ON SECOND ORDER DERIVATIVES OF CONVEX FUNCTIONS ON INFINITE DIMENSIONAL SPACES WITH MEASURES

VLADIMIR I. BOGACHEV AND BEN GOLDYS

ABSTRACT. We consider convex functions on infinite dimensional spaces equipped with measures. Our main results give some estimates of the first and second derivatives of a convex function, where second derivatives are considered from two different points of view: as point functions and as measures.

INTRODUCTION

In recent years several works have been published on infinite dimensional extensions of the classical result of A.D. Alexandroff on the second order differentiability of convex functions and related problems (see [4], [9], [10]). It turns out that the Alexandroff theorem has no direct extension to infinite dimensions, although a number of interesting positive results have been proved. One of the negative results is that, given a nice measure μ on an infinite dimensional separable Hilbert space X, one can find a convex function that has no second derivative at almost every point with respect to μ. The situation is similar to that of the Fréchet differentiability of a Lipschitzian function; moreover, a convex Lipschitzian function that fails to be Fréchet differentiable μ-a.e. provides a counter-example for the second order differentiability. It is known, however, that the situation with the Fréchet differentiablity of Lipschitzian functions changes if one considers the differentiability along a compactly embedded subspace $E \subset X$. Then, for any reasonable measure μ, one obtains the Fréchet differentiablity along E almost everywhere with respect to μ (see, e.g., [3]). In this paper, we make an attempt to investigate along the same lines the second order differentiability of convex functions. A study of convexity along a smaller subspace has been undertaken in [6], [11], where H-convex functions have been introduced in the case of a Gaussian measure with the Cameron–Martin space H. Here we are concerned with more general measures and mostly deal with convexity on the lines parallel to a given vector along which the measure is differentiable. Our main results give some estimates of the first and second derivatives of a convex function, where second derivatives are considered from two different points of view: as point functions and as measures.

1. TERMINOLOGY AND AUXILIARY RESULTS

Throughout the term *a Radon measure* μ means a bounded (possibly signed) Borel measure which is compact inner regular and $\|\mu\|$ stands for the total variation of μ. A function f on a locally convex space X is called smooth cylindrical if

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

The space of all functions of such a form is denoted by $\mathcal{FC}_b^\infty$. We recall that a Radon measure μ on a locally convex space X is called differentiable along a vector $h \in X$ in the sense of Skorohod

1991 *Mathematics Subject Classification.* 28C20, 46G05, 46G12, 60H07.

Key words and phrases. Convex function, second derivative, differentiable measure, Gaussian measure.

(or Skorohod differentiable) if there exists a measure $d_h\mu$ (called the Skorohod derivative of μ along h) such that for every smooth cylindrical function f one has

$$\int \partial_h f(x)\,\mu(dx) = -\int f(x)\,d_h\mu(dx).$$

If $d_h\mu \ll \mu$, then μ is called Fomin differentiable. In that case, the Radon–Nikodym derivative β_h^μ of $d_h\mu$ with respect to μ is called the logarithmic derivative of μ along h. The terminology is explained by the fact that in the case when $X = \mathbb{R}$ and $h = 1$, the measure μ is Fomin differentiable precisely when μ has an absolutely continuous density ϱ with $\varrho' \in L^1(\mathbb{R})$; then $\beta_1^\mu = \varrho'/\varrho$. The existence of a Skorohod derivative in the one dimensional case is equivalent to the existence of a density ϱ of bounded variation. Then $d_1\mu$ is the derivative of μ in the sense of generalized functions. The situation is similar in $\mathbb{R}^n$, e.g., μ is Fomin differentiable along n linearly independent directions if and only if μ has a density ϱ which belongs to the Sobolev class $W^{1,1}(\mathbb{R}^n)$; then $d_h\mu = \partial_h\varrho\,dx$ and $\beta_h^\mu = \partial_h\varrho/\varrho$.

The shift of a measure μ on X along a vector h, i.e., the measure $B \mapsto \mu(B-h)$, is denoted by μ_h. If μ_{th} is equivalent to μ for all real t, then μ is called quasi-invariant along h. The quasi-invariance along h implies the continuity of μ along h, i.e., the equality $\lim\limits_{t\to 0} \|\mu - \mu_{th}\| = 0$. A measure on $\mathbb{R}^n$ is continuous along n linearly independent vectors precisely when it is absolutely continuous. The quasi-invariance along n linearly independent vectors is equivalent to the existence of a density that does not vanish almost everywhere.

Higher order derivatives are definited inductively; e.g., $d_h^2\mu := d_h(d_h\mu)$. If $h, k \in X$ are such that both $d_hd_k\mu$ and $d_kd_h\mu$ exist, then one can show that $d_hd_k\mu = d_kd_h\mu$, and we say that $d_{hk}^2\mu := d_hd_k\mu = d_kd_h\mu$ exists. It is worth noting that if μ is differentiable (in Skorohod's or Fomin's sense) along h and k, then it is differentiable in the same sense along any linear combination of h and k and $d_{sh+tk}\mu = sd_h\mu + td_k\mu$. If μ is twice Fomin differentiable along h, then the density of $d_h^2\mu$ with respect to μ is denoted by $\beta_{h,h}^\mu$. We observe that if β_h^μ is in $L^2(\mu)$ and has a μ-integrable partial derivative $\partial_h\beta_h^\mu$, then μ is twice Fomin differentiable along h and $d_h^2\mu = [\partial_h\beta_h^\mu + (\beta_h^\mu)^2]\mu$.

Another useful fact that we employ below is that if μ is differentiable along h and k, then one can find differentiable (in the same sense) conditional measures on the planes parallel to the linear span L of h and k. More precisely, let Y be any closed linear subspace in X such that X is a topological sum of L and Y. Let ν be the image of μ under the natural projection to Y. Then one can find measures μ^y, $y \in Y$, on the subspaces $y + L$ that are differentiable along h and k in the same sense as μ and

$$\mu(B) = \int_Y \mu^y(B)\,\nu(dy), \quad B \in \mathcal{B}(X).$$

The same is true in the case of quasi-invariance or continuity along h and k (see, e.g. [1]).

A Radon probability measure μ on X is called centered Gaussian if, for every continuous linear functional f on X, the induced measure $\mu \circ f^{-1}$ on the line is centered Gaussian. Given $h \in X$, let us set

$$|h|_H := \sup\biggl\{ f(h)\colon\ f \in X^*, \int f(x)^2\,\mu(dx) \le 1 \biggr\}.$$

The space

$$H = H(\mu) := \bigl\{ h \in X\colon\ |h|_H < \infty \bigr\}$$

is called the Cameron–Martin space of μ. It is known that H with the norm $|\cdot|_H$ is a separable Hilbert space and its natural embedding into X is compact. If X is a Hilbert space, then

$H = R(X)$, where R is the Hilbert–Schmidt operator given by $(Rh, Rk) = \int (h,x)(k,x)\,\mu(dx)$. To every $h \in H$, there corresponds a unique element $\widehat{h}$ from the closure of X^* in $L^2(\mu)$ specified by the equality

$$\int_X \widehat{h}(x) f(x)\,\mu(dx) = f(h), \quad f \in X^*.$$

The element $\widehat{h}$ has a linear version and is called the measurable linear functional generated by h. It is known that a centered Gaussian measure μ is infinitely differentiable along all vectors $h \in H = H(\mu)$ and $\beta_h^\mu = -\widehat{h}$. In addition, μ is quasi-invariant along h and the density of μ_h with respect to μ is given by $\varrho_h(x) = \exp\bigl[\widehat{h}(x) - |h|_H^2/2\bigr]$. The Ornstein–Uhlenbeck semigroup $(T_t)_{t\geq 0}$ on $L^1(\mu)$ is defined by the formula

$$T_t\psi(x) = \int \psi\Bigl(e^{-t}x + \sqrt{1-e^{-2t}}y\Bigr)\,\mu(dy).$$

For more details on Gaussian measures, see, e.g., [2].

Now let F be a real valued μ-measurable function on X such that the restriction of F to $y + L$ is convex for ν-a.e. y, where L is the linear span of the vectors h and k along which μ is continuous. According to the Alexandroff theorem, this restriction is twice differentiable at almost every point $x \in L$, where L is equiped with Lebesgue measure induced by the isomorphism of L and $\mathbb{R}^2$. To be more precise, the limit

$$\begin{aligned}\partial_h\partial_k F(x) :=& \frac{1}{2}\lim_{t\to 0}\, t^{-2}\Bigl[F(x+th+tk) + F(x-th-tk) \\ &- F(x+th) - F(x-th) - F(x+tk) - F(x-tk) + 2F(x)\Bigr]\end{aligned}$$

exists a.e. on L. Hence, $\partial_h\partial_k F$ exists μ-a.e. For the same reason,

$$\partial_h^2 F(x) := \lim_{t\to 0} t^{-2}[F(x+th) + F(x-th) - F(x+th) - 2F(x)]$$

exists μ-a.e. In addition, $\partial_h^2 F(x) \geq 0$.

We shall call F convex along a linear subspace E (or E-convex) if $h \mapsto F(x+h)$ is convex on E for every $x \in X$. If E is endowed with a norm $|\cdot|_E$, then we say that an E-convex function F is second order differentiable along E at a point x if there exist $l_x \in \partial F(x)$, where $\partial F(x)$ is the sub-differential of F at x, and a bounded linear operator $T_x\colon E \to E^*$ such that for each $h \in E$ one has

$$F(x+th) - F(x) = tl_x(h) + \frac{t^2}{2}T_x(h)(h) + o(t^2), \quad (t \to 0).$$

If F is Gâteaux differentiable along E at the point x, then $\partial F(x)$ consists of a single element $D_E F(x) \in E^*$.

It should be noted that $\partial_h^2 F$ may be different from the second derivative of F in the sense of generalized functions. For example, if f is the usual Cantor function on $[0,1]$ and $F(x) = \int_0^x f(t)\,dt$, then $F'' = 0$ a.e., but F'' is not zero in the sense of distributions. It is known that if F is a finite convex function on $\mathbb{R}^n$, then there exist locally bounded Borel measures F_{ij} on $\mathbb{R}^n$ such that the generalized derivative $\partial_{x_i}\partial_{x_j} F$ is the measure F_{ij}, and the matrix $(F_{ij}(B))_{i,j\leq n}$ is nonnegative for every Borel set B. One can write the decomposition $F_{ij} = F_{ij}^{ac}\,dx + F_{ij}^{sing}$ into the absolutely continuous and singular parts and then almost everywhere

$$F_{ij}^{ac} = \lim_{t\to 0} t^{-2}\Big[F(x+te_i+te_j)+F(x-te_i-te_j)$$
$$-F(x-te_i)-F(x-te_j)+2F(x)\Big],$$

where $e_1,\dots,e_n$ is the standard basis in $\mathbb{R}^n$. One of our goals is to obtain a similar decomposition in the infinite dimensional case. The main point is, of course, to find a suitable analogue of second generalized derivatives. Our approach is as follows. Suppose that in the finite dimensional case we replace Lebesgue measure by some probability measure μ with a nice density ϱ. Then the generalized second derivative of F can be expressed via ϱ. Namely, we can consider the measure $F\mu$ (i.e., the measure with density F with respect to μ), look at its generalized second order derivatives $\partial^2_{x_i}(F\mu)$, which are locally finite measures. If F is twice differentiable in the usual sense, then we can recover $\partial^2_{x_i}F$ from the expression

$$(\partial^2_{x_i}F)\mu = \partial^2_{x_i}(F\mu) - 2\partial_{x_i}F\partial_{x_i}\mu - F\partial^2_{x_i}\mu,$$

where the right hand side exists as a locally bounded measure for any convex F. In general, however, the right hand side is not absolutely continuous and $\partial^2_{x_i}F$ has to be recovered from its absolutely continuous part. We shall follow this approach also in infinite dimensions.

Although the pointwise second order derivative does not completely characterize the function, it is of interest to have the integrability of the function $\partial^2_h F$. The next result gives a sufficient condition for the integrability along with a bit stronger version of the above mentioned finite dimensional differentiability.

Lemma 1.1. (i) *Let h_n be a sequence of vectors along which μ is differentiable and let L be the linear span of $\{h_n\}$. Suppose a measurable function F is convex along L. Then μ-a.e. F has the second derivative along every finite dimensional subspace in L.*

(ii) *Let $F\in L^p(\mu)$ for some $p>1$. Assume that μ is quasi-invariant along h and that for the Radon–Nikodym derivative ϱ_{th} of μ_{th} with respect to μ we have*

$$\big|t^{-2}[\varrho_{th}(x)+\varrho_{-th}(x)-2]\big| \le G(x),$$

where $G\in L^{p'}(\mu)$. Then the limit

$$\partial^2_h F(x) = \lim_{t\to 0}\frac{F(x+th)+F(x-th)-2F(x)}{t^2},$$

which exists μ-a.e., defines a nonnegative μ-integrable function. In particular, this assertion is true if μ is a centered Gaussian measure with the Cameron–Martin space H and $h\in H$.

Proof. (i) In the proof of the Alexandroff theorem in [5], it has been verified that the second order derivative of a convex function on $\mathbb{R}^n$ exists at a point x provided that x is a Lebesgue point for the first derivative of F and for the absolutely continuous part of the second order derivative and, in addition, the singular component ν of the second order derivative satisfies the condition $\lim_{r\to 0}|\nu|(B(x,r))r^{-n}=0$, where $B(x,r)$ is the closed ball of radius r centered at x. By using the conditional measures, one obtains all the three conditions a.e. on every finite dimensional subspace which is a shift of the linear span of $h_1,\dots,h_n$.

(ii) The fact that $\partial^2_h F(x)\ge 0$ whenever it exists, follows by convexity. In order to show that $\partial^2_h F$ is integrable, it suffices, by Fatou's theorem, to obtain an upper bound on the integrals of

$$g_n(x) := n^2[F(x+n^{-1}h)+F(x-n^{-1}h)-2F(x)].$$

We have

$$\int g_n(x)\,\mu(dx) = \int F(y)n^2[\varrho_h(y-n^{-1}h)+\varrho_h(y+n^{-1}h)-2]\,\mu(dy)$$
$$\leq \int |F(y)|G(y)\,\mu(dy).$$

In the Gaussian case, we recall that $\varrho_{th}(x)=\exp(t\widehat{h}(x)-t^2|h|_H^2/2)$, where $\widehat{h}$ is the measurable linear functional generated by h. Now it suffices to note that $\exp|\widehat{h}|\in L^s(\mu)$ for all $s<\infty$. □

2. Main results

We shall now see that if a function F is convex along h and is integrable together with $\partial_h F$ with respect to a reasonable measure μ, then the measure $F\mu$ is twice differentiable along h. This fact enables one to consider a generalized second derivative of F along h. Moreover, we shall see that $\partial_h F$ is actually better integrable than we assume a priori.

Theorem 2.1. *Suppose that a Radon probability measure μ on X is twice Skorohod differentiable along a vector $h\in X$ and that F is convex on almost all lines $x+\mathbb{R}^1h$. Assume also that F is integrable with respect to the measures μ and $d_h^2\mu$ and that $\partial_h F$ is integrable with respect to $d_h\mu$. Then the measure $F\mu$ is twice Skorohod differentiable along h. In addition, one has*

$$\|d_h^2(F\mu)\|\leq 2\|F\|_{L^1(d_h^2\mu)}+2\|\partial_h F\|_{L^1(d_h\mu)}. \tag{2.1}$$

If $F\geq 0$ and one has $F^p, F|\beta_h^\mu|^p\in L^1(\mu)$ for some $p>1$, then

$$\int |\partial_h F|^r\,d\mu<\infty \tag{2.2}$$

for some $r>1$. Finally, if $F\in L^\alpha(\mu)$ for all $\alpha\in[1,\infty)$ and $\beta_h^\mu\in L^2(\mu)$, then $\partial_h F\in L^r(\mu)$ for every $r<2$.

Proof. Let us consider first the one dimensional case. In addition, we shall assume that the support of μ belongs to some bounded interval $[a,b]$. Clearly, F is Lipschitzian on $[a,b]$. The measure μ has an absolutely continuous density ϱ such that ϱ' has bounded variation. Hence the measure $F\mu$ is differentiable and $d_1(F\mu)=F'\mu+Fd_1\mu$. Assume, in addition, that F and ϱ are smooth. Then, certainly, the measure $d_1(F\mu)$ is Skorohod differentiable, but we need an estimate of the variation of its derivative. By convexity, $F''\geq 0$. Therefore,

$$0\leq \int F''(x)\varrho(x)\,dx=\int F(x)\varrho''(x)\,dx\leq \int |F(x)|\,|d_1^2\mu|(dx). \tag{2.3}$$

As we have

$$d_1^2(F\mu)=F''\mu+2F'd_1\mu+Fd_1^2\mu,$$

we obtain from (2.3) the estimate

$$\|d_1^2(F\mu)\|\leq\|F''\mu\|+2\|F'd_1\mu\|+\|Fd_1^2\mu\|\leq 2\|F\,d_1^2\mu\|+2\|F'd_1\mu\|.$$

Therefore, estimate (2.1) is established in the present special case. Now, still assuming that F is smooth and μ has bounded support, but ϱ is only absolutely continuous with ϱ' of bounded variation, we can find a sequence of smooth probability densities ϱ_j with support in a fixed interval such that

$$\lim_{j\to\infty}\int\Big[|\varrho_j(x)-\varrho(x)|+|\varrho_j'(x)-\varrho'(x)|\Big]\,dx=0$$

and $\sup_j \|d_1^2\mu_j\| < \infty$. By (2.1) we have

$$\limsup_{j\to\infty} \|d_1^2(F\mu_j)\| \le 2\|Fd_1^2\mu\| + 2\|F'd_1\mu\|.$$

This yields (see [1]) that $d_1^2(F\mu)$ exists and

$$\|d_1^2(F\mu)\| \le \limsup_{j\to\infty} \|d_1^2(F\mu_j)\|.$$

Indeed, it suffices to note that for every $\psi \in C_0^\infty(\mathbb{R})$ with $|\psi| \le 1$, one has

$$\begin{aligned}\int \psi\, d_1^2[F\mu](dx) &= \lim_{j\to\infty} \int \psi''\, F\mu_j(dx)\\ &= \lim_{j\to\infty} \psi\, d_1^2[F\mu_j](dx) \le \|d_1^2(F\mu_j)\|.\end{aligned}$$

The next step is to relax the smothness assumption on F still assuming that μ has bounded support in some $[a, b]$. To this end, it suffices to note that there exists a sequence of smooth convex functions F_j which converge uniformly to F on $[a, b]$ such that the functions F_j' converge to F' in $L^1[a, b]$. In the same manner as above, one verifies that (2.1) still holds. Now let us drop the assumption that μ has bounded support. Let ζ_j be smooth compactly supported functions such that $0 \le \zeta_j \le 1$, $\zeta_j(x) = 1$ if $|x| \le j$, $\zeta_j(x) = 0$ if $|x| \ge j+1$, $\sup_j \sup_x \Big[|\zeta_j'(x)|+|\zeta_j''(x)|\Big] < \infty$. Let $\mu_j := \zeta_j\mu$. Then the measures $F\mu_j$ converge to $F\mu$ in the variation norm. In addition,

$$d_1\mu_j = \zeta_j'\mu + \zeta_j d_1\mu, \quad d_1^2\mu_j = \zeta_j''\mu + 2\zeta_j' d_1\mu + \zeta_j d_1^2\mu.$$

It is readily seen from this expression that $F'd_1\mu_j \to Fd_1\mu$ and $Fd_1^2\mu_j \to Fd_1^2\mu$ in the variation norm. Thus, we arrive at (2.1) in the general case.

We can write X as a topological sum $X = \mathbb{R}^1h + Y$ for some closed hyperplane Y in X. Let ν denote the image of μ under the natural projection to Y. It is known that there exist conditional measures μ^y the lines $y + \mathbb{R}^1h$, $y \in Y$, which are twice Skorohod differentiable along h (see [1, Ch. 2]). For ν-almost every $y \in Y$, the restriction of the function F to $y + \mathbb{R}^1h$ is integrable with respect to $d_h^2\mu^y$ and the restriction of $\partial_h F$ is integrable with respect to $d_h\mu^y$. Therefore, by using the one dimensional case, we arrive at the estimate

$$\|d_h^2(F\mu)\| \le 2\|F\, d_h^2\mu\| + 2\|\partial_h F d_h\mu\|,$$

which is (2.1).

Now suppose $|F|^p, |\beta_h^\mu|^p|F| \in L^1(\mu)$ for some $p > 1$. According to Krugova's inequality [7], we have the estimate

$$\left(\int |\beta_h^{F\mu}(x)|^{2-\varepsilon}\, F(x)\, \mu(dx)\right)^{1/(2-\varepsilon)} \le (1+\varepsilon^{-1})\|d_h(F\mu)\| + \frac{1-\varepsilon}{\varepsilon}\|d_h^2(F\mu)\| \qquad (2.4)$$

for every $\varepsilon \in (0, 1)$. We observe that $\beta_h^{F\mu} = \beta_h^\mu + \partial_h F/F$ a.e. with respect to the measure $F\mu$. Since by our hypothesis $\beta_h^\mu \in L^p(F\mu)$ with some $p \in (1, 2)$, then also $\partial_h F/F \in L^p(F\mu)$. Now let $r \in (1, p)$. Set

$$s = \frac{p}{r}, \quad \alpha = \frac{p-1}{s} = r\frac{p-1}{p}.$$

Let $t = s(s-1)^{-1}$. Then $\alpha t = r(p-1)(p-r)^{-1}$. Since $p(p-r)(p-1)^{-1} \to p$ as $r \to 1$, there is $r > 1$ such that $r \le p(p-r)(p-1)^{-1}$. With this r, one has $\alpha t \le p$, hence $|F|^{\alpha t} \in L^1(\mu)$ and we obtain by Hölder's inequality

$$\int |\partial_h F|^r \, \mu(dx) = \int \frac{|\partial_h F|^r}{F^\alpha} F^\alpha \, \mu(dx) \\ \leq \Bigl(\int \frac{|\partial_h F|^{rs}}{F^{\alpha s}} \, \mu(dx)\Bigr)^{1/s} \Bigl(\int F^{\alpha t} \, \mu(dx)\Bigr)^{1/t} < \infty.$$

The proof is complete. □

Corollary 2.2. *Let μ be a centered Gaussian measure and let $h \in H(\mu)$. Suppose that F is convex along h and the functions $F(1 + |\widehat{h}|^2)$ and $\partial_h F \widehat{h}$ are in $L^1(\mu)$. Then the measure $F\mu$ is twice Skorohod differentiable along h and* (2.1) *is true. If, in addition, $F \in L^p(\mu)$ with some $p > 1$, then* (2.2) *is true.*

Proof. It suffices to recall that $\beta_h^\mu = -\widehat{h}$ and $\beta_{h,h}^\mu = |\widehat{h}|^2 - |h|_H^2$. □

As it has already been mentioned, it may happen that the pointwise second order derivative is almost everywhere zero, but the corresponding part of $d_h(F\mu)$ is nontrivial. Let us explain how $\partial_h^2 F$ can be interpreted in the generalized sense in analogy with the case of Lebesgue measure. Let us set

$$F_{hh} := d_h^2(F\mu) - 2\partial_h F d_h\mu - F d_h^2\mu,$$

provided that each of the three measures on the right exists separately. Heuristically, $F_{hh} = \partial_h^2 F\mu$, since if F is twice differentiable along h in the usual sense and $\partial_h^2 F \in L^1(\mu)$, then

$$d_h^2(F\mu) = \partial_h^2 F\mu + F d_h^2\mu + 2\partial_h F d_h\mu. \tag{2.5}$$

We know that $\partial_h^2 F$ exists μ-a.e. and is μ-integrable. However, (2.5) may fail (as it happens in the above mentioned one dimensional example). In general, F_{hh} can be regarded as a derivative of $\partial_h F$ along h in the sense of distributions over (X, μ).

Proposition 2.3. *Suppose that the hypotheses of Theorem* 2.1 *are fulfilled and that $\partial_h F \in L^1(\mu)$. Then the measure F_{hh} is finite and nonnegative. In addition,*

$$F_{hh} = d_h(\partial_h F\mu) - \partial_h F d_h\mu. \tag{2.6}$$

Proof. Let ζ be a nonnegative smooth cylindrical function. We have

$$\int \zeta(x)\, F_{hh}(dx) = \int \zeta(x) d_h^2(F\mu)(dx) - 2\int \zeta(x)\partial_h F(x)\, d_h\mu(dx) - \int \zeta(x) F(x)\, d_h^2\mu(dx) \\ = -\int \partial_h\zeta(x) d_h(F\mu)(dx) - 2\int \zeta(x)\partial_h F(x)\, d_h\mu(dx) + \int [\partial_h\zeta(x)F(x) + \zeta(x)\partial_h F(x)]\, d_h\mu(dx) \\ = -\int \partial_h\zeta(x)\partial_h F(x)\, \mu(dx) - \int \zeta(x)\partial_h F(x)\, d_h\mu(dx) \\ = -\int \partial_h F(x)\, d_h(\zeta\mu)(dx).$$

The right-hand side is nonnegative. This is verified by using the one dimensional conditional measures and noting that if ϱ is an absolutely continuous probability density on the real line and G is a convex function such that $G'\varrho'$ and $G'\varrho$ are integrable, then

$$\int G'(t)\varrho'(t)\, dt \leq 0.$$

Indeed, for any fixed a and b one has

$$\int_a^b G'(t)\varrho'(t)\,dt \le G'(b)\varrho(b) - G'(a)\varrho(a),$$

which follows by the integration by parts formula, since $\varrho \ge 0$ and the measure G'' is nonnegative. The integrability of $G'\varrho$ enables us to pick $a \to -\infty$ and $b \to +\infty$ in such a way that $G'(b)\varrho(b)$ and $G'(a)\varrho(a)$ tend to zero. Equality (2.6) is seen from the above chain of equalities. □

Proposition 2.4. *Let Ψ be a μ-integrable μ-a.e. finite nonnegative function such that the sets $\{\Psi \le c\}$ have compact closure, let $h \in X$, and let F_n, $n \in \mathbb{N}$, be μ-measurable functions such that the sequence $\{F_n\}$ is bounded in $L^p(\mu)$ for some $p > 1$. Assume that the functions F_n are convex and twice differentiable along h on almost all lines parallel to h and*

$$\partial_h F_n \beta_h^\mu,\, F_n\beta_h^\mu,\, F_n\beta_{h,h}^\mu \in L^1(\mu).$$

Suppose also that the measure $d_h^2(\Psi\mu)$ has a density $g \in L^{p'}(\mu)$ with respect to μ. Then the sequence of measures $(\partial_h^2 F_n)\mu$ is uniformly tight.

Proof. According to Theorem 2.1 the measures $F_n\mu$ are twice Skorohod differentiable along h. Since the measures $\partial_h F d_h\mu$ and $F d_h^2\mu$ have bounded variations, the functions $\partial_h^2 F_n$ are μ-integrable. Therefore,

$$\int \partial_h^2 F_n(x)\,\Psi(x)\,\mu(dx) = \int F_n(x)\,d_h^2(\Psi\mu)(dx)$$
$$= \int F_n(x) g(x)\,\mu(dx) \le \|\Psi\|_{L^p(\mu)}\|g\|_{L^{p'}(\mu)}.$$

As $\partial_h^2 F_n(x) \ge 0$ a.e., the integrals on the left are uniformly bounded, hence the sequence of measures $(\partial_h^2 F_n)\mu$ is uniformly tight. □

Corollary 2.5. *Let μ be a centered Radon Gaussian measure on a sequentially complete locally convex space X and let H be the Cameron–Martin space of μ. Let $F \in L^p(\mu)$, where $p > 1$, be an H-convex function. Then, for any $h \in H$, the measures $(\partial_h^2 T_\varepsilon f)\mu$, $\varepsilon \in (0,1)$, are uniformly tight and converge weakly to F_{hh} as $\varepsilon \to 0$.*

Proof. We shall construct a nonnegative function Ψ that is finite on a linear space of full measure such that the sets $\{\Psi \le c\}$ have compact closure and the functions $\Psi, \partial_h\Psi, \partial_h^2\Psi$ belong to all $L^p(\mu)$. It follows by our hypotheses that there exists a balanced convex compact set K of positive μ-measure. The Minkowski functional q_K of K is defined by the formula $q_K(x) = \inf\{t > 0\colon\ t^{-1}x \in K\}$ on the linear span E_K of K and $q_K(x) = +\infty$ if $x \notin E_K$. The function q_K is H-Lipschitzian and belongs to all $L^p(\mu)$. Clearly, the sets $\{q_K \le c\} = cK$ are compact. However, this function may not be sufficiently differentiable. Let us consider the function $\Psi = T_1 q_K$. This function is infinitely differentiable along H and all its partial derivatives along directions from H are in all $L^p(\mu)$. In addition, any set $\{\Psi \le c\}$ has compact closure, because it is contained in the set $(6c + m)K$, where $m > 0$ is such that $\mu(mK) \ge 1/2$. □

We recall that a countably additive measure m on a measurable space $(X, \mathcal{B})$ with values in a normed space E is said to have bounded semivariation if

$$\|m\|_E := \sup\{\|l(m)\|\colon\ l \in E^*, \|l\| \le 1\} < \infty,$$

where $\|l(m)\|$ is the usual total variation of the scalar measure $l(m)$. It is easily seen that

$$\sup_{B\in\mathcal{B}} |m(B)|_E \le \|m\|_E \le 2\sup_{B\in\mathcal{B}} |m(B)|_E.$$

The total variation of m is defined as

$$v(m) := \sup\Bigl\{\sum_{i=1}^n |m(B_i)|_E\Bigr\},$$

where the supremum is taken over all finite partitions of X into disjoint sets $B_i \in \mathcal{B}$. If E is infinite dimensional, then m may have finite semivariation, but infinite total variation. We recall that if m is a countably additive measure of bounded total variation with values in a Hilbert space E, then there exist a probability measure μ and a μ-integrable mapping $f\colon\ X \to E$ such that $m = f\mu$, i.e.

$$m(B) = \int_B f\, d\mu, \quad \forall\, B \in \mathcal{B}.$$

Corollary 2.6. *Let μ be a centered Gaussian measure on X with the Cameron–Martin space H and let $F \in L^2(\mu)$ be H-convex. Then the formula*

$$(T_B h, h)_H := F_{hh}(B), \quad B \in \mathcal{B}(X),$$

defines a countably additive measure $B \mapsto T_B$ with values in the space HS of all Hilbert–Schmidt operators on H equipped with the Hilbert–Schmidt norm. In addition, this measure has bounded semivariation such that

$$\|T_B\|_{HS} \le \|F\|_{L^2(\mu)}.$$

Proof. Suppose first that $F \in W^{2,2}(\mu)$. Then the second derivative $D_H^2 F(x)$ is a nonnegative Hilbert–Schmidt operator on H. Let $B \in \mathcal{B}(X)$. Then the formula

$$(T_B h, h)_H = \int_B D_H^2 F(x)(h,h)\, \mu(dx)$$

defines a nonnegative Hilbert–Schmidt operator whose Hilbert–Schmidt norm is majorized by $\int \|D_H^2 F(x)\|_{HS}^2\, \mu(dx)$. Let $\{e_j\}$ be an orthonormal basis such that $T_B e_j = t_j e_j$. We observe that the functions

$$\xi_j := \partial_{e_j}\beta^\mu_{e_j} + |\beta^\mu_{e_j}|^2 = -1 + |\widehat{e_j}|^2$$

are mutully orthogonal in $L^2(\mu)$ and have equal norms in $L^2(\mu)$. Since the functions $\partial_{e_j}^2 F$ are nonnegative, the Hilbert–Schmidt norm of T_B can be estimated as follows:

$$\begin{aligned}
\sum_{j=1}^\infty t_j^2 &= \sum_{j=1}^\infty \Bigl|\int_B \partial_{e_j}^2 F(x)\,\mu(dx)\Bigr|^2 \\
&\le \sum_{j=1}^\infty \Bigl|\int_X \partial_{e_j}^2 F(x)\,\mu(dx)\Bigr|^2 = \sum_{j=1}^\infty \Bigl|\int_X F(x)\xi_j(x)\,\mu(dx)\Bigr|^2 \\
&\le \int_X |F(x)|^2\,\mu(dx).
\end{aligned}$$

Therefore, the claim is true for the functions $T_{1/k}F$. Letting $k \to \infty$ and noting that $T_{1/k}F \to F$ in $L^2(\mu)$, we obtain the claim by the previous corollary. ☐

It is worth noting that analogous results hold for a broader class of a-convex functions considered in [6]. To be more specific, recall that given a centered Radon Gaussian measure μ, a μ-measurable function f with values in the extended real line is said to be a-convex along a vector $h \in H = H(\mu)$, where $a \in \mathbb{R}^1$, if f is finite a.e. and the function $f + \frac{a}{2}\widehat{h}^2$ is convex on the lines $x + \mathbb{R}^1 h$. It is clear from the proofs that some of the above results extend also to such functions.

Acknowledgements. This work was started during a visit of the first author to the University of New South Wales. Our research was supported by the RFBR project 04-01-00748, the DFG Grant 436 RUS 113/343/0(R), the INTAS project 03-51-5018, the DFG–Forschergruppe "Spectral Analysis, Asymptotic Distributions, and Stochastic Dynamics", and the University of New South Wales.

References

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

[2] Bogachev V.I., Gaussian measures, Amer. Math. Soc., Rhode Island, 1998.

[3] Bogachev V.I., Mayer-Wolf E., Some remarks on Rademacher's theorem in infinite dimensions, Potential Anal. 5 (1996), no. 1, 23–30.

[4] Borwein J.M., Noll D., Second order differentiability of convex functions in Banach spaces, Trans. Amer. Math. Soc. 342 (1994), 43–82.

[5] Evans L.C., Gariepy R.F., Measure theory and fine properties of functions. CRC Press, Boca Raton, Florida, 1992.

[6] Feyel D., Üstünel A.S., The notion of convexity and concavity on Wiener space, J. Funct. Anal. 176 (2000), 400–428.

[7] Krugova E.P., On the integrability of logarithmic derivatives of measures, Mat. Zametki 53 (1993), no. 5, 76–86 (in Russian); English transl.: Math. Notes 53 (1993), no. 5, 506–512.

[8] Krugova E.P., On shifts of convex measures, Mat. Sb. 188 (1997), no. 2, 57–66 (in Russina); English transl.: Sbornik Math. 188 (1997), no. 2, 227–236.

[9] Matoušek J., Matouškovà E., A highly non-smooth norm on Hilbert space, Israel J. Math. 112 (1999), 1–27.

[10] Matoušková E., Zajiček L., Second order differentiability and Lipschitz smooth points of convex functionals, Czech. Math. J. 48 (1998), 617–640.

[11] Üstünel A.S., Zakai M., Transformation of measure on Wiener space. Springer-Verlag, Berlin, 2000.

Department of Mechanics and Mathematics, Moscow State University, 119992 Moscow, Russia

School of Mathematics, University of New South Wales, Sydney 2052, Australia
E-mail address: beng@maths.unsw.edu.au