SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

ALI SÜLEYMAN ÜSTÜNEL
MOSHE ZAKAI

The change of variables formula on Wiener space

Séminaire de probabilités (Strasbourg), tome 31 (1997), p. 24-39
<http://www.numdam.org/item?id=SPS_1997__31__24_0>

The change of variables formula on Wiener space

A.S. Üstünel and M. Zakai

Abstract

The transformations of measure induced by a not-necessarily adapted perturbation of the identity is considered. Previous results are reviewed and recent results on absolute continuity and related Radon-Nikodym densities are derived under conditions which are 'as near as possible' to the conditions of Federer's area theorem in the finite dimensional case.

I. Introduction

Let $x \in \mathbb{R}^n$ and T a C^1 map from $\mathbb{R}^n$ to $\mathbb{R}^n$. The classical Jacobi formula yields

$$\int_{\mathbb{R}^n} \rho(x) g(Tx) |J(x)| dx = \int_{\mathbb{R}^n} g(x) \sum_{\theta \in T^{-1}\{x\}} \rho(\theta) dx \tag{1.1}$$

where J is the Jacobian determinant of T and ρ and g are bounded, positive and of compact support. Consider now the formulation of the same result with the Lebesgue measure replaced by the standard Gaussian measure on $\mathbb{R}^n$. Replacing, in (1.1), $g(x)$ with $(2\pi)^{-n/2} e^{-|x|^2/2} g(x)$ and setting

$$T(x) = x + f(x)$$

yields

$$E\left[\rho(x) g(Tx) \,|\, \Lambda(x)|\right] = E\left[g(x) \sum_{\theta \in T^{-1}\{x\}} \rho(\theta)\right] \tag{1.2}$$

where $E\psi(x) = \int_{\mathbb{R}^n} \psi(x) (2\pi)^{-n/2} e^{-|x|^2/2} dx$ and

$$\Lambda(x) = J(x) \cdot \exp - \sum_{i=1}^{n} f_i(x) \cdot x_i - \frac{1}{2} \sum_{1}^{n} f_i^2(x)\,. \tag{1.3}$$

Equation (1.1) or (1.2) under the C^1 condition may be considered as "the first year calculus change of variables formula". The conditions under which (1.1) is valid have been considerably extended by Federer [9] by replacing the C^1 requirement on T with the requirement that T be Lipschitz and more generally by the condition

(A) $\mathbb{R}^n$ is the countable union of measurable sets such that the restriction of T to each set is Lipschitz and then (1.2) holds in the sense that if one of the sides of (1.2) is finite so is the other side and equality holds. Equation (1.1) or (1.2) under (A) may be referred to as the "Federer change of variables formula".

The extension of equation (1.2) to the infinite dimensional case was considered, first, by Cameron and Martin in 1949 [4] and, since then, by many authors. The purpose of this paper is twofold: to survey the results on this topic and to present an extension of (1.2) to the abstract Wiener space under conditions which are "as near as currently possible" to the Federer condition (A).

In the finite dimensional case the proof of Federer's formula is by starting with (1.1) under the C^1 condition and applying the following result of Federer which is based on Whitney's extension theorem:

Theorem 1.1. ([9]) *If $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz then for any $\epsilon > 0$ there exists a function $\psi_\epsilon : \mathbb{R}^n \to \mathbb{R}^n$ which is C^1 on $\mathbb{R}^n$ and satisfies $\mathrm{Leb}_n A_\epsilon \le \epsilon$ where $A_\epsilon \subset \mathbb{R}^n$ and*

$$A_\epsilon = \{x : \psi(x) \neq \psi_\epsilon(x)\} \bigcup \{x : \nabla\psi(x) \neq \nabla\psi_\epsilon(x)\}.$$

Recall the Rademacher theorem which states that if ψ is Lipschitz on $\mathbb{R}^n$ then $\nabla\psi$ exists for Leb_n a.a. x in $\mathbb{R}^n$ and $|\nabla\psi|$ is bounded by the Lipschitz constant of ψ. Note that each of these two theorems characterizes the Lipschitz property of ψ and could be used as a (very non-elegant) definition for this property.

Let (W, H, μ) be an abstract Wiener space. The notions of continuity, Lipschitz continuity and C^1 which turn out to be relevant in the problem of transformation of measure are as follows.

Definition 1.1. *Let (W, H, μ) be an abstract Wiener space, $\mathcal{X}$ a separable Hilbert space and $F(w)$ an $\mathcal{X}$ valued random variable, then*

(a) $F(w)$ is H-continuous if for almost all w, the map $h \mapsto F(w+h)$ is continuous on H.

(b) $F(w)$ is H-Lipschitz continuous with Lipschitz constant c if, for a.a. w

$$|F(w+h) - F(w)|_{\mathcal{X}} \le c|h|_H$$

for all $h \in H$.

(c) $F(w)$ is $H - C^1$ if for almost all w, the map $h \mapsto F(w+h)$ is continuously a Fréchet differentiable function of $h \in H$. A related notion of locally $(H - C^1)$ will be defined later.

As will be pointed out in section III, the results obtained till now for the change of variables formula on abstract Wiener space can be considered as the infinite dimensional extension of the "first-year calculus change of variables formula". The problem of extending the Federer change of variables formula to the abstract Wiener case is delicate because an extension of theorem 1.1 to infinite dimensions (with the Lebesgue measure replaced by the Wiener measure) is not available. Several properties of Lipschitz functions have been extended to the abstract Wiener space (e.g., [8, 19]) but not theorem 1.1. In order to overcome (or bypass) this difficulty we will follow, here, the following path. Let us replace (A) by ($\tilde{\text{A}}$):

($\tilde{\text{A}}$) There exists a countable sequence of measurable sets B_k and C^1 functions $\psi_k(x)$ such that $\mu(\cup B_k) = 1$ and $T(x) = \psi_k(x)$ whenever $x \in B_k$.

Note that by theorem 1.1, condition ($\tilde{\text{A}}$) is equivalent to (A). Instead of extending equation (1.2) under (A) to the abstract Wiener space we will extend (1.2) under ($\tilde{\text{A}}$) to this space.

In the next section we will summarize some results of stochastic analysis which are needed in later sections. Previous results on the change of variables formula will be reviewed in section III. This will represent the path from 1945 till recent years and can be considered as the extension of the "first year calculus formula" to the infinite dimensional case. The extension of Federer's change of variables formula (under ($\tilde{\text{A}}$)) to the infinite dimensional case will be formulated and proved in section IV and an example of a case where this result is applicable while previous results are not, will be given.

II. Preliminaries

Let (W, H, μ) be an abstract Wiener space. We start with a short summary of the notations of the Malliavin calculus. For $h \in H^* = H$, the Wiener integral $w(h)$ will also be denoted $\langle h, w\rangle$, $w \in W$. Let $\mathcal{X}$ be a real separable Hilbert space; smooth, $\mathcal{X}$-valued functionals on (W, H, μ) are functionals of the form

$$a(w) = \sum_{1}^{N} \eta_i(\langle h_1, w\rangle, \cdots, \langle h_m, w\rangle)\, x_i$$

with $x_i \in \mathcal{X}$ and $\eta_i \in C_b^\infty(\mathbb{R}^m)$, $h_i \in W^* \subset H$. For smooth $\mathcal{X}$-valued functionals, define

$$\nabla a(w) = \sum_{i=1}^{N} \sum_{j=1}^{m} \partial_j \eta_i \Big(\langle h_1, w\rangle, \cdots, \langle h_j, w\rangle\Big) \cdot x_i \otimes h_j \,,$$

and $\nabla^k, k = 2, 3, \cdots$ are defined recursively. For $p > 1$, $k \in \mathrm{N}$ the Sobolev space $\mathbb{D}_{p,k}(\mathcal{X})$ is the completion of $\mathcal{X}$-valued smooth functionals with respect to the norm

$$\| a \|_{p,k} = \sum_{i=0}^{k} \| \nabla^i a \|_{L^p(\mu, \mathcal{X} \otimes H^{\otimes i})} \,. \tag{2.1}$$

The gradient $\nabla : \mathbb{D}_{p,k}(\mathcal{X}) \to \mathbb{D}_{p,k-1}(\mathcal{X} \otimes H)$ denotes the closure of ∇ as defined for smooth functionals under the norm of (2.1). The gradient ∇a is considered as a mapping from H to $\mathcal{X}$ and $(\nabla a)^*$ will denote the adjoint of ∇a and is a mapping from $\mathcal{X}^*$ to H. The adjoint of ∇ under the Wiener measure μ is denoted by δ and called the divergence or the Skorohod integral or the Ito-Ramer integral (recall that it is defined by the "integration by parts formula" $E(G\delta u) = E\langle \nabla G, u \rangle_H$ for smooth real valued G and H-valued u). Also recall that if F is in $\mathbb{D}_{p,1}(H)$, for some $p > 1$, then for a.e. w, $\nabla F(w)$ is a Hilbert-Schmidt operator from H to H and for any smooth H-valued F and any complete orthonormal basis of H, say $\{e_i,\ i = 1, 2, \cdots\}$ we have

$$\delta F = \sum_{i=0}^{\infty} \langle F, e_i \rangle_H \langle e_i, w \rangle - \Big\langle \nabla(\langle F, e_i \rangle_H), e_i \Big\rangle_H \,, \tag{2.2}$$

and the Ogawa integral, if it exists, is given by

$$\delta \circ F = \sum_{i=1}^{\infty} \langle F, e_i \rangle_H \langle e_i, w \rangle \,. \tag{2.3}$$

An $\mathcal{X}$-valued random variable F is said to be in $\mathbb{D}^{\mathrm{loc}}_{p,k}(\mathcal{X})$ if there exists a sequence (A_n, F_n) where A_n are measurable subsets of $W, \cup_n A_n = W$ almost surely, $F_n \in \mathbb{D}_{p,k}(\mathcal{X})$ and for every n, $F_n = F$ almost surely on A_n. It was shown in [10] that if $F(w)$ is H valued and $H - C^1$, then $F \in \mathbb{D}^{\mathrm{loc}}_{\infty,1}(H)$.

Let K be a linear operator from H to H with discrete spectrum and let λ_i, $i = 1, 2, \cdots$ be the sequence of eigen-values of K repeated according to their multiplicity. The Carleman-Fredholm determinant of K is defined as:

$$\det\nolimits_2(I + K) = \prod_{i=1}^{\infty} (1 + \lambda_i) e^{-\lambda_i} \tag{2.4}$$

and the product is known to converge for Hilbert-Schmidt operators. For $F \in \mathbb{D}^{\mathrm{loc}}_{p,1}(H)$, ∇F is Hilbert-Schmidt and define

$$\Lambda_F(w) = \det\nolimits_2(I + \nabla F) \exp(-\delta F - \frac{1}{2} \| F \|_H^2) \,. \tag{2.5}$$

The following lemma will be needed in section IV:

Lemma 2.1. *Let F_1, F_2, F_3 belong to $\mathbb{D}^{\text{loc}}_{p,1}(H)$ and let $T_i w = w + F_i(w)$, $i = 1, 2, 3$. Assume that: (i) $\mu \circ T_2^{-1} \ll \mu$ and (ii) $T_3 = T_1 \circ T_2$ (i.e. $F_3 = F_2 + F_1 \circ T_2$). Then*

(a) $I + \nabla F_3 = [I + (\nabla F_1)(T_2)](I + \nabla F_2)$

(b) $\Lambda_{F_3} = (\Lambda_{F_1} \circ T_2) \cdot \Lambda_{F_2}$.

The proof is straightforward (*cf.* lemma 6.1 of [10] or lemma 1.5 of [11]) and uses the fact that for $T(w) = w + u(w)$

$$(\delta F) \circ T = \delta(F \circ T) + \langle F \circ T, u \rangle_H + \text{Trace } ((\nabla F) \circ T \cdot \nabla u\,.$$

Remark: Recall that for any measurable set A on W there exists a *σ-compact modification of* A, i.e. there exists a σ-compact set G such that $G \subset A$ and $\mu(G) = \mu(A)$.

Following Kusuoka [10] we associate with every measurable subset A of W the following random variable $\rho_A(w)$ which plays an important role in the construction of a class of mollifiers:

Definition 2.1. *Let A be a measurable subset of W, set*

$$\rho_A(w) = \inf_{h \in H} \{\| h \|_H : \; w + h \in A\} \tag{2.6}$$

and $\rho_A(w) = \infty$ if $w \notin A + H$.

Clearly, $\rho_A(w) = 0$ if $w \in A$, moreover [10], $\rho_A(w)$ is a measurable random variable and:

(i) If $A \subset B$ then $\rho_A(w) \geq \rho_B(w)$.

(ii) $|\rho_A(w) - \rho_A(w + h)| \leq \| h \|_H$.

(iii) $A_n \nearrow A$ implies $\rho_{A_n}(w) \searrow \rho_A(w)$.

(iv) If G is σ-compact and $\varphi \in C_0^\infty(\mathbb{R})$ (compact support) then $\varphi(\rho_G(w)) \in \mathbb{D}_{p,1}$ for all p and

$$\begin{aligned} \| \nabla \varphi(\rho_G(w)) \|_H &\leq \| \varphi' \|_\infty \cdot \mathbf{1}_{\{\varphi'(\rho_G) \neq 0\}} \\ &\leq \| \varphi' \|_\infty \,. \end{aligned} \tag{2.7}$$

(v) Let $Z = \{w : \; \rho_A(w) < \infty\}$. It is straightforward to see that, $A \subset Z$, and that, if $w \in Z$, then so does $w + h$, for any $h \in H$. Consequently, the distributional derivative (*cf.* e.g. [1] or [16]) $\nabla \mathbf{1}_Z = 0$, hence $\mathbf{1}_Z$ is almost surely a constant. Consequently $\mu(Z) = 1$ if $\mu(A) > 0$.

III. Review of previous results

In this section we present a short guided tour in the research on the change of variables formula along the "main" or "central" research path in the last 45 years, other directions will be mentioned very briefly later. Let μ be the classical Wiener measure on $C_0([0,1])$, and $f_\cdot = \int_0^\cdot f_s' ds$, with $\int_0^1 (f_s')^2 ds < \infty$ denote the elements of the Cameron Martin space. For any $w \in C_0([0,1])$ set

$$\begin{aligned}(Tw)_\cdot &= w_\cdot + \int_0^\cdot f_s'(w) ds \\ &= w + f(w),\end{aligned} \tag{3.1}$$

where $f(w)$ is an H-valued measurable random variable and T is said to be the shift induced by f on the Wiener space . Let $T^*\mu$ denote the measure defined on $C_0([0,1])$

$$T^*\mu(A) = \mu(T^{-1}A)\,. \tag{3.2}$$

Otherwise stated, for any bounded measurable function on $C_0[0,1]$, $E_\mu[g \circ T] = E_{T^*\mu}[g]$.

The first problem associated with the 'change of variables formula' is whether $T^*\mu$ is absolutely continuous with respect to μ (or equivalent to μ) and to calculate the associated Radon Nikodym derivative $L(w)$

$$\frac{dT^*\mu}{d\mu}(w) = L(w)\,.$$

A related, but not equivalent, problem is the following: A measure ν is said to be a Girsanov measure associated with T if $T^*\nu = \mu$, i.e., $\nu(T^{-1}A) = \mu(A)$ or $E_\nu[g \circ T] = E_\mu[g]$; namely, Tw is Wiener under ν. If such ν exists and $\nu \ll \mu$, we will denote $(d\nu/d\mu)(w)$ by $|\Lambda(w)|$, where we denote the density as the absolute value of some random variable Λ since the random variable Λ itself plays an important role in the degree theory on the Wiener space (cf. [21]). Note that if T is (left) invertible

$$E_\mu[g(T^{-1}Tw)] = E_{(T^{-1})^*\mu}[g \circ T]$$

and

$$\nu = (T^{-1})^*\mu\,.$$

The case of T as defined by equation 3.1 where f is non-random was first considered by Cameron and Martin in 1944. This was followed in 1945 with a treatment of the case where f is linearly dependent on w and in 1949 with the case where f may depend non-linearly on w [4].

A short outline of the case where $f(w)$ is finite dimensional is as follows. Let $e_i, i = 1, \cdots$ be a complete orthonormal base on H ($e_i = \int_0^\cdot e_i'(s)ds, \int_0^1 e_i'(s)e_j'(s)ds = \delta_{ij}$) and $w(e_i) = \int_0^1 e_i'(s)dw_s$. Assume that

$$(Tw)_t = w_t + \sum_1^n \psi_i(w(e_1), \cdots, w(e_n))e_i(t)$$

As is well known the Wiener process $w(t)$ has the representation

$$w(t) = \sum_1^\infty \eta_i e_i(t).$$

where $\eta_i = w(e_i)$. Therefore we have in this case

$$(Tw)_t = \sum_1^\infty w(e_i)e_i(t) + \sum_1^n \psi_i(w(e_1), \cdots w(e_n)) \cdot e_i(t) \tag{3.3}$$

and only the first n-coordinates participate in the transformation. Consequently, from equations (1.2) and (1.3) and assuming that T is bijective, it follows that

$$E\left[|\Lambda| \cdot G \circ T\right] = E[G] \tag{3.4}$$

where

$$\begin{aligned}\Lambda(w) &= \det\left(I_{\mathbb{R}^n} + \left(\frac{\partial\psi_i(w(e_1), \cdots, w(e_n))}{\partial x_j}\right)\right) \cdot \\ &\cdot \exp\left(-\sum_1^n \psi_i(w(e_1), \cdots, w(e_n)) \cdot w(e_i) - \frac{1}{2}\sum_1^n \psi_i^2\right). \end{aligned} \tag{3.5}$$

Hence

$$E\left[\mathbf{1}_A \cdot |\Lambda|\right] = \nu(A)$$

is the Girsanov measure. Note that the first sum in the exponent is an Ogawa integral (*cf.* equation (2.3)). Denoting

$$f(\cdot) = \int_0^\cdot f_s'(w)ds = \sum_i \psi_i(w(e_1), \cdots, w(e_n)) \cdot e_i(\cdot)$$

and denoting by λ_i the eigenvalues of the $(n \times n)$ matrix $\partial\psi_i/\partial x_j$, equation (3.5) can be rewritten as

$$\Lambda(w) = \prod_1^n (1 + \lambda_i) \exp -\delta \circ f - \frac{1}{2}\int_0^1 f_s^2(w)ds \tag{3.6}$$

$$= \det{}_1(I_H + \nabla f) \exp -\delta \circ f - \frac{1}{2}\int_0^1 f_s^2(w)ds \tag{3.7}$$

where $\det_1(I_H + \nabla f)$ is the Fredholm determinant of $(I_H + \nabla(\sum_1^n \psi_i(w(e_1), \cdots, w(e_n))e_i)$, i.e., the Fredholm determinant of $\left(I_H + \sum_{i,j=1}^n \frac{\partial \psi_i}{\partial x_j} e_j \otimes e_i\right)$.

Many papers were written in the period from 1949 till 1974 devoted to proving the validity of equation (3.4) with Λ as given by (3.7) for the infinite dimensional classical and abstract Wiener spaces. Two difficulties stood in the way of such an extension. The first being the fact that the Fredholm determinant of $(I_H + K)$ where K is a Hilbert-Schmidt operator on H, may not exist since $\det_1(1 + K) = \prod_1^\infty (1 + \lambda_i)$ where λ_i are the eigenvalues of K and the product may not converge or the convergence may depend on the order of λ_i. In order to assure the existence of the Fredholm determinant of $(I_H + K)$, K has to be of trace class and this is a strong restriction. The second serious difficulty is the Ogawa integral appearing in the exponent since this integral is not a closable operation and strong conditions are needed in order to assume its existence.

It was Ramer who pointed out in his 1974 paper [13] that equation (3.7) is the 'wrong' prototype. Following the 1965 paper of L. Shepp [14], dealing with the absolute continuity of Gaussian measures with respect to the Wiener measure, Ramer noticed that the right prototype for the change of variables formula induced by a bijective transformation is obtained by first rewriting (3.5), (3.6) as

$$\Lambda(w) = \left(\prod_1^n (1 + \lambda_i) e^{-\lambda_i}\right) \exp\left(-\sum_1^n \psi_i(w(e_1), \cdots, w(e_n)) w(e_i) + \sum_1^n \lambda_i - \frac{1}{2}\sum_1^n \psi_i^2\right). \tag{3.8}$$

Note that what is achieved by multiplying and dividing by $\exp - \sum \lambda_i$ is, (a) the Fredholm determinant becomes a Carleman-Fredholm determinant which exists for all Hilbert-Schmidt operators and (b) since $\sum \lambda_i$ is the trace of the $n \times n$ matrix $\partial\psi_i/\partial x_j$, the first two sums in the exponent of (3.8) can be written as a Skorohod or Ito-Ramer integral since by (2.2):

$$\Lambda(w) = \det_2(1 + \nabla f) \cdot \exp\left(-\sum \psi_i \cdot w(e_i) + \operatorname{trace} \partial\psi_i/\partial x_j - \frac{1}{2}\sum \psi_i^2\right) \tag{3.9}$$

hence it can be written in short as

$$\Lambda = \det_2(1 + \nabla f) \exp[-\delta f - \frac{1}{2}|f|_H^2] \tag{3.10}$$

where the elements $\det_1$ and $\delta \circ f$ are replaced by $\det_2$ and δf which exists under considerably weaker assumptions. Note that Ramer's paper appeared in 1974, Skorohod's paper introducing the Skorohod integral appeared in 1975, the Malliavin calculus made its appearance in 1975 but the fact that the Skorohod integral is the dual to the gradient was shown by Gaveau and Trauber in 1982. Ramer's paper showed very convincingly that the right the end of what may be called "the romantic period" and starts "the modern period" of research on this subject. The results of Ramer required some strong continuity assumptions, his work was considerably improved by Kusuoka [10]. The main result of Kusuoka is the following:

Theorem 3.1. **([10])** *: Let $F(w) \in H - C^1$. Further assume that $Tw = w + F(w)$ is bijective and $(I_H + \nabla F)$ is a.s. invertible, then*

$$E\left[g \circ T \cdot |\Lambda|\right] = E[g].$$

The results of [10] were generalized by Üstünel and Zakai [18] in two directions; first, the shifts $T = w + F(w)$ were not required to be invertible and the condition that $F(w)$ be $H - C^1$ was replaced by the following weaker condition.

Definition 3.1. *An H-valued random variable is said to be $(H - C^1)_{\text{loc}}$ if there exists a random variable $q(w) > 0$ a.s. such that the map $h \mapsto F(w + h)$ is continuously Fréchet differentiable for all $h \in H$ satisfying $|h| < q(w)$.*

It was shown in [18] that if F is $(H - C^1)_{\text{loc}}$ then $F \in \mathbb{D}^{\text{loc}}_{\infty,1}(H)$. The result of [18] is the following

Theorem 3.2. *Let $F : W \to H$ be an $H - C^1_{\text{loc}}$ map, $Tw = w + F(w)$. Let M denote the set*

$$M = \{w : det_2(I_H + \nabla F(w)) \neq 0\}$$

or, what is the same, M is the set on which $I_H + \nabla F$ is invertible. Then there exists a measurable partition of $(M_n;\ n = 1, 2, \cdots)$ of M and a sequence of shifts $(T_n; n = 1, 2, \cdots)$ with $T_n w = w + F_n(w)$, $F_n \in \mathbb{D}^{\text{loc}}_{p,1}$ for some $p > 1$ such that, for each n, $T_n = T$ almost surely on M_n and $T_n : W \to W$ is bijective. Moreover

$$E[g \circ T_n |\Lambda_n|] = E[g],$$

for any $g \in C_b(W)$. Consequently

(i) For almost all w, the cardinal of the set $T^{-1}\{w\} \cap M$, denoted by $N(w, M)$ is at most countably infinite.

(ii) For any $g, \rho \in C_b^+(W)$, we have

$$E[g \circ T|\Lambda|] = E[g\,.N(w, M)],$$

and

$$E[g(Tw)\rho(w)|\Lambda|] = E\left[g(w) \cdot \sum_{\theta \in T^{-1}\{w\} \cap M} \rho(\theta)\right].$$

(iii) $(\mu|_M) \circ T^{-1} = T^*(\mu|_M) \ll \mu$ *with*

$$\frac{dT^*(\mu|_M)}{d\mu}(w) = \sum_{\theta \in T^{-1}\{w\} \cap M} \frac{1}{|\Lambda_F(\theta)|}.$$

Theorem 3.2 was proved in [18] using the decomposition technique developed in [10] and the following result

Theorem 3.3. **([18])** *Let* $F : W \to H$ *be a measurable map belonging to* $\mathbb{D}_{p,1}(H)$ *for some* $p > 1$. *Assume that there exist constants* c, d *(with* $c > 1$*) such that for almost every* $w \in W$

$$\| \nabla F(w) \| \leq c < 1$$

and

$$\| \nabla F(w) \|_2 \leq d < \infty$$

where $\| \cdot \|$ *denotes the operator norm and* $\| \cdot \|_2 = \| \cdot \|_{H \otimes H}$ *denotes the Hilbert-Schmidt (or* $H \otimes H$*) norm (in other words, for almost all* $w \in W$, $\| F(w+h) - F(w) \|_H \leq c \| h \|_H$ *for all* $h \in H$ *where* c *is a constant,* $c < 1$ *and* $\nabla F \in L^\infty(\mu, H \otimes H)$*). Then:*

(a) Almost surely $w \mapsto T(w) = w + F(w)$ *is bijective.*

(b) the measures μ *and* $T^*\mu$ *are mutually absolutely continuous.*

(c) $E[f] = E[f \circ T \cdot |\Lambda_F|]$
for all bounded and measurable f *on* W *and in particular* $E[|\Lambda_F|] = 1$.

Theorem 3.3 extends previous results ([10, 2, 3]) and its proof is based on the result of [15] (*cf.* also [16]) that $\| \nabla F \|_2 \leq d$ implies that $E \exp \lambda |F|^2 < \infty$ for all $\lambda < 1/2d^2$.

A result similar to that of theorem 3.3 for the case where Tw is a monotone shift $((T(w+h) - T(w)), h)_H > 0$ a.s. has recently been derived in [20].

At the beginning of this section we referred to the problem of the change of variables formula for $Tw = w + F(w)$ as the 'central model'. Another direction of active research, considered here, is the following: Assume that $F(w)$ is parameterized by some parameter $\alpha \in [0,1]$, $T_\alpha w = w + F_\alpha(w)$, where $F_{\alpha=0}(w) = 0$ and T_α is considered as a flow ([5, 3, 17, 7]).

The reader must have noticed that we have not mentioned the celebrated Girsanov, or Cameron-Martin-Maruyama-Girsanov, theorem. This is not because of lack of respect to this result, it would be difficult to overestimate the importance of this theorem. The Girsanov theorem and its extension to martingale setup play a most important role in both the theory (e.g., weak solutions to stochastic differential equations and extending to semimartingales the results known for quasimartingales) and to applications (e.g. non linear filtering and stochastic control theory) of stochastic processes. The success of the Girsanov theorem was perhaps an important spur for the derivation of analogous results for the non-adapted case. The reason for not mentioning it here was that the techniques are quite different. For some considerations related to or motivated by Girsanov theorem and, in particular, for an explanation why the multiplicity $N(w, H)$ and $\det_2$ do not appear in the Girsanov theorem, *cf.* [12] and [22].

IV. The change of variables formula

Theorem 4.1. *Let $F \in \mathbb{D}^{\mathrm{loc}}_{p,1}(H)$ for some $p > 1$. Suppose that there exists a sequence of measurable sets B_m such that $\mu(\cup B_m) = 1$, and a sequence of $(H - C^1)_{\mathrm{loc}}$ random H valued functions F_m such that*

$$\mathbf{1}_{B_m}(w) \cdot (F(w) - F_m(w)) = 0 \quad a.s.$$

Let $M = \{w : \det_2(I_H + \nabla F) \neq 0\}$. Then:

(i) The cardinal of the set $T^{-1}\{w\} \cap M$, denoted $N(w, M)$ is, at most countably infinite.

(ii) For any positive measurable bounded real random variables ρ and g

$$E[\rho(Tw)g(w) \cdot |\Lambda_F|] = E\left[\rho(w) \sum_{\theta \in T^{-1}\{w\} \cap M} g(\theta)\right]$$

in the sense that if one side is finite, so is the other side and equality holds.

(iii) $T^*(\mu|_M) \ll \mu$ and

$$\frac{dT^*(\mu|_M)}{d\mu}(w) = \sum_{\theta \in T^{-1}\{w\} \cap M} \frac{1}{|\Lambda_F(\theta)|}$$

The proof of this theorem follows along the same lines as the proof of theorem 3.2 i.e. the decomposition technique developed in [10] and theorem 3.3, it will be given after the following example which presents a case which is covered by theorem 4.1 but is not covered by previous results.

Let r_n denote the rationals in $(0, 1)$ arranged in some order and $\eta(x) = \exp -|x|$. Set

$$\theta(x) = \sum_{n=1}^{\infty} 2^{-n} \eta(x - r_n) .$$

The function $\theta(x)$ is Lipschitz with Lipschitz constant 1 and is non differentiable on all the rationals x in $[0, 1]$. Let $e_i, i = 1, 2, \cdots$ be a complete orthonormal base on H and set

$$F(w) = \sum_{i=1}^{\infty} 2^{-i} \theta(\delta e_i) \cdot e_i \tag{4.1}$$

then $F(w+h)$ is Lipschitz in h. Note first that even the case where $F(w) = \theta(\delta e_1) \cdot e_1$ is not covered by the results of the previous section, however it can be deduced from the finite dimensional Federer formula (1.1) or (1.2). Returning to $F(w)$ as defined by (4.1), let ϵ_n, $\epsilon > 0$ and $\Sigma \epsilon_n \leq \epsilon$. Let a_n denote a subset of $(-\infty, \infty)$ in which $\theta(\cdot)$ is C^1 and Lebesgue measure of a_n^c is bounded by ϵ_n. Therefore $F(w)$ is $H - C^1$ on $\{w : \delta e_i \in a_n,\, i \in \mathrm{N}\}$ and the conditions of theorem (4.1) are satisfied.

In order to prove theorem 4.1 we prepare the following:

Proposition 4.1. *Under the assumptions of the theorem, there exists a measurable partition* $M_{m,n}$ $m, n = 1, 2, \cdots$, *of* M *and shifts* $T_{m,n}w = w + F_{m,n}, F_n \in \mathbb{D}^{\mathrm{loc}}_{p,1}(H)$ *for some* $p > 0$ *and such that for each* m *and* n, $T_{m,n} = T$ *on* $M_{m,n}$ *a.s. and the* $T_{m,n} : W \to W$ *are bijective. Moreover*

$$E\left[\rho(w) g(T_{m,n}w) \cdot |\Lambda_{m,n}(w)|\right] = E\left[g(w)\rho(T_{m,n}^{-1}w)\right] . \tag{4.2}$$

for all bounded and measurable ρ *and* g.

Proof of the proposition: (*cf.* p. 495 of [18]] for a heuristic outline of the proof):

Let $e_i, i = 1, 2, \cdots$ be a complete orthonormal basis of H. Let $\lambda = \{\lambda_{i,j},\, 1 \leq i, j \leq n\}$ be a real valued $n \times n$ matrix such that $I_{\mathbb{R}^n} + \lambda$ is invertible. Set $T_\lambda w = w + F_\lambda(w)$ where

$$F_\lambda(w) = \sum_{i,j=1}^{n} \lambda_{i,j} \cdot \delta e_j \cdot e_i$$

and note that $\nabla F_\lambda = \sum_{i,j} \lambda_{i,j} e_j \otimes e_i$ is deterministic, $T_\lambda^* \mu \sim \mu$ and T_λ is bijective.

Let $\gamma(\lambda)$ be the inverse of the operator norm of $(I_{\mathbb{R}^n} + \lambda)^{-1}$ and define

$$A(m,n,v,\lambda) = \left\{ w : w \in B_m, q_m(w) > \frac{4}{n}, \sup_{\|h\| < \lambda} \| F_m(w+h) - F_\lambda(w+h) - v \| \leq a\frac{\gamma(\lambda)}{n} \right.$$

$$\left. \text{and} \sup_{\|h\| < \frac{1}{n}} \| \nabla F_m(w+h) - \nabla F_\lambda \|_{H-S} \leq a\gamma(\lambda) \right\},$$

where a is a constant to be chosen below and q_m is a random variable which is the radius of the set of $H - C^1$-property of F_m. Let $G(m,n,v,\lambda)$ be a σ-compact modification of $A(m,n,v,\lambda) \cap M$. Let $\rho_A(w)$ be as defined by equation (2.6) and let $\varphi \in C_o^\infty(\mathbb{R})$, $|\varphi(x)| \leq 1$, $|\varphi(x)| = 1$ for $|x| < \frac{1}{3}$, $\varphi(x) = 0$, for $|x| > 2/3$ and $\| \varphi' \|_\infty \leq 4$. Assume, now, that v and the elements of λ are rational and λ is non-singular. The collection of such four-tuples (m,n,v,λ) is countable and $G(m,n,v,\lambda)$ will be denoted by $G_\nu, \nu = 1, 2, \cdots$. Set $F_\nu(w) = v + \varphi(n\rho_{G_\nu}(w))[F(w) - F_\lambda(w) - v] + F_\lambda(w)$ and note that for $w \in G_\nu$, $\rho_{G_\nu}(w) = 0$

and $F_\nu(w) = F(w)$. On the other hand, setting

$$\begin{aligned} T_\nu(w) &= w + F_\nu(w), \\ T_\lambda(w) &= w + F_\lambda(w) \\ T_c(w) &= w + \varphi(n\rho_{G_\nu}(T_\lambda^{-1}w))[F(T_\lambda^{-1}w) - F_\lambda(T_\lambda^{-1}w) - v] \\ T_v(w) &= w + v \end{aligned}$$

it is easily verified that $T_v \circ T_c \circ T_\lambda = T_\nu$. Now, T_λ and T_v are bijective and from the definition of G_ν for $a < 1/3 \parallel \nabla F_c(w) \parallel_2 < 1$. Therefore by theorem 3.3, T_c is also bijective and consequently T_ν is bijective. Moreover, $T_\nu^* \mu \sim \mu$ since T_λ, T_v and T_c induce equivalent transformations of measure.

Now, let $T_i, i = 1, 2, 3$ be measurable transformation of W to W, $T_i^* \mu \ll \mu$ and

$$E\left[\eta_i(w) f(T_i(w))\right] = E[f]$$

then

$$\begin{aligned} E\left[g \circ T_3 \circ T_2 \circ T_1 \cdot \eta_3 \circ T_2 \circ T_1 \cdot \eta_2 \circ T_1 \cdot \eta_1\right] &= E\left[g(T_3 \circ T_2 w)\eta_3(T_2 w)\eta_2(w)\right] \\ &= E\left[g(T_3 w)\eta_3(w)\right] \\ &= E[g] . \end{aligned}$$

Therefore

$$Eg(w) = Eg(T_\nu w) \cdot \Lambda_{F_v}(T_c \circ T_\lambda w) \cdot \Lambda_{F_c}(T_\lambda(w))\Lambda_{F_\lambda}(w) .$$

By a direct calculation using

$$\det{}_2(1 + a)(1 + b) = \det{}_2(I + a) \cdot \det{}_2(I + b) \cdot \exp - \text{trace } (ab)$$

and for $Tw = w + f(w)$

$$(\delta G(w)) \circ Tw = \delta G(Tw) - (G(Tw), f)_H - \text{trace } ((\nabla G \circ Tw) \cdot \nabla f) ,$$

we get that

$$\Lambda_{F_v}(T_c \circ T_\lambda w) \circ \Lambda_{F_c}(T_\lambda w) \cdot \Lambda_{F_\lambda}(w) = \Lambda_{F_\nu}(w)$$

(for details *cf.* lemma 6.1 of [10] or the lemma of [11]) and (4.2) follows. □

Proof of theorem 4.1: Since $\cup_{m,n} M_{m,n} = M$ we may assume without loss of generality that $M_{m,n}$ are disjoint and

$$\begin{aligned} T^{-1}\{w\} \cap M &= \{\theta \in M : T(\theta) = w\} \\ &= \bigcup_{m,n} \{\theta \in M_{m,n} : T_{m,n}\theta = w\} . \end{aligned}$$

Since the shifts are bijective, the cardinality of the above set is at most countably infinite. Now

$$\begin{aligned} E\left[\rho(w)g(Tw)\cdot|\Lambda_F|\right] &= \sum_{m,n} E\mathbf{1}_{M_{m,n}}(w)\rho(w)g(T_{m,n}w)\cdot|\Lambda_{m,n}(w)| \\ &= \sum_{m,n} E\mathbf{1}_{T_{m,n}M_{m,n}}(T_{m,n}w)\cdot\rho(w)g(T_{m,n}w)|\Lambda_{m,n}|\,. \end{aligned}$$

Applying equation (4.2) yields

$$\begin{aligned} E\left[\rho(w)g(Tw)|\Lambda_F|\right] &= \sum_{m,n} E\mathbf{1}_{T_{m,n}M_{m,n}}(w)\rho(T_{m,n}^{-1}w)g(w) \\ &= E\left[g(w)\sum_{\theta\in T^{-1}\{w\}\cap M}\rho(\theta)\right] \end{aligned}$$

which proves (ii) and (iii) follows by a similar argument (cf. [18]). □

References

[1] Bouleau, N., Hirsch, F.: Dirichlet Forms and Analysis on Wiener Space. De Gruyter Studies in Math., Vol. 14, Berlin, New York; De Gruyter 1991.

[2] Buckdahn, R.: Anticipative Girsanov transforms. Probab. Theory Relat. Fields **89**, 211–238 (1991).

[3] Buckdahn, R.: Anticipative Girsanov transformations and Skorohod stochastic differential equations. Seminarbericht No. 92–2 (1992).

[4] Cameron, R.H., Martin, W.T.: The transformation of Wiener integrals by nonlinear transformations. Trans. Amer. Math. Soc. **66**, 253–283 (1949).

[5] Cruzeiro, A.B.: Equations differentielles sur l'espace de Wiener et formules de Cameron-Martin nonlinéaires. J. Functional Anal., Vol. 52, 335–347, (1984).

[6] Enchev, O.: Nonlinear transformations on the Wiener space. The Annals of Probability, Vol. 21, 2169–2188 (1993).

[7] Enchev, O., D.W. Stroock: Anticipative diffusions and related change of measure. J. Functional Anal. **116**, 449–477 (1993).

[8] Enchev, O., D.W. Stroock: Rademacher's theorem for Wiener functionals. Annals of Probability **21** 25–33 (1993).

[9] Federer, H.: Geometric Measure Theory, Springer-Verlag, 1969.

[10] Kusuoka, S.: The nonlinear transformation of Gaussian measures on Banach space and its absolute continuity, I.J. Fac. Sci. Univ. Tokyo, Sect. IA, math. **29**, 567–598 (1982).

[11] Nualart, D.: Markov fields and transformations of the Wiener measure. The Proceedings of the Oslo-Silivri Conference on stochastic analysis and related topics, p. 45–88, Gordon and Breach Stochastic Monograph, Vol. 8, T. Lindstrom,B. Oksendal and A. S. Üstünel (editors).

[12] Nualart, D., Zakai, M.: Generalized stochastic integrals and the Malliavin calculus. Probab. Theory Relat. Fields **73**, 255–280 (1986).

[13] Ramer, R.: On nonlinear transformations of Gaussian measures. J. Funct. Anal. **15**, 166–187 (1974).

[14] Shepp, L.A.: Radon Nikodym derivatives of Gaussian measures. The Annals of Math. Stat. **37**, 321–354 (1966).

[15] Üstünel, A.S.: Integrabilité exponentielle de fonctionnelles de Wiener. CRAS, Paris, Série I **305**, 279-282 (1992).

[16] Üstünel, A.S.: An Introduction to Analysis on Wiener Space. Lecture Notes in Math. Vol. 1610, Springer, 1995.

[17] Üstünel, A.S., Zakai, M.: Transformations of Wiener measure under anticipative flows. Probab. Theory Relat. Fields **93**, 91–136 (1992).

[18] Üstünel, A.S., Zakai, M.: Transformation of the Wiener measure under non-invertible shifts, Prob. Theory Related Fields, **99**, 485–500, 1994.

[19] Üstünel, A.S., Zakai, M.: Extensions of Lipschitz functions on Wiener space. To appear.

[20] Üstünel, A.S., Zakai, M.: Measures induced on Wiener space by monotone shifts. To appear.

[21] Üstünel, A.S., Zakai, M.: Applications of the degree theorem to absolute continuity on Wiener space. Probab. Theory Relat. Fields, 95, 509-520, 1993.

[22] Zakai, M., Zeitouni, O.: When does the Ramer formula look like the Girsanov formula? The Annals of Probability, Vol., 20, pp. 1436–1440, 1992.

A.S. Üstünel, ENST, Dépt. Réseaux, 46 rue Barrault, 75013, Paris, France.

M. Zakai, Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel.