Vladimir I. Bogachev

Surface measures in infinite-dimensional spaces

2.1 Introduction

There are three main constructions of surface measures in $\mathbb{R}^n$ with the standard Lebesgue measure. The most general one is based on the Hausdorff measure H^{n-1} of dimension $n-1$, which is a special case of the p-dimensional Hausdorff measure H^p with $0 \le p \le n$. Its role of a "surface measure" is explained by the following two factors: one, typical sets of finite positive H^{n-1}-measure are surfaces of dimension $n-1$, and two, this measure coincides with other natural candidates in cases where comparisons are possible (in particular, it coincides with the usual Lebesgue measure on hyperplanes). For reasonable sets, this surface measure can be obtained as a limit of normalized volumes of metric ε-neighborhoods of these sets. There is a much older construction that is closer to the intuitive understanding of what a surface measure must be: it is a natural measure on a regular surface $S \subset \mathbb{R}^n$, say, on the graph of a smooth function f on $\mathbb{R}^{n-1}$. This means that this surface measure arises as a limit of flat measures on small pieces of tangent hyperplanes approximating the given surface. Finally, one more construction deals with surfaces that are level sets of regular functions and defines the surface measure of the set $\{F = t\}$ as a certain limit of suitably normalized volumes of "neighborhoods" $\{t - \varepsilon \le F \le t + \varepsilon\}$. Locally, if $\nabla F \neq 0$ on the level set, this construction coincides with the previous one; moreover, all the three constructions coincide in this case. However, in general, a set of finite positive H^{n-1}-measure need not be located on a surface (neither a graph nor a level set); certainly, a level set need not be a graph even locally.

When discussing surface measures in infinite-dimensional spaces, it is customary to recall that there are no exact analogs of Lebesgue measures in infinite-dimensions. This is indeed but not a major problem: there are exact infinite-dimensional analogs of other important measures on $\mathbb{R}^n$, for example, Gaussian, and the local theory of surface measures associated with the standard Gaussian measure on $\mathbb{R}^n$ does not differ much from the classical construction. Apparently, the principal difficulty in constructing surface measures in infinite dimensions is that such measures are related to some intrinsic geometry of the measure but not of the space. In other words, it seems that in many cases there is no natural canonical geometry on the space determining surface measures. For example, we shall see below that the countable power $\mathbb{R}^\infty$ of the

Vladimir I. Bogachev: Department of Mechanics and Mathematics, Moscow State University, 119991 Moscow, Russia, St.-Tikhon's Orthodox Humanitarian University, Moscow, Russia, and National Research University Higher School of Economics, Moscow, Russia

real line equipped with the countable power of the standard Gaussian measure carries surface measures associated with Sobolev smooth functions; however, these surface measures have nothing to do with the standard metric on the space $\mathbb{R}^\infty$ (making it a Polish space); moreover, the measure can be restricted to many (continuum of incomparable) weighted Hilbert spaces in $\mathbb{R}^\infty$ of full measure and the surface measures will be unchanged.

Surface measures on general spaces have become a popular subject of study in recent years due to development of the Malliavin calculus, geometric measure theory in metric measure spaces, and infinite-dimensional stochastic analysis, see [1], [2], [3], [4], [5], [6], [7], [8], [13], [15], [20], [21], [22], [24], [25], [26], [29], [37], [43], [44], [62] and [63], where one can find discussions of diverse problems explicitly or implicitly connected with surface measures in infinite dimensions. In the Gaussian case, surface measures related to Gaussian volume measures by versions of the Gauss–Ostrogradskii formula were considered in the 1960–70s by a number of authors, see, e.g., [69], [66], [67], [39], and [47]. Actually, Skorohod [66], [67] considered surface measures for more general quasi-invariant measures. A rich theory of surface measures on infinite-dimensional spaces equipped with differentiable measures was worked out by A.V. Uglanov in the 1970–80s and presented in his book [75] (see also [73], [74], [77], and [32]). In the same years, an approach to surface measures for Gaussian volume measures was developed in the framework of the Malliavin calculus which provided efficient tools for the study of induced measures. For this approach, see [1], [49], [9], [10], and [11]; far reaching generalizations to the case of differentiable measures were obtained in [56], [57], [58], [59], and [60]. A close construction for configuration spaces was presented in [31]. Hausdorff measures associated with Gaussian measures were studied in [36] (see also [35]); more references for the Gaussian case can be found in [11] and [13].

The goal of this survey is to discuss several approaches to surface measures in infinite dimensions with a particular emphasis on the construction from the recent paper [18] that follows Malliavin's idea, but applies to nonlinear spaces and requires less regularity of the function F generating the surface. This surface measure on $F^{-1}(y)$ is a weak limit of the measures $r^{-1}I_{\{y<F<y+r\}} \cdot \theta_F \cdot \mu$ as $r \to 0$, where θ_F is some weight function (for a sufficiently regular surface, one can think of θ_F as the derivative of F along the "normal to the surface"). In the Gaussian case this construction applies to one-fold Malliavin differentiable functions with gradients having divergences. In the nondegenerate case, these surface measures are equivalent to the standard ones. However, this approach leads to much shorter and simpler proofs; in particular, we shall see that the existence of surface measures is proved in a few lines. We also mention some open problems related to infinite-dimensional surface measures. In particular, surface measures on zero sets of polynomials have not been sufficiently studied and this is an interesting direction of research. There are interesting connections between surface measures and Sobolev and BV functions on infinite-dimensional spaces, connections which have been intensively studied in the last decade, see [2], [3], [6], [7],

[8], [20], [21], [24], [48], and [51]. Finally, surface measures are important for the study of boundary value and variational problems in infinite dimensions, see [28] and [76].

The main construction discussed below is related to the concept of conditional measure (recalled in the next section). It makes sense in great generality while surface measures are usually defined in a more special situation, where one can consider suitable neighborhoods of the "surfaces" $\{F = y\}$ and obtain a reasonable limit after appropriate scaling. For example, the usual surface measure in $\mathbb{R}^d$ arises as a limit of the ratio of the volume of the ε-neighborhood of the surface and 2ε, as $\varepsilon \to 0$. The discussed construction of a surface measure σ^y on the level set $F^{-1}(y)$ is this: we introduce a certain weight function θ_F and set

$$\int f(x)\,\sigma^y(dx) := \lim_{r\to 0}\frac{1}{r}\int\limits_{\{y<F<y+r\}} f(x)\theta_F(x)\,\mu(dx)$$

for a suitable class of functions f (say, bounded Lipschitzian). Under our assumptions the surface measure will be actually a weak limit of the measures

$$r^{-1}I_{\{y<F<y+r\}}\theta_F\cdot\mu.$$

Unlike the case of conditional measures, such constructions require certain constraints on measures and functions in question. In the case of a Gaussian measure μ on a locally convex space X this construction applies to a function F in the second Sobolev class $W^{2,2}(\mu)$ and we take $\theta_F = |D_H F|^2$ (or $\theta_F = |D_H F|$ in a modified construction), where $D_H F$ is the Sobolev gradient of F along the Cameron–Martin space H of the measure μ (or to a function F in the first Sobolev class $W^{1,1}(\mu)$ if $D_H F/|D_H F|_H$ has divergence). The weight function θ_F can be later dismissed provided it is sufficiently nondegenerate; its purpose is to allow degenerate F and lower the required order of differentiability of F. This approach can be also of interest for the study of surface measures on metric measure spaces (see [27], [40], [46], and [71]).

Why is it not enough to deal with conditional measures that exist in much greater generality? The reason is essentially the same as in the finite-dimensional case: the Gauss–Ostrogradskii–Stokes formula and integration by parts. This explains at once why certain smoothness restrictions on the volume measure and the function generating level sets are needed. Another reason is that conditional measures μ^y depend not only on the level sets $F^{-1}(y)$, but also on the image-measure $\mu\circ F^{-1}$ (though, for the induced measures with positive densities this dependence reduces to a constant factor for each fixed y). The presented construction shares this property, but allows a modification that does not.

The paper is organized as follows. Section 2 contains notation and terminology. In Section 3 we first discuss Gaussian surface measures in the finite-dimensional case and then explain how the same method works in infinite dimensions. Section 4 is devoted to the main construction in an abstract setting and its relation to conditional measures. Section 5 provides an additional step needed to obtain surface measures on every level set of a given function (not merely on almost every), which involves a brief discussion of capacities. In Section 6 we consider examples, in particular, returning to Gaussian measures. Finally, surface measures of surfaces of higher codimension are discussed in Section 7.

I am grateful to H. Airault, L. Ambrosio, G. Da Prato, N. Gigli, A. Lunardi, I.I. Malofeev, M. Röckner, and G.I. Zelenov for useful discussions. This research was sup-ported by the Russian Foundation for Basic Research Grant 17-01-00662 at Moscow State University.

2.2 Notation and terminology

Let $C_b^\infty(\mathbb{R}^d)$ be the class of all bounded infinitely differentiable functions on $\mathbb{R}^d$ with bounded derivatives, and let $C_0^\infty(\mathbb{R}^d)$ be its subclass consisting of functions with compact support.

Let X be a completely regular topological space with its Borel σ-field $\mathcal{B}$. Let μ be a bounded nonnegative Radon measure on $\mathcal{B}$, i.e., for every Borel set B and every $\varepsilon > 0$, there is a compact set $K_\varepsilon \subset B$ such that $\mu(B\backslash K_\varepsilon) < \varepsilon$ (see [12] for a discussion of such measures). Below we also use signed Radon measures, i.e., Borel measures m such that $|m|$ is a Radon measure, where $|m| = m^+ + m^-$ is the usual total variation of m.

Functions measurable with respect to the Lebesgue completion of the measure μ are called μ-measurable; such a function can be defined μ-almost everywhere μ-a.e., i.e., outside of a set of measure zero.

Given a a measure μ, the measure with density θ with respect to μ is denoted by $\theta \cdot \mu$, i.e.,

$$\theta \cdot \mu(B) = \int_B \theta \, d\mu.$$

In some assertions we shall assume that μ is concentrated on a countable union of metrizable compact sets; this is always the case if the space X is Souslin or metrizable, or if μ is Gaussian. The main definition does not use this assumption and actually applies to general probability spaces; however, the assumption becomes important in order to compare surface measures with conditional measures, and to ensure that our surface measures are indeed concentrated on the level sets.

Given a measurable function $F\colon\ X \to \mathbb{R}$ or a measurable mapping $F\colon\ X \to Y$ with values in a topological space Y, we can take the image-measure $\mu \circ F^{-1}$ defined by the formula

$$\mu \circ F^{-1}(B) := \mu(F^{-1}(B))$$

on the Borel σ-field in $\mathbb{R}$ or Y, respectively.

We recall that a Radon probability measure γ on a locally convex space X is called Gaussian if its one-dimensional images $\gamma \circ l^{-1}$ for all $l \in X^*$ are Gaussian measures on the real line, that is, are either given by densities $(2\pi\sigma)^{-1/2} \exp(-(t-a)^2/(2\sigma))$ with

$\sigma > 0$ or are Dirac point measures. If all these measures are symmetric, then γ is called symmetric or centered; the latter is equivalent to the identity $\gamma(B) = \gamma(-B)$ for all Borel sets B.

The most important examples of Gaussian measures are the standard Gaussian measure γ_n on $\mathbb{R}^n$ with density $(2\pi)^{-n/2}\exp(-|x|^2/2)$, the standard Gaussian measure γ on $\mathbb{R}^\infty$ that is the countable power of the standard Gaussian measure γ_1 on $\mathbb{R}$ (which is defined on the countable power $\mathbb{R}^\infty$ of the real line, i.e., on the space of all real sequences), and the Wiener measure on $C[0,1]$ or $L^2[0,1]$, which can be defined as the image of γ under the mapping

$$(x_n) \mapsto \sum_{n=1}^{\infty} x_n \int_0^t e_n(s)\,ds,$$

where $\{e_n\}$ is an orthonormal basis in $L^2[0,1]$. Certainly, such a definition needs a justification (it turns out that this series converges in $L^2[0,1]$ and even in $C[0,1]$ for almost all $x = (x_n)$). Other equivalent definitions of the Wiener measure can be found in [11] (see also [13] and [14]).

Let γ be a centered Radon Gaussian measure on X. The Cameron–Martin space H of γ consists of all vectors h with finite norm

$$|h|_H := \sup\{l(h):\ l \in X^*,\ \|l\|_{L^2(\gamma)} \le 1\}.$$

It is known that H with this norm is a separable Hilbert space compactly embedded into X; the corresponding inner product is denoted by $(\cdot,\cdot)_H$. A typical example: if γ is the standard Gaussian measure on $\mathbb{R}^\infty$, then $H = l^2$ with the usual Hilbert norm.

For every $h \in H$, there is a measurable linear functional $\widehat{h}$, belonging to the closure of X^* in $L^2(\gamma)$, such that

$$l(h) = \int_X l(x)\widehat{h}(x)\,\gamma(dx) \quad \forall\, l \in X^*.$$

The inner product in H can be defined by the formula

$$(h,k)_H = \int_X \widehat{h}\widehat{k}\,d\gamma.$$

An important role of the Cameron–Martin space is that it is precisely the set of all vectors h such that the shifted measure γ_h defined by $\gamma_h(B) = \gamma(B-h)$ is equivalent to γ; the corresponding Radon–Nikodym derivative is given by the Cameron–Martin formula

$$\exp(\widehat{h} - |h|_H^2/2). \tag{2.1}$$

If a Radon probability measure μ is concentrated on a countable union of metrizable compact sets, then, for any μ-measurable mapping F with values in a complete separable metric space Y, we can find the so-called conditional measures μ^y on X such that the function $y \mapsto \mu^y(B)$ is μ-measurable for each $B \in \mathcal{B}$, μ^y is concentrated on $F^{-1}(y)$ for every y (or $\mu \circ F^{-1}$-a.e. y) and μ is the integral of μ^y against $\mu \circ F^{-1}$, which is written as

$$\mu = \mu^y \cdot \mu \circ F^{-1}(dy),$$

in the sense that

$$\int_X f(x)\,\mu(dx) = \int_Y \int_X f(x)\,\mu^y(dx)\,\mu \circ F^{-1}(dy)$$

for every bounded Borel function f on X; the integral exists due to the assumption of measurability for μ^y, see [12, Chapter 10] or [13, Chapter 1] for details. Actually, conditional measures in a weaker sense exist under more general assumptions about μ and F.

2.3 Surface measures in the Gaussian case

We first consider a smooth function F on $\mathbb{R}^n$ equipped with the standard Gaussian measure γ_n. Suppose that $\nabla F(x) \neq 0$. In this case, the level sets $\{F = t\}$ are smooth surfaces that locally look like graphs of smooth functions. They can be equipped with usual surface (Lebesgue) measures and then Gaussian surface measures can be introduced by simply multiplying these Lebesgue surface measures by the standard Gaussian density. However, we are interested in globally finite Gaussian surface measures. To this end, we shall assume that the function $|\nabla F(x)|^{-1}$ belongs to certain $L^p(\gamma_n)$; the required value of p will become clear soon.

The first step is to verify that the distribution function

$$t \mapsto \gamma_n(x\colon\ F(x) < t)$$

is continuously differentiable. To show this we employ the Malliavin calculus. Let us consider the gradient vector field

$$v(x) = \nabla F(x)$$

and the corresponding differentiation

$$\partial_v g = (\nabla g, v).$$

Let ϱ_n be the standard Gaussian density. For every function $\phi \in C_b^\infty(\mathbb{R})$ we have

$$\int_{\mathbb{R}} \phi'(t)\,\gamma_n \circ F^{-1}(dt) = \int_{\mathbb{R}^n} \phi'(F(x))\,\gamma_n(dx)$$

$$= \int_{\mathbb{R}^n} \phi'(F(x))\partial_v F(x) \frac{1}{\partial_v F(x)}\, \gamma_n(dx) = \int_{\mathbb{R}^n} \partial_v(\phi \circ F) \frac{1}{\partial_v F}\, d\gamma_n. \quad (2.2)$$

Integrating by parts by means of the formula

$$\int_{\mathbb{R}^n} \partial_v f\psi\, d\gamma_n = -\int_{\mathbb{R}^n} f\partial_v\psi\, d\gamma_n \int_{\mathbb{R}^n} f\psi[\operatorname{div} v + (v, \nabla\varrho_n/\varrho_n)]\, d\gamma_n,$$

we represent the right-hand side as

$$-\int_{\mathbb{R}^n} \phi \circ F\Big[\partial_v\Big(\frac{1}{\partial_v F}\Big) + \frac{1}{\partial_v F}\big(\operatorname{div} v + (v, \nabla\varrho_n/\varrho_n)\big)\Big]\, d\gamma_n.$$

We have

$$\partial_v F = |\nabla F|^2, \quad \partial_v(|\nabla F|^2) = 2(D^2F \cdot \nabla F, \nabla F),$$

$$\partial_v\Big(\frac{1}{\partial_v F}\Big) = -\frac{\partial_v|\nabla F|^2}{|\partial_v F|^2} = -\frac{2(D^2F \cdot \nabla F, \nabla F)}{|\nabla F|^4},$$

where

$$\operatorname{div} v + (v, \nabla\varrho_n/\varrho_n) = LF$$

and L is the Ornstein–Uhlenbeck operator defined by

$$Lf(x) = \Delta f(x) - (\nabla f(x), x) = \sum_{i=1}^{n}[\partial^2_{x_i} f(x) - x_i\partial_{x_i} f(x)].$$

Therefore, letting

$$g = \frac{2(D^2F \cdot \nabla F, \nabla F)}{|\nabla F|^4} + \frac{LF}{|\nabla F|^2},$$

we obtain

$$\int_{\mathbb{R}} \phi'(t)\, \gamma_n \circ F^{-1}(dt) = -\int_{\mathbb{R}} \phi(t)\, \eta(dt),$$

where η is the image under F of the measure with density g with respect to γ_n, provided that

$$\int_{\mathbb{R}^n} |g|\, d\gamma_n < \infty.$$

This means that the generalized derivative of $\gamma_n \circ F^{-1}$ is the measure η. Therefore, $\gamma_n \circ F^{-1}$ is an absolutely continuous measure and its density is $\eta((-\infty, t))$. Moreover, the measure η is also absolutely continuous, since it is obviously absolutely continuous with respect to $\gamma_n \circ F^{-1}$. Hence $\gamma_n \circ F^{-1}$ has a continuous density ϱ_1 and

$$\|\varrho_1\| \le \|g\|_{L^1(\gamma_n)}.$$

The integrability of the function g with respect to the standard Gaussian measure is ensured by the integrability with respect to γ_n of the functions

$$\|D^2F(x)\|/|\nabla F(x)|^2 \quad \text{and} \quad |LF(x)|/|\nabla F(x)|^2,$$

where $\|D^2F(x)\|$ is the operator norm of the second derivative (one can also use the Hilbert–Schmidt operator). Therefore, if the function $|\nabla F(x)|^{-2-\varepsilon}$ is γ_n-integrable for some $\varepsilon > 0$ then, by Hölder's inequality, it suffices to have the γ_n-integrability of $\|D^2F(x)\|^{1+2/\varepsilon}$, because it is equivalent to the γ_n-integrability of $|LF(x)|^{1+2/\varepsilon}$, so that applying Hölder's inequality to the second function above we see that both functions will be integrable.

Under these assumptions, the function $\gamma_n(F < t)$ is continuously differentiable and we can assign the value

$$\varrho_1(t) := \frac{d}{dt}\gamma_n(F < t)$$

to the surface $S_t := F^{-1}(t)$. However, it is still not a surface measure but just its value on S_t.

Our next step is to observe that we can define a measure in a similar manner: this measure will be the weak limit of the measures

$$\mu_r(B) := (2r)^{-1}\gamma_n(B \cap \{t - r < F < t + r\})$$

as $r \to 0+$. Let us recall that a sequence of Borel measures μ_j on $\mathbb{R}^n$ converges weakly to a Borel measure μ if

$$\int_{\mathbb{R}^n} f\, d\mu = \lim_{j\to\infty} \int_{\mathbb{R}^n} f\, d\mu_j$$

for every bounded continuous function f. Moreover, for nonnegative measures it suffices if this limit exists for all bounded Lipschitz functions f (see, e.g., [13, Chapter 8]); furthermore, it is enough if it exists for all nonnegative bounded Lipschitz functions f.

It remains to observe that our construction also works if we replace the initial measure γ_n by $\psi\gamma_n$, where ψ is a nonnegative bounded Lipschitz function. Indeed, in our previous calculations we replace ϱ_n with $\psi\varrho_n$ noting that $\nabla\psi(x)$ exists almost everywhere and is bounded. Hence, we obtain the new functions

$$\partial_v\Big(\frac{\psi}{\partial_v F}\Big) = \psi\partial_v\Big(\frac{1}{\partial_v F}\Big) + \frac{(v, \nabla\psi)}{|\nabla F|^2},$$

$$g = \frac{2\psi(D^2F \cdot \nabla F, \nabla F)}{|\nabla F|^4} + \frac{\psi LF}{|\nabla F|^2} + \frac{(\nabla F, \nabla\psi)}{|\nabla F|^2}.$$

This function is integrable under the previous assumptions because ψ and $|\nabla\psi|$ are bounded. Thus, the measures μ_r converge weakly to a bounded Borel measure σ^t (when t is fixed) that we can take for a surface measure. However, this is not a true geometric surface measure; because it depends on the function F and not on only on the level set S_t. In particular, this is not the limit of normalized measures of S_t metric neighborhoods. To obtain a true geometric surface measure, we can consider the surface measure

$$\sigma_0^t := |\nabla F| \cdot \sigma^t,$$

which exists at least locally and is finite if, for example, F satisfies the conditions used above (integrability of certain rations with the second derivative and LF) to ensure that σ^t is finite. Moreover, these conditions can be relaxed: we need

$$\|D^2F\|/|\nabla F| + |LF|/|\nabla F| \in L^1(\gamma_n),$$

and for example, it suffices that $\|D^2F\| \in L^2(\gamma_n)$ and $1/|\nabla F| \in L^2(\gamma_n)$. Actually, the same reasoning as above applies to the measure $|\nabla F| \cdot \gamma_n$ in place of γ_n, so that in (2.2) we have to multiply and divide by $|\nabla F|$ in place of $|\nabla F|^2$. This leads to a weaker integrability condition on $1/|\nabla F|$.

Let us show that σ_t^0 is locally the weak limit of the measures $(2r)^{-1}\gamma_n(\cdot \cap S_t^r)$, where S_t^r is the metric r-neighborhood of S_t. This means that for every Lipschitz function $\phi \geq 0$ with bounded support we have

$$\int_{S_t} \phi \, d\sigma_0^t = \lim_{r\to 0}(2r)^{-1} \int_{S_t^r} \phi \, d\gamma_n. \tag{2.3}$$

We can assume that $t = 0$. Let us fix $\varepsilon > 0$. Taking a smooth partition of unity, we can assume that the support of ϕ is contained in a ball K centered at x_0 so small that $1-\varepsilon \leq |\nabla F(x)|/|\nabla F(x_0)| \leq 1+\varepsilon$ on K. We can also assume, changing coordinates, that $\partial_{x_n}F(x_0) = |\nabla F(x_0)|$, $1 - \varepsilon \leq \partial_{x_n}F(x)/\partial_{x_n}F(x_0) \leq 1 + \varepsilon$ on K and $S_0 \cap K$ is the graph of a smooth function of variables $x_1, \dots, x_{n-1}$. For $r > 0$ small enough, the integral of ϕ against σ_0^0 is $(2r)^{-1}|\nabla F(x_0)|\phi \cdot \mu(|F| < r)$ up to a factor $q \in (1-\varepsilon, 1+\varepsilon)$. The metric r-neighborhood of S_0 in K is contained in the intersection of K with the set $\{|F| \leq (1+\varepsilon)|\nabla F(x_0)|r\}$. Therefore, for $r > 0$ small enough we have

$$(2r)^{-1} \int_{S_0^r} \phi \, d\gamma_n \leq (1+2\varepsilon)|\nabla F(x_0)| \int_K \phi \, d\sigma^0 \leq (1+2\varepsilon)(1+\varepsilon) \int_K \phi \, d\sigma_0^0.$$

On the other hand, the set $\{|F| < r\} \cap K$ is contained in S_0^{qr} with $q = (1-\varepsilon)^{-1}/|\nabla F(x_0)|$, since if $x \in S_0 \cap K$ and $F(x + se_n) > r$, then $s > r(1-\varepsilon)^{-1}|\nabla F(x_0)|^{-1}$ by the equality $F(x + se_n) - F(x) = s\partial_{x_n}F(x + \theta e_n)$. Therefore,

$$(2r)^{-1}|\nabla F(x_0)| \int_{|F|<r} \phi \, d\gamma_n \leq (2r)^{-1}|\nabla F(x_0)| \int_{S_0^{qr}} \phi \, d\gamma_n,$$

which yields that

$$|\nabla F(x_0)| \int_{S_0} \phi \, d\sigma^0 \leq (1-\varepsilon)^{-1} \lim_{r\to 0}(2r)^{-1} \int_{S_0^r} \phi \, d\gamma_n.$$

These bounds yield (2.3).

We now consider a slightly modified construction an advantage of which is exclusion of the non-degeneracy assumption about ∇F. To this end, from the very beginning we replace the original measure γ_n by the measure

$$\nu := \partial_v F \cdot \gamma_n,$$

assuming that $\partial_v F \in L^1(\gamma_n)$; however, the latter assumption can be dismissed if we agree to deal with local surface measures. This trick enables us to write

$$\int_{\mathbb{R}} \phi'(t)\, \nu \circ F^{-1}(dt) = \int_{\mathbb{R}^n} \phi'(F(x))\partial_v F(x)\, \gamma_n(dx)$$
$$= \int_{\mathbb{R}^n} \partial_v(\phi \circ F)\, d\gamma_n = -\int_{\mathbb{R}^n} \phi \circ F \partial_v \varrho_n\, dx.$$

This shows that the generalized derivative of $\nu \circ F^{-1}$ is the image under F of the measure with density $-(\nabla F(x), x)\varrho_n(x)$ which is finite when $|\nabla F| \in L^2(\gamma_n)$ and absolutely continuous with respect to ν. Repeating the construction above we arrive at different surface measures σ_1^t without any non-degeneracy conditions on ∇F. Again, the obtained surface measures are not "geometric". To return to usual surface measures we have to assume that $\nabla F \neq 0$ and in that case we can consider (at least locally) the surface measures $|\nabla F|^{-1} \cdot \sigma_1^t$. The measures σ^t and σ_1^t are related by the equality $\sigma_1^t = |\nabla F|^2 \cdot \sigma^t$.

Example 2.3.1. Suppose that F is a polynomial of $\mathbb{R}^n$ such that $\nabla F(x) \neq 0$. It is known (see, e.g., [72]) there are numbers $c > 0$ and $\alpha > 0$ such that

$$|\nabla F(x)|^2 \geq c(1 + |x|^2)^{-\alpha};$$

hence the function $|\nabla F|^{-2}$ belongs to all $L^p(\gamma_n)$, $p < \infty$. In this case both constructions apply and yield globally finite surface measures on all surfaces $F^{-1}(t)$.

The described second construction has advantages also in the infinite-dimensional case because it does not involve division by $\partial_v F$. We shall now discuss it still in the Gaussian case but in infinite dimensions, and then present it in full generality.

Now let γ be a centered Radon Gaussian measure on a locally convex space X. Without loss of generality, one can assume this is the standard Gaussian measure on $\mathbb{R}^\infty$ or its restriction to a weighted Hilbert space of sequences $x = (x_n)$ with finite norm

$$\Big(\sum_{n=1}^{\infty} c_n x_n^2\Big)^{1/2}, \quad c_n > 0, \sum_{n=1}^{\infty} c_n < \infty.$$

The latter condition ensures this space has measure 1.

Let H be the Cameron–Martin space of γ, i.e., the usual Hilbert space l^2 for the standard Gaussian measure on $\mathbb{R}^\infty$.

Let $\mathcal{FC}$ be the class of all functions of the form

$$f(x) = f_0(l_1(x), \dots, l_n(x)), \quad f_0 \in C_b^\infty(\mathbb{R}^n),\ l_i \in X^*,$$

which for $\mathbb{R}^\infty$ is just the union of all classes $C_b^\infty(\mathbb{R}^n)$. Functions in this class are called smooth cylindrical.

Using the Radon–Nikodym density expression of the shifted measure (see (2.1)) we obtain the equality

$$\int_X t^{-1}[f(x+th) - f(x)]\,\gamma(dx) = \int_X t^{-1}\Big[\exp\big(t\widehat{h}(x) - t^2|h|_H^2/2\big) - 1\Big] f(x)\,\gamma(dx)$$

for all $f \in \mathcal{FC}$; hence, it follows by letting $t \to 0$ that

$$\int_X \partial_h f(x)\,\gamma(dx) = \int_X f(x)\widehat{h}(x)\,\gamma(dx), \tag{2.4}$$

where

$$\partial_h f(x) := \lim_{t\to 0} t^{-1}(f(x+th) - f(x)).$$

This simple formula is the basis for our construction.

The Sobolev class $W^{p,1}(\gamma)$, $p \in [1, +\infty)$, is defined as the completion of the class $\mathcal{FC}$ with respect to the Sobolev norm

$$\|f\|_{p,1} = \|f\|_{L^p(\gamma)} + \|D_H f\|_{L^p(\gamma)} = \|f\|_{L^p(\gamma)} + \Big(\int_X |D_H f(x)|_H^p\,\gamma(dx)\Big)^{1/p},$$

where the gradient $D_H f(x) \in H$ (which now plays the role of $\nabla f(x)$) is defined by

$$(D_H f(x), h)_H = \partial_h f(x).$$

If $\{e_n\}$ is an orthogonal basis in H the vector $D_H f(x)$ has coordinates $\partial_{e_n} f(x)$. For the standard Gaussian measure on $\mathbb{R}^\infty$ functions of class $\mathcal{FC}$ are just smooth functions with bounded derivatives in finitely many variables, and $D_H f(x) = \nabla f(x)$.

One defines similarly the Sobolev classes $W^{p,1}(\gamma, E)$ of mappings with values in a separable Hilbert space E; in this case, $D_H f(x)$ is an operator between H and E, and the Hilbert–Schmidt norm $\|\cdot\|_{HS}$ is used to define the Sobolev norm. This means that in place of $|D_H f(x)|_H$ in the previous formula we use the quantity

$$\|D_H f(x)\|_{HS} = \Big(\sum_{n=1}^\infty |\partial_{e_n} f(x)|_E^2\Big)^{1/2}.$$

As a result of completion, every Sobolev function $f \in W^{p,1}(\gamma)$ obtains a gradient $D_H f$, an L^p-mapping with values in H. On account of (2.4) it satisfies the integration by parts formula

$$\int_X \psi(x)(D_H f(x), h)_H\,\gamma(dx) = -\int_X f(x)[\partial_h \psi(x) - \psi(x)\widehat{h}(x)]\,\gamma(dx)$$

for all $\psi \in \mathcal{FC}$. Actually, this equality extends to $\psi \in W^{q,1}(\gamma)$, $q = p/(p-1)$. By using this directional integration by parts formula, one can show that γ is differentiable along vector fields $v \in W^{p,1}(\gamma, H)$ in the following sense: there is a function $\beta_v \in L^p(\gamma)$ such that for all functions $f \in W^{p',1}(\gamma)$ with $p' = /(p-1)$ one has

$$\int_X (D_H f(x), v(x))_H \, \gamma(dx) = -\int_X \beta_v(x) f(x) \, \gamma(dx). \tag{2.5}$$

In this case for $v(x) = \sum_{n=1}^{\infty} v_n(x) e_n$ we have

$$\beta_v(x) = \sum_{n=1}^{\infty} (\partial_{e_n} v(x) - v_n(x)\widehat{e}_n(x)),$$

where the series converges in $L^p(\gamma)$. The function β_v is called the logarithmic derivative, or divergence of v with respect to γ. If $v(x) = h \in H$ is constant then $\beta_v = \widehat{h}$. Moreover, we can go a step further:

$$\int_X \psi(x)(D_H f(x), v(x))_H \, \gamma(dx) = -\int_X \psi(x)\beta_v(x) f(x) \, \gamma(dx) - \int_X f(x)(D_H \psi(x), v(x))_H \, \gamma(dx), \tag{2.6}$$

where $f, \psi \in W^{2p',1}(\gamma)$ so that $f\psi \in W^{p',1}(\gamma)$ and (2.5) can be applied to $f\psi$.

It should be noted that (2.5) can hold (with some function β_v) for a vector field v not belonging to a Sobolev class; for example, there are irregular vector fields on the plane with zero divergence in the sense of distributions. If (2.5) holds for all smooth cylindrical functions f then γ is called differentiable along the vector field v.

Inductively one defines higher Sobolev classes $W^{p,k}(\gamma, E)$ with derivatives up to order k; actually, we only need $k = 1, 2$. For example, the class $W^{p,2}(\gamma)$ consists of all functions $f \in W^{p,1}(\gamma)$ such that $D_H f \in W^{p,1}(\gamma, H)$. Therefore, the measure γ is differentiable (in the sense explained above) along the gradient field $v = D_H F$ once $F \in W^{p,2}(\gamma)$. In this case

$$\beta_v = LF,$$

where L is the Ornstein–Uhlenbeck operator; for $\mathbb{R}^\infty$ it is defined as the closure of the operator

$$Lf(x) = \sum_i [\partial_{x_i}^2 f(x) - x_i \partial_{x_i} f(x)]$$

on smooth cylindrical functions. However, belonging to the second Sobolev class is not necessary for the existence of divergence of $D_H F$. This happens already in the finite-dimensional case.

Now, given a function $F \in W^{p,2}(\gamma)$, where γ is the standard Gaussian measure on $\mathbb{R}^\infty$, with some $p > 1$ such that

$$\frac{1}{|D_H F|^2} \in L^{p'}(\gamma),$$

we consider the exact analogs of the two constructions of surface measures on the level sets $S_t = F^{-1}(t)$ considered above for $\mathbb{R}^n$. In the first construction we have the following exact analog of (2.2):

$$\int_{\mathbb{R}} \phi'(t)\,\gamma \circ F^{-1}(dt) = \int_X \partial_v(\phi \circ F)\frac{1}{\partial_v F}\,d\gamma.$$

By using (2.6) and the equality

$$\partial_v(D_H F(x), D_H F(x))_H = 2(D_H^2 F(x) \cdot D_H F(x), D_H F(x))_H,$$

where $D_H^2 F(x) \cdot D_H F(x)$ is the action of the operator $D_H^2 F(x)$ on the vector $D_H F(x)$, we write the right-hand side as

$$-\int_X \phi(F)\Big[\frac{\beta_v}{\partial_v F} - \frac{\partial_v^2 F}{|\partial_v F|^2}\Big]\,d\gamma = -\int_X \phi(F)\Big[\frac{LF}{|D_H F|_H^2} - \frac{2(D_H^2 F \cdot D_H F, D_H F)_H}{|D_H F|_H^4}\Big]\,d\gamma.$$

The integral on the right exists since we have LF, $\|D_H^2 F\|_{HS} \in L^p(\gamma)$ by the assumption that $F \in W^{p,2}(\gamma)$; hence, $|LF|/|D_H F|_H^2$, $\|D_H^2 F\|_{HS}/|D_H F|_H^2 \in L^1(\gamma)$ by Hölder's inequality. Next we show that similar equalities hold if we replace the measure γ by $\psi \cdot \gamma$, where ψ is a bounded function that is Lipschitz on $\mathbb{R}^\infty$ with respect to the metric

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \max(|x_n - y_n|, 1).$$

To this end, we observe that such a function is Lipschitz along H, i.e.,

$$|\psi(x + h) - \psi(x)| \le L|h|_H \quad \text{for all } x \in \mathbb{R}^\infty \text{ and } h \in H = l^2.$$

Indeed, $\max(|h_n|, 1) \le |h|_H$ for all n, hence $d(x + h, x) = d(h, 0) \le |h|_H$. It is known (see [11, Section 5.11]) that this yields the inclusion $\psi \in W^{s,1}(\gamma)$ for all $s \in [1, +\infty)$ and $|D_H \psi|_H \le L$. Hence, in the calculations above we have to replace $1/\partial_v F$ by $\psi/\partial_v F$ so that in place of $\partial_v(1/\partial_v F)$ we have $\partial_v \psi/\partial_v F + \psi \partial_v(1/\partial_v F)$. By the equality

$$\partial_v \psi/\partial_v F = (D_H \psi, D_H F)_H/|D_H F|_H$$

and boundedness of $D_H \psi$, this yields finite integrals in analogous calculations. Therefore, as in the finite-dimensional case, the measures

$$B \mapsto (2r)^{-1}\gamma(B \cap \{t - r < F < t + r\})$$

converge weakly to finite measures σ^t as $r \to 0+$.

Finally, in place of sets $\{t - r < F < t + r\}$ we could deal with sets $\{t < F < t + r\}$ or $\{t \le F < t + r\}$ and divide by r in place of $2r$ in the appropriate places; this will be done in the next section for the sake of some minor technical simplifications.

The measures σ_0^t can again be defined by the equality

$$\sigma_0^t = |D_H F| \cdot \sigma^t$$

provided that $|D_H F|$ is σ^t-integrable. This condition automatically fulfilled if $|D_H F|$ is bounded or under the assumptions $F \in W^{2,p}(\gamma)$ and $1/|D_H F|^2 \in L^{p'}(\gamma)$ used above. Moreover, as in the finite-dimensional case, even weaker assumptions are sufficient: it is enough to have

$$F \in W^{2,2}(\gamma), \quad 1/|D_H F| \in L^2(\gamma).$$

This is verified by the same method with the measure $|D_H F| \cdot \gamma$ in place of γ.

The second construction with the measure $|D_H F|^2 \cdot \gamma$ in place of γ is completely analogous; we obtain finite measures σ_1^t as weak limits of the measures

$$B \mapsto (2r)^{-1}(|D_H F|^2 \cdot \gamma)(B \cap \{t - r < F < t + r\}), \quad r \to 0+.$$

We could also in this case deal with measures of the sets $\{t < F < t+r\}$ or $\{t \le F < t+r\}$ divided by r.

However, now we have a problem that arises also in $\mathbb{R}^n$ if we do not assume the continuity of F (which does not follow from the membership in $W^{p,2}(\mathbb{R}^n)$ for large n). Namely, we cannot assert that σ^t or σ_1^t is concentrated on S_t. We shall solve this problem in the next section in a general setting. To do so we compare our surface measures with conditional measures and show that they are concentrated on the level sets S_t for almost all t (with respect to the image measure). Moreover, in Section 5 we involve Sobolev capacities to construct surface measures concentrated on the corresponding surfaces for all t.

It is worth noting that an exact analog of the finite-dimensional situation considered above arises under the following two conditions: (i) we restrict our measure γ to a weighted Hilbert space Y of sequences $x = (x_n)$ with finite norm $\|x\|_Y$ defined by $\|x\|_Y^2 = \sum_{n=1}^\infty c_n x_n^2$, where $c_n > 0$ and $\sum_{n=1}^\infty c_n < \infty$ (in this case $\gamma(Y) = 1$), (ii) F is least twice continuously Fréchet differentiable on Y. In that case, the measures σ^t and σ_1^t will be concentrated on S_t for each t. However, in the definition of an analog of the measure σ_0^t we now have two non-equivalent options: we can take the measures $|D_H F|_H \cdot \sigma^t$ or the measures $\|\nabla F\|_Y \cdot \sigma^t$. The former corresponds to r-neighborhoods of S_t with respect to the norm of H, i.e., to the sets $S_t + rU_H$, where U_H is the unit ball of H; while the latter corresponds to the norm of Y. The relation to Y is, on the one hand, natural, since γ is concentrated on Y (but not on H); however, on the other hand, there is no preference in our choice of Y (there are too many suitable spaces).

Consider continuously Fréchet differentiable function F on Y with $\nabla F \neq 0$. We can define surface measures locally by considering small neighborhoods U in which S_t looks like the graph of a continuously differentiable function G on a closed hyperplane Y_0 in Y. This is always possible by the implicit function theorem; moreover, it is possible to choose such a hyperplane Y_0 in such a way that it is orthogonal in Y to a vector h from the Cameron–Martin space H. This simplifies the previous construction, since we can take a constant vector field: $v(x) = h$. To minimize changes in the discussed construction, it is convenient to replace γ by a measure of the form $g \cdot \gamma$, where $g \geq 0$ is a Lipschitz function with support in a ball U_0 of radius r_0 such that the ball with the twice larger radius belongs to the neighborhood U. In addition, functions ψ with support in U will be taken. We can assume that $\partial_h F \geq c > 0$ in U, so the function $1/\partial_h F$ is bounded in U. Hence, both constructions are applicable in this case. In principle, if F is continuously Fréchet differentiable and $\nabla F \neq 0$, we can introduce global surface measures as possibly σ-finite measures by summing local surface measures.

Example 2.3.2. Let us consider the case where $F = \widehat{h}$ is a measurable linear functional. Typically, it has no continuous version. For example, in the case of the standard Gaussian measure on $\mathbb{R}^\infty$ only finite linear combinations of coordinate functions are continuous. Series $\sum_{n=1}^\infty c_n x_n$ with infinitely many nonzero coefficients c_n have no continuous versions; the stochastic integral

$$\int_0^1 \psi(t)\, dx(t)$$

on the Wiener space, where $\psi \in L^2[0,1]$, has a continuous version precisely when ψ has a version of bounded variation. We can assume that $|h|_H = 1$. Then $|D_H \widehat{h}|_H = |h|_H = 1$ and all the three surface measures σ^t, σ_0^t and σ_1^t coincide. They can be calculated by using the connection with conditional measures that are known (see the next section), but this can be also done directly. Let us recall that any Radon measure is uniquely determined by its Fourier transform, i.e., the integrals of the functions $\exp(il)$ for $l \in X^*$. According to our construction, the integral of $\exp(il)$ against σ^t is the limit of the expressions

$$(2r)^{-1} \int_{\{t-r<\widehat{h}<t+r\}} \exp(il)\, d\gamma.$$

We can write $l = c\widehat{h} + \xi$, where $c \in \mathbb{R}$, $\xi = \widehat{u}$ for some $u \in H$ such that ξ and $\widehat{h}$ are orthogonal in $L^2(\gamma)$ (equivalently, $(h,u)_H = 0$). We shall use a proper linear version of ξ (which exists); in this case it is known that $\xi(h) = (u,h)_H = 0$. The orthogonal measurable linear functionals ξ and $\widehat{h}$ are independent Gaussian random variables, hence

$$\begin{aligned}(2r)^{-1} \int_{\{t-r<\widehat{h}<t+r\}} \exp(il)\, d\gamma &= (2r)^{-1} \int_{\{t-r<\widehat{h}<t+r\}} \exp(ic\widehat{h})\, d\gamma \int_X \exp(i\xi)\, d\gamma \\ &= (2r)^{-1}(2\pi)^{-1/2} \int_{t-r}^{t+r} \exp(ics)\exp(-s^2/2)\, ds \int_X \exp(i\xi)\, d\gamma,\end{aligned}$$

which tends to

$$(2\pi)^{-1/2}\exp(ict - t^2/2)\exp(-\|\xi\|_2^2/2)$$

as $r \to 0$. If $t = 0$, then we see that σ^0 coincides up to the factor $(2\pi)^{-1/2}$ with the Gaussian measure η that is the image of γ under the linear mapping $Px = x - \widehat{h}(x)h$. Indeed, the Fourier transform of $\gamma\circ P^{-1}$ at the functional l represented as above equals

$$\int_X \exp\big(il(x - \widehat{h}(x)h)\big)\,\gamma(dx) = \int_X \exp\big(i(c\widehat{h} - \xi)(x - \widehat{h}(x)h)\big)\,\gamma(dx)$$
$$= \int_X \exp\big(i\xi(x)\big)\,\gamma(dx) = \exp(-\|\xi\|_2^2/2),$$

since $\xi(h) = (u, h)_H = 0$. Therefore, the measure σ^t is the shift of the measure σ^0 by the vector th.

We conclude this section by considering an analog of “Gaussian” Hausdorff measures associated with the geometry of the Cameron–Martin space as proposed by Feyel and de La Pradelle in [36]. This construction begins from $\mathbb{R}^n$. Recall (see [12], [33], [34], [79]) that for every $m \in (0, n]$ the classical Hausdorff measure H_m is generated by means of the outer measure H_m^δ defined for each set A by

$$H_m^\delta(A) = \inf\sum_{j=1}^\infty C_m(\operatorname{diam} A_j)^m, \qquad C_m = 2^{-m}\Gamma(1/2)^m/\Gamma(1+m/2),$$

where inf is taken over all sequences of closed sets A_j of diameter at most δ with union containing A. The values $H_m^\delta(A)$ increase as $\delta \to 0+$ and have a limit (possibly, infinite) denoted by $H_m(A)$. If we apply this method with balls in place of arbitrary closed sets the result will be the spherical Hausdorff measure $\mathcal{S}_m$ that is larger than H_m (on sufficiently regular sets they coincide). Through this approach, the following “Gaussian spherical Hausdorff” measures θ_k on $\mathbb{R}^n$ were introduced:

$$\theta_k(B) = \varrho_n \cdot \mathcal{S}_{n-k}(B),$$

where $\mathcal{S}_{n-k}(B)$ is the limit as $\delta > 0+$ of the infimum of $\sum_{j=1}^\infty \lambda_{n-k}(B_i)$ over all covers of B by closed balls of radius at most δ and $\lambda_{n-k}(B_i)$ is the $(n-k)$-dimensional volume of B_i measured as the $(n-k)$-dimensional volume of the section of B_i by a subspace of dimension $n-k$ passing through the center of B_i (which equals $const(n-k)r_i^{n-k}$, where r_i is the radius of B_i). The number k in this notation refers to “codimension”.

The next step is to fix k and take an n-dimensional subspace H_n in H. The orthogonal projection $P_n\colon H \to H_n$ admits a measurable linear extension $\widehat{P}_n \to X \to H_n$: if $e_1, \dots, e_n$ is an orthonormal basis in H_n, then

$$\widehat{P}_n x := \widehat{e}_1(x)e_1 + \cdots \widehat{e}_n(x)e_n.$$

Let γ_n be the image of γ under the measurable linear operator $I - \widehat{P}_n$. For every Borel (or Souslin) set A in X set

$$\eta_k^{H_n}(A) := \int_X \theta_k(A_x)\,\gamma_n(dx), \quad A_x = \{y \in H_n\colon\ x + y \in A\}.$$

The section A_x is Borel in H_n if A is Borel (and is Souslin for Souslin A). In addition, the function $x \mapsto \theta_k(A_x)$ is measurable with respect to all Borel measures. Finally, the Gaussian Hausdorff measure η_k of codimension k on X is defined as follows: $\eta_k(A)$ is the supremum of $\eta_k^{H_n}(A)$ over all n-dimensional subspaces in H with $n \geq k$.

These Gaussian Hausdorff measures are related to the initial Gaussian measure γ by the following formula established in [36] and presented here for simplicity in the case $k = 1$: if $f \in W^{p,2}(\gamma)$ for all $p \in [1, +\infty)$ and f is continuous (actually, it suffices that f be quasi-continuous with respect to the Sobolev capacity corresponding to the class $W^{p,1}(\gamma)$, see the next section), then

$$\int_A |D_H f|\,d\gamma = \int_{\mathbb{R}} \eta_1(A \cap \{|D_H f| > 0\} \cap f^{-1}(t))\,dt.$$

If $|D_H f| > 0$, then we integrate $\eta_1(A \cap f^{-1}(t))$ on the right.

Similarly to the finite-dimensional case, an obvious advantage of Gaussian Hausdorff measures is that their construction is absolutely independent of any particular representations of sets. The other side of this universality is that it is rather difficult to calculate surface measures of given sets; this happens already in $\mathbb{R}^n$. Actually, the last formula can help in such calculations: if $|D_H f| > 0$ and $A = B \cap \{f < s\}$, where B is a Borel set and $s \in \mathbb{R}$, then we have

$$\int_{\{f<s\}\cap B} |D_H f|\,d\gamma = \int_{-\infty}^{s} \eta_1(B \cap f^{-1}(t))\,dt.$$

Letting $\eta_{1,t}(B) = \eta_1(B \cap f^{-1}(t))$, for bounded Borel functions ϕ we obtain

$$\int_{\{f<s\}} \phi |D_H f|\,d\gamma = \int_{-\infty}^{s} \int_X \phi\,d\eta_{1,t}\,dt.$$

Therefore, $\eta_1(B \cap f^{-1}(t))$ can be obtained as the derivative of the distribution function $I_B|D_H f| \cdot \gamma(f < t)$; this identifies η_1 on $f^{-1}(t)$ with the surface measure σ_0^t considered above.

It is not clear whether every surface of the form $S = f^{-1}(0)$, where f is a continuous polynomial on a Hilbert space with a Gaussian measure γ such that $\nabla f(x) \neq 0$, has a finite surface measure; certainly, locally all these approaches give nice surface measures. The problem is that it is not known whether the function $|\nabla f(x)|^{-p}$ is integrable (it is even unknown whether there is $p > 0$ for which it is integrable). Unlike the finite-dimensional case, there is no lower bound of the form $Q(x) \geq c(1+|x|^2)^{-\alpha}$ for any continuous polynomial $Q > 0$. It is also worth noting that zero sets of continuous polynomials on a Hilbert space are more complicated sets than in $\mathbb{R}^n$. For example, the cardinality of the set of disjoint connected components of $f^{-1}(0)$ can be continuum and the class of orthogonal projections of such sets coincides with the class of all Souslin sets (see [11, Exercise 6.11.19]). A survey of results on distributions of polynomials is given in [16].

2.4 Surface measures for differentiable measures

Here we described an abstract approach to surface measures suggested in [18], where the proofs of some technical assertions can be found. For the reader's convenience we include justifications of the most important steps.

Let $\mu \geq 0$ be a fixed Radon measure on a completely regular space X and let $\mathcal{B}$ be the Borel σ-algebra of X. Let $\mathcal{F}$ be a class of bounded $\mathcal{B}$-measurable real functions. Recall that a class of functions *separates measures* if two measures coincide whenever they assign equal integrals to all functions in this class. We assume throughout that $\mathcal{F}$ satisfies the following conditions:

(F1) $\mathcal{F}$ is a linear space separating Radon measures on X, and $\phi(f) \in \mathcal{F}$ for all $f \in \mathcal{F}$ and all $\phi \in C_b^\infty(\mathbb{R})$.

For example, if X is a metric space the class of all bounded Lipschitzian functions on X satisfies conditions (F1). We shall see in Section 4 that this class is indeed convenient for many applications.

Another example: given some class $\mathcal{F}_0$ of $\mathcal{B}$-measurable functions, we take for $\mathcal{F}$ the class of all compositions $\phi(f_1, \dots, f_n)$, where $f_i \in \mathcal{F}_0$ and $\phi \in C_b^\infty(\mathbb{R}^n)$. This class is a linear space and is stable under compositions with C_b^∞-functions; certainly, we still need the additional condition that it must separate measures (which trivially holds if $\mathcal{F}_0$ is separating). See also the modification of (F1) for multidimensional mappings considered in Section 7.

It follows from (F1) that $1 \in \mathcal{F}$ and that $fg \in \mathcal{F}$ for all $f, g \in \mathcal{F}$. Indeed, $f^2 \in \mathcal{F}$ for all $f \in \mathcal{F}$, because we can take for ϕ a function in $C_b^\infty(\mathbb{R})$ that coincides with x^2 on the bounded range of f, so it remains to use the equality $2fg = (f+g)^2 - f^2 - g^2$.

Definition 2.4.1. *A vector field on X (or an $\mathcal{F}$-vector field if we need to indicate its relation to $\mathcal{F}$) is a linear mapping (differentiation)*

$$v\colon\ \mathcal{F} \to L^1(\mu), \quad f \mapsto \partial_v f,$$

such that

$$\partial_v(\phi \circ f) = \phi'(f)\partial_v f \quad \mu\text{-a.e.} \tag{2.7}$$

for all $f \in \mathcal{F}$ and $\phi \in C_b^\infty(\mathbb{R})$.

Similarly we can define more general vector fields for which functions $\partial_v f$ belong to the space $L^0(\mu)$ of μ-measurable functions.

Applying (2.7) to $f, g, f+g$ and ϕ such that $\phi(t) = t^2$ on a sufficiently large interval we obtain the Leibniz rule

$$\partial_v(fg) = f\partial_v g + g\partial_v f \text{ a.e.} \quad \forall f, g \in \mathcal{F}. \tag{2.8}$$

It is worth noting that $\partial_v 1 = 0$ because we can take $\phi = 1$ in (2.7) or, alternatively, we can take $f = g = 1$ in (2.8).

Below a fixed vector field v will play a role of a normal field on level sets (and in some cases one can use indeed unit normal fields).

Definition 2.4.2. *The measure μ is called Skorohod differentiable along v (with respect to $\mathcal{F}$) if there is a Radon measure $d_v\mu$ on $\mathcal{B}$, called the Skorohod derivative of μ along v, such that*

$$\int_X \partial_v f(x)\,\mu(dx) = -\int_X f(x)\,d_v\mu(dx) \quad \forall f \in \mathcal{F}. \tag{2.9}$$

We say that μ is Fomin differentiable along v if $d_v\mu \ll \mu$; in that case the Radon–Nikodym density of $d_v\mu$ with respect to μ is denoted by β_v and is called the logarithmic derivative of μ along v or divergence of v with respect to μ.

For example, let μ be a measure on $\mathbb{R}^d$ with a smooth density ϱ, $\mathcal{F}$ be the class of all bounded Lipschitz functions or the classes $C_b^\infty(\mathbb{R}^d)$, and v be a nonzero constant vector (so that $\partial_v f$ is the usual partial derivative). Then $d_v\mu$ is given by density $\partial_v\varrho$ and $\beta_v = (\partial_v\varrho)/\varrho$, which explains the terminology. If $v = 1$ on the real line the usual Lebesgue measure λ on $[0, 1]$ regarded as a measure on $\mathbb{R}$ is Skorohod differentiable, and $d_1\lambda = \delta_0 - \delta_1$ is the difference of two Dirac measures and Fomin differentiable.

In the case of $\mathbb{R}^d$ with $\mathcal{F}$ as above, a measure μ is Skorohod differentiable along all constant vectors precisely when it has a density ϱ belonging to the class $BV(\mathbb{R}^d)$ of functions of bounded variation, that is, functions in $L^1(\mathbb{R}^d)$ whose first order partial derivatives in the sense of distributions are bounded measures (see [33], [79]); in that case, for a constant vector v, the measure $d_v\mu$ is the partial derivative of μ along v in the sense of distributions. The measure μ is Fomin differentiable along all constant vectors in $\mathbb{R}^d$ precisely when it has a density in the Sobolev class $W^{1,1}(\mathbb{R}^d)$ of integrable functions possessing integrable first order partial derivatives in the sense of distributions. If $\mu = \varrho\,dx$ and $\varrho \in W^{1,1}(\mathbb{R}^d)$, then for any constant vector v one also has $\beta_v = \partial_v\varrho/\varrho$, where we set $\beta_v = 0$ on the set of zeros of ϱ. If a vector field v on $\mathbb{R}^d$ is not constant, but is bounded and Lipschitzian, then

$$\beta_v = \operatorname{div} v + \partial_v\varrho/\varrho.$$

This is also true for vector fields belonging to appropriate Sobolev classes. For a survey of the theory of differentiable measures, see [13].

If μ is a centered Gaussian measure with the Cameron–Martin space H the measure μ is Fomin differentiable along the constant vector field h and $\beta_h = -\widehat{h}$, which is a trivial corollary of the Cameron–Martin formula (2.1).

The original definition of Fomin dealt with constant vector fields on linear spaces. Differentiability of measures along non-constant vector fields was already considered in the 1980–1990s (sometimes implicitly) in the Malliavin calculus and its modifications (see [70], [30], [68], [9], and [31]); close constructions arise in relation to the so-called “carré du champ” operators (see [23]). Obviously, to be differentiable depends on $\mathcal{F}$. However, in reasonable situations differentiability with respect to small classes (separating measures) often yields differentiability with respect to larger classes. For example, in the case of $v(x) = 1$ on the real line, differentiability with respect to C_b^∞ yields differentiability with respect to the class of bounded Lipschitz functions.

Observe that $d_v\mu(X) = 0$, which follows from (2.9) applied to $f = 1$, so $d_v\mu$ is necessarily a signed measure.

We need an extension of ∂_v to functions outside of $\mathcal{F}$.

Definition 2.4.3. *Let μ be Skorohod differentiable along a vector field v. We say that a $\mathcal{B}$-measurable function Ψ belongs to $\mathfrak{D}_v$ if $\Psi \in L^1(\mu) \cap L^1(d_v\mu)$ and there is a sequence of functions $f_n \in \mathcal{F}$ converging to Ψ in $L^1(\mu)$ and in $L^1(d_v\mu)$ such that the functions $\partial_v f_n$ converge in $L^1(\mu)$ to some function w and the functions $f_n\partial_v g$ converge in $L^1(\mu)$ for each $g \in \mathcal{F}$. Then we set $\partial_v\Psi := w$.*

Note that by convergence of $\{f_n\}$ in $L^1(\mu)$ the sequence $\{f_n\partial_v g\}$ converges in $L^1(\mu)$ precisely when it is uniformly μ-integrable. This condition holds if $\{f_n\}$ converges to Ψ in $L^p(\mu)$ for some $p > 1$ and all functions $\partial_v g$ for $g \in \mathcal{F}$ belong to $L^q(\mu)$, $q = p/(p-1)$.

This definition is somewhat technical because we want to make it sufficiently broad. Similar technicalities already arise on the real line if one wants to integrate by parts unbounded functions with respect to measures with densities of bounded variation (but not absolutely continuous) such that the derivatives can be singular. In the case of a Fomin differentiable measure μ with $\beta_v \in L^2(\mu)$, it would be natural to say that a function $\Psi \in L^2(\mu)$ has a Sobolev derivative $\partial_v\Psi \in L^2(\mu)$ if there is a sequence of functions f_n converging to f in $L^2(\mu)$ such that $\{\partial_v f_n\}$ also converges in $L^2(\mu)$. In that case, the limiting function for $\{\partial_v f_n\}$ would satisfy the integration by parts formula. The definition above follows a similar idea under weaker integrability conditions.

The function w (if it exists) is uniquely defined. Indeed, for each $g \in \mathcal{F}$ we have

$$\int_X g(x)w(x)\,\mu(dx) = \lim_{n\to\infty}\int_X g(x)\partial_v f_n(x)\,\mu(dx)$$
$$= \lim_{n\to\infty}\int_X [\partial_v(gf_n)(x) - f_n(x)\partial_v g(x)]\,\mu(dx)$$

$$= -\lim_{n\to\infty}\int_X (gf_n)(x)\, d_v\mu(dx) - \int_X \Psi\partial_v g\, \mu(dx)$$

$$= -\int_X (g\Psi)(x)\, d_v\mu(dx) - \int_X \Psi(x)\partial_v g(x)\, \mu(dx).$$

Therefore, the integral of gw is determined for each $g \in \mathcal{F}$, which uniquely determines w according to Condition (F1).

Remark 2.4.4. If f is bounded a sequence $\{f_n\}$ with the properties mentioned in the definition can be replaced by a uniformly bounded sequence; the technical condition of uniform integrability will therefore be fulfilled automatically. Indeed, taking $\zeta \in C_b^\infty(\mathbb{R})$ such that $\zeta(t) = t$ on an interval containing the range of f, we obtain a new sequence $g_n = \zeta(f_n)$ that is uniformly bounded and converges to f in $L^1(\mu)$ and $L^1(d_v\mu)$. In addition, the functions $\partial_v g_n = \zeta'(f_n)\partial_v f_n$ converge to $\partial_v f$ in $L^1(\mu)$.

Let μ be Skorohod differentiable along v. We shall assume that $F\colon X \to \mathbb{R}$ is a $\mathcal{B}$-measurable function such that

(F2) $\psi(F) \in \mathfrak{D}_v$ for each function $\psi \in C_0^\infty(\mathbb{R})$ and there is a $\mathcal{B}$-measurable function $\partial_v F$ such that $\partial_v(\psi \circ F) = \psi'(F)\partial_v F$ a.e. for each $\psi \in C_0^\infty(\mathbb{R})$. Moreover,

$$\partial_v F \geq 0, \quad \partial_v F \in L^1(\mu).$$

Set

$$\nu := (\partial_v F)\cdot\mu, \quad \eta := d_v\mu \circ F^{-1}. \tag{2.10}$$

The measure ν is finite and nonnegative (it can be zero). The conditional measures on the level sets $F^{-1}(y)$ generated by the measure ν will be denoted by ν^y (in the case where μ is concentrated on a countable union of metrizable compact sets).

We now introduce our surface measures σ^y; it might be reasonable to use the symbol σ_v^y to emphasize dependence on v, but we omit this indication for notational simplicity. The definition employs only the differentiability of the distribution functions

$$\Phi_f(y) := \int_{\{F<y\}} f(x)\,\nu(dx)$$

at a given point. In this respect, no topological structures are needed. However, for deriving further properties of our surface measures we shall need some additional assumptions of topological nature. We set

$$\varrho_f(y) = \Phi_f'(y) = \lim_{h\to 0}\frac{\Phi_f(y+h) - \Phi_f(y)}{h} \tag{2.11}$$

if a finite limit exists.

Definition 2.4.5. *Let $y \in \mathbb{R}$. Suppose that Φ_f is differentiable at y for each $f \in \mathcal{F}$ and there is a Radon measure σ^y on $\mathcal{B}$ such that*

$$\int_X f(x)\,\sigma^y(dx) = \varrho_f(y) \quad \forall f \in \mathcal{F}. \tag{2.12}$$

Then σ^y is called the surface measure associated with the level set $F^{-1}(y)$.

Remark 2.4.6. Note that we do not require that the surface measure be concentrated at the level set $F^{-1}(y)$, but under broad assumptions (see the next theorem) it is indeed a measure on $F^{-1}(y)$. In this respect, the situation is similar with conditional measures.

However, if F is continuous (which we do not assume) and, for every $z \in X \backslash F^{-1}(y)$ and every neighborhood U of z, there is a nonnegative function $f \in \mathcal{F}$ with support in U and positive in a neighborhood of z (which is fulfilled if $\mathcal{F}$ contains all bounded Lipschitz functions on a metric space) then σ^y is automatically concentrated on $F^{-1}(y)$. This is readily seen from (2.12) because we can take U such that $|F(x)-y| > |F(z)-y|/2$ for all $x \in U$; by (2.11) this yields $\varrho_f(y) = 0$. Hence, the integral of f against σ^y vanishes, so that z does not belong to the topological support of σ^y.

A similar reasoning shows that σ^y is automatically concentrated on $F^{-1}(y)$ provided that F satisfies $\phi(F) \in \mathcal{F}$ for all $\phi \in C_b^\infty(\mathbb{R})$. In this case, considering suitable functions $f = \phi(F)$, we obtain that the sets $\{F > y + 1/n\}$ and $\{F < y - 1/n\}$ have σ^y-measure zero for all $n \in \mathbb{N}$. The latter condition with compositions is fulfilled if in the Gaussian case (see Section 3 and Section 5) we take for $\mathcal{F}$ the class of all bounded functions in the Sobolev space $W^{2,2}(\mu)$ and $F \in W^{2,2}(\mu)$.

We shall see that the hypothesis of differentiability of Φ_f is fulfilled if μ is Fomin differentiable along v and F satisfies (F2). In typical cases, the assumptions of this definition are ensured by the following condition: the measures $\nu_r = r^{-1} I_{\{y<F<y+r\}} \cdot \nu$ converge weakly as $r \to 0$, which in turn is ensured by their uniform tightness and convergence of the integrals against ν_r of bounded functions from a measure separating class. Exactly this will be implemented below.

In the case of measures on locally convex spaces differentiable along constant vectors, this construction is close to the ones described in [9], [10], [11] and later developed in [58]; however, in our case it requires only one-fold differentiability of F. In [58] the membership of F in the second Sobolev class is required and in [29], in the Gaussian case, also the second derivative is used (the function F is in the first Sobolev class, but its appropriately scaled Malliavin gradient must be also in the first Sobolev class).

The following result from [18] gives broad sufficient conditions for the existence of surface measures and describes their connections with conditional measures.

Theorem 2.4.7. *Let μ be Fomin differentiable along v with the logarithmic derivative β_v. Suppose that* (F1) *holds, a function $F\colon X \to \mathbb{R}$ satisfies* (F2) *and $\mu \circ F^{-1}$ has no atoms* (*i.e.*, $\mu(F^{-1}(y)) = 0$ *for all* y). *Assume also that at least one of the following conditions holds*:

(i) X *is a complete metric space and* $\mathcal{F}$ *contains all bounded Lipschitzian functions*;

(ii) *the measure* μ *has compact support*;

(iii) *there exists a nonnegative function* $W \in \mathfrak{D}_v$ *such that* $W\beta_v$, $W\partial_v F \in L^1(\mu)$ *and the sets* $\{W \leq R\}$ *are compact for all* $R \geq 0$.

Then, for each $y \in \mathbb{R}$, *the Radon surface measure* σ^y *associated with* $F^{-1}(y)$ *exists.*

In addition, if μ *is concentrated on a countable union of metrizable compact sets, then, for* $\nu \circ F^{-1}$-*a.e.* y, *where* $\nu = (\partial_v F) \cdot \mu$, *the surface measure* σ^y *is concentrated on* $F^{-1}(y)$ *and we have the equality*

$$\sigma^y = \varrho_1(y) \cdot \nu^y,$$

where ϱ_1 *is the density of* $\nu \circ F^{-1}$ *and* $\{\nu^y\}$ *is the system of conditional measures for* ν.

It follows that σ^y is absolutely continuous also with respect to the conditional measure μ^y for μ.

The proof will be given below after a number of auxiliary results. However, we can say right now that in all these cases the measure σ^y will be obtained as the limit of the measures $r^{-1}I_{\{y<F<y+r\}} \cdot \nu$ in the week topology; in cases (i) and (ii) this will be an immediate corollary of our assumptions and in case (iii) some little extra work will be needed.

The main point is that, under the assumptions of the theorem, for every $f \in \mathcal{F}$ the function $\Phi_f(y)$ is continuously differentiable. This can be explained immediately in the case $F \in \mathcal{F}$. The function $\Phi_f(y)$ is the distribution function of the bounded measure

$$m_f := (f \cdot \nu) \circ F^{-1}$$

on the real line. Therefore, it suffices to show that this measure has a continuous density ϱ_f with respect to Lebesgue measure. This will be done if we show that the derivative of m_f in the sense of generalized functions is a bounded measure η_f without points of nonzero measure. Using the standard reasoning in the Malliavin calculus, we now show that

$$\eta_f = (f \cdot d_v\mu + \partial_v f \cdot \mu) \circ F^{-1}$$

is the generalized derivative of m_f. Let $\psi \in C_0^\infty(\mathbb{R})$. We have

$$\begin{aligned}\int \psi'(t)\, m_f(dt) &= \int \psi'(t)\, (f \cdot \nu) \circ F^{-1}(dt) = \int_X \partial_v(\psi(F))(x) f(x)\, \mu(dx)\\ &= \int_X \partial_v(f\psi(F))(x)\, \mu(dx) - \int_X \psi(F(x))\partial_v f(x)\, \mu(dx)\\ &= -\int_X \psi(F(x)) f(x)\, d_v\mu(dx) - \int_X \psi(F(x))\partial_v f(x)\, \mu(dx) = -\int \psi(t)\, \eta_f(dt). \end{aligned} \qquad (2.13)$$

Finally, η_f has no points of nonzero measure if this is true for $\mu \circ F^{-1}$ and $d_v\mu \ll \mu$ (the latter is true in the case of Fomin differentiability). The proof below is similar, we just need to extend (2.13) to more general functions F satisfying (F2).

Remark 2.4.8. It is worth noting that under our assumptions (F1) and (F2) the measure $\nu \circ F^{-1}$ is absolutely continuous (see below), hence $\mu \circ F^{-1}$ is also absolutely continuous provided that $\partial_v F > 0$ μ-a.e. (then μ and ν are equivalent). In particular, in the latter case $\mu \circ F^{-1}$ has no atoms. In general, of course, this is not true, since F can be constant on a positive measure set.

Remark 2.4.9. It will be clear from the proof that if we are interested only in surface measures on $F^{-1}(y)$ for y in some interval I it suffices to have that $\mu(F^{-1}(y)) = 0$ only for $y \in I$. In addition, in many cases the construction can be "localized" by replacing μ with the measure $f \cdot \mu$, where $f \in \mathcal{F}$ has an appropriate support. In this way, one can make assumptions about also local ν, say, replacing ν by $f \cdot \nu$ (note that the surface measures in [11] are constructed locally).

The following classical concepts and facts are crucial for the proof of Theorem 2.4.7.

A sequence of Radon measures μ_n converges weakly to a Radon measure μ if, for each bounded continuous function f, we have

$$\int_X f(x)\,\mu(dx) = \lim_{n\to\infty} \int_X f(x)\,\mu_n(dx).$$

By Aleksandrov's theorem, weak convergence of a sequence of Radon probability measures to a Radon probability measure μ is equivalent to the relation $\mu(W) \le \liminf_{n\to\infty} \mu_n(W)$ for every open set W (see [12, Section 8.2]).

By LeCam's theorem (see [12, Corollary 8.6.3]), ofr complete metric spaces, if a sequence of nonnegative Radon measures μ_n is such that the integrals of each bounded Lipschitzian function with respect to these measures converge then this sequence converges weakly to some Radon measure.

Finally, it follows from Prohorov's theorem (see [12, Section 8.6]) that if a sequence of Radon measures ν_n on X is uniformly bounded in variation and uniformly tight, i.e., for every $\varepsilon > 0$ there is a compact set K_ε such that $|\mu_n|(X \backslash K_\varepsilon) < \varepsilon$ for all n, and there is a class of bounded Borel functions on X separating Radon measures such that the integrals of f against μ_n converge for each f in this class, then the measures μ_n converge weakly to some Radon measure μ on X (Prohorov's theorem ensures the existence of a Radon measure μ that is a limit point of $\{\mu_n\}$ in the weak topology and the second condition says that this limit point is unique, hence the sequence converges weakly to it).

Note that in order to have the uniform tightness of nonnegative Radon measures ν_n it suffices to have a nonnegative Borel function W on X such that the sets $\{W \le R\}$ are compact for all $R \ge 0$ and the integrals of W with respect to the measures ν_n are uniformly bounded by some number C. In that case, by the Chebyshev inequality we have

$$\nu_n(X\backslash\{W \le R\}) \le CR^{-1}.$$

Lemma 2.4.10. *Let μ be Skorohod differentiable along v. Then*

(i) *for every $f \in \mathfrak{D}_v$, the measure $(\partial_v f \cdot \mu) \circ f^{-1}$ is absolutely continuous and its distributional derivative is $(f \cdot d_v\mu) \circ f^{-1}$, hence for all t we have*

$$\int_{\{f<t\}} \partial_v f(x)\,\mu(dx) = \int_{-\infty}^{t} \int_{\{f<s\}} f(x)\,d_v\mu(dx)\,ds. \tag{2.14}$$

(ii) *$\phi(f) \in \mathfrak{D}_v$ for each Lipschitzian function ϕ and each $f \in \mathcal{F}$.*

In addition, $\phi(f) \in \mathfrak{D}_v$ for any continuously differentiable function ϕ with a bounded derivative and any $f \in \mathfrak{D}_v$. In both cases, $\partial_v(\phi \circ f) = \phi'(f)\partial_v f$ μ-a.e.

For a proof, see [18].

Lemma 2.4.11. *We have $\psi(F) \in \mathfrak{D}_v$ for each bounded Lipschitzian function ψ on the real line. In addition, $\partial_v(\psi(F)) = \psi'(F)\partial_v F$.*

The proof of this lemma is easy and can be also found in [18].

Corollary 2.4.12. *Under assumptions* (F1) *and* (F2) *we have*

$$(\nu \circ F^{-1})' = \eta = d_v\mu \circ F^{-1}$$

in the sense of distributions, where ν and η are defined by (2.10). *Hence the measure $\nu \circ F^{-1}$ has a density ϱ_1 of bounded variation, moreover,*

$$\varrho_1(t) = \eta((-\infty, t)) = d_v\mu(x\colon\ F(x) < t).$$

If $\mu \circ F^{-1}$ has no atoms and μ is Fomin differentiable along v, then $|d_v\mu|(\{F = t\}) = 0$ for every t, hence this density is continuous.

Moreover, for every $f \in \mathcal{F}$ we have

$$((f \cdot \nu) \circ F^{-1})' = (f \cdot d_v\mu) \circ F^{-1} + (\partial_v f \cdot \mu) \circ F^{-1} \tag{2.15}$$

and

$$\|((f \cdot \nu) \circ F^{-1})'\| \le \|f \cdot d_v\mu + \partial_v f \cdot \mu\|.$$

If $\mu \circ F^{-1}$ has no atoms then the measure $(f \cdot \nu) \circ F^{-1}$ has a continuous density ϱ_f of bounded variation and

$$|\varrho_f(y)| \le \|f \cdot d_v\mu + \partial_v f \cdot \mu\| \le \|d_v\mu\| \cdot \|f\|_\infty + \|\partial_v f\|_{L^1(\mu)}.$$

Finally, if $d_v\mu = \beta_v \cdot \mu$, where $\beta_v \in L^q(\mu)$, $q = p/(p-1)$, then

$$|\varrho_f(y)| \le \|\beta_v\|_{L^q(\mu)}\|f\|_{L^p(\mu)} + \|\partial_v f\|_{L^1(\mu)}. \tag{2.16}$$

Proof. Extending (2.13), we find that

$$\int \psi'(t)\,(f\cdot v)\circ F^{-1}(dt) = \int_X \partial_v(\psi\circ F)(x)f(x)\,\mu(dx)$$
$$= -\int_X \psi(F(x))f(x)\,d_v\mu(dx) - \int_X \psi(F(x))\partial_v f(x)\,\mu(dx),$$

which gives (2.15), hence all our assertions follow. □

Remark 2.4.13. It should be noted that the "non-normalized" surface measures introduced above are still not true "surface measures"; they depend not only on the level sets $F^{-1}(y)$ but also on the whole function F. Obviously, the whole measure depends also on our choice of the vector field v. However, there is some scaling invariance of the construction: for example, if we replace F by kF with some number $k > 0$, the set $F^{-1}(0)$ does not change and our measure σ^0 respects this. The sets $\{0 < kF < r\}$ are the old sets $\{0 < F < r/k\}$, so, when evaluating the derivative of the distribution function at zero, we have the factor k coming from $\partial_v(kF)$ and obtain the same quantity.

Nevertheless, if v also depends on F, e.g., if we take for v a suitable gradient of F without normalization, then we loose this invariance. This is a certain disadvantage of our definition which will be partially overcome below (by passing to surface measures normalized by weights or by taking normalized vector fields). One should bear in mind that, even dealing with very nice functions F on infinite-dimensional spaces, the known constructions do not really define surface measures on individual level sets $F^{-1}(y)$. As it happens with usual nice surfaces in $\mathbb{R}^d$, it is still necessary that each fixed surface be included in a special family of level sets. An important exception is the Gaussian Hausdorff measure mentioned in Section 3; as we have seen this is a non-σ-finite measure on all Borel sets and it is not easy to calculate its value on individual surfaces. Another exception is a surface determined by a nondegenerate Fréchet differentiable function on a Banach space (but typical infinitely Sobolev differentiable functions on infinite-dimensional spaces are not even continuous). On the other hand, by using weight functions one can obtain "geometric surface measures" on the basis of our surface measures for a reasonable individual surface.

Proof of Theorem 2.4.7. We can assume that $y = 0$. We know from Corollary 2.4.12 that for every $f \in \mathcal{F}$ the distribution function of the measure $(f\cdot v)\circ F^{-1}$ is differentiable at zero and its derivative is $\varrho_f(0)$. Clearly,

$$\varrho_f(0) = \lim_{n\to\infty} n\int_{B_n} f(x)\,v(dx) = \lim_{n\to\infty}\int_X f(x)\,v_n(dx),$$

where

$$B_n = \{0 < F < n^{-1}\} \quad \text{and} \quad v_n := nI_{B_n}\cdot v.$$

Certainly, in place of n^{-1} we can take numbers $h_n > 0$ decreasing to zero (then the factor n is replaced by h_n^{-1}). The nonnegative measures ν_n are uniformly bounded since their values on the whole space X converge to $\varrho_1(0)$.

If either (i), (ii) or (iii) is fulfilled, it follows that there is a bounded nonnegative Radon measure σ^0 on X such that

$$\int_X f(x)\,\sigma^0(dx) = \lim_{n\to\infty} \int_X f(x)\,\nu_n(dx).$$

Indeed, in case (i) we apply Le Cam's theorem. Note that the measures ν_n are Radon and are concentrated on a common separable subspace. In case (ii) we obviously have the uniform tightness of the measures ν_n, which gives a Radon limit, as explained above.

The same is also true in case (iii) because the integrals

$$\int_X W(x)\,\nu_n(dx) = n \int_{B_n} W(x)\,\nu(dx)$$

are uniformly bounded. This follows by the same reasoning that proves the existence of $\varrho_1(0)$, just in place of μ we take $W \cdot \mu$; our assumptions in (iii) are such that this works. This completes the proof of Theorem 2.4.7.

Let us compare the constructed measures σ^y with the conditional measures ν^y, assuming that μ is concentrated on a countable union of metrizable compact sets. It follows from the definition of $\varrho_f(y)$ that

$$\int_{-\infty}^{+\infty} \int_X f(x)\,\sigma^y(dx)\,dy = \int_X f(x)\,\nu(dx) = \int_{\mathbb{R}} \int_X f(x)\,\nu^y(dx)\,\nu\circ F^{-1}(dy).$$

The integral on the left can be written as

$$\int_{-\infty}^{+\infty} \int_X f(x)\frac{1}{\varrho_1(y)}\,\sigma^y(dx)\,\varrho_1(y)dy = \int_{\mathbb{R}} \int_X f(x)\frac{1}{\varrho_1(y)}\,\sigma^y(dx)\,\nu\circ F^{-1}(dy).$$

Hence the measure $\sigma^y/\varrho_1(y)$ coincides with the conditional measure ν^y for $\nu\circ F^{-1}$-a.e. y due to our assumption that $\mathcal{F}$ separates measures on $\mathcal{B}$, and the essential uniqueness of conditional measures. □

Remark 2.4.14. (i) The assumption that μ is concentrated on metrizable compact sets has not been used for the proof of existence of σ^y; it is only needed if we wish to compare surface measures with conditional measures and localize σ^y on $F^{-1}(y)$.

(ii) In place of the sets $\{y < F < y+h\}$ we could deal with the sets $\{y-h < F < y+h\}$, but then the factor h^{-1} must be replaced by $(2h)^{-1}$.

(iii) It follows from our construction that in cases (i) – (iii) the mapping $y \mapsto \sigma^y$ is continuous provided that the space of probability measures is equipped with the weak topology. Indeed, according to (2.12), whenever $y_j \to y$, for each $f \in \mathcal{F}$ the integral of f against σ^{y_j} converges to the integral of f against σ^y; by the respective assumption this yields weak convergence.

(iv) Although our construction of surface measures is topological in the sense that it involves weak convergence of measures and the latter depends on the initial topology, the resulting measures possess certain topological invariance. Obviously, they do not change if we continuously embed our space into a larger space. Moreover, in many practical situation they do not change even if we consider them on a smaller topological space of full measure continuously embedded into the original space in such a way that the restriction of our measure μ to this smaller space is Radon on it.

(v) Surface measures are closely related to the classes BV of functions of bounded variation. In $\mathbb{R}^n$ the functions of this class are precisely those functions $f \in L^1(\mathbb{R}^n)$ for which the derivatives $\partial_{x_i} f$ in the sense of distributions are measures of bounded variation. The indicator function I_V of a domain belongs to BV precisely when the boundary ∂V of V has finite perimeter. For the infinite-dimensional case, see [2]–[8], [20], [21], [22], [24], [25], [43], [44], [48], [51], [62], and [63]. An interesting question is to determine when level sets of a convex function (for example, a seminorm) have finite surface measures. This question is not trivial even for Gaussian measures and becomes especially challenging for convex (or logarithmically concave) measures, i.e., probability measures whose finite-dimensional projections are measures given by densities of the form e^{-V} with a convex function V with respect to Lebesgue measures on affine subspaces. One of the problems with such measures in infinite dimensions is that it is not known whether they always have vectors of Skorohod differentiability. Returning to the Gaussian case, recall that if γ is a centered Gaussian measure and A is a Borel set of measure $\gamma(A) \geq 1/2$ and $\Phi(a) = \gamma(A)$, where Φ is the standard Gaussian distribution function, then the following isoperimetric inequality holds (see, e.g., [11]):

$$\gamma(A + tU_H) \geq \Phi(a + t) \quad \forall\, t \geq 0,$$

where U_H is the closed unit ball in the Cameron–Martin space H. Therefore,

$$\frac{\gamma(A + tU_H) - \gamma(A)}{t} \geq \frac{\Phi(a + t) - \Phi(t)}{t}.$$

Hence letting $t \to 0$ we conclude that the liminf of the left-hand side is at least $\Phi'(a) = \varrho(a)$, which is the surface measure of the half-space of γ-measure a. Therefore, half-spaces possess minimal surface measures among sets of γ-measure a.

2.5 Fine versions of surface measures controlled by capacities

The surface measures constructed above on the basis of the Malliavin calculus have the property that they are defined on "almost all" level sets similarly to conditional measures. In this section we give some additional conditions under which there is a more canonical version of σ^y sitting on $F^{-1}(y)$ for each y. Here we assume that μ is Fomin differentiable along v (also some higher integrability of β_v will be assumed). Note that the measures σ^y do not change if we take a different version of F, but the sets $F^{-1}(y)$ can change. We recall that, as noted above, there is no problem if F is continuous and for every point z in the complement of $F^{-1}(y)$ and every neighborhood U of z there is a nonnegative continuous function of class $\mathcal{F}$ positive at z and having support in U.

Another concept coming along with surface measures is capacity (see [11] or [13]). Suppose that $\mathcal{F}$ is equipped with a norm $\|\cdot\|_{\mathcal{F}}$ such that convergence in this norm yields convergence in $L^1(\mu)$. In practical situations, this will be often the norm of a suitable Sobolev space $W^{p,1}(\mu)$, but so far no Sobolev spaces are needed. This norm generates a capacity: for every open set $U \subset X$ we define its capacity associated with $\mathcal{F}$ by the formula

$$C_{\mathcal{F}}(U) = \inf\{\|f\|_{\mathcal{F}} \colon\ f \in \mathcal{F},\ f \geq 0,\ f \geq 1\ \ \mu\text{-a.e. on } U\}.$$

For any set $B \subset X$ let

$$C_{\mathcal{F}}(B) = \inf\{C_{\mathcal{F}}(U) \colon\ U \supset B \text{ is open}\}.$$

Typically, capacities of the sort we consider are tight (see [61], [59], and [60]), i.e., for each $\varepsilon > 0$ there is a compact set K_ε such that $C_{\mathcal{F}}(X\backslash K_\varepsilon) < \varepsilon$. However, we do not assume this property.

Recall that a function f is called $C_{\mathcal{F}}$-quasi-continuous if, for each n, there is a closed set A_n such that $C_{\mathcal{F}}(X\backslash A_n) < 1/n$ and $f|_{A_n}$ is continuous.

It is known that each function $f \in \mathcal{F}$ has a $C_{\mathcal{F}}$-quasi-continuous version (see [13, Section 8.13]), provided that the norm $\|\cdot\|_{\mathcal{F}}$ is strictly convex. This is the case with the L^p-norm with $p \in (1, +\infty)$, and, more generally, with any norm of the form $\|f\|_{\mathcal{F}} = \|T^{-1}f\|_{L^p(m)}$, where m is a probability measure and T is a bounded injective linear operator from $L^p(m)$ to $L^1(\mu)$; in particular, the latter case covers most of Sobolev classes such as $W^{p,1}(\gamma)$ for a Gaussian measure γ (see Section 3). However, in place of such assumptions we simply assume in addition to (F1) and (F2) that

(F3) *F has a quasi-continuous version.*

We now fix a quasi-continuous version of F; the results below refer to this version.

Lemma 2.5.1. *Let μ be Fomin differentiable along v with respect to $\mathcal{F}$ and let* (F1), (F2) *and* (F3) *hold. Suppose that there is $p > 1$ such that*

$$\|f\|_{L^p(\mu)} + \|\partial_v f\|_{L^1(\mu)} \leq \|f\|_{\mathcal{F}}, \quad f \in \mathcal{F}. \tag{2.17}$$

Assume also that $\beta_v \in L^{p/(p-1)}(\mu)$. Then, for every open set $W \subset X$ and any $r > 0$, the measure $\nu = (\partial_v F) \cdot \mu$ satisfies the estimate

$$\nu(W \cap \{y < F < y + r\}) \leq rC(\mu)C_{\mathcal{F}}(W), \quad C(\mu) = 1 + \|\beta_v\|_{L^q(\mu)},\ q = \frac{p}{p-1}. \tag{2.18}$$

Proof. Let $f \in \mathcal{F}$, $f \geq 0$ and $f \geq 1$ μ-a.e. on W. Then $f \geq 1$ ν-a.e. on W, hence on account of (2.16) and (2.17) we obtain

$$\nu(W \cap \{y < F < y + r\}) \leq \int\limits_{W\cap\{y<F<y+r\}} f(x)\,\nu(dx)$$
$$\leq \int\limits_{y<F<y+r} f(x)\,\nu(dx) \leq rC(\mu)(\|f\|_{L^p(\mu)} + \|\partial_\nu f\|_{L^1(\mu)}) \leq rC(\mu)\|f\|_{\mathcal{F}},$$

which yields the announced estimate by taking inf in f. □

Certainly, we can always equip $\mathcal{F}$ with the norm given by the left-hand side of (2.17). Moreover, this norm is strictly convex (since so is the L^p-norm) and convergence in this norm obviously yields convergence in $L^p(\mu)$, hence in $L^1(\mu)$. However, in concrete examples there might be other natural norms on $\mathcal{F}$ not related to ν, e.g., certain Sobolev norms. Quasi-continuous versions of F depend on our choice of $C_{\mathcal{F}}$, hence on our choice of a norm on $\mathcal{F}$.

Theorem 2.5.2. *Suppose that in Theorem* 2.4.7 *we have $\beta_\nu \in L^{p/(p-1)}(\mu)$ for some $p > 1$ and that* (2.17) *and* (F3) *hold* (*which can be ensured by taking the norm on $\mathcal{F}$ defined by the left-hand side of* (2.17)). *Then each σ^y is concentrated on the set $F^{-1}(y)$ and vanishes on all sets of $C_{\mathcal{F}}$-capacity zero.*

Proof. Let us show that $\sigma^y(X\backslash F^{-1}(y)) = 0$. We can assume again that $y = 0$. It suffices to show that σ^0 vanishes on each set $U := \{|F| > \delta\}$, where $\delta > 0$. By assumption, for each n, there is a closed set A_n such that $C_{\mathcal{F}}(X\backslash A_n) < 1/n$ and $F|_{A_n}$ is continuous. The sets

$$U_n = U \cap (X\backslash A_n)$$

are open because $\{|F| \leq \delta\} \cap A_n$ is closed by the continuity of the restriction $F|_{A_n}$. We have $U \subset \bigcap_{n=1}^\infty U_n$. Let $k > 1/\delta$. Then $\nu_k(U) = 0$, where, as above, $\nu_k = kI_{\{0<F<k^{-1}\}} \cdot \nu$. By the lemma we have

$$\nu_k(U_n) = \nu_k(X\backslash A_n) \leq C(\mu)n^{-1},$$

hence $\sigma^0(U_n) \leq C(\mu)n^{-1}$, which yields that $\sigma^0(U) = 0$. Note that we could not derive this directly from the equality $\nu_k(U) = 0$, because U need not be open.

We now prove that $\sigma^y(B) = 0$ for every set $B \in \mathcal{B}$ of zero $C_{\mathcal{F}}$-capacity. Again it suffices to consider the case $y = 0$. Let $\varepsilon > 0$. By definition, there is an open set W containing B such that $C_{\mathcal{F}}(W) < \varepsilon$. Therefore, there is a function $f \in \mathcal{F} \geq 0$ such that $f \geq 1$ μ-a.e. on W and $\|f\|_{\mathcal{F}} < \varepsilon$. It follows from the lemma that $|\nu_n(W)| \leq \varepsilon C(\mu)$, which yields that $|\sigma^0(W)| \leq \varepsilon C(\mu)$. Letting $\varepsilon \to 0$ we arrive at the desired conclusion. □

In the framework described above there is no natural way of normalizing our surface measures. One way of making the construction more invariant is this: assuming that there is some intrinsic norm $|v|_H$ (as is the case for Gaussian measures when we use the norm of the Cameron–Martin space H) and $|v(x)|_H > 0$ μ-a.e., one could use the unit field $v/|v|_H$, which leads to the weight $\partial_v F/|v|_H$ in place of $\partial_v F$. However, to avoid possible problems with differentiation along this new field, we just assume that the measure $|v|_H^{-1} \cdot \nu$ is finite and take new measures

$$\sigma_0^y := |v|_H^{-1} \cdot \nu^y$$

on the same level sets $F^{-1}(y)$. These measures are finite for $\nu \circ F^{-1}$-a.e. y, hence also for $\mu \circ F^{-1}$-a.e. y. Finally, if no weight is used, we arrive at surface measures that coincide with conditional measures.

Remark 2.5.3. Note that for any bounded continuous function g the measure $g \cdot \sigma^y$ is naturally defined for every y (not for almost every y). It is readily seen from Theorem 2.5.2 that the same is true for any bounded quasi-continuous function g. In order to assign finite integrals with respect to all surface measures σ^y to some unbounded functions g one can use the following trick: use surface measures generated by the new measure $g \cdot \mu$. Obviously, this requires some additional assumptions about g because our approach is based on positive differentiable measures. However, it works if there is $p > 1$ such that $\partial_v f \in L^p(\mu)$ for all $f \in \mathcal{F}$, $\beta_v \in L^p(\mu)$, $\partial_v F \in L^p(\mu)$, $g \geq 0$ belongs to $\mathfrak{D}_v \cap L^{p'}(\mu)$, $\partial_v g \in L^{p'}(\mu)$, where $p' = p/(p-1)$. Using this trick separately for g^+ and g^-, one can extend this to certain functions of variable sign. One can show that if in this situation g is bounded continuous this procedure yields the usual products $g \cdot \sigma^y$.

It is also worth noting that if we apply the same construction to the original measure μ in place of ν in order to integrate by parts in the equality with $\psi'(F)$ we must artificially add the factor $\partial_v F$ to obtain the expression $\partial_v(\psi(F))$. The effect is that we must impose the assumption of differentiability not on μ, but on the measure $(\partial_v F)^{-1} \cdot \mu$. In principle, this is quite possible but requires some extra assumptions.

In the considered situation we have the following version of the Gauss–Ostrogradskii–Stokes formula with our non-normalized surface measure. Set

$$V_r = F^{-1}(-\infty, r), \quad S_r = F^{-1}(r).$$

Theorem 2.5.4. *Let u be another vector field along which μ is differentiable, satisfying the same hypotheses as v. Then*

$$\int_{V_r} \beta_u(x)\,\mu(dx) = -\int_{S_r} \frac{\partial_u F(x)}{\partial_v F(x)}\,\sigma^r(dx),$$

provided that either the function $\xi := \partial_u F/\partial_v F$ is bounded quasi-continuous or ξ^+ and ξ^- satisfy the additional conditions mentioned in the previous remark.

Proof. Let $\psi_h(s) = 1$ if $s \leq r$, $\psi_h(s) = 0$ if $s \geq r + h$, $\psi_h(s) = C - s/h$ if $r < s < r + h$, $C = 1 + r/h$. Then $\psi_h'(s) = -1/h$ in the interval $(r, r + h)$ and $\psi_h' = 0$ outside the closure of this interval. We have $\partial_u(\psi_h \circ F) = -h^{-1}\partial_u F$ on the set $\{r < F < r + h\}$ and

$$\int_X \psi_h(F(x))\beta_u(x)\,\mu(dx) = -\int_X \partial_u(\psi_h \circ F)(x)\,\mu(dx) = h^{-1} \int_{r<F<r+h} \partial_u F(x)\,\mu(dx).$$

As $h \to 0$, the left-hand side of this identity tends to the integral of β_u over V_r and the right-hand side tends to the surface integral of the function $\partial_u F/\partial_v F$ against the surface measure σ^r. The latter holds if either ξ is bounded quasi-continuous or there are surface measures associated with the new measures $\xi^+ \cdot \mu$ and $\xi^- \cdot \mu$. □

If $u = v$, then we have

$$\int_{V_r} \beta_v(x)\,\mu(dx) = -\int_{S_r} \sigma^r(dx),$$

which gives the total mass of the surface.

2.6 Examples and comments

It is clear from the comments above that all our assumptions are rather general except for the requirement of differentiability of μ along a suitable vector field. Actually, also vector fields of differentiability can be found for quite general measures (see [13, Chapter 11]). The only serious restriction arises if we wish to find this field in a such a way that $\partial_v F$ is not very degenerate in order to connect our surface measures with more traditional surface measures as explained above. For this reason we include $\partial_v F$ in our measure. In particular, identically zero v fits our construction pretty well and produces zero surface measures. In order to avoid such meaningless situations we now consider some examples where for a given F one can find a suitable v with $\partial_v F > 0$ μ-a.e.

Example 2.6.1. Suppose that X is a Banach space, μ is Fomin differentiable along a nonzero constant vector v, and F is a continuous function on X differentiable along v such that $\partial_v F$ is continuous and $c_1 \le \partial_v F \le c_2$ for some positive numbers c_1 and c_2. Then there exist surface measures σ^y on the level sets $F^{-1}(y)$. In addition, one can use equivalent "traditional" surface measures $|\partial_v F|^{-2} \cdot \sigma^y$. One can also use a local version of this construction multiplying μ by a Lipschitzian bump function with a small support in a neighborhood of a point x_0 where $\partial_v F(x_0) > 0$.

In this situation we can apply both (i) and (iii) in Theorem 2.4.7. Applicability of (i) follows from the fact that any Lipschitzian function on X is μ-a.e. differentiable along v, which in turn follows from the one-dimensional case and the existence of differentiable conditional measures on the straight lines $x + \mathbb{R}v$ (see [13, Chapter 3]). Case (iii) applies here if we take for $\mathcal{F}$ the same class of bounded Lipschitzian functions or the class $\mathcal{FC}$ of smooth cylindrical functions. Finally, for W we can take a function of the form $W(x) = w(\|x\|)$, where w is an unbounded Lipschitzian sufficiently slowly increasing function on $\mathbb{R}$ such that W is μ-integrable (one can always find such a function).

An obvious disadvantage of constant vector fields is that they give less chances to obtain positive $\partial_v F$. For example, if X is a Hilbert space and F is Gâteaux differentiable, then it would be optimal in this respect to take $v = \nabla F$, which gives $\partial_v F(x) = \|\nabla F(x)\|^2$. However, even for very nice functions F there might be no natural measures differentiable along ∇F. For example, if we take $F(x) = (x, x)$ and want to define surface measures on the spheres, we have to ensure differentiability of μ along the field $v(x) = x$. However, say, Gaussian measures on infinite-dimensional spaces are not differentiable along this field (see the example below). For this reason, one has to consider vector fields with values in suitable analogs of the Cameron–Martin space.

Example 2.6.2. Let γ be a centered Radon Gaussian measure on a locally convex space X such that its Cameron–Martin space H is infinite-dimensional. Then γ is not Fomin differentiable along the vector field $v(x) = x$. Indeed, it suffices to prove this for the standard Gaussian measure γ on $\mathbb{R}^\infty$. Suppose that $\beta_v \in L^1(\gamma)$ is the divergence of v. Then for every smooth cylindrical function f in variables $x_1, \dots, x_n$ we have

$$\int_X \sum_{i=1}^n x_i \partial_{x_i} f(x)\,\gamma(dx) = -\int_X f(x)\beta_v(x)\,\gamma(dx).$$

The left-hand side equals

$$-\int_X f(x) \sum_{i=1}^n (1 - x_i^2)\,\gamma(dx)$$

by the integration by parts formula. Therefore, the function $S_n(x) = \sum_{i=1}^n (1 - x_i^2)$ equals the conditional expectation of the function β_v with respect to the σ-field generated by $x_1, \dots, x_n$. We recall that the conditional expectation of an integrable function ξ with respect to a smaller σ-algebra $\mathcal{A}$ is an integrable function $\mathrm{E}^{\mathcal{A}}\xi$ that is measurable with respect to $\mathcal{A}$ and satisfies the identity

$$\int_X \eta\xi\, d\gamma = \int_X \eta \mathrm{E}^{\mathcal{A}}\xi\, d\gamma$$

for all bounded functions η measurable with respect to $\mathcal{A}$. By the martingale convergence theorem (see, e.g., [12, Chapter 10]) the sequence of functions S_n converges in $L^1(\gamma)$, but the functions $1 - x_i^2$ are second order polynomials, hence the series of $1 - x_i^2$ converges in all $L^p(\gamma)$ (see [11, Chapter 5]). However, there is no convergence in $L^2(\gamma)$, because these functions are mutually orthogonal in $L^2(\gamma)$ and have equal norms.

Example 2.6.3. Let us return to Section 3 and consider the case of a centered Gaussian measure μ with the Cameron–Martin space H and F belonging to the Sobolev class $W^{2,2}(\mu)$. In this case the vector field $v = D_H F \in W^{2,1}(\mu, H)$ has divergence $\beta_v = LF \in L^2(\mu)$ and $\partial_v F = |D_H F|_H^2$ is in $L^1(\mu)$.

If $F \in W^{p,2}(\mu)$ with some $p \in (1, 2)$, then we can use the vector field

$$v = D_H F/|D_H F|_H^2,$$

for which $\partial_v F = 1$ (and we obtain surface measures from [29]), or the vector field

$$v = D_H F/|D_H F|_H,$$

so that v is a unit normal on the surface (but with respect to the Cameron–Martin norm) and $\partial_v F = |D_H F|_H$; now in both cases it becomes necessary to require that v must have divergence. In the first case it suffices to have $\|D_H^2 F\|_{HS}/|D_H F|_H^2 \in L^p(\mu)$ and in the second case it suffices to have $\|D_H^2 F\|_{HS}/|D_H F|_H \in L^p(\mu)$ (the corresponding surface measures will be different). It is also possible to use a less constructive (but weaker) assumption that $F \in W^{1,1}(\mu)$ and one of these two vector fields has divergence $\beta_v \in L^1(\mu)$.

If X is a Banach space, then we can take for $\mathcal{F}$ the class of all bounded Lipschitzian functions. Conditions (F1)–(F3) are readily verified in this case. Actually, the case of a general locally convex space with a Radon Gaussian measure reduces to this one by the Tsirelson linear isomorphism theorem (see [11, Chapter 3]). If $\beta_v \in L^p(\mu)$, then we can take $\mathcal{F} = \mathcal{FC}$ (smooth cylindrical functions) and apply case (iii) in Theorem 2.4.7. If X is sequentially complete, then for W we can take the Minkowski functional p_Q of an absolutely convex compact set Q of positive measure. It is known that $p_Q \in W^{r,1}(\mu)$ for all $r \geq 1$.

Actually, if one is interested only in a Gaussian measure μ on a Banach space, then the construction of surface measures along these lines becomes straightforward as explained in Section 3.

In particular, if $D_H F \neq 0$ a.e., our construction can be compared with surface measures considered in [1], [49], [11], [13], and [29] (note that the latter work develops a construction based on distribution functions related to the measure μ itself, so that it leads to surface measures that are not invariant under scaling of F, as discussed above). In order to come to usual surface measures one should either deal with a unit field $D_H F/|D_H F|_H$ or deal with $v = D_H F$, and then multiply the obtained surface measures by $1/|D_H F|_H$.

We emphasize that for better surfaces (existing individually such as level sets of continuously Fréchet differentiable functions with nondegenerate derivatives) there is no need to involve variable vector fields $D_H F$: it becomes much simpler to define surface measures locally by using only constant vector fields of differentiability of μ as in Example 2.6.1. In that case no second derivatives of F appear at all and in this way we recover the existence results of [75] (even under weaker assumptions).

In place of a Gaussian measure μ, it is possible to consider a Radon probability measure μ on a locally convex space X that is Fomin differentiable along a continuously embedded dense Hilbert space H. A simple example is the countable power of a probability measure on $\mathbb{R}$ with a smooth density with compact support; then one can take $X = \mathbb{R}^\infty$ and $H = l^2$ is a natural choice (the same measure can be also considered on a weighted Hilbert space of sequences (x_n) such that $\sum_{n=1}^\infty \alpha_n x_n^2 < \infty$, where $\alpha_n > 0$ and $\sum_{n=1}^\infty \alpha_n < \infty$). Then one can also define Sobolev classes. This situation (studied in [58], [59], and [60]) has been the most general considered so far in the linear case. In [29], similar results have been reproved in the Gaussian case.

Remark 2.6.4. (i) It is worth noting that the case of a Fréchet space reduces to that of a separable reflexive Banach space, since every Radon measure on a Fréchet space is concentrated on a compactly embedded separable reflexive space (see [12, Theorem 7.12.4]). For many measures on Banach spaces (Gaussian, differentiable), the class of bounded Lipschitzian functions is a suitable candidate for $\mathcal{F}$, since such functions are almost everywhere differentiable with respect to such measures.

(ii) The class $\mathcal{FC}$ of smooth cylindrical functions and the larger class of bounded Lipschitzian cylindrical functions satisfy condition (F1), but in general they do not have the property of the whole class of bounded Lipschitzian functions that convergence of integrals of such functions with respect to a sequence of probability measures ensures weak convergence of these measures (say, this is not true for infinite-dimensional Hilbert spaces, although is true for $\mathbb{R}^\infty$). This is why we considered cases (ii) and (iii) in Theorem 2.4.7. As already noted, it is possible to define surface measures locally in a suitable sense (for example, on compact sets) by replacing μ by $\zeta \cdot \mu$, where $\zeta \geq 0$ is a bump function whose support gives the desired localization. For example, in the Gaussian case or in the case of a differentiable measure on a Banach space, it is always possible to choose ζ in a such a way that its support will be compact and will contain a given compact set, and the measure $\zeta \cdot \mu$ will remain Fomin differentiable along the same directions as μ. This approach can give local surface measures in more general situations where there are no global surface measures. A possible way of gluing these local surface measures is based on establishing their uniform tightness.

If X is equipped with a suitable tangent structure enabling us to consider v not as a differentiation, but as a true vector field possessing the corresponding norm $|v(x)|$, one might try to use fields of unit length; again the question of their choice arises.

The choice $v = D_H F$ in the Gaussian case mentioned above is connected with another natural object related to Gaussian Hausdorff measures mentioned at the end of Section 3: H-neighborhoods of sets. Given a Borel set B, we take the set $B^r = B + rU_H$, where U_H is the unit ball in the Cameron–Martin space H. The set B^r in general is much smaller than the usual metric r-neighborhood of B. Then, for certain "surfaces" B, the surface measure of B can be obtained as a limit of $\mu(B^r)/r$ as $r \to 0$. However, a precise definition the surface measure of B is more involved.

Among various restrictions on μ and F imposed above, certainly the most stringent one is the existence of vector fields of differentiability for μ. For example, in many cases, given a measure μ on a metric space, one can take for $\mathcal{F}$ the space of bounded Lipschitzian functions; in many cases, such functions possess appropriate gradients μ-almost everywhere, so if F is locally Lipschitzian, then the only problem is to find suitable differentiability fields for the measure. It is not always possible to build such fields from constant vector fields (this happens already for distributions of diffusion processes with non-constant diffusion coefficients, see [13, Chapter 4]). It would be interesting to study vector fields of differentiability of measures in the framework of metric measure spaces.

Remark 2.6.5. It would be interesting to study a possible analog of the Radon transform associated with surface measures in the spirit of the construction developed in [41], [42], [50], [17] for conditional measures on hyperplanes. Recall that the classical Radon transform reconstructs a function on the plane by its integrals over all straight lines (with Lebesgue measure). A natural infinite-dimensional analog (considered in the cited papers) is this: given a Radon probability measure μ, to get some information about a function with given integrals with respect to conditional measures μ^{L+y} on the set $L + y$ for all possible hyperplanes L.

2.7 Surface measures of higher codimension

In this section we consider surface measures on surfaces of higher codimension. Note that Gaussian Hausdorff measures (see Section 3) are defined in a unified way for arbitrary codimension; other approaches exploited the fact that F was a real function. Nevertheless, the construction developed in the previous sections also works in the case of surfaces of higher codimension, but requires a bit more regularity of the mapping

$$F = (F_1, \dots, F_d)\colon\ X \to \mathbb{R}^d$$

on the level sets of which we wish to define surface measures. We recall that conditional measures are not sensitive at all to this change, they exist even for mappings with values in quite general infinite-dimensional spaces.

Now we need d vector fields $v_1, \dots, v_d$ along which the measure μ is differentiable. However, in the multidimensional case it is reasonable to modify our conditions on $\mathcal{F}$ as follows: $\mathcal{F}$ is a linear space separating Radon measures on X such that

$$\phi(f_1, \dots, f_n) \in \mathcal{F} \quad \forall f_1, \dots, f_n \in \mathcal{F}$$

for all functions $\phi \in C_b^\infty(\mathbb{R}^n)$ with arbitrary n and

$$\partial_{v_i}(\phi(f_1, \dots, f_n)) = \sum_{j=1}^n \partial_{x_j}\phi(f_1, \dots, f_n)\partial_{v_i} f_j.$$

In order to modify condition (F2) we shall suppose that $\psi(F) \in \mathfrak{D}_{v_i}$ for all functions $\psi \in C_0^\infty(\mathbb{R}^d)$ and each $i = 1, \dots, d$. This enables us to define functions $\partial_{v_i} F_j$ as we have done in the one-dimensional case in (F2).

In place of $\partial_v F$ we now take the determinant Δ_F of the so-called Malliavin matrix

$$(\sigma_{ij})_{i,j\le d}, \quad \sigma_{ij} := \partial_{v_i} F_j.$$

Let $M^{ij}(x) = (-1)^{i+j} m^{ji}$, where $m^{ji}(x)$ is the minor in the Malliavin matrix corresponding to the element $\sigma_{ji}(x)$. Thus, M^{ij} is the transposed matrix of cofactors of the Malliavin matrix. If the matrix $(\sigma_{ij}(x))_{i,j\le d}$ is invertible, the inverse matrix will be denoted by $(\gamma^{ij}(x))_{i,j\le d}$. In that case

$$\gamma^{ij} = M^{ij}/\Delta_F.$$

Therefore,

$$\sum_{j\le d} M^{ij}(x)\sigma_{jk}(x) = \Delta_F(x)\delta_{ik}, \tag{2.19}$$

where δ_{ik} is Kronecker's symbol. Indeed, this is true for invertible matrices, but remains valid for any matrix by approximation by invertible matrices.

In the Gaussian case considered above we take $v_i = D_H F_i$, so that

$$\sigma_{ij} = (D_H F_i, D_H F_j)_H$$

and the matrix $(\sigma_{ij})_{i,j\le d}$ is nonnegative definite.

In place of (F2) we suppose that $\Delta_F \ge 0$ and $\Delta_F \in L^1(\mu)$. Set

$$\nu = \Delta_F \cdot \mu.$$

Let

$$U_r = \{x \in \mathbb{R}^d : \ |x| < r\} \quad \text{and} \quad W_r = \{|F| < r\}.$$

Let $BV(U_r)$ be the space of functions of bounded variation on U_r and let $W^{p,1}(U_r)$ be the Sobolev class of functions belonging to $L^p(U_r)$ along with their generalized first order partial derivatives.

Theorem 2.7.1. (i) *Suppose that $fM^{ij} \in \mathfrak{D}_{v_j}$ for all $i, j \le d$ and all $f \in \mathcal{F}$ vanishing outside of W_r. Then the measure $\nu \circ F^{-1}$ is absolutely continuous on U_r and has a density ϱ of class $BV(U_r)$. In particular, $\varrho \in L^{d/(d-1)}(U_r)$.*

If $\Delta_F(x) \neq 0$ μ-a.e. on W_r, then $\varrho \in W^{1,1}(U_r)$.

(ii) *If, in addition,*

$$u_i := \frac{I_{W_r}}{\Delta_F} \sum_{j\le d} [\partial_{v_j} M^{ij} + M^{ij}\beta_{v_j}] \in L^s(\nu) \quad \text{for some } s > d, \tag{2.20}$$

then this density ϱ belongs to $W^{p,1}(U_r)$ with some $p > d$ and has a continuous version.

(iii) *If $s > 2d$, then for any $f \in \mathcal{F}$ the measure $(f \cdot \nu) \circ F^{-1}$ is absolutely continuous on U_r and has a bounded continuous density ϱ_f such that*

$$\sup_{y \in U_r} |\varrho_f(y)| \le C\Big(\|f\|_{L^{2d}(\nu)} + \sum_{j=1}^{d} \|\partial_{v_j} f\|_{L^{2d}(\nu)}\Big), \tag{2.21}$$

where C is a number that depends only on $d, s, r_0, \|u_i\|_{L^s(\nu)}$, and $\|I_{W_r} M^{ij}/\Delta_F\|_{L^s(\nu)}$, whenever $r \le r_0$ and $r_0 > 0$ is fixed.

Proof. (i) Let $\psi \in C_0^\infty(U_r)$. By (2.19) we have

$$\begin{aligned}
\int_{U_r} \partial_{y_i}\psi(y)\, \nu \circ F^{-1}(dy) &= \int_{W_r} \partial_{y_i}\psi(F(x))\Delta_F(x)\, \mu(dx) \\
&= \int_X \sum_{j,k \le d} M^{ij}(x)\sigma_{jk}(x)[\partial_{y_k}\psi(F(x))]\, \mu(dx) = \int_X \sum_{j \le d} \partial_{v_j}(\psi \circ F)(x) M^{ij}(x)\, \mu(dx) \\
&= -\sum_{j \le d} \int_X (\psi \circ F)(x) M^{ij}(x)\, d_{v_j}\mu(dx) - \sum_{j \le d} \int_X (\psi \circ F)(x)\partial_{v_j} M^{ij}(x)\, \mu(dx) \\
&= -\sum_{j \le d} \int_X (\psi \circ F)(x)[\partial_{v_j} M^{ij}(x) + M^{ij}(x)\beta_{v_j}(x)]\, \mu(dx).
\end{aligned}$$

The right-hand side can be written as the integral of ψ with respect to a bounded measure on U_r, hence the measure $\nu \circ F^{-1}$ on U_r has a density ϱ of class $BV(U_r)$. By the Sobolev embedding theorem $\varrho \in L^{d/(d-1)}(U_r)$, see, e.g., [13, Chapter 2], [33] or [79].

If the measure ν is equivalent to μ, which is the case where $\Delta_F > 0$ μ-a.e. and μ can be written as $\Delta_F^{-1} \cdot \nu$, the right-hand side can be written as the integral of $\psi g_i \varrho$, where g_i is the conditional expectation of the ν-integrable function $-u_i$ with respect to the measure ν and the σ-field generated by F (here we also take into account that $\psi(F) = \psi(F)I_{W_r}$, because ψ has support in U_r). Therefore, $\varrho \in W^{1,1}(U_r)$.

(ii) Note that for some function g_i we have

$$\partial_{y_i}\varrho = g_i \varrho.$$

By Jensen's inequality for conditional expectations the inclusion $|u_i|^s \in L^1(\nu)$ yields the inclusion $|g_i|^s \varrho \in L^1(U_r)$. We recall that this inequality states that

$$V(\mathrm{E}^{\mathcal{A}}\xi) \le \mathrm{E}^{\mathcal{A}}[V(\xi)]$$

for every convex function V. In particular, $|\mathrm{E}^{\mathcal{A}}\xi|^p \le \mathrm{E}^{\mathcal{A}}[|\xi|^p]$ for all $p \ge 1$.

We now show that $\partial_{y_i}\varrho$ is better integrable under the assumptions of the last assertion. Suppose that $\varrho \in L^p(U_r)$ for some $p \ge 1$. By Hölder's inequality we have $g_i\varrho \in L^{sp/(p+s)}(U_r)$. Therefore, $\varrho \in W^{p_1,1}(U_r)$ with $p_1 = sp/(p+s)$; if $p_1 < d$ by the Sobolev embedding this yields that $\varrho \in L^{p_2}(U_r)$ with

$$p_2 = \frac{dp_1}{d - p_1} = p\frac{ds}{ds - p(s-d)} \ge p\frac{ds}{ds - s + d} = \lambda p, \quad \lambda = \frac{ds}{ds - s + d} > 1.$$

If $p_1 = d$ then $\varrho \in L^q(U_r)$ for any $q < \infty$, hence $\partial_{y_i}\varrho \in W^{s-\varepsilon,1}(U_r)$ for any $\varepsilon > 0$. Therefore, in finitely many steps we arrive at the situation where $\partial_{y_i}\varrho \in W^{p,1}(U_r)$ with some $p > d$. Therefore, the Sobolev embedding ensures a continuous density.

(iii) It follows from the reasoning above that if we replace ν by the measure $f \cdot \nu$, the generalized partial derivative of the measure $(f \cdot \nu) \circ F^{-1}$ with respect to y_i will equal the function $\widehat{g}_i\varrho$, where $\widehat{g}_i$ is the conditional expectation of the function

$$\widehat{u}_i = fu_i + \frac{I_{W_r}}{\Delta_F}\sum_{j\le d} M^{ij}\partial_{v_j}f$$

with respect to the measure ν and the σ-field generated by F. Thus, we have

$$\partial_{y_i}\varrho_f = \widehat{g}_i\varrho.$$

However, now we already know that ϱ is bounded continuous on U_r and it follows from the previous step that its sup on U_r is estimated by a constant that depends on d, s, r and the $L^s(\nu)$-norm of u_i. Therefore, choosing $\varepsilon > 0$ such that $s = 2d(d+\varepsilon)/(d-\varepsilon)$ and letting $t = d + \varepsilon > d$, we conclude that the $L^t(U_r)$-norm of $\partial_{y_i}\varrho_f$ is estimated by a constant depending on the indicated quantities and the $L^t(\nu)$-norm of $\widehat{u}_i$. By Hölder's inequality

$$\|uw\|_t \le \|u\|_s\|w\|_{st/(s-t)}, \qquad \frac{st}{s-t} = 2d.$$

We apply this inequality with $u = u_i$ and $w = f$ and also with $u = I_{W_r}M^{ij}/\Delta_F$ and $w = \partial_{v_j}f$. This gives a bound on the $W^{t,1}(U_r)$-norm of ϱ_f via

$$\|f\|_{L^{2d}(\nu)} + \sum_{j=1}^{d}\|\partial_{v_j}f\|_{L^{2d}(\nu)}$$

multiplied by a constant, which yields the announced estimate by the Sobolev embedding theorem. □

We now give a constructive sufficient condition for the continuity of densities of multidimensional distributions related to μ rather than ν. This requires, however, second derivatives of F. In the next proposition we assume that $\partial_{v_k}\partial_{v_j}F_i$ can be defined in the same sense as $\partial_{v_j}F_i$ above by using that $\psi(\partial_{v_j}F) \in \mathfrak{D}_{v_k}$ for smooth functions on $\mathbb{R}^d$ with compact support.

Theorem 2.7.2. (i) *Suppose that for every $r \in \mathbb{N}$ there is $\varepsilon_r > 0$ such that the functions*

$$\exp\Big(\varepsilon_r\Big|\sum_k(\gamma^{ik}\beta_{v_k} + \partial_{v_k}\gamma^{ik})\Big|\Big) \tag{2.22}$$

are μ-integrable on the set $\{|F| < r\}$. Then the measure $\mu \circ F^{-1}$ has a continuous density without zeros.

(ii) *Suppose that for every $r \in \mathbb{N}$ there is $p_r > d$ such that the functions*

$$\Big|\sum_k \gamma^{ik}\beta_{v_k}\Big|^{p_r}, \quad \Big|\sum_k \partial_{v_k}\gamma^{ik}\Big|^{p_r}$$

are μ-integrable on the set $\{|F| < r\}$. Then the measure $\mu \circ F^{-1}$ has a continuous density.

Proof. (i) We shall use the following result (see [19] or [13, Proposition 6.4.1]): if a nonnegative function ϱ on a ball $U \subset \mathbb{R}^d$ belongs to the Sobolev class $W^{1,1}(U)$ and there is $\varepsilon > 0$ such that $\varrho \exp(\varepsilon|\nabla\varrho|/\varrho) \in L^1(U)$, where we define $\nabla\varrho/\varrho = 0$ on the set $\{\varrho = 0\}$, then ϱ has a continuous version that is either identically zero or positive.

Let us fix $r \in \mathbb{N}$ and let U be the open ball of radius r in $\mathbb{R}^d$ centered at the origin. Let $\phi \in C_0^\infty(U)$. We have

$$\begin{aligned}
\int_U \partial_{y_i}\phi(y)\,\mu\circ F^{-1}(dy) &= \int_X \partial_{y_i}\phi(F(x))\,\mu(dx) = \int_X \sum_{k,j\le d} \gamma^{ik}\sigma_{kj}(\partial_{y_j}\phi)\circ F\,d\mu \\
&= \int_X \sum_{k\le d} \gamma^{ik}\partial_{v_k}(\phi\circ F)\,d\mu = -\int_{|F|<r} \sum_{k\le d} \phi\circ F[\partial_{v_k}\gamma^{ik} + \gamma^{ik}\beta_{v_k}]\,d\mu \\
&= -\int_U \phi(y)\eta_i(y)\,\mu\circ F^{-1}(dy),
\end{aligned}$$

where η_i is the conditional expectation of the function

$$\sum_k [\partial_{v_k}\gamma^{ik} + \gamma^{ik}\beta_{v_k}] I_{\{|F|<r\}}$$

with respect to the measure μ and the σ-field generated by F. It follows that the generalized derivative of the measure $\mu \circ F^{-1}$ on U in the variable y_i is the measure $\eta_i \cdot (\mu \circ F^{-1}) \ll \mu \circ F^{-1}$. Therefore, $\mu \circ F^{-1}$ on U has a density $\varrho \in W^{1,1}(U)$ and $\partial_{y_i}\varrho/\varrho = \eta_i$. By our assumption (2.22) and Jensen's inequality for conditional expectations (now applied to exp), we arrive at the condition mentioned above.

(ii) If we are given that $\mu \circ F^{-1}$ has a locally bounded density ϱ, then the previous relation can be written as

$$\int_U \partial_{y_i}\phi(y)\varrho(y)\,dy = -\int_U \phi(y)\eta_i(y)\varrho(y)\,dy,$$

which means that $\partial_{y_i}\varrho = \eta_i\varrho$ on U in the sense of distributions. We obtain again that $\varrho \in W^{1,1}(U)$, but now we conclude that $\partial_{y_i}\varrho \in L^{p_r}(U)$ by the same iteration of the Sobolev embedding theorem as in the previous proposition. Therefore, by the Sobolev embedding theorem ϱ has a continuous density (now it is not asserted that it is positive). □

In order to ensure (2.22) in terms of the original Malliavin matrix, we note that

$$\partial_{v_k}\gamma^{ik} = \partial_{v_k}(M^{ik}\Delta_F^{-1}) = (\partial_{v_k}M^{ik})\Delta_F^{-1} - (\partial_{v_k}\Delta_F)\Delta_F^{-2}.$$

The first term is a sum of functions of the form $\Delta_F^{-1}\partial_{v_k}\sigma^{ij}w$, where w is a product of $d-2$ matrix elements of the Malliavin matrix. The second term is a sum of functions of the form $\Delta_F^{-2}\partial_{v_k}\sigma^{ij}w$ with w as above. Therefore, it suffices to have the μ-integrability of the functions

$$\exp\Big(\frac{\varepsilon_r}{\Delta_F^2}|\partial_{v_k}\partial_{v_j}F_i|\,|\partial_{v_l}F_m|^{d-1}\Big),\ \exp\Big(\frac{\varepsilon_r}{\Delta_F}|\beta_{v_k}|\,|\partial_{v_l}F_m|^{d-1}\Big),$$

on the set $\{|F|<r\}$. For example, this holds if for some $\delta_r>0$ the exponents of

$$\delta_r\Delta_F^{-4},\quad \delta_r|\partial_{v_k}\partial_{v_j}F_i|^4,\quad \delta_r|\partial_{v_l}F_m|^{4d-4},\quad \delta_r\beta_{v_k}^2$$

are integrable.

Definition 2.7.3. *The surface measure σ^y is defined as a Radon measure such that*

$$\int_X f(x)\,\sigma^y(dx)=\varrho_f(y)\quad \forall f\in\mathcal{F}.$$

This definition means that

$$\int_X f(x)\,\sigma^y(dx)=\lim_{r\to 0}\frac{1}{|U_r|}\int_{\{|F-y|<r\}} f(x)\,\nu(dx),$$

where $|U_r|$ is the usual volume of the ball U_r. The existence of the limit in the right-hand side is the only condition required by the definition, and this condition is fulfilled in the situation of Theorem 2.7.1.

As in Section 4, we have to show that this relation defines a Radon measure.

We need also an analog of Lemma 2.5.1.

Lemma 2.7.4. *Suppose that the hypotheses of case* (iii) *of Theorem* 2.7.1 *hold and*

$$\|f\|_{L^{2d}(\mu)}+\sum_{j=1}^d\|\partial_{v_j}f\|_{L^{2d}(\mu)}\le C_0\|f\|_{\mathcal{F}}\quad \forall f\in\mathcal{F}. \tag{2.23}$$

Then, for every open set $W\subset X$ and any $r>0$, we have

$$\nu(W\cap\{|F-y|<r\})\le C_0C_1r^dC_{\mathcal{F}}(W), \tag{2.24}$$

where C_1 depends on the same quantities as in assertion (iii) *of Theorem* 2.7.1.

Proof. Let $f\in\mathcal{F}$, $f\ge 0$ and $f\ge 1$ μ-a.e. on W. Then $f\ge 1$ ν-a.e. on W, hence as in the proof of Lemma 2.5.1 we obtain

$$\nu(W\cap\{|F-y|<r\})\le r^d\sup_{z:\ |z-y|<r}|\varrho_f(z)|,$$

where ϱ_f is the density of the measure $(f \cdot \nu) \circ F^{-1}$. According to assertion (iii) of Theorem 2.7.1 we can estimate the maximum of the continuous version of ϱ_f on W_r by

$$\|f\|_{L^{2d}(\nu)} + \sum_{j=1}^{d} \|\partial_{v_j} f\|_{L^{2d}(\nu)}$$

multiplied by some constant depending on the quantities indicated in that proposition. □

Theorem 2.7.5. *In the situation of Theorem* 2.7.1(iii)*, the assertion of Theorem* 2.4.7 *is true. The assertion of Theorem* 2.5.2 *is true as well if the condition of the previous lemma holds and the respective quasi-continuous versions of* F_i *are considered.*

The proof is essentially the same, however, we should note that the assumptions are now much stronger.

As in case $d = 1$ we can equip $\mathcal{F}$ with the norm given by the left-hand side of (2.23). However, this is not always convenient since this norm depends on F and ν. If the Gaussian case it may be preferable to use some Sobolev norm on $\mathcal{F}$. For example, if we take $v_i = D_H F_i$ as in Example 2.6.3, then $\partial_{v_i} f = (D_H f, D_H F_i)_H$, so that $\partial_{v_i} f \in L^{2d}(\nu)$ provided that $f \in W^{4d,1}(\mu)$, $F_i \in W^{8d,1}(\mu)$ and $\Delta_F^{-1} \in L^{8d-4}(\mu)$.

For Fomin differentiable measures, the above construction applies under much broader assumptions than in [78].

Remark 2.7.6. Since $\sigma^y = \varrho_1(y)\nu^y$, every ν-integrable $\mathcal{B}$-measurable function g is σ^y-integrable for $\nu \circ F^{-1}$-almost every y. This enables us to define surface measures for $g \cdot \nu$. Alternatively, we can use the trick described in Remark 2.5.3.

It should be noted the construction presented is chiefly oriented towards infinite-dimensional spaces, where typical measures are not doubling and differ also in other respects from measures usual in measure metric spaces. Nevertheless, it would be interesting to compare suitable Hausdorff measures on metric measure spaces and surface measures described above; measures differentiable along vector fields and all other objects considered above (differentiations, gradients, Sobolev classes, etc.) are meaningful on such spaces (see, e.g., [27], [40], [46], [64], and [65]). In particular, if we have a differentiation of the form $f \mapsto \Gamma(f, g)$ defined by a fixed function in the Dirichlet space (see [64]) built on a probability space (X, μ) admitting a “carré du champ” $\Gamma(f, g)$ such that $2\Gamma(f, g) = L(fg) - fLg - gLf$, where L is a Markov symmetric generator such that $\mathrm{E}\Gamma(f, g) = -\mathrm{E}fLg$, then we see that Lg is precisely the divergence of the considered field (a similar framework is considered in [70]). In relation to constructing vector fields possessing divergences, which is crucial for the presented approach, the recent paper [38] is of interest.

Finally, it would be interesting to continue investigation of surface measures and surface Sobolev and Besov classes connected, in particular, with restrictions of

Sobolev and Besov classes on the whole space. About fractional Sobolev classes over infinite-dimensional spaces, see [25], [52]–[54], and [55].

Bibliography

[1] H. Airault, P. Malliavin: Intégration géometrique sur l'espace de Wiener. Bull. Sci. Math. (2) 112(1), 3–52 (1988).

[2] L. Ambrosio, G. Da Prato, D.Pallara: *BV* functions in a Hilbert space with respect to a Gaussian measure. Rend. Acc. Lincei 21 (2010), 405–414.

[3] L. Ambrosio, S. Di Marino: Equivalent definitions of *BV* space and of total variation on metric measure spaces. J. Funct. Anal. 266, 4150–4188 (2014).

[4] L. Ambrosio, A. Figalli: Surface measures and convergence of the Ornstein–Uhlenbeck semigroup in Wiener spaces. Ann. Fac. Sci. Toulouse Math. (6) 20(2), 407–438 (2011).

[5] L. Ambrosio, A. Figalli, E. Runa: On sets of finite perimeter in Wiener spaces: reduced boundary and convergence to halfspaces. Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. 24(1), 111–122 (2013).

[6] L. Ambrosio, M. Miranda, S. Maniglia, D. Pallara: BV functions in abstract Wiener spaces. J. Funct. Anal. 258(3), 785–813 (2010).

[7] L. Ambrosio, M. Miranda, D. Pallara: Sets with nice perimeter in Wiener spaces, perimeter measure and boundary rectifiability. Discrete Contin. Dyn. Syst. 28 (2010), 591–606.

[8] L. Ambrosio, M. Miranda, D. Pallara: Some nice properties of *BV* functions on Wiener spaces. Anal. Geom. Metr. Spaces 3 (2015), 212–230.

[9] V.I. Bogachev: Smooth measures, the Malliavin calculus and approximation in infinite dimensional spaces. Acta Univ. Carolinae, Math. et Phys. 31(2), 9–23 (1990).

[10] V.I. Bogachev: Differentiable measures and the Malliavin calculus. J. Math. Sci. (New York) 87(4), 3577–3731 (1997).

[11] V.I. Bogachev: Gaussian measures. Amer. Math. Soc., Providence, Rhode Island, 1998.

[12] V.I. Bogachev: Measure theory. V. 1, 2. Springer, Berlin, 2007.

[13] V.I. Bogachev: Differentiable measures and the Malliavin calculus. Amer. Math. Soc., Providence, Rhode Island (2010).

[14] V.I. Bogachev: Gaussian measures on infinite-dimensional spaces. In: Real and Stochastic Analysis. Current Trends (M.M. Rao ed.), pp. 1–83. World Sci., Singapore (2014).

[15] V.I. Bogachev: Sobolev classes on infinite-dimensional spaces. In: Geometric Measure Theory and Real Analysis (L. Ambrosio ed.), pp. 1–56, Publications of the Scuola Normale Superiore, V. 17, Pisa, Edizioni della Normale, 2014.

[16] V.I. Bogachev: Distributions of polynomials on multidimensional and infinite-dimensional spaces with measures. Uspehi mat. Nauk 71(3) (2016) (in Russian); English transl.: Russian Math. Surveys, 71(3) (2016).

[17] V.I. Bogachev, M.N. Lukintsova: The Radon transform in infinite-dimensional spaces. Dokl. Akad. Nauk 443(3), 279–282 (2012), (in Russian); English transl.: Dokl. Math. 85(2), 208–211 (2012).

[18] V.I. Bogachev, I.I. Malofeev: Surface measures generated by differential measures. Potential Analysis 44(4), 767–792 (2016).

[19] V.I. Bogachev, E. Mayer-Wolf: Absolutely continuous flows generated by Sobolev class vector fields in finite and infinite dimensions. J. Funct. Anal. 167(1), 1–68 (1999).

[20] V.I. Bogachev, A.Yu. Pilipenko, E.A.Rebrova: Classes of functions of bounded variation on infinite-dimensional domains. Dokl. Akad. Nauk 451(2), 127–131 (2013) (in Russian); English transl.: Dokl. Math. 88(1), 391–395 (2013).
[21] V.I. Bogachev, A.Yu. Pilipenko, A.V.Shaposhnikov: Sobolev functions on infinite-dimensional domains. J. Math. Anal. Appl. 419, 1023–1044 (2014).
[22] V.I. Bogachev, E.A. Rebrova: Functions of bounded variation on infinite-dimensional spaces with measures. Dokl. Russian Acad. Sci. 449(2), 131–135 (2013) (in Russian); English transl.: Dokl. Math. 87(2), 144–147 (2013).
[23] N. Bouleau, F. Hirsch: Dirichlet forms and analysis on Wiener space. De Gruyter, Berlin – New York (1991).
[24] V. Caselles, A. Lunardi, M. Miranda, M. Novaga: Perimeter of sublevel sets in infinite dimensional spaces. Adv. Calc. Var. 5(1), 59–76 (2012).
[25] P. Celada, A. Lunardi: Traces of Sobolev functions on regular surfaces in infinite dimensions. J. Funct. Anal. 266, 1948–1987 (2014).
[26] S. Chaari, F. Cipriano, H.-H. Kuo, H. Ouerdiane: Surface measures on the dual space of the Schwartz space. Comm. Stoch. Anal. 4, 467–480 (2010).
[27] J. Cheeger: Differentiability of Lipschitz functions on metric measure spaces. Geom. Funct. Anal. 9, 428–517 (1999).
[28] G. Da Prato: Lunardi A. Maximal Sobolev regularity in Neumann problems for gradient systems in infinite dimensional domains. Ann. Inst. H. Poincaré Probab. Stat. 51(3), 1102–1123 (2015).
[29] G. Da Prato, A. Lunardi, L. Tubaro: Surface measures in infinite dimension. Rendic. Lincei 25(3), 309–330 (2014).
[30] Yu.L. Daletskiĭ: Stochastic differential geometry. Uspehi Matem. Nauk. 38(3), 87–111 (1983) (in Russian); English transl.: Russian Math. Surv. 38(3), 97–125 (1983).
[31] Yu.A. Davydov, M.A. Lifshits, N.V.Smorodina: Local properties of distributions of stochastic functionals. Translated from the Russian. Amer. Math. Soc., Providence, Rhode Island (1998) (Russian ed.: Moscow (1995)).
[32] E.I. Efimova, A.V. Uglanov: Green’s formula on a Hilbert space. Mat. Sb. 119(2), 225–232 (1982) (in Russian); English transl.: Math. USSR Sb. 47, 215–222 (1984).
[33] C. Evans, R.F. Gariepy: Measure theory and fine properties of functions. CRC Press, Boca Raton – London, 1992.
[34] H. Federer: Geometric measure theory. Springer-Verlag, New York, 1969.
[35] D. Feyel: Hausdorff–Gauss measures. In: Stochastic Analysis and Related Topics, VII. Kusadasi 1998, pp. 59–76. Progr. in Probab. V. 98, Birkhäuser (2001).
[36] D. Feyel, A. de La Pradelle: Hausdorff measures on the Wiener space. Potential. Anal. 1(2), 177–189 (1992).
[37] M. Fukushima, M. Hino: On the space of BV functions and a related stochastic calculus in infinite dimensions. J. Funct. Anal. 183(1), 245–268 (2001).
[38] N. Gigli: Nonsmooth differential geometry – an approach tailored for spaces with Ricci curvature bounded from below. arXiv:1407.0809 (to appear in Memoirs of the AMS).
[39] V. Goodman: A divergence theorem for Hilbert space. Trans. Amer. Math. Soc. 164, 411–426 (1972).
[40] J. Heinonen: Nonsmooth calculus. Bull. Amer. Math. Soc. 44(2), 163–232 (2007).
[41] A. Hertle: Gaussian surface measures and the Radon transform on separable Banach spaces. Lecture Notes in Math. 794, 513–531 (1980).
[42] A. Hertle: Gaussian plane and spherical means in separable Hilbert spaces. Lecture Notes in Math. 945, 314–335 (1982).
[43] M. Hino: Sets of finite perimeter and the Hausdorff–Gauss measure on the Wiener space. J. Funct. Anal. 258(5), 1656–1681 (2010).

[44] M. Hino: Dirichlet spaces on *H*-convex sets in Wiener space. Bull. Sci. Math. 135, 667–683 (2011); Erratum: ibid. 137, 688–689 (2013).
[45] I. Holmes, A.N. Sengupta: A Gaussian Radon transform for Banach spaces. J. Funct. Anal. 263(11), 3689–3706 (2012).
[46] S. Keith: A differentiable structure for metric measure spaces. Adv. Math. 183(2), 271–315 (2004).
[47] H.-H. Kuo: Gaussian measures in Banach spaces. Springer-Verlag, Berlin – New York (1975).
[48] A. Lunardi, M. Miranda, D. Pallara: *BV* functions on convex domains in Wiener spaces. Potential Anal. 43 (2015), 23–48.
[49] P. Malliavin: Stochastic analysis. Springer-Verlag, Berlin (1997).
[50] V. Mihai, A.N. Sengupta: The Radon-Gauss transform. Soochow J. Math. 33(3), 415–433 (2007).
[51] M. Miranda, M. Novaga, D. Pallara: An introduction to *BV* functions in Wiener spaces. In: Variational methods for evolving objects (L. Ambrosio, Y. Giga, P. Rybka, Y. Tonegawa eds.), Advanced Studies in Pure Mathematics 67 (2015), 245–293.
[52] E.V. Nikitin: Fractional order Sobolev clsses on infinite-dimensional spaces. Dokl. Russian Acad. Sci. 452(2), 130–135 (2013) (in Russian); English tranls.: Dokl. Math. 88(2), 518–523 (2013).
[53] E.V. Nikitin: Besov classes on infinite-dimensional spaces. Matem. Zametki 93(6), 951–953 (2013) (in Russian); English tranls.: Math. Notes 93(6), 936–939 (2013).
[54] E.V. Nikitin: Comparison of two definitions of Besov classes on infinite–dimensional spaces. Matem. Zametki 95(1), 150–153 (2014) (in Russian); English tranls.: Math. Notes 95(1), 133–135 (2014).
[55] E. Pineda, W. Urbina: Some results on Gaussian Besov–Lipschitz spaces and Gaussian Triebel—Lizorkin spaces. J. Approx. Theory 161(2), 529–564 (2009).
[56] O.V. Pugachev: Surface measures in infinite-dimensional spaces. Matem. Zametki 63(1), 106–114 (1998) (in Russian); English transl.: Math. Notes 63(1), 94–101 (1998).
[57] O.V. Pugachev: The Gauss–Ostrogradskii formula in infinite-dimensional space. Matem. Sb. 189(5), 115–128 (1998) (in Russian); English transl.: Sb. Math. 189(5), 757–770 (1998).
[58] O.V. Pugachev: Construction of non-Gaussian surface measures by the Malliavin method. Matem. Zametki 65(3), 377–388 (1999) (in Russian); English transl.: Math. Notes 65(3-4), 315–325 (1999).
[59] O.V. Pugachev: Tightness of Sobolev capacities in infinite dimensional spaces. Infinite Dimensional Analysis, Quantum Probab. Relat. Topics 2(3), 427–440 (1999).
[60] O.V. Pugachev: Capacities and surface measures in locally convex spaces. Teor. Verojatn. i Primen. 53(1), 178–188 (2008) (in Russian); English transl.: Theory Probab. Appl. 53(1), 179–190 (2009).
[61] M. Röckner, B. Schmuland: Tightness of general $C_{1,p}$ capacities on Banach spaces. J. Funct. Anal. 108(1), 1–12 (1992).
[62] M. Röckner, R.-Ch. Zhu, X.-Ch. Zhu: The stochastic reflection problem on an infinite dimensional convex set and BV functions in a Gelfand triple. Ann. Probab. 40(4), 1759–1794 (2012).
[63] M. Röckner, R. Zhu, X. Zhu: BV functions in a Gelfand triple for differentiable measures and its applications. Forum Math. 27(3), 1657–1687 (2015).
[64] G. Savaré: Self-improvement of the Bakry–Émery condition and Wasserstein contraction of the heat flow in RCD(K, ∞) metric measure spaces. Discrete Contin. Dyn. Syst. 34(4), 1641–1661 (2014).
[65] A. Schioppa: On the relationship between derivations and measurable differentiable structures. Ann. Acad. Sci. Fenn. Math. 39(1), 275–304 (2014).
[66] A.V. Skorohod: Surface integrals and the Green formula in Hilbert space. Teor. Verojatnost. i Mat. Statist. N2, 172–175 (1970) (in Russian).

[67] A.V. Skorohod: Integration in Hilbert space. Transl. from the Russian. Springer-Verlag, Berlin – New York (1974).
[68] N.V. Smorodina: Differential calculus in configuration space and stable measures. I. Teor. Verojatn. i Primen. 33(3), 522–534 (1988) (in Russian); English transl.: Theory Probab. Appl. 33(3), 488–499 (1988).
[69] G. Stengle: A divergence theorem for Gaussian stochastic process expectations. J. Math. Anal. Appl. 21, 537–546 (1968).
[70] D. Stroock: The Malliavin calculus, a functional analytic approach. J. Funct. Anal. 44(2), 212–257 (1981).
[71] K.-T. Sturm: On the geometry of metric measure spaces. I, II. Acta Math. 196(1), 65–131, 133–177 (2006).
[72] J.F. Trèves: Linear partial differential equations with constant coefficients: existence, approximation and regularity of solutions. Gordon and Breach, New York, 1966.
[73] A.V. Uglanov: Surface integrals in a Banach space. Matem. Sb. 110(2), 189–217 (1979) (in Russian); English transl.: Math. USSR Sb. 38, 175–199 ().
[74] A.V. Uglanov: Surface integrals in Fréchet spaces. Matem. Sb. 189(11), 139–157 (1998) (in Russian); English transl.: Sb. Math. 189(11), 1719–1737 (1998).
[75] A.V. Uglanov: Integration on infinite-dimensional surfaces and its applications. Kluwer Acad. Publ., Dordrecht (2000).
[76] A.V. Uglanov: Calculus of variations on Banach spaces. Mat. Sb. 191(10), 105–118 (2000) (in Russian); English transl.: Sbornik Math. 191(10), 1527–1540 (2000).
[77] A.V. Uglanov: Surface integrals in locally convex spaces. Tr. Mosk. Mat. Ob-va. 62, 262–285 (2001) (in Russian); English transl.: Trans. Mosc. Math. Soc. 249–270 (2001).
[78] V.Yu. Yakhlakov: Surface measures on surfaces of finite codimension in a Banach space. Matem. Zametki 47(4), 147–156 (1990) (in Russian); English transl.: Math. Notes 47(3-4), 414–421 (1990).
[79] W. Ziemer: Weakly differentiable functions. Springer-Verlag, New York – Berlin (1989).