JOURNAL OF FUNCTIONAL ANALYSIS **149**, 226–244 (1997)
ARTICLE NO. FU963092

The Sard Inequality on Wiener Space

A. S. Üstünel

Département Réseaux, ENST, 46, rue Barrault, 75013 Paris, France

and

M. Zakai

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

Received November 20, 1996; accepted December 18, 1996

Let $T(w) = w + u(w)$ be a Cameron–Martin perturbation of the identity. The formal infinite dimensional extension of the Sard inequality,

$$\mu(TA) \leqslant \int_A |\Lambda| \, d\mu,$$

is shown to hold and applications to absolute continuity on Wiener space are presented. © 1997 Academic Press

I. INTRODUCTION

The Sard lemma on $\mathbb{R}^n$ states that if $D \subset \mathbb{R}^n$ is open, T is a continuously differentiable function from D to $\mathbb{R}^n$ and $E_0 = \{x \in D\colon \det \nabla T(x) = 0\}$, then the Lebesgue measure of E_0 is zero. This result is useful in many applications as it often avoids the need to consider what happens on E_0. In [9], J. T. Schwartz presented a generalization of this result:

Theorem 1.1 [9]. *Let D and T be as above and let $J(x)$ denote the Jacobian determinant of T at x; also, let E be a measurable subset of D, then $T(E)$ is measurable and*

$$\int_{\mathbb{R}^n} \mathbf{1}_{\mathrm{TE}}(x)\, dx \leqslant \int_{\mathbb{R}^n} \mathbf{1}_{\mathrm{E}}(x)\, |J(x)|\, dx. \tag{1.1}$$

0022-1236/97 $25.00

In order to represent this result for the case where the Lebesgue measure is replaced with the standard Gaussian measure on $\mathbb{R}^n$, note that if $\psi(x)$ is measurable and nonnegative, then (1.1) implies that

$$\int_{\mathbb{R}^n} \psi(x)\, \mathbf{1}_{\mathrm{TE}}(x)\, dx \leqslant \int_{\mathbb{R}^n} \psi(Tx)\, \mathbf{1}_{\mathrm{E}}(x)\, |J(x)|\, dx. \tag{1.2}$$

In particular, setting

$$\psi(x) = (2\pi)^{-n/2} \exp -|x|^2/2$$
$$\mu(dx) = \psi(x)\, dx$$
$$Tx = x + f(x)$$

and

$$\Lambda(x) = J(x) \exp(\langle x, f(x)\rangle - \tfrac{1}{2}\, |f(x)|^2)$$

yields

$$\int_{\mathbb{R}^n} \mathbf{1}_{\mathrm{TE}}(x)\, d\mu(x) \leqslant \int_{\mathbb{R}^n} \mathbf{1}_{\mathrm{E}}(x)\, |\Lambda(x)|\, \mu(dx)$$

or

$$\mu(\mathrm{TE}) \leqslant \int_{\mathrm{E}} |\Lambda(x)|\, \mu(dx). \tag{1.3}$$

An extension of (1.1) where the condition of T being continuously differentiable is replaced by a weaker assumption is a part of Federer's area theorem for $m = n$, (Theorem 3.2.3 of [3]). Cf., also, Theorem 5.6 of [4].

An infinite dimensional extension of Sard's lemma (with zero Lebesgue measure replaced by first category) was presented by Smale [10]. In the context of Wiener space, Kusuoka presented a Sard-type result (Theorem 8.1 of [6]) and indicated the validity of the Sard lemma under certain restrictions in [7], Cf. also Getzler [5]. The purpose of this paper is to present detailed proofs of the measurability of the forward images of Borel sets under the perturbation of identity maps, the Sard inequality and some applications of these results. Some of these results are applied in [15] to degree theory on the Wiener space.

In the next section we will summarize some definitions and results of stochastic analysis that will be needed in the paper. The measurability problem will be discussed in Section 3. Section 4 is devoted to the Sard inequality. The strategy of the proof follows Smale [10]: T is shown to be representable locally as $T = T_S \circ T_G$ where T_G is invertible and T_S is finite dimensional. This is done in Lemma 4.1 following the technique of Kusuoka [6]. It is then shown, Lemma 4.2, that the Sard inequality for T follows from the application of the finite dimensional Sard inequality to T_S. Section 5 is devoted to a certain extension of the Sard inequality and the infinite dimensional extension of (1.2) is also given there. Some applications to the question of absolute continuity are discussed in Section 6.

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 $\mathscr{X}$ be a real separable Hilbert space; smooth, $\mathscr{X}$-valued functionals on (W, H, μ) are functionals of the form

$$a(w) = \sum_1^N \eta_i(\langle h_1, w \rangle, ..., \langle h_m, w \rangle)\, x_i$$

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

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

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

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

The gradient $\nabla: \mathbb{D}_{p,k}(\mathscr{X}) \to \mathbb{D}_{p,k-1}(\mathscr{X} \otimes H)$ 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 $\mathscr{X}$ and $(\nabla a)^*$ will denote the adjoint of ∇a and is a mapping from $\mathscr{X}^*$ 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 $\{e_i, i=1,2,...\}$ we have

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

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, $\bigcup_n A_n = W$ 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 [6] 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, $i=1,2,...$ be the sequence of eigen-values of K repeated according to their multiplicity. The Carleman–Fredholm determinant of K is defined as:

$$\det{}_2(I+K) = \prod_{i=1}^{\infty} (1+\lambda_i)\, e^{-\lambda_i} \tag{2.3}$$

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

$$\Lambda_F(w) = \det{}_2(I+\nabla F) \exp(-\delta F - \tfrac{1}{2} \|F\|^2_H). \tag{2.4}$$

The following lemma will be needed in section IV:

Lemma 2.1. *Let F_1, F_2, F_3 belong to $\mathbb{D}^{\mathrm{loc}}_{p,1}(H)$ 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 [6] or [8] and uses the fact that for $T(w) = w + u(w)$

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

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 $\mu(G) = \mu(A)$.

With every measurable subset A of W we associate the 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

$$\rho(w, A) = \inf_{h \in H} \{ \|h\|_H : w + h \in A \} \tag{2.5}$$

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

Clearly, $\rho(w, A) = 0$ if $w \in A$, moreover [6], $\rho(w, A)$ is a measurable random variable and:

(i) If $A \subset B$, then $\rho(w, A) \geqslant \rho(w, B)$

(ii) $|\rho(w, A) - \rho(w+h, A)| \leqslant \|h\|_H$.

(iii) $A_n \nearrow A$ implies $\rho(w, A_n) \searrow \rho(w, A)$.

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

$$\begin{aligned} \|\nabla \varphi(\rho(w, G))\|_H &\leqslant \|\varphi'\|_\infty \cdot \mathbf{1}_{\{\varphi(\rho G) \neq 0\}} \\ &\leqslant \|\varphi'\|_\infty . \end{aligned} \tag{2.6}$$

(v) Let $Z = \{w : \rho(w, A) < \infty\}$. 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 $\nabla \mathbf{1}_Z = 0$, hence $\mathbf{1}_Z$ is almost surely a constant. Consequently $\mu(Z) = 1$ if $\mu(A) > 0$.

The following result will be needed in Section IV, cf. [6, 12].

THEOREM 2.1. *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 $w \in W$*

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

and

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

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 $w \in W$, $\|F(w+h) - F(w)\|_H \leqslant c \, \|h\|_H$ 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, the inverse T^{-1} satisfies $T^{-1}w = w + L(w)$ where $\|L(w)\|_H \leqslant \|F(w)\|_H / 1 - c$ and $\|\nabla L\|_2 \leqslant d/1 - c$.*

(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 $E[|\Lambda_F|] = 1$.*

DEFINITION 2.2. Let $u(w)$ be an H-valued random variable

(a) $u(w)$ is said to be an $H-C$ map if, for almost all $w \in W$, $h \mapsto u(w+h)$ is a continuous function of $h \in H$.

(b) $u(w)$ is said to be $H-C^1$ if it is $H-C$ and for almost all $w \in W$, $h \mapsto u(w+h)$ is continuously Fréchet differentiable on H.

(c) $u(w)$ is said to be "locally $H-C^1$" if there exists an almost surely strictly positive random variable ρ such that $h \mapsto u(w+h)$ is C^1 on the set $\{h \in H\colon |h| < \rho(w)\}$.

(d) $u(w)$ will be said to be $\eta - H - C^1$, if there exists a non-negative random variable $\eta(w)$ such that $\mu\{\eta(w) > 0\} > 0$ and for all $w \in Q = \{w\colon \eta(w) > 0\}$, $u(w+h)$ is Fréchet differentiable on $\{h \in H, \|h\|_H < \eta(w)\}$.

III. THE MEASURABILITY OF THE FORWARD IMAGE

THEOREM 3.1. *Suppose that $u\colon W \to H$ is a measurable map. Then for any measurable $A \subset W$, $(I_W + u)(A) = T(A)$ is in the universally completed Borel sigma algebra of W.*

Proof. If $w \in T(A)$, then $w = \theta + u(\theta)$ where $\theta \in A$. Otherwise stated, setting $\theta = w + h$, h satisfies

$$0 = h + u(w+h)$$

and

$$w + h \in A.$$

Let $\Gamma(w)$ be the multifunction taking values in subsets of H:

$$\Gamma(w) = \{h\colon h + u(w+h) = 0 \text{ and } (w+h) \in A\}.$$

Then

$$T(A) = \{w \in W\colon \Gamma(w) \neq \phi\} = \pi_W(G(\Gamma)),$$

where $G(\Gamma)$ is the graph of Γ: $G(\Gamma) = \{(h, w)\colon h \in \Gamma(w)\}$ and $\pi_W(h, w) = w$. Since $(w, h) \mapsto w + h$ is measurable, $G(\Gamma)$ is measurable in $W \times H$ hence $\pi_W G(\Gamma)$ is universally measurable (c.f. Theorem 23, p. 75 of [1]).

IV. THE SARD INEQUALITY

The following result is the infinite dimensional version of the Sard inequality which implies the Sard lemma.

THEOREM 4.1. *Suppose that* $u: W \to H$ *is a measurable map in some* $\mathbb{D}_{p,1}(H)$ *and is* $\eta - H - C^1$, *i.e. there exists a non-negative random variable* η, *with* $\mu(Q) = \mu\{\eta > 0\} > 0$ *and the map* $h \mapsto u(w+h)$ *is continuously Fréchet differentiable on the random open ball* $\{h \in H: \|h\|_H < \eta(w)\}$. *Then we have, for any* $A \in \mathscr{B}(W)$,

$$\mu(T(A \cap Q)) \leqslant \int_{A \cap Q} |\Lambda_u| \, d\mu.$$

The proof of the theorem will follow from the following two lemmas.

LEMMA 4.1. *Under the assumptions of Theorem* 4.1, *there exists a countable cover* $Q_{m,n}$ *of* Q *and two sequences in* $\mathbb{D}_{p,1}(H)$, *denoted by* $K_{m,n}(w)$ *and* $S_{m,n}(w)$ *such that*

1. $$\|\nabla K_{m,n}\|_2 \leqslant \lambda_{m,n} < 1$$

for almost all $w \in W$, *where* $\|\cdot\|_2$ *denotes the Hilbert–Schmidt norm.*

2. $S_{m,n}(w)$ *is finite dimensional on* $Q_{m,n}$, *i.e. there exists a finite dimensional subspace of* H, *say* $H_{m,n}$, *such that* $S_{m,n}(w) \in H_{m,n}$ *for all* $w \in Q_{m,n}$.

3. $T = T_{S_{m,n}} \circ T^{-1}_{K_{m,n}}$.

Proof of Lemma 4.1. Let $(\pi_n; n \in \mathbb{N})$ be a sequence of orthogonal projections of H increasing to I_H. Let α be a fixed positive number (to be specified later), set

$$Q_{m,n} = \left\{ w \in W: \|\nabla u(w+h) - \nabla u(w)\|_2 \leqslant \alpha, \text{ for all } |h|_H \leqslant \frac{1}{m} \right\}$$

$$\cap \left\{ w \in W: |\pi_n^\perp u(w)|_H < \frac{\alpha}{m}, \ \|\pi_n^\perp \nabla u(w)\|_2 \leqslant \alpha, \right.$$

$$\left. \|\nabla u(w)\|_2 \leqslant m, \eta(w) > \frac{4}{m} \right\},$$

where $\|\cdot\|_2$ denotes the Hilbert-Schmidt norm. By the $H-C^1$-property, $(Q_{m,n}; n, m \in \mathbb{N})$ covers Q almost surely (here, if necessary, we add a negligible set to have equality everywhere instead of almost everywhere but we keep the same notation). Let us denote $Q_{m,n}$ by q. It is easy to see that for $w \in q$ and any $h \in H$, $\|h\|_H \leqslant 1/m$

$$\|\pi_n^\perp \nabla u(w+h)\|_2 \leqslant 2\alpha \tag{4.1}$$

and, assuming that $\alpha < 1$,

$$\begin{aligned}\|\pi_n^\perp u(w+h)\|_H &\leqslant \|\pi_n^\perp u(w)\|_H + \int_0^1 \|\pi_n^\perp \nabla u(w+th)\|_2 \cdot \|h\|_H \cdot dt \\ &\leqslant \frac{\alpha}{m} + \frac{2\alpha}{m} \leqslant \frac{3\alpha}{m}.\end{aligned} \tag{4.2}$$

Let φ be a smooth function on $\mathbb{R}$ such that $|\varphi(t)| \leqslant 1$ and $|\varphi'(t)| \leqslant 2$ for all $t \in \mathbb{R}$, furthermore assume that $\varphi(t) = 1$ on $|t| \leqslant 1/2$ and $\varphi(t) = 0$ on $|t| \geqslant 2$. Let $\rho(w, q) = \inf\{\|h\|_H : h \in H, w+h \in q\}$.

Set

$$g(w) = \varphi(m\rho(w, q))$$

and

$$G(w) = g(w)\, \pi_n^\perp u(w).$$

Therefore, if $g(w) \neq 0$, then $m \cdot \rho(w, q) < 1$, hence for some $w_0 \in q$, $\|w - w_0\|_H < 1/m$. Therefore, by (4.1) and (4.2), for all $w \in W$,

$$\|G(w)\|_H \leqslant \frac{3\alpha}{m} \tag{4.3}$$

and

$$\begin{aligned}\|\nabla G(w)\|_2 &\leqslant \|\nabla g \otimes \pi_n^\perp u\|_2 + \|g \cdot \nabla \pi_n^\perp u\|_2 \\ &\leqslant 2m \cdot \frac{3\alpha}{m} + 2\alpha = 8\alpha.\end{aligned} \tag{4.4}$$

Setting, now, $\alpha = 0.5 \cdot 10^{-2}$, it follows from Theorem 2.1 that $T_G = I_w + G$ is a.s. bijective. Let $E = T_G(q)$, then by the result of the previous section, E is measurable and for any w satisfying $\rho(w, E) \leqslant 1/3m$ there exists some

$w_0 \in q$, such that $w - T_G w_0 \in H$ and $\|w - T_G w_0\|_H < \frac{1+\varepsilon}{3m}$, $\varepsilon > 0$. Therefore, by (a) of Theorem (2.1) and (4.4)

$$\|T_G^{-1} w - w_0\|_H \leqslant \frac{\|w - T_G w_0\|_H}{1 - 8\alpha} \leqslant \frac{1}{2m}.$$

Hence, $\rho(T_G^{-1} w, q) < 1/2m$ and $\varphi(m \cdot \rho(T_G^{-1} w, q)) = 1$, i.e. $G(w) = \pi_n u(w)$ and consequently

$$(I + \pi_n^\perp u) \circ T_G^{-1} w = w \tag{4.5}$$

for any w such that $\rho(w, E) < 1/3m$ and in particular to any $w \in E$. Now set

$$\begin{aligned} -K(w) &= \varphi(8m\rho(w, E))(w - (I+G)^{-1} w) \\ &= \varphi(8m\rho(w, E))\, G((I+G)^{-1} w). \end{aligned} \tag{4.6}$$

Hence by Theorem 2.1 and (4.3)

$$\|K(w)\|_H \leqslant \frac{3\alpha}{m}$$

and

$$\begin{aligned} \|\nabla K\|_1 &\leqslant 16m\, \|G(w)\|_H + \|\nabla G \circ (I+G)^{-1} w\|_2 \cdot (1 + \|\nabla(I - (I-G)^{-1})\|_2) \\ &\leqslant \frac{48m\alpha}{m} + 8\alpha\left(1 + \frac{8\alpha}{1-8\alpha}\right) \\ &< 0.3. \end{aligned}$$

Setting $I_W + S = T \circ T_K$, i.e., $S(w) = K(w) + u(T_K(w))$, if $\rho(w, E) < 1/8m$ (in particular, if $w \in E$) then by (4.5), (4.6) $T_K(w) = T_G^{-1} w$ and

$$\begin{aligned} w &= (I_w + \pi_n^\perp u)\, T_K(w) \\ &= w + K(w) + \pi_n^\perp u(T_K(w)). \end{aligned}$$

Therefore

$$S(w) = -\pi_n^\perp u(T_K(w)) \qquad \text{and}$$

$$\begin{aligned} S(w) &= K(w) + u(T_K(w)) \\ &= (1 - \pi_n^\perp)\, u(T_K(w)) \\ &= \pi_n u(T_K(w)). \end{aligned}$$

Consequently, for $\rho(w, E) < 1/8m$, $S(w)$ is in a finite dimensional space. Setting $K = K_{m,n}$ and $S = S_{m,n}$ completes the proof of the lemma.

LEMMA 4.2. *Let A be any measurable subset of W and let $Q_{m,n}$ be as defined in Lemma* 4.1, *then*

$$\mu(T(A \cap Q_{m,n})) \leqslant \int_{A \cap Q_{m,n}} |\Lambda_u(w)| \, \mu(dw).$$

Proof. Let $\tilde{A} = A \cap Q_{m,n}$; $S = S_{m,n}$ and $K = K_{m,n}$ are as defined in Lemma 4.1. By Theorem 3.1, $T\tilde{A}$ is measurable. Set $E = T_G \tilde{A}$, then E is also measurable since T_G satisfies the conditions of Theorem 3.1. Now, $T_S = T \circ T_K$ on E, therefore by Lemma 2.1, $|\Lambda_S(w)| = |\Lambda(T_K(w)| \cdot |\Lambda_K(w)|$ on E. Let h_i, $i = 1, 2, ...$ be a C.O.N.B. on H and $\pi_{n,m}$ is the projection on $H_{m,n}$ defined in Lemma 4.1.

$$w = \{\delta h_1, \delta h_2, ...\}$$

$$w_a = \{\delta h_i, i \leqslant n\}$$

$$w_b = \{\delta h_i, i \geqslant n+1\}$$

$$w = w_a \oplus w_b,$$

where $w_a \oplus w_b$ denotes the concatenation of w_a with w_b.

Define $\mathscr{F}_a = \sigma\{\delta h_i, i \leqslant n\}$, $\mathscr{F}_b = \sigma\{\delta h_i, i \geqslant n+1\}$ and μ_a, μ_b the restriction of μ to $\mathscr{F}_a$ and $\mathscr{F}_b$ respectively. Then

$$EF(w) = \int_W F(w_a \oplus w_b) \, \mu_a(dw_a) \cdot \mu_b(dw_b).$$

Note that $\rho(w, A)$ is Lipschitz continuous (cf. property (ii) of $\rho(w, A)$). Consequently for all $w \in E$, $K(w+h)$ and $S(w+h)$ are Lipschitz continuous on $(w+h) \in Q_{m,n}$ and for any $(w_a \oplus w_b)$ in E, $S(w_a \oplus w_b)$ is Lipschitz continuous in the w_a variables. Now, the area theorem of Federer (cf. [3, p. 243, Theorem 3.2.3]), for a Lipschitz function $f\colon \mathbb{R}^n \to \mathbb{R}^n$ yields

$$\int_A J_m f(x) \, dx = \int_{\mathbb{R}^n} \text{Cardinality}(A \cap f^{-1}(y)) \, dy \geqslant \int_{\mathbb{R}^n} \mathbf{1}_{f(A)}(y) \, dy$$

which extends the Sard inequality to Lipschitz functions. Therefore, setting

$$\pi_n(w_a \oplus w_b) = w_a$$

we have

$$E(\mathbf{1}_{T_S E}(w) \mid \mathscr{F}_b) \leqslant \int_{E \cap \pi_n W} |\Lambda_S(w_a \oplus w_b)| \, \mu_a(dw_a).$$

Consequently

$$\begin{aligned} \mu(T_S E) &\leqslant \int_E |\Lambda_S(w)| \, \mu(dw) \\ &= \int_E |\Lambda_u(T_K(w))| \cdot |\Lambda_K(w)| \, \mu(dw) \\ &= \int_W [|\mathbf{1}_{\tilde{A}}(\cdot) \cdot \Lambda_u(\cdot)| \circ (T_K w)] \cdot |\Lambda_K(w)| \, \mu(dw). \end{aligned}$$

Applying part (c) of Theorem 2.1 to T_K yields

$$\mu(T_S E) \leqslant \int_{\tilde{A}} |\Lambda_u(w)| \cdot \mu(dw),$$

which completes the proof of the lemma, since $T_S E = T\tilde{A}$.

Turning to the proof of Theorem 4.1, cutting and pasting $Q_{m,n}$ to form a partition of $A \cap Q$ (keeping the same notation),

$$\begin{aligned} \mu(T(A \cap Q)) &= \mu(T(\cup Q_{m,n})) \\ &= \mu(\cup T(Q_{m,n})) \\ &\leqslant \Sigma \mu(T(Q_{m,n})) \\ &\leqslant \sum \int_{Q_{m,n}} |\Lambda_u| \, \mu(dw) \\ &= \int_{A \cap Q} |\Lambda_u| \, \mu(dw), \end{aligned}$$

which completes the proof of the theorem.

V. APPLICATION: THE CHANGE OF VARIABLES FORMULA

If in Theorem 4.1, the set Q has full measure then we have

$$\mu(T(A \cap Q)) \leqslant \int_A |\Lambda_u| \, d\mu,$$

we would like to have in this case that

$$\mu(T(A)) \leqslant \int_A |\Lambda_u| \, d\mu.$$

However, due to adding negligeable sets to A in the course of the proof of Lemma 4.1, this result is not true unless the things are reinterpreted as explained in the following extension of Theorem 5.2 of [12].

THEOREM 5.1. 1. *Suppose that* $u: W \to H$ *is locally in some* $\mathbb{D}_{p,1}(H)$ *and that it is* $\eta - H - C^1$ *with* $\mu(Q) = \mu\{\eta > 0\} > 0$. *Let* $T = I_W + u$. *For any positive, bounded, measurable functions* f *and* g *on* W, *we have*

$$E[f \circ T g \, \mathbf{1}_Q \, |\Lambda_u|] = E\left[f \sum_{y \in T^{-1}\{w\} \cap Q} g(y) \right],$$

where $\Lambda_u = \det_2(I_H + \nabla u) \exp[-\delta u - \frac{1}{2} |u|_H^2]$.

2. *Furthermore, if* u *is* $H - C^1_{\text{loc}}$, *then there exists a modification of* u, *denoted by* u' (*i.e.,* $\mu\{u = u'\} = 1$), *such that the corresponding shift* T' *satisfies*

$$E[f \circ T' g \, |\Lambda_{u'}|] = E\left[f \sum_{y \in T'^{-1}\{w\}} g(y) \right].$$

In particular, we have

$$\mu(T'(A)) \leqslant \int_A |\Lambda_{u'}| \, d\mu,$$

for any $A \in \mathscr{B}(W)$.

3. *If moreover* $Q + H \subset Q$, *then the restriction of* T *to the set* Q *satisfies the conclusion of* (2) *where* T' *is replaced by* $T|_Q$. *In other words we can replace* (W, H, μ) *by* (Q, H, μ) *and think of it as an abstract Wiener space on which it holds that*

$$\mu(T(A)) \leqslant \int_A |\Lambda_u| \, d\mu,$$

for any $A \in \mathscr{B}(Q)$, *where* $\mathscr{B}(Q)$ *denotes the trace of* $\mathscr{B}(W)$ *on* Q.

Proof. From Theorem 5.2 of [12], we have

$$E[f \circ T g \, \mathbf{1}_Q \, |\Lambda_u|] = E\left[f \sum_{y \in T^{-1}\{w\} \cap M \cap Q} g(y) \right].$$

Therefore, if $g=g'$ almost surely on Q then

$$\sum_{y\in T^{-1}\{w\}\cap Q\cap M} g(y)=\sum_{y\in T^{-1}\{w\}\cap Q\cap M} g'(y)$$

almost surely. Moreover, we have

$$E\left[f\sum_{y\in T^{-1}\{w\}\cap M\cap Q} g(y)\right]=E\left[f\,\mathbf{1}_{(T(M^c\cap Q))^c}\sum_{y\in T^{-1}\{w\}\cap Q} g(y)\right]$$

and the first part of the theorem follows from Theorem 4.1. For the second part, it suffices to define $u'(w)=\mathbf{1}_Q(w)\,u(w)$ and to note that $\mu(Q)=1$. Since $\mathbf{1}_{T'(A)}(w)\leqslant N'(w,A)$, where $N'(w,A)$ is the cardinal of the set $T'^{-1}\{w\}\cap A$, which is equal to $N(w,Q\cap A)$ almost surely, we have

$$\begin{aligned}\mu(T'(A))&\leqslant E[N'(w,A)]\\&=E[N(w,A\cap Q)]\\&\leqslant E[\mathbf{1}_A\,|\Lambda_u|]\\&=E[\mathbf{1}_A\,|\Lambda_{u'}|].\end{aligned}$$

The third claim follows from the fact that $T(Q)\subset Q$ whenever $Q+H\subset Q$ (note that in this case Q^c is a slim set).

Below we give the proof of the inequality (1.2) in the setting of the abstract Wiener space:

COROLLARY 5.1. *Let u be a $H-C^1_{\mathrm{loc}}$. Then there exists $u'=u$ almost surely and $T'=I_W+u'$ satisfies*

$$E[\psi\,\mathbf{1}_{T'(A)}]\leqslant E[\psi\circ T'\,\mathbf{1}_A\,|\Lambda_{u'}|],$$

for any $A\in\mathscr{B}(W)$ and $\psi\geqslant 0$ any measurable function on W. If u is $H-C^1$, then we can take $T=T'$ above provided that the triple (W,H,μ) is replaced by (Q,H,μ).

Proof. Set $u'=\mathbf{1}_Q u$ and let $M=\{w\in W\colon \det_2(I+\nabla u(w))\neq 0\}$. From Theorem 5.1, we have $\mu(T(M^c\cap Q))=0$, hence

$$E[\psi\,\mathbf{1}_{T'(A)}]=E[\psi\,\mathbf{1}_{T(A\cap M)}].$$

M has a countable partition (M_n) such that on each M_n, $T=I_W+u$ is equal to a bijective transformation, say T_n (cf. [7, 13]) such that $d(T_n^{-1})^*\,\mu=|\Lambda_n|\,d\mu$. Hence

$$
\begin{aligned}
E[\psi\,\mathbf{1}_{T(A\cap M)}] &\leqslant \sum_n E[\psi\,\mathbf{1}_{T_n(M_n\cap A)}] \\
&= \sum_n E[\psi\,\mathbf{1}_{M_n\cap A}\circ T_n^{-1}] \\
&= \sum_n E[\psi\circ T_n\,\mathbf{1}_{M_n\cap A}\,|\Lambda_n|] \\
&= \sum_n E[\psi\circ T\,\mathbf{1}_{M_n\cap A}\,|\Lambda|] \\
&= E[\psi\circ T\,|\Lambda_u|\,\mathbf{1}_A] \\
&= E[\psi\circ T'\,|\Lambda_{u'}|\,\mathbf{1}_A].
\end{aligned}
$$

VI. APPLICATIONS TO ABSOLUTE CONTINUITY

In the following three propositions we show how the Sard property and the existence of a right inverse yield new results on the absolute continuity of certain measures. The results will be presented under some general assumptions.

DEFINITION 6.1. Let $(W, \mathscr{B}(W), \mu)$ be any probability space and T a measurable transformation on W. The pair (T, μ) will be said to possess the Sard property with respect to $Q \in \mathscr{B}(W)$ if for every $V \in \mathscr{B}(W)$

(i) $T(V \cap Q)$ is universally measurable.

(ii) $\mu(T(V \cap Q)) = 0$ whenever $\mu(V \cap Q) = 0$.

PROPOSITION 6.1. *Let (T, μ) possess the Sard property with respect to Q and ν another probability measure on $(W, \mathscr{B}(W))$ for which $\nu(Q) > 0$ such that $\nu|_Q$ and μ are mutually singular; then $(T^*(\nu|_Q))$ and μ are mutually singular.*

Proof. Let N denote the set $N \subset Q$, $\mu(N) = 0$, $\nu(N) = \nu(Q)$, then $\mu(TN) = 0$ and

$$
\begin{aligned}
T^*(\nu|_Q)(TN) &\geqslant \nu(N \cap Q) \\
&= \nu(N) \\
&= \nu(Q),
\end{aligned}
$$

which completes the proof.

PROPOSITION 6.2. *Assume that* (T, μ) *possesses the Sard property with respect to* Q. *Further assume that* T *has a measurable right inverse* (*i.e.* $TSw = w$ *for almost all* w) *then*

$$\mu|_{S^{-1}(Q)} \ll T^*(\mu|_Q).$$

Therefore $\mu|_{S^{-1}(Q)} \ll T^*\mu$.

Proof. $S^*\mu = \nu_1 + \nu_2$ where $\nu_1 \ll \mu$, $\nu_2 \perp \mu$, then $\nu_2|_Q \perp \mu$; hence by Proposition 6.1,

$$T^*(\nu_2|_Q) \perp \mu.$$

On the other hand

$$T^*(\nu_1|Q) + T^*(\nu_2|Q) = T^*((S^*\mu)|_Q)$$
$$= \mu|_{S^{-1}(Q)}.$$

Hence $T^*(\nu_2|_Q) \ll \mu|_{S^{-1}(Q)}$. Consequently $T^*(\nu_2|_Q) = 0$ and $\mu|_{S^{-1}(Q)} = T^*(\nu_1|_Q) \ll T^*(\mu|_Q)$, since $\mu_1 \ll \mu_2$ implies $T^*\mu_1 \ll T^*\mu_2$.

DEFINITION 6.2. (T, μ) is said to possess the strong Sard property if, for any measurable V, TV is universally measurable and there exists a non-negative a.s. finite random variable Λ such that

$$\mu(TV) \leqslant \int_V \Lambda \, d\mu.$$

EXAMPLE VI.1. In the case of the abstract Wiener space, if $u: W \to H$ is locally $H - C^1$ such that the set $Q = \{\eta > 0\}$ is H-invariant, i.e., $Q + H \subset Q$, then $T = I_W + u$ satisfies the strong Sard property. In particular this is true if u is $H - C^1$.

PROPOSITION 6.3. *Let* T *possess the strong Sard property, set* $M = \{w: \Lambda(\omega) \neq 0\}$. *Assume that* T *possesses a measurable right inverse, then*

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

and

$$S^*\mu \ll \mu|_M.$$

Proof. Note that, since S is injective, the set $S(A)$ is measurable for any measurable subset A of W. We have

$$\mu(A)=\mu(TSA)\leqslant\int_{SA}\Lambda\,d\mu$$

$$=\int_W \mathbf{1}_{SA}\Lambda\,d\mu\leqslant\int \mathbf{1}_A(T\omega)\,\Lambda\,d\mu$$

which proves the first part. In order to prove the second part

$$\mu(S^{-1}(A))\leqslant\mu(TA)$$

$$\leqslant\int_A \Lambda\,d\mu,$$

hence $S^*\mu \ll \mu|_M$ which completes the proof.

From here on, we shall be working again in the frame of an abstract Wiener space (W, H, μ).

PROPOSITION 6.4. *Suppose that u is $\eta-H-C^1$ with the corresponding set Q and that there exists a measurable map S: $T(W)\mapsto W$ s.t. $S(T(w))=w\mu$-a.s. (i.e., S is a left inverse). Then $S^*(\mu|_{T(Q)})\approx\mu|_{M\cap Q}$ where $M=\{w: \det_2(I+\nabla u(w))\neq 0\}$.*

Proof. From the change of variables formula, we have, for any $f\in C_b^+(W)$,

$$E[f\circ S\circ T\,\mathbf{1}_Q\,|\Lambda|]=E[f\circ S\cdot N(w, Q)]$$

where $N(w, Q)$ is the multiplicity of T on Q and note that in this case we have $N(w, Q)=\mathbf{1}_{T(Q)}(w)$. Hence we have

$$E[f\cdot\mathbf{1}_Q\,|\Lambda|]=E[f\circ S\cdot\mathbf{1}_{T(Q)}]$$

and the proof follows.

COROLLARY 6.1. *Suppose moreover that u is $H-C^1_{\text{loc}}$, then we have*

$$S^*(\mu|_{T(Q)})\approx\mu|_M.$$

We say that a shift $T=I_W+u$ is *locally monotone* if there exists an increasing sequence (W_n) of measurable subsets of W which covers it almost surely and some $(u_n; n\in\mathbb{N})\subset\bigcup_{p>1}\mathbb{D}_{p,1}(H)$ such that $u=u_n$ almost surely on W_n and $\langle(I_H+\nabla u_n(w))\,h, h\rangle\geqslant 0$ almost surely for any $h\in H$ (the negligeable set may depend on h). For such a shift T (cf., [14]) it is known that

$$E[f\circ T\,|\Lambda|]\leqslant E[f],$$

for any $f\in C_b^+(W)$.

PROPOSITION 6.5. *Let $u\colon W\mapsto H$ be $H-C^1_{\mathrm{loc}}$ and $T=I_W+u$ be locally monotone. Then T possesses a left inverse S and we have*

$$S^*(\mu|_{T(Q)})\approx\mu|_M.$$

In fact

$$E[f\circ S\,\mathbf{1}_{T(Q)}]=E[f\,|\Lambda|],$$

for any $f\in C_b(W)$.
Moreover

$$\frac{dT^*(\mu|_M)}{d\mu}(w)=\mathbf{1}_{T(Q)}\frac{1}{|\Lambda_u(Sw)|},$$

μ almost surely.

Proof. Let us show that T possesses a measurable left inverse on Q. In fact, from Theorem 5.1 and from the monotonicity assumption, we have (c.f. [14]),

$$E[f\circ T\cdot|\Lambda|]=E[f\circ N(w,Q)]\leqslant E[f],$$

for any $f\in C_b^+(W)$. Hence $0\leqslant N(w,Q)\leqslant 1$. We have $T(Q)=\{w\colon N(w,Q)=1\}$ almost surely. Let T_Q be the restriction of T to Q and denote by U the set

$$U=T_Q(Q)\cap\{w\colon N(w,Q)=1\}.$$

Define $S\colon U\to Q$ as $S(T_Q y)=y$. Note that, if $w=T_Q y=T_Q y'$ then $y=y'$ since $N(w,Q)=1$, hence S is well-defined on U. If $A\in\mathscr{B}(W)$, then

$$S^{-1}(A\cap Q)=\{z\in W\colon N(z,Q)=1\}\cap T(A\cap Q),$$

as $T(A\cap Q)$ is in the universal sigma algebra by Theorem 3.1, S is measurable with respect to the trace of this sigma algebra on U. To show the equivalence, note that we have

$$E[f\circ T\cdot|\Lambda|]=E[f\,\mathbf{1}_{T(Q)}],$$

for any positive, bounded, measurable function f on W. Using this and the construction of S,

$$\begin{aligned} E[f\,\mathbf{1}_U \circ T\,|\Lambda|] &= E[f\,\mathbf{1}_U \circ T\,\mathbf{1}_Q\,|\Lambda|] \\ &= E[f \circ S \circ T\,\mathbf{1}_U \circ T\,\mathbf{1}_Q\,|\Lambda|] \\ &= E[f \circ S\,\mathbf{1}_U\,\mathbf{1}_{T(Q)}] \\ &= E[f \circ S\,\mathbf{1}_{T(Q)}], \end{aligned}$$

since $U = T(Q)$ almost surely. Moreover

$$\begin{aligned} E[\mathbf{1}_U \circ T\,|\Lambda|] &= E[\mathbf{1}_U\,N(w, Q)] \\ &= E[N(w, Q)] \\ &= E[|\Lambda|] \end{aligned}$$

and this implies that $\mathbf{1}_U \circ T = 1$ almost surely on the set $\{\Lambda \neq 0\}$. Combining this with the above relation, we obtain

$$E[f\,|\Lambda|] = E[f \circ S\,\mathbf{1}_{T(Q)}].$$

Note that $f \circ S$ is well-defined on the set $T(Q)$ since it is almost surely equal to U. Let us now calculate the Radon–Nikodym density of $T^*(\mu|_M)$:

$$\begin{aligned} E[f \circ T\,\mathbf{1}_M] &= E\left[f \circ T\,\mathbf{1}_M \frac{|\Lambda|}{|\Lambda|}\right] \\ &= E\left[f \sum_{y \in T^{-1}\{w\} \cap Q} \frac{1}{|\Lambda(y)|}\right] \\ &= E\left[f\,\mathbf{1}_U \sum_{y \in T^{-1}\{w\} \cap Q} \frac{1}{|\Lambda(y)|}\right] \\ &= E\left[f \frac{1}{|\Lambda(Sw)|}\mathbf{1}_{T(Q) \cap Q}\right] \\ &= E\left[f \frac{1}{|\Lambda(Sw)|}\mathbf{1}_{T(Q)}\right]. \end{aligned}$$

This completes the proof.

Remark. If $H + Q \subset Q$, then one can replace Q by W in the proposition.

REFERENCES

1. C. Castaing and M. Valadier, "Convex Analysis and Measurable Multifunctions," Lecture Notes in Math., Vol. 580, Springer-Verlag, Berlin, 1977.
2. C. Dellacherie and P.-A. Meyer, "Probabilities and Potential," Chaps. I–IV, North-Holland, Amsterdam, 1978.
3. H. Federer, "Geometric Measure Theory," Die Grund. Mathematischen Wissen., Vol. 153, Springer-Verlag, Berlin, 1969.
4. I. Fonseca and W. Gangbo, "Degree Theory in Analysis and Applications," Oxford Lecture Series in Mathematics and Its Applications 2, Clarendon Press, Oxford, 1955.
5. E. Getzler, Degree theory for Wiener maps, *J. Funct. Anal.* **68** (1986), 388–403.
6. S. Kusuoka, The nonlinear transformation of Gaussian measures on Banach space and its absolute continuity, *J. Fac. Sci., Tokyo Univ., Sect. IA* **29** (1982), 567–597.
7. S. Kusuoka, Some remarks on Getzler's degree theorem, *in* "Proc. of 5th Japan USSR Symp.," Lect. Notes Math, Vol. 1229, pp. 239–249, Springer-Verlag, Berlin/Heidelberg/New York, 1988.
8. D. Nualart, "The Malliavin Calculus and Related Topics. Probability and Its Applications," Springer-Verlag, Berlin, 1995.
9. J. T. Schwartz, "Nonlinear Functional Analysis," Gordon and Breach, New York, 1969.
10. S. Smale, An infinite dimensional version of Sardi's theorem, *Amer. J. Math.* **87** (1965), 861–866.
11. A. S. Üstünel, "An Introduction to Analysis on Wiener Space," Lecture Notes in Math., Vol. 1610, Springer-Verlag, Berlin, 1995.
12. A. S. Üstünel and M. Zakai, Transformation of the Wiener measure under noninvertible shifts, *Prob. Theory Rel. Fields* **99** (1994), 485–500.
13. A. S. Üstünel and M. Zakai, The change of variables formula on Wiener spaces, "Seminaire de Probabilités XXXI," Lecture Notes in Math., pp. 24–39, Springer-Verlag, Berlin, 1997.
14. A. S. Üstünel and M. Zakai, Measures induced on Wiener space by monotone shifts, *Prob. Theory Rel. Fields* **105** (1996), 545–563.
15. A. S. Üstünel and M. Zakai, Degree theory on Wiener space, *Prob. Theory Rel. Fields* **108** (1997), 259–279.