On the convergence in the variation norm for the images of measures under differentiable mappings

D. E. Alexandrova[1], V. I. Bogachev[2], A. I. Pilipenko[3]

[1], [2]: Department of Mechanics and Mathematics,
Moscow State University, 119899 Moscow, Russia

[3]: Department of Mathematics, Kiev University,
Kiev 252601, Ukraine

Abstract

Let F_j, $F\colon \mathbf{R}^n \to \mathbf{R}^n$ be measurable mappings such that $F_j \to F$ and $\partial_{x_i} F_j \to \partial_{x_i} F$ in measure on a measurable set E. We give conditions ensuring that the images of Lebesgue measure $\lambda|_E$ on E under the mappings F_j converge in the variation norm to the image of $\lambda|_E$ under F. For example, a sufficient condition is the convergence of F_j to F in the Sobolev space $W^{p,1}(\mathbf{R}^n, \mathbf{R}^n)$ with $p > n$ and the inclusion $E \subset \{\det DF \neq 0\}$. Analogous results are obtained for mappings between Riemannian manifolds and mappings from infinite dimensional spaces.

AMS 1991 Subject classification:
Primary: 28A33, 60B10
Secondary: 46E35, 46G12, 60B11, 60E05
Keywords and phrases: convergence in variation, images of measures under differentiable mappings, Sobolev classes
Running head: Convergence in variation

In this work we investigate the following problem. Let ξ_j be a sequence of random vectors in $\mathbf{R}^n$ convergent in probability to a random vector ξ. When do the distributions of the vectors ξ_j converge to the distribution of ξ in the variation norm? This problem arises in the study of the limit theorems in probability theory (see, e.g., [7]); it is of independent interest as well. A simple and convenient sufficient condition for the convergence in variation has been obtained by Yu. A. Davydov in [8], where

$\{\xi_j\}$ are absolutely continuous functions on the interval $[0,1]$, convergent uniformly to an absolutely continuous function ξ such that $\|\xi_j' - \xi'\|_{L^1[0,1]} \to 0$. Under these assumptions, the restriction $\lambda|_E$ of Lebesgue measure to the set $E = \{\xi' \neq 0\}$ is transformed by the functions ξ_j to measures that converge in the variation norm to the image of the measure $\lambda|_E$ under the mapping ξ. Below we give a simple proof of a considerably more general result. In particular, the above assertion is true if ξ_j and ξ are measurable mappings in $\mathbf{R}^n$ with Lusin's property (N) such that they have derivatives (or at least regular approximate derivatives) $D\xi_j$ and $D\xi$ almost everywhere, these deriavtives are locally uniformly integrable, the mappings ξ_j converge uniformly on compact sets to ξ, and the mappings $D\xi_j$ converge in measure on a bounded measurable set E to the mapping $D\xi$, which is nondegenerate on E. Thus, also in the one dimensional case the result of the work [8] is reinforced. Analogous results are obtained for mappings between finite dimensional Riemannian manifolds, where the nondegeneracy of the derivative $D\xi$ of the limiting mapping is replaced by the surjectivity of $D\xi$. In addition, we obtain corollaries of the main result for mappings from infinite dimensional spaces, applicable to wide classes of measures, in particular, to Gaussian measures and to Gibbs distributions.

1 Notation and terminology

Let us recall some definitions. For any open set $\Omega \subset \mathbf{R}^d$ and any number $p \geq 1$, the Sobolev space $W^{p,1}(\Omega)$ consists of the functions $f \in L^p(\Omega)$, whose generalized partial derivatives are also elements of $L^p(\Omega)$. The Sobolev space $W^{p,1}(\Omega, \mathbf{R}^n)$ consists of the mappings $f = (f_1, \dots, f_n)\colon\ \Omega \to \mathbf{R}^n$ such that $f_i \in W^{p,1}(\Omega)$. This space is equipped with the norm

$$\|f\|_{p,1} = \left\| \Bigl(\sum_{i=1}^n f_i^2 \Bigr)^{1/2} \right\|_{L^p(\Omega)} + \left\| \Bigl(\sum_{i\leq d,\, j\leq n} (\partial_{x_i} f_j)^2 \Bigr)^{1/2} \right\|_{L^p(\Omega)}.$$

The symbol $W^{p,1}_{loc}(\Omega, \mathbf{R}^n)$ stands for the class of all mappings $f\colon\ \Omega \to \mathbf{R}^n$ such that $\zeta f \in W^{p,1}(\Omega, \mathbf{R}^n)$ for all $\zeta \in C_0^\infty(\Omega)$, where $C_0^\infty(\Omega)$ is the class of all smooth functions with compact supports in Ω.

Let $A \subset \mathbf{R}^d$ and let $x \in A$. If

$$\lim_{r\to 0} \frac{\lambda\bigl(K(x,r) \cap A\bigr)}{\lambda\bigl(K(x,r)\bigr)} = 0,$$

where $K(x,r)$ is the closed ball of radius $r > 0$ centered at x, then the point x is called a point of density 0 for the set A. A point $y \in Y$ is called the approximate limit of the mapping $f\colon\ A \to \mathbf{R}^n$ at the point x if the density of the set $\mathbf{R}^d \setminus f^{-1}(U)$ at x equals 0 for every neighborhood U of the point $y \in Y$. Clearly, there might be at most one such y; we denote it by $\operatorname{ap}\lim_{z\to x} f(z)$. If f is a real function, then its approximate upper limit $\operatorname{ap}\limsup\limits_{x\to a} f(x)$ at the point a is defined as the infimum of all numbers t such that the point a is a point of density 0 for the set $\{z\colon\ f(z) > t\}$.

A mapping $f\colon\ A \to \mathbf{R}^n$ is called approximately differentiable at a point a if there exists a linear mapping $L\colon\ \mathbf{R}^d \to \mathbf{R}^n$ such that

$$\operatorname{ap}\lim_{x\to a} \|f(x) - f(a) - L(x-a)\|/\|x-a\| = 0.$$

It follows that $\mathbf{R}^d \setminus A$ has density 0 at the point a and, in addition, that such L is unique (see [10, 3.1.2]). The operator L is denoted by $\operatorname{ap}Df(a)$ and called the approximate derivative of f at the point a.

The approximate Jacobian $J(f)$ of the mapping f is defined as the determinant of the approximate derivative of f.

Let us define also the approximate partial derivatives of a mapping $f\colon\ A \to \mathbf{R}^n$ by the formula

$$\operatorname{apD}_i f(a) = \operatorname{ap}\lim_{t\to 0} [f(\ldots, a_{i-1}, a_i + t, a_{i+1}, \ldots) - f(a)]/t.$$

The almost everywhere existence of the approximate partial derivatives of a measurable mapping $f\colon\ A \to \mathbf{R}^n$ implies its approximate differentiability almost everywhere on A (see [10, Theorem 3.1.4]), whence, in turn, it follows that

$$\operatorname{ap}\limsup_{x\to a} \|f(x) - f(a)\|/\|x-a\| < \infty \tag{1.1}$$

for almost all $a \in A$. The usual differentiability of f at a point implies the approximate differentiability at this point, and, in addition, the usual derivative Df serves in this case as the approximate one. Unlike the usual derivative, in the definition of the approximate derivative one does not need that the mapping f be defined in a neighborhood of the point x. We recall that every mapping $f \in W^{1,1}_{loc}(\mathbf{R}^d, \mathbf{R}^n)$ has a version that possesses almost everywhere the partial derivatives; hence this version is approximately differentiable almost everywhere.

If a mapping f from an open set $U \subset \mathbf{R}^d$ to $\mathbf{R}^n$ has the approximate derivative L at a point $x \in U$ and there exists a sequence $r_j \to 0$ such that

$$\lim_{j\to\infty} \sup_{\|h\|=1} \left\| \frac{f(x + r_j h) - f(x)}{r_j} - Lh \right\| = 0,$$

then L is called the regular approximate derivative of f at the point x (in this case L is called also the K-derivative and the weak total differential). Clearly, the usual differentiability at a point implies the existence of the regular approximate derivative at this point and its coincidence with the usual derivative. It is known that every mapping $f \in W^{p,1}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$ with $p > n-1$ has a modification $\widetilde{f}$ (i.e., $f = \widetilde{f}$ a.e.) possessing almost everywhere the regular approximate derivative, which coincides almost everywhere with the Sobolev derivative of f (see [20, §2], [25, p. 170]). If $p > n$, then f has a continuous modification, which almost everywhere possesses the usual derivative (see [12, Ch. 2, §5.4, Theorem 5.2]).

Let us recall that a mapping $f\colon \mathbf{R}^n \to \mathbf{R}^n$ satisfies Lusin's condition (N) (or has Lusin's (N)-property) if it takes every measure zero set to a measure zero set. According to the Rademacher–Ellis theorem, a measurable mapping f has Lusin's (N)-property precisely when it takes every measurable set to a measurable one (see [9] or [2, Problem 6.8.7], where the formulation is given for Borel f, which implies the claim also for measurable mappings). It is known that the continuous modification of any mapping $F \in W^{p,1}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$ with $p > n$ (such a modification exists as noted above) satisfies Lusin's condition (N) (see [4, Lemma 8.1]). It is shown in [17], that for every $n > 1$, there exists a homeomorphism $h\colon \mathbf{R}^n \to \mathbf{R}^n$ such that $h \in W^{p,1}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$ for all $p < n$, but h does not satisfy Lusin's condition (N). However, every homeomorphism $h \in W^{n,1}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$ satisfies Lusin's condition (N) (see [21, Ch. II, §2, Corollary 1 of Theorem 6.2]; for more general results, see [22]).

Let μ be a measure on a measurable space $(X, \mathcal{B})$. We denote by $\mathcal{B}_\mu$ the Lebesgue completion of $\mathcal{B}$ with respect to μ. If f is a μ-measurable mapping with values in a measurable space $(Y, \mathcal{F})$, i.e., $f^{-1}(A) \in \mathcal{B}_\mu$ for any $A \in \mathcal{F}$, then the image of the measure μ under the mapping f is denoted by $\mu \circ f^{-1}$ and is defined by the formula

$$\mu \circ f^{-1}(A) = \mu\big(f^{-1}(A)\big), \qquad A \in \mathcal{F}.$$

We observe that the induced measure $\mu \circ f^{-1}$ does not change if we replace f by its modification.

The restriction of a measure μ to a measurable set E is denoted by $\mu|_E$. Lebesgue measure is denoted by λ.

2 Main results

Theorem 2.1. *Let $F\colon \mathbf{R}^n \to \mathbf{R}^n$ and $F_j\colon \mathbf{R}^n \to \mathbf{R}^n$ be measurable mappings, let λ be Lebesgue measure on $\mathbf{R}^n$, and let E be a measurable set of finite Lebesgue measure in $\mathbf{R}^n$. Suppose that at every point $x \in E$, there exist the approximate partial derivatives $\mathrm{ap}D_iF(x)$ and $\mathrm{ap}D_iF_j(x)$, $i = 1, \dots, n$, the mappings F_j converge in measure on the set E to F, and their approximate partial derivatives $\mathrm{ap}D_iF_j$, $i = 1, \dots, n$, converge in measure on the set E to the corresponding approximate partial derivatives of F. Assume, in addition, that the approximate Jacobian $J(F)$ of the mapping F does not vanish on the E. Then the following conditions are equivalent:*

(i) *for every measurable set* $A \subset E$, *the measures* $\lambda|_A \circ F_j^{-1}$ *converge in the variation norm to the measure* $\lambda|_A \circ F^{-1}$;

(ii) *for every measurable set* $A \subset E$ *and every* $\delta > 0$, *there exists a compact set* $K_\delta \subset A$ *such that* $\lambda(A \backslash K_\delta) \leq \delta$ *and* $\lim\limits_{j\to\infty} \lambda\big(F_j(K_\delta)\big) = \lambda\big(F(K_\delta)\big)$.

Proof. Assume that condition (ii) is fulfilled. It suffices to establish that every subsequence $\{G_j\}$ in $\{F_j\}$ contains a subsequence $\{G_{j_k}\}$, for which our claim is true. Hence it suffices to find such a subsequence in $\{F_j\}$. In particular, by the Riesz theorem, one can assume that the sequence $\{F_j\}$ converges together with the approximate partial derivatives almost everywhere on A. In addition, in the process of the proof we shall repeatedly employ the following simple observation: it suffices to show that for every $\varepsilon > 0$, there exists a compact set $A_\varepsilon \subset A$ such that $\lambda(A \backslash A_\varepsilon) < \varepsilon$ and the claim to be proven is true for A_ε.

Note now that the theorem reduces to the case where the mappings F_j and F are the restrictions to A of continuously differentiable mappings. Indeed, according to [10, Theorem 3.1.16], if $f\colon A \to \mathbf{R}^n$ satisfies condition (1.1) for almost every $a \in A$, then for every number $\varepsilon > 0$, there exists a continuously differentiable mapping $g\colon \mathbf{R}^n \to \mathbf{R}^n$ such that

$$\lambda\big(x \in A\colon\ f(x) \neq g(x)\big) < \varepsilon.$$

Given $\varepsilon > 0$, this theorem is applied to every mapping F_j and the number $\varepsilon 4^{-j}$, and also to F and the number $\varepsilon/2$. Moreover, by virtue of Egorov's theorem and the observation made above, the claim reduces to the case, where E is compact, the mappings F_j converge to F uniformly on E, and their derivatives converge uniformly on E to the derivative of F. By the inverse function theorem, every point in E has a neighborhood, in which F is a diffeomorphism. Hence E can be partitioned into finitely many disjoint measurable subsets possessing neighborhoods, in which F is a diffeomorphism. Since the measure $\lambda|_E$ equals the sum of the restrictions of Lebesgue measure to those subsets, everything reduces to the case where F is a diffeomorphism in some neighborhood U of the compact set E. One can assume that F is the identity mapping, since the claim to be proven is invariant with respect to diffeomorphisms. Making use of the uniform convergence of $\det DF_j$ to 1 on E and passing to a subsequence once again, we may assume that for all $x \in E$ and all j one has the inequality $|\det DF_j(x) - 1| < 2^{-j}$. According to [10, Corollary 3.2.4], for every j, there exists a compact set $Q_j \subset E$ such that F_j is injective on Q_j and

$$\lambda\big(F_j(E) \backslash F_j(Q_j)\big) < 2^{-j}. \tag{2.2}$$

Making use of condition (ii), by an arbitrary small decrease in measure of E one can pass to the situation where there holds the equality

$$\lim_{j\to\infty} \lambda\big(F_j(E)\big) = \lambda\big(F(E)\big) = \lambda(E). \tag{2.3}$$

Let us take a subsequence F_{j_k} such that

$$\lambda\Big(F_{j_k}(E) \bigtriangleup E\Big) < \frac{1}{2^k}. \tag{2.4}$$

Such a subsequence exists by virtue of (2.3) and the uniform convergence of the F_j's to F on E, by which for any open set $U \supset E$ with $\lambda(U \backslash E) < 4^{-k}$ there exists a number n_k with $F_j(E) \subset U$ whenever $j \geq n_k$. We may assume further on that $j_k = k$. According to [10, Theorem 3.2.3], for any measurable set A, on which F_j is injective and $\det DF_j > 0$, there holds the equality

$$\lambda\big(F_j(A)\big) = \int\limits_A \det DF_j(x)\, dx. \tag{2.5}$$

Then from (2.5) with $A = Q_j$, (2.2), and (2.4) we obtain

$$\begin{aligned} \lambda(Q_j) &\geq \lambda\big(F_j(Q_j)\big) - 2^{-j}\lambda(E) \geq \lambda\big(F_j(E)\big) - 2^{-j} - 2^{-j}\lambda(E) \\ &\geq \lambda(E) - 2^{1-j} - 2^{-j}\lambda(E) = \lambda(E) - 2^{-j}\big(2 + \lambda(E)\big). \end{aligned}$$

The sets $E_N = \bigcap_{j=N}^{\infty} Q_j$ are compact, the mappings F_j with $j \geq N$ are injective on E_N, and by the previous estimate one has $\lambda(E_N) \to \lambda(E)$, since

$$\lambda(E \backslash E_N) \leq \sum_{j=N}^{\infty} \lambda(E \backslash Q_j) \leq \big(2 + \lambda(E)\big) \sum_{j=N}^{\infty} 2^{-j}.$$

Therefore, it suffices to prove our claim for every E_N. Thus, we shall suppose that E coincides with one of the E_N's. Passing to a subsequence once again and replacing E_N by its compact part of close measure, we arrive as above to relationship (2.4) with $j_k = k$ for our new set E. Let us denote by S the set of all $y \in E$, for which there exists a number $m = m(y)$ with

$$y \in \bigcap_{j \geq m} F_j(E).$$

By (2.4) we have $\lambda(S) = \lambda(E)$, since

$$\lambda\Big(E \bigtriangleup \bigcap_{j=m}^{\infty} F_j(E)\Big) \leq \sum_{j=m}^{\infty} 2^{-j} = 2^{-m+1}.$$

For completing the proof it remains to show that the densities ϱ_j of the induced measures $\lambda|_E \circ F_j^{-1}$ converge almost everywhere on E to the density $\varrho = I_E$ of the measure $\lambda|_E \circ F^{-1}$ (see [23, Theorem 3 §10 Ch. II]). Note that

$$\varrho_j(y) = \frac{I_{F_j(E)}(y)}{\det DF_j\big(F_j^{-1}(y)\big)}.$$

Indeed, by the univalence of F_j on E, for any bounded measurable function φ, we have by the change of variable formula (see [10, Theorem 3.2.5]) that

$$\int_{\mathbf{R}^n} \varphi(y)\,\lambda|_E \circ F_j^{-1}(dy) = \int_E \varphi\big(F_j(x)\big)\,dx = \int_{F_j(E)} \varphi(y)\frac{1}{\det DF_j\big(F_j^{-1}(y)\big)}\,dy.$$

Let $y \in S$. Then $F_j^{-1}(y) \to F^{-1}(y) = y$. Indeed, otherwise some subsequence $F_{j_k}^{-1}(y)$ of the sequence $F_j^{-1}(y)$ in the compact set E converges to a point $x \neq y$, whence by the uniform convergence of F_j to F we get $y = F_{j_k}\big(F_{j_k}^{-1}(y)\big) \to F(x)$, which contradicts the univalence of F. Further, by the uniform convergence of derivatives we obtain

$$\det DF_j\big(F_j^{-1}(y)\big) \to \det DF\big(F^{-1}(y)\big) = 1.$$

If $y \notin E$, then for all j sufficiently large, we have $I_{F_j(E)}(y) = 0$, so that we obtain $\varrho_j(y) \to 0$.

Let now condition (i) be satisfied. It is clear from the reasoning above that the general case reduces to the one where $A \subset E$ is compact, F is a continuously differentiable mapping in a neighborhood of A, the F_j's are continuous on A, and $F_j \to F$ uniformly on A. Let $\varepsilon > 0$ and let U be a neighborhood of $F(A)$ such that $\lambda\big(U\backslash F(A)\big) < \varepsilon$. Then for all sufficiently large j there holds the relationship $F_j(A) \subset U$, whence $\lambda\big(F_j(A)\backslash F(A)\big) < \varepsilon$. On the other hand, $\lambda\big(F(A)\backslash F_j(A)\big) \to 0$ as $j \to \infty$. Indeed, the set $X_j = F(A)\backslash F_j(A)$ has measure zero with respect to $\lambda|_A \circ F_j^{-1}$, and its measure with respect to $\lambda|_A \circ F^{-1}$ is estimated from below by $c\lambda(X_j)$, since it follows from the proof of the first part of this theorem that the density of the measure $\lambda|_A \circ F^{-1}$ on $F(A)$ is uniformly separated from zero. □

Remark 2.2. (i) We observe that under the hypotheses of the theorem proven the measure $\lambda|_E \circ F^{-1}$ is absolutely continuous (this is seen from the proof), hence the absolutely continuous components of the measures $\lambda|_E \circ F_j^{-1}$ converge in variation to $\lambda|_E \circ F^{-1}$, and their singular components converge in variation to zero.

(ii) Note that the almost sure nondegeneracy of apDF on E is also necessary for the absolute continuity of the induced measure $\lambda|_E \circ F^{-1}$ (see, e.g., [16, Theorem 3 and the remark following Theorem 4]).

(iii) It is seen from the above proof that if the mappings F_j are injective on E, then condition (i) in Theorem 2.1 (hence also condition (ii)) follows from the hypotheses of that theorem. Indeed, in this case one has at once the above expression for ϱ_j. However, in the general case, the equivalent conditions (i) and (ii) do not follow automatically from the other hypotheses of Theorem 2.1. Let us consider the following simple example. Let $E \subset [0,1]$ be a Cantor-type set (nowhere dense compact) of positive Lebesgue measure and let $F(x) = x$. For every natural number

k, we partition $[0,1]$ in the intervals $I_{k,j}$ by the points of the form j/k. Inside every interval $I_{k,j}$ that meets E in a positive measure set, we take two disjoint closed intervals $\Delta^i_{k,j} = [a^i_{k,j}, b^i_{k,j}]$, $i = 1, 2$, such that

$$\lambda\Big(E \cap \Delta^1_{k,j}\Big) = \lambda\Big(E \cap \Delta^2_{k,j}\Big) \geq \lambda\Big(E \cap I_{k,j}\Big)/3.$$

We may assume that $\Delta^2_{k,j}$ borders from the left and from the right with certain intervals $L_{k,j}$ and $R_{k,j}$ from the complement to E. Now the function F_k is defined as follows: $F_k(x) = x$ if $x \notin L_{k,j} \cup \Delta^2_{k,j} \cup R_{k,j}$, $F_k(x) = x - c_{k,j}$ if $x \in \Delta^2_{k,j}$, where $c_{k,j} = b^2_{k,j} - b^1_{k,j}$, and on the sets $L_{k,j}$ and $R_{k,j}$ we define F_k making it a smooth function on $[0,1]$. Then $F'_k = 1$ on E and $F_k \to F$ uniformly on E (as it is seen from the construction, one can get the uniform convergence on all of $[0,1]$), but the images of Lebesgue measure on E with respect to F_k do not converge to the image of Lebesgue measure on E with respect to F. This is clear from the fact that $\lambda\big(F_k(E) \triangle E\big) \geq \lambda(E)/3$.

Note also that condition (ii) does not imply that $\lambda\big(F_j(E)\big) \to \lambda\big(F(E)\big)$. For example, let E be the classical Cantor set of measure zero, let $F(x) = 0$, and let $F_j = G$, $j \geq 1$, where G is an arbitrary function with $G(E) = [0,1]$. Then $\lambda|_E \circ F^{-1} = \lambda|_E \circ F_j^{-1} = 0$ for all j, and all hypotheses of the above theorem are fulfilled, since $\lambda(E) = 0$.

Let us recall that if one does not wish to deal with the approximate derivatives, then, in the formulation of the theorem just proven, one can replace them by the usual partial derivatives (although then one gets a somewhat weaker statement). For example, the theorem applies to the mappings from the class $W^{1,1}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$ that meet the required convergence conditions on E.

Corollary 2.3. *Let μ be an absolutely continuous probability measure on $\mathbf{R}^n$. Assume that in the situation of Theorem* 2.1*, where E may now have infinite Lebesgue measure, either of the two equivalent conditions* (i) *or* (ii) *is fulfilled. Then the measures $\mu|_E \circ F_j^{-1}$ converge in the variation norm to the measure $\mu|_E \circ F^{-1}$.*

Proof. Let $\mu = \varrho^\mu\, dx$. By Theorem 2.1, the statement is true if ϱ^μ assumes only finitely many values. Then the general case readily follows by approximating μ in the variation norm by measures μ_i with densities of the above type and noting that $\|\mu_i|_E \circ F_j^{-1} - \mu|_E \circ F_j^{-1}\| \to 0$ as $i \to \infty$ uniformly in j. □

Let us now give conditions that are sufficient for the validity of (ii). These conditions are not necessary, but for practical verification might be more useful.

Corollary 2.4. *Let continuous mappings $F_j\colon \mathbf{R}^n \to \mathbf{R}^n$ converge uniformly on compact sets to a continuous mapping $F\colon \mathbf{R}^n \to \mathbf{R}^n$ and let the F_j's and F have*

Lusin's (N)*-property and almost everywhere possess the regular approximate derivatives (e.g., just usual derivatives) DF_j and DF such that the DF_j's converge in measure on some set E of finite Lebesgue measure to DF. Assume that* $\det DF \neq 0$ *on E and that on every compact set the sequence $\{|\det DF_j|\}$ is uniformly integrable. Then the measures $\lambda|_E \circ F_j^{-1}$ converge in the variation norm to the measure $\lambda|_E \circ F^{-1}$. In addition, if μ is an absolutely continuous probability measure on $\mathbf{R}^n$, then the measures $\mu|_E \circ F_j^{-1}$ converge in the variation norm to the measure $\mu|_E \circ F^{-1}$ (this is true also for E of infinite Lebesgue measure).*

Proof. It follows from the uniform integrability of the functions $\{\det DF_j\}$ on compact sets and the estimate (see [12, Ch. 5, §1.2, Theorem 1.4])

$$\lambda\big(F_j(A)\big) \leq \int\limits_A |\det DF_j(x)|\, dx \tag{2.6}$$

that $\sup\limits_j \lambda\big(F_j(A)\big) \to 0$ as $\lambda(A) \to 0$, where A is a bounded measurable set. Let us verify that $\lim\limits_{j\to\infty} \lambda\big(F_j(K)\big) = \lambda\big(F(K)\big)$ for every compact set $K \subset E$. It is seen from the proof of the previous theorem that it suffices to verify our claim for the compact sets K, on which F is the restriction of a diffeomorphism in some neighborhood of K. Therefore, as above, we may assume that F on K coincides with the identity mapping. In addition, we may assume that the mappings F_j have the regular approximate derivatives at every point of K. Let $\varepsilon > 0$ and $4\varepsilon < \lambda(K)$. It is clear that for all sufficiently large j there holds the estimate $\lambda\big(F_j(K)\big) \leq \lambda\big(F(K)\big) + \varepsilon$. Suppose that $\lambda\big(F_j(K)\big) \leq \lambda\big(F(K)\big) - \varepsilon$ for all j. There exists a connected set Q that is a finite union of closed cubes such that $K \subset Q$ and $\lambda\big(F_j(Q)\big) \leq \lambda\big(F_j(K)\big) + \varepsilon/8$ for all F_j, and also for F in place of F_j. Let ∂Q be the boundary of Q. Then $F(\partial Q)$ has measure zero. Let us take a sufficiently small closed neighborhood G of the compact set $F(\partial Q)$ such that $\lambda(G) < \varepsilon/8$ and $F_j(\partial Q) \subset G$ for all sufficiently large j. One can take for G a finite union of closed cubes. Then $\mathbf{R}^n \backslash G$ has finitely many connected components $W_1, \dots, W_m$, in which the restrictions of F_j and F to Q have the constant degrees of a map $\deg_Q F_j$ and $\deg_Q F$ (see [12, Ch. 5, §1.5]), and by the uniform convergence of F_j to F we have for all sufficiently large j that

$$\deg_Q F_j(y) = \deg_Q F(y) = d_i \qquad \text{for all } y \in W_i.$$

If $y \in Y := F(K)\backslash\Big(G \cup F(Q\backslash K)\Big)$, then $\deg_Q F(y) = 1$. This follows from the fact that for any such point y, there exists a ball $K(y,r)$ with $\deg_{K(y,r)} F(y) = 1$ (see [18, V.2.2, p. 329]). Therefore, $\deg_Q F_j(y) = 1$ for all $y \in Y$ and all sufficiently large j. Since $|\deg_Q F_j(y)|$ does not exceed the cardinality of $F_j^{-1}(y) \cap Q$ (see [12, Ch. 5,

§1.5, Remark]), one has $Y \subset F_j(Q)$, whence

$$\lambda\big(F_j(Q)\big) \geq \lambda\big(F(K)\big) - \lambda(G) - \lambda\big(F(Q\backslash K)\big) \geq \lambda\big(F(K)\big) - \varepsilon/4$$

which is a contradiction. The assertion for μ in place of λ follows by the same reasoning as in Corollary 2.3. $\square$

Corollary 2.5. *In the situation of Corollary* 2.4*, one has*

$$\lim_{j\to\infty} \lambda\big(F_j(A) \triangle F(A)\big) = 0 \qquad \textit{and} \qquad \lim_{j\to\infty} \lambda\big(F_j(A)\big) = \lambda\big(F(A)\big) \tag{2.7}$$

for any measurable set $A \subset E$. The same is true for μ in place of λ.

Proof. Suppose first that A is compact. Then the second equality in (2.7) has been established in the proof of Corollary 2.4. Let $\varepsilon > 0$ and let V be an open set such that $F(A) \subset V$ and $\lambda\big(V\backslash F(A)\big) < \varepsilon$. Then for all j sufficiently large, one has $F_j(A) \subset V$ and

$$\lambda\big(F_j(A)\big) > \lambda\big(F(A)\big) - \varepsilon \geq \lambda(V) - 2\varepsilon.$$

Hence $\lambda\big(F_j(A) \triangle F(A)\big) < 3\varepsilon$, so that we obtain the first equality in (2.7). Both equalities extend to arbitrary measurable $A \subset E$ by (2.6) and the uniform integrability of $\{\det DF_j\}$. The claim for μ follows by analogy with Corollary 2.3. $\square$

As it has already been noted, a sufficient condition for the almost everywhere existence of the regular approximate derivative is the membership in the Sobolev class $W^{p,1}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$ with $p > n - 1$. However, this condition does not guarantee the existence of a continuous modification, which requires the estimate $p > n$ (this estimate, as note above, ensures also Lusin's condition (N)). In the latter case, as we shall now see, the formulation of the main result is simplified. We now consider in more detail the mappings of the class $W^{p,1}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$.

In the proof of a one more corollary we need the following technical result.

Lemma 2.6. *Let $f \in W^{p,1}(U, \mathbf{R}^n)$, where $U \subset \mathbf{R}^n$ is the unit ball and $p \geq 1$, have a continuous modification on the closure of U, e.g., let $p > n$. Then for any $\varepsilon > 0$, there exists a Lipschitzian mapping $\psi\colon U \to \mathbf{R}^n$ such that*

1) $\lambda\big(x \in U |\, f(x) \neq \psi(x)\big) \leq \varepsilon$;
2) $\operatorname*{esssup}_{x\in U} \|f(x) - \psi(x)\| \leq \varepsilon$;
3) $\displaystyle\int_U \|Df(x) - D\psi(x)\|^p \, dx \leq \varepsilon.$

Proof. We can deal with a continuous modification of f. Let us take a smooth mapping β on $\mathbf{R}^n$ such that $0 \leq \|\beta\| \leq 1$, $\beta(x) = x$ if $\|x\| \leq 1$, and $\beta(x) = 0$ if $\|x\| \geq 2$. Let $M = \sup_x \|D\beta(x)\|$. Let us take a sequence $\{f_j\} \subset C^1(U, \mathbf{R}^n)$ convergent to

f in $W^{p,1}(U,\mathbf{R}^n)$ and uniformly on U, which is possible by the continuity of f. According to [10, Theorem 3.1.16], there exists a mapping $g \in C^1(U,\mathbf{R}^n)$ such that

$$\lambda\big(x \in U |\, g(x) \neq f(x)\big) < \frac{\varepsilon}{2}.$$

By using the uniform integrability of the sequence $\|Df_j\|$ and the absolute continuity of the Lebesgue integral, we can choose $\delta \in (0,\varepsilon)$ such that

$$\int\limits_{\delta/4 \le \|f-g\| \le 2\delta} \big[\|Df(x)\| + M\|Dg(x)\| + M\|Df_j(x)\|\big]^p \, dx \le \frac{\varepsilon}{4}$$

$$\int\limits_{\|f-g\| \le \delta} \|Df(x) - Dg(x)\|^p \, dx < \frac{\varepsilon}{4}.$$

Let us pick j such that

$$\operatorname*{esssup}_{x \in U} \|f_j(x) - f(x)\| < \frac{\delta}{4}, \qquad \int\limits_U \|Df_j(x) - Df(x)\|^p \, dx < \frac{\varepsilon}{2}.$$

For every $\alpha \in (0,1)$, we define a mapping $\beta_\alpha \colon \mathbf{R}^n \to \mathbf{R}^n$ as follows:

$$\beta_\alpha(x) = \alpha\beta(x/\alpha).$$

Then the mapping

$$\psi(x) = \beta_{\delta/2}\big(g(x) - f_j(x)\big) + f_j(x)$$

has the desired properties. Conditions 1) and 2) are clearly fulfilled; in order to verify condition 3) we note that

$$\{\|g - f_j\| < \delta/2\} \subset \{\|f-g\| < \delta\}, \quad \{\delta/2 \le \|g - f_j\| \le \delta\} \subset \{\delta/4 \le \|f-g\| \le 2\delta\},$$

$$D\psi(x) = \begin{cases} Dg & \text{if } \|f_j(x) - g(x)\| < \delta/2 \\ Df_j & \text{if } \|f_j(x) - g(x)\| \ge \delta, \end{cases}$$

and that

$$\|D\psi(x)\| \le M\big(\|Dg(x)\| + \|Df_j(x)\|\big) \qquad \text{if } \|f_j(x) - g(x)\| < \delta.$$

Therefore,

$$\int_U \|Df(x) - D\psi(x)\|^p\,dx \le \int_{\|f_j - g\| < \delta/2} \|Df(x) - Dg(x)\|^p\,dx$$

$$+ \int_{\delta/2 \le \|f_j - g\| \le \delta} \big[\|Df(x)\| + M\|Dg(x)\| + M\|Df_j(x)\|\big]^p\,dx$$

$$+ \int_U \|Df(x) - Df_j(x)\|^p\,dx \le \int_{\|f-g\| < \delta} \|Df(x) - Dg(x)\|^p\,dx$$

$$+ \int_{\delta/4 \le \|f-g\| \le 2\delta} \big[\|Df(x)\| + M\|Dg(x)\| + M\|Df_j(x)\|\big]^p\,dx + \frac{\varepsilon}{2} < \varepsilon.$$

The lemma is proven. □

Corollary 2.7. *Let $F_j, F \in W^{p,1}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$, where $p > n$, and let the mappings F_j converge to F with respect to the Sobolev norm $\|\cdot\|_{p,1}$ on every ball. Assume that $E \subset \{\det F \neq 0\}$ is a measurable set of finite Lebesgue measure. Then the measures $\lambda|_E \circ F_j^{-1}$ converge in the variation norm to the measure $\lambda|_E \circ F^{-1}$. The same is true if instead of the convergence with respect to the norm $\|\cdot\|_{p,1}$ on every ball one requires the boundedness of $\{F_j\}$ in $W^{p,1}(U, \mathbf{R}^n)$ for every ball U and the convergence of F_j to F almost everywhere. Finally, our assertion is true in the case $p = n$ if in addition to the convergence in the norm $\|\cdot\|_{p,1}$ on every ball U, the mappings F_j have continuous modifications and one has* $\operatorname{esssup}_U \|F_j - F\| \to 0$.

Proof. Recall that in the case $p > n$, the mappings F_j and F have continuous modifications that satisfy Lusin's condition (N). By the Sobolev embedding theorem, the continuous modifications of the mappings F_j converge to the continuous modification of F uniformly on compact sets. Thus, the previous corollary applies. Note that, as one can see from the reasoning below for the case $p = n$, we could prove our claim without use of the fact that the continuous modifications of the mappings F_j and F satisfy Lusin's condition (N).

In the case of the local boundedness of $\{F_j\}$ in the norm $\|\cdot\|_{p,1}$ with $p > n$ and the almost sure convergence to F together with the derivatives, the same reasoning applies. Indeed, by the compactness of the embedding of $W^{p,1}(U, \mathbf{R}^n)$ to $C(U, \mathbf{R}^n)$ for every ball U and the almost sure convergence, the sequence of continuous modifications of the mappings F_j converges uniformly on U to the continuous modification of F_0. In addition, the sequence $\{\det DF_j\}$ is uniformly integrable on U by the boundedness of $\{\|DF_j\|^p\}$ in $L^1(U)$ and the condition $p > n$.

Let us pass to the case $p = n$ and consider a sequence of continuous mappings $F_j \in W^{n,1}(U, \mathbf{R}^n)$ convergent to F_0 uniformly on U and in $W^{p,1}(U, \mathbf{R}^n)$. Assume that $\Omega \subset \{x \in U |\ \det DF_0(x) \neq 0\}$. Let us put $c_j = \sup_{x\in U} \|F_j(x) - F_0(x)\|$. Note that $\lim_{j\to\infty} c_j = 0$. Let us fix $\varepsilon > 0$ and take mappings $\{\psi_j\}$, $j \geq 0$, according to Lemma 2.6 applied to every mapping F_j, $j \geq 0$, in place of f and the number $\varepsilon_j = \varepsilon 2^{-j}$ in place of ε. The mappings

$$G_j = \beta_{c_j}(\psi_j - \psi_0) + \psi_0,$$

where β_α has been defined in the proof of Lemma 2.6, are Lipschitzian and converge uniformly to $G_0 = \psi_0$ as $j \to \infty$. The sequence $\{\|DG_j\|^n\}$ is uniformly integrable, whence the uniform integrability of the sequence $\{\det DG_j\}$ follows, since $|\det DF_j| \leq \|DF_j\|^n$. Let us put

$$A = \bigcup_{j=0}^{\infty} (F_j \neq \psi_j).$$

Then $G_j = \psi_j = F_j$ outside A for all $j \geq 0$ and $\lambda(A) < 2\varepsilon$. Since $D\beta_c(y) = I$ if $\|y\| < c$ and $D\psi_j(x) = DF_j(x)$ for almost all $x \in \Omega\backslash A$ by the coincidence of ψ_j and F_j on $\Omega\backslash A$, one has $DG_j(x) = D\psi_j(x) = DF_j(x)$ if $x \in \Omega\backslash A$, whence it follows that the mappings DG_j converge in measure on the set $\Omega\backslash A$ to $DF_0|_{\Omega\backslash A} = DG_0|_{\Omega\backslash A}$. Since the Lipschitzian mappings G_j satisfy Lusin's condition (N), we get by the previous corollary that the measures $\lambda|_{\Omega\backslash A} \circ F_j^{-1} = \lambda|_{\Omega\backslash A} \circ G_j^{-1}$ converge in the variation norm to the measure $\lambda|_{\Omega\backslash A} \circ G_0^{-1} = \lambda|_{\Omega\backslash A} \circ F_0^{-1}$. Since $\varepsilon > 0$ is arbitrary, we obtain that the measures $\lambda|_\Omega \circ F_j^{-1}$ converge in the variation norm to the measure $\lambda|_\Omega \circ F_0^{-1}$. $\square$

Let us recall that the Sobolev space $W^{p,r}(U)$, where $p > 1$, r is a natural number, and U is an open set in $\mathbf{R}^n$, is defined as the set of all functions $f \in L^p(U)$ whose generalized partial derivatives up to order r belong to $L^p(U)$. We denote by $W^{p,r}_{loc}(U)$ the set of all functions f such that $\zeta f \in W^{p,r}(U)$ for every $\zeta \in C_0^\infty(U)$. Finally, let $W^{p,r}_{loc}(U, \mathbf{R}^n)$ be the class of all mappings $F = (f_1, \ldots, f_n)$ on U such that $f_i \in W^{p,r}_{loc}(U)$, $i = 1, \ldots, n$. The r-th derivative $D^r f$ is a mapping with values in the space $\mathcal{H}_r$ of r-linear mappings from $(\mathbf{R}^n)^r$ to $\mathbf{R}^n$ equipped with its natural inner product. The Sobolev space $W^{p,r}(U, \mathbf{R}^n)$ is complete with respect to the norm

$$\|F\|_{p,r} = \Big\| \Big(\sum_{i=1}^n f_i^2 \Big)^{1/2} \Big\|_{L^p(U)} + \Big\| \Big(\sum_{1\leq k\leq r, i\leq n} (\partial_{x_{i_1}} \cdots \partial_{x_{i_k}} f_i)^2 \Big)^{1/2} \Big\|_{L^p(U)}.$$

By the Sobolev embedding theorem (see [24, Theorem 4.6.1]), if $rp > n$ and U is a ball, then every mapping $F \in W^{p,r}(U, \mathbf{R}^n)$ has a continuous modification that satisfies Hölder's condition of order $\alpha = \min(1, r - n/p)$. Moreover, there exists $C(p,r) > 0$ such that for the continuous version of any $F \in W^{p,r}(U, \mathbf{R}^n)$, one has

$$\|F(x) - F(y)\| \leq C(p,r) \|D^r F\|_{L^p(U)} \|x - y\|^\alpha, \qquad \forall x,\, y \in U, \tag{2.8}$$

where $\|D^r F\|_{L^p(U)}$ stands for the L^p-norm of the function $\|D^r F\|_{\mathcal{H}_r}$ on U. By the same reasoning as in [4, Lemma 8.1] one gets that the continuous version of $F \in W^{p,r}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$ with $pr > n$ satisfies Lusin's condition (N). For the reader's convenience, we include a proof. Let $A \subset \mathbf{R}^n$ have Lebesgue measure zero. In order to show that $\lambda\bigl(F(A)\bigr) = 0$, we may assume that A is contained in a ball U and that $F \in W^{p,r}(\mathbf{R}^n, \mathbf{R}^n)$. It suffices to consider the case $\alpha < 1$, since for Lipschitzian mappings the statement is true. Let $\varepsilon > 0$. There exists a sequence of cubes Q_i with disjoint interiors such that $A \subset \bigcup_{i=1}^\infty Q_i \subset U$ and $\lambda\bigl(\bigcup_{i=1}^\infty Q_i\bigr) < \varepsilon$. On the account of (2.8), the diameter of the set $F(Q_i)$ does not exceed $Cr_i^\alpha \|D^r F\|_{L^p(Q_i)}$, where r_i is the length of an edge of Q_i and C is independent of i and F. Hence,

$$\lambda\bigl(F(A)\bigr) \le C^n \sum_{i=1}^\infty r_i^{n\alpha} \|D^r F\|^n_{L^p(Q_i)}.$$

By Hölder's inequality applied to $s = \alpha^{-1}$ and $t = s(s-1)^{-1} = (1-\alpha)^{-1}$ and the estimate $nt \ge p$, which follows by the inequality $n \ge p(1-\alpha) = p - pr + n$, we get

$$\begin{aligned}
\lambda\bigl(F(A)\bigr) &\le C^n \Bigl(\sum_{i=1}^\infty r_i^n\Bigr)^\alpha \Bigl(\sum_{i=1}^\infty \|D^r F\|^{nt}_{L^p(Q_i)}\Bigr)^{1/t} \\
&\le C^n \Bigl(\sum_{i=1}^\infty r_i^n\Bigr)^\alpha \Bigl(\sum_{i=1}^\infty \bigl(\|D^r F\|_{L^p(Q_i)}/\|D^r F\|_{L^p(U)}\bigr)^{nt} \|D^r F\|^{nt}_{L^p(U)}\Bigr)^{1/t} \\
&\le C^n \Bigl(\sum_{i=1}^\infty r_i^n\Bigr)^\alpha \Bigl(\sum_{i=1}^\infty \|D^r F\|^p_{L^p(Q_i)}\Bigr)^{1/t} \|D^r F\|^{n-p/t}_{L^p(U)} \\
&\le C^n \lambda\Bigl(\bigcup_{i=1}^\infty Q_i\Bigr)^\alpha \|D^r F\|^{p/t}_{L^p(U)} \|D^r F\|^{n-p/t}_{L^p(U)} \le C^n \varepsilon^\alpha \|D^r F\|^n_{L^p(U)},
\end{aligned}$$

whence the desired conclusion.

Now it is obvious that the statement of Corollary 2.7 remains valid if the convergence in $W^{p,1}(U, \mathbf{R}^n)$ is replaced by the following conditions: the mappings F_j belong to the class $W^{p,r}_{loc}(\mathbf{R}^n, \mathbf{R}^n)$ with $rp > n$ and converge to F in the norm of $W^{p,r}(U, \mathbf{R}^n)$ for every ball U, and, in addition, the sequence $\{|\det DF_j|\}$ is uniformly integrable on U.

Finally, note that the hypothesis of the nondegeneracy of DF on the set E is essential: for example, the sequence of functions $F_j(x) = j^{-1}$ on $[0, 1]$ converges to zero uniformly together with the derivatives, but the induced measures are Dirac's measures at the points j^{-1}, which do not converge to Dirac's measure at 0 in the variation norm. In the next section we consider mappings between spaces of different dimensions; then the nondegeneracy of DF is replaced by the surjectivity of DF.

3 Applications and generalizations

Let us recall that a Borel measure μ (possibly, signed) on a locally convex space X is called Radon if for every Borel set $B\subset X$ and every $\varepsilon>0$, there exists a compact set $K_\varepsilon\subset B$ such that $|\mu|(B\backslash K_\varepsilon)<\varepsilon$, where $|\mu|$ is the total variation of μ. Let us also recall (see, e.g., [13, §21.2] or [1, §1]) that if X is the direct sum of a finite-dimensional linear subspace Z and a closed linear subspace Y, then there exists a system of measures μ^y, $y\in Y$, called the conditional measures, which for every Borel set B satisfy the equality

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

where ν is the image of $|\mu|$ under the natural projection of X to Y.

We shall call a measurable function ϱ on $\mathbf{R}^n$ locally strictly positive if $\operatorname*{ess\,inf}_U \varrho>0$ for every ball U.

Theorem 3.1. *Let X be a locally convex space, let Z be a finite-dimensional linear subspace in X generated by vectors $e_1,\dots,e_n$, and let Y be a closed linear subspace that complements Z to X. Suppose that μ is a Radon probability measure on X such that conditional measures μ^y on the subspaces $Z+y$, $y\in Y$, are absolutely continuous with respect to the natural Lebesgue measures on Z generated by some linear isomorphism between Z and $\mathbf{R}^n$ and that the corresponding densities are either locally strictly positive or continuous. Let F_j, $F\colon X\to\mathbf{R}^n$ be μ-measurable mappings, which for μ-almost all x are absolutely continuous on the lines $x+\mathbf{R}^1e_i$, $i=1,\dots,n$, and let $F_j\to F$ and $\partial_{e_i}F_j\to\partial_{e_i}F$ in $L^p(\mu,\mathbf{R}^n)$, where $p>n$. If E is a μ-measurable set such that $\det\Bigl((\partial_{e_i}F,\partial_{e_k}F)\Bigr)_{i,k=1}^n\neq 0$ on E, then the measures $\mu|_E\circ F_j^{-1}$ converge in the variation norm to the measure $\mu|_E\circ F^{-1}$.*

Proof. As noted in the proof of Theorem 2.1, it suffices to verify our claim for some subsequence in $\{F_j\}$. Hence we may assume that

$$\|F_j-F\|_{L^p(\mu,\mathbf{R}^n)}+\|\partial_{e_i}F_j-\partial_{e_i}F\|_{L^p(\mu,\mathbf{R}^n)}\le 2^{-j}$$

for all $j\ge 1$ and $i=1,\dots,n$. Let ν be the image of μ under the natural projection of X onto Y. By Fubini's theorem and the integrability of the series

$$S_i(x)=\sum_{j=1}^\infty\Bigl[\|\partial_{e_i}F_j(x)-\partial_{e_i}F(x)\|^p+\|F_j(x)-F(x)\|^p\Bigr],\quad i=1,\dots,n,$$

it follows that for ν-almost every y, the series

$$Z_i(y,u)=\sum_{j=1}^\infty\Bigl[\|\partial_{e_i}F_j(y+u)-\partial_{e_i}F(y+u)\|^p+\|F_j(y+u)-F(y+u)\|^p\Bigr],\quad i=1,\dots,n,$$

are integrable in u with respect to the measure μ^y. For every such y and every $i = 1, \dots, n$, we get the convergence of the sequence $\{F_j(y + \cdot)\}$ to $F(y + \cdot)$ and the convergence of the sequence $\{\partial_{e_i} F_j(y + \cdot)\}$ to $\partial_{e_i} F(y + \cdot)$ in $L^p(\mu^y, \mathbf{R}^n)$. In the case of the strict local positivity of the density ϱ^y of the measure μ^y, this yields the convergence in $L^p(U, \mathbf{R}^n)$ for every ball $U \subset Z$, and in the case of continuity of ϱ^y the same takes place for every ball, whose closure is contained in the set $\{\varrho^y \neq 0\}$. Hence Corollary 2.7 applies. We observe that for every Borel set $B \subset X$, one has

$$\mu|_E(B) = \int\limits_Y \mu^y|_E(B)\, \nu(dy).$$

It follows that for every Borel mapping $G\colon X \to \mathbf{R}^n$ and every Borel set $B \subset \mathbf{R}^n$, one has

$$\mu|_E \circ G^{-1}(B) = \int\limits_Y \mu^y|_E \circ G^{-1}(B)\, \nu(dy). \tag{3.9}$$

It remains to note that

$$\left\| \mu|_E \circ F_j^{-1} - \mu|_E \circ F^{-1} \right\| \le \int\limits_Y \left\| \mu^y|_E \circ F_j^{-1} - \mu^y|_E \circ F^{-1} \right\| \nu(dy), \tag{3.10}$$

which converges to zero as $j \to \infty$ by Lebesgue's dominated convergence theorem. □

Corollary 3.2. *Assume that in the situation of the previous theorem ν is a measure absolutely continuous with respect to μ. Then the measures $\nu \circ F_j^{-1}$ converge to $\nu \circ F^{-1}$ in the variation norm.*

Proof. The same reasoning as in Corollary 2.3 applies. □

Note that in the situation of the theorem proven the measure $\mu|_E \circ F^{-1}$ is absolutely continuous (this is obvious from the proof; see also [2, Corollary 6.8.2], [5, Ch.II, §5], [15, Lemma 2.1.2]).

We recall that a Radon measure μ on a locally convex space X is called continuous along the direction of a vector $h \in X$ if for every Borel set B, the function $t \mapsto \mu(B + th)$ is continuous. If $X = \mathbf{R}^n$, then the measure μ is continuous along n independent directions precisely when it is absolutely continuous with respect to Lebesgue measure. The continuity of μ along the directions of the vectors from a finite-dimensional linear subspace Z is equivalent to the existence of absolutely continuous conditional measures on the subspaces parallel to Z. A more detailed information can be found in [1]. The logarithmic derivative of the measure μ along a vector h is a μ-integrable function β_h^μ such that the integration by parts formula

$$\int\limits_X \partial_h \varphi(x)\, \mu(dx) = -\int\limits_X \varphi(x) \beta_h^\mu(x)\, \mu(dx)$$

is valid for any function φ of the form $\varphi(x) = \psi\big(l_1(x), \ldots, l_n(x)\big)$, where $\psi \in C_b^\infty(\mathbf{R}^n)$ and the l_i's are continuous linear functionals on X. The existence of the logarithmic derivative along h is equivalent to the differentiability of the functions $t \mapsto \mu(A + th)$ for all Borel sets A; then the mapping $d_h\mu\colon A \mapsto \frac{d}{dt}\mu(A + th)|_{t=0}$ is a countably additive signed measure, which is absolutely continuous with respect to μ, and the Radon–Nikodym density of $d_h\mu$ with respect to μ coincides with the logarithmic derivative along h (see [1] or [6]). A measure μ on $\mathbf{R}^n$ is differentiable along n linearly independent vectors if and only if it has a density $\varrho \in W^{1,1}(\mathbf{R}^n)$. A sufficient condition for the existence of continuous strictly positive densities of conditional measures μ^y is the existence of the logarithmic derivatives $\beta_{e_i}^\mu$, $i = 1, \ldots, n$, and the integrability of the functions $\exp\big(\varepsilon|\beta_{e_i}^\mu|\big)$ for some $\varepsilon > 0$ (see [3, Proposition 1]).

Concerning the notions and notation from the theory of Gaussian measures used in the following assertion, see [2]. In particular, $W^{p,1}(\gamma, \mathbf{R}^n)$ stands for the Sobolev space of mappings $f = (f_1, \ldots, f_n)\colon X \to \mathbf{R}^n$ such that the f_j's belong to the Sobolev space $W^{p,1}(\gamma)$ defined as the completion of the class of all functions of the form $\psi(x) = \psi_0(l_1, \ldots, l_k)$, where $\psi_0 \in C_b^\infty(\mathbf{R}^k)$ and the l_j are continuous linear functionals on X, equipped with the norm

$$\|\psi\|_{W^{p,1}} = \|\psi\|_{L^p(\mu)} + \|D_H\psi\|_{L^p(\mu,H)},$$

where $D_H\psi$ is the gradient of ψ along H.

Corollary 3.3. *Let γ be a Radon Gaussian measure on a locally convex space X having the Cameron–Martin space H, let F_j, $F \in W^{p,1}(\gamma, \mathbf{R}^n)$, where $p > n$, and let $F_j \to F$ with respect to the norm of $W^{p,1}(\gamma, \mathbf{R}^n)$. Then for any γ-measurable set $E \subset \Big\{x\colon\ D_H F(x)(H) = \mathbf{R}^n\Big\}$, the measures $\gamma|_E \circ F_j^{-1}$ converge in the variation norm to the measure $\gamma|_E \circ F^{-1}$.*

Proof. The set E is contained in the union of the sets

$$E_{k_1,\ldots,k_n} = \big\{x \in E\colon\ \det(\partial_{e_{k_i}}F, \partial_{e_{k_j}}F) \neq 0\big\},$$

where $\{e_i\}$ is a fixed orthonormal basis in H. Hence it suffices to prove our claim for $E \subset E_{k_1,\ldots,k_n}$. It remains to use the existence of modifications of the mappings F_j and F that are absolutely continuous on the lines $x + \mathbf{R}^1 e_i$, $i = 1, \ldots, n$ (see [2, Ch. 5]) and apply the previous theorem. $\square$

Note that Theorem 3.1 applies to Gibbs measures (see [11]) possessing absolutely continuous conditional distributions with locally strictly positive densities on the finite-dimensional subspaces. Clearly, Theorem 3.1 applies to the spaces $X = \mathbf{R}^d$ with $d \geq n$ and suitable measures μ (e.g., measures with locally strictly positive densities). This enables us to extend the results in the previous section to the case of manifolds. The Sobolev classes $W^{p,r}(U,M)$ of mappings between Riemannian manifolds U and M are defined by analogy with the case of $\mathbf{R}^n$; the corresponding local classes $W^{p,r}_{loc}(U,M)$ can be defined merely in local coordinates.

Corollary 3.4. *Let M_1 and M_2 be two finite dimensional Riemannian manifolds, let $p > \dim M_2$, and let $F_j \in W^{p,1}_{loc}(M_1, M_2)$ be a sequence of mappings convergent to a mapping $F \in W^{p,1}_{loc}(M_1, M_2)$ in the following sense: for every relatively compact local chart $U \subset M_1$, one has $F_j|_U \to F|_U$ in the Sobolev space $W^{p,1}(U, M_2)$. Suppose that μ is a Radon measure on M_1 that is absolutely continuous with respect to the Riemannian volume on M_1 and E is a set of finite μ-measure. If $DF(x)\colon T_xM_1 \to T_{F(x)}M_2$ is surjective for almost every $x \in E$, then the measures $\mu|_E \circ F_j^{-1}$ converge to $\mu|_E \circ F^{-1}$ in the variation norm.*

Proof. As above, it suffices to prove our claim for compact E. Then everything reduces to finitely many charts with compact closures, hence we may assume that $M_1 = \mathbf{R}^d$ and $M_2 = \mathbf{R}^n$, where $d \geq n$. Let $e_1, \ldots, e_d$ be a basis in $\mathbf{R}^d$. The set E is covered by the subsets $E_{k_1,\ldots,k_n}$ of the points x such that the vectors $DF(x)(e_{k_i})$, $i = 1, \ldots, n$, span $\mathbf{R}^n$. It suffices to prove our assertion for $E \subset E_{k_1,\ldots,k_n}$. Now Corollary 3.2 applies. □

Let us prove an analogue of Corollary 2.4 for mappings between spaces of different dimensions.

Proposition 3.5. *Let F, $F_j\colon \mathbf{R}^d \to \mathbf{R}^n$ be continuous mappings with Lusin's* (N)*-property such that they have regular approximate derivatives (e.g., just usual derivatives) DF and DF_j almost everywhere and let $E \subset \mathbf{R}^d$ be a set of finite Lebesgue measure. Suppose that the mappings F_j converge to F uniformly on compact sets, that the mappings DF_j converge in measure on E to DF, and that the minors of order n of the matrices DF_j either converge in $L^1(K)$ for every compact set K or are majorized in the absolute value by a locally integrable function. If $DF(x)$ is surjective for almost every $x \in E$, then the measures $\lambda|_E \circ F_j^{-1}$ converge to $\lambda|_E \circ F^{-1}$ in variation. Furthermore, if μ is an absolutely continuous finite measure, then the measures $\mu|_E \circ F_j^{-1}$ converge to $\mu|_E \circ F^{-1}$ in variation (this is true also for E of infinite Lebesgue measure).*

Proof. The reasoning is very similar to that of Theorem 3.1. It suffices to prove our assertion for $E \subset E_{k_1,\ldots,k_n}$, where $E_{k_1,\ldots,k_n}$ is the set of all points x in a fixed ball U such that the vectors $DF(x)(e_{k_i})$, $i = 1, \ldots, n$, span $\mathbf{R}^n$. In addition, passing to a subsequence, we may assume that the mappings DF_j converge almost everywhere on E and that $\|\Lambda_{F_j} - \Lambda_F\|_{L^1(U)} \leq 2^{-j}$, where $\Lambda_G(x) = \det\Bigl(DG(x)(e_{k_i}), DG(x)(e_{k_l})\Bigr)_{i,l=1}^{n}$. Then by the same arguments as in the proof of Theorem 3.1 we obtain that for a.e. $y \in \mathbf{R}^{d-n}$, the mappings $z \mapsto F_j(y,z)$ and $z \mapsto F(y,z)$ on $\mathbf{R}^n$ satisfy the hypotheses of Corollary 2.4 with the set $E_y = \bigl\{z \in$

$\mathbf{R}^n\colon\ (y,z)\in E\big\}$ in place of E. Clearly, all those hypotheses, except for the uniform integrability of the Jacobians, are satisfied in a trivial way. The above estimate on Λ_{F_j} yields the L^1-convergence for a.e. y, hence the desired uniform integrability. Certainly, in the case of a locally integrable majorant, one also has the uniform integrability. Therefore, estimate (3.10) and Corollary 2.4 complete our proof. $\square$

Obviosly, the above proposition extends to mappings between Riemannian manifolds, provided the corresponding assumptions are fulfilled in local charts.

Remark 3.6. (i) It follows from relation (3.9) and Corollary 2.5 applied to the conditional measures μ^y that in the situations of Theorem 3.1, Corollary 3.3, Corollary 3.4, and Proposition 3.5, one has

$$\lim_{j\to\infty}\mu\big(F_j(A)\bigtriangleup F(A)\big)=0 \qquad \text{and} \qquad \lim_{j\to\infty}\mu\big(F_j(A)\big)=\mu\big(F(A)\big)$$

for any measurable set $A\subset E$. The same is true for any μ-absolutely continuous measure ν in place of μ.

(ii) It is worth noting that if the mappings F_j and F have stronger differentiability and nondegeneracy properties (with certain uniformity in j), then one has a stronger convergence of the induced measures. For example, suppose that in the situation of Corollary 3.3 the mappings $F_j=(f_j^1,\ldots,f_j^n)$ are uniformly bounded in every space $W^{p,r}(\gamma,\mathbf{R}^n)$ and the functions $\Delta_j=\big|\det(D_{_H}f_j^i,D_{_H}f_j^k)_{i,k=1}^n\big|^{-1}$ are uniformly bounded in every $L^p(\gamma)$, $1<p<\infty$. If $F_j\to F$ almost everywhere, then the measures $\gamma\circ F_j^{-1}$ have densities ϱ_j which converge to the density ϱ of the measure $\gamma\circ F^{-1}$ in every space $W^{p,r}(\mathbf{R}^n)$, in particular, locally uniformly together with all derivatives. This is a well-known result in the Malliavian calculus (see, e.g., [2, §6.9], [1], [14, Ch. III, §6], [15], [19]). Certainly, results of this sort can be easily proved also for measures on $\mathbf{R}^n$ with sufficiently nice densities (see, e.g., [2, §6.9]). However, methods of proof of such results are completely different from those used above in the case of low regularity and weaker nondegeneracy.

The authors are grateful to P.V. Paramonov for useful discussions. This work has been supported in part by the Russian Foundation of Fundamental Research Grants N 96–15–96865 and N 97–01–00932, the INTAS project 94–378, and the DFG Grant N 436 RUS 113/343/0(R).

References

[1] Bogachev V.I., Differentiable measures and the Malliavin calculus, J. Math. Sci., 1997, v. 87, n 5, p. 3577–3731.

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

[3] Bogachev V.I., Mayer-Wolf E., The flows generated by vector fields of the Sobolev type and the corresponding transformations of probability measures, Dokl. Russian Acad. Sci., 1998, t. 358, n 4, p. 442–446 (in Russian); English transl.: Dokl. Math., 1998, v. 57, n 1, p. 64–68.

[4] Bojarski B., Iwaniec T., Analytical foundations of the theory of quasiconformal mappings in R^n, Annales Acad. Sci. Fennicae, Ser. A. I. Math., 1983, v. 8, p. 257–324.

[5] Bouleau N., Hirsch F., Dirichlet forms and analysis on Wiener space, Walter de Gruyter, Berlin – New–York, 1991.

[6] Daletskii Yu.L., Fomin S.V., Measures and Differential equations in infinite dimensional spaces, Nauka, Moscow, 1983 (in Russian); English transl.: Kluwer Academic Publ., Dordrecht, 1993.

[7] Davydov Yu.A., Lifshits M.A., Smorodina N.V., Local properties of distributions of stochastic functionals, Fizmatlit, Moscow, 1995 (in Russian); English transl.: Amer. Math. Soc., Providence, Rhode Island, 1998.

[8] Davydov Yu.A., On convergence in variation of one-dimensional image measures, Zap. Nauch. Semin. POMI, 1992, t. 194, Problemi teorii veroyatnostnih raspredelenii XII, p. 48–58 (in Russian); English transl.: J. Math. Sci., 1995, v. 75, n 5, p. 1903–1909.

[9] Ellis H.W. Darboux properties and applications to non-convergent integrals, Canad. J. Math., 1951, v. 3, p. 471–485.

[10] Federer H., Geometric measure theory, Springer, Berlin, 1969.

[11] Georgi H.-O., Gibbs measures and phase transitions, de Gruyter, Berlin — New York, 1988.

[12] Gol'dshtein V.M., Reshetnyak Yu.G., Quasiconformal mappings and Sobolev spaces, Nauka, Novosibirsk, 1983 (in Russian); English transl.: Kluwer Academic Publ., Dordrecht, 1990.

[13] Hennequin P.-L., Tortrat A., Théorie des probabilités et quelques applications, Masson et Gie, 1965.

[14] Malliavin P., Stochastic analysis, Springer, Berlin – New York, 1997.

[15] Nualart D. The Malliavin calculus and related topics, Springer, Berlin – New York, 1995.

[16] Ponomarev S.P., Submersions and preimages of sets of measure zero, Sib. Matem Zhurn., 1987, t. 28, n 1, p. 199–210 (in Russian); English transl.: Sib. Math. J., 1987, v. 28, n 1, p. 153–163.

[17] Ponomarev S.P., Property N of homeomorphisms of the class W_p^1, Sibirsk. Matem. Zhurn., 1987, t. 28, n 2, p. 140–148 (in Russian); English transl.: Sib. Math. J., 1987, v. 28, n 2, p. 291–298.

[18] Rado T., Reichelderfer P.V. Continuous transformations in analysis, Berlin, Springer–Verlag, 1955.

[19] Ren J., Watanabe S., A convergence theorem for probability densities and conditional expectations of Wiener functionals, In: Dirichlet forms and stochastic processes (Beijing, 1993), 335–344, de Gruyter, Berlin, 1995.

[20] Reshetnyak Yu.G., Space mappings with bounded distortion, Sibirsk. Matem. Zhurn., 1967, t. 8, n 3, p. 629–658 (in Russian); English transl.: Sib. Math. J., 1967, v. 8, n 3, p. 466–487.

[21] Reshetnyak Yu.G., Space mappings with bounded distortion, Nauka, Moscow, 1982 (in Russian); English transl.: Amer. Math. Soc., Providence, Rhode Island, 1989.

[22] Reshetnyak Yu.G., Property N for the space mappings of class $W_{n,loc}^1$, Sibirsk. Matem. Zhurn., 1987, t. 28, n 5, p. 149–153 (in Russian); English transl.: Sib. Math. J., 1987, v. 28, n 5, p. 810–813.

[23] Shiryaev A.N., Probability theory, Springer, Berlin, 1984.

[24] Triebel H., Interpolation theory, function spaces, differential operators, North-Holland, Amsterdam — New York — Oxford, 1978.

[25] Ziemer W. Weakly differentiable functions, Springer–Verlag, New York – Berlin, 1989.