CHAPTER 3

Differentiable measures on linear spaces

In this chapter we consider the basic properties of differentiable measures on linear spaces. As everywhere in this book, the reader may assume that we are concerned with measures on Banach spaces or Souslin locally convex spaces or even on l^2 or $\mathbb{R}^\infty$, which simplifies some technical details. However, most of the definitions are given in their natural generality.

3.1. Directional differentiability

Let X be a linear space equipped with some σ-algebra $\mathcal{A}$ invariant with respect to the shifts along a given vector $h \in X$, i.e., $A + th \in \mathcal{A}$ for all $A \in \mathcal{A}$ and $t \in \mathbb{R}^1$. We are going to define certain properties of measures on $\mathcal{A}$ such as the differentiability, continuity, and analyticity along h. The most important example is the case where X is a locally convex space with one of our two standard σ-algebras $\mathcal{B}(X)$ or $\sigma(X)$. Most of the special results below will be obtained for these cases. The only purpose for our consideration in this generality is to emphasize that only linear and measurable structures are intrinsically connected with our basic concepts.

If μ is a measure on $\mathcal{A}$, then we set

$$\mu_h(A) := \mu(A + h).$$

The measure μ_h is the image of the measure μ under the mapping $x \mapsto x - h$. It is called the *shift of the measure* μ by h. The following identity is valid:

$$(f \cdot \mu)_h = f(\,\cdot + h) \cdot \mu_h, \quad f \in \mathcal{L}^1(\mu).$$

For the proof we use the relations

$$\int_X \varphi(x)\,(f \cdot \mu)_h(dx) = \int_X \varphi(x - h) f(x)\,\mu(dx) = \int_X \varphi(y) f(y + h)\,\mu_h(dy),$$

which are valid for every bounded $\mathcal{A}$-measurable function φ.

The following fundamental concept is due to S.V. Fomin [**440**]–[**443**].

3.1.1. Definition. *A measure μ on X is called differentiable along the vector h in Fomin's sense if, for every set $A \in \mathcal{A}$, there exists a finite limit*

$$d_h\mu(A) := \lim_{t\to 0} \frac{\mu(A + th) - \mu(A)}{t}. \tag{3.1.1}$$

Since in place of A in (3.1.1) one can substitute $A + t_0 h$, the function

$$t \mapsto \mu(A + th)$$

is differentiable at every point.

The set function $d_h\mu$ defined by (3.1.1) can be written as a pointwise limit of the sequence of measures $A \mapsto n\bigl(\mu(A + n^{-1}h) - \mu(A)\bigr)$. Therefore, by the Nikodym theorem recalled in §1.3 it is automatically a countably additive measure on $\mathcal{A}$ of bounded variation. Here $d_h\mu$ is always a signed measure: $d_h\mu(X) = 0$. Hence it is reasonable from the very beginning to admit signed measures.

The measure $d_h\mu$ defined by (3.1.1) is called the *derivative of the measure* μ *along the vector* h or *in the direction* h (it is also called *Fomin's derivative*).

Higher order derivatives $d_h^n\mu$ as well as mixed derivatives $d_{h_1}\cdots d_{h_n}\mu$ are defined inductively. This will be discussed below in §3.7.

3.1.2. Definition. *A measure* μ *on* $\mathcal{A}$ *is called continuous along the vector* h *if we have*

$$\lim_{t\to 0} \|\mu_{th} - \mu\| = 0. \tag{3.1.2}$$

It is clear that if the measure μ is continuous along the vector h, then for every set $A \in \mathcal{A}$ the function $t \mapsto \mu(A + th)$ is continuous. As we shall see below, the converse is also true.

3.1.3. Definition. *A measure* μ *on* $\mathcal{A}$ *is called quasi-invariant along the vector* h *if the measures* μ_{th} *and* μ *are equivalent for every real* t.

It is readily verified that a measure μ continuous or quasi-invariant along two vectors h and k has the same property also along their linear combinations.

Two other concepts of differentiability were introduced by Albeverio and Høegh-Krohn [**41**] and Skorohod [**1046**].

3.1.4. Definition. *We shall say that a nonnegative measure* μ *on* $\mathcal{A}$ *is differentiable along the vector* h *in the sense of Albeverio and Høegh-Krohn (or* L^2*-differentiable) if there exists a measure* $\lambda \geqslant 0$ *on* $\mathcal{A}$ *such that* $\mu_{th} \ll \lambda$, $\mu_{th} = f_t \cdot \lambda$ *and the mapping*

$$t \mapsto f_t^{1/2}$$

from $\mathbb{R}^1$ *to* $L^2(\lambda)$ *is differentiable (with respect to the norm of* $L^2(\lambda)$*).*

More generally, given $p \in [1, +\infty)$, *let us define* L^p*-differentiability of the measure* μ *as the differentiability of the mapping* $t \mapsto f_t^{1/p}$ *from* $\mathbb{R}^1$ *to the space* $L^p(\lambda)$, *where we have* $\mu_t = f_t \cdot \lambda$ *for all* t.

Let us note at once that L^2-differentiability implies Fomin's differentiability and the equality

$$d_h\mu = 2f_0^{1/2}\psi \cdot \lambda, \quad \text{where} \quad \psi := \frac{d}{dt}\sqrt{f_t}\Big|_{t=0}. \tag{3.1.3}$$

Indeed, for every $A \in \mathcal{A}$, the function

$$t \mapsto \mu_{th}(A) = \int_A f_t(x)\,\lambda(dx) = (I_A f_t^{1/2}, f_t^{1/2})_{L^2(\lambda)}$$

is differentiable and its derivative at zero is $2(I_A\psi, f_0^{1/2})_{L^2(\lambda)}$. It will be seen from the results in §3.3 that Fomin's differentiability of a nonnegative measure is equivalent to its L^1-differentiability.

3.1.5. Definition. *A Baire measure μ on a locally convex space X is called Skorohod differentiable or S-differentiable along the vector h if, for every function $f \in C_b(X)$, the function*

$$t \mapsto \int_X f(x - th)\,\mu(dx)$$

is differentiable. Skorohod differentiability of a Borel measure is understood as the differentiability of its restriction to the Baire σ-algebra.

It follows from this definition that, for every sequence $t_n \to 0$, the sequence of Baire measures $t_n^{-1}(\mu_{t_n h} - \mu)$ is fundamental in the weak topology in the space of measures. By a theorem due to A.D. Alexandroff there exists a Baire measure ν which is the weak limit of the indicated sequence (see [**193**, Theorem 8.7.1]). The measure ν is independent of our choice of $\{t_n\}$ since the integral of any bounded continuous function f with respect to the measure ν equals the derivative at zero of the function indicated in the definition. Thus, for every bounded continuous function f on X we have

$$\lim_{t\to 0} \int_X \frac{f(x - th) - f(x)}{t}\,\mu(dx) = \int_X f(x)\,\nu(dx). \tag{3.1.4}$$

3.1.6. Definition. *A Baire measure ν on X satisfying identity* (3.1.4) *is called the Skorohod derivative* (*or the weak derivative*) *of the measure μ along h.*

Thus, every Skorohod differentiable Baire measure has a weak derivative. However, the space of Radon measures need not be sequentially complete in the weak topology (certainly, for most of the spaces encountered in real applications such problems do not arise since in such spaces the Baire and Borel σ-algebras coincide and all measures are Radon). Nevertheless, it will be shown below that for any Radon measure μ the Skorohod derivative is Radon as well.

The natural concept of analyticity of a measure was introduced by Albeverio and Høegh-Krohn [**41**] and Bentkus [**131**], [**132**].

3.1.7. Definition. *A measure μ on $\mathcal{A}$ is called analytic along the vector h if, for every set $A \in \mathcal{A}$, the function $t \mapsto \mu(A + th)$ admits a holomorphic extension to a circle of the form $\{z\colon\ |z| < c\}$ independent of A.*

Replacing A by $A + t_0h$ we see that for any analytic measure μ the function $t \mapsto \mu(A + th)$ has a holomorphic extension to the strip $\{z\colon\ |\mathrm{Im}\, z| < c\}$ independent of A.

3.1.8. Example. Let μ be a measure on $\mathbb{R}^1$ with a continuously differentiable density p and bounded support. Then the measure μ is continuous along every vector $h \in \mathbb{R}^1$ and is differentiable both in the sense of Fomin and Skorohod.

PROOF. Let $h = 1$. Then

$$\frac{\mu(A+t)-\mu(A)}{t} = \int_A \frac{p(x+t)-p(x)}{t}\,dx \to \int_A p'(x)\,dx \text{ as } t \to 0.$$

So the measure μ is Fomin differentiable. In addition, the measure $d_1\mu$ has density p'/p with respect to μ. Similarly, it is Skorohod differentiable. It is clear that

$$\lim_{t\to 0} \|\mu_t - \mu\| = \lim_{t\to 0} \int |p(x+t) - p(x)|\,dx = 0,$$

i.e., the measure μ is continuous. □

Below we shall encounter less trivial examples (and also completely characterize continuous and differentiable measures on the real line). However, even this example exhibits two interesting properties of Fomin differentiable measures: the variation $|\mu|$ of a differentiable measure is again differentiable (in our example one can verify that the measure with density $|p|$ is differentiable) and the derivative $d_h\mu$ is absolutely continuous with respect to μ. It turns out that these two properties have a very general character.

Everywhere below the term "differentiable measure" will be used for the differentiability in the sense of Fomin and all other kinds of differentiability will be properly specified.

Now we give an important example of a differentiable measure on an infinite dimensional space.

3.1.9. Theorem. *Let γ be a centered Radon Gaussian measure on a locally convex space X. Then*

(i) *its Cameron–Martin space $H(\gamma)$ coincides with the collection of all vectors of continuity as well as with the collection of all vectors of differentiability;*

(ii) *for every $h \in H(\gamma)$ we have $d_h\gamma = -\widehat{h}\cdot\gamma$, where $\widehat{h}$ is the measurable linear functional generated by h (see Chapter 1).*

PROOF. By Theorem 1.4.2 for any $h \in H(\gamma)$ the measure γ_h is given by density $\exp(-\widehat{h} - |h|_H^2/2)$ with respect to the measure γ and for any other shift we have $\gamma_h \perp \gamma$. The ratio $t^{-1}\bigl[\exp(-t\widehat{h} - t^2|h|_H^2/2) - 1\bigr]$ has a limit $-\widehat{h}$ in $L^1(\gamma)$ as $t \to 0$. This is seen from the fact that for $|t| \leqslant 1$ this ratio is majorized in the absolute value by the integrable function $|\widehat{h}|\exp|\widehat{h}|$ and its pointwise limit is $-\widehat{h}$. Hence the measures $t^{-1}(\gamma_{th} - \gamma)$ converge in variation to the measure $-\widehat{h}\cdot\gamma$, which means that $d_h\gamma = -\widehat{h}\cdot\gamma$. □

3.2. Properties of continuous measures

Here we establish a number of elementary properties of continuous measures. Let us observe that by the estimate from Proposition 1.2.1 the continuity of μ along h is characterized in terms of the Hellinger integral as the relationship

$$H(\mu, \mu_{th}) \to \|\mu\| \quad \text{as} \quad t \to 0.$$

3.2.1. Proposition. *Suppose that a measure μ on $\mathcal{A}$, where $\mathcal{A}$ is one of our main σ-algebras $\mathcal{B}a(X)$, $\sigma(X)$ or $\mathcal{B}(X)$ (where in the case of $\mathcal{B}(X)$ the measure is Radon) is continuous along h. Then every measure ν that is absolutely continuous with respect to μ is continuous along h as well.*

Proof. By the Radon–Nikodym theorem $\nu = f \cdot \mu$, where $f \in \mathcal{L}^1(\mu)$. Since $(f\cdot\mu)_{th} = f(\,\cdot + th)\cdot\mu_{th}$, as noted in §3.1, assuming that the function f is bounded and continuous along the vector h, we obtain

$$\begin{aligned}\|\nu_{th} - \nu\| &\leqslant \|f(\,\cdot + th)\cdot(\mu_{th} - \mu)\| + \|(f(\,\cdot + th) - f\cdot\mu\| \\ &\leqslant \sup_x |f(x)|\,\|\mu_{th} - \mu\| + \|f(\,\cdot + th) - f)\|_{L^1(|\mu|)},\end{aligned}$$

which, as $t \to 0$, tends to zero by the continuity of μ along h and the Lebesgue dominated convergence theorem. In the general case, given $\varepsilon > 0$, it suffices to find a bounded function g continuous along the vector h such that $\|f - g\|_{L^1(|\mu|)} \leqslant \varepsilon$. This is possible in all three cases mentioned in the theorem. Then it remains to employ the estimate

$$\begin{aligned}\|(f\cdot\mu)_{th} - f\cdot\mu\| &\leqslant \|(f\cdot\mu)_{th} - (g\cdot\mu)_{th}\| + \|(g\cdot\mu)_{th} - g\cdot\mu\| + \|g\cdot\mu - f\cdot\mu\| \\ &= 2\|g\cdot\mu - f\cdot\mu\| + \|(g\cdot\mu)_{th} - g\cdot\mu\|,\end{aligned}$$

which follows from the invariance of the variation with respect to shifts. □

3.2.2. Corollary. *A measure μ on the σ-algebra $\mathcal{A}$ from the previous proposition is continuous along h if and only if the measure $|\mu|$ is also.*

Let ϱ be a fixed probability density on the real line with Lebesgue measure. Set $\sigma := \varrho\, dx$.

3.2.3. Corollary. *Suppose that for every $A \in \mathcal{A}$ the map $(x, t) \mapsto x - th$ from $X \times \mathbb{R}^1$ to X is measurable with respect to the pair of σ-algebras $\mathcal{A}\otimes\mathcal{B}(\mathbb{R}^1)$ and $\mathcal{A}$ (which is fulfilled if $\mathcal{A} = \sigma(X)$ or $\mathcal{A} = \mathcal{B}(X)$). A measure μ on $\mathcal{A}$ is continuous along h precisely when it is absolutely continuous with respect to the measure*

$$|\mu| * \sigma(A) := \int_{-\infty}^{+\infty} |\mu|(A - th)\,\varrho(t)\,dt.$$

Proof. We observe that due to our assumption $|\mu| * \sigma$ is a nonnegative measure on $\mathcal{A}$. If the measure μ is continuous along h and $|\mu| * \sigma(A) = 0$, then $|\mu|(A - th) = 0$ for almost all t, which by the continuity of $|\mu|$ along h yields the equality $\mu(A) = 0$. So $\mu \ll |\mu| * \sigma$. For the proof of the converse it suffices to verify the continuity of $|\mu| * \sigma$ along h. Let ϱ be a smooth function with support in the interval $[-M, M]$. Then, for every $A \in \mathcal{A}$ and $|t| \leqslant 1$, we have

$$|\mu| * \sigma(A - th) - |\mu| * \sigma(A) = \int_{-\infty}^{+\infty} [|\mu|(A - th - sh) - |\mu|(A - sh)]\varrho(s)\, ds$$

$$= \int_{-\infty}^{+\infty} |\mu|(A - zh)[\varrho(z - t) - \varrho(z)]\, dz \leqslant 2(M + 1)|t|\|\mu\| \sup_z |\varrho'(z)|.$$

So $\|(|\mu| * \sigma)_{th} - |\mu| * \sigma\| \to 0$ as $t \to 0$. For completing the proof it remains to approximate the density ϱ by smooth densities with bounded support in the norm of $L^1(\mathbb{R}^1)$. Finally, the measurability of the mapping $(x, t) \mapsto x - th$ in the case $\mathcal{A} = \sigma(X)$ or $\mathcal{A} = \mathcal{B}(X)$ follows from Lemma 1.2.15. □

3.2.4. Corollary. *A measure μ on the σ-algebra $\sigma(X)$ in a locally convex space X is continuous along h precisely when it is absolutely continuous with respect to the measure $|\mu| * \gamma$, where γ is the image of the standard Gaussian measure on the real line under the mapping $t \mapsto th$. The same is true for Radon measures on $\mathcal{B}(X)$.*

3.2.5. Corollary. *The measure μ on $\sigma(X)$ is continuous along h if and only if the function $t \mapsto \mu(A + th)$ is continuous for every $A \in \sigma(X)$. In the case of Radon measures the same is true for $\mathcal{B}(X)$ in place of $\sigma(X)$.*

In addition, the continuity of a Radon measure μ along h follows from the continuity of all functions $t \mapsto \mu(A + th)$, where $A \in \sigma(X)$.

Proof. The continuity of the measure μ along h implies the continuity of the functions $t \mapsto \mu(A + th)$ for all $A \in \sigma(X)$ (in the case of a Radon measure also for $A \in \mathcal{B}(X)$). Suppose that all these functions are continuous. Then $\mu \ll |\mu| * \sigma$ (where the measure σ is the same as above). Indeed, if $|\mu| * \sigma(A) = 0$, then there exists a sequence $t_n \to 0$ with $|\mu|(A - t_n h) = 0$, whence we find $\mu(A - t_n h) = 0$ and so $\mu(A) = 0$. Hence the measure μ is continuous along h. In the case of a Radon measure the continuity along h on $\sigma(X)$ yields at once the continuity on $\mathcal{B}(X)$. □

3.2.6. Corollary. *Let a σ-algebra $\mathcal{A}$ be the same as in the proposition above and let a sequence of measures μ_n continuous along h be such that for every $A \in \mathcal{A}$ there exists a finite limit $\mu(A) = \lim_{n\to\infty} \mu_n(A)$. Then the measure μ is continuous along h as well.*

Proof. By the Nikodym theorem μ is indeed a measure and the measures μ_n are uniformly bounded in variation. The measure $\nu = \sum_{n=1}^{\infty} 2^{-n}|\mu_n|$ is continuous along h. It is clear that $\mu \ll \nu$. □

3.2.7. Remark. If a measure μ is continuous along h, then so is the measure $I_A \cdot |\mu|$ for every $A \in \mathcal{B}(X)$. Hence, for every set $A \in \mathcal{B}(X)$ with $|\mu|(A) > 0$, we obtain $(A + th) \cap A \neq \varnothing$ for sufficiently small t. In particular, h belongs to the linear span of every compact set K with $|\mu|(K) > 0$. It follows that the property of a measure to be continuous along h is invariant with respect to the space carrying the measure: if X is embedded by means of an injective continuous linear operator to a locally convex space Y, then the measure μ considered on Y is continuous along the same vectors along which it is continuous on X. This is true, of course, for other differential properties introduced in §3.1.

Let us mention the following result from Yamasaki, Hora [**1198**].

3.2.8. Proposition. *There is no nonzero Borel measure on an infinite dimensional locally convex space that is continuous or quasi-invariant along all vectors.*

The proof follows from the next proposition, which is of interest in its own right and shows that in infinite dimensions a measure has no minimal linear subspace of full measure unless it is concentrated on a finite dimensional subspace.

3.2.9. Proposition. *Let X be an infinite dimensional locally convex space and let μ be a Borel probability measure on X. Then there is a Borel proper linear subspace $L \subset X$ with $\mu(L) > 0$.*

Proof. Suppose that $\mu(L) = 0$ for every Borel proper linear subspace. Let us equip X^* with the following metric corresponding to convergence in measure μ:

$$d(f, g) := \int_X \min(|f(x) - g(x)|, 1)\, \mu(dx).$$

Note that d is indeed a metric since if $l \in X^*$ vanishes μ-a.e. we have $l = 0$ due to our assumption (otherwise $L = l^{-1}(0)$ is a Borel linear subspace of positive measure). Set

$$V_n := \{l \in X^*\colon\ d(l, 0) \leqslant n^{-1}\}, \quad U_n := \{x \in X\colon \sup_{l \in V_n} |l(x)| \leqslant 1\}.$$

The sets U_n are closed. We observe that $X = \bigcup_{n=1}^{\infty} U_n$. Indeed, otherwise we would have an element $x_0 \in X$ and functionals l_n with $d(l_n, 0) \leqslant n^{-1}$ and $l_n(x_0) > 1$. The sequence $\{l_n\}$ converges to zero in measure, hence there is a subsequence $\{l_{n_i}\}$ convergent to zero μ-a.e. The linear subspace L consisting of all $x \in X$ such that $\lim\limits_{i\to\infty} l_{n_i}(x) = 0$ is Borel and has positive measure, but $x_0 \notin L$, which is impossible. There is n such that $\mu(U_n) > 0$. This yields that the metric d on X^* generates the same convergence as the norm from $L^2(\mu|_{U_n})$. Indeed, let $l_i \to 0$ with respect to this norm. Then there is a subsequence $\{l_{i_j}\}$ convergent to zero μ-a.e. on U_n. As above, this shows that $\{l_{i_j}\}$ converges to zero pointwise, hence in metric d. It follows that the whole sequence converges to zero in metric d. Conversely, if $d(l_i, 0) \to 0$, then we may assume that $\{l_i\} \subset V_n$, hence $\sup_{x \in U_n} |l_i(x)| \leqslant 1$, which yields convergence $l_i \to 0$ in $L^2(\mu|_{U_n})$ by the dominated convergence theorem. Now we arrive at a contradiction by showing that X^* must be finite dimensional. Indeed, we have proved that there is $R > 0$ such that the unit ball B of X^* with the $L^2(\mu|_{U_n})$-norm is contained in RV_n. Hence, for any collection $l_1, \ldots, l_k \in X^*$ orthonormal in $L^2(\mu|_{U_n})$, we have

$$\begin{aligned} k &= \sum_{i=1}^{k} \int_{U_n} |l_i|^2 \, d\mu \leqslant \sup_{x \in U_n} \sum_{i=1}^{k} |l_i(x)|^2 \\ &= \sup_{x \in U_n} \Bigl| \Bigl\langle \Bigl(\sum_{i=1}^{k} |l_i(x)|^2 \Bigr)^{-1/2} \sum_{i=1}^{k} l_i(x) l_i, x \Bigr\rangle \Bigr|^2 \leqslant \sup_{x \in U_n} \sup_{l \in B} |l(x)|^2 \\ &\leqslant \sup_{x \in U_n} \sup_{l \in RV_n} |l(x)|^2 \leqslant R^2, \end{aligned}$$

since for any fixed x, the $L^2(\mu|_{U_n})$-norm of $\sum_{i=1}^{k} l_i(x) l_i$ is $\bigl(\sum_{i=1}^{k} |l_i(x)|^2\bigr)^{1/2}$. This shows that $\dim X^* \leqslant R^2$. □

3.3. Properties of differentiable measures

We assume that $\mathcal{A}$ is one of our three main σ-algebras and that in the case of the Borel σ-algebra the considered measures are Radon, although some of the results below are also valid in more general cases. In particular, the following theorem is valid under the only assumption of invariance of $\mathcal{A}$ with respect to the shifts along the vectors th, which will be obvious from the proof.

3.3.1. Theorem. *Suppose that a measure μ on $\mathcal{A}$ is Fomin differentiable along the vector h. Then, its positive part μ^+, its negative part μ^-, and its variation $|\mu|$ are Fomin differentiable along h as well. In addition, if X^+ and X^- are the sets from the Hahn–Jordan decomposition for μ, then*

$$d_h(\mu^+) = d_h\mu(\,\cdot \cap X^+), \quad d_h(\mu^-) = d_h\mu(\,\cdot \cap X^-).$$

Proof. Let $X = X^+ \cup X^-$ be the Hahn decomposition for the measure μ. Then the restriction of μ to X^+ is nonnegative and its restriction to X^- is nonpositive. The function $\varphi(t) = \mu(X^+ + th)$ is differentiable and attains its maximum at zero since $\mu(X^+)$ is the maximal value of μ. Hence its derivative at zero vanishes, i.e., as $t \to 0$ we have

$$\frac{\varphi(t) - \varphi(0)}{t} = t^{-1}\mu\bigl((X^+ + th)\backslash X^+\bigr) - t^{-1}\mu\bigl(X^+\backslash(X^+ + th)\bigr) \to 0.$$

Since $\mu\bigl((X^+ + th)\backslash X^+\bigr) \leqslant 0$ and $\mu\bigl(X^+\backslash(X^+ + th)\bigr) \geqslant 0$, we obtain as $t \to 0$ that

$$t^{-1}\mu\bigl((X^+ + th)\backslash X^+\bigr) \to 0, \quad t^{-1}\mu\bigl(X^+\backslash(X^+ + th)\bigr) \to 0.$$

Set $X_t := X^+ + th$. It is easy to verify the equality

$$X^+ = \bigl(X_t \cup (X^+\backslash X_t)\bigr)\backslash(X_t\backslash X^+).$$

Therefore, for every $A \in \mathcal{A}$, as $t \to 0$ we have

$$\frac{\mu^+(A+th) - \mu^+(A)}{t} = \frac{\mu\big((A+th)\cap X^+\big) - \mu(A\cap X^+)}{t}$$
$$= \frac{\mu\big((A+th)\cap X_t\big) - \mu(A\cap X^+)}{t} + \frac{\mu\big(A\cap (X^+\backslash X_t)\big) - \mu\big(A\cap (X_t\backslash X^+)\big)}{t}$$
$$\to d_h\mu(A\cap X^+)$$

since the second term in the second equality tends to zero according to what has been said above. The reasoning is similar for μ^-. □

3.3.2. Corollary. *If a measure μ is Fomin differentiable along h, then one has $d_h\mu \ll \mu$.*

Proof. The previous theorem reduces the general case to the case of a nonnegative measure μ, in which our claim is obvious since if $\mu(A) = 0$, then the function $t \mapsto \mu(A + th)$ has minimum at zero, whence it follows that its derivative vanishes at zero. □

3.3.3. Corollary. *If a measure μ is Fomin differentiable along a vector h, then the measure $d_h\mu$ is continuous along h. In particular, all functions $t \mapsto \mu(A + th)$, where $A \in \mathcal{A}$, are continuously differentiable.*

3.3.4. Proposition. *A measure μ is Fomin differentiable along the vector h precisely when there exists a measure ν on $\mathcal{A}$ continuous along h such that*

$$\mu(A+th) = \mu(A) + \int_0^t \nu(A+sh)\,ds, \quad \forall\, A\in\mathcal{A}, t\in\mathbb{R}^1. \tag{3.3.1}$$

In this case $d_h\mu = \nu$. In addition, for every bounded $\mathcal{A}$-measurable function f for all $t \in \mathbb{R}^1$ the following equality holds:

$$\int_X f(x)\,(\mu_{th} - \mu)(dx) = \int_0^t \int_X f(x-sh)\,d_h\mu(dx)\,ds. \tag{3.3.2}$$

Finally, yet another necessary and sufficient condition for Fomin's differentiability of μ is the existence of a measure $\nu \ll \mu$ such that one has (3.3.1).

Proof. Let the measure μ be Fomin differentiable along h. Then the derivative of the function $\varphi\colon\, t \mapsto \mu(A + th)$ is $d_h\mu(A + th)$. So we have $|\varphi'(t)| \leqslant \|d_h\mu\|$, i.e., φ is Lipschitzian, which yields equality (3.3.1) with $\nu = d_h\mu$. Then we have (3.3.2) for simple functions f, which gives (3.3.2) for bounded functions f by means of uniform approximations by simple functions. As we have proved earlier, $d_h\mu \ll \mu$ and the measure $d_h\mu$ is continuous along h. The converse is obvious. If we are given identity (3.3.1) with some measure $\nu \ll \mu$, then we obtain the continuity of the measure μ along h, which implies the continuity of ν along h. □

3.3.5. Corollary. *Suppose that in the situation of the previous proposition for μ-a.e. x there exists* $\partial_h f(x) := \lim\limits_{t\to 0} t^{-1}[f(x+th) - f(x)]$ *and*

$$|t^{-1}[f(x+th) - f(x)]| \leqslant g(x), \quad \textit{where } g \in \mathcal{L}^1(\mu).$$

Then

$$\int_X \partial_h f(x)\,\mu(dx) = -\int_X f(x)\,d_h\mu(dx). \tag{3.3.3}$$

PROOF. On the left in (3.3.2) we have the integral of $f(x-th) - f(x)$, which enables us to apply the Lebesgue dominated convergence theorem to the sequence of functions $n[f(x-h/n) - f(x)]$. □

3.3.6. Corollary. *If a measure μ is differentiable along two vectors h and k, then it is differentiable along every linear combination of these vectors and*

$$d_{\alpha h+\beta k}\mu = \alpha d_h\mu + \beta d_k\mu.$$

PROOF. The cases $k = 0$ and $h = 0$ are obvious. So it suffices to consider the case $\alpha = \beta = 1$. Let us recall that the measure $d_h\mu$ is continuous along h. In addition, it is continuous also along k since it is a limit of the sequence of measures $n(\mu_{n^{-1}h} - \mu)$ that are continuous along k by the differentiability along k. We have

$$\begin{aligned}\mu(A+th+tk) - \mu(A) &= \mu(A+th+tk) - \mu(A+tk) + \mu(A+tk) - \mu(A)\\ &= \int_0^t d_h\mu(A+tk+sh)\,ds + \int_0^t d_k\mu(A+sk)\,ds.\end{aligned}$$

The right-hand side divided by t tends to $d_h\mu(A) + d_k\mu(A)$ as $t \to 0$. This follows from the continuity of the function $(t,s) \mapsto d_h(A+tk+sh)$ of two variables, which is clear from the estimate

$$\begin{aligned}&|d_h\mu(A+tk+sh) - d_h\mu(A+t_0k+s_0h)|\\ &\qquad\leqslant |d_h\mu(A+tk+sh) - d_h\mu(A+t_0k+sh)|\\ &\qquad+ |d_h\mu(A+t_0k+sh) - d_h\mu(A+t_0k+s_0h)|\\ &\qquad\qquad\leqslant \|(d_h\mu)_{tk} - (d_h\mu)_{t_0k}\| + \|(d_h\mu)_{sh} - (d_h\mu)_{s_0h}\|\end{aligned}$$

and the continuity of $d_h\mu$ along h and k. □

3.3.7. Theorem. *Suppose that a measure μ is differentiable along the vector h. Then*

(i) *one has* $\|\mu_h - \mu\| \leqslant \|d_h\mu\|$;

(ii) *for every $t \in \mathbb{R}^1$ one has*

$$\|\mu_{th} - \mu - t d_h\mu\| \leqslant |t| \sup_{0<\tau<t} \|(d_h\mu)_{\tau h} - d_h\mu\|;$$

(iii) *one has* $\big\|(\mu_{th} - \mu)/t - d_h\mu\big\| \to 0$ *as* $t \to 0$.

PROOF. Assertion (i) is readily seen from equality (3.3.2). The same relationship yields (ii). Finally, (iii) follows from (ii) by the continuity of $d_h\mu$ along h. □

3.3.8. Definition. *The Radon–Nikodym density of the measure $d_h\mu$ with respect to μ is denoted by β_h^μ and is called the logarithmic derivative of μ along h.*

This terminology is justified by Example 3.1.8. As we have seen in Theorem 3.1.9, if γ is a Gaussian measure and $h \in H(\gamma)$, then $\beta_h^\gamma = -\widehat{h}$ is a measurable linear function.

It follows from what has been said above that

$$d_h\mu = \beta_h^{\mu^+} \cdot \mu^+ - \beta_h^{\mu^-} \cdot \mu^-, \quad d_h|\mu| = \beta_h^{\mu^+} \cdot \mu^+ + \beta_h^{\mu^-} \cdot \mu^-.$$

Since the measures μ^+ and μ^- are mutually singular, $|\mu|$-a.e. we have

$$|\beta_h^\mu| = |\beta_h^{\mu^+}| + |\beta_h^{\mu^-}| = |\beta_h^{|\mu|}|. \tag{3.3.4}$$

This relationship is analogous to the equality $|\nabla f| = \bigl|\nabla|f|\bigr|$ a.e. for functions from Sobolev classes.

If a nonnegative measure μ is L^2-differentiable along h, then equality (3.1.3) yields the following representation of the logarithmic derivative:

$$\beta_h^\mu = 2\psi f_0^{-1/2}, \tag{3.3.5}$$

where $\mu_{th} = f_t \cdot \lambda$ and ψ is the derivative at zero of the mapping $t \mapsto f_t^{1/2}$ to $L^2(\lambda)$. Since $\psi \in L^2(\lambda)$, one has $\beta_h^\mu \in L^2(\mu)$. The latter condition characterizes L^2-differentiability. An analogous assertion is true for all numbers $p \in (1, +\infty)$.

3.3.9. Theorem. *A measure $\mu \geqslant 0$ is L^p-differentiable along h precisely when it is Fomin differentiable along h and $\beta_h^\mu \in L^p(\mu)$. If $p > 1$, this is also equivalent to the estimate*

$$\|f_t^{1/p} - f_0^{1/p}\|_{L^p(\lambda)} \leqslant C|t|.$$

In this case one can take $C = \|\beta_h^\mu\|_{L^p(\mu)}/p$.

A short proof of a more general fact (for differentiable families of measures) is given in Theorem 11.2.13 in Chapter 11, so here we do not reproduce this proof.

For $p = 2$, by using the so-called Hellinger integral

$$H(\mu_{th}, \mu) := \int_X f_t^{1/2}(x) f_0^{1/2}(x)\, \lambda(dx)$$

this estimate can be written in the form

$$H(\mu_{th}, \mu) \geqslant 1 - t^2 \|\beta_h^\mu\|_{L^2(\mu)}^2 / 8.$$

Let us observe that the quantity $H(\mu_{th}, \mu)$ is independent of the measure λ with the property that $\mu_{th} \ll \lambda$ for all t.

3.3.10. Proposition. *A measure μ on $\mathcal{A}$ is analytic along h if and only if there exists a number $\delta > 0$ such that for all complex numbers z with $|z| < \delta$ the series*

$$\sum_{n=1}^{\infty} \frac{1}{n!} z^n d_h^n \mu$$

converges in variation in the Banach space of complex measures on X. In this case, for some C, $r > 0$ one has the estimate

$$\|d_h^n \mu\| \leqslant C r^n n!, \quad n \in \mathbb{N},$$

and μ is quasi-invariant along h.

Proof. Convergence of the series obviously yields the analyticity of the measure. Conversely, if the measure is analytic, then the series with terms $z^n d_h^n \mu(A)/n!$ converges whenever $|z| < c$, $A \in \mathcal{A}$, which for any fixed z with $|z| < c$ implies the uniform boundedness of the quantities $|z|^n \|d_h^n \mu\|/n!$. This gives convergence of the series indicated in the theorem and the estimate $\|d_h^n \mu\| \leqslant C(c/2)^{-n} n!$ for some number $C > 0$. If $|\mu|(A) = 0$ for some $A \in \mathcal{A}$, then by the relationships $d_h^{k+1}\mu \ll d_h^k \mu$ we obtain $d_h^k \mu(A) = 0$ for all $k \geqslant 1$, whence $\mu(A + th) = 0$ for all $t \in \mathbb{R}^1$. Hence $\mu_{th} \ll \mu$. Therefore, we have the relationship $\mu_{th} \sim \mu$. □

3.3.11. Corollary. *If a Radon measure μ on a locally convex space X is analytic along a vector $h \in X$, then $d_h^n \mu \sim \mu$ for all n.*

Proof. Let $|d_h \mu|(A) = 0$ for some set $A \in \mathcal{B}(X)$. Then $|d_h \mu|(K) = 0$ for every compact set $K \subset A$. According to what has been proved above, $|d_h^n \mu|(K) = 0$ for all $n \geqslant 1$, whence $\mu(K + th) = \mu(K)$ for all $t \in \mathbb{R}^1$. Since μ is Radon, we obtain the equality $\lim_{t \to \infty} \mu(K + th) = 0$. Thus, $\mu(K) = 0$ and so $|\mu|(A) = 0$, i.e., $\mu \sim d_h \mu$. Then $\mu \sim d_h^n \mu$ for all n. □

Let us discuss some elementary operations on differentiable measures.

3.3.12. Proposition. *Let μ be a measure differentiable along h and let f be a bounded measurable function possessing a uniformly bounded partial derivative $\partial_h f$. Then, the measure $\nu = f \cdot \mu$ is differentiable along h as well and one has*

$$d_h \nu = \partial_h f \cdot \mu + f \cdot d_h \mu.$$

Proof. For every $A \in \mathcal{A}$ we have

$$\begin{aligned}
\nu(A + th) - \nu(A) &= \int_X [I_A(x - th) - I_A(x)] f(x)\, \mu(dx) \\
&= \int_X I_A(y) f(y + th)\, \mu_{th}(dy) - \int_X I_A(x) f(x)\, \mu(dx) \\
&= \int_X I_A(y)[f(y + th) - f(y)]\, \mu_{th}(dy) + \int_X I_A(x) f(x)\, (\mu_{th} - \mu)(dx).
\end{aligned}$$

Substituting in place of t points $t_n \to 0$ and dividing by t_n we obtain in the limit the sum of the integrals of $I_A \partial_h f$ and $I_A f \beta_h^\mu$ against the measure μ. This follows from our assumptions and Theorem 1.2.19. □

3.3.13. Proposition. *Let X be a locally convex space equipped with one of our three σ-algebras $\mathcal{A}$. Let μ be a measure on $\mathcal{A}$ possessing one of the properties along h described by our definitions from* §3.1 *and let ν be a measure on a locally convex space Y equipped with a σ-algebra of the same kind as $\mathcal{A}$. Then the following assertions are true.*

(i) *If $T\colon\ X \to Y$ is a continuous linear mapping (or, in the case of Fomin differentiability, a linear mapping measurable with respect to μ), then the measure $\mu \circ T^{-1}$ has the same property along Th that μ has along h. In the case of differentiability one has*

$$d_{Th}(\mu\circ T^{-1}) = (d_h\mu)\circ T^{-1}.$$

In the case of a Fomin differentiable probability measure μ, the logarithmic derivative $\beta_{Th}^{\mu\circ T^{-1}}$ is a Borel function β on Y such that the function $\beta\circ T$ coincides with the conditional expectation $\mathbb{E}^T\beta_h^\mu$ of the function β_h^μ with respect to the σ-algebra σ_T generated by the mapping T. Conversely, every Borel function β on Y for which $\beta\circ T = \mathbb{E}^T\beta_h^\mu$ μ-a.e. coincides with $\beta_{Th}^{\mu\circ T^{-1}}$ almost everywhere with respect to the measure $\mu\circ T^{-1}$.

(ii) *The measure $\mu\otimes\nu$ on $X\times Y$ has the same property along $(h,0)$ that μ has along h.*

(iii) *If $X = Y$, then the measure $\mu * \nu$ has the corresponding property along h and in the case of differentiability*

$$d_h(\mu * \nu) = d_h\mu * \nu.$$

Proof. The verification is direct by using the definitions and Fubini's theorem. Let us only explain the last claim in (i). For definiteness we shall assume that we deal with Borel measures. For every $B \in \mathcal{B}(Y)$ on account of the equality in (i) we have

$$\int_{T^{-1}(B)} \beta_h^\mu(x)\,\mu(dx) = \int_B \beta_{Th}^{\mu\circ T^{-1}}(y)\,\mu\circ T^{-1}(dy) = \int_{T^{-1}(B)} \beta_{Th}^{\mu\circ T^{-1}}(Tx)\,\mu(dx).$$

It follows that the function $\beta_{Th}^{\mu\circ T^{-1}}\circ T$ coincides with the conditional expectation of β_h^μ with respect to the σ-algebra σ_T since the latter is the class of all sets of the form $T^{-1}(B)$, where $B \in \mathcal{B}(Y)$. Conversely, if a Borel function β on Y is such that we have $\beta\circ T = \mathbb{E}^T\beta_h^\mu$ μ-a.e., then

$$\begin{aligned}\int_B \beta(y)\,\mu\circ T^{-1}(dy) &= \int_{T^{-1}(B)} \beta(Tx)\,\mu(dx) = \int_{T^{-1}(B)} \mathbb{E}^T\beta_h^\mu(x)\,\mu(dx)\\ &= \int_{T^{-1}(B)} \beta_h^\mu(x)\,\mu(dx) = d_h\mu\bigl(T^{-1}(B)\bigr) = d_{Th}(\mu\circ T^{-1}),\end{aligned}$$

whence it follows that β coincides with $\beta_{Th}^{\mu\circ T^{-1}}$. □

3.3.14. Example. Let X and $\mathcal{A}$ be the same as in the proposition above, let μ be a measure on $\mathcal{A}$, and let f be a bounded $\mathcal{A}$-measurable function on X. If μ is Fomin differentiable along h, then the function

$$f * \mu(x) := \int_X f(x-y)\,\mu(dy)$$

has a partial derivative

$$\partial_h(f*\mu)(x) = \int_X f(x-y)\,d_h\mu(dy).$$

Proof. If f is the indicator function of a measurable set, then this is the definition. Therefore, our claim is true for linear combinations g of indicator functions. Let x be fixed. According to Corollary 3.3.3, the functions $t \mapsto d_h\mu(A+x+th)$ are continuous. Hence so are the functions

$$t \mapsto \int_X g(x+th-y)\,d_h\mu(dy).$$

It remains to choose a sequence of functions g_n of the above type which converges to f uniformly on X. Then the function $t \mapsto f*\mu(x+th)$ is differentiable being a uniform limit of differentiable functions whose derivatives converge uniformly. □

3.4. Differentiable measures on $\mathbb{R}^n$

In this section we discuss differentiable measures on the space $\mathbb{R}^n$. The term "measure on $\mathbb{R}^n$" will always mean a bounded Borel measure on $\mathbb{R}^n$ unless we deal with Lebesgue measure (which will always be explicitly explained). First we consider the case $n=1$.

3.4.1. Proposition. *Let μ be a measure on $\mathbb{R}^1$.*

(i) *The measure μ is continuous along $h \neq 0$ if and only if it is absolutely continuous with respect to Lebesgue measure.*

(ii) *The measure μ is Skorohod differentiable along $h \neq 0$ if and only if it has a density of bounded variation. In particular, any Skorohod differentiable measure on $\mathbb{R}^1$ admits a bounded density.*

(iii) *The measure μ is Fomin differentiable along $h \neq 0$ if and only if it has an absolutely continuous density ϱ whose derivative ϱ' is integrable. In this case $d_1\mu = \varrho'dx$.*

(iv) *The measure μ is quasi-invariant along $h \neq 0$ if and only if it has a density ϱ such that $\varrho \neq 0$ almost everywhere.*

(v) *If $\mu \geqslant 0$, then μ is L^2-differentiable along $h \neq 0$ if and only if it admits an absolutely continuous density ϱ for which $|\varrho'|^2/\varrho \in L^1(\mathbb{R}^1)$, where we set $0/0 := 0$. The quantity*

$$I(\varrho) := \int_{-\infty}^{+\infty} \frac{|\varrho'(t)|^2}{\varrho(t)}\,dt$$

is called the Fisher information of the measure μ or the Fisher information of the density ϱ.

PROOF. To prove (i) we observe that if μ is continuous along 1, then $|\mu|$ is continuous as well, which easily yields that $|\mu| \ll |\mu| * \sigma$, where σ is any probability measure with a smooth density (if $|\mu|(A-x) = 0$ for a.e. x, then $|\mu|(A) = 0$ by the continuity). Similar justification of (iv) is left as an easy exercise (Exercise 3.9.1). Suppose that the measure μ is Skorohod differentiable along $h = 1$. Then it has a density ϱ. Denote by $J(f)$ the derivative at zero of the function

$$t \mapsto \int_{\mathbb{R}^1} f(x-t)\varrho(x)\,dx, \quad \text{where } f \in C_b(\mathbb{R}^1).$$

Let $\|f\| = \sup_x |f(x)|$. By the Banach–Steinhaus theorem there exists a number C such that $|J(f)| \leqslant C\|f\|$ for all $f \in C_b(\mathbb{R}^1)$. Therefore, there exists a bounded measure ν on $\mathbb{R}^1$, for which

$$J(f) = \int_{\mathbb{R}^1} f(x)\,\nu(dx), \quad \forall\, f \in C_0^1(\mathbb{R}^1).$$

Thus, ν is the derivative of ϱ in the sense of generalized functions. Now it is readily verified that ϱ coincides almost everywhere with the distribution function of ν. Hence ϱ is of bounded variation. Conversely, suppose that μ admits a density ϱ of bounded variation. Then its derivative in the sense of generalized functions is a bounded measure ν. For every $f \in C_0^\infty(\mathbb{R}^1)$ this yields

$$\int_{\mathbb{R}^1} [f(x-t) - f(x)]\,\mu(dx) = \int_0^t \int_{\mathbb{R}^1} f(x-s)\,\nu(dx)\,ds$$

since the derivatives of both sides coincide. Therefore, this integral identity is true for all $f \in C_b(\mathbb{R}^1)$, which is Skorohod differentiability.

If the measure μ is differentiable, then according to what has been said above the measure $d_1\mu$ is absolutely continuous, hence has a density g. From our previous consideration we conclude that the indefinite integral of g (taken from $-\infty$) serves as a density for μ. Conversely, if the measure μ admits an absolutely continuous density ϱ, then it is Skorohod differentiable and ϱ' is a density of the measure $d_1\mu$. Therefore, μ is also Fomin differentiable (see Proposition 3.3.4).

Finally, if a nonnegative measure μ is L^2-differentiable, then, as noted above, it is Fomin differentiable and has an absolutely continuous density ϱ. It is clear from Theorem 3.3.9 that the function $|\varrho'|^2/\varrho$ belongs to $L^1(\mathbb{R}^1)$. The same theorem gives the converse. □

In terms of the theory of generalized functions (distributions) the obtained result has the following interpretation: Skorohod differentiability of a measure μ is equivalent to that its derivative in the sense of distributions is a bounded measure, and Fomin differentiability is equivalent to that its generalized derivative is given by an integrable function.

3.4.2. Example. (i) Let $p(t) = 1$ on $[0, 1]$ and $p(t) = 0$ at all other points. Then the measure μ with density p is Skorohod differentiable (but not Fomin differentiable) and we have the equality

$$d_1\mu = \delta_0 - \delta_1.$$

(ii) There exists a probability measure μ with an infinitely differentiable density on $\mathbb{R}^1$ which is not quasi-invariant, but $\mu(U) > 0$ for every open set U.

Proof. Assertion (i) is verified directly. For proving (ii) it suffices to find a smooth probability density whose zero set is nowhere dense, but is of positive Lebesgue measure. □

3.4.3. Proposition. (i) *A measure μ on $\mathbb{R}^n$ is continuous along n linearly independent directions precisely when μ is absolutely continuous.*

(ii) *A measure μ on $\mathbb{R}^n$ is Skorohod differentiable along n linearly independent vectors if and only if its generalized partial derivatives are bounded measures.*

(iii) *A measure μ on $\mathbb{R}^n$ is Fomin differentiable along n linearly independent vectors if and only if its generalized partial derivatives are integrable functions. This is equivalent to the following: $\mu = \varrho\, dx$, where $\varrho \in W^{1,1}(\mathbb{R}^n)$.*

Proof. (i) If the measure μ is absolutely continuous, then it is absolutely continuous with respect to the standard Gaussian measure γ, the continuity of which along all vectors is obvious. If we are given the continuity of μ along n linearly independent vectors, then according to what has been said in §3.2 we see that $\mu \ll |\mu| * \gamma$, which yields the absolute continuity of the measure μ since the measure $|\mu| * \gamma$ is absolutely continuous. Justification of (ii) and (iii) is much the same as in the case $n = 1$ considered above. Yet, in the multidimensional case one cannot assert the existence of a bounded density. □

The following result is a direct corollary of Theorem 2.2.1 and Proposition 2.3.1.

3.4.4. Corollary. *Let a measure μ on $\mathbb{R}^n$ be Fomin or Skorohod differentiable along n linearly independent vectors. Then it has a density from the class $L^p(\mathbb{R}^n)$, where $p = n/(n-1)$.*

3.4.5. Remark. (i) In the language of Sobolev spaces Skorohod differentiability of a measure μ on $\mathbb{R}^n$ along n linearly independent vectors is equivalent to the representation $\mu = \varrho\, dx$, where $\varrho \in BV(\mathbb{R}^n)$ (the definition of the class BV is given in §2.3).

(ii) There exists an example of a function $\varrho \in W^{1,1}(\mathbb{R}^2)$ without locally bounded modifications (see Example 2.1.9). Thus, differentiable measures on $\mathbb{R}^n$ with $n \geqslant 2$ may have densities that are essentially locally unbounded.

The next result follows from the properties of functions of the class $W^{1,1}(\mathbb{R}^n)$ mentioned in Theorem 2.1.10 and Remark 2.1.11.

3.4.6. Proposition. *Let a measure μ on $\mathbb{R}^n$ be Fomin differentiable along all vectors. Then it has a density ϱ which is absolutely continuous on almost every straight line parallel to the ith coordinate line for each fixed index $i = 1, \ldots, n$. In addition, there is a version of ϱ which is absolutely continuous on the straight lines generated by almost all points in the sphere equipped with Lebesgue surface measure. Finally, for every two-dimensional plane L and every orthogonal operator S which is a rotation in L and the identity operator on the orthogonal complement of L one can find a version of ϱ which is absolutely continuous on almost all orbits of S.*

The theory of Sobolev spaces provides some additional information about the properties of differentiable measures on $\mathbb{R}^n$.

3.4.7. Theorem. *Let $\{e_i\}$ be the standard basis in $\mathbb{R}^n$, let $k \in \mathbb{N}$, and let μ be a measure on $\mathbb{R}^n$ such that, for every k vectors $h_1, \ldots, h_k$, there exists a derivative $d_{h_1} \cdots d_{h_k}\mu$ in the sense of Fomin. Then*

(i) *if $k = n$, then μ has a bounded continuous density ϱ with respect to Lebesgue measure and $\|\varrho\|_{L^\infty} \leqslant \|d_{e_1} \cdots d_{e_n}\mu\|$;*

(ii) *if $k > n$, then μ has a $(k - n - 1)$-times continuously differentiable density ϱ with respect to Lebesgue measure such that the partial derivatives of ϱ up to order $k - n$ exist a.e. and are integrable and the derivatives up to order $k - n - 1$ are bounded;*

(iii) *the measure μ has derivatives of all orders along all vectors if and only if it admits an infinitely differentiable density ϱ such that all partial derivatives of ϱ belong to $L^1(\mathbb{R}^n)$.*

Let us observe that a density ϱ of a measure μ on $\mathbb{R}^n$ which has partial derivatives up to order n (in the Fomin or Skorohod sense) can be written as

$$\varrho(x) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n}\mu(y\colon \quad y_i < x_i,\ i = 1, \ldots, n).$$

In the case of Skorohod differentiability ϱ may fail to have a continuous density, but the estimate $\|\varrho\|_{L^\infty} \leqslant \|d_{e_1} \cdots d_{e_n}\mu\|$ still holds.

3.4.8. Theorem. *Let $1 < p < \infty$. A Borel probability measure μ on $\mathbb{R}^n$ is L^p-differentiable along all vectors precisely when it is Fomin differentiable with $\beta_h^\mu \in L^p(\mu)$ for all $h \in \mathbb{R}^n$. This is also equivalent to that μ is given by a density ϱ with $\varrho^{1/p} \in W^{p,1}(\mathbb{R}^n)$.*

Proof. Suppose that the measure μ is Fomin differentiable along all vectors and has logarithmic derivatives from $L^p(\mu)$. Hence it has a density $\varrho \in W^{1,1}(\mathbb{R}^n)$ and $\beta_h^\mu = (\nabla\varrho/\varrho, h)$, whence $|\nabla\varrho|/\varrho \in L^p(\mu)$. Set $\psi := \varrho^{1/p}$. Then $\psi \in L^p(\mathbb{R}^n)$. On every ball U the functions $\psi_\varepsilon := (\varrho + \varepsilon)^{1/p}$, where $\varepsilon > 0$, belong to $W^{p,1}(U)$ since

$$|\nabla\psi_\varepsilon|^p = p^{-p}|\nabla\varrho|^p(\varrho + \varepsilon)^{1-p} \leqslant p^{-p}|\nabla\varrho|^p\varrho^{1-p} \in L^p(U).$$

Letting $\varepsilon \to 0$ we get $\psi \in W^{p,1}(U)$ and $\nabla\psi = p^{-1}\varrho^{1/p-1}\nabla\varrho$. By the integrability of ψ^p and $|\nabla\psi|^p$ we obtain the inclusion $\psi \in W^{p,1}(\mathbb{R}^n)$. Conversely, if we know that $\psi = \varrho^{1/p} \in W^{p,1}(\mathbb{R}^n)$, then we obtain the equality $\nabla\varrho = p\psi^{p-1}\nabla\psi$, whence the equality $|\nabla\varrho/\varrho|^p = p^p|\nabla\psi|^p\varrho^{-1} \in L^p(\mu)$ follows. Let us show that the inclusion $\psi := \varrho^{1/p} \in W^{p,1}(\mathbb{R}^n)$ gives L^p-differentiability of μ along every $h \in \mathbb{R}^n$. It suffices to have the differentiability of the mapping $t \mapsto \psi(\,\cdot + th)$ with values in $L^p(\mathbb{R}^n)$. The latter holds by Theorem 2.1.13. Finally, if μ is L^p-differentiable, then $\beta_h^\mu \in L^p(\mu)$ by Theorem 3.3.9. □

3.4.9. Corollary. *Suppose that a measure μ on $\mathbb{R}^n$ is Fomin differentiable along linearly independent vectors $e_1, \ldots, e_n$ and $\beta_{e_i}^\mu \in L^p(|\mu|)$, where $p \in (1, +\infty)$. Then $|\mu|$ is L^p-differentiable along all vectors and $\mu = \varrho\, dx$, where $\varrho \in W^{1,1}(\mathbb{R}^n)$ is locally Hölder continuous and $|\varrho|^{1/p} \in W^{p,1}(\mathbb{R}^n)$.*

Proof. We already know that $\mu = \varrho\, dx$, where $\varrho \in W^{1,1}(\mathbb{R}^n)$. Hence we have $|\mu| = |\varrho|\, dx$, $|\varrho| \in W^{1,1}(\mathbb{R}^n)$, $|\beta_{e_i}^{|\mu|}(x)| = |\beta_{e_i}^\mu(x)|$ for $|\mu|$-a.e. x, whence we obtain $\beta_{e_i}^{|\mu|} \in L^p(|\mu|)$. Now the previous theorem applies and gives $|\varrho|^{1/p} \in W^{p,1}(\mathbb{R}^n)$. Hence the function $|\varrho|^{1/p}$ has a Hölder continuous version. Therefore, $|\varrho|$ is locally Hölder continuous. Since $\varrho \in W^{1,1}(\mathbb{R}^n)$ and $|\nabla\varrho/\varrho|^p|\varrho| \in L^p(\mathbb{R}^n)$, this implies that $|\nabla\varrho| \in L^p_{\rm loc}(\mathbb{R}^n)$, whence we obtain $\varrho \in W^{p,1}_{\rm loc}(\mathbb{R}^n)$. Consequently, ϱ has a locally Hölder continuous version. □

3.4.10. Corollary. *Let μ be a measure on $\mathbb{R}^n$ such that $\beta_{e_i}^\mu \in L^p(|\mu|)$ for some $p > n$. Then μ admits a bounded locally Hölder continuous density ϱ satisfying the estimate*

$$\sup_x |\varrho(x)| \leqslant M(n,p)\|\mu\|^{1-n/p} \sum_{i=1}^n \|\beta_{e_i}^\mu\|^n_{L^p(|\mu|)}, \tag{3.4.1}$$

where $M(n,p)$ is a constant that depends only on n and p.

Proof. The previous corollary yields the existence of a density ϱ in the class $W^{1,1}(\mathbb{R}^n)$ such that $f := |\varrho|^{1/p} \in W^{p,1}(\mathbb{R}^n)$, whence by the Sobolev embedding theorem we obtain a bounded continuous modification of ϱ. Now Theorem 2.2.9 gives a constant $C(n,p)$ such that

$$\sup_x |\varrho(x)| \leqslant C(n,p)\|\nabla f\|^n_{L^p(\mathbb{R}^n)}\|f\|^{p-n}_{L^p(\mathbb{R}^n)}.$$

Here we have the equalities $\|f\|_{L^p(\mathbb{R}^n)} = \|\mu\|^{1/p}$, $|\nabla f| = p^{-1}|\varrho|^{1/p-1}|\nabla\varrho|$, so $\|\nabla f\|_{L^p(\mathbb{R}^n)} = p^{-1}\|\nabla\varrho/\varrho\|_{L^p(|\mu|)}$, which gives (3.4.1). □

3.4.11. Corollary. *Let μ be a measure on $\mathbb{R}^n$ with a bounded density ϱ and let a measure ν on $\mathbb{R}^n$ be such that $\nu = \psi \cdot \mu$ and $d_{e_i}\nu = g_i \cdot \mu$, where $\psi, g_i \in L^p(|\mu|)$ and $p > n$. Then ν admits a bounded locally Hölder continuous density ϱ_ν satisfying the estimate*

$$\sup_x |\varrho_\nu(x)| \leqslant C(n,p)\|\varrho\|_\infty^{1-1/p}\|\psi\|^{1-n/p}_{L^p(|\mu|)} \sum_{i=1}^n \|g_i\|^{n/p}_{L^p(|\mu|)}, \tag{3.4.2}$$

where $C(n,p)$ is a constant that depends only on n and p. In addition, if ϱ is Hölder continuous in a neighborhood of some point x_0 and $\varrho(x_0)=0$, then $\varrho_\nu(x_0)=0$.

PROOF. Our hypotheses yield that ν is Fomin differentiable along all e_i, hence ν has a density $\varrho_\nu \in W^{1,1}(\mathbb{R}^n)$. Therefore, $\varrho_\nu=\psi\varrho$, $\partial_{x_i}\varrho_\nu = g_i\varrho$ and

$$\|\varrho_\nu\|_{L^p(\mathbb{R}^n)} \leqslant \|\varrho\|_\infty^{1-1/p}\|\psi\|_{L^p(|\mu|)}, \quad \|\partial_{x_i}\varrho_\nu\|_{L^p(\mathbb{R}^n)} \leqslant \|\varrho\|_\infty^{1-1/p}\|g_i\|_{L^p(|\mu|)}.$$

It remains to use again Theorem 2.2.9. Suppose now that ϱ is Hölder continuous of order $\kappa>0$ in a ball centered at x_0 and $\varrho(x_0)=0$. We may assume that $x_0=0$ and that $p<n+\kappa$: one can always make p smaller in order that it be in $(n,n+\kappa)$. Let us fix a smooth function η with support in the ball of radius 1 centered at the origin such that $0\leqslant\eta\leqslant 1$, $\eta(0)=1$, and $|\nabla\eta/\eta|^r\eta\in L^1(\mathbb{R}^n)$. Given $\varepsilon>0$, let $\eta_\varepsilon(x)=\eta(\varepsilon^{-1}x)$. Let us consider the measure ν_ε with density $\eta_\varepsilon\varrho_\nu$. It has support in the ball U_ε of radius ε and its density with respect to $\mu_\varepsilon := \eta_\varepsilon\cdot\mu$ equals ψ. We have

$$d_{e_i}\nu_\varepsilon = g_i\cdot\mu_\varepsilon + \frac{\partial_{x_i}\eta_\varepsilon}{\eta_\varepsilon}\,\mu_\varepsilon.$$

Therefore, on account of (3.4.2) we obtain

$$\sup_{x\in U_\varepsilon}|\eta_\varepsilon(x)\varrho_\nu(x)| \leqslant C\sup_{x\in U_\varepsilon}|\eta_\varepsilon(x)\varrho(x)|^{1-1/p}\sum_{i=1}^n \|g_i+\partial_{x_i}\eta_\varepsilon/\eta_\varepsilon\|_{L^p(|\mu_\varepsilon|)}^{n/p},$$

where C is independent of ε. Since

$$\int_{U_\varepsilon}\Bigl|\frac{\partial_{x_i}\eta_\varepsilon}{\eta_\varepsilon}\Bigr|^p\eta_\varepsilon|\varrho|\,dx \leqslant \varepsilon^{n-p}\sup_{x\in U_\varepsilon}|\varrho(x)|\int_{U_1}\Bigl|\frac{\partial_{x_i}\eta}{\eta}\Bigr|^p\eta\,dx,$$

we obtain that $\|g_i+\partial_{x_i}\eta_\varepsilon/\eta_\varepsilon\|_{L^p(|\mu_\varepsilon|)}$ is estimated by

$$\|g_i\|_{L^p(|\mu_\varepsilon|)} + \|\partial_{x_i}\eta_\varepsilon/\eta_\varepsilon\|_{L^p(|\mu_\varepsilon|)} \leqslant C_1 + C_1\varepsilon^{n-p}\sup_{x\in U_\varepsilon}|\varrho(x)|,$$

where C_1 is independent of ε. Since $\varrho(0)=0$ and ϱ is κ-Hölder continuous in U_ε for sufficiently small $\varepsilon>0$, we have $\sup_{x\in U_\varepsilon}|\varrho(x)|\leqslant C_2\varepsilon^\kappa$. Therefore, the previous estimates show that

$$|\varrho_\nu(0)|\leqslant C_3\varepsilon^\alpha, \quad \alpha := \kappa(1-1/p)+(n+\kappa-p)n/p>0,$$

where C_3 is independent of ε. Letting $\varepsilon\to 0$, we obtain $\varrho_\nu(0)=0$. □

3.4.12. Corollary. *Let μ be a measure on $\mathbb{R}^n$ with a bounded density ϱ and let a measure ν on $\mathbb{R}^n$ be such that one has $\partial^{(\alpha)}\nu = g_\alpha\cdot\mu$ for every multi-index α with $0<|\alpha|\leqslant m$, where $g_\alpha\in L^p(|\mu|)$ and $p>n$. Then ν has a density $\varrho_\nu\in C_b^{m-1}(\mathbb{R}^n)$ for any α with $0<|\alpha|<m$ satisfying the estimates*

$$\sup_x|\partial^{(\alpha)}\varrho_\nu(x)| \leqslant C(n,p,m)\|\varrho\|_\infty^{1-1/p}\|g_\alpha\|_{L^p(|\mu|)}^{1-n/p}\sum_{j=1}^n\|g_{\alpha+\delta_j}\|_{L^p(|\mu|)}^{n/p},$$

where $C(n,p,m)$ depends only on n, p, and m, and the multi-index $\alpha+\delta_j$ is defined by $\partial^{(\alpha+\delta_j)} := \partial_{x_j}\partial^{(\alpha)}$. In addition, the functions $\partial^{(\alpha)}\varrho_\nu$ with $|\alpha|<m$ are locally Hölder continuous.

PROOF. We apply the previous corollary to $\partial^{(\alpha)}\nu$ with $|\alpha|<m$ and observe that $d_{e_j}\partial^{(\alpha)}\nu = g_{\alpha+\alpha_j}\cdot\mu$. □

3.5. Characterization by conditional measures

The examples in the previous section hint to characterize directional differential properties of measures by means of conditional measures. Let μ be a Radon measure on a locally convex space X and $h\in X$, $h\neq 0$. Then there exists a closed hyperplane Y in X for which $X=\mathbb{R}^1 h\oplus Y$. In addition, there exists a continuous linear mapping $\pi\colon X\to Y$, $th+y\mapsto y$, called the natural projection. Let ν denote the image of the measure $|\mu|$ under this projection.

Let us recall that according to §1.3 one can choose Borel measures μ^y on the straight lines $y+\mathbb{R}^1 h$, $y\in Y$, such that for every Borel set $B\subset X$ the following equality holds:

$$\mu(B)=\int_Y \mu^y(B)\,\nu(dy). \tag{3.5.1}$$

These measures are determined uniquely up to a redefinition for points y from a set of ν-measure zero.

3.5.1. Theorem. *Let μ be a Radon measure on a locally convex space X and let $h\in X$. Then*

(i) *the measure μ is continuous along h if and only if the measures μ^y on the straight lines $y+\mathbb{R}^1 h$ are continuous along h for ν-a.e. y;*

(ii) *the measure μ is Fomin differentiable along h if and only if the measures μ^y on the straight lines $y+\mathbb{R}^1 h$ are Fomin differentiable along h for ν-a.e. y and one has*

$$\int_Y \|d_h\mu^y\|\,\nu(dy)<\infty;$$

(iii) *the measure μ is Skorohod differentiable along h if and only if for ν-a.e. y the measures μ^y on the straight lines $y+\mathbb{R}^1 h$ are Skorohod differentiable along h and one has*

$$\int_Y \|d_h\mu^y\|\,\nu(dy)<\infty;$$

(iv) *the measure μ is L^p-differentiable along h if and only if the measures μ^y on the straight lines $y+\mathbb{R}^1 h$ are L^p-differentiable along h for ν-a.e. y and one has*

$$\int_Y\int_{y+\mathbb{R}^1 h} |\beta_h^{\mu^y}(x)|^p\,|\mu^y|\,(dx)\nu(dy)<\infty;$$

(v) *the measure μ is quasi-invariant along h if and only if the measures μ^y on the straight lines $y+\mathbb{R}^1 h$ are quasi-invariant along h for ν-a.e. y.*

In the case of Fomin or Skorohod differentiability we have

$$d_h\mu(B) = \int_Y d_h\mu^y(B)\,\nu(dy) \tag{3.5.2}$$

for all Borel sets B.

PROOF. The sufficiency of the listed requirements on conditional measures for the respective properties of μ is easily verified using formulae (1.3.1) and (1.3.3). For example, if almost all conditional measures μ^y are continuous along h, then, for every $A \in \mathcal{B}(X)$, we have $\mu^y(A + t_n h) \to \mu^y(A)$ as $t_n \to 0$, which by the Lebesgue dominated convergence theorem yields that $\mu(A + t_n h) \to \mu(A)$.

The necessity of these conditions is not obvious. Let the measure μ be Fomin differentiable along h. Since $d_h\mu \ll \mu$, according to (1.3.1) and (1.3.4) there exist measures μ^y and $\xi^y \ll \mu^y$ on $y + \mathbb{R}^1 h$ such that

$$\mu(A) = \int_Y \mu^y(A)\,\nu(dy),$$

$$d_h\mu(A) = \int_Y \xi^y(A)\,\nu(dy)$$

for all Borel sets A. Hence for all $s \in \mathbb{R}^1$ we have

$$d_h\mu(A + sh) = \int_Y \xi^y(A + sh)\,\nu(dy),$$

$$\mu(A + th) - \mu(A) = \int_Y [\mu^y_{th}(A) - \mu^y(A)]\,\nu(dy),$$

whence by (3.3.1) we find

$$\int_Y [\mu^y_{th}(A) - \mu^y(A)]\,\nu(dy) = \int_Y \int_0^t \xi^y(A + sh)\,ds\,\nu(dy).$$

Since A was arbitrary and conditional measures are determined uniquely ν-a.e., we obtain the equality

$$\mu^y_{th} - \mu^y = \int_0^t \xi^y_{sh}\,ds$$

for ν-a.e. y. For such y, taking into account the relationship $\xi^y \ll \mu^y$, we obtain Fomin differentiability of μ^y along h and the equality $d_h\mu^y = \xi^y$. Therefore, assertion (ii) is proven, and (iv) is its direct corollary. Similarly we verify (v), which gives (i) due to the fact that any measure continuous along h is absolutely continuous with respect to some measure quasi-invariant along h. Finally, (iii) is proven analogously to (ii) with the only difference being that now the role of ν is played by the projection of the measure $|\mu| + |d_h\mu|$ to the hyperplane Y. □

A straightforward modification of our reasoning leads to the following more general result.

3.5.2. Theorem. *Let μ be a Radon measure on a locally convex space X that is a direct topological sum of a locally convex space Y and a Souslin locally convex space F and let S be a finite or countable subset of F. Suppose that for some n there exist derivatives $d_{h_1}\cdots d_{h_k}\mu$ (in the sense of Fomin or Skorohod) for every $h_1,\ldots,h_k \in S$, $k \leqslant n$. Then one can find conditional measures μ^y, $y \in Y$, on $y+F$ such that, for every $h_1,\ldots,h_k \in S$, $k \leqslant n$, the derivatives $d_{h_1}\cdots d_{h_k}\mu^y$ exist in the corresponding sense and we have*

$$d_{h_1}\cdots d_{h_k}\mu(B) = \int_Y d_{h_1}\cdots d_{h_k}\mu^y(B)\,\nu(dy) \quad \text{for all } B \in \mathcal{B}(X),$$

where ν is the image of $|\mu|$ under the natural projection to Y. Analogous assertions are true for the infinite differentiability, continuity, and quasi-invariance.

Proof. In order to extend (ii) of Theorem 3.5.1 to this case, we denote by ν the projection of $|\mu|$ on F and choose conditional measures μ^y and $\xi^{h,y}$ for μ and $d_h\mu$, where $h \in S$, on the sets $y+F$, $y \in Y$. Since S is at most countable, we arrive as above to the equality

$$\mu^y_{th} - \mu^y = \int_0^t \xi^{h,y}_{sh}\,ds$$

for ν-a.e. $y \in Y$ simultaneously for all $h \in S$ and all t. As above, this shows that μ^y is Fomin differentiable along any $h \in S$ and $d_h\mu^y = \xi_h$. Other assertions from the previous theorem are extended similarly. The case of higher differentiability follows by induction. For example, if, for some $h_1, h_2 \in S$, the measure $d_{h_2}\mu$ is differentiable along h_1, the same reasoning as above shows that for ν-a.e. y one has

$$d_{h_2}\mu^y_{th_1} - d_{h_2}\mu^y = \int_0^t \eta^{h_1,h_2,y}_{sh_1}\,ds$$

for all t, where $\{\eta^{h_1,h_2,y}\}_{y\in Y}$ is the family of conditional measures on the sets $y+F$ corresponding to the measure $d_{h_1}d_{h_2}\mu$. □

3.5.3. Corollary. *In the situation of the previous theorem, the measure μ is analytic along all vectors $h \in S$ precisely when one can find conditional measures μ^y on $y+F$ such that, for each $h \in S$, the measures μ^y are analytic along h and for some $c(h) > 0$ one has*

$$\sum_{n=0}^{\infty} \int_Y \frac{t^n}{n!}\|d_h^n\mu^y\|\,\nu(dy) < \infty \quad \forall t \in [0, c(h)].$$

Proof. If this condition is fulfilled, then one readily verifies that the series $\sum_{n=0}^{\infty} z^n d_h^n\mu/n!$ converges in variation if $|z| < c(h)$. Conversely, convergence of this series yields convergence of the series $\sum_{n=0}^{\infty} t^n\|d_h^n\mu\|/n!$ whenever $t \in [0, c(h)]$. Since we have

$$\|d_h^n\mu\| = \int_Y \|d_h^n\mu^y\|\,\nu(dy),$$

the conclusion follows by the monotone convergence theorem. □

Let us recall that conditional measures (in the form of conditional expectations) arise also in expressions for logarithmic derivatives of linear images of differentiable measures (Proposition 3.3.13).

3.6. Skorohod differentiability

Here we consider measures differentiable in the Skorohod sense. By using our characterization in terms of conditional measures we shall derive a result from [**174**] which guarantees the existence of a Radon (and not just Baire, as noted above) weak derivative for every Skorohod differentiable Radon measure on an arbitrary locally convex space. The proof below differs from that of [**174**].

3.6.1. Theorem. *Let a Radon measure μ on a locally convex space X be Skorohod differentiable along a vector $h \neq 0$. Then, there exists a Radon measure ν which is its Skorohod derivative, i.e., satisfies* (3.1.4) *for all bounded continuous functions f on X.*

PROOF. Let us decompose X in the sum $X = Y \oplus \mathbb{R}^1 h$, where Y is a closed hyperplane, denote by μ_0 the projection of $|\mu|$ to Y. On the straight lines $y + \mathbb{R}^1 h$ we can take conditional measures μ^y that are Skorohod differentiable along h. Now it suffices to show that the measure $d_h\mu$ is given by the equality

$$d_h\mu(A) = \int_Y d_h\mu^y(A)\,\mu_0(dy),$$

where the right side generates a Radon measure. We first observe that the right side defines a Baire measure that agrees with $d_h\mu$ on all Baire sets A. To see this, we recall that for every $f \in C_b(X)$, the integral of f against $d_h\mu^y$ is the limit of the integral of $n\big[(f(x + n^{-1}h) - f(x)\big]$ against μ^y as $n \to \infty$. In addition, the latter integral is majorized in the absolute value by $\sup_x |f(x)|\,\|d_h\mu^y\|$, which is a μ_0-integrable function. On the other hand, the integral of $n\big[(f(x + n^{-1}h) - f(x)\big]$ against μ tends to the integral of f against $d_h\mu$. Since μ_0 is a Radon measure and the function $\|d_h\mu^y\|$ is μ_0-integrable, we conclude that the measure $d_h\mu$ is tight. Indeed, we can take increasing compact sets $K_n \subset Y$ such that $\mu_0(Y\backslash K_n) \to 0$. Then the sets $S_n = K_n + \{th\colon\, t \in [-n, n]\}$ are compact and it is easily seen from the established identity on $\mathcal{B}a(X)$ that $|d_h\mu|^*(X\backslash S_n) \to 0$. Now we take for ν the Radon extension of $d_h\mu$ (see [**193**, Corollary 7.3.3] on its existence). Finally, we obtain that the above disintegration formula holds for all Borel sets A. Indeed, the Radon measure ν admits some disintegration with respect to the projection of $|\nu|$ to Y. The latter is absolutely continuous with respect to μ_0 (if $A_0 \subset Y$ is a Baire set of μ_0-measure zero, then we have $d_h\mu(A_0 \oplus \mathbb{R}^1 h) = 0$). Hence there exist measures σ^y on the straight lines $y + \mathbb{R}^1 h$ such that

$$\nu(A) = \int_Y \sigma^y(A)\,\mu_0(dy)$$

for all Borel sets A. Then we obtain $\sigma^y = d_h\mu^y$ for μ_0-a.e. y, which yields the desired equality and completes the proof. □

It is readily verified that if a measure μ on the real line is Skorohod differentiable along 1, then the measure $|\mu|$ is also. Indeed, the measure μ is given by a density p of bounded variation, and the function $|p|$ is also of bounded variation. In addition, we have $\operatorname{Var}|p| \leqslant \operatorname{Var} p$ (Exercise 3.9.2). The functions p^+ and p^- have the same property. In terms of Skorohod differentiability this means Skorohod differentiability of the measures $|\mu|$, μ^+, μ^- and the validity of the inequalities $\|d_1|\mu|\| \leqslant \|d_1\mu\|$, $\|d_1\mu^+\| \leqslant \|d_1\mu\|$, $\|d_1\mu^-\| \leqslant \|d_1\mu\|$. Note that $\|d_1|\mu|\|$ can be strictly smaller than $\|d_1\mu\|$: take μ on $\mathbb{R}^1$ with the density $I_{[0,1]}-I_{[1,2]}$; then $d_1\mu = \delta_0 - 2\delta_1 + \delta_2$, $d_1|\mu| = \delta_0 - \delta_2$. It is clear from this example that in general there is no estimate of the type $\|d_1\mu\| \leqslant C\|d_1|\mu|\|$. Note also that $d_1|\mu| \ll d_1\mu$ (see the next lemma), but these two measures need not be equivalent unlike the case of the Fomin derivative.

3.6.2. Lemma. *For any Borel set $B \subset \mathbb{R}^1$, the following inequality holds*: $|d_1|\mu||(B) \leqslant |d_1\mu|(B)$.

PROOF. Clearly, it suffices to show that the desired estimate holds for every finite collection of disjoint intervals. In turn, it is enough to show that

$$|d_1|\mu||\big([a,b)\big) \leqslant |d_1\mu|\big([a,b)\big)$$

for all a, b. It remains to observe that the left side equals $\big||\varrho(b)| - |\varrho(a)|\big|$ and the right side equals $|\varrho(b) - \varrho(a)|$, where ϱ is a left-continuous density of μ of bounded variation. □

By using conditional measures we easily extend this to the general case and obtain the following result.

3.6.3. Theorem. *Suppose that a Radon measure μ on a locally convex space X is Skorohod differentiable along a vector h. Then, its positive part μ^+, its negative part μ^-, and its variation $|\mu|$ are Skorohod differentiable along h as well and*

$$\|d_h|\mu|\| \leqslant \|d_h\mu\|, \quad \|d_h\mu^+\| \leqslant \|d_h\mu\|, \quad \|d_h\mu^-\| \leqslant \|d_h\mu\|.$$

PROOF. Let Y and μ_0 be the same as in the proof of the previous theorem. We take Skorohod differentiable conditional measures μ^y. The measures $d_h[(\mu^y)^+]$ and $d_h[(\mu^y)^-]$ exist and their variations do not exceed $\|d_h\mu^y\|$. Note that the measures μ^+ and μ^- are mutually singular and possess conditional measures $(\mu^+)^y$ and $(\mu^-)^y$, which are mutually singular for μ_0-a.e. y. Then it follows that $(\mu^y)^+ = (\mu^+)^y$ and $(\mu^y)^- = (\mu^-)^y$ for μ_0-a.e. y. So the only technicality is to show that $d_h[(\mu^y)^+]$ and $d_h[(\mu^y)^-]$ depend measurably on y. □

The next result was also obtained in [**174**]; we give another proof.

3.6.4. Theorem. *A Radon measure μ on a locally convex space X is Skorohod differentiable along a vector h if and only if for every $A \in \mathcal{B}(X)$ there exists a number $c(A)$ such that*

$$|\mu(A+th) - \mu(A)| \leqslant c(A)|t|. \tag{3.6.1}$$

In this case, the numbers $c(A)$ can be chosen uniformly bounded.

PROOF. It is clear that Skorohod differentiability gives estimate (3.6.1) with a common constant $C = \|d_h\mu\|$. Conversely, if we have (3.6.1), then by the Nikodym theorem one can take a common constant, whence we obtain $\|\mu_{th} - \mu\| \leqslant C|t|$ with some number C. Let us show that this ensures Skorohod differentiability.

First we consider the case $X = \mathbb{R}^1$. We observe that it suffices to have an estimate of the type

$$\|\mu_{t_i h} - \mu\| \leqslant C|t_i|, \quad \text{where } t_i \to 0.$$

Indeed, this estimate yields that for every $\varphi \in C_0^\infty(\mathbb{R}^1)$ one has

$$\Bigl|\int_{\mathbb{R}} t_i^{-1}[\varphi(x-t_i) - \varphi(x)]\,\mu(dx)\Bigr| = \Bigl|t_i^{-1}\int_{\mathbb{R}} \varphi(x)\,[\mu_{t_i} - \mu](dx)\Bigr| \leqslant C\sup_x|\varphi(x)|,$$

hence

$$\Bigl|\int_{\mathbb{R}} \varphi'(x)\,\mu(dx)\Bigr| \leqslant C\sup_x|\varphi(x)|.$$

This shows that the generalized derivative of μ is a bounded measure. To be more precise, the previous estimate ensures the existence of measures on all intervals $(-n,n)$ that are generalized derivatives of μ on $(-n,n)$, but the same estimate shows that the total variations of these derivatives on $(-n,n)$ are uniformly bounded. Hence μ is Skorohod differentiable. In addition, $\|d_1\mu\| \leqslant C$.

Let us consider the general case. Let $h \neq 0$. Let us write X in the form $X = Y \oplus \mathbb{R}^1 h$, where Y is a closed hyperplane, denote by μ_0 the projection of $|\mu|$ to Y and take conditional measures μ^y on the straight lines $y + \mathbb{R}^1 h$, $y \in Y$. We may assume that μ_0 is a probability measure. Points $x \in X$ will be written in the form $x = (y,s)$, $y \in Y$, $s \in \mathbb{R}^1$. Then the conditional measures μ^y can be regarded as measures on $\mathbb{R}^1$. Let us set $t_n = n^{-1}$. Then

$$\int_Y t_n^{-1}\|\mu^y_{t_n h} - \mu^y\|\,\mu_0(dy) \leqslant C.$$

By Fatou's theorem

$$\int_Y \liminf_n t_n^{-1}\|\mu^y_{t_n h} - \mu^y\|\,\mu_0(dy) \leqslant C.$$

Hence for μ_0-a.e. y we have

$$C(y) := \liminf_n t_n^{-1}\|\mu^y_{t_n h} - \mu^y\| < \infty.$$

By the one-dimensional case for every such y the measure μ^y is Skorohod differentiable and $\|d_h\mu^y\| \leqslant C(y)$. Now we apply Theorem 3.5.1. □

Thus, most of the properties introduced in §3.1 can be described in terms of the functions $t \mapsto \mu(A+th)$.

A direct corollary of Theorem 3.6.4 is the fact that if X is linearly and continuously embedded into a locally convex space Y, then the collection of vectors of Skorohod differentiability for μ remains the same independently of whether μ is considered as a measure on X or on Y, although the classes of continuous functions in these two cases may differ. It is clear that μ extends to $\mathcal{B}(Y)$ since $X \cap B$ is a Borel set in X for every $B \in \mathcal{B}(Y)$. For other differential properties considered above this is obvious.

The following result gives a characterization of Skorohod differentiability in the integral form.

3.6.5. Theorem. *Let μ be a Radon measure on a locally convex space X and let $\mathcal{A} = \mathcal{B}(X)$. A measure μ is Skorohod differentiable along a vector h if and only if there exists a Radon measure ν such that for every set $A \in \mathcal{B}(X)$ one has*

$$\mu(A+th) - \mu(A) = \int_0^t \nu(A+sh)\, ds. \tag{3.6.2}$$

This is also equivalent to the equality

$$\int_X [f(x-th) - f(x)]\, \mu(dx) = \int_0^t \int_X f(x-sh)\, \nu(dx)\, ds \tag{3.6.3}$$

for all bounded Borel functions f or for all functions $f \in C_b(X)$. In this case $\nu = d_h\mu$. An analogous assertion with Baire functions in place of Borel ones is true for the class of Baire measures. For Radon measures or measures on $\sigma(X)$ it suffices to have (3.6.3) *for $f \in \mathcal{FC}_b^\infty$.*

Proof. The equivalence of (3.6.2) and (3.6.3) is obvious since the former is just the latter for $f = I_A$. Moreover, (3.6.3) for all $f \in C_b(X)$ yields (3.6.3) for all bounded Borel functions f in the case of a Radon measure and all bounded Baire functions f in the case of a Baire measure (or for $f \in \mathcal{FC}_b^\infty$ in the cases of measures on $\sigma(X)$ and Radon measures). Finally, let us show that equality (3.6.3) for all $f \in C_b(X)$ is equivalent to the fact that ν is the Skorohod derivative of the measure μ. Indeed, dividing both sides of (3.6.3) by t and letting $t \to 0$, in the limit we obtain on the left the integral of f against ν since the integrals of $t^{-1}f(x-sh)$ in s over $[0,t]$ are uniformly bounded and converge to $f(x)$. On the other hand, if ν is the Skorohod derivative of μ along h, then both sides of (3.6.3) are differentiable in t and their derivatives at any point t_0 coincide with the integral of $-f(x-t_0h)$ against ν. Vanishing at $t=0$, they are equal everywhere. □

By the proven theorem the Skorohod derivative of any Fomin differentiable measure coincides with its Fomin derivative. This justifies our usage of the symbol $d_h\mu$ also for weak derivatives.

Note that, by Theorem 3.6.4, under the hypotheses of Theorem 3.6.5 and its corollary in the sufficiency part, in place of the assumption that ν is a Radon measure it suffices to require that it be defined only on $\sigma(X)$

and satisfy the corresponding equalities only for $\sigma(X)$-measurable sets and functions. In that case, such a measure can be chosen Radon, which follows from Theorem 3.6.4 and the necessary condition from Theorem 3.6.5.

3.6.6. Corollary. *If a Radon measure μ on a locally convex space X is Skorohod differentiable along a vector h, then it is Fomin differentiable along this vector if and only if at least one of the following conditions is fulfilled* (*then all of them are fulfilled*):

(i) *its weak derivative ν is absolutely continuous with respect to μ*;

(ii) *its weak derivative ν is continuous along h.*

The same is true for measures on $\sigma(X)$.

3.6.7. Corollary. *Suppose that a Radon measure μ on a locally convex space X is Skorohod differentiable along a vector h. Then, for every bounded Borel function f possessing a uniformly bounded partial derivative $\partial_h f$ we have*

$$\int_X \partial_h f(x)\,\mu(dx) = -\int_X f(x)\,d_h\mu(dx). \tag{3.6.4}$$

Applying (3.6.4) to the functions $f(x) = \exp\bigl(il(x)\bigr)$ with $l \in X^*$ and recalling that the Fourier transform $\widetilde{\mu}$ of the measure μ is defined by the formula

$$\widetilde{\mu}\colon\ X^* \to \mathbb{C}, \quad \widetilde{\mu}(l) = \int_X \exp\bigl(il(x)\bigr)\,\mu(dx),$$

we obtain the identity

$$\widetilde{d_h\mu}(l) = -il(h)\widetilde{\mu}(l). \tag{3.6.5}$$

As observed in [**166**], this identity is actually a criterion of differentiability in terms of the Fourier transform (which solves a problem posed by Fomin [**443**]). Namely, we have the following necessary and sufficient condition for differentiability.

3.6.8. Theorem. *Let μ be a measure on the σ-algebra $\sigma(X)$ in a locally convex space X and let $h \in X$. Then*

(i) *the measure μ has a Skorohod derivative along the vector h if and only if there exists a measure ν on $\sigma(X)$ such that*

$$\widetilde{\nu}(l) = -il(h)\widetilde{\mu}(l), \quad \forall l \in X^*; \tag{3.6.6}$$

in this case ν coincides with the weak derivative;

(ii) *the measure μ is Fomin differentiable along h if and only if there exists a measure ν on $\sigma(X)$ satisfying equality* (3.6.6) *and absolutely continuous with respect to μ. In this case $\nu = d_h\mu$.*

Proof. From (3.6.6) we obtain (3.6.3) for all functions $f = \exp(il)$, where $l \in X^*$, which gives (3.6.3) for all $f \in \mathcal{FC}_b^\infty$ since two measures on $\sigma(X)$ with equal Fourier transforms coincide. □

3.6.9. Corollary. *Suppose that a Radon measure μ is Skorohod differentiable along vectors h and k. Then it is Skorohod differentiable along* *every linear combination of these vectors and for the weak derivatives we have*

$$d_{sh+tk}\mu = sd_h\mu + td_k\mu.$$

A trivial corollary of Theorem 3.6.8 is the uniform boundedness of the function $\varphi\colon l \mapsto -il(h)\widetilde{\mu}(l)$ on X^*. Let us observe that if X is a Hilbert space, then this function is always continuous in the Sazonov topology (even for nondifferentiable measures μ). We recall that the Sazonov topology is generated by the family of seminorms of the form $x \mapsto \|Tx\|$, where T is a Hilbert–Schmidt operator, and the Fourier transform of any Radon measure on X is continuous in this topology (see [**193**, §7.13]). So it is natural to ask which conditions on the function φ guarantee that it is the Fourier transform of a bounded measure. Unfortunately, such a measure must be signed. For this reason the classical Minlos–Sazonov theorem (which describes the Fourier transforms of nonnegative measures as positive definite functions continuous in the Sazonov topology) is not applicable to this situation. There is an analogue of the Minlos–Sazonov theorem for signed measures (see Tarieladze [**1115**]). However, the corresponding hypotheses include the assumption that φ is the Fourier transform of a cylindrical measure ν of bounded variation. In our special case such analogues are useless since the boundedness of variation of ν is precisely what is needed for Skorohod differentiability. Indeed, if we know that there exists a cylindrical measure ν of bounded variation such that $\varphi = \widetilde{\nu}$, then its finite dimensional projections are the weak derivatives of the corresponding projections of μ. This gives the uniform Lipschitzness of the functions $t \mapsto \mu(C + th)$ for all cylinders, which by Theorem 3.6.4 implies Skorohod differentiability.

One might try to find other conditions of differentiability in terms of the Fourier transform not involving a priori assumptions of bounded variation type. Here, for example, it is not enough that the function φ be uniformly bounded.

3.6.10. Example. Let μ be the countable product of Cauchy probability measures μ_n on the real line given by densities

$$p_n(y) = \frac{2^n}{\pi}\frac{1}{1+2^{2n}t^2}.$$

Since $\mu(l^2) = 1$ (Exercise 3.9.6), this measure can be regarded as a measure on l^2. Let $h = (h_n)$, where $h_n = 2^{-n}$. Then the measure μ is not Skorohod differentiable along h, but the function $\varphi(y) = -i(y,h)\widetilde{\mu}(y)$ is uniformly bounded and continuous in the Sazonov topology on l^2.

Proof. By Corollary 4.1.2 below, the set $D_C(\mu)$ of all vectors of S-differentiability of μ is the space of all sequences (x_n) such that $\sum_{n=1}^{\infty} 2^{2n}x_n^2 < \infty$. Hence $h \notin D_C(\mu)$. The Fourier transform of μ is given by the equality

$$\widetilde{\mu}(y) = \exp\Bigl(\sum_{n=1}^{\infty} 2^{-n}|y_n|\Bigr), \quad y \in l^2.$$

It is easily verified that the function φ defined above is uniformly bounded and continuous in the Sazonov topology. However, this function cannot be the Fourier transform of a cylindrical measure of bounded variation. $\square$

The same example shows that even exponential decay of the function $\widetilde{\mu}$ does not guarantee the differentiability of the measure μ along h. However, as the following nice result due to A.A. Belyaev shows, the measure μ is analytic along h provided that $\widetilde{\mu}$ has support in a suitable strip. This result can be regarded as an infinite dimensional version of the classical Paley–Wiener theorem (see Rudin [**976**, Ch. 7]).

3.6.11. Theorem. *Let μ be a measure on $\sigma(X)$ and $h \in X$. Let $R > 0$ be such that $\widetilde{\mu}$ vanishes outside the strip $\Pi_R = \{l \in X^*\colon\ |l(h)| \leqslant R\}$. Then the measure μ is analytic and quasi-invariant along h. In addition, for every bounded $\sigma(X)$-measurable function f the function*

$$t \mapsto \int_X f(x - th)\,\mu(dx)$$

extends to an entire function ψ_f on the complex plane satisfying the estimate

$$|\psi_f(z)| \leqslant \|\mu\| \sup_X |f(x)| \exp(R \cdot \operatorname{Im} z). \tag{3.6.7}$$

Conversely, if estimate (3.6.7) *is valid for every bounded $\sigma(X)$-measurable function f, then $\widetilde{\mu}$ has support in Π_R.*

PROOF. 1) Let $\varepsilon > 0$ and let $g \in C_0^\infty(\mathbb{R}^1)$ be a function with support in $[-R-\varepsilon, R+\varepsilon]$ and equal to 1 on $[-R, R]$. Then g is the Fourier transform of a function $p_0 \in \mathcal{S}(\mathbb{R}^1)$. Set $p := p_0/\sqrt{2\pi}$. We have

$$p^{(n)}(x) = \frac{i^n}{2\pi} \int_{-R-\varepsilon}^{R+\varepsilon} e^{itx} g(t)\,dt,$$

hence

$$x^2 p^{(n)}(x) = -\frac{i^n}{2\pi} \int_{-R-\varepsilon}^{R+\varepsilon} e^{itx} [t^n g''(t) + 2nt^{n-1} g'(t) + n(n-1) g(t)]\,dt,$$

which yields the estimate $|p^{(n)}(x)| \leqslant C_1 n^2 R^n (1+x^2)^{-1}$ with some C_1 independent of n. Consequently, $\|p^{(n)}\|_{L^1} \leqslant C_2 n^2 R^n$. Therefore, the measure λ with density p is analytic (note that from the Paley–Wiener theorem it only follows that p is analytic). Note also that for all complex numbers z we have

$$|p(x+z)| \leqslant \frac{1}{2\pi} \int_{-R-\varepsilon}^{R+\varepsilon} \bigl|e^{itz}\bigr|\,|g(t)|\,dt \leqslant C_3 \exp\bigl[(R+\varepsilon)|\operatorname{Im} z|\bigr].$$

Let ψ be a bounded Borel function on $\mathbb{R}^1$ with bounded support. Then the function $\psi(-\,\cdot\,) * p$ is analytic and the previous estimate yields

$$\Bigl|\int \psi(x-z) p(x)\,dx\Bigr| = \Bigl|\int \psi(x) p(x+z)\,dx\Bigr| \leqslant C(\psi,\lambda) \exp\bigl[(R+\varepsilon)|\operatorname{Im} z|\bigr].$$

However, we need a more explicit bound on $C(\psi,\lambda)$. Since for real z we have

$$\Bigl|\int \psi(x-z)\,\lambda(dx)\Bigr| \leqslant \|\lambda\|\cdot\|\psi\|_\infty,$$

the Phragmén–Lindelöf theorem gives the estimate

$$\Bigl|\int \psi(x)p(x+z)\,dx\Bigr| \leqslant \|\lambda\|\cdot\|\psi\|_\infty \exp\bigl[(R+\varepsilon)|\mathrm{Im}\,z|\bigr]. \tag{3.6.8}$$

Indeed, denoting the convolution on the left by φ and considering the function $\varphi_1(z) := \varphi(z)\exp[i(R+\varepsilon)z]$ in the upper half-plane, we obtain a bounded holomorphic function such that $|\varphi_1(x)| \leqslant \|\lambda\|\cdot\|\psi\|_\infty$ on the real line. By the Phragmén–Lindelöf principle (see [**778**, Ch. 7, §1]) this estimate holds in the whole upper half-plane, which yields our claim, and the same applies to the lower half-plane.

Now if ψ is an arbitrary bounded Borel function, then it also satisfies the above estimate. Indeed, letting $\psi_n(t) = \psi(t)I_{[-n,n]}(t)$, we have a common estimate (3.6.8) for the corresponding convolutions $\psi_n(-\,\cdot\,)*p$, which for real z converge to $\psi(-\,\cdot\,)*p$. Hence the entire functions $\psi_n(-\,\cdot\,)*p$ converge uniformly on compact sets in $\mathbb{C}$ to an entire function that coincides with $\psi(-\,\cdot\,)*p$ on the real line and satisfies (3.6.8).

Let ν be the image of λ under the mapping $t\mapsto th$. Then $\widetilde{\nu}(l)=1$ if $|l(h)|\leqslant R$ since by the change of variables formula one has

$$\widetilde{\nu}(l) = \widehat{p_0}\bigl(-l(h)\bigr) = g\bigl(-l(h)\bigr).$$

Since $\widetilde{\mu}(l)=0$ if $|l(h)|\geqslant R$, we arrive at the equality $\widetilde{\mu}(l)=\widetilde{\mu}(l)\widetilde{\nu}(l)$, which means that $\mu=\mu*\nu$. In particular, the measure μ is analytic. Moreover, one has

$$\int_X f(x-th)\,\mu(dx) = \int_X f(x-th)\,\mu*\nu(dx)$$

$$= \int_X\int_X f(x+y-th)\,\mu(dy)\,\nu(dx) = \int\int_X f(sh-th+y)\,\mu(dy)\,\lambda(ds),$$

where the function

$$F\colon\ t\mapsto \int_X f(th+y)\,\mu(dy)$$

is Borel measurable and its absolute value is estimated by $\|\mu\|\cdot\sup_x|f(x)|$. By what has already been proven, the function

$$\psi(t) = \int F(s-t)\,\lambda(ds) = \int_X f(x-th)\,\mu(dx)$$

extends to an entire function satisfying (3.6.8). Applying the Phragmén–Lindelöf theorem once more we arrive at the estimate

$$|\psi(z) \leqslant \|\mu\|\cdot\sup_x|f(x)|\exp\bigl[(R+\varepsilon)|\mathrm{Re}\,z|\bigr].$$

Since ε is arbitrary, we obtain the desired estimate.

2) We prove the converse. Let $l\in X^*$, $l(h)=R$, $|\alpha|\geqslant 1$. We shall show that $\widetilde{\mu}(\alpha l)=0$. Considering the measure $\lambda=\mu\circ l^{-1}$ we reduce the assertion

to the one-dimensional case. The measure λ has analytic density p and by the Paley–Wiener theorem (see Rudin [**976**, Ch. 7]) the function

$$\varphi(t) = \int f(x - tR)p(x)\,dx, \quad f(x) = \exp(-x^2),$$

is the Fourier transform of a generalized function $G \in \mathcal{S}'$ with support in $[-R, R]$. Since $\varphi \in L^1(\mathbb{R}^1)$ and $\widehat{f} > 0$ we see that the support of $\widehat{p}$ is contained in $[-R, R]$. $\square$

It is worth noting that the following property of a measure μ along h has recently been studied by A.V. Shaposhnikov [**992**]: the absolute continuity of all functions $t \mapsto \mu(A + th)$. He has constructed an example of a probability measure on the real line showing that this property does not imply Skorohod differentiability. Namely, such a measure is given by the density

$$\varrho(x) := \sum_{m=1}^{\infty} (m^{2/3+1}2^m)^{-1} I_{[2\pi 4^m, 2\pi(4^m+2^m)]}(x) \sin^2 mx.$$

On the other hand, he proved the following positive assertions.

3.6.12. Theorem. *A Radon measure μ on a locally convex space X is Skorohod differentiable along a vector h if and only if the mapping $t \mapsto \mu_{th}$ with values in the space of measures equipped with the variation norm is absolutely continuous on* $[0, 1]$.

3.6.13. Theorem. *Let X be a locally convex space, $h \in X$, and let μ be a bounded Borel measure on X. Then the following conditions are equivalent*:

(1) *for every Borel set A the function $t \mapsto \mu(A + th)$ is absolutely continuous on* $[0, 1]$*;*

(2) *for every open set $U \subset X$ the function $t \mapsto \mu(U + th)$ is absolutely continuous on* $[0, 1]$*;*

(3) *for every bounded Borel function f the function*

$$\varphi_f(t) = \int_X f(x)\,\mu_{th}(dx)$$

is absolutely continuous on $[0, 1]$.

The next result of A.V. Shaposhnikov shows that Skorohod differentiability follows from the "global" absolute continuity of the considered functions.

3.6.14. Theorem. *Let h be a fixed vector in a locally convex space X. Suppose that a Radon measure μ on X has the following property*: *for every $A \in \mathcal{B}(X)$ the function $t \mapsto \mu(A + th)$ is absolutely continuous on the whole real line in the sense that, for every $\varepsilon > 0$, there is $\delta > 0$ such that for each finite collection of pairwise disjoint closed intervals $[s_1, t_1], \ldots, [s_k, t_k]$ of total length $\sum_{i=1}^k |t_i - s_i| < \delta$ one has $\sum_{i=1}^k |\mu_{t_i h}(A) - \mu_{s_i h}(A)| < \varepsilon$. Then μ is Skorohod differentiable along h.*

3.7. Higher order differentiability

For Skorohod or Fomin differentiability, higher order derivatives $d_h^n\mu$ and mixed derivatives $d_{h_1}\cdots d_{h_n}\mu$ are defined inductively. However, there is an alternative to call a measure μ n times Fomin differentiable along $h_1,\ldots,h_n$ if the functions

$$F\colon\ (t_1,\ldots,t_n)\mapsto \mu(A+t_1h_1+\cdots+t_nh_n) \tag{3.7.1}$$

have partial derivatives $\partial_{t_1}\cdots\partial_{t_n}F$ for all sets $A\in\mathcal{B}(X)$. Here one can take $h_1=\cdots=h_n$. In the case of the Skorohod derivative one considers the functions

$$(t_1,\ldots,t_n)\mapsto\int_X f(x+th_1+\cdots+t_nx)\,\mu(dx),\quad f\in C_b(X). \tag{3.7.2}$$

3.7.1. Proposition. *Let μ be a measure on a locally convex space X (Radon or defined on $\sigma(X)$) and let $h\in X$.*

(i) *Suppose that for every $A\in\sigma(X)$ the function $t\mapsto\mu(A+th)$ is n-fold differentiable. Then the measure μ is n-fold Fomin differentiable along h.*

(ii) *Let μ be a Baire measure such that, for every function $f\in C_b(X)$, the function*

$$t\mapsto\int_X f(x-th)\,\mu(dx)$$

is n-fold differentiable. Then the measure μ is n-fold Skorohod differentiable along h and $(n-1)$-fold Fomin differentiable along h.

(iii) *Let $h_1,\ldots,h_n\in X$. The mixed Fomin derivatives $d_{h_{i_1}}\cdots d_{h_{i_k}}\mu$, $1\leqslant i_j\leqslant n$, exist if and only if the functions* (3.7.1) *have partial derivatives $\partial_{t_{i_1}}\cdots\partial_{t_{i_k}}F$ for all sets $A\in\mathcal{B}(X)$. In this case, these functions on $\mathbb{R}^n$ have continuous derivatives of order n. An analogous assertion is true in the case of Skorohod differentiability and functions* (3.7.2).

PROOF. (i) The assertion follows by induction since the derivative of the indicated function is $d_h\mu(A+th)$. (ii) Here induction also applies since the derivative of the indicated function equals the integral of $f(x-th)$ with respect to $d_h\mu$. The $(n-1)$-fold Fomin differentiability follows from the fact that the Skorohod derivative of $d_h^{n-2}\mu$ is continuous being Skorohod differentiable. (iii) Clearly, the existence of all mixed Fomin derivatives yields the existence and continuity of the corresponding mixed partial derivatives of F, which gives its n-fold differentiability on $\mathbb{R}^n$. Conversely, if the indicated partial derivatives of F exist for all A, then, arguing by induction, we obtain the existence of the mixed Fomin partial derivatives of μ up to order $n-1$ and the equality

$$\partial_{t_{i_1}}\cdots\partial_{t_{i_{k-1}}}F(t_1,\ldots,t_n)=d_{h_{i_1}}\cdots d_{h_{i_{k-1}}}\mu(A+t_1h_1+\cdots+t_nh_n).$$

The case of Skorohod differentiability is similar. $\square$

It will be shown in Proposition 6.3.3 in Chapter 6 that for any Radon measure $\mu\geqslant 0$ Fomin differentiable along h, a sufficient condition for the differentiability of the measure $d_h\mu$ is the existence of a version of β_h^μ which is locally absolutely continuous (or everywhere differentiable) on the straight lines $x + \mathbb{R}^1 h$ such that $\partial_h \beta_h^\mu$ is integrable with respect to μ. Then we have $\beta_h^\mu \in L^2(\mu)$ and $d_h^2\mu = [\partial_h \beta_h^\mu + (\beta_h^\mu)^2] \cdot \mu$. According to Proposition 4.3.12, for any convex measure μ this is equivalent to the inclusion $\beta_h^\mu \in L^2(\mu)$. Under some additional conditions the equality for $d_h^2\mu$ can be differentiated, which yields certain expressions for the subsequent derivatives.

3.7.2. Remark. It is readily seen that if a measure μ is n-fold differentiable, then its nth derivative is symmetric, i.e., the measure $d_{h_1} \cdots d_{h_n}\mu$ does not depend on the ordering of $h_1, \ldots, h_n$ (this follows at once from the characterization of differentiability by means of the Fourier transform). Moreover, if, for example, the derivatives $d_h\mu$, $d_k\mu$, and $d_h d_k\mu$ exist in the Fomin or Skorohod sense, then $d_k d_h\mu$ exists in the same sense and equals $d_h d_k\mu$ since the function $l \mapsto -il(k)\widetilde{d_h\mu}(l)$ coincides with the Fourier transform of the measure $d_h d_k\mu$. However, the existence of a partial derivative $d_h d_k\mu$ does not imply the existence of $d_h\mu$. One can construct the corresponding example on the plane (Exercise 3.9.7).

3.8. Convergence of differentiable measures

The situation with the differentiability of limits of sequences of differentiable measures is similar to the one for functions: the limit is differentiable provided that we are given some convergence or boundedness of the derivatives of the convergent elements. The following results illustrate this general rule.

3.8.1. Theorem. *Let $\{\mu_n\}$ be a sequence of Radon measures on a locally convex space X possessing the Skorohod derivatives $d_{h_n}\mu_n$. Suppose that there exists a Radon measure on X such that*

$$\widetilde{\mu}(l) = \lim_{n\to\infty} \widetilde{\mu_n}(l), \quad \forall l \in X^*.$$

Suppose, in addition, that the sequence $\{h_n\}$ converges to a vector h in the weak topology. Then

(i) *if $\sup_n \|d_{h_n}\mu_n\| \leqslant C < \infty$, then the measure μ is Skorohod differentiable along h and $\|d_h\mu\| \leqslant C$;*

(ii) *if $h_n \equiv h$, the measures μ_n are Fomin differentiable along the vector h and the limit $\lim\limits_{n\to\infty} d_h\mu_n(A)$ exists for every set $A \in \mathcal{B}(X)$, then the measure μ is Fomin differentiable;*

(iii) *if $\mu_n \geqslant 0$ and $\sup_n \|\beta_{h_n}^{\mu_n}\|_{L^p(\mu_n)} \leqslant C < \infty$, where $p > 1$, then the measure μ is L^p-differentiable along h and $\beta_h^\mu \in L^p(\mu)$. In this case one has $\|\beta_h^\mu\|_{L^p(\mu)} \leqslant C$.*

Proof. It suffices to prove (i) in the one-dimensional case due to Theorem 3.6.8. Then the claim follows at once from Theorem 3.5.2 and Theorem 3.6.4. Assertion (ii) follows from (i) and Corollary 3.6.6 since the measure $d_h\mu$ turns out to be absolutely continuous with respect to μ. (iii) Let us consider a functional F on the space $\mathcal{FC}_b^\infty(X)$ defined by the formula

$$F(f) = -\int_X \partial_h f(x)\,\mu(dx).$$

By convergence of the Fourier transforms of the nonnegative measures μ_n we have weak convergence of their finite dimensional projections, whence we obtain the equalities

$$\begin{aligned} F(f) &= -\lim_{n\to\infty}\int_X \partial_{h_n} f(x)\mu_n(dx) = \lim_{n\to\infty}\int_X f(x)\,d_{h_n}\mu_n(dx) \\ &= \lim_{n\to\infty}\int_X f(x)\beta_{h_n}^{\mu_n}(x)\,\mu_n(dx). \end{aligned}$$

Therefore, for $q = p/(p-1)$ one has

$$|F(f)| \leqslant \limsup_{n\to\infty} C\biggl(\int_X |f(x)|^q\,\mu_n(dx)\biggr)^{1/q}.$$

The right side equals $C\|f\|_{L^q(\mu)}$. Thus, $|F(f)| \leqslant C\|f\|_{L^q(\mu)}$. Therefore, there exists a function $G \in L^p(\mu)$ such that

$$F(f) = \int_X f(x)G(x)\,\mu(dx) \quad \text{and} \quad \|G\|_{L^p(\mu)} \leqslant C.$$

This yields L^p-differentiability of the measure μ along h and the equality $\beta_h^\mu = G$ (see Theorem 3.6.8). □

An analogous theorem is true for measures on $\sigma(X)$ if we define Skorohod differentiability of μ along h as the existence of a measure ν on $\sigma(X)$ such that $\widetilde{\nu}(l) = -il(h)\widetilde{\nu}(l)$. Then the above proof of (i) covers the finite dimensional case, which gives the desired measure ν of bounded variation on the algebra of cylindrical sets. Its countable additivity follows from the case $X = \mathbb{R}^\infty$, to which the claim reduces since any countable set of cylinders is determined by a countable family of functionals $l_i \in X^*$, which define a map into $\mathbb{R}^\infty$.

It is clear from the proof that this theorem extends to nets in the following formulation.

3.8.2. Theorem. *Assertions* (i)–(iii) *of the previous theorem are valid also for nets $\{\mu_\alpha\}$ and $\{h_\alpha\}$ in place of countable sequences with the difference that in* (ii) *one should require additionally that the measures $d_h\mu_\alpha$ be uniformly bounded and the limit $\lim_\alpha d_{h_\alpha}\mu_\alpha(A)$ define a measure absolutely continuous with respect to μ.*

Similarly one proves the following result.

3.8.3. Proposition. *Let a Radon measure μ on X be Fomin differentiable along vectors h_α which converge weakly to a vector h. If the measures $d_{h_\alpha}\mu$ are uniformly bounded and converge on every Borel set to a measure $\nu \ll \mu$ (for example, converge in variation), then μ is Fomin differentiable along h.*

3.9. Comments and exercises

Differentiable measures were introduced more than 40 years ago by S.V. Fomin [**440**]–[**443**]. Close ideas in terms of the distributions of random processes differentiably depending on parameters had been earlier developed by Pitcher [**905**]–[**908**]. For Gaussian measures, integration by parts was applied in Daletskiĭ [**316**], Gross [**504**]. The first detailed investigation of differentiable measures on linear spaces was undertaken in Averbukh, Smoljanov, Fomin [**82**], [**83**]. In particular, the existence of logarithmic derivatives of Fomin differentiable measures and the important Theorem 3.3.1 were established in [**82**]. One should also note Skorohod's book [**1046**], in which one more important kind of differentiability of measures was introduced. The principal results of all these investigations are briefly presented in Smolyanov [**1050**] and Dalecky, Fomin [**319**]. In the 1970s-80s considerable progress in this area was achieved in the works of Smolyanov's school. A thorough survey of the main achievements in the theory of differentiable measures over the first 20 years of its existence and a detailed discussion of connections with the conceptually close Malliavin calculus was given in Bogachev, Smolyanov [**233**]. These two directions along with subsequent developments are discussed in the extensive survey Bogachev [**189**]. The book Dalecky, Fomin [**319**] includes many results of Yu.L. Daletskiĭ's students. Some aspects of the theory of differentiable measures are presented in the books Norin [**834**] and Uglanov [**1141**], concerned with applications of this theory. At present, hundreds of papers are published on the theory of differentiable measures and its applications. Many of these works are mentioned in different chapters of this book in relation with concrete results or in comments. Continuous measures were implicitly introduced in Averbukh, Smoljanov, Fomin [**82**], where the continuity of differentiable measures was shown; explicitly the term was introduced in Romanov [**965**], where a study of continuous measures was initiated (see also Romanov [**966**]–[**971**]). Theorem 3.5.1 accumulates the results of several works. Its assertion (v), which implies (i), was obtained in Skorohod [**1045**], [**1046**] (see also Gihman, Skorohod [**477**, Ch. VII, §2]). The similar assertion (ii) was proved in Yamasaki, Hora [**1198**], and (iv) is its direct corollary (later it was also proved in Albeverio, Kusuoka, Röckner [**49**]). Finally, (iii) is analogous to (ii) and was noted in Bogachev, Smolyanov [**233**], Khafizov [**602**]. Similar questions are considered in Uglanov [**1138**]. Theorem 3.5.2 was obtained in [**233**]. Among other works related to general problems of the theory of differentiable measures, we mention Bogachev [**166**], [**169**], [**179**], [**181**], Kuo [**663**]–[**670**], Shimomura [**1027**], Uglanov [**1130**], [**1131**]. The absence of Radon probability measures on infinite dimensional spaces equivalent to all its shifts was established in Girsanov, Mityagin [**481**] and Sudakov [**1087**], [**1088**] under some restrictions on measures or spaces; Proposition 3.2.8 generalizes these results. On applications of quasi-invariant measures to representations of the canonical commutation relations in quantum field theory and spectral theory, see Gelfand, Vilenkin [**473**], Samoilenko [**981**], Araki [**77**], [**78**], Hegerfeld [**527**]–[**529**], Hegerfeld, Melsheimer [**530**], and Shimomura [**1030**].

Exercises

3.9.1. Justify assertion (iv) of Proposition 3.4.1.

3.9.2. Let f be a function of bounded variation on $[a,b]$. Prove that the function $|f|$ is of bounded variation as well and $\operatorname{Var}|f| \leqslant \operatorname{Var} f$.

3.9.3. Let $p>1$. Prove that a function f on $[a,b]$ is absolutely continuous and $f' \in L^p[a,b]$ precisely when $\sup \sum_{i=1}^n |f(b_i)-f(a_i)|^p |b_i - a_i|^{1-p} < \infty$, where sup is taken over all disjoint intervals $[a_i,b_i] \subset [a,b]$, and this supremum equals $\|f'\|_{L^p}^p$.

3.9.4. (Hora [**553**], Shimomura [**1028**]) Let ξ be an integrable random variable on a probability space with $\mathbb{E}\xi = 0 < \mathbb{E}|\xi|$. Prove that there exists a probability measure μ with an absolutely continuous density ϱ on the real line such that its logarithmic derivative ϱ'/ϱ as a random variable on the real line with the measure μ has the same distribution as ξ.

3.9.5. Let μ be a Radon probability measure on a locally convex space X that is Skorohod differentiable along two vectors h and k such that $d_h\mu = d_k\mu$. Prove that $h=k$.

HINT: Let $l \in X^*$; taking a sufficiently small $t>0$ with $\widetilde{\mu}(tl) > 0$ we obtain $l(h) = l(k)$ since $\widetilde{d_h\mu}(tl) = \widetilde{d_k\mu}(tl)$, i.e., $l(th)\widetilde{\mu}(tl) = l(tk)\widetilde{\mu}(tl)$.

3.9.6. Let μ be the countable product of Cauchy probability measures μ_n on the real line given by densities $p_n(y) = 2^n\pi^{-1}(1+2^{2n}t^2)^{-1}$. Prove that $\mu(l^2)=1$.

3.9.7. (Yamasaki, Hora [**1198**]) Construct an example of a probability measure μ on $\mathbb{R}^2$ for which the Fomin derivative $d_{e_1}d_{e_2}\mu$ exists, but $d_{e_1}\mu$ does not.

HINT: Take a probability density $\varphi \in C_0^\infty(\mathbb{R}^2)$ and consider a measure with density $\varrho(x,y) = c\sum_{n=1}^\infty n^{-2}\varphi(n^{-1}x + a_n, ny+b_n)$, where a_n, b_n are chosen in such a way that the functions $\varphi(n^{-1}x+a_n, ny+b_n)$ have disjoint supports.

3.9.8. Let a Radon probability measure μ on a locally convex space X be Skorohod differentiable along $h \in X$ and let $l \in X^*$, $l(h)=1$. Prove the inequality

$$\|d_h\mu\| \cdot \int_X |l(x)|\,\mu(dx) \geqslant 1/8.$$

3.9.9. Let μ be a nonzero Radon measure on a locally convex space X such that, for some nonzero $h \in X$, one has $\mu \perp \mu_{th}$ for every $t \neq 0$. Show that $\mu \perp \nu$ for every Radon measure ν that is continuous along h.

HINT: Otherwise there is a Radon probability measure ν continuous along h such that $\mu = \sigma + \eta$, $\sigma \ll \nu$, $\sigma \perp \eta$, and $\|\sigma\| > 0$; then σ is continuous along h, hence $\|\sigma_{th} - \sigma\| \to 0$ as $t \to 0$, which is impossible since

$$2\|\mu\| = \|\mu_{th} - \mu\| \leqslant \|\sigma_{th} - \sigma\| + \|\eta_{th} - \eta\| \leqslant \|\sigma_{th} - \sigma\| + 2\|\eta\|, \quad \|\eta\| < \|\mu\|.$$

3.9.10. Let μ be a Radon measure on a locally convex space X such that $\mu \sim \mu_h$ for some h. Suppose that a sequence of μ-measurable functions converges in measure μ. Show that the functions $f_n(\,\cdot - h)$ converge in measure μ.

HINT: Use Exercise 1.6.23.