_{L^2(H, K)}\} + \mathbb{E}\{\langle DF, \dot{DG} \rangle_{\mathcal{L}_2\left(H, L^2(H, K)\right)}\}.$$ 

e) The operator $\delta$, as an operator densely defined from $L^2\left(E, \mu, \mathcal{L}_2(H, K)\right)$ into $L^2(\Omega, \mu, K)$ is closed.
We now briefly introduce the Ogawa integral.

Let \( P : H \rightarrow H \) be an orthogonal projector with finite rank: \( P(h) = \sum_{j=1}^{n} \langle h, e_j \rangle_H e_j \).

We denote \( \tilde{P} \) the random variable with values in \( H \):

\[
\tilde{P}(\cdot) := \sum_{j=1}^{n} \tilde{e}_j(\cdot) e_j.
\]

Now let \( F \in L^0(E, \mu, H) \) be a random variable with values in \( H \). We shall say that \( F \) is "**Ogawa integrable**", if there exists \( G \in L^0(E, \mu) \) such that, for every increasing sequence \((P_n)\) of orthogonal projectors converging simply to \( \text{Id}_H \), the sequence of real random variables \( (\langle F, \tilde{P}_n \rangle_H)_n \) converges to \( G \) in probability.

We shall denote by \( \delta(F) \) the Ogawa integral \( G \) of \( F \).

If \( F \in L^2(E, \mu, H) \) is such that, for every \( a \in H \):

\[
\langle F, a \rangle_H \tilde{a}(\cdot) \text{ belongs to } L^2(E, \mu),
\]

we shall say that \( F \) is "**2-Ogawa integrable**" when there exists \( G \in L^2(E, \mu) \) such that

\[
\langle F, \tilde{P}_n \rangle_H \longrightarrow G \text{ in quadratic mean.}
\]

(The \( P_n \) being as above).

**Example:** \((E, \mu) = (\mathbb{R}^n, \gamma_n)\). The Ogawa integral is equal to \( \langle \cdot, F(\cdot) \rangle_{\mathbb{R}^n} \).

In this case, we have:

\[
\delta(F) = \delta(F) + \text{Trace } (\nabla F).
\]

**Remark:** There exists elements of \( \mathcal{D}^{2,1}(H) \) which do not possess an Ogawa integral (Rosinski).

For instance, in the case of the Brownian motion, the function: \( \omega \rightsquigarrow J(B(T - \cdot)(\omega)) \) where \( J \) denotes the indefinite integral null at zero, belongs to \( \mathcal{D}^{2,1}(H) \) but is not Ogawa integrable.
Next we give a necessary and sufficient condition for Ogawa integrability:

Let \( F \in D^{2,1}(H) \); \( F \) is Ogawa integrable if and only if, for almost every \( x \):

\[
DF \in L_1(H, H) \quad (\iff DF \text{ is nuclear})
\]

and we have:

\[
\overset{\circ}{\delta}(F) = \delta(F) + \text{Trace}(DF).
\]

Sketch of the proof:

Suppose \( P : H \to H \) is an orthogonal projector with finite rank. We know that:

\[
\delta(PF) = \langle F, \tilde{P} \rangle - \text{Trace}(D(PF)).
\]

Let \( P_n \uparrow \text{Id} \). We know that

\[
\delta(P_nF) \to \delta(F).
\]

It is trivial that:

\[
\langle F, \tilde{P}_n \rangle \to \overset{\circ}{\delta}(F)
\]

(if \( \overset{\circ}{\delta}(F) \) exists) and

\[
\text{Trace}(D(P_nF)) \to \text{Trace}(DF)
\]

— Q.E.D.—

6 - Extensions and remarks - Localization

Now we shall consider the case where \((E, H, \mu)\) is the Wiener space. If \( F \in D^{2,1} \), then \( \nabla F \) is a random variable with values in the Cameron-Martin space. Therefore, if \( t \in [0, T] \) we can speak of the value of \( \nabla F(\omega) \) at \( t \), denoted \( \nabla_t F(\omega) \). Analogously, time derivative of \( \nabla F(\omega) \) at time \( t \) (defined for almost every \( t \)) makes sense. We shall denote it: \( \dot{\nabla}_t F(\omega) \). We have the equality:

\[
||\nabla F(\cdot)||^2_{L^2(H)} = \mathbb{E}(\int_0^t |\dot{\nabla}_t F(\omega)|^2 dt).
\]

Lemma 1: Let \( F \in D^{2,1} \). Then \( 1\{F=0\} \dot{\nabla}_t F = 0 \) almost everywhere on \([0, T] \times \Omega \).

For the proof see Nualart-Pardoux.

This results in a localization theorem: if \( F \) is null (almost everywhere) on a set, so is its derivative. The derivation is a "local operator".
Definition 1: A random variable $F$ will be said to belong to $\mathbb{D}^{2,1}_{\text{loc}}$ if there exist
- a sequence of measurable sets of $E$, $E_k \uparrow E$
and
- a sequence $(F_k) \subset \mathbb{D}^{2,1}$ such that $F|_{E_k} = F_k|_{E_k}$ a.s. \( \forall k \in \mathbb{N} \).

Thanks to the preceding lemma we can define the derivation operator for an element of $\mathbb{D}^{2,1}_{\text{loc}}$.

Definition 2: Let $F \in \mathbb{D}^{2,1}_{\text{loc}}$ localized by the sequence $(E_k, F_k)$. $DF$ is the unique equivalence class of $dt \times dP$ a.e. equal processes such that
$$DF|_{E_k} = DF_k|_{E_k}, \quad \text{for all } k \in \mathbb{N}.$$ This generalized derivative has the usual properties of composition:

- Let $\varphi : \mathbb{R}^m \to \mathbb{R}$ of the class $C^1$; suppose $F = (F^1, ..., F^m)$ is a random vector whose components belong to $\mathbb{D}^{2,1}_{\text{loc}}$; then
  $$\varphi(F) \in \mathbb{D}^{2,1}_{\text{loc}}$$

and
  $$\nabla \varphi(F) = \sum_{i=1}^{m} \frac{\partial \varphi}{\partial x_i}(F) \cdot DF^i.$$ 

In the same manner we define $(\text{Dom } \delta)_{\text{loc}}$ as follows:

$F : E \to H$ belongs to $(\text{Dom } \delta)_{\text{loc}}$ if there exists a sequence $E_k \uparrow E$, and a sequence $F_k : E \to H$ such that $F_k \in (\text{Dom } \delta)$ for every $k$, such that

- $F = F_k$ on $E_k$
- $\delta(F_k) = \delta(F_\ell)|_{F_k}$ a.s. if $k < \ell$;

we shall say that $F$ is “localized” by $(E_k, F_k)$.

For sufficiently reasonable integrands on $(\text{Dom } \delta)$ Nualart-Pardoux have shown that $\delta$ is local.

Definition 3: Let $F \in (\text{Dom } \delta)_{\text{loc}}$ localized by $(E_k, F_k)$, $\delta(F)$ is defined as the unique equivalence class on random variables on $E$ such that
$$\delta(F)|_{E_k} = \delta(F_k)|_{E_k}, \quad \text{for all } k \in \mathbb{N}.$$ (Note that $\delta(F)$ may depend on the localizing sequence.)
We shall need another notion of stochastic derivatives and Skorokhod integrals for some functions not necessarily belonging to $\mathbb{D}^{2,1}$, nor Skorokhod integrable, introduced by Buckdahn:

Let $T : E \to E$ be a measurable mapping of the form:

$$x \mapsto x + Fx$$

where $F \in \mathbb{D}^{2,1}(H)$.

Let $\xi \in \mathbb{D}^{2,1}$ and suppose that for every sequence of smooth random variables $(\xi_n) \in \mathbb{D}^{2,1}$ converging to $\xi$ in $\mathbb{D}^{2,1}$, the following limit exists and is independent of the approximating sequence chosen:

$$\lim_{n \to \infty} \nabla (\xi_n \circ T)$$

where the limit is taken in probability.

Let us remark that $\xi_n \circ T$ belongs to $\mathbb{D}^{2,1}$ since the $\xi_n$ are smooth.

The common limit of the above sequences is denoted by $\tilde{\nabla} (\xi \circ T)$.

Lemma 2: Suppose that $T(\mu) \ll \mu$, then the limit exists and we have, $\mu$-almost surely:

$$\tilde{\nabla} (\xi \circ T) = (I_H + (\nabla F)^*) ((\nabla \xi) \circ T) = (I_H + \nabla F)^* ((\nabla \xi) \circ T)$$

(where $(\cdot)^*$ denotes the adjoint of the bounded operator).

Moreover, if $\xi \circ T \in \mathbb{D}^{2,1}$: $\tilde{\nabla} (\xi \circ T) = \nabla (\xi \circ T)$.

Proof:

We have, since the $(\xi_n)$ are smooth:

$$\nabla (\xi_n \circ T) = (I_H + \nabla F)^* ((\nabla \xi_n) \circ T).$$

Moreover, $\nabla \xi_n$ converges in probability, and since $T(\mu)$ is absolutely continuous with respect to $\mu$, $(\nabla \xi_n) \circ T$ converges in probability, so does $\nabla (\xi_n \circ T)$.

It now remains to prove that the limit does not depend upon the approximating sequence $(\xi_n)$.

Let $\xi_n \to \xi$ and $\eta_n \to \xi$ in $\mathbb{D}^{2,1}$. Since the operator $\nabla$ is closed we have:

$$\lim_n \nabla (\xi_n \circ T) = \lim_n \nabla (\eta_n \circ T).$$

Therefore, $\tilde{\nabla}$ is well defined by what precedes. It is obvious that:

$$\tilde{\nabla} = \nabla \quad \text{if} \quad \xi \circ T \in \mathbb{D}^{2,1}.$$

By duality, we can define a generalized Skorokhod integral of $\xi \circ T$, for $\xi \in D^{2,1}(H)$:

— Lemma 2 is proven.—
**Definition**: Let \((e_i)_{i \in \mathbb{N}}\) be a fixed orthonormal basis of \(H\). We define
\[
\tilde{\delta}(\xi \circ T) := \sum_i (\langle \xi \circ T, e_i \rangle_H \tilde{e}_i - \nabla e_i (\langle \xi \circ T, e_i \rangle_H),
\]
if the limit of the right member is taken in probability.
(\(\nabla e_i\) denotes the generalized derivative in the \(e_i\)-direction introduced just above).

**Lemma 3**: Suppose \(T = I + F\) as above is such that \(T(\mu) \ll \mu\). Then \(\tilde{\delta}(\xi \circ T)\) exists and satisfies the following identity:
\[
(\delta(\xi)) \circ T = \tilde{\delta}(\xi \circ T) + \langle \xi \circ T, F \rangle_H + \text{Trace} \left( (\nabla\xi) \circ T \cdot \nabla F \right) \mu\text{-almost surely}.
\]

**Proof**:

Let \(\xi^N = \sum_{i=1}^N \langle \xi, e_i \rangle_H e_i\), then
\[
\tilde{\delta}(\xi^N \circ T) = \sum_{i=1}^N (\langle \xi \circ T, e_i \rangle_H \tilde{e}_i - \sum_{i=1}^N \nabla e_i (\langle \xi \circ T, e_i \rangle_H).
\]
But
\[
\tilde{e}_i \circ T = \tilde{e}_i + \langle F, e_i \rangle_H,
\]
therefore:
\[
\delta(\xi^N \circ T) = \sum_{i=1}^N \left\{ (\langle \xi, e_i \rangle_H \tilde{e}_i \circ T - \langle F, e_i \rangle_H) - (\langle I_H + \nabla F \rangle^* (\nabla (\langle \xi, e_i \rangle_H)) \circ T, e_i \rangle_H \right. \]
(by the preceding lemma)
\[
= \sum_{i=1}^N \left\{ (\langle \xi, e_i \rangle_H \circ T - \langle \xi, e_i \rangle_H (F, e_i)_H) - (\langle I_H + \nabla F \rangle^* (\nabla (\langle \xi, e_i \rangle_H)) \circ T, e_i \rangle_H \right. \]
\[
= \sum_{i=1}^N \left[ (\langle \xi, e_i \rangle_H \tilde{e}_i - (\nabla e_i \xi, e_i)_{H} \circ T - (\xi^N \circ T, F)_H - \text{Trace} \left( (\nabla F^*, (\nabla \xi^N) \circ T \right).\right.
\]

Now \(\xi^N \to \xi\) in \(D^{2,1}(H)\); then the right member of this last equality converges in \(L^0(E, \mu)\). Hence the sum is convergent in \(L^0(E, \mu)\) and
\[
\sum_{i=1}^\infty (\langle \xi \circ T, e_i \rangle_H \tilde{e}_i - \nabla e_i (\langle \xi \circ T, e_i \rangle_H) \text{ is convergent in } L^0(E, \mu).
\]

--- *Lemma 3 is proven.* ---
CHAPTER TWO

Transformation of a Gaussian measure

Given an abstract Wiener space \((H, E, \mu)\) and \(T : E \rightarrow E\) of the form:

\[ Tx = x + F(x), \quad F : E \rightarrow H. \]

We shall examine when \(T(\mu) \ll \mu\). We shall consider the following cases:

- \(F\) is linear continuous from \(E\) into \(H\),
- \(F\) is regular (i.e., possesses stochastic derivatives).

We shall give some expressions for the Radon-Nikodym density \(\frac{dT(\mu)}{d\mu}\).

In the following chapter we shall study a family of flows: \(T_t = I + F_t\) where \(F_t : E \rightarrow H, \quad (t \in [0, 1])\) and shall study the work of Cruzeiro, Buckdahn and Ustunel-Zakai on this subject. We shall only give the statements of the results and from time to time sketch of the proofs.

1 - Preliminary results on equivalence and orthogonality of product measures

Let \((E_k, B_k)_{k \in \mathbb{N}^*}\) be a sequence of measurable spaces and for every \(k\), let \(\mu_k\) and \(\nu_k\) be two probabilities on \((E_k, B_k)\) such that \(\mu_k \ll \nu_k\). Let us set \(\rho_k = \frac{d\mu_k}{d\nu_k}\).

Let us consider the product measures:

\[ \mu = \prod_{k=1}^{\infty} \mu_k \]

and

\[ \nu = \prod_{k=1}^{\infty} \nu_k \]

and let

\[ \alpha_k = \int_{E_k} \sqrt{\rho_k(x_k)} \; \nu_k (dx_k). \]

These notations having been fixed we have the following result of Kakutani:
THEOREM 1: We have the dichotomy:

\[ \mu \ll \nu \text{ or } \mu \perp \nu. \]

a) \( \mu \ll \nu \iff \prod \alpha_k \) converges; and in this case the density is equal to \( \rho(x) = \prod_{1}^{\infty} \rho_k(x_k) \) (convergence in mean).

b) \( \mu \perp \nu \iff \prod \alpha_n \) diverges to zero. (We cannot have divergence to infinity since \( \alpha_k^2 \leq 1 \)).

Applications: \( E_k = \mathbb{R} \) for every \( k \)

\[ \nu_k(dx_k) = \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left\{ -\frac{(x_k - \gamma_k)^2}{2\sigma_k^2} \right\} \, dx_k \]

\[ \mu_k(dx_k) = \frac{1}{\lambda_k \sqrt{2\pi}} \exp\left\{ -\frac{(x_k - \beta_k)^2}{2\lambda_k^2} \right\} \, dx_k. \]

Then

\[ \rho_k(x_k) = \frac{\sigma_k}{\lambda_k} \exp\left\{ -\frac{1}{2\sigma_k^2 \lambda_k^2} \left\{ (x_k - \beta_k)^2 \sigma_k^2 - (x_k - \gamma_k)^2 \lambda_k^2 \right\} \right\} \]

and

\[ \alpha_k = \int_{\mathbb{R}} \sqrt{\rho_k(x_k)} \, d\nu_k(x_k) = \sqrt{\frac{2\lambda_k \sigma_k}{\lambda_k^2 + \sigma_k^2}} \exp\left\{ -\frac{(\beta_k - \gamma_k)^2}{4(\lambda_k^2 + \sigma_k^2)} \right\}. \]

We now give some particular cases:

- Same covariance (\( \lambda_k = \sigma_k \) for every \( k \)). \( \mu \) and \( \nu \) are equivalent if and only if

\[ \sum_k \frac{(\beta_k - \gamma_k^2)^2}{\sigma_k^2} < \infty \]

and the density is then equal to

\[ \exp\left\{ \sum_{k=1}^{\infty} \frac{x_k(\beta_k - \gamma_k)}{\sigma_k^2} - \frac{\beta_k^2 - \gamma_k^2}{2\sigma_k^2} \right\}. \]

Otherwise, we have orthogonality of measures.
- **Same mean** \( \beta_k = \gamma_k = 0 \) for every \( k \).

\( \mu \) and \( \nu \) are equivalent if and only if:

\[
\sum_{k=1}^{\infty} \frac{(\lambda_k - \sigma_k)^2}{\lambda_k \sigma_k} < \infty
\]

and in this case the density is equal to:

\[
\frac{d\mu}{d\nu}(x) = \lim_{n \to \infty} \prod_{k=1}^{n} \frac{\sigma_k}{\lambda_k} \exp\left\{-\frac{x_k^2}{2} \left(\frac{\sigma_k^2 - \lambda_k^2}{\sigma_k^2 \lambda_k^2}\right)\right\}.
\]

If this condition is not satisfied we have orthogonality.

### 2 - Affine transformations of Gaussian measures

Now let \((E, H, \mu)\) be an abstract Wiener space. If \((e_n)\) is an orthonormal basis of \(H\), the random variables \(\tilde{e}_n\) are independent Gaussian variables on \(E\), with mean zero and variance one. The law of the sequence \((\tilde{e}_n)\) is therefore a product measure on \(\mathbb{R}^\mathbb{N}\):

\[
\gamma_\mathbb{N} = \bigotimes_{n=0}^{\infty} \gamma_n
\]

where \(\gamma_n = \gamma\) (Gaussian measure on \(\mathbb{R}\)) for every \(n\).

Now we have a measurable (defined almost everywhere) map \(\theta\) of \(E\) into \(\mathbb{R}^\mathbb{N}\):

\[
x \mapsto (\tilde{e}_n(x))_n.
\]

If the \(e_n\) belong to \(E'\), the \(\tilde{e}_n\) are everywhere defined and \(\theta\) is continuous from \(E\) into \(\mathbb{R}^\mathbb{N}\).

It is clear now that the image of \(\mu\) under \(\theta\) is equal to \(\gamma_\mathbb{N}\). We have \(\theta(H) = \ell^2\) as we can see immediately (the \(\tilde{e}_n(x)\) are defined in a unique way on \(H\)).

**Proposition 1**: Let \(a \in E\) and \(\tau_a(\mu)\) be the translate of \(\mu\) by \(a\). Then we have the dichotomy:

\(\tau_a(\mu) \sim \mu\) or \(\tau_a(\mu) \perp \mu\),

\(\tau_a(\mu) \sim \mu\) if and only if \(a \in H\) and the density is equal to \(\exp\{\tilde{a}(\ast) - \frac{1}{2} \|a\|_H^2\}\).
Proof :

\( \tau_a(\mu) \) is a Gaussian (non centered if \( a \neq 0 \)) measure with the same covariance than \( \mu \).
Let \( (e_n) \subset E' \) (orthonormal in \( H \)). It suffices to prove the same result for \( \theta(\mu) \) and \( \theta(\tau_a(\mu)) \). But \( \theta(\tau_a(\mu)) \) is the product of Gaussian measures on \( \mathbb{R} \) with variances one and mean \( e_n(a) \). Therefore it suffices to apply the result of the previous paragraph.

\[ \text{--- Q.E.D. --} \]

Now let \( T = I + F \) be a linear continuous transform of \( E \) into \( E \). Let us suppose that \( F(E) \subset H \). In this case \( F \) is continuous for the topology of \( H \) by closed graph theorem.
Suppose moreover, that \( T|_H = I + F|_H \) is an \textit{invertible operator}. Then \( T : E \to E \) is also invertible and

\[ T^{-1} = I - (T|_H)^{-1} \circ F. \]

**Proposition 2 :** Suppose \( T = I + F \) with the above properties and that \( F|_H \) is nuclear. Then \( T^{-1}(\mu) \) and \( \mu \) are equivalent and

\[ \frac{dT^{-1}(\mu)}{d\mu}(x) = \exp\left\{ -(Fx, x)_H - \frac{1}{2} \|Fx\|_H^2 \right\} \left| \det T \right|. \]

Proof :

Let us explain what this formula means. Indeed, \( F|_H \) being nuclear, admits the decomposition : \( F|_H(x) = \sum_n \lambda_n (x, e_n)_H f_n \), \((e_n, f_n \text{ orthonormal in } H)\) and we can define \( \langle F(x), x \rangle_H \) on \( E \) by \( \sum_n \lambda_n \tilde{e}_n(x) \tilde{f}_n(x) \). We set : \( \det (I + F) = \prod_n (1 + \lambda_n) \). (This has sense since \( \sum_n |\lambda_n| < \infty \).

- Let us suppose first that \( F \) is symmetrical :

\[ F(x) = \sum_n \lambda_n (x, e_n)_H e_n \]

where \( e_n \) is an orthonormal basis composed of eigenvectors of \( F \).
Let \( \theta : E \to \mathbb{R}^N \) associated to these \( e_n \). We have seen that \( \theta(\mu) = \gamma_N \) (product measure).
Now $\theta((I + F)^{-1}\mu)$ is the product of measures with densities:

$$\frac{1}{\sqrt{2\pi}} (1 + \lambda_n) \exp\left\{ -\frac{1}{2} (1 + \lambda_n)^2 x_n^2 \right\}.$$  

We have

$$d((1 + 03BB_n)^{-1} \varepsilon_n(\mu)) (x_n) = (1 + \lambda_n) \exp\left\{ -\lambda_n x_n^2 - \frac{1}{2} \lambda_n^2 x_n^2 \right\}$$

$$d(\theta((I + F^{-1})(\mu))) (x) = \prod (1 + \lambda_n) \exp\left\{ -(Fx, x)_H - \frac{1}{2} \|Fx\|_H^2 \right\}.$$  

- Now let us consider the general case ($F$ non necessarily symmetrical)

$$(I + F) \circ i$$ is an operator from $H$ into $H$. There exists a unitary operator $U : H \to H$ “diagonalizing” $F \circ i$, therefore $(I + F) \circ i$. Let $\tilde{U}$ its extension to $E \to E$. We apply the result for $\tilde{U}(I + F) \tilde{U}^{-1}$.

--- Q.E.D. ---

Now we shall consider the case where $F|_H$ is not nuclear.

We know that in any case $F|_H$ is Hilbert-Schmidt.

- Suppose at first that rank ($F$) is finite.

Then the formula of Proposition 2 gives:

$$n \prod_{i=1}^{n} (1 + \lambda_i) \exp\left\{ -\sum_{i=1}^{n} \lambda_i x_i^2 - \frac{1}{2} \sum_{i=1}^{n} \lambda_i^2 x_i^2 \right\}$$

$$= \prod_{i=1}^{n} (1 + \lambda_i) e^{-\lambda_i} \exp\left\{ -\left(\sum_{i=1}^{n} \lambda_i x_i^2 - \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \|Fx\|_H^2 \right) \right\}.$$  

- Now suppose $F$ Hilbert-Schmidt with infinite rank:

$$\prod_{i} (1 + \lambda_i) e^{-\lambda_i}$$ converges since $\sum_{i} |\lambda_i|^2 < \infty$.

The limit is called the “Carleman determinant”. 
Now we can prove that
\[
\lim_{n\to\infty} \exp\left\{-\sum_{i=1}^{n} \lambda_i x_i^2 - \sum_{i=1}^{n} \lambda_i \right\} - \frac{1}{2} \|Fx\|_H^2 \} \text{ exists in } L^1(\mu) \text{ if } F \text{ is } H-S.
\]

We denote it by:
\[
\exp\left\{-\left[(Fx,x)_H - \text{Trace } F\right] - \frac{1}{2} \|Fx\|_H^2 \right\}.
\]

Therefore we have the following theorem:

**THEOREM 2**: Let \( T : E \to E \) linear continuous, such that \( Tx = x + Fx \) with \( F(E) \subseteq H \).

Then \( F|_H \) defines a Hilbert-Schmidt operator from \( H \) into \( H \). Suppose that \( T|_H \) is invertible then \( T : E \to E \) is invertible. Moreover, \( T^{-1}(\mu) \) is absolutely continuous with respect to \( \mu \) and we have
\[
\frac{d(T^{-1}(\mu))}{d\mu} (x) = \tilde{\Delta}(I + F) \exp\left\{-\left[(Fx,x)_H - \text{Trace } F\right] - \frac{1}{2} \|Fx\|_H^2 \right\}
\]

with
\[
\tilde{\Delta}(I + F) = \prod_{1}^{\infty} (1 + \lambda_i) e^{-\lambda_i},
\]
the \( \lambda_i \) being the eigenvalues of \( F \).

We have seen the affine case.

Now we may give the result for the general case announced in the beginning.

**THEOREM 3**: Let \( F \in D^{2,1}(H) \). Suppose that \( (I + F) \) is invertible and that for every \( x \in E \), the operator \( I_H + \nabla F(x) \) from \( H \) to \( H \) is invertible, then \( (I + F)^{-1}(\mu) \) is absolutely continuous with respect to \( \mu \) and we have:
\[
\frac{d((I + F)^{-1}\mu)}{d\mu} (x) = \tilde{\Delta} \left(I_H + \nabla F(x)\right) \exp\left\{-\delta(F)(x) - \frac{1}{2} \|Fx\|_H^2\right\}.
\]
CHAPTER THREE

Transformation of Gaussian measures under anticipative flows

Let $(\Omega, H, P)$ be an abstract Wiener space and let $T$ be an invertible transformation of $\Omega$ into $\Omega$ (the only interesting case will be of the form : $T := Id + F$ with $F \in D^{2,1}(H)$).

Definition: A family of transformations $(T_t)_{t \in [0,1]}$ from $\Omega$ to $\Omega$ will be called an "interpolation" of the invertible transformation $T$ if

a) $T_0 = Id$, $T_1 = T$,
b) each $T_t$ is invertible,
c) for each $\omega$, $t \leadsto T_t\omega$ and $t \leadsto T_t^{-1}\omega$ are strongly continuous.

Moreover, if
d) for each $\omega$, $t \leadsto T_t\omega$ and $t \leadsto T_t^{-1}\omega$ are strongly continuously differentiable, the interpolation will be said to be "smooth".

Example 1: $T_t(\omega) = \omega + tA(\omega)$ where $A$ is a function from $\Omega$ to $H$, such that

$$\omega \leadsto \omega + tA(\omega)$$
is invertible for every $t$.

Example 2: Suppose $A : \Omega \rightarrow H$ is continuous and suppose that we have defined a family of transformations $(T_t)$ from $\Omega$ into $\Omega$ by:

$$T_t\omega = \omega + \int_0^t A(T_s\omega) \, ds \quad \text{(time homogeneous case)}$$

i.e.

$$\left| \frac{dT_t}{dt}(\omega) \right| = A(T_t\omega)$$

$$\left| \frac{dT_t}{dt}(\omega) \right| = A(T_t\omega)$$

we have then:

$$\frac{dT_t}{dt}(T_t^{-1}(\omega)) = A(\omega).$$
Example 3: \( T_t(\omega) = \omega + \int_0^t \sum(s, T_s(\omega)) \, ds. \)

If \( \sum(r, \omega) \) is continuous on \([0,1] \times \Omega\) into \( \Omega \) or into \( H \) and satisfies a global Lipschitz condition:

\[
|\sum(t, \omega_1) - \sum(t, \omega_2)| \leq L\|\omega_1 - \omega_2\|
\]

We can consider \( T_t(\omega) \) as the solution of the ordinary differential equation

\[
\begin{cases}
\frac{dT_t}{dt}(\omega) = \sum(t, T_t(\omega)) \\
T_0(\omega) = \omega
\end{cases}
\]

on the Banach space \( \Omega \).

If for every \( t \in [0,1], \sum(t, \omega) \) is Fréchet differentiable, with Fréchet differential denoted by \( \partial \sum(t, \omega) \), and if we assume that \( \partial \sum(t, \omega) \) is bounded continuous on \([0,1] \times \Omega\), then the equation

\[
T_t\omega = \omega + \int_0^t \sum(r, T_r(\omega)) \, dr
\]

has a unique solution.

Moreover, \( \omega \rightarrow T_t(\omega) \) is Fréchet differentiable and \( \partial T_t(\omega) \) is continuous, invertible on \([0,1] \times \Omega\), and satisfies the differential equation:

\[
\frac{d}{dt}(\partial T_t\omega) = (\partial \sum(t, \omega) \circ T_t(\omega)) \cdot \partial T_t(\omega).
\]

Its inverse \( \partial^{-1}T_t\omega \) satisfies:

\[
\frac{d}{dt}(\partial^{-1}T_t\omega) = -\partial^{-1}T_t(\omega) \cdot (\partial \sum(t, \omega) \circ T_t(\omega)).
\]

Consequently, by the global inverse theorem, \( T_t(\omega) \) is a \( C_1 \)-diffeomorphism. Therefore, we have an interpolation of \( T \) defined by

\[
T(\omega) = \omega + \int_0^1 \sum(r, T_r\omega) \, dr.
\]

Later on we shall come back to this example. Now let us return to the general situation.
THEOREM 1 : Let $T$ be a transformation from $\Omega$ to $\Omega$ and $(T_t, t \in [0,1])$ be an interpolation of $T$. Let us assume moreover that

(a) $T_t(P) \ll P$, $\forall t \in [0,1]$ and let $X_t(\omega) = \frac{dT_t(P)}{dP}(\omega)$,

(b) $G_t = T_t^{-1} - I \in \mathbb{D}^{2,1}(H)$ and $\frac{dT_t^{-1}}{dt} \in H$,

(c) $\frac{dT_t^{-1}}{dt}$ as a function from $[0,1] \times \Omega$ into $H$ is almost surely continuous in $(t, \omega)$ (for $dt \otimes dP$) and $\nabla T_t^{-1}(\omega)$ will be assumed to possess a continuous extension $[0,1] \times \Omega$,

(d) $\frac{dT_t^{-1}}{ds} \circ T_t \in \mathbb{D}^{2,1}(H)$.

Then

$$X_t(\omega) = \exp\left\{ - \int_0^t \left( \delta \left[ \frac{dT_t^{-1}}{ds} \circ T_s \right] \right) \circ T_t^{-1}(\omega) \, ds \right\}$$

(1)

This implies that the measures $T_t(P), T_t^{-1}(P)$ and $P$ are equivalent.

Moreover

$$X_t = \exp\left\{ - \int_0^t \tilde{\delta} \left[ \frac{dG_s}{ds} \right] \, ds \\
- \frac{1}{2} \langle G_t, G_t \rangle_H \\
- \int_0^t \text{Trace} \left[ \left( \nabla \left[ \frac{dG_s}{ds} \circ T_s \right] \circ T_s^{-1} \right) \cdot \nabla G_s \right] \, ds \right\}$$

(2)

where $\tilde{\delta}$ was defined previously by:

$$\tilde{\delta} (\xi \circ T) = (\delta \xi) \circ T - \langle \xi \circ T, F \rangle_H - \text{Trace} \left( (\nabla \xi) \circ T \cdot \nabla F \right).$$

Moreover, if $\frac{dG_s}{ds}$ and $G_s$ are in $\mathbb{D}^{2,1}(H)$, then the formula (2) becomes:

$$X_t = \exp\left\{ - \delta(G_t) - \frac{1}{2} \langle G_t, G_t \rangle_H \\
- \int_0^t \text{Trace} \left[ \left( \nabla \left[ \frac{dG_s}{ds} \circ T_s \right] \circ T_s^{-1} \right) \cdot \nabla G_s \right] \, ds \right\}.$$  (3)
Proof of (1):

We have:

\[
0 = \frac{1}{\varepsilon} \left[ T_{t+\varepsilon}^{-1} \circ T_t - T^{-1}_t \circ T_t \right]
\]

\[
= \frac{1}{\varepsilon} \left[ T_{t+\varepsilon}^{-1} \circ T_t - T_{t+\varepsilon}^{-1} \circ T_t \right] + \frac{1}{\varepsilon} \left[ T_{t+\varepsilon}^{-1} \circ T_t - T^{-1}_t \circ T_t \right].
\]

Therefore by (c)

\[
\left[(\nabla T^{-1}_t) \circ T_t(\omega)\right] \cdot \frac{dT_t}{dt}(\omega) + \frac{dT^{-1}_t}{dt} \circ T_t(\omega) = 0
\]  

(4)

Let now \(a: \Omega \to \mathbb{R}\) smooth and let \(h \in H\). By (d) we have:

\[
\langle (\nabla a) \circ T_t(\omega), h \rangle_H = \lim_{\varepsilon \to 0} \frac{\partial}{\partial \varepsilon} a(T_t(\omega + \varepsilon h))
\]

\[
= \lim_{\varepsilon \to 0} \frac{\partial}{\partial \varepsilon} \left((a \circ T_t)(T^{-1}_t(T_t(\omega + \varepsilon h)))\right)
\]

\[
= \lim_{\varepsilon \to 0} \frac{\partial}{\partial \varepsilon} \left((a \circ T_t)(\omega + \varepsilon(\nabla T^{-1}_t) \circ T_t(\omega) \cdot h + o(\varepsilon))\right)
\]

\[
= \langle \nabla(a \circ T_t), (\nabla T^{-1}_t) \circ T_t(\omega) \cdot h \rangle_H.
\]

Now if we set \(h = \frac{d}{dt} T_t(\omega)\), comparing with (4), we obtain:

\[
\langle (\nabla a) \circ T_t(\omega), \frac{d}{dt} T_t(\omega) \rangle_H = -\langle (\nabla a \circ T_t)(\omega), \frac{dT^{-1}_t}{dt} \circ T_t(\omega) \rangle_H.
\]

But the left-hand member of this equality is equal to \(\frac{d}{dt} (a \circ T_t)(\omega)\). Therefore we obtain:

\[
\mathbb{E}\{a \circ T_t(\omega) - a(\omega)\} = \mathbb{E} \left( \int_0^t \frac{d}{ds} (a \circ T_s)(\omega) \, ds \right)
\]

\[
= -\mathbb{E} \left( \int_0^t \langle \nabla (a \circ T_s)(\omega), \frac{dT^{-1}_s}{ds} \circ T_s(\omega) \rangle \, ds \right).
\]

But from condition (d) , \(\left(\frac{dT^{-1}_s}{ds} \circ T_s \in \mathbb{D}^{2,1}(H)\right)\), and integrating by parts we obtain:

\[
\mathbb{E}\{a \circ T_t(\omega) - a(\omega)\} = -\int_0^t \mathbb{E}\left\{ (a \circ T_s(\omega)) \delta \left[ \frac{dT^{-1}_s}{ds} \circ T_s \right](\omega) \right\} \, ds
\]
and

$$\mathbb{E} \{a(\omega).(X_t(\omega) - 1)\} = -\mathbb{E} \left( \int_0^t a(\omega) \ X_s(\omega) \left( \delta \left[ \frac{dT^{-1}_s}{ds} \circ T_s \right] \right) \circ T_s^{-1} \omega \ ds \right).$$

Since this last inequality is true for smooth functions we have:

$$X_t(\omega) = 1 - \int_0^t X_s(\omega) \left( \delta \left[ \frac{dT^{-1}_s}{ds} \circ T_s \right] \right) \circ T_s^{-1} \omega \ ds.$$

Finally, since $X_t$ is $P$-almost surely positive, $T_tP$ and $P$ are equivalent.

On the other hand, if $a : \Omega \to \mathbb{R}$ is smooth, then:

$$\mathbb{E} \{a \circ T^{-1}_t \ X_t\} = \mathbb{E}a.$$

Hence if $B$ is a Borelian subset of $\Omega$, then

$$P(B) = 0 \iff \mathbb{E}\{1_B \circ T^{-1}_t \ X_t\} = 0 \iff 1_B \circ T^{-1}_t = 0, \ a.s.$$

Therefore, $T^{-1}_t (P)$ and $P$ are equivalent.

**Proof of (2) :**

We start from

$$(\delta \xi) \circ T = \delta (\xi \circ T) + (\xi \circ T, F)_H + \text{Trace} \left( (\nabla \xi) \circ T \cdot \nabla F \right)$$

with

$$\xi = \frac{dT^{-1}_s}{ds} \circ T_s, \ T = T_s^{-1}, \ F = T - \text{Id} = G_s$$

and

$$\frac{dG_s}{ds} = \frac{dT^{-1}_s}{ds}.$$

Then

$$\delta \left[ \frac{dT^{-1}_s}{ds} \circ T_s \right] \circ T_s^{-1} = \delta \left( \frac{dG_s}{ds} \right) + \left( \frac{dG_s}{ds}, G_s \right) + \text{Trace} \left( \left( \nabla \left[ \frac{dG_s}{ds} \circ T_s \right] \right) \circ T_s^{-1} \cdot \nabla G_s \right)$$

and we integrate from 0 to $t$. 
Proof of (3):

It is immediate from (2) since \( \delta = \delta \) under this hypothesis.

We have expressed the density \( X_s \) in terms of \( \frac{dT_s^{-1}}{dt} \). (The next result will give an expression of \( X_t \) in terms of \( \frac{dT_s}{ds} \)).

\[ - \text{Q.E.D. -} \]

Corollary: Under the assumptions and conditions of the theorem 1 let us replace \( T, T_t, T_s \) and \( X_t \) by \( T^{-1}, T_t^{-1}, T_s^{-1}, \frac{dT_t^{-1}(P)}{dP} = Y_t \). Then we have:

\[
X_t(\omega) = \frac{dT_t(P)}{dP}(\omega)
= \exp\left\{ \int_0^t \left( \delta \left[ \frac{dT_s}{ds} \circ T_s^{-1}(\omega) \right] \right) \circ T_s T_t^{-1}(\omega) \, ds \right\}
\]

and

\[
X_t(\omega) = \exp\left\{ - \delta(G_t)(\omega) - \frac{1}{2} \langle G_t, G_t \rangle_H(\omega)
+ \int_0^t \text{Trace} \left[ \left( \nabla \left[ \frac{dT_s}{ds} \circ T_s^{-1} \right] \circ T_s T_t^{-1}(\omega) \right) \bullet \nabla \left( G_t - G_s (T_s T_t^{-1}) \right)(\omega) \right] \, ds \right\}.
\]

Proof:

By Theorem 1:

\[
Y_t(\omega) = \exp\left\{ - \int_0^t \left( \delta \left[ \frac{dT_s}{ds} \circ T_s^{-1} \right] \right) \circ T_s(\omega) \, ds \right\}.
\] (A)

On the other hand, if \( a \) is a smooth functional:

\[
\mathbb{E}\{a(\omega) Y_t^{-1} (T_t^{-1}(\omega))\} = \mathbb{E}\{a(T_t T_t^{-1}(\omega)) Y_t^{-1} (T_t^{-1}(\omega))\}
= \mathbb{E}\{a(T_t(\omega)) Y_t^{-1}(\omega) Y_t(\omega)\}
= \mathbb{E}\{a(\omega) X_t(\omega)\}.
\]

Therefore:

\[
X_t(\omega) = Y_t^{-1} (T_t^{-1}(\omega)) = \exp\left\{ \int_0^t \left( \delta \left[ \frac{dT_s}{ds} \circ T_s^{-1}(\omega) \right] \right) \circ T_s \circ T_t^{-1}(\omega) \, ds \right\},
\]

which proves the first formula.
To prove the second formula let us start from

\[ T_s \omega = \omega + F_s(\omega) \]

which implies

\[ T_s T_t^{-1} \omega = T_t^{-1} \omega + F_s(T_t^{-1} \omega), \]

and if \( s = t \)

\[ \omega = T_t^{-1} \omega + F_t(T_t^{-1} \omega). \]

Therefore

\[ T_s T_t^{-1} \omega = \omega + F_s(T_t^{-1} \omega) - F_t(T_t^{-1} \omega). \]

Now

\[ G_t(\omega) = T_t^{-1}(\omega) - \omega = -F_t(T_t^{-1} \omega). \]

Therefore:

\[ T_s T_t^{-1} \omega = \omega + G_t(\omega) - G_s(T_s T_t^{-1} \omega). \]

In the formula

\[ X_t(\omega) = \exp \left\{ \int_0^t \left( \delta \left[ \frac{dT_s}{ds} \circ T_t^{-1} \right] \right) \circ T_s T_t^{-1} \omega \, ds \right\}, \]

let us apply the formula given \( \delta \) in terms of \( \tilde{\delta} \). We obtain:

\[ X_t(\omega) = \exp \left\{ \int_0^t \left( \tilde{\delta} \left[ \frac{dT_s}{ds} \circ T_t^{-1} \right] \right) \circ T_s T_t^{-1} \omega \, ds \right\} \]

\[ + \left( \frac{dT_s}{ds} \circ T_t^{-1}(\omega), G_t(\omega) - G_s(T_s T_t^{-1} \omega) \right)_H \]

\[ + \text{Trace} \left[ \left( \nabla \left[ \frac{dT_s}{ds} \circ T_t^{-1} \right] \circ T_s T_t^{-1}(\omega) \right) \cdot \nabla \left( G_t - G_s(T_s T_t^{-1} \omega) \right) \right] \, ds \}

Now we integrate with respect to \( s \), by using:

\[ \frac{d}{ds} (T_s \circ T_t^{-1}(\omega)) = - \frac{d}{ds} (G_s(T_s T_t^{-1} \omega)) = \frac{d}{ds} (G_t(\omega) - G_s(T_s T_t^{-1} \omega)). \]

— We obtain the second formula.—
Now we give an integral equation satisfied by $X_t$.

**THEOREM 2**: Let $T : \Omega \rightarrow \Omega$ and $T_t : \Omega \rightarrow \Omega$ ($t \in [0,1]$) be an interpolation of $T$. Assume that for each $t \in [0,1]$, $T_t(P) \ll P$ and that $X_s \left[ \frac{dT_s}{ds} \circ T_s^{-1} \right] \in \mathbb{D}^{2,1}_{\text{loc}}(H)$ (this condition is satisfied if $\frac{dT_s}{ds} \circ T_s^{-1} \in \mathbb{D}^{2,1}(H)$ and $X_s \in \mathbb{D}^{2,1}_{\text{loc}}$), then $X_t$ satisfies:

$$X_t = 1 + \int_0^t \delta[X_s \frac{dT_s}{ds} \circ T_s^{-1}] \, ds.$$

**Proof:**

Let $a$ be a smooth functional. Then

$$\mathbb{E}\{X_t(\omega)a(\omega)\} = \mathbb{E}\{a(T_t(\omega))\}$$

$$= \mathbb{E}\{a(\omega) + \int_0^t \frac{d}{ds}(T_s(\omega)) \, ds\}$$

$$= \mathbb{E}\{a(\omega) + \int_0^t \langle(\nabla a) \circ T_s \omega, \frac{d}{ds} T_s(\omega) \rangle \, ds\}$$

$$= \mathbb{E}\{a(\omega) + \int_0^t \mathbb{E}\{X_s(\omega)(\nabla a)(\omega), T_s^{-1}(\omega)\} \} \, ds$$

$$= \mathbb{E}\{a(\omega) + \int_0^t \mathbb{E}\{a(\omega) [X_s \frac{dT_s}{ds} \circ T_s^{-1}](\omega)\} \} \, ds$$

$$= \mathbb{E}\{a(\omega) + \int_0^t \mathbb{E}\{a(\omega) \} \} \, ds$$

--- Q.E.D. ---

**Applications of these formulas.**

- **In the example (1)**: $T_t(\omega) = \omega + t \ A(\omega)$,

$$X_t(\omega) = \exp\left\{ \int_0^t \left( \delta[A(T_s^{-1}(\omega)) \circ T_s T_t^{-1}(\omega)] \right) \, ds \right\}$$

(this result was obtained by Bell).

- **In the example (2)**: $T_t(\omega) = \omega + \int_0^t A(T_s(\omega)) \, ds$

$$\frac{dT_s}{ds} (T_s^{-1}(\omega)) = A(\omega)$$
and

\[ X_t(\omega) = \exp\left\{ \int_0^t \left( \delta(A) \right) \circ T_sT_t^{-1}(\omega) \, ds \right\}. \]

- We shall now study the example three:

\[ T_t(\omega) = \omega + \int_0^t \sum (r, T_r(\omega)) \, dr. \]  

(B)

We have given some hypotheses insuring that \( T_t \omega \) is a solution of the ODE with values in the Banach space \( \Omega \)

\[
\begin{align*}
\frac{dT_t}{dt}(\omega) &= \sum(t, T_t(\omega)) \\
T_0(\omega) &= \omega
\end{align*}
\]

and that \( \omega \sim T_t(\omega) \) and \( \omega \sim T_t^{-1}(\omega) \) are Fréchet differentiable (in \( \omega \)). Then:

\[ I_H + \nabla \int_0^t \sum(s, T_s\omega) \, ds \]

is invertible and satisfies the hypotheses of Ramer's theorem.

As a consequence the probabilities

\[ T_t P, \ P \text{ and } T_t^{-1} P \]

are equivalent.

Now in (B) we replace \( \omega \) by \( T_s^{-1}\omega \):

\[ T_t T_s^{-1}(\omega) = T_s^{-1}(\omega) + \int_0^t \sum(r, T_r T_s^{-1}(\omega)) \, dr. \]

Setting \( T_t T_s^{-1}(\omega) = \varphi_{s,t}(\omega) \) and \( T_s T_t^{-1}(\omega) = \psi_{s,t}(\omega), \ t \geq s \), we have:

\[ \psi_{s,t} \circ \varphi_{s,t} = \varphi_{s,t} \circ \psi_{s,t} = \text{Id} \]

and:

\[ \varphi_{s,t}(\omega) = \omega + \int_s^t \sum(r, \varphi_{s,r}(\omega)) \, dr \]

\[ \psi_{s,t}(\omega) = \omega - \int_s^t \sum(r, \psi_{r,t}(\omega)) \, dr. \]

Note that \( \varphi_{(1-s)t,t}, \ s \in [0,1] \) is, for \( t \) fixed, an interpolation of \( T_t \) and naturally \( (T_t)_{t \in [0,1]} \) is an interpolation of \( T_1 : \varphi_{s,t} \) is a \textit{two-parameter} interpolation of \( T \).
Now we shall specialize the example in the case $\Omega = C_0[0,1]$, with the Wiener measure and we shall use the following notations in this case:

If $U$, $U_1$ and $U_2$ are random functions with values in $H$; if $H$ is the Cameron-Martin space, then

$$U(\omega)(\cdot) = \int_0^\infty u(\theta, \omega) \, d\theta$$

$$\delta(U) = \int_0^1 u(\theta, \omega) \, \delta_\theta(W)$$

$$\langle U_1, U_2 \rangle = \int_0^1 u_1(\theta, \omega) u_2(\theta, \omega) \, d\theta.$$ 

But if $H$ is the $L^2[0,1]$ space

$$U(\omega)(\cdot) = u(\cdot, \omega)$$

$$\delta U = \int_0^1 u(\theta, \omega) \, \delta_\theta(W)$$

$$\langle U_1, U_2 \rangle = \int_0^1 u_1(\theta, \omega) u_2(\theta, \omega) \, d\theta.$$

$$(T_t\omega)(\cdot) = \omega(\cdot) + \int_0^t \rho(r, \cdot) \, \sigma(r, T_r \omega) \, dr \quad (C)$$

where $\rho$ is a smooth function on $[0,1]^2$ and $\sigma : [0,1] \times \Omega \to \mathbb{R}$ is assumed to satisfy Lipschitzian and differentiability conditions.

In terms of $\varphi_{s,t}$ and $\psi_{s,t}$, $(s \leq t)$ we have:

$$\varphi_{s,t}(\omega)(\cdot) = \omega(\cdot) + \int_s^t \rho(r, \cdot) \, \sigma(r, \varphi_{s,r}(\omega)) \, dr$$

$$\psi_{s,t}(\omega)(\cdot) = \omega(\cdot) - \int_s^t \rho(r, \cdot) \, \sigma(r, \psi_{r,t}(\omega)) \, dr.$$ 

We consider these equations as ODE in Banach space (the first in $t$ with $s$ fixed; the second in $s$ for $t$ fixed), we have existence and unicity of solutions with

$$\varphi_{s,s}(\omega) = \omega, \quad \psi_{t,t}(\omega) = \omega \quad \text{and} \quad \varphi_{s,t} \circ \psi_{s,t}(\omega) = \omega.$$ 

Then $\psi_{s,t}(\omega)$ and $\varphi_{s,t}(\omega)$ are Fréchet differentiable in $\omega \in C_0([0,1])$. 
Consequently, $\partial \varphi_{s,t}$ and $\partial \psi_{s,t}$ restricted to $H$ are invertible, and by Ramer's theorem: $\varphi_{s,t}(P)$, $\psi_{s,t}(P)$ and $P$ are equivalent.

Set

$$L_{s,t}(\omega) = \frac{d\varphi_{s,t}(P)}{dP}$$

and

$$\Lambda_{s,t} = \frac{d\psi_{s,t}(P)}{dP}.$$

Now let us fix $t$ in the equation:

$$T_t \omega (\cdot) = \omega(\cdot) + \int_0^t \rho(r, \cdot) \sigma(r, T_r \omega) \, dr.$$ 

Let $s = t - \lambda$ and $\lambda \in [0, t]$ be the interpolation parameters.

Now let us recall that (cf (3))

$$X_t = \exp\left\{ -\delta(G_t) - \frac{1}{2} \langle G_t, G_t \rangle_H ight. \
- \int_0^t \text{Trace} \left( \nabla \left[ \frac{dG_s}{ds} \circ T_s \circ T_s^{-1} \bullet \nabla G_s \right] \right) ds \right\} \quad (D)$$

where $G_t = T_t^{-1} - Id$, and apply the result for $T_t$ satisfying the relation:

$$T_t \omega (\cdot) = \omega(\cdot) + \int_0^t \rho(r, \cdot) \sigma(r, T_r \omega) \, dr.$$ 

Then we obtain an expression for $X_t$:

$$X_t = \exp\left\{ \int_0^1 \left[ \int_0^t \frac{\partial \rho}{\partial \theta} (r, \theta) \sigma(r, \psi_{0,r}) \, dr \right] \delta_\theta(W) \right. \
- \frac{1}{2} \int_0^1 \left[ \int_0^t \frac{\partial \rho(r, \theta)}{\partial \theta} \sigma(r, \psi_{0,r}) \, dr \right]^2 \, d\theta \right. \
- \int_0^t \int_0^t \int_0^t \left[ \int_0^\lambda \frac{\partial \rho(r, \eta)}{\partial \eta} D_\theta \sigma(r, \psi_{0,r}) \, dr \right] \circ \frac{\partial \rho(\lambda, \theta)}{\partial \theta} (D_\eta \sigma(\lambda, \cdot)) \circ \psi_{0,\lambda} \, d\lambda \, d\theta \, d\eta \right\}.$$ 

We can obtain another formula for the Radon-Nikodym density using the relation:

$$\delta(aU) = a \delta U - (\nabla a, U)_H$$
in the expression:

\[ X_t(\omega) = \exp \left\{ \int_0^t \left( \delta \left[ \frac{dT_s}{ds} \circ T_s^{-1} \right] \right) \circ T_s T_t^{-1}(\omega) \, ds \right\}. \]

We then obtain:

\[ L_{s,t} = \exp \left\{ \int_s^t \sigma(r, \psi_{r,t}) \left[ \delta \rho(r, \cdot) - \int_s^r \sigma(u, \psi_{u,t}) \langle \rho(r, \cdot), \rho(u, \cdot) \rangle_H \, du \right] \, dr 
- \int_s^t \langle \nabla \sigma \rangle(r, \psi_{r,t}), \rho(r, \cdot) \rangle_H \, dr \right\}. \]

**REFERENCES**


