

Differentiable Shifts for Measures on Infinite Dimensional Spaces

By

Yasuo YAMASAKI* and Akihito HORA*

Introduction

The purpose of the present paper is to develop differentiation theory of measures on infinite dimensional vector spaces. Differentiable measures have been studied by several authors. See, for instance, Averbuh-Smolyanov-Fomin [1], Kuo [3] and Skorohod [4]. The derivative of a measure μ is defined by considering $\lim_{t \rightarrow 0} \frac{1}{t} (\mu_{ta} - \mu)$ for a vector a with respect to a certain topology on the space of measures. In general a measure μ on an infinite dimensional space is not differentiable along every direction. So we shall study the space of vectors along which μ is differentiable (such vectors are called differentiable shifts of μ).

In Chapter 1, taking ASF [1] as the starting point, we develop general theories on differentiation of measures on vector spaces. In Chapter 2, we apply these theories to product-type measures on \mathbb{R}^∞ and estimate the space of differentiable shifts. In particular differentiability w.r.t. (l^2) is discussed in detail.

Contents

Chapter 1.	General Theory on Differentiation of Measures
§1.	Several Topologies on the Space of Measures
§2.	Differentiation of Measures
§3.	Properties of a Differential Operator $\partial_a \mu$
§4.	Higher Derivatives of Measures
§5.	Translations and Linear Transformations
§6.	Conditional Measures and Product Measures
Chapter 2.	Applications to Measures on Product Spaces of \mathbb{R}
§7.	Finite Dimensional Case

Received July 16, 1986.

* Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606, Japan.

- §8. Product Measures on \mathbb{R}^∞
- §9. Differentiability w.r.t. (I^2)
- §10. Higher Order Differentiability
- §11. An Example — Gaussian Measures on \mathbb{R}^∞

Appendix
References

Chapter 1. General Theory on Differentiation of Measures

§1. Several Topologies on the Space of Measures

Let X be a real vector space and \mathcal{E} a linear subspace of X^* (algebraic dual of X) separating X . Then $\langle X, \mathcal{E} \rangle$ is called a dual pairing. \mathcal{B} denotes the minimal σ -field on X w.r.t. which all elements of \mathcal{E} are measurable. A real-valued σ -additive set function defined on a measurable space (X, \mathcal{B}) is called a real measure on (X, \mathcal{B}) . $M(X)$ denotes the set of all real measures on (X, \mathcal{B}) .

We consider four topologies τ_v, τ_b, τ_c and τ_s on $M(X)$ which are defined respectively by the following families of seminorms;

- τ_v : $\|\mu\|$ (total variation norm)
- τ_b : $\{|\int f(x)d\mu(x)|; f \text{ is a bounded } \mathcal{B}\text{-measurable function on } X\}$
- τ_c : $\{|\int f(x)d\mu(x)|; f \text{ is a bounded } \mathcal{E}\text{-weakly continuous and } \mathcal{B}\text{-measurable function on } X\}$
- τ_s : $\{|\mu(A)|; A \in \mathcal{B}\}$.

Now we give some relations between the above topologies.

Proposition 1.1. (1) $\tau_v \rightarrow \tau_b \begin{matrix} \nearrow \tau_c \\ \searrow \tau_s \end{matrix}$,

where “ \rightarrow ” implies that the left-hand side is stronger than the right-hand side;

(2) For a net $\{\mu_\alpha\}_{\alpha \in A}$ in $M(X)$, if there exists a positive measure μ satisfying

- (i) $\mu_\alpha \leq \mu$ (absolutely continuous) for $\forall \alpha \in A$, and
- (ii) $\sup_{\alpha \in A} \left\| \frac{d\mu_\alpha}{d\mu} \right\|_\infty < \infty$ ($\|\cdot\|_\infty$ is the essential supremum), then convergence w.r.t.

τ_c implies convergence w.r.t. τ_b ;

(3) For a net $\{\mu_\alpha\}_{\alpha \in A}$ in $M(X)$ which is bounded w.r.t. τ_s , convergence w.r.t.

τ_s implies convergence w.r.t. τ_b .

Proof. (1) is evident. (2) follows from an inequality

$$\left| \int f(x)d\mu_\alpha(x) \right| \leq \left| \int g(x)d\mu_\alpha(x) \right| + \left\| \frac{d\mu_\alpha}{d\mu} \right\|_\infty \|f-g\|_{L^1(\mu)}$$

where f is an arbitrary bounded \mathcal{B} -measurable function and g is a bounded continuous function approximating f w.r.t. $L^1(\mu)$ -topology.

Noticing that boundedness w.r.t. τ_s is equivalent to that w.r.t. τ_v , we see that (3) follows from an inequality

$$\left| \int f(x) d\mu_a(x) \right| \leq \left| \int g(x) d\mu_a(x) \right| + \|f-g\|_\infty \|\mu_a\|$$

where f is an arbitrary bounded \mathcal{B} -measurable function and g is a simple function approximating f w.r.t. $\|\cdot\|_\infty$. q.e.d.

The following fact is well-known. See, for instance, Dunford-Schwartz [2].

Proposition 1.2. *$M(X)$ is sequentially complete w.r.t. τ_s .*

Next we consider continuity of an element of $M(X)$ under shifts along a line. For $\mu \in M(X)$ and $a \in X$, μ_a is defined as $\mu(\cdot + a)$.

Theorem 1.1. *Let $\mu \in M(X)$, $a \in X$, $a \neq 0$ and L_a be the one-dimensional subspace generated by a . Put $\psi(t) = \mu_{ta}$ for $t \in \mathbb{R}$. The following conditions are equivalent;*

- (1) ψ is continuous w.r.t. τ_v ,
- (2) ψ is continuous w.r.t. τ_s ,
- (3) there exists an L_a -quasi-invariant positive measure λ such that $\mu \lesssim \lambda$.

Proof. (1) \Rightarrow (2) is evident.

(2) \Rightarrow (3): Let $\{t_k\}_{k=1}^\infty$ be a countable dense subset of \mathbb{R} . Put $\lambda = \sum_{k=1}^\infty 2^{-k} |\mu|_{t_k a}$. It follows from the continuity of μ_{ta} in τ_s that, for $A \in \mathcal{B}$,

$$\begin{aligned} \lambda(A) = 0 &\Leftrightarrow |\mu|_{t_k a}(A) = 0 \text{ for } \forall k \in \mathbb{N} \\ &\Leftrightarrow \mu_{t_k a}(B) = 0 \text{ for } \forall B \subset A, \forall k \in \mathbb{N} \\ &\Leftrightarrow \mu_{ta}(B) = 0 \text{ for } \forall B \subset A, \forall t \in \mathbb{R} \\ &\Leftrightarrow |\mu|_{ta}(A) = 0 \text{ for } \forall t \in \mathbb{R} \\ &\Leftrightarrow \lambda_{ta}(A) = 0 \text{ for } \forall t \in \mathbb{R}, \end{aligned}$$

namely λ is L_a -quasi-invariant. $\mu \lesssim \lambda$ is evident.

(3) \Rightarrow (1): It is sufficient to prove the continuity at $t=0$. Let m be the image measure of the one-dimensional Lebesgue measure under the embedding $t \mapsto ta$. Put $\nu = m * \lambda$. Then ν is L_a -invariant and $\nu \sim \lambda$. We denote the density $d\mu/d\nu$ by ρ . We can approximate ρ w.r.t. $L^1(\nu)$ -topology with a function g in the form of $g(x) = \varphi((\xi_1, x), \dots, (\xi_n, x))$ where $\xi_1, \xi_2, \dots, \xi_n \in \mathcal{E}$ and $\varphi \in C_0(\mathbb{R}^n)$.

Then we have

$$\begin{aligned} \|\mu_{ta} - \mu\| &= \int_X |\rho(x+ta) - \rho(x)| d\nu(x) \\ &\leq \int_X |g(x+ta) - g(x)| d\nu(x) + 2 \int_X |g(x) - \rho(x)| d\nu(x). \end{aligned}$$

For a fixed g , the first term in the right-hand side tends to zero as $t \rightarrow 0$. This completes the proof. q.e.d.

Definition. An element a of X is called a continuous shift of μ if it satisfies the conditions of Theorem 1.1. We denote by C_μ the set of all continuous shifts of μ .

C_μ is a linear subspace of X .

Remark. We see $C_\mu = C_{|\mu|}$ by Theorem 1.1.

§2. Differentiation of Measures

In this section we define derivatives of measures according to ASF [1]. Proposition 2.1, Proposition 2.2 and Theorem 2.2 are due to them. Theorem 2.1 is a slight modification of a result by them, to which, however, we give here a quite different proof from the original one.

Definition. Let $\mu \in M(X)$ and $a \in X$. a is called a differentiable shift of μ if $\frac{1}{t}(\mu_{ta} - \mu)$ converges w.r.t. τ_s as $t \rightarrow 0$ ($t \in \mathbb{R}$). Then the limit is denoted by $\partial_a \mu$. We denote by D_μ the set of all differentiable shifts of μ .

It is clear that C_μ contains D_μ .

Remark 1. In view of Proposition 1.2, $\partial_a \mu$ exists iff $\frac{1}{t}(\mu_{ta} - \mu)$ is of Cauchy w.r.t. τ_s .

Remark 2. Some authors define differentiability of μ w.r.t. other topologies on $M(X)$. See Kuo [3] and Skorohod [4].

Proposition 2.1. $a \in D_\mu$ implies

- (1) $a \in D_{\mu^+} \cap D_{\mu^-}$ where $\mu = \mu^+ - \mu^-$ is the Jordan-Hahn decomposition of μ ,
- (2) $\partial_a \mu \leq |\mu|$.

Proof is omitted. See ASF [1].

Proposition 2.2. Let $a \in D_\mu$. Then an inequality

$$(2.1) \quad \|\mu_a - \mu\| \leq \|\partial_a \mu\|$$

holds.

Proof. Since $\mu_a(A) - \mu(A) = \partial_a \mu(A + \theta a)$ ($0 < \theta < 1$) for $A \in \mathcal{B}$,

$$(2.2) \quad \sup_{A \in \mathcal{B}} |\mu_a(A) - \mu(A)| \leq \sup_{A \in \mathcal{B}} |\partial_a \mu(A)|$$

holds. Noting $\|\nu\| = 2 \sup_{A \in \mathcal{B}} |\nu(A)|$ for an element ν of $M(X)$ satisfying $\nu(X) = 0$, we see (2.2) implies (2.1). q.e.d.

Theorem 2.1. $a \in D_\mu$ and $b \in C_\mu$ imply $b \in C_{\partial_a \mu}$. In particular, $a \in D_\mu$ implies $a \in C_{\partial_a \mu}$.

Proof. In view of Proposition 2.1 (2) and Theorem 1.1, there exists an L_b -quasi-invariant positive measure λ such that $\partial_a \mu \leq \lambda$. q.e.d.

Using this theorem, we prove that the definition of differentiable shifts does not change even if we adopt τ_ν -topology on $M(X)$ instead of τ_s .

Theorem 2.2. $a \in D_\mu$ iff $\frac{1}{t}(\mu_{ta} - \mu)$ converges w.r.t. τ_ν as $t \rightarrow 0$.

Proof. Since $(\partial_a \mu)_{ta}$ is continuous in t w.r.t. τ_ν by Theorem 2.1, it is continuous also w.r.t. τ_s , so that we can express set-wisely

$$(2.3) \quad \mu_{ta} - \mu = \int_0^t (\partial_a \mu)_{sa} ds.$$

Then, using $\|\nu\| = 2 \sup_{A \in \mathcal{B}} |\nu(A)|$ again, we have

$$\left\| \frac{1}{t}(\mu_{ta} - \mu) - \partial_a \mu \right\| \leq \left| \frac{1}{t} \int_0^t \|(\partial_a \mu)_{sa} - \partial_a \mu\| ds \right| \xrightarrow{t \rightarrow 0} 0. \quad \text{q.e.d.}$$

§3. Properties of Differential Operator $\partial_a \mu$

Proposition 3.1. (1) D_μ is a linear subspace of X , (2) $\partial_a \cdot$ ($a \neq 0$) and $\partial_a \mu$ ($\mu \neq 0$) are one-to-one linear operators.

Proof. Let $a, b \in D_\mu$. We have

$$\begin{aligned} \left\| \frac{\mu_{t(a+b)} - \mu}{t} - (\partial_a \mu + \partial_b \mu) \right\| &\leq \left\| \frac{\mu_{ta+tb} - \mu_{tb} - \partial_a \mu}{t} \right\| + \left\| \frac{\mu_{tb} - \mu}{t} - \partial_b \mu \right\| \\ &\leq \left\| \frac{\mu_{ta} - \mu}{t} - \partial_a \mu \right\| + \left\| \partial_a \mu - (\partial_a \mu)_{-tb} \right\| + \left\| \frac{\mu_{tb} - \mu}{t} - \partial_b \mu \right\| \xrightarrow{t \rightarrow 0} 0 \end{aligned}$$

by virtue of Theorem 2.1 and Theorem 2.2. Therefore $a+b \in D_\mu$ and $\partial_{a+b} \mu = \partial_a \mu + \partial_b \mu$. As for scalar multiplications, the proof is trivial.

“One-to-one” part follows from Lemma 3.1 below. q.e.d.

Lemma 3.1. *Let $\mu \in M(X)$ and $a \in D_\mu$. $\partial_a \mu = 0$ implies $a = 0$ or $\mu = 0$.*

Proof. By (2.3) it follows from $\partial_a \mu = 0$ that $\mu_{ta} = \mu$ for $\forall t \in \mathcal{R}$. Then $(\mu^+)_{ta} = \mu^+$ and $(\mu^-)_{ta} = \mu^-$. Therefore it is sufficient to prove that, if $a \neq 0$ and μ is a positive measure satisfying $\mu_{ta} = \mu$ for $\forall t \in \mathcal{R}$, then $\mu = 0$. Choose $\xi \in \mathcal{E}$ (see the opening of §1) such that $\langle a, \xi \rangle = 1$. We denote by $\xi \circ \mu$ the image measure of μ under ξ . Then we have for $\forall t \in \mathcal{R}$

$$(3.1) \quad (\xi \circ \mu)_t = (\xi \circ \mu)_{\langle ta, \xi \rangle} = \xi \circ (\mu_{ta}) = \xi \circ \mu.$$

Since $\xi \circ \mu$ is a finite positive measure on \mathcal{R} , (3.1) implies $\xi \circ \mu = 0$, therefore $\mu = 0$. q.e.d.

Theorem 3.1. *Let X be given a topology stronger than or equal to the \mathcal{E} -weak topology and $M(X)$ given τ_ν -topology. Assume that $a_\alpha \xrightarrow{\alpha \in \Lambda} a$ (Λ is a directed set) in X , $\mu_n \xrightarrow{n \rightarrow \infty} \mu$ in $M(X)$, $a_\alpha \in D_{\mu_n}$ for $\forall \alpha \in \Lambda$ and $\forall n \in \mathcal{N}$, and $\{\partial_{a_\alpha} \mu_n\}_{(\alpha, n) \in \Lambda \times \mathcal{N}}$ is a Cauchy net in $M(X)$ along the directed set $\Lambda \times \mathcal{N}$. Then $a \in D_\mu$ and*

$$(3.2) \quad \partial_a \mu = \lim_{(\alpha, n)} \partial_{a_\alpha} \mu_n$$

holds.

Corollary. $\partial_a \cdot$ and $\partial_a \mu$ are closed operators.

Proof of Theorem 3.1. First we show that $(\mu_n)_{ta_\alpha} \xrightarrow{(\alpha, n)} \mu_{ta}$ for $\forall t \in \mathcal{R}$. Using Proposition 2.2, we have

$$\begin{aligned} \|(\mu_m)_{ta_\alpha} - (\mu_n)_{ta_\beta}\| &\leq \|(\mu_m)_{ta_\alpha} - (\mu_m)_{ta_\beta}\| + \|(\mu_m)_{ta_\beta} - (\mu_n)_{ta_\beta}\| \\ &\leq |t| \|\partial_{a_\alpha} \mu_m - \partial_{a_\beta} \mu_m\| + \|\mu_m - \mu_n\| \end{aligned}$$

and so $\{(\mu_n)_{ta_\alpha}\}_{(\alpha, n) \in \Lambda \times \mathcal{N}}$ is a Cauchy net in $M(X)$. Then we see

$$(3.3) \quad \mu_{ta} = \lim_{(\alpha, n)} (\mu_n)_{ta_\alpha}$$

by checking the characteristic functionals $(\mu_n(\xi) \rightarrow \mu(\xi))$ and $\exp(i(\xi, ta_\alpha)) \rightarrow \exp(i(\xi, ta))$ for $\forall \xi \in \mathcal{E}$.

Put $\nu = \lim_{(\alpha, n)} \partial_{a_\alpha} \mu_n$. For $\alpha \in \Lambda$ and $m \in \mathcal{N}$, we have

$$(3.4) \quad \left\| \frac{\mu_{ta} - \mu}{t} - \nu \right\| \leq \left\| \frac{\mu_{ta} - \mu}{t} - \frac{(\mu_m)_{ta_\alpha} - \mu_m}{t} \right\| + \left\| \frac{(\mu_m)_{ta_\alpha} - \mu_m}{t} - \partial_{a_\alpha} \mu_m \right\| + \|\partial_{a_\alpha} \mu_m - \nu\|.$$

The first term in the right-hand side of (3.4) is estimated uniformly w.r.t. t as follows:

$$\begin{aligned}
 (3.5) \quad & \lim_{(n, \beta)} \left\| \frac{(\mu_n)_{t\alpha\beta} - \mu_n}{t} - \frac{(\mu_m)_{t\alpha\beta} - \mu_m}{t} \right\| \\
 & \leq \lim_{(n, \beta)} \left\{ \left\| \frac{(\mu_n - \mu_m)_{t\alpha\beta} - (\mu_n - \mu_m)}{t} \right\| + \left\| \frac{(\mu_m)_{t\alpha\beta} - (\mu_m)_{t\alpha\beta}}{t} \right\| \right\} \\
 & \leq \lim_{(n, \beta)} \{ \|\partial_{\alpha\beta}\mu_n - \partial_{\alpha\beta}\mu_m\| + \|\partial_{\alpha\beta}\mu_m - \partial_{\alpha\beta}\mu_m\| \}.
 \end{aligned}$$

Since the second term in the right-hand side of (3.4) tends to zero as $t \rightarrow 0$ for a fixed (α, m) , (3.4) and (3.5) imply $\partial_a\mu = \nu$. q.e.d.

§4. Higher Derivatives of Measures

Proposition 4.1. *Let $\partial_a\mu$, $\partial_b\mu$ and $\partial_b\partial_a\mu$ exist. Then $\partial_a\partial_b\mu$ exists and is equal to $\partial_b\partial_a\mu$.*

Proof. We have

$$\begin{aligned}
 (4.1) \quad & \frac{(\partial_b\mu)_{ta} - \partial_b\mu}{t} = \partial_b \left(\frac{\mu_{ta} - \mu}{t} \right) = \partial_b \left(\frac{1}{t} \int_0^t (\partial_a\mu)_{ua} du \right) \\
 & = \lim_{s \rightarrow 0} \frac{1}{s} \frac{1}{t} \int_0^t \{ (\partial_a\mu)_{ua+sb} - (\partial_a\mu)_{ua} \} du = \frac{1}{t} \int_0^t \partial_b(\partial_a\mu)_{ua} du
 \end{aligned}$$

(the last equality follows from Lebesgue’s convergence theorem). Since $\partial_b(\partial_a\mu)_{ua}$ is continuous in u by Theorem 2.1, (4.1) converges to $\partial_b\partial_a\mu$. q.e.d.

Corollary. *Higher derivatives of measures are independent of the order of the differentiation, if they exist.*

Remark. We cannot omit the condition that $\partial_b\mu$ exists in Proposition 4.1. A counterexample is given in §7.

Definition. Let E be a linear subspace of X . An element μ of $M(X)$ is called n -differentiable w.r.t. E if $\partial_{a_1} \cdots \partial_{a_n}\mu$ exists for arbitrary n elements a_1, \dots, a_n of E .

Theorem 4.1. *Let μ be n -differentiable w.r.t. E , a_1, \dots, a_n elements of X and $\{a_{1,\alpha_1}\}_{\alpha_1 \in A_1}, \dots, \{a_{n,\alpha_n}\}_{\alpha_n \in A_n}$ nets in E . Assume that $a_{k,\alpha_k} \xrightarrow{\tau_k} a_k$ in X (w.r.t. a topology stronger than or equal to the \mathcal{E} -weak topology) for $k=1, \dots, n$ and $\{\partial_{a_{n,\alpha_n}} \cdots \partial_{a_{1,\alpha_1}} \mu\}_{(\alpha_m, \dots, \alpha_n) \in A_m \times \dots \times A_n}$ is a Cauchy net in $M(X)$ (w.r.t. τ_ν -topology) for $m=1, \dots, n$. Then $\partial_{a_1} \cdots \partial_{a_n}\mu$ exists and*

$$\partial_{a_1} \cdots \partial_{a_n}\mu = \lim_{(\alpha_1, \dots, \alpha_n)} \partial_{a_{1,\alpha_1}} \cdots \partial_{a_{n,\alpha_n}} \mu$$

holds.

Proof is by induction on n and Theorem 3.1.

§5. Translations and Linear Transformations

Proposition 5.1. *Let X be a measurable vector space, μ an element of $M(X)$, τ a translation on X (i.e. $\forall z \in X, \tau x = x + z$ for $\forall x \in X$). Then we have*

$$(1) \quad D_{\tau \circ \mu} = D_\mu \text{ and}$$

$$(5.1) \quad \partial_a(\tau \circ \mu) = \tau \circ \partial_a \mu \quad \text{for } \forall a \in D_\mu,$$

$$(2) \quad \text{for } E, \text{ a linear subspace of } X,$$

μ is n -differentiable w.r.t. $E \Leftrightarrow \tau \circ \mu$ is n -differentiable w.r.t. E .

Proposition 5.2. *Let X, Y be measurable vector spaces, μ an element of $M(X)$, T a measurable linear map $X \rightarrow Y$. Then we have*

$$(1) \quad T(D_\mu) \subset D_{T \circ \mu} \text{ and}$$

$$(5.2) \quad \partial_{Ta}(T \circ \mu) = T \circ \partial_a \mu \quad \text{for } \forall a \in D_\mu,$$

$$(2) \quad \text{for } E, \text{ a linear subspace of } X,$$

μ is n -differentiable w.r.t. $E \Rightarrow T \circ \mu$ is n -differentiable w.r.t. $T(E)$,

$$(5.3) \quad \partial_{Ta_1} \cdots \partial_{Ta_n}(T \circ \mu) = T \circ (\partial_{a_1} \cdots \partial_{a_n} \mu) \quad \text{for } a_1, \dots, a_n \in E.$$

Remark. If T is a measurable linear isomorphism in Proposition 5.2, the same holds also for T^{-1} . Thus we have the equality $T(D_\mu) = D_{T \circ \mu}$ in (1), and the equivalence \Leftrightarrow in (2).

Above propositions are easily proved from definitions.

§6. Conditional Measures and Product Measures

Consider (X, \mathcal{B}) , X being a real vector space and \mathcal{B} the σ -field defined by \mathcal{E} where $\langle X, \mathcal{E} \rangle$ is a dual pairing. Let $a \in X, a \neq 0$, and take $\xi \in \mathcal{E}$ such that $\langle a, \xi \rangle = 1$. Then X can be decomposed as $X = L_a \oplus \{\xi\}^\perp$, where L_a is the one-dimensional subspace generated by a and $\{\xi\}^\perp$ is the annihilator of ξ ,

$$(6.1) \quad X \ni x = \lambda a + y; \quad \lambda = \langle x, \xi \rangle, \quad y = x - \lambda a \in \{\xi\}^\perp.$$

Similarly \mathcal{E} can be decomposed as $\mathcal{E} = L_\xi \oplus \{a\}^\perp$. Then the measurable space (X, \mathcal{B}) is isomorphic with the product measurable space $(L_a, \mathcal{B}_{\{\xi\}}) \times (\{\xi\}^\perp, \mathcal{B}_{\{a\}^\perp})$ where $\mathcal{B}_{\{\xi\}}$ [resp. $\mathcal{B}_{\{a\}^\perp}$] is the σ -field defined by $\{\xi\}$ [resp. $\{a\}^\perp$]. Note that $(L_a, \mathcal{B}_{\{\xi\}})$ is isomorphic with the Borel measurable space of \mathbf{R} , so we shall identify them.

Consider $\mu \in M(X)$. Identifying X with $\mathbf{R} \times \{\xi\}^\perp$, we identify the restriction of $|\mu|$ on $\mathcal{B}_{\{a\}^\perp}$ with a positive measure $\bar{\mu}_{\{a\}^\perp}$ on $(\{\xi\}^\perp, \mathcal{B}_{\{a\}^\perp})$ given by

$$(6.2) \quad \bar{\mu}_{(a)}(A) = |\mu|(\mathbb{R} \times A) \quad \text{for } A \in \mathcal{B}_{(a)}^\perp.$$

For any Borel set B of \mathbb{R} , the measure $\mu_B(A) = \mu(B \times A)$ is absolutely continuous w.r.t. $\bar{\mu}_{(a)}$, so that we get $d\mu_B = f_B d\bar{\mu}_{(a)}$ with some $\mathcal{B}_{(a)}^\perp$ -measurable function f_B . Since the Borel field of \mathbb{R} is standard (namely since it is isomorphic with the infinite product measurable space of $\{0, 1\}$), $\{f_B(x)\}_{B: \text{Borel}}$ determines a measure $\mu^x(B)$ on the Borel field of \mathbb{R} . Therefore we have

$$(6.3) \quad \mu(B \times A) = \int_A \mu^x(B) d\bar{\mu}_{(a)}(x).$$

From this we see that for any $C \in \mathcal{B}$,

$$(6.4) \quad \mu(C) = \int \mu^x(C(x)) d\bar{\mu}_{(a)}(x)$$

where $C(x) = \{\lambda \in \mathbb{R}; \lambda a + x \in C\}$.

μ^x is called the conditional measure of μ at $x \in \{\xi\}^\perp$. Note that μ^x has meaning only for $\bar{\mu}_{(a)}$ -almost all x .

Theorem 6.1. *Let $\mu \in M(X)$ and $a \in X, a \neq 0$. Consider the representation of μ as (6.4).*

- (1) $a \in C_\mu \Leftrightarrow 1 \in C_{\mu^x}$ for $\bar{\mu}_{(a)}$ -a.a.x;
- (2) Under the assumption that μ is positive, μ is quasi-invariant under translation by $L_a \Leftrightarrow \mu^x$ is quasi-invariant (on \mathbb{R}) for $\bar{\mu}_{(a)}$ -a.a.x;
- (3) $a \in D_\mu \Leftrightarrow 1 \in D_{\mu^x}$ for $\bar{\mu}_{(a)}$ -a.a.x and

$$(6.5) \quad \int \|\partial \mu^x\| d\bar{\mu}_{(a)}(x) < \infty$$

where $\partial \mu^x$ is the derivative of μ^x along 1.

Proof. (1) From (6.4), we see

$$(6.6) \quad \|\mu\| = \int \|\mu^x\| d\bar{\mu}_{(a)}(x).$$

Hence $\|\mu_{ia} - \mu\| = \int \|\mu^x_{ia} - \mu^x\| d\bar{\mu}_{(a)}(x)$ assures the “ \Leftarrow ” part of (1).

Conversely if $a \in C_\mu$, then we have $\mu \lesssim \nu$ for some L_a -quasi-invariant positive measure ν (see Theorem 1.1.). Consider the conditional decomposition of ν :

$$(6.7) \quad \nu = \int \nu^x d\bar{\nu}_{(a)}(x).$$

Putting $d\mu/d\nu = f(\lambda, x)$ ($(\lambda, x) \in \mathbb{R} \times \{\xi\}^\perp$), we have

$$\mu(C) = \iint_{C(x)} f(\lambda, x) d\nu^x(\lambda) d\bar{\nu}_{(a)}(x) \quad \text{for } C \in \mathcal{B}.$$

On the other hand, putting $d\bar{\mu}_{(a)}/d\bar{\nu}_{(a)}=g$, we have

$$\mu(C) = \int \mu^x(C(x))g(x)d\bar{\nu}_{(a)}(x) \quad \text{for } C \in \mathcal{B}.$$

Comparing these, we get $\mu^x(\cdot) \sim \int f(\lambda, x)d\nu^x(\lambda) \lesssim \nu^x$ for $\bar{\mu}_{(a)}$ -a.a. x , by virtue of the uniqueness of the conditional decomposition.

If we admit (2), ν^x is quasi-invariant for $\bar{\nu}_{(a)}$ -a.a. x . Thus, noting that $\mu^x \lesssim \nu^x$, we see μ^x is translationally continuous for $\bar{\mu}_{(a)}$ -a.a. x .

(2) “ \Leftarrow ” part is evident.

If $\mu \sim \nu$, then as in the second paragraph of the proof of (1), we see $\mu^x \sim \nu^x$ for $\bar{\mu}_{(a)}$ -a.a. x . Suppose that μ is L_a -quasi-invariant. We have $\mu \sim \mu * \lambda$ for any positive measure λ on L_a . Then $\mu^x \sim (\mu * \lambda)^x = \mu^x * \lambda$, so that if λ is quasi-invariant, so is μ^x .

(3) “ \Leftarrow ” part is derived from (6.6) as

$$\left\| \frac{\mu_{ta} - \mu}{t} - \int \partial \mu^x d\bar{\mu}_{(a)}(x) \right\| = \left\| \frac{\mu^x_t - \mu^x}{t} - \partial \mu^x \right\| d\bar{\mu}_{(a)}(x).$$

Conversely assume $a \in D_\mu$. Consider the conditional decomposition of $\partial_a \mu$. Since $\partial_a \mu \lesssim |\mu|$, we can write as

$$(6.8) \quad \partial_a \mu = \int \nu^x d\bar{\mu}_{(a)}(x)$$

for some measure ν^x on \mathcal{R} . Then

$$\mu_{ta} - \mu = \int_0^t (\partial_a \mu)_{sa} ds = \iint_0^t \nu^x_s ds d\bar{\mu}_{(a)}(x).$$

From the uniqueness of the decomposition, for any fixed t , we have

$$(6.9) \quad \mu^x_t - \mu^x = \int_0^t \nu^x_s ds \quad \text{for } \bar{\mu}_{(a)}\text{-a.a.}x.$$

Since μ is continuous along a , μ^x and ν^x are translationally continuous. Therefore (6.9) holds for all t for $\bar{\mu}_{(a)}$ -a.a. x . This implies that μ^x is differentiable and that the derivative is ν^x . From (6.8) we have

$$\|\partial_a \mu\| = \int \|\nu^x\| d\bar{\mu}_{(a)}(x) < \infty$$

which implies (6.5).

q.e.d.

Similar discussions are valid even if we consider a finite dimensional subspace Y instead of single a . The results are written as follows:

Let Y be an n -dimensional subspace of X . Then X can be decomposed as $X = Y \oplus Z$, so that as a set we have $X \simeq \mathcal{R}^n \times Z$. (X, \mathcal{B}_X) is isomorphic with the product measurable space of $(\mathcal{R}^n, \text{Borel field})$ and $(Z, \mathcal{B}_{Y^\perp})$. A measure

$\mu \in M(X)$ can be written uniquely as

$$(6.10) \quad \mu = \int \mu^x d\bar{\mu}_Y(x),$$

where μ^x is a measure on \mathbb{R}^n , and $\bar{\mu}_Y$ is the restriction of $|\mu|$ on \mathcal{B}_Y^\perp .

Theorem 6.2. *Under the representation (6.10), we have the following: Let $a \in Y$ and \bar{a} be the corresponding element of \mathbb{R}^n .*

- (1) $a \in C_\mu \Leftrightarrow \bar{a} \in C_{\mu^x}$ for $\bar{\mu}_Y$ -a.a.x
- (2) Under the assumption that μ is positive, μ is L_a -quasi-invariant $\Leftrightarrow \mu^x$ is $L_{\bar{a}}$ -quasi-invariant for $\bar{\mu}_Y$ -a.a.x
- (3) $a \in D_\mu \Leftrightarrow \bar{a} \in D_{\mu^x}$ for $\bar{\mu}_Y$ -a.a.x and

$$(6.11) \quad \int \|\partial_{\bar{a}} \mu^x\| d\bar{\mu}_Y(x) < \infty.$$

Of course this theorem can be applied to the case where μ is continuous or differentiable along every direction in Y . If μ is differentiable along every direction in Y , then μ^x can be written as Theorem 7.1 (see §7).

Next we consider the case that (X, \mathcal{B}) is the product measurable space of (Y, \mathcal{B}_1) and (Z, \mathcal{B}_2) (Y and Z may be infinite dimensional).

Theorem 6.3. *Let μ_1 and μ_2 be non-zero elements of $M(Y)$ and $M(Z)$ respectively. Consider the product measure $\mu = \mu_1 \times \mu_2$. Then we have*

$$(6.12) \quad D_\mu = D_{\mu_1} \times D_{\mu_2}$$

$$(6.13) \quad \partial_{(y,z)} \mu = \partial_y \mu_1 \times \mu_2 + \mu_1 \times \partial_z \mu_2 \text{ for } y \in D_{\mu_1} \text{ and } z \in D_{\mu_2}$$

Proof. First note the equality $\|\mu\| = \|\mu_1\| \|\mu_2\|$. Then for $y \in D_{\mu_1}$,

$$\left\| \frac{\mu_{t(y,0)} - \mu}{t} - \partial_y \mu_1 \times \mu_2 \right\| = \left\| \frac{(\mu_1)_{ty} - \mu_1}{t} - \partial_y \mu_1 \right\| \|\mu_2\|.$$

This implies $(y, 0) \in D_\mu$ and $\partial_{(y,0)} \mu = \partial_y \mu_1 \times \mu_2$. In a similar way, if $z \in D_{\mu_2}$, then $(0, z) \in D_\mu$ and $\partial_{(0,z)} \mu = \mu_1 \times \partial_z \mu_2$. Since $(y, z) = (y, 0) + (0, z)$, these imply $D_{\mu_1} \times D_{\mu_2} \subset D_\mu$ and (6.13).

We must prove $D_\mu \subset D_{\mu_1} \times D_{\mu_2}$. Let $(a, b) \in D_\mu$. For any $A \in \mathcal{B}_1$, since

$$\frac{\mu_{t(a,b)} - \mu}{t}(A \times Z) = \frac{(\mu_1)_{ta} - \mu_1}{t}(A) \mu_2(Z),$$

if $\mu_2(Z) \neq 0$, then $(a, b) \in D_\mu$ implies $a \in D_{\mu_1}$. Thus the proof is completed for the case $\mu_2(Z) \neq 0$ and $\mu_1(Y) \neq 0$.

We shall consider the case $\mu_1(Y) = 0$ or $\mu_2(Z) = 0$. Again let $(a, b) \in D_\mu$.

Since $(a, b) \in D_{|\mu|}$ and $|\mu| = |\mu_1| \times |\mu_2|$, we get at least $a \in D_{|\mu_1|}$ and $b \in D_{|\mu_2|}$. So $a \in C_{|\mu_1|}$ [resp. $b \in C_{|\mu_2|}$], hence $a \in C_{\mu_1}$ [resp. $b \in C_{\mu_2}$] (see Remark in §1). Since $\mu_1(A+ta)\mu_2(B+tb)$ is differentiable in t for any $A \in \mathcal{B}_1$ and $B \in \mathcal{B}_2$, for the differentiability of $\mu_1(A+ta)$ at $t=0$ it is sufficient to prove that there exists $B \in \mathcal{B}_2$ such that $\mu_2(B) \neq 0$ and $\mu_2(B+tb)$ is differentiable at $t=0$.

Let $Y = P_1 \cup P_1^c$ [resp. $Z = P_2 \cup P_2^c$] be the Hahn decomposition w.r.t. μ_1 [resp. μ_2]. We can assume that $\mu_1(P_1) > 0$ and $\mu_2(P_2) > 0$, because we can consider $-\mu_1$ or $-\mu_2$ instead of μ_1 or μ_2 if necessary. Since $\mu_1(P_1+ta)\mu_2(P_2+tb)$ is maximal at $t=0$, we get from the differentiability

$$(6.14) \quad \lim_{t \rightarrow 0} \frac{1}{t} \{ \mu_1(P_1+ta)\mu_2(P_2+tb) - \mu_1(P_1)\mu_2(P_2) \} = 0.$$

But we have

$$(6.15) \quad \begin{aligned} & \frac{1}{t} \{ \mu_1(P_1+ta)\mu_2(P_2+tb) - \mu_1(P_1)\mu_2(P_2) \} \\ &= \frac{1}{t} \{ \mu_1(P_1+ta) - \mu_1(P_1) \} \mu_2(P_2+tb) + \frac{1}{t} \{ \mu_2(P_2+tb) - \mu_2(P_2) \} \mu_1(P_1). \end{aligned}$$

For sufficiently small t we have $\mu_2(P_2+tb) > 0$ from $b \in C_{\mu_2}$. Thus two terms in the right-hand side of (6.15) have the same sign. Therefore each term must tend to zero as $t \rightarrow 0$. This means that $\mu_1(P_1+ta)$ and $\mu_2(P_2+tb)$ are differentiable with the derivative zeros at $t=0$. q.e.d.

Chapter 2. Applications to Measures on Product Spaces of \mathbf{R}

§7. Finite Dimensional Case

We see in this section that differentiation of measures on finite dimensional spaces is attributed to that of functions. Theorem 7.1 and Theorem 7.2 below are due to ASF [1].

We denote the Borel field on \mathbf{R}^d by \mathcal{B}_d , the Lebesgue measure on \mathbf{R}^d by m^d , the canonical basis of \mathbf{R}^d by $\{e_1, \dots, e_d\}$, and by $\partial_k f$ the derivative of f (in distribution sense) along e_k .

Theorem 7.1. *Let $\mu \in M(\mathbf{R}^d)$. The following (1) and (2) are necessary and sufficient for $D_\mu = \mathbf{R}^d$:*

$$(1) \quad \mu \leq m^d$$

$$(2) \quad \partial_k f \in L^1(m^d) \text{ for } k=1, \dots, d \text{ where } f = \frac{d\mu}{dm^d} \text{ (density of } \mu \text{ w.r.t. } m^d).$$

And then $\frac{d\partial_{e_k}\mu}{dm^d} = \partial_k f$ ($k=1, \dots, d$) holds.

Remark. Under the assumption (1) of Theorem 7.1, we have $\partial_k f \in L^1(m^d) \Leftrightarrow e_k \in D_\mu$ for each k , even if μ is not differentiable along other directions.

Theorem 7.2. *Let $\mu \in M(\mathbb{R}^d)$. The following (1) and (2) are necessary and sufficient for μ to be infinitely differentiable along all directions:*

- (1) $\mu \lesssim m^d$
- (2) $\partial^\alpha f \in L^1(m^d)$ for every multi-index α where $f = \frac{d\mu}{dm^d}$.

For the proofs of above theorems, see ASF [1].

Now we give a counterexample notified in the Remark after Proposition 4.1.

Example. Let ψ be a C^2 -function on \mathbb{R}^2 with compact support. For arbitrarily given positive sequences $\{\alpha_n\}_{n=1}^\infty$ and $\{\beta_n\}_{n=1}^\infty$, we can take real sequences $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ such that the supports of the functions $\{\psi(\alpha_n x + a_n, \beta_n y + b_n)\}_{n=1}^\infty$ are mutually disjoint and separated enough. Taking another positive sequence $\{r_n\}_{n=1}^\infty$, we put

$$(7.1) \quad f(x, y) = \sum_{n=1}^\infty r_n \psi(\alpha_n x + a_n, \beta_n y + b_n).$$

We see that f is a C^2 -function on \mathbb{R}^2 and

$$\begin{aligned} \|f\| &= \sum_{n=1}^\infty \frac{r_n}{\alpha_n \beta_n} \|\psi\|, \quad \|\partial_x f\| = \sum_{n=1}^\infty \frac{r_n}{\beta_n} \|\partial_x \psi\|, \\ \|\partial_y f\| &= \sum_{n=1}^\infty \frac{r_n}{\alpha_n} \|\partial_y \psi\|, \quad \|\partial_y \partial_x f\| = \sum_{n=1}^\infty r_n \|\partial_y \partial_x \psi\| \end{aligned}$$

where $\|\cdot\|$ is the L^1 -norm w.r.t. m^2 . Therefore, determining $\{\alpha_n\}$, $\{\beta_n\}$ and $\{r_n\}$ appropriately, we have $f, \partial_x f, \partial_y \partial_x f \in L^1(m^2)$ and $\partial_y f \notin L^1(m^2)$ (for example $\alpha_n = \frac{1}{n}, \beta_n = n$, and $r_n = \frac{1}{n^2}$). Now we put $a = (1, 0), b = (0, 1)$ and $\mu = f m^2$. Remark after Theorem 7.1 shows that $\partial_a \mu, \partial_b \partial_a \mu$ exist and $\partial_b \mu$ does not.

§8. Product Measures on \mathbb{R}^∞

From now on we study differentiable shifts for a product measure $\mu = \prod_{k=1}^\infty \mu_k$ on \mathbb{R}^∞ where each μ_k is a probability measure on \mathbb{R} . If every μ_k is the same measure μ_1 , we call $\mu = \mu_1^\infty$ a stationary product measure.

We denote the canonical basis of \mathbb{R}^∞ by $\{e_1, e_2, \dots\}$ and put $\mathbb{R}_0^\infty = \{(x^k)_{k=1}^\infty \in \mathbb{R}^\infty; \exists n \in \mathbb{N} \text{ st. } x^k = 0 \text{ for } k \geq n\}$. We give \mathbb{R}^∞ the usual weak topology and denote by $\mathcal{B}(\mathbb{R}^\infty)$ the weak Borel field on \mathbb{R}^∞ .

Definition. Let E be a linear subspace of X . We say μ to be E -differentiable if $E \subset D_\mu$ and exactly E -differentiable if $E = D_\mu$.

Proposition 8.1. *The following (1) and (2) are necessary and sufficient for a product measure $\mu = \prod_{k=1}^\infty \mu_k$ to be \mathbf{R}_0^∞ -differentiable:*

- (1) $\mu_k \lesssim m^1$ for $\forall k \in \mathbf{N}$
- (2) $f'_k \in L^1(m^1)$ for $\forall k \in \mathbf{N}$ where $f_k = \frac{d\mu_k}{dm^1}$.

Then we have

$$(8.1) \quad \partial_{e_k} \mu = \mu_1 \times \cdots \times \mu_{k-1} \times \partial \mu_k \times \mu_{k+1} \times \cdots .$$

Proof is trivial. See Theorem 7.1 and Theorem 6.3.

Proposition 8.2. *Let μ be an \mathbf{R}_0^∞ -differentiable measure on \mathbf{R}^∞ (not necessarily positive or product-type). Put $E = \{a = (a^k)_{k=1}^\infty \in \mathbf{R}^\infty; \sum_{k=1}^\infty |a^k| \|\partial_{e_k} \mu\| < \infty\}$. Then $E \subset D_\mu$ holds.*

Proof. Let $a \in E$ and put $a_n = (a^1, \dots, a^n, 0, \dots)$. Then $a_n \xrightarrow{n \rightarrow \infty} a$ in \mathbf{R}^∞ and $\{\partial_{a_n} \mu = \sum_{k=1}^n a^k \partial_{e_k} \mu\}_{n=1}^\infty$ is a Cauchy sequence in $M(\mathbf{R}^\infty)$ w.r.t. τ_v . Therefore, from Corollary to Theorem 3.1, we have $a \in D_\mu$ and

$$(8.2) \quad \partial_a \mu = \sum_{k=1}^\infty a^k \partial_{e_k} \mu \quad (\tau_v\text{-convergent}). \quad \text{q.e.d.}$$

Corollary. *If an \mathbf{R}_0^∞ -differentiable product measure $\mu = \prod_{k=1}^\infty \mu_k$ satisfies $\sup_k \|\partial \mu_k\| < \infty$ (in particular if μ is an \mathbf{R}_0^∞ -differentiable stationary product measure), then $(l^1) \subset D_\mu$.*

Definition (Skorohod [4]). Let μ be a positive measure and $a \in D_\mu$. The density function $\frac{d\partial_a \mu}{d\mu}(x)$ is called the logarithmic derivative of μ along a and denoted by $l_\mu(a; x)$.

The following properties of logarithmic derivatives are evident: for $a, b \in D_\mu$ and $s, t \in \mathbf{R}$,

$$(8.3) \quad l_\mu(sa + tb; x) = sl_\mu(a; x) + tl_\mu(b; x) \quad (\mu\text{-a.e.}),$$

$$(8.4) \quad \|l_\mu(a; \cdot)\|_{L^1(\mu)} = \|\partial_a \mu\| .$$

Proposition 8.3. *Let $\mu = \prod_{k=1}^\infty \mu_k$ be \mathbf{R}_0^∞ -differentiable. Then $(l^\infty) \cdot D_\mu \subset D_\mu$ holds, where for $t \in (l^\infty)$ and $a \in D_\mu$ their product is defined by $ta = (t^k a^k)_{k=1}^\infty$.*

Proof. Let $t \in (I^\infty)$ and $a \in D_\mu$. We may assume $|t^k| \leq 1$ for $\forall k \in \mathbb{N}$ since D_μ is a linear space. Put $t_n = (t^1, \dots, t^n, 0, \dots)$ and $a_n = (a^1, \dots, a^n, 0, \dots)$. Noting Corollary to Theorem 3.1 and (8.4), we have only to show that $l_\mu(t_n a_n; \cdot)$ converges in $L^1(\mu)$ as $n \rightarrow \infty$. Since $l_\mu(t_n a_n; x) = \sum_{k=1}^n t^k a^k \frac{d\partial \mu_k}{d\mu_k}(x^k)$, Lemma 8.1 below completes the proof. q.e.d.

Lemma 8.1. *Let $\{g_k\}_{k=1}^\infty$ be a system of independent random variables on (X, μ) with mean zero. If $\sum_k g_k$ converges in $L^1(\mu)$, then $\sum_k t^k g_k$ also converges in $L^1(\mu)$ for $\forall t = (t^k)$ st. $|t^k| \leq 1$.*

Proof. In general, $\|x\| \leq \|x+y\|$ implies $\|x+ty\| \leq \|x+y\|$ for $0 \leq t \leq 1$. For,

$$\|x+ty\| = \|t(x+y) + (1-t)x\| \leq t\|x+y\| + (1-t)\|x\| \leq \|x+y\|.$$

Now that for $S_1, S_2 \subset \mathbb{N}$ such that $S_1 \subset S_2$ and $\#S_2 < \infty$

$$(8.5) \quad \left\| \sum_{k \in S_1} g_k \right\|_{L^1(\mu)} \leq \left\| \sum_{k \in S_2} g_k \right\|_{L^1(\mu)}$$

holds, putting $\epsilon^k = 1$ [resp. -1] if $t^k \geq 0$ [resp. < 0], we have

$$(8.6) \quad \begin{aligned} \left\| \sum_{k=1}^n t^k g_k \right\| &\leq \left\| \sum_{k=1}^{n-1} t^k g_k + \epsilon^n g_n \right\| \leq \dots \leq \left\| \sum_{k=1}^n \epsilon^k g_k \right\| \\ &\leq \left\| \sum_{\epsilon^k=1} g_k \right\| + \left\| \sum_{\epsilon^k=-1} g_k \right\| \leq 2 \left\| \sum_{k=1}^n g_k \right\|. \end{aligned}$$

(8.6) shows that $\sum_{k=1}^\infty t^k g_k$ converges in $L^1(\mu)$. q.e.d.

§9. Differentiability w.r.t. (I^2)

In this section we study the conditions for \mathbb{R}_0^∞ -differentiable measures to be (I^2) -differentiable. In the sequel we consider the product measures such that

$$(9.1) \quad \mu = \prod_{k=1}^\infty \mu_k, \mu_k = f_k m^1, f_k \geq 0, \|f_k\|_{L^1(m^1)} = 1, f'_k \in L^1(m^1),$$

or the stationary product measures such that

$$(9.2) \quad \mu = \mu_1^\infty, \mu_1 = f m^1, f \geq 0, \|f\|_{L^1(m^1)} = 1, f' \in L^1(m^1).$$

When we say “ μ as (9.1) [resp. μ as (9.2)]” below, μ_k and f_k [resp. μ_1 and f] are always assumed to satisfy (9.1) [resp. (9.2)].

Proposition 9.1. *Let μ be a product measure as (9.1). If $\sup_k \left\| \frac{f'_k}{f_k} \right\|_{L^2(\mu_k)} < \infty$, then we have $(I^2) \subset D_\mu$ and, for $a \in (I^2)$,*

$$(9.3) \quad I_\mu(a; x) = \sum_{k=1}^\infty a^k \frac{f'_k(x^k)}{f_k(x^k)} \quad \text{in } L^1(\mu).$$

Proof. Let $a \in (l^2)$ and put $a_n = (a^1, \dots, a^n, 0, \dots)$. Then

$$\|I_\mu(a_n; x) - I_\mu(a_m; x)\|_{L^2(\mu)}^2 = \sum_{k=m+1}^n (a^k)^2 \left\| \frac{f'_k}{f_k} \right\|_{L^2(\mu_k)}^2 \xrightarrow{m, n \rightarrow \infty} 0.$$

Therefore $\{I_\mu(a_n; \cdot)\}_{n=1}^\infty$ is a Cauchy sequence in $L^1(\mu)$. This completes the proof by Corollary to Theorem 3.1. q.e.d.

Corollary. Let μ be a stationary product measure as (9.2). If $\frac{f'}{f} \in L^2(\mu_1)$, then $(l^2) \subset D_\mu$ holds.

For stationary product measures, the converse of the above result holds.

Theorem 9.1. Let μ be a stationary product measure as (9.2). If $(l^2) \subset D_\mu$, then $\frac{f'}{f} \in L^2(\mu_1)$ holds.

Proof. Corollary to Theorem 3.1 and the closed graph theorem show that the map $a \mapsto \partial_a \mu$ of (l^2) into $M(\mathbf{R}^\infty)$ is continuous ((l^2) is given the usual Hilbertian topology), namely

$$\forall \varepsilon (0 < \varepsilon < 1), \exists \delta > 0 \text{ st. } \forall a \in (l^2), \|a\| < \delta \Rightarrow \|\partial_a \mu\| < \varepsilon.$$

For arbitrary $n \in \mathbf{N}$ and $t^k \in \mathbf{R}$ such that $|t^k| < 1$ ($k=1, 2, \dots$), we put

$$a = \left(\frac{\delta}{\sqrt{n}} t^1, \dots, \frac{\delta}{\sqrt{n}} t^n, 0, \dots \right).$$

Since $\|a\| < \delta$, putting $\varphi = \frac{f'}{f}$, we have

$$(9.4) \quad \varepsilon > \|\partial_a \mu\| = \|I_\mu(a; \cdot)\|_{L^1(\mu)} = \left\| \sum_{k=1}^n \frac{\delta}{\sqrt{n}} t^k \varphi(x^k) \right\|_{L^1(\mu)}.$$

Since $|e^{i\theta} - 1| \leq |\theta|$ for $\theta \in \mathbf{R}$, (9.4) implies

$$(9.5) \quad \int_{\mathbf{R}^\infty} \left| \exp\left(i \sum_{k=1}^n \frac{\delta}{\sqrt{n}} t^k \varphi(x^k)\right) - 1 \right| d\mu(x) < \varepsilon.$$

Integrating (9.5) w.r.t. m^n on $\left[-\frac{1}{2}, \frac{1}{2}\right]^n$ and estimating it from below, we have

$$\varepsilon > \left| \int_{\mathbf{R}^\infty} \int_{[-1/2, 1/2]^n} \left\{ \exp\left(i \sum_{k=1}^n \frac{\delta}{\sqrt{n}} t^k \varphi(x^k)\right) - 1 \right\} dt^1 \dots dt^n d\mu(x) \right|$$

$$\begin{aligned} &= \left| \int_{\mathbb{R}^\infty} \left(\prod_{k=1}^n \frac{\sin \frac{\delta}{2\sqrt{n}} \varphi(x^k)}{\frac{\delta}{2\sqrt{n}} \varphi(x^k)} - 1 \right) d\mu(x) \right| \\ &= 1 - \left(\int_{\mathbb{R}} \frac{\sin \frac{\delta}{2\sqrt{n}} \varphi(t)}{\frac{\delta}{2\sqrt{n}} \varphi(t)} d\mu_1(t) \right)^n, \end{aligned}$$

namely

$$(9.6) \quad \left(\int_{\mathbb{R}} \frac{\sin \frac{\delta}{2\sqrt{n}} \varphi(t)}{\frac{\delta}{2\sqrt{n}} \varphi(t)} d\mu_1(t) \right)^n > 1 - \varepsilon \quad \text{for } \forall n \in \mathbb{N}.$$

Assuming that n is odd, we see that the integral in (9.6) is positive. Taking “log” of (9.6) and noting $\log s \leq s - 1$ for $s > 0$, we have

$$(9.7) \quad n \left(1 - \int_{\mathbb{R}} \frac{\sin \frac{\delta}{2\sqrt{n}} \varphi(t)}{\frac{\delta}{2\sqrt{n}} \varphi(t)} d\mu_1(t) \right) < \log \frac{1}{1 - \varepsilon}.$$

Therefore, by Fatou’s inequality, we see

$$\log \frac{1}{1 - \varepsilon} \geq \int_{\mathbb{R}} \lim_{\substack{n \rightarrow \infty \\ n: \text{ odd}}} n \left(1 - \frac{\sin \frac{\delta}{2\sqrt{n}} \varphi(t)}{\frac{\delta}{2\sqrt{n}} \varphi(t)} \right) d\mu_1(t) = \int_{\mathbb{R}} \frac{1}{6} \left(\frac{\delta}{2} \varphi(t) \right)^2 d\mu_1(t),$$

which shows $\varphi \in L^2(\mu_1)$. q.e.d.

Next we study the converse implication: $D_\mu \subset (I^2)$. We prepare two Lemmas.

Lemma 9.1. *Let $\{(X_n, \mathcal{B}_n)\}_{n=1}^\infty$ be a projective system of measurable vector space, (X, \mathcal{B}) its projective limit, $p_n: X \rightarrow X_n$ its projection and μ a positive measure on X . Then $a \in D_\mu$ implies that $p_n a \in D_{p_n \circ \mu}$ and*

$$(9.8) \quad \partial_{p_n a}(p_n \circ \mu) = p_n \circ \partial_a \mu,$$

$$(9.9) \quad I_{p_n \circ \mu}(p_n a; p_n x) \xrightarrow[n \rightarrow \infty]{} I_\mu(a; x) \text{ in } L^1(\mu) \text{ and } \mu\text{-a.e.}$$

Proof. Let $a \in D_\mu$. Proposition 5.2 shows $p_n a \in D_{p_n \circ \mu}$ and (9.8). Since p_n ’s are projections, the $L^1(\mu)$ -convergence of (9.9) is evident, so that, by martingale convergence theorem, (9.9) converges also μ -a.e. q.e.d.

Lemma 9.2. *Let $\{g_k\}_{k=1}^\infty$ be a system of independent random variables on (X, μ) with mean zero. If $\sum_k g_k$ converges in $L^1(\mu)$, then $\sum_k g_{\sigma(k)}$ also converges in $L^1(\mu)$ for any $\sigma \in \mathfrak{S}_\infty$ (the set of all bijections of \mathbb{N}).*

Proof is easy. Note (8.5) in the proof of Lemma 8.1.

Here we introduce a notation. Let g be a real-valued function and $C > 0$. Put

$$(9.10) \quad g^C(x) = \begin{cases} C & \text{if } g(x) > C \\ g(x) & \text{if } |g(x)| \leq C \\ -C & \text{if } g(x) < -C. \end{cases}$$

g^C is called the truncation of g at C .

Theorem 9.2. *Let μ be a product measure as (9.1). Put $\varphi_k = \frac{f'_k}{f_k}$. If $\lim_{k \rightarrow \infty} \|\varphi_k^C\|_{L^2(\mu_k)} > 0$ for some $C > 0$, then $D_\mu \subset (I^2)$ holds.*

Corollary. *For a stationary product measure μ as (9.2), $D_\mu \subset (I^2)$ holds.*

Proof of Theorem 9.2. Let $a = (a^k)_{k=1}^\infty \in D_\mu$. Applying Lemma 9.1 where $X_n = \mathbb{R}^n$ and $X = \mathbb{R}^\infty$, we see $\sum_k a^k \varphi_k(x^k)$ converges in $L^1(\mu)$ and μ -a.e. Therefore, by Kolmogorov's three series theorem,

$$(9.11) \quad \sum_k V_\mu([a^k \varphi_k(x^k)]^C) < \infty,$$

$$(9.12) \quad \sum_k E_\mu([a^k \varphi_k(x^k)]^C) \text{ converges}$$

where V_μ is the variance and E_μ is the mean w.r.t. μ . Since $\sum_k a^{\sigma(k)} \varphi_{\sigma(k)}(x^{\sigma(k)})$ converges in $L^1(\mu)$ for any $\sigma \in \mathfrak{S}_\infty$ by Lemma 9.2, it also converges μ -a.e. Then, again by the three series theorem, $\sum_k E_\mu([a^{\sigma(k)} \varphi_{\sigma(k)}(x^{\sigma(k)})]^C)$ converges. This implies that (9.12) is absolutely convergent. Therefore,

$$(9.13) \quad \sum_k \{E_\mu([a^k \varphi_k(x^k)]^C)\}^2 < \infty$$

holds. (9.11) and (9.13) imply

$$(9.14) \quad \sum_k E_\mu(\{[a^k \varphi_k(x^k)]^C\}^2) < \infty.$$

Noting that $[\alpha\beta]^C \geq \alpha^1\beta^C$ for $\alpha, \beta \geq 0$, we see from (9.14) that

$$\sum_k \{[a^k]^1\}^2 E_\mu(\{\varphi_k(x^k)\}^C)^2 < \infty.$$

Therefore, from the assumption of Theorem, we have $\sum_k \{[a^k]^1\}^2 < \infty$, which is

equivalent to $a \in (l^2)$. q.e.d.

Corollary to Proposition 9.1, Theorem 9.1 and Corollary to Theorem 9.2 characterize the (l^2) -differentiability of stationary product measures:

Theorem 9.3. *For a stationary product measure μ as (9.2),*

$$\frac{f'}{f} \in L^2(\mu_1) \Leftrightarrow D_\mu = (l^2)$$

holds.

§10. Higher Order Differentiability

Proposition 10.1. *Let μ be a product measure as (9.1) satisfying $f_k^{(m)} \in L^1(m^1)$ for $k=1, 2, \dots$ and $m=1, \dots, n$. If $\sup_k \left\| \frac{f_k^{(m)}}{f_k} \right\|_{L^1(\mu_k)} < \infty$ for $m=1, \dots, n$, then μ is n -times differentiable w.r.t. (l^1) .*

Proof. Note that n -times differentiability w.r.t. \mathbb{R}_0^∞ is assured by the assumption that $f^{(m)} \in L^1(m^1)$ for $k=1, 2, \dots$ and $m=1, \dots, n$. Using Theorem 4.1, we easily see the assertion. q.e.d.

Theorem 10.1. *Let μ be a product measure as (9.1) satisfying $f_k^{(m)} \in L^1(m^1)$ for $k=1, 2, \dots$ and $m=1, \dots, n$ ($n \geq 2$). If $\sup_k \left\| \frac{f_k^{(m)}}{f_k} \right\|_{L^2(\mu_k)} < \infty$ for $m=1, \dots, n-1$ and $\sup_k \left\| \frac{f_k^{(n)}}{f_k} \right\|_{L^1(\mu_k)} < \infty$, then μ is n -times differentiable w.r.t. (l^2) .*

Proof. In view of Theorem 4.1, it is sufficient to show that, for $a_1, \dots, a_m \in (l^2)$, $m \leq n$,

$$(10.1) \quad \sum_{j_1, \dots, j_m} a_1^{j_1} \dots a_m^{j_m} \frac{d\partial_{e_{j_1}} \dots \partial_{e_{j_m}} \mu}{d\mu}(x) \text{ converges in } L^1(\mu).$$

We shall consider only the case $m=n$ because the case $m < n$ can be discussed in the same way. The summation (10.1) is divided in $\sum_{j=1}^\infty a_1^j a_2^j \dots a_n^j d(\partial_{e_j}^n \mu) / d\mu(x)$ and the sum \sum' of the other terms (containing at least two different j_k 's). The inequality

$$\begin{aligned} \sum_{j=1}^\infty \left\| a_1^j \dots a_n^j \frac{d(\partial_{e_j}^n \mu)}{d\mu} \right\|_{L^1(\mu)} &= \sum_{j=1}^\infty |a_1^j| \dots |a_n^j| \left\| \frac{f_j^{(n)}}{f_j} \right\|_{L^1(\mu_j)} \\ &\leq \|a_1\| \dots \|a_n\| \sup_k \left\| \frac{f_k^{(n)}}{f_k} \right\|_{L^1(\mu_k)} \end{aligned}$$

assures the $L^1(\mu)$ -convergence of the first sum.

To prove the $L^1(\mu)$ -convergence of Σ' , it is sufficient to show its $L^2(\mu)$ -convergence:

$$(10.2) \quad \left\| \sum'_{(j)} a_{j_1}^{j_1} \cdots a_{j_n}^{j_n} \frac{d\partial_{e_{j_1}} \cdots \partial_{e_{j_n}} \mu}{d\mu} \right\|_{L^2(\mu)}^2 \\ = \sum'_{(j)} \sum'_{(k)} a_{j_1}^{j_1} \cdots a_{j_n}^{j_n} a_{k_1}^{k_1} \cdots a_{k_n}^{k_n} \int_{\mathbb{R}^\infty} \frac{d\partial_{e_{j_1}} \cdots \partial_{e_{j_n}} \mu}{d\mu}(x) \frac{d\partial_{e_{k_1}} \cdots \partial_{e_{k_n}} \mu}{d\mu}(x) d\mu(x).$$

However the integral is equal to zero if j_1 does not appear in k_1, \dots, k_n , because

$$\int_{\mathbb{R}} \frac{f_{j_1}^{(m)}(t)}{f_{j_1}(t)} d\mu_{j_1}(t) = 0 \quad \text{for } m = 1, \dots, n-1.$$

The same holds for j_2, \dots, j_n . The integral in (10.2) is, if it is not equal to zero, of the form of

$$\prod_k \int_{\mathbb{R}} \frac{f_k^{(m_k)}(t)}{f_k(t)} \frac{f_k^{(m'_k)}(t)}{f_k(t)} d\mu_k(t) \quad (1 \leq m_k, m'_k \leq n-1)$$

where the number of factors of the product is at most n , so that its absolute value is less than

$$\left(\sup \left\{ \left\| \frac{f_k^{(m)}}{f_k} \right\|_{L^2(\mu_k)} ; m = 1, \dots, n-1, k = 1, 2, \dots \right\} \right)^{2n}.$$

Thus, it is sufficient to prove

$$\Sigma'' |a_{j_1}^{j_1}| \cdots |a_{j_n}^{j_n}| |a_{k_1}^{k_1}| \cdots |a_{k_n}^{k_n}| < \infty,$$

Σ'' being the sum over those terms such that every j_p appears in k_1, \dots, k_n and every k_p appears in j_1, \dots, j_n . In general we see that, for $b_1, \dots, b_r \in (l^2)$,

$$\sum'_{(j)} |b_{j_1}^{j_1}| \cdots |b_{j_r}^{j_r}| \leq C_r \|b_1\| \cdots \|b_r\|$$

for some $C_r > 0$, where \sum^0 means the sum over those terms such that every j_p appears in j_1, \dots, j_r at least twice. q.e.d.

§11. An Example—Gaussian Measures on \mathbb{R}^∞

In this section we consider Gaussian measures on \mathbb{R}^∞ . Let $m = (m^k)_{k=1}^\infty$ and $c = (c^k)_{k=1}^\infty$ be real and positive sequences respectively. Put $f_k(t) = \frac{1}{\sqrt{2\pi c^k}} \exp\left\{-\frac{(t-m^k)^2}{2c^k}\right\}$ and $\mu_k = f_k m^k$. $g_{m,c} = \prod_{k=1}^\infty \mu_k$ is called a Gaussian measure on \mathbb{R}^∞ with mean m and variance c . When $m = (0, 0, \dots)$ and $c = (1, 1, \dots)$, we denote simply by g and call it the standard Gaussian measure.

From Theorem 9.3 and Theorem 10.1, we have the following

Proposition 11.1. *The standard Gaussian measure on \mathbb{R}^∞ is infinitely differentiable exactly w.r.t. (l^2) .*

We define A_c and τ_m , operators on \mathbb{R}^∞ , by $A_c: (x^k) \rightarrow (\sqrt{c^k} x^k)$ and $\tau_m: (x^k) \rightarrow (x^k + m^k)$. We see $g_{m,c} = \tau_m A_c \circ g$. Then, from Proposition 5.1, Proposition 5.2 and Proposition 11.1, we have the following

Proposition 11.2. *Put $H_c = \{x \in \mathbb{R}^\infty; \sum_{k=1}^{\infty} \frac{(x^k)^2}{c^k} < \infty\}$. A Gaussian measure $g_{m,c}$ is infinitely differentiable exactly w.r.t. H_c .*

Appendix

In this appendix we prove that a measure on an infinite dimensional space cannot be continuous (therefore cannot be differentiable) along every direction.

Theorem A. *Let (X, \mathcal{B}) be a real measurable vector space where \mathcal{B} is the σ -field defined by E , a linear subspace of X^n , separating X . Consider a probability measure μ on (X, \mathcal{B}) . If $\mu(Z) = 0$ for any measurable linear proper subspace Z of X , then we have $\dim X < \infty$.*

Proof. We give E the characteristic topology of μ (which is identical with the restriction of the topology of measure convergence on E) and denote it by E_μ . First we show $X = E_\mu^*$. Let $\xi_{k_j} \xrightarrow{j \rightarrow \infty} 0$ in E_μ . Then, taking a subsequence, we have $\xi_{k_j} \xrightarrow{j \rightarrow \infty} 0$ μ -a.e., namely $\mu(Z) = 1$ where $Z = \{x \in X; \langle x, \xi_{k_j} \rangle \xrightarrow{j \rightarrow \infty} 0\}$. This implies by the assumption that $Z = X$, namely $\langle x, \xi_{k_j} \rangle \xrightarrow{j \rightarrow \infty} 0$ for $\forall x \in X$. Since $\{\xi_{k_j}\}$ is an arbitrary sequence which tends to zero in E_μ , we see $x \in E_\mu^*$ for $\forall x \in X$. Thus $X = E_\mu^*$, so that X is a countable union of the polars of neighborhoods of zero in E_μ .

Now we take V , a neighborhood of zero in E_μ , satisfying $\mu^*(V^0) > 0$ (μ^* is the outer measure and V^0 is the polar of V). Put $\|\xi\|_V^2 = \int_{V^0} |\langle x, \xi \rangle|^2 d\mu(x)$ for $\xi \in E$. We show that $\|\cdot\|_V$ determines the characteristic topology of μ . Let $\|\xi_{k_j}\|_V \xrightarrow{j \rightarrow \infty} 0$. Then, taking a subsequence, we have $\xi_{k_j} \xrightarrow{j \rightarrow \infty} 0$ μ -a.e. on V^0 . Putting $Z = \{x \in X; \langle x, \xi_{k_j} \rangle \xrightarrow{j \rightarrow \infty} 0\}$, we see $\mu(Z) > 0$ because $\mu^*(V^0) > 0$. This implies by the assumption that $Z = X$, therefore $\xi_{k_j} \xrightarrow{j \rightarrow \infty} 0$ in E_μ . Since $\{\xi_{k_j}\}$ is arbitrary, we see that the topology defined by $\|\cdot\|_V$ is stronger than the characteristic topology. The converse is evident since V is a neighborhood of zero in E_μ .

Let $\delta > 0$ be such that $\|\xi\|_V \leq \delta \Rightarrow \xi \in V$. Take an orthonormal system

$\{\xi_k\}_{k=1}^N$ w.r.t. $\|\cdot\|_V$ then,

$$\begin{aligned} N &= \sum_{k=1}^N \|\xi_k\|_V^2 \leq \sup_{x \in V^0} \sum_{k=1}^N |\langle x, \xi_k \rangle|^2 = \sup_{x \in V^0} \left| \left\langle x, \frac{\sum_{k=1}^N \langle x, \xi_k \rangle \xi_k}{\left(\sum_{k=1}^N |\langle x, \xi_k \rangle|^2\right)^{1/2}} \right\rangle \right|^2 \\ &\leq \sup_{x \in V^0} \sup_{\|\xi\|_V=1} |\langle x, \xi \rangle|^2 \leq \sup_{x \in V^0} \sup_{\xi \in (1/\delta)V} |\langle x, \xi \rangle|^2 \leq 1/\delta^2. \end{aligned}$$

Thus E_μ is finite dimensional, therefore so is X . q.e.d.

Corollary. *If X is infinite dimensional, μ is not continuous along every direction.*

Proof. Take a measurable linear proper subspace Z of X such that $\mu(Z) > 0$ and a vector a such that $a \notin Z$. If μ is continuous along a , $\mu(Z+ta) > 0$ for sufficiently small t . On the other hand $\{Z+ta\}_{t \in \mathbb{R}}$ is mutually disjoint. This is a contradiction. q.e.d.

Remark. The continuity of μ implies that of μ^+ and μ^- . So the above Corollary is valid even if μ is a real measure.

From Theorem A some measurable linear proper subspace Z satisfies $\mu(Z) > 0$ in an infinite dimensional X . But we cannot claim $\mu(Z)=1$.

Example. We define a probability measure μ on $\mathbf{R}_0^\infty = \bigcup_{n=1}^\infty (\mathbf{R}^n \times (0))$ as follows. Let μ_n be a probability measure on \mathbf{R}^n which is equivalent to the Lebesgue measure. Consider μ_n as a measure on $\mathbf{R}^n \times (0)$ and put $\mu = \sum_{n=1}^\infty \frac{1}{2^n} \mu_n$. If a measurable linear subspace Z satisfies $\mu(Z)=1$, Z must be equal to \mathbf{R}_0^∞ . For,

$$\begin{aligned} \mu(Z) = 1 &\Leftrightarrow \forall n \in \mathbf{N}, \mu_n(Z) = 1 \\ &\Leftrightarrow \forall n \in \mathbf{N}, \{x \in \mathbf{R}^n; (x, 0) \in Z\} = \mathbf{R}^n \\ &\Leftrightarrow Z = \mathbf{R}_0^\infty. \end{aligned}$$

References

- [1] Averbuh, V.I., Smolyanov, O.G., and Fomin, S.V., Generalized functions and differential equations in linear spaces, I. Differentiable measures, *Trans. Moscow Math. Soc.*, **24** (1971), 140–184.
- [2] Dunford, N. and Schwartz, J.T., *Linear operators, Part I. General theory*, Interscience, New York, 1967.
- [3] Kuo, H.H., Differentiable measures, *Chinese J. Math.*, **2** (1974), 189–199.
- [4] Skorohod, A.V., *Integration in Hilbert spaces*, Springer, Berlin-New York, 1974.