STUDIA MATHEMATICA 107 (1) (1993)

Molecular decompositions and embedding theorems for vector-valued Sobolev spaces with gradient norm

by

A. PEŁCZYŃSKI and M. WOJCIECHOWSKI (Warszawa)

Abstract. Let E be a Banach space. Let $L^1_{(1)}(\mathbb{R}^d, E)$ be the Sobolev space of E-valued functions on $\mathbb{R}^d$ with the norm

$$\int_{\mathbb{R}^d} \|f\|_E\,dx + \int_{\mathbb{R}^d} \|\nabla f\|_E\,dx = \|f\|_1 + \|\nabla f\|_1 .$$

It is proved that if $f \in L^1_{(1)}(\mathbb{R}^d, E)$ then there exists a sequence $(g_m) \subset L^1_{(1)}(\mathbb{R}^d, E)$ such that $f = \sum_m g_m$; $\sum_m(\|g_m\|_1 + \|\nabla g_m\|_1) < \infty$; and $\|g_m\|_\infty^{1/d}\|g_m\|_1^{(d-1)/d} \le b\|\nabla g_m\|_1$ for $m = 1, 2, \ldots$, where b is an absolute constant independent of f and E. The result is applied to prove various refinements of the Sobolev type embedding $L^1_{(1)}(\mathbb{R}^d, E) \hookrightarrow L^2(\mathbb{R}^d, E)$. In particular, the embedding into Besov spaces

$$L^1_{(1)}(\mathbb{R}^d, E) \hookrightarrow B^{\theta(p,d)}_{p,1}(\mathbb{R}^d, E)$$

is proved, where $\theta(p, d) = d(p^{-1} + d^{-1} - 1)$ for $1 < p \le d/(d-1)$, $d = 1, 2, \ldots$

The latter embedding in the scalar case is due to Bourgain and Kolyada.

Introduction. This paper is devoted to the study of Sobolev spaces of differentiable functions in d real variables, mainly on the d-dimensional Euclidean space $\mathbb{R}^d$, taking values in an arbitrary Banach space E. We are primarily interested in the limit case of L^1-norm, i.e. when the norm is defined to be the integral of the norm of the gradient of the function plus the integral of the norm of the function. Denote this space by $L^1_{(1)}(\mathbb{R}^d, E)$. We exploit the following phenomenon called the molecular decomposition:

Every function in the Sobolev space endowed with the L^1-gradient norm can be written as an absolutely convergent series of uniformly bounded functions; the sup norm of each term is controlled by the ratio of dth power of the L^1-norm of the gradient of the term and the $(d-1)$th power of the L^1-norm of the term (where d is the dimension of the underlying Euclidean space).

1991 *Mathematics Subject Classification*: 42B15, 46E35, 46E40.

Supported in part by KBN grant 2-1055-91-01.

The molecular decomposition is closely related to the coarea formula due to Federer and Kronrod (cf. [F], [Kr] and [BZ]) and to some isoperimetric inequalities. Implicitly it goes back to Federer and Fleming [FF] and Maz'ya [M1] who used similar decomposition ideas to evaluate the best constant in the Sobolev type embedding theorem $L^1_{(1)}(\mathbb{R}^d,\mathbb{C}) \hookrightarrow L^2(\mathbb{R}^d,\mathbb{C})$ due to Gagliardo [G] and Nirenberg [N]. In integral form the molecular decomposition was stated by Bourgain [Br1] who then employed it to prove the analogue of the Hardy inequality for analytic functions in Sobolev spaces with L^1-norms as well as a refinement of the Gagliardo–Nirenberg embedding (cf. [Br2] and [Po]). The Bourgain–Hardy type inequality says, roughly speaking, that if $f \in L^1_{(1)}(\mathbb{R}^d,\mathbb{C})$ for some $d \ge 2$ then the Fourier Transform of f belongs to some weighted L^1-space; the right weight is $(1+|\xi|)^{1-d}$ $\xi \in \mathbb{R}^d$. This weight function can be regarded as a linear functional on (the Fourier Transforms of) the Besov space $B^0_{d/(d-1),1}(\mathbb{R}^d,\mathbb{C})$. Therefore it suggests the embedding

$$(*)\qquad L^1_{(1)}(\mathbb{R}^d,\mathbb{C}) \hookrightarrow B^0_{d/(d-1),1}(\mathbb{R}^d,\mathbb{C})$$

(cf. [Br2] and Kolyada [K]). Although it is not clear for what Banach spaces E the Bourgain–Hardy type inequality holds for E-valued functions (not for all—the space c_0 is a counterexample) the embedding $(*)$ can be extended to all Banach spaces. To this end we use the molecular decomposition for $L^1_{(1)}(\mathbb{R}^d,E)$ as well as the fact that a large portion of the theory of Besov spaces can be carried over to Banach space-valued functions.

There are no essential difficulties in extending the molecular decomposition to Sobolev spaces of vector-valued functions on tori, on nice domains in $\mathbb{R}^d$, and on Euclidean manifolds. The required tools are an analog of the coarea formula and the right isoperimetric inequality. The connection of the latter with the Sobolev embedding theorem has been recently extensively discussed by Ledoux (cf. [Le] and the references there). The right generality for embedding theorems for Besov spaces seems to be that discussed in a recent book of Coulhon, Saloff-Coste and Varopoulos [CSV].

Briefly about the organization of the paper. In Section 1 we introduce vector-valued Sobolev spaces and the gradient norm in the space $L^1_{(1)}(\mathbb{R}^d,E)$. We show (Theorem 1.1) that f in $L^1_{(1)}(\mathbb{R}^d,E)$ always implies $\|f(\cdot)\|_E \in L^1_{(1)}(\mathbb{R}^d,\mathbb{R})$ (even when the norm $\|\cdot\|_E$ is nowhere differentiable!). Moreover, $|\nabla\|f(\cdot)\|_E| \le \|\nabla f\|$ a.e. Section 2 is devoted to the proof of the molecular decomposition theorem in the E-valued case (Theorem 2.1). In the scalar case we get it with the best possible constants. In Section 3 we apply the results of the previous sections to prove various embedding theorems; in some of them we get best constants. Part of the section is devoted to extending

some results on equivalence of various norms on scalar-valued Besov spaces to the Banach space-valued case.

Acknowledgements. The authors would like to express their gratitude to several mathematicians: Jean Bourgain made accessible to us his unpublished manuscripts [Br1] and [Br2] which had crucial impact on the birth of the present paper. S. Kashin called our attention to Kolyada's work. P. Lizorkin in a private letter proposed the idea of using the embedding $B^{\theta(p,d)}_{p,1}(\mathbb{R}^d) \hookrightarrow B^0_{p,1}(\mathbb{R}^d)$ to move away from the limit case (cf. Lemma 3.3 and Corollary 3.1). We are also indebted to K. Ball, T. Coulhon and M. Ledoux for valuable discussions.

Preliminaries. $\mathbb{R}$ stands for the reals, $\mathbb{C}$ for the complex plane, $\mathbb{R}^d$ for the d-dimensional real vector space ($d = 1,2,\ldots$), $\mathbb{Z}^d$ is the lattice in $\mathbb{R}^d$ consisting of all points with integer coordinates. For $x = (x(j))$, $y = (y(j))$ in $\mathbb{R}^d$ we put $\langle x,y\rangle = \sum_{j=1}^d x(j)y(j)$ and $|x| = \langle x,x\rangle^{1/2}$. B_d stands for the closed Euclidean unit ball of $\mathbb{R}^d$, i.e. $B_d = \{x \in \mathbb{R}^d : |x| \le 1\}$. λ_d denotes the d-dimensional Lebesgue measure of $\mathbb{R}^d$ normalized by $\lambda_d(\mathbb{I}^d) = 1$ where

$$\mathbb{I}^d = \{x = (x(j)) \in \mathbb{R}^d : |x(j)| \le 1/2 \text{ for } j = 1,\ldots,d\}.$$

Unless otherwise stated, integration is against λ_d for appropriate d.

If E is a complex Banach space and $f : \mathbb{R}^d \to E$ a measurable function then the Fourier transform of f is defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-i\langle x,\xi\rangle}\,dx \quad \text{for } \xi \in \mathbb{R}^d$$

provided the right hand side makes sense, say exists for λ_d-a.e. ξ.

1. Vector-valued Sobolev spaces and the gradient norm

Sobolev spaces of vector-valued functions on $\mathbb{R}^d$. Let E be a Banach space. Partial derivatives of a function $f : \mathbb{R}^d \to E$ are defined with respect to the norm topology of E. For example,

$$D^{(1,0,\ldots,0)}f(x_0) = \frac{\partial f}{\partial x(1)}(x_0) = a \in E$$
$$:\Leftrightarrow \lim_{t\to 0}\|a - t^{-1}(f(x_0 + t(1,0,\ldots,0)) - f(x_0))\|_E = 0.$$

$C^{(k)}(\mathbb{R}^d,E)$ stands for the space of functions $f : \mathbb{R}^d \to E$ such that f vanishes at infinity and has continuous partial derivatives of order $\le k$ vanishing at infinity. $C^{(k)}(\mathbb{R}^d,E)$ is a Banach space in the topology of uniform convergence of functions with all partial derivatives of order $\le k$ ($k = 1,2,\ldots$). $\mathcal{D}(\mathbb{R}^d,E) = \bigcap_{k=1}^\infty C^{(k)}(\mathbb{R}^d,E) \cap \{f : \mathbb{R}^d \to E : \operatorname{supp} f \text{ compact}\}$ where $\operatorname{supp} f = \operatorname{closure}\{x \in \mathbb{R}^d : f(x) \ne 0\}$.

Given $f:\mathbb{R}^d\to E$ and a multiindex α we say that $g:\mathbb{R}^d\to E$ is the αth distributional partial derivative of f (in symbols $g=\widetilde{D}^\alpha f$) provided

$$\int\limits_{\mathbb{R}^d}\psi(x)g(x)\,dx=(-1)^{|\alpha|}\int\limits_{\mathbb{R}^d}D^\alpha\psi(x)f(x)\,dx\quad \text{for } \psi\in\mathcal{D}(\mathbb{R}^d,\mathbb{R}),$$

where $|\alpha|=\sum_{j=1}^d\alpha(j)$.

$L^p(\mathbb{R}^d,E)$ stands for the Banach space of equivalence classes of E-valued functions on $\mathbb{R}^d$ which are p-Bochner integrable (cf. [DS], Chapt. III, §3, Definition 4) with respect to the Lebesgue measure on $\mathbb{R}^d$, equipped with the norm

$$\|f\|_p=\Big(\int\limits_{\mathbb{R}^d}\|f(x)\|_E^p\,dx\Big)^{1/p}\quad \text{for } 1\le p<\infty,$$

$$\|f\|_\infty=\operatorname*{ess\,sup}_{x\in\mathbb{R}^d}\|f\|_E\quad \text{for } p=\infty.$$

Now we are ready to define (similarly to the scalar case, cf. [St], Chapt. V, §2) the Sobolev spaces $L^p_{(k)}(\mathbb{R}^d,E)$ for $1\le p\le\infty$ and for $k=1,2,\ldots$

$L^p_{(k)}(\mathbb{R}^d,E)$ is the space of all functions $f:\mathbb{R}^d\to E$ having all distributional partial derivatives of order $\le k$ such that $f\in L^p(\mathbb{R}^d,E)$ and $\widetilde{D}^\alpha f\in L^p(\mathbb{R}^d,E)$ for $|\alpha|\le k$. The convergence in $L^p_{(k)}(\mathbb{R}^d,E)$ is the convergence of functions with all distributional partial derivatives of order $\le k$ in the norm topology of $L^p(\mathbb{R}^d,E)$. The space $L^p_{(k)}(\mathbb{R}^d,E)$ is banachable in this topology ($1\le p\le\infty$, $k=1,2,\ldots$).

If $1\le p<\infty$ then $\mathcal{D}(\mathbb{R}^d,E)$ as well as $C^{(k)}(\mathbb{R}^d,E)$ can be regarded as dense subsets of $L^p_{(k)}(\mathbb{R}^d,E)$. The proof is similar to that of [St], Chapt. V, Prop. 1.

The gradient norm in $L^p_{(1)}(\mathbb{R}^d,E)$. By $\mathcal{L}(\mathbb{R}^d,E)$ we denote the Banach space of all bounded real-linear operators from $\mathbb{R}^d$ into $E_\mathbb{R}$ where $E_\mathbb{R}$ is E regarded as a real Banach space. The norm in $\mathcal{L}(\mathbb{R}^d,E)$ is the usual operator norm

$$\|A\|=\|A:\mathbb{R}^d\to E\|=\sup\{\|Ay\|_E:|y|\le 1\}.$$

Note that if E is a complex Banach space then so is $\mathcal{L}(\mathbb{R}^d,E)$ with multiplication by complex scalars $(z,A)\in\mathbb{C}\oplus\mathcal{L}(\mathbb{R}^d,E)\mapsto zA\in\mathcal{L}(\mathbb{R}^d,E)$ defined by $(zA)(y)=z(A(y))$ for $y\in\mathbb{R}^d$.

Given $f:\mathbb{R}^d\to E$ and $x\in\mathbb{R}^d$ we denote by $\nabla f(x)$ the operator in $\mathcal{L}(\mathbb{R}^d,E)$ (if it exists) such that

$$\lim_{|y|\to 0}\|f(x+y)-f(x)-\nabla f(x)(y)\|_E|y|^{-1}=0.$$

Let $U\subset\mathbb{R}^d$ be the set of points x such that $\nabla f(x)$ exists. The function $\nabla f:U\to\mathcal{L}(\mathbb{R}^d,E)$ defined by $x\mapsto\nabla f(x)$ is called the *gradient* of f.

The *distributional gradient* of a function $f:\mathbb{R}^d\to E$ is the function $\widetilde{\nabla}f:\mathbb{R}^d\to\mathcal{L}(\mathbb{R}^d,E)$ such that

$$\int\limits_{\mathbb{R}^d}\widetilde{\nabla}f(x)(\psi(x))\,dx=-\int\limits_{\mathbb{R}^d}\operatorname{tr}J(\psi(x))f(x)\,dx\quad \text{for } \psi\in\mathcal{D}(\mathbb{R}^d,\mathbb{R}^d),\tag{1.1}$$

where $J(\psi)$ denotes the Jacobi matrix of the map $\psi:\mathbb{R}^d\to\mathbb{R}^d$ and "tr" stands for trace. If $\psi=(\psi_1,\ldots,\psi_d)$ then the right hand side of (1.1) equals

$$-\int\limits_{\mathbb{R}^d}\sum_{j=1}^d\frac{\partial\psi_j}{\partial x(j)}f\,dx.$$

The *gradient norm* in the Sobolev space $L^p_{(1)}(\mathbb{R}^d,E)$ for $1\le p\le\infty$ is defined by

$$\|f\|_{(1),p}=(\|f\|_p^p+\|\widetilde{\nabla}f\|_p^p)^{1/p}\quad \text{for } 1\le p<\infty,$$

$$\|f\|_{(1),\infty}=\max\,(\|f\|_\infty,\|\widetilde{\nabla}f\|_\infty).$$

Remarks. 1. It is easy to verify that $f:\mathbb{R}^d\to E$ has a gradient in distribution sense belonging to $L^p(\mathbb{R}^d,\mathcal{L}(\mathbb{R}^d,E))$ iff f has all first order partial derivatives in distribution sense and these derivatives belong to $L^p(\mathbb{R}^d,E)$. Thus the gradient norm $\|\cdot\|_{(1),p}$ is well defined on $L^p_{(1)}(\mathbb{R}^d,E)$ and it is compatible with the topology of L^p-convergence of functions together with their first order partial derivatives; precisely, $\lim_n\|f_n-f\|_{(1),p}=0$ iff $\lim_n\|f_n-f\|_p=0$ and $\lim_n\|\widetilde{\partial}f_n/\partial x(j)-\widetilde{\partial}f/\partial x(j)\|_p=0$ for $1\le j\le d$.

2. If $f:\mathbb{R}^d\to E$ has a distributional derivative, say $\widetilde{\nabla}f$, then $\nabla f(x)$ exists a.e. and $\nabla f(x)=\widetilde{\nabla}f(x)$ a.e.

3. Sobolev spaces of vector-valued functions on an Euclidean manifold and on a closed domain of $\mathbb{R}^d$ with sufficiently regular boundary are defined similarly to vector-valued Sobolev spaces on $\mathbb{R}^d$, analogously to scalar-valued Sobolev spaces on such sets (cf. [St], [M2]).

A few more details in the case of the d-dimensional torus $\mathbb{T}^d$. It is convenient to take as a model of $\mathbb{T}^d$ the cube $\mathbb{I}^d=[-1/2,1/2]^d$ in $\mathbb{R}^d$ whose boundary points are identified modulo the unit vector basis. The space $L^p(\mathbb{T}^d,E)$ of E-valued functions on $\mathbb{T}^d$ which are p-Bochner integrable with respect to the normalized Haar measure of $\mathbb{T}^d$ is identified with the space of E-valued functions on $\mathbb{R}^d$ which are one-periodic in each coordinate and locally p-Bochner integrable. We equip $L^p(\mathbb{T}^d,E)$ with the norm

$$\|f\|_p=\Big(\int\limits_{\mathbb{I}^d}\|f(x)\|_E^p\,dx\Big)^{1/p},\quad 1\le p\le\infty.$$

The class $\mathcal{D}(\mathbb{R}^d,E)$ of "test functions" in the case $\mathbb{R}^d$ is replaced by the class $C^\infty_\pi(\mathbb{T}^d,E)$ of E-valued C^∞ functions on $\mathbb{R}^d$ one-periodic in each coordinate.

For $f \in C^\infty_\pi(\mathbb{T}^d, E)$ and $1 \le p \le \infty$ the gradient norm on $L^p_{(1)}(\mathbb{T}^d, E)$ is defined by

$$\|f\|_{(1),p} = \Big(\int_{\mathbb{T}^d} \|f(x)\|^p_E\,dx + \int_{\mathbb{T}^d} \|\nabla f(x)\|^p_E\,dx\Big)^{1/p}.$$

For $1 \le p < \infty$ the space $L^p_{(1)}(\mathbb{T}^d, E)$ can be defined as the completion of $C^\infty_\pi(\mathbb{T}^d, E)$ in that norm.

Now we pass to the proof of the main result of this section.

THEOREM 1.1. *Let E be a Banach space and let $d = 1, 2, \ldots$ Then $f \in L^1_{(1)}(\mathbb{R}^d, E)$ implies $\|f(\cdot)\|_E \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$. Moreover,*

$$(1.2)\qquad |\widetilde{\nabla}\|f(\cdot)\|_E(x)| \le \|(\widetilde{\nabla} f)(x)\| \quad x\text{-a.e. on } \mathbb{R}^d.$$

Proof. Step 1: $d = 1$. Let $f \in L^1_{(1)}(\mathbb{R}, E)$. Let h be the distributional derivative of f. Then

$$(1.3)\qquad h \in L^1(\mathbb{R}, E) \quad\text{and}\quad \int_{\mathbb{R}} \psi h\,dx = -\int_{\mathbb{R}} \psi' f\,dx \quad \text{for } \psi \in \mathcal{D}(\mathbb{R}, \mathbb{R}).$$

It follows from (1.3) that

$$f(t) = \int_{-\infty}^{t} h(s)\,ds \quad \text{for } t \in \mathbb{R}.$$

(Obviously $h(x) = f'(x)$ for $x \in \mathbb{R}$ whenever $f \in \mathcal{D}(\mathbb{R}, E)$ because $\mathcal{D}(\mathbb{R}, E)$ is dense in $L^1_{(1)}(\mathbb{R}, E)$; in the general case the formula follows by approximation.)

Hence f is continuous, bounded, and

$$f(b) - f(a) = \int_a^b h(x)\,dx \quad \text{for } -\infty < a \le b < \infty.$$

Thus

$$(1.4)\qquad \big|\|f(b)\|_E - \|f(a)\|_E\big| \le \|f(b) - f(a)\|_E \le \int_a^b \|h(x)\|_E\,dx \quad \text{for } -\infty < a \le b < \infty.$$

It follows from (1.4) (recall that $h \in L^1(\mathbb{R}, E)$) that $\|f(\cdot)\|_E : \mathbb{R} \to \mathbb{R}$ is a (continuous) function of bounded variation. Thus there exists a unique Borel measure μ on $\mathbb{R}$ such that

$$\int_a^b d\mu = \|f(b)\|_E - \|f(a)\|_E \quad \text{for } -\infty < a \le b < \infty.$$

Moreover, μ is absolutely continuous with respect to the Lebesgue measure λ on $\mathbb{R}$. If $g = d\mu/d\lambda$ is the Radon–Nikodym derivative of μ with respect to λ then $g \in L^1(\mathbb{R}, \mathbb{R})$ and

$$(1.5)\qquad \int_a^b |g(x)|\,dx = \int_a^b d|\mu| \le \int_a^b \|h(x)\|_E\,dx \quad \text{for } -\infty < a \le b < \infty.$$

Therefore

$$(S)\int_{\mathbb{R}} \psi\,d\|f\|_E = \int_{\mathbb{R}} \psi\,d\mu = \int_{\mathbb{R}} \psi g\,dx \quad \text{for } \psi \in \mathcal{D}(\mathbb{R}, \mathbb{R}).$$

(The symbol $(S)\int_{\mathbb{R}} \ldots dw$ stands for the Riemann–Stieltjes integral with respect to a function w of bounded variation.) Using the integration by parts formula for the Stieltjes integral we get

$$(S)\int_{\mathbb{R}} \psi\,d\|f\|_E = -(S)\int_{\mathbb{R}} \|f\|_E\,d\psi \quad \text{for } \psi \in \mathcal{D}(\mathbb{R}, \mathbb{R}).$$

Obviously, if $\psi \in \mathcal{D}(\mathbb{R}, \mathbb{R})$ then $\psi' \in L^1(\mathbb{R}, \mathbb{R})$, hence

$$(S)\int_{\mathbb{R}} \|f\|_E\,d\psi = \int_{\mathbb{R}} \|f\|_E \psi'\,dx.$$

Combining the last three identities we infer that $g \in L^1(\mathbb{R}, \mathbb{R})$ is the distributional derivative of $\|f(\cdot)\|_E$. Hence $\|f(\cdot)\|_E \in L^1_{(1)}(\mathbb{R}, \mathbb{R})$. Moreover, it follows from (1.5) and the Lebesgue theorem on density points that $|g(x)| \le \|h(x)\|_E$ for λ-a.e. $x \in \mathbb{R}$.

Step 2. *If $f \in L^1_{(1)}(\mathbb{R}^d, E)$ then $\|f(\cdot)\|_E \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$, $d = 2, 3, \ldots$*

Let h be a first order distributional partial derivative of f, say $h = \widetilde{\partial} f/\partial x(1)$. Then

$$(1.6)\qquad h \in L^1(\mathbb{R}^d, E),$$

$$(1.7)\qquad \int_{\mathbb{R}^d} \psi h\,dx = -\int_{\mathbb{R}^d} \frac{\partial \psi}{\partial x(1)} f\,dx \quad \text{for } \psi \in \mathcal{D}(\mathbb{R}^d, \mathbb{R}).$$

Put $x = (t, y)$ where $t = x(1) \in \mathbb{R}$ and $y = (x(2), \ldots, x(d)) \in \mathbb{R}^{d-1}$. It follows from (1.6) and the vector-valued Fubini Theorem ([DS], Chapt. III, §11, Theorem 9) that there is a set $Z_0 \subset \mathbb{R}^{d-1}$ of full Lebesgue measure on $\mathbb{R}^{d-1}$ such that

$$(1.8)\qquad h(\cdot, y) \in L^1(\mathbb{R}, E) \quad \text{for } y \in Z_0.$$

Next consider functions $\psi \in \mathcal{D}(\mathbb{R}^d, \mathbb{R})$ depending on t and y separately, i.e. $\psi(x) = \psi_1(t)\psi_2(y)$ for $x = (t, y) \in \mathbb{R}^d$ ($\psi_1 \in \mathcal{D}(\mathbb{R}, \mathbb{R})$, $\psi_2 \in \mathcal{D}(\mathbb{R}^{d-1}, \mathbb{R})$). Fix $\psi_1 \in \mathcal{D}(\mathbb{R}, \mathbb{R})$. Combining (1.7) with the Fubini Theorem we get, for all $\psi_2 \in \mathcal{D}(\mathbb{R}^{d-1}, \mathbb{R})$,

$$\int_{\mathbb{R}^{d-1}} \psi_2(y)\Big[\int_{\mathbb{R}} \psi_1(t)h(t,y)\,dt\Big]\,dy = -\int_{\mathbb{R}^{d-1}} \psi_2(y)\Big[\int_{\mathbb{R}} \psi_1'(t)f(t,y)\,dt\Big]\,dy\,.$$

Hence there exists a set $Z(\psi_1) \subset \mathbb{R}^{d-1}$ of full Lebesgue measure such that for $y \in Z(\psi_1)$,

$$\int_{\mathbb{R}} \psi_1(t)h(t,y)\,dt = -\int_{\mathbb{R}} \psi_1'(t)f(t,y)\,dt\,. \tag{1.9}$$

Thus for every countable set $W \subset \mathcal{D}(\mathbb{R},\mathbb{R})$ there is a set $Z(W) \subset \mathbb{R}^{d-1}$ of full Lebesgue measure such that (1.9) holds for every $\psi_1 \in W$ and for every $y \in Z(W)$. If we choose W to be dense in $\mathcal{D}(\mathbb{R},\mathbb{R})$ in the topology of uniform convergence on bounded sets with all derivatives, then the standard approximation technique yields that (1.9) holds for every $\psi_1 \in \mathcal{D}(\mathbb{R},\mathbb{R})$ and for every $y \in Z(W)$.

Fix W dense in $\mathcal{D}(\mathbb{R},\mathbb{R})$ and put $Z = Z_0 \cap Z(W)$. Clearly Z is a subset of $\mathbb{R}^{d-1}$ of full Lebesgue measure. Furthermore, $h(\cdot,y) \in L^1(\mathbb{R},E)$ is the distributional derivative of $f(\cdot,y) \in L^1(\mathbb{R},E)$ for every $y \in Z$. Hence, by Step 1, the function $\|f(\cdot,y)\|_E \in L^1(\mathbb{R},\mathbb{R})$ has the distributional derivative, say $g_y \in L^1(\mathbb{R},\mathbb{R})$, for $y \in Z$. It is not clear, however, whether the function $(y,t) \to g_y(t)$ is measurable in $\mathbb{R}^d$. To bypass this obstruction we need the next construction.

For $k,n,m = 1,2,\ldots$ we put

$$A_{k,n,m} = \Big\{(t,y) \in \mathbb{R}^d : \Big|n\Big(\Big\|f\Big(t+\frac{1}{2n}\Big)\Big\|_E - \Big\|f\Big(t-\frac{1}{2n}\Big)\Big\|_E\Big) - m\Big(\Big\|f\Big(t+\frac{1}{2m}\Big)\Big\|_E - \Big\|f\Big(t-\frac{1}{2m}\Big)\Big\|_E\Big)\Big| \le \frac{1}{k}\Big\},$$

$$A = \bigcap_{k=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{m>n} A_{k,n,m}\,.$$

Obviously, for $(t,y) \in A$ the sequence $(2n(\|f(t+\frac{1}{2n})\|_E - \|f(t-\frac{1}{2n})\|_E))_{n=1}^{\infty}$ satisfies the Cauchy condition. Define $g : \mathbb{R}^d \to \mathbb{R}$ by

$$g(t,y) = \begin{cases} \lim_n n(\|f(t+\frac{1}{2n})\|_E - \|f(t-\frac{1}{2n})\|_E) & \text{for } (t,y) \in A,\\ 0 & \text{for } (t,y) \notin A.\end{cases}$$

Since $\|f(\cdot)\|_E$ is measurable, so is g. Note that if $y \in Z$ then it follows from the considerations in Step 1 that the distributional derivative g_y (defined for $y \in Z$) satisfies

$$\|f(b,y)\|_E - \|f(a,y)\|_E = \int_a^b g_y(t)\,dt\,, \quad -\infty < a \le b < \infty\,.$$

Thus, by the Lebesgue Theorem on density points, for each $y \in Z$ there is a set $Z(y) \subset \mathbb{R}$ of full Lebesgue measure on $\mathbb{R}$ such that for all $t \in Z(y)$ the point $(t,y) \in A$ and $g_y(t) = g(t,y)$. Invoking the Fubini Theorem for positive functions and taking into account that Z is of full Lebesgue measure in $\mathbb{R}^{d-1}$ we get

$$\begin{aligned}\int_{\mathbb{R}^d} |g|\,dx &= \int_{\mathbb{R}^{d-1}}\Big(\int_{\mathbb{R}} |g(t,y)|\,dt\Big)\,dy = \int_Z\Big(\int_{Z(y)} |g(t,y)|\,dt\Big)\,dy\\ &= \int_Z\Big(\int_{Z(y)} |g_y(t)|\,dt\Big)\,dy\\ &\le \int_Z\Big(\int_{Z(y)} \|h(t,y)\|_E\,dt\Big)\,dy \quad \text{(by Step 1)}\\ &= \|h\|_1\,.\end{aligned}$$

Thus $|g| \in L^1(\mathbb{R}^d,\mathbb{R})$ and therefore $g \in L^1(\mathbb{R}^d,\mathbb{R})$ because g is measurable. Hence we can use the Fubini Theorem for ψg with $\psi \in \mathcal{D}(\mathbb{R}^d,\mathbb{R})$. We have, for $\psi \in \mathcal{D}(\mathbb{R}^d,\mathbb{R})$,

$$\begin{aligned}\int_{\mathbb{R}^d} \psi g\,dx &= \int_{\mathbb{R}^{d-1}}\Big(\int_{\mathbb{R}} \psi(t,y)g(t,y)\,dt\Big)\,dy = \int_Z\Big(\int_{Z(y)} \psi(t,y)g_y(t)\,dt\Big)\,dy\\ &= -\int_Z\Big(\int_{Z(y)} \frac{\partial}{\partial t}\psi(t,y)\|f(t,y)\|_E\,dt\Big)\,dy\\ &= -\int_{\mathbb{R}^d} \frac{\partial}{\partial t}\psi(x)\|f(x)\|_E\,dx\,.\end{aligned}$$

Thus $g \in L^1(\mathbb{R}^d,\mathbb{R})$ is the distributional partial derivative of $\|f(\cdot)\|_E$ with respect to the first coordinate. Using the inequality $|g_y(t)| \le \|h(t,y)\|_E$ for $y \in Z$ and for $t \in Z_1(y)$ where $Z_1(y) \subset Z(y)$ is a set of full Lebesgue measure in $\mathbb{R}$ we also have, for every nonnegative $\psi \in \mathcal{D}(\mathbb{R}^d,\mathbb{R})$,

$$\begin{aligned}\int_{\mathbb{R}^d} \psi|g|\,dx &= \int_{\mathbb{R}^{d-1}}\Big(\int_{\mathbb{R}} \psi(t,y)|g(t,y)|\,dt\Big)\,dy\\ &= \int_Z\Big(\int_{Z_1(y)} \psi(t,y)|g_y(t)|\,dt\Big)\,dy\\ &\le \int_Z\Big(\int_{Z_1(y)} \psi(t,y)\|h(t,y)\|_E\,dt\Big)\,dy\\ &= \int_{\mathbb{R}^d} \psi(x)\|h(x)\|_E\,dx\,.\end{aligned}$$

Hence $|g(x)| \le \|h(x)\|_E$ for λ_d-a.e. $x \in \mathbb{R}^d$.

It follows from the above considerations that if $f \in L^1_{(1)}(\mathbb{R}^d, E)$ then obviously $\|f(\cdot)\|_E \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$ and the function $\|f(\cdot)\|_E$ has all distributional partial derivatives of the first order belonging to $L^1(\mathbb{R}^d, \mathbb{R})$. Hence $\|f(\cdot)\|_E \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$.

Step 3. Proof of the inequality (1.2). Fix a unit vector $\xi = (\xi(j)) \in \mathbb{R}^d$ and consider the function $f_\xi = f \circ T_\xi$ where T_ξ is the rotation of $\mathbb{R}^d$ such that $T_\xi((1,0,\ldots,0)) = \xi$. Then $f \in L^1_{(1)}(\mathbb{R}^d, E)$ implies $f_\xi \in L^1_{(1)}(\mathbb{R}^d, E)$ and the distributional partial derivative $\widetilde{\partial} f_\xi/\partial x(1)$ coincides with the distributional partial derivative in the direction of ξ, $\widetilde{\partial} f/\partial\xi = \sum_{j=1}^d \xi_j \widetilde{\partial} f/\partial x(j)$. It follows from Step 2 that $\|f(\cdot)\|_E$ has the distributional partial derivative in the direction of ξ,

$$\frac{\widetilde{\partial}\|f\|}{\partial\xi} = \sum_{j=1}^d \frac{\widetilde{\partial}\|f\|}{\partial x(j)}\xi_j,$$

and there exists a set $S(\xi) \subset \mathbb{R}^d$ of full Lebesgue measure such that

$$\left|\frac{\widetilde{\partial}\|f\|}{\partial\xi}(x)\right| \le \left\|\frac{\widetilde{\partial} f}{\partial\xi}(x)\right\|_E \quad \text{for } x \in S(\xi).$$

Let Ω be a countable set of unit vectors which is dense in the Euclidean unit sphere of $\mathbb{R}^d$. Put $S = \bigcap_{\xi\in\Omega} S(\xi)$. Then S is a subset of $\mathbb{R}^d$ of full Lebesgue measure and

$$\left|\frac{\widetilde{\partial}\|f\|}{\partial\xi}(x)\right| \le \left\|\frac{\widetilde{\partial} f}{\partial\xi}(x)\right\|_E \quad \text{for } x \in S \text{ and } \xi \in \Omega.$$

For $x \in S$ we have

$$\begin{aligned}|\nabla\|f\|(x)| &= \sup_{\xi\in\Omega}|\nabla\|f\|(x)(\xi)| = \sup_{\xi\in\Omega}\left|\frac{\widetilde{\partial}\|f\|}{\partial\xi}(x)\right| \\ &\le \sup_{\xi\in\Omega}\left\|\frac{\widetilde{\partial} f}{\partial\xi}(x)\right\|_E = \sup_{\xi\in\Omega}\|\nabla f(x)(\xi)\|_E = \|\nabla f(x)\|,\end{aligned}$$

which proves (1.2) and completes the proof of the theorem. ■

COROLLARY 1.1. *Let $1 < p < \infty$. Then $f \in L^p_{(1)}(\mathbb{R}^d, E)$ implies $\|f(\cdot)\|_E \in L^p_{(1)}(\mathbb{R}^d, \mathbb{R})$. Moreover,*

$$|\widetilde{\nabla}\|f\|(x)| \le \|\widetilde{\nabla} f(x)\| \quad \textit{for } \lambda_d\textit{-a.e. } x. \tag{1.10}$$

Proof. Pick $\eta \in \mathcal{D}(\mathbb{R}^d, \mathbb{R})$ so that $\eta \ge 0$ and $\|\eta\|_\infty = \eta(0) = 1$. Put $\eta_n(x) = \eta(xn^{-1})$ for $x \in \mathbb{R}^d$ and $n = 1, 2, \ldots$ Then $\eta_n(x) \to 1$ and $\nabla\eta_n(x) \to 0$ for $x \in \mathbb{R}^d$. Thus $(\eta_n f)(x) \to f(x)$ for $x \in \mathbb{R}^d$, and $\widetilde{\nabla}(\eta_n f)(x) \to \widetilde{\nabla} f(x)$ for a.e. $x \in \mathbb{R}^d$, because of the identity $\widetilde{\nabla}(\eta_n f) = (\nabla\eta_n) f + \eta_n \widetilde{\nabla} f$. The same relations also hold for $\|f(\cdot)\|_E$. Since $\eta_n f \in L^1_{(1)}(\mathbb{R}^d, E) \cap L^p_{(1)}(\mathbb{R}^d, E)$, Theorem 1.1 yields that $\|(\eta_n f)(\cdot)\|_E = \eta_n\|f(\cdot)\|_E \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$. Moreover, (1.2) applied to $\eta_n f$ for $n = 1, 2, \ldots$ implies (1.10) by letting $n \to \infty$. Thus $\|f(\cdot)\|_E \in L^p_{(1)}(\mathbb{R}^d, E)$. ■

Remark. After this paper has been submitted for publication S. Kwapień observed that our proof of Theorem 1.1 in fact gives

THEOREM 1.1′. *Let E be a Banach space, let $1 \le p < \infty$ and let $\varphi: E \to \mathbb{R}$ satisfy the Lipschitz condition with constant 1. Then $f \in L^p_{(1)}(\mathbb{R}^d, E)$ implies that $\varphi \circ f \in L^p_{(1)}(\mathbb{R}^d, \mathbb{R})$ and*

$$\|\widetilde{\nabla}(\varphi\circ f)(x)\| \le \|\widetilde{\nabla} f(x)\| \quad \textit{for a.e. } x \in \mathbb{R}^d.$$

He also observed that the proof of Theorem 1.1 can be simplified. Indeed, the function $g = d\mu/d\lambda$ appearing in Step 1 satisfies

$$g(x) = \limsup_{n\to\infty} n\left[\left\|f\left(x + \frac{1}{n}\right)\right\|_E - \|f(x)\|_E\right] \quad x\text{-a.e. on } \mathbb{R}.$$

This observation essentially simplifies the argument in Step 2 that the function $(t, u) \to g(t, u)$ is λ_d-measurable.

COROLLARY 1.2 (well known). *Let $f \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$. Then $f_+ = \frac{1}{2}(|f| + f)$ and $f_- = \frac{1}{2}(|f| - f)$ belong to $L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$ and the following identities hold:*

$$\begin{gathered} f = f_+ - f_-, \quad \|f\|_1 = \|f_+\|_1 + \|f_-\|_1, \\ \|\widetilde{\nabla} f\|_1 = \|\widetilde{\nabla} f_+\|_1 + \|\widetilde{\nabla} f_-\|_1. \end{gathered} \tag{1.11}$$

Proof. To prove the third identity of (1.11) note that $\widetilde{\nabla} f_+$, $\widetilde{\nabla} f_-$ and $\widetilde{\nabla} f$ are a.e. equal to the derivatives of f_+, f_- and f. Moreover, $\nabla f_-(x) = 0$ for a.e. $x \in \{f \ge 0\}$ and $\nabla f_+(x) = 0$ for a.e. $x \in \{f \le 0\}$. The other assertions of the corollary follow directly from Theorem 1.1 and linearity of $L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$. ■

Remark. Because of the "local character" of inequality (1.2) there is no difficulty in extending Theorem 1.1 and its corollaries to Sobolev spaces of E-valued functions on an Euclidean manifold as well as a domain in $\mathbb{R}^d$ with sufficiently regular boundary.

2. Molecular decomposition in $L^1_{(1)}(\mathbb{R}^d, E)$.

The main result of this section is

THEOREM 2.1. *Let $d = 1, 2, \ldots$ There exist positive constants $a = a(d)$ and $b = b(d)$ such that for every Banach space E and for every $\varepsilon > 0$, given* *a function* $f \in L^1_{(1)}(\mathbb{R}^d, E)$ *there exists a sequence* $(g_m) \subset L^1_{(1)}(\mathbb{R}^d, E)$ *such that*

$$\sum_m g_m(x) = f(x), \quad \sum_m \widetilde{\nabla} g_m(x) = \widetilde{\nabla} f(x) \quad (\lambda_d\text{-}a.e.)\,, \tag{2.1}$$

$$\sum_m \|g_m\|_1 \le (1+\varepsilon)a\|f\|_1, \quad \sum_m \|\widetilde{\nabla} g_m\|_1 \le (1+\varepsilon)a\|\widetilde{\nabla} f\|_1\,, \tag{2.2}$$

$$\|g_m\|_\infty^{1/d}\|g_m\|_1^{(d-1)/d} \le (1+\varepsilon)b\|\widetilde{\nabla} g_m\|_1 \quad \textit{for } m = 1,2,\ldots \tag{2.3}$$

The g_m's are called *molecules*.

The existence of a molecular decomposition (2.1)–(2.3) reduces to establishing it for f in an auxiliary subset of $L^1_{(1)}(\mathbb{R}^d, E)$. To define the subset we first introduce

DEFINITION 2.1. A continuous nonnegative $f : \mathbb{R}^d \to \mathbb{R}$ with compact support is called λ_d-*nonflat* provided

$$\lambda_d(\operatorname{supp} f \setminus \{f > 0\}) = \lambda_d(\{f = c\}) = 0 \quad \text{for every } c > 0\,.$$

Remark. The definition of λ_d-nonflat functions on $\mathbb{T}^d$ is the same. The role of λ_d is played by the Haar measure of $\mathbb{T}^d$. Clearly in this case every function has compact support.

We begin with two lemmas involving the concept of λ_d-nonflatness. Put

$$W^+(\mathbb{R}, d) = \{\psi \in \mathcal{D}(\mathbb{R}^d, \mathbb{R}) : \psi\ \lambda_d\text{-nonflat}\}\,,$$
$$W(E, d) = \{\psi \in \mathcal{D}(\mathbb{R}^d, E) : \|\psi(\cdot)\|_E \in W^+(\mathbb{R}, d)\}\,.$$

LEMMA 2.1. *For every Banach space* E *and for* $d = 1, 2, \ldots$ *the set* $W(E, d)$ *is dense in* $L^1_{(1)}(\mathbb{R}^d, E)$.

Proof. Pick $\psi \in \mathcal{D}(\mathbb{R}^d, \mathbb{R})$ so that $\psi \ge 0$ and $\int_{\mathbb{R}^d} \psi\, dx = 1$. Next define $\eta \in \mathcal{D}(\mathbb{R}^d, \mathbb{R})$ by $\eta(x) = h(|x|^2)$ for $x \in \mathbb{R}^d$ where $h \in \mathcal{D}(\mathbb{R}, \mathbb{R})$ is nonnegative, symmetric with respect to the origin, $h(0) = \|h\|_\infty = 1$ and for every $c > 0$ and $\gamma > 0$ the equation $h(t) = ch(\gamma t)$ has at most finitely many solutions (for instance $h(t) = \exp(-(t^2-1)^{-1}+1)$ for $|t| < 1$ and $h(t) = 0$ for $|t| \ge 1$). For $\delta > 0$ and $\varrho > 0$ define ψ_δ and η_ϱ by

$$\psi_\delta(x) = \delta^d \psi(\delta^{-1}x), \quad \eta_\varrho(x) = \eta(\varrho x) \quad \text{for } x \in \mathbb{R}^d\,.$$

For $f \in L^1_{(1)}(\mathbb{R}^d, E)$ put $f_{\varrho,\delta} = \eta_\varrho(\psi_\delta * f)$ (here and in the sequel $f * g$ denotes convolution). Repeating the argument in [St], Chapt. V, §1, Proposition 1, we infer that $f_{\varrho,\delta} \in \mathcal{D}(\mathbb{R}^d, E)$ and $\|f_{\varrho,\delta} - f\|_{(1),1} \to 0$ as $\delta \to 0$ and $\varrho \to 0$. Moreover, the nonnegativity of ψ and η combined with Theorem 1.1 yield

$$\|f_{\varrho,\delta}(\cdot)\|_E = \eta_\varrho(\psi_\varrho * \|f(\cdot)\|_E) \in \mathcal{D}(\mathbb{R}^d, \mathbb{R}) \quad (\varrho, \delta > 0)\,.$$

Thus given $\varepsilon > 0$ there exists $\delta_0 > 0$ such that for $0 < \delta < \delta_0$, $0 < \varrho < \delta_0$ one has

$$\|f_{\varrho,\delta} - f\|_{(1),1} < \varepsilon \quad \text{and} \quad \|f_{\varrho,\delta}(\cdot)\|_E \in \mathcal{D}(\mathbb{R}^d, \mathbb{R})\,.$$

Furthermore, since $\psi_\delta * \|f(\cdot)\|_E$ is strictly positive, $\|f_{\varrho,\delta}(x)\|_E = 0$ iff $\eta_\varrho = 0$. Thus

$$\lambda_d(\operatorname{supp}\|f_{\varrho,\delta}(\cdot)\|_E \setminus \{\|f_{\varrho,\delta}(\cdot)\|_E > 0\}) = \lambda_d(\operatorname{supp}\eta_\varrho \setminus \{\eta_\varrho > 0\}) = 0$$

because $\operatorname{supp}\eta_\varrho \setminus \{\eta_\varrho > 0\}$ is a $(d-1)$-dimensional Euclidean sphere in $\mathbb{R}^d$. To complete the proof we show that for every fixed δ with $0 < \delta < \delta_0$ there exists a ϱ with $0 < \varrho < \delta_0$ such that

$$\lambda_d(\{\|f_{\varrho,\delta}(\cdot)\|_E = c\}) = 0 \quad \text{for every } c > 0\,.$$

Indeed, if it were not true then for every ϱ with $0 < \varrho < \delta_0$ there would exist a $c(\varrho) > 0$ such that $\lambda_d(\{\|f_{\varrho,\delta}(\cdot)\|_E = c(\varrho)\}) > 0$. Thus there would exist ϱ and τ with $0 < \varrho < \tau < \delta_0$ such that $\lambda_d(X_{\varrho,a} \cap X_{\tau,b}) > 0$ where $a = c(\varrho)$, $b = c(\tau)$, $X_{\varrho,a} = \{\|f_{\varrho,\delta}(\cdot)\|_E = a\}$, $X_{\tau,b} = \{\|f_{\tau,\delta}(\cdot)\|_E = b\}$. Note that

$$X_{\varrho,a} \cap X_{\tau,b} = \{x \in \mathbb{R}^d : \eta_\varrho(x)(\psi_\delta * \|f(\cdot)\|_E)(x) = a \text{ and}$$
$$\eta_\tau(x)(\psi_\delta * \|f(\cdot)\|_E)(x) = b\}\,.$$

Hence $X_{\varrho,a} \cap X_{\tau,b} \subset \{\eta_\varrho\eta_\tau^{-1} = ab^{-1}\} = \{x \in \mathbb{R}^d : \eta(\varrho x)\eta^{-1}(\tau x) = ab^{-1}\}$. The latter set is, by the choice of η, either empty or a finite union of $(d-1)$-dimensional Euclidean spheres in $\mathbb{R}^d$. Hence $\lambda_d(X_{\varrho,a} \cap X_{\tau,b}) = 0$, a contradiction. ∎

The analysis of the proof of Lemma 2.1 immediately gives

COROLLARY 2.1. *The set* $W^+(\mathbb{R}, d)$ *is dense, in the norm* $\|\cdot\|_{(1),1}$, *in the cone* $\{f \in L^1_{(1),1}(\mathbb{R}^d, \mathbb{R}) : f \ge 0\}$. ∎

LEMMA 2.2. *Let* $f \ne 0$ *be a* λ_d-*nonflat function on* $\mathbb{R}^d$. *Define* $\Phi : [0, \|f\|_\infty] \to \mathbb{R}$ *by*

$$\Phi(c) = \begin{cases} \lambda_d(\{f > c\}) & \textit{for } 0 \le c < \|f\|_\infty, \\ 0 & \textit{for } c = \|f\|_\infty. \end{cases}$$

Then Φ *is continuous and strictly decreasing. Moreover, for every* $\delta > 0$ *there exists a strictly increasing sequence* $(c_m)_{m=0}^\infty$ *with* $c_0 = 0$ *and* $\lim_{m\to\infty} c_m = \|f\|_\infty$ *such that*

$$\textit{if} \quad c \in [c_{m-1}, c_m] \quad \textit{then} \quad \Phi(c) \le (1+\delta)\Phi(c_m)\,. \tag{2.4}$$

Proof. Φ is finite because f has compact support. $\Phi(t) > 0$ for $0 < t < \|f\|_\infty$ and Φ is strictly decreasing because f is continuous and therefore $\{f > t\}$ is an open set for $0 \le t < \|f\|_\infty$. Clearly $0 < t < s < \|f\|_\infty$ implies $\{f > s\} \subset \{f > t\}$ and $\{f > t\} = \bigcup_{s>t}\{f > s\}$. Thus

$$t_k \downarrow t \quad \text{implies} \quad \Phi(t_k) \uparrow \Phi(t) \quad \text{as } k \to \infty\,. \tag{2.5}$$

If $0 < s < \|f\|_\infty$ then we have $\bigcap_{t<s}\{f > t\} = \{f > s\} \cup \{f = s\}$ and $\bigcap_{t<\|f\|_\infty}\{f > t\} = \{f = \|f\|_\infty\}$. Thus it follows from λ_d-nonflatness of f that

$$t_k \uparrow t \quad \text{implies} \quad \Phi(t_k) \downarrow \Phi(t) \quad \text{as } k \to \infty\,. \tag{2.6}$$

Clearly (2.5) and (2.6) yield the continuity of Φ.

Now fix a strictly increasing sequence $(t_k)_{k=0}^{\infty}$ with $t_0 = 0$ and $\lim_k t_k = \|f\|_\infty$. Fix $\delta > 0$. Since Φ is strictly positive and continuous on $[t_{k-1}, t_k]$, the function $\log\Phi$ is continuous and therefore uniformly continuous on $[t_{k-1}, t_k]$ for $k = 1, 2, \ldots$ Thus, for $k = 1, 2, \ldots$, there exists a strictly increasing finite sequence $(t_{k,n})_{1\le n\le n(k)}$, with $t_{k,1} = t_{k-1}$ and $t_{k,n(k)} = t_k$, such that if $c \in [t_{k,n}, t_{k,n+1}]$ then $|\log\Phi(c) - \log\Phi(t_{k,n})| \le \log(1+\delta)$ for $n = 1, \ldots, n(k) - 1$. The desired sequence $(c_m)_{m=0}^{\infty}$ is the unique arrangement into a strictly increasing sequence of elements of the set $\bigcup_{k=1}^{\infty}\bigcup_{n=1}^{n(k)-1}\{t_{k,n}\}$. ∎

The next lemma reduces the proof of Theorem 2.1 to functions with "nice" properties; precisely, those from $W(E,d)$ (or from $W^+(\mathbb{R},d)$ in the scalar case).

LEMMA 2.3. *Let W denote either $W(E,d)$ or $W^+(\mathbb{R},d)$, and let $\mathrm{cl}\,W$ denote the closure of W in the norm $\|\cdot\|_{(1),1}$. If there are constants a and b such that for every $f \in W$ and every $\varepsilon > 0$ there exists (g_m) such that* (2.1)–(2.3) *hold, then the same is true for every $f \in \mathrm{cl}\,W$.*

Proof. Let $f \in \mathrm{cl}\,W$. Then for every $\varepsilon > 0$ there exists a sequence $(f_k) \subset W$ such that $f = \sum_{k=1}^{\infty} f_k$, $\sum_k \|f_k\|_1 \le (1+\varepsilon)\|f\|_1$ and $\sum_k \|\widetilde{\nabla} f_k\|_1 \le (1+\varepsilon)\|\widetilde{\nabla} f\|_1$, and moreover, for each k there is a sequence $(g_{m,k}) \subset L^1_{(1)}(\mathbb{R}^d, E)$ satisfying (2.1)–(2.3) with f_k in place of f. The construction of (f_k) is trivial for $W = W(E,d)$. For $W = W^+(\mathbb{R},d)$ we proceed as follows. For fixed $f \ne 0$, $f \in \mathrm{cl}\,W = \{f \in L^1_{(1)}(\mathbb{R}^d,\mathbb{R}) : f \ge 0\}$, and $\varepsilon > 0$ pick $\varphi_{1,1} \in W$ so that $\|f_{1,1}\| < \varepsilon\|f\|_{(1),1}$ where $f_{1,1} = f - \varphi_{1,1}$. Next define inductively for $n > 1$ and for $j = 1, \ldots, 2^{n-1}$ the functions $f_{j,n} \in \mathrm{cl}\,W \cup -\mathrm{cl}\,W$ and $\varphi_{j,n} \in W \cup -W$ so that $f_{2s-1,n} = (f_{s,n-1} - \varphi_{s,n-1})_+$, $f_{2s,n} = -(f_{s,n-1} - \varphi_{s,n-1})_-$, $\varphi_{2s-1,n} \in W$, $\varphi_{2s,n} \in -W$, and

$$\|f_{2s-1,n} - \varphi_{2s-1,n}\|_{(1),1} < 8^{-n}\varepsilon\|f\|_{(1),1}\,,$$
$$\|f_{2s,n} - \varphi_{2s,n}\|_{(1),1} < 8^{-n}\varepsilon\|f\|_{(1),1} \quad (s = 1, \ldots, 2^{n-2})\,.$$

Then

$$\|\varphi_{j,n}\|_{(1),1} < \varepsilon(8^{-n+1} + 8^{-n})\|f\|_{(1),1} \quad (j = 1, \ldots, 2^{n-1},\ n = 1, 2, \ldots)\,.$$

Hence

$$\sum_{n=1}^{\infty}\sum_{j=1}^{2^{n-1}} \|\varphi_{j,n}\|_{(1),1} < \Big(1 + \frac{2}{3}\varepsilon\Big)\|f\|_{(1),1}\,.$$

Moreover, $f = \sum_{n=1}^{\infty}\sum_{j=1}^{2^{n-1}} \varphi_{j,n}$.

Let (g_m) be an arrangement of the elements of $\bigcup_{k=1}^{\infty}\bigcup_{m=1}^{\infty}\{g_{m,k}\}$ in a sequence. Then (g_m) together with f satisfy (2.1)–(2.3) with ε replaced by $2\varepsilon + \varepsilon^2$. Note that (2.1) is a consequence of the general fact: if $\sum_m \|h_m\|_1 < \infty$ then $\sum_m h_m$ converges almost everywhere. ∎

Now we are ready to prove Theorem 2.1 for real-valued functions. Here and in the sequel we put $\beta(d) = \Gamma(d/2+1)^{1/d}/(\sqrt{\pi}d)$ for $d = 1, 2, \ldots$

PROPOSITION 2.1. *Let $f \in L^1_{(1)}(\mathbb{R}^d,\mathbb{R})$. Then for every $\varepsilon > 0$ there exists a sequence $(g_m) \subset L^1_{(1)}(\mathbb{R}^d,\mathbb{R})$ satisfying* (2.1)–(2.3) *with $a = 1$ and $b = \beta(d)$ $(d = 1, 2, \ldots)$.*

Proof. Assume first that $f \in W^+(\mathbb{R},d)$, i.e. $f \in \mathcal{D}(\mathbb{R}^d,\mathbb{R})$ and f is λ_d-nonflat. Let $\Phi : [0, \|f\|_\infty] \to \mathbb{R}$ and let the sequence $(c_k)_{k=0}^{\infty}$ satisfying (2.4) be as in Lemma 2.2 for some $\delta = \delta(\varepsilon)$. Let

$$A_m = \{f > c_{m-1}\} \setminus \{f > c_m\} \quad \text{for } m = 1, 2, \ldots$$

and define $g_m : \mathbb{R}^d \to \mathbb{R}$ by

$$g_m(x) = \begin{cases} 0 & \text{for } x \in \{f \le c_{m-1}\},\\ f(x) - c_{m-1} & \text{for } x \in A_m,\\ c_m - c_{m-1} & \text{for } x \in \{f > c_m\}. \end{cases}$$

Then $g_m \in L^1_{(1)}(\mathbb{R}^d,\mathbb{R})$. Indeed, if $\psi \in \mathcal{D}(\mathbb{R}^d,\mathbb{R})$ satisfies $\psi(x) = 1$ for $x \in \operatorname{supp} f$ then $(f - c_{m-1})\psi$ and $(f - c_m)\psi$ belong to $\mathcal{D}(\mathbb{R}^d,\mathbb{R})$. Hence, by Corollary 1.2, $g_m = ((f - c_{m-1})\psi)_+ - ((c_m - f)\psi)_- \in L^1_{(1)}(\mathbb{R}^d,\mathbb{R})$. In view of λ_d-nonflatness of f the distributional derivative of g_m is

$$\widetilde{\nabla} g_m(x) = \begin{cases} 0 & \text{for a.e. } x \in \{f \le c_{m-1}\} \setminus \{f > c_{m-1}\},\\ \nabla f(x) & \text{for a.e. } x \in A_m. \end{cases} \tag{2.7}$$

It follows from the definition of g_m that

$$\sum_{j=1}^{m} g_j(x) = \begin{cases} f(x) & \text{for } x \notin \{f > c_m\},\\ c_m & \text{for } x \in \{f > c_m\}. \end{cases}$$

Thus

$$\sum_{j=1}^{\infty} g_j(x) = f(x) \quad \text{uniformly on } \mathbb{R}^d. \tag{2.8}$$

Hence, using the positivity of the g_m's we get

$$\|f\|_1 = \int_{\mathbb{R}^d} f(x)\,dx = \sum_{m=1}^{\infty}\int_{\mathbb{R}^d} g_m(x)\,dx = \sum_{m=1}^{\infty}\|g_m\|_1\,.$$

Taking into account (2.7) and remembering that f is λ_d-nonflat we have

$$\|\nabla f\|_1 = \int_{\{f>0\}} |\nabla f(x)|\,dx = \sum_{m=1}^{\infty} \int_{A_m} |\nabla f(x)|\,dx = \sum_{m=1}^{\infty} \int_{A_m} |\widetilde{\nabla} g_m(x)|\,dx = \sum_{m=1}^{\infty} \|\widetilde{\nabla} g_m\|_1 . \tag{2.9}$$

Thus (g_m) and f satisfy (2.2) (with $a=1$) and therefore, taking also into account (2.8), they satisfy (2.1).

The verification of (2.3) is less trivial. Fix m ($m=1,2,\ldots$). Clearly $\|g_m\|_\infty = c_m - c_{m-1}$ and $g_m(x) = 0$ for $x \notin \{f > c_{m-1}\}$. Then

$$\begin{aligned} \|g_m\|_\infty^{1/d} \|g_m\|_1^{(d-1)/d} &= \|g_m\|_\infty^{1/d} \Big(\int_{\{f>c_{m-1}\}} g_m(x)\,dx \Big)^{(d-1)/d} \\ &\le \|g_m\|_\infty \lambda_d(\{f > c_{m-1}\})^{(d-1)/d} \\ &= (c_m - c_{m-1}) \Phi(c_{m-1})^{(d-1)/d} \\ &\le (c_m - c_{m-1}) \Phi(c_m)^{(d-1)/d} (1+\delta)^{(d-1)/d} \quad \text{(by (2.4))}. \end{aligned}$$

The crucial estimate in our argument follows from a version of the Federer–Kronrod coarea formula (cf. [F], Theorem 3.2.12, [BZ], Chapt. II, §2.4, [M2], 1.2.4).

If $\Psi \ge 0$ is a continuous function on an open set Ω and $\mathcal{H}_{d-1}$ denotes the $(d-1)$-dimensional Hausdorff measure in $\mathbb{R}^d$ then for every $f \in \mathcal{D}(\mathbb{R}^d, \mathbb{R})$ one has

$$\int_\Omega \Psi(x) |\nabla f(x)|\,dx = \int_0^\infty \int_{\{f=t\}\cap\Omega} \Psi(x)\, \mathcal{H}_{d-1}(dx)\,dt .$$

In particular, for $\Psi \equiv 1$ and $\Omega = \Omega_m = \{f > c_{m-1}\} \setminus \{f \ge c_m\}$ one has

$$\int_{\Omega_m} |\nabla f(x)|\,dx = \int_{c_{m-1}}^{c_m} \mathcal{H}_{d-1}(\{f=t\})\,dt \quad (m = 1,2,\ldots).$$

Obviously $\Omega_m \subset A_m$ and $\lambda_d(A_m \setminus \Omega_m) = 0$ because f is λ_d-nonflat. Hence, by (2.7),

$$\|\widetilde{\nabla} g_m\|_1 = \int_{A_m} |\widetilde{\nabla} g_m(x)|\,dx = \int_{\Omega_m} |\nabla f(x)|\,dx = \int_{c_{m-1}}^{c_m} \mathcal{H}_{d-1}(\{f=t\})\,dt . \tag{2.10}$$

Next use the observation (cf. [M2], 1.2.2) that for a.e. t in $\mathbb{R}^d$ the Hausdorff measure $\mathcal{H}_{d-1}(\{f=t\})$ equals the $(d-1)$-dimensional measure of the surface $\{f=t\}$ (the set $\{f=t\}$ is a $(d-1)$-dimensional manifold for a.e. t). Now we use the second ingredient in our proof—the isoperimetric inequality (cf. [BZ], Chapt. 2, §3.1). For a.e. $t \in \mathbb{R}$ we have

$$\begin{aligned} \mathcal{H}_{d-1}(\{f=t\}) &= \lambda_{d-1}(\{f=t\}) = \lambda_{d-1}(\partial\{f \ge t\}) \ge \lambda_{d-1}(\partial\{f > t\}) \\ &\ge \frac{\lambda_{d-1}(S_{d-1})}{\lambda_d(B_d)^{(d-1)/d}} \lambda_d(\{f>t\})^{(d-1)/d} \quad \text{(isoperimetric inequality)}, \end{aligned}$$

where $B_d = \{x \in \mathbb{R}^d : |x| \le 1\}$, $S_{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$, and ∂A denotes the boundary of a set $A \subset \mathbb{R}^d$.

A direct calculation gives

$$\lambda_{d-1}(S_{d-1}) \lambda_d(B_d)^{-(d-1)/d} = \sqrt{\pi}\, d\, \Gamma\Big(\frac{d}{2}+1\Big)^{-1/d} = \beta(d)^{-1} .$$

Thus remembering that $\lambda_d(\{f>t\}) = \Phi(t)$ and that $\Phi : [0, \|f\|_\infty] \to \mathbb{R}$ is decreasing, we get

$$\begin{aligned} \int_{c_{m-1}}^{c_m} \mathcal{H}_{d-1}(\{f=t\})\,dt &\ge \beta(d)^{-1} \int_{c_{m-1}}^{c_m} \Phi(t)^{(d-1)/d}\,dt \\ &\ge \beta(d)^{-1} \Phi(c_m)^{(d-1)/d} (c_m - c_{m-1}) . \end{aligned} \tag{2.11}$$

Combining (2.9) with (2.10) and (2.11) we get

$$\|g\|_\infty^{1/d} \|g_m\|_1^{(d-1)/d} \le \beta(d)(1+\delta)^{d/(d-1)} \|\nabla g_m\|_1 \quad (m=1,2,\ldots). \tag{2.12}$$

Choosing $\delta > 0$ so that $(1+\delta)^{d/(d-1)} < 1+\varepsilon$ we get (2.3), which completes the proof for $f \in W^+(\mathbb{R}, d)$.

Now applying Lemma 2.3 and Corollary 2.1 we infer that the assertion of Proposition 2.1 holds for nonnegative $f \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$. Finally, every $f \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$ can be written as $f = f_+ - f_-$. For fixed $\varepsilon > 0$, pick sequences (g_m^+) and (g_m^-) in $L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$ for nonnegative f_+ and f_- to satisfy (2.1)–(2.3) with $a=1$ and $b = \beta(d)$. Let $(g_m) \subset L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$ be defined by

$$g_{2k-1} = g_k^+ , \quad g_{2k} = -g_k^- \quad (k = 1,2,\ldots).$$

Then, using Corollary 1.2, we easily check that (g_m) satisfies (2.1)–(2.3) with $a=1$ and $b = \beta(d)$. ■

Remarks. 1. The constants $a=1$ and $b=\beta(d)$ are the best possible. This is clear for a. For $b = \beta(d)$ see Remark 2 after Theorem 3.2.

2. If one does not insist on getting the smallest possible constants one can prove Proposition 2.1 in a more elementary way with $a=1$ and $b=1/2$ (much larger than $\beta(d) = o(d^{-1/2})$) using the Loomis–Whitney inequality. We are indebted to Keith Ball for suggesting to us this approach.

We have to estimate the quantity $(c_m - c_{m-1})\Phi(c_m)^{(d-1)/d}$ by $\|\nabla g_m\|_1$. Put $C_m = \{f > c_m\}$ for $m = 0,1,\ldots$ Note that $\lambda_d(C_m) = \Phi(c_m)$. For $j =$

$1,2,\ldots$ denote by ${}_jC_m$ the orthogonal projection of C_m onto the hyperplane ${}_j\mathbb{R}^d = \{x \in \mathbb{R}^d : x(j) = 0\}$. Then by the Loomis–Whitney inequality (cf. [LW]; [BZ], Chapt. 2, §4.3)

$$\lambda_d(C_m)^{d-1} \le \Big(\prod_{j=1}^{d} \lambda_{d-1}({}_jC_m)\Big)^{1/d}.$$

Hence there exists a j such that

$$\lambda_d(C_m)^{(d-1)/d} \le \lambda_{d-1}({}_jC_m).$$

Denoting by e_j the unit vector perpendicular to ${}_j\mathbb{R}^d$, we have

$$\|\nabla g_m\|_1 = \int_{\mathbb{R}^d} |\widetilde{\nabla} g_m(x)|\,dx = \int_{{}_j\mathbb{R}^d} \int_{-\infty}^{\infty} |\widetilde{\nabla} g_m(u+te_j)|\,dt\,\lambda_{d-1}(du)$$
$$\ge \int_{{}_jC_m} \int_{-\infty}^{\infty} |\widetilde{\nabla} g_m(u+te_j)|\,dt\,\lambda_{d-1}(du).$$

Now fix $u \in {}_jC_m$ and put, for $k = 0,1$,

$$t_k^+ = \sup\{t : u + te_j \in C_{m-k}\}, \qquad t_k^- = \inf\{t : u+te_j \in C_{m-k}\}.$$

Since $u_j \in {}_jC_m$ and $C_m \subset C_{m-1}$, the line $\{u+te_j : -\infty < t < \infty\}$ intersects C_m and C_{m-1}. Clearly C_{m-1} and C_m are bounded and the boundary of C_{m-1} is disjoint from the closure of C_m. Hence

$$-\infty < t_0^- < t_1^- \le t_1^+ < t_0^+ < \infty.$$

Clearly $g_m(t_0^-) = g_m(t_0^+) = c_{m-1}$ and $g_m(t_1^-) = g_m(t_1^+) = c_m$. Hence the increment of g_m on the intervals $[t_0^-, t_1^-]$ and $[t_1^+, t_0^+]$ equals $c_m - c_{m-1}$ and $c_{m-1} - c_m$ respectively. Hence for every $u \in {}_jC_m$,

$$\int_{-\infty}^{\infty} |\nabla g_m(u+te_j)|\,dt \ge \int_{t_0^-}^{t_1^-} |\nabla g_m(u+te_j)|\,dt + \int_{t_1^+}^{t_0^+} |\nabla g_m(u+te_j)|\,dt$$
$$\ge 2(c_m - c_{m-1}).$$

Thus

$$\int_{{}_jC_m} \int_{-\infty}^{\infty} |\nabla g_m(u+te_j)|\,dt\,\lambda_{d-1}(du) \ge 2\lambda_{d-1}({}_jC_m)(c_m - c_{m-1})$$
$$\ge 2\Phi(c_m)^{(d-1)/d}(c_m - c_{m-1}).$$

Hence $2^{-1}\|\nabla g_m\|_1 \ge \Phi(c_m)^{(d-1)/d}(c_m - c_{m-1})$, which combined with (2.9) yields (2.3) with $b = 1/2$. ■

Now we are ready to prove Theorem 2.1 in full generality for vector-valued functions.

Proof of Theorem 2.1. By Lemmas 2.1 and 2.3 it is enough to establish the assertion of the theorem for $F \in W(E,d)$. Note that if $F \in W(E,d)$ then $f = \|F(\cdot)\|_E \in W^+(\mathbb{R},d)$. Fix $\delta > 0$ and define $(g_m)_{m=1}^{\infty} \subset L^1(\mathbb{R}^d,\mathbb{R})$ as in the proof of Proposition 2.1. Define $G_m : \mathbb{R}^d \to E$ by

$$G_m = \begin{cases} (g_m f^{-1})F & \text{for } f \ne 0,\\ 0 & \text{for } f = 0,\end{cases}$$

for $m = 1,2,\ldots$

First we check that $G_m \in L^1_{(1)}(\mathbb{R}^d, E)$ for $m = 1,2,\ldots$

For $m \ge 2$ put $g = g_m$, $G = G_m$ and $a = 2^{-1}c_m > 0$. It has been established in the proof of Proposition 2.1 that $g \in L^1_{(1)}(\mathbb{R}^d,\mathbb{R})$. Hence given $\varepsilon > 0$ there exists $\psi \in \mathcal{D}(\mathbb{R}^d,\mathbb{R})$ such that $\|g-\psi\|_1 < \varepsilon$, $\|\widetilde{\nabla} g - \nabla\psi\|_1 < \varepsilon$, $\psi(x) = 0$ whenever $f(x) \le a$.

Put $\Psi := (\psi f^{-1})F$ for $f \ne 0$ and $\Psi = 0$ for $f = 0$. Then $\Psi \in \mathcal{D}(\mathbb{R}^d, E)$ and (taking into account that ∇G exists almost everywhere) we have

$$\|\nabla\Psi - \nabla G\|_1 = \int_{\{f\ge a\}} \Big\| \frac{\nabla\psi - \nabla g}{f} F - \frac{(\psi-g)\nabla f}{f^2} F + \frac{(\psi - g)\nabla F}{f}\Big\|_E\,dx$$
$$\le \varepsilon + (\varepsilon\|\nabla f\|_\infty + \varepsilon\|\nabla F\|_\infty)a^{-1}.$$

(Let $h : \mathbb{R}^d \to \mathbb{C}$, $F : \mathbb{R}^d \to E$. We use the formula $\nabla(hF) = \nabla h\cdot F + h\cdot\nabla F$; the meaning of the second term is obvious: $(h\nabla F)(x) = h(x)(\nabla F)(x) \in \mathcal{L}(\mathbb{R}^d, E)$; the operator $(\nabla h\cdot F)(x)$ is defined by

$$(\nabla h\cdot F)(x)(\xi) = (\nabla h)(x)(\xi)\cdot F(x) \quad \text{for } \xi \in \mathbb{R}^d.)$$

On the other hand, $\|\Psi - G\|_1 = \|\psi - g\|_1 < \varepsilon$.

Thus G can be approximated in the norm $\|\cdot\|_{(1),1}$ by functions from $\mathcal{D}(\mathbb{R}^d, E)$. Hence $G \in L^1_{(1)}(\mathbb{R}^d, E)$.

Next consider the case of G_1. Pick $h \in \mathcal{D}(\mathbb{R}^d,\mathbb{R})$ so that $h(x) = 0$ for $x \in \{f \le 4^{-1}c_1\}$ and $h(x) = 1$ for $x \in \{f \ge 2^{-1}c_1\}$. Write $G_1 = hG_1 + (1-h)G_1$. Note that $(1-h)G_1 = (1-h)F$. Thus $(1-h)G_1 \in \mathcal{D}(\mathbb{R}^d, E) \subset L^1(\mathbb{R}^d, E)$. Put $g = hg_1$, $G = gf^{-1}F$ for $f \ne 0$ and $G = 0$ for $f = 0$. Note that $g \in L^1_{(1)}(\mathbb{R}^d,\mathbb{R})$. (Indeed, if $\phi \in \mathcal{D}(\mathbb{R}^d,\mathbb{R})$ satisfies $\operatorname{supp} f \subset \{\phi = 1\}$ then $\phi(hf - c_1) \in \mathcal{D}(\mathbb{R}^d,\mathbb{R})$ and $g = (hf)_+ - (\phi(hf - c_1))_-$; the desired relation follows from Corollary 1.2.) Now the proof that $G \in L^1_{(1)}(\mathbb{R}^d, E)$ is the same as that for $G = G_m$ for $m \ge 2$ (with $a = 4^{-1}c_1$). Finally, note that $hG_1 = G$. Hence $G_1 \in L^1_{(1)}(\mathbb{R}^d, E)$.

Now we are ready to verify that (G_m) is the desired decomposition of F. Observe that

$$\|G_m(x)\|_E = g_m(x) \quad \text{for } x \in \mathbb{R}^d \ (m = 1,2,\ldots).$$

Thus $\|G_m\|_\infty = \|g_m\|_\infty$ and $\|G_m\|_1 = \|g_m\|_1$. Hence, by Propositon 2.1 (more specifically, by (2.12)) and by Theorem 1.1, if $1+\varepsilon > (1+\delta)^{(d-1)/d}$ then

$$\|G_m\|_\infty^{1/d}\|G_m\|_1^{(d-1)/d} = \|g_m\|_\infty^{1/d}\|g_m\|_1^{(d-1)/d} \le \beta(d)(1+\varepsilon)\|\widetilde{\nabla} g_m\|_1 \le \beta(d)(1+\varepsilon)\|\widetilde{\nabla} G_m\|_1 \quad (m=1,2,\ldots).$$

Hence the sequence (G_m) satisfies (2.3) with $b=\beta(d)$.

Since $F \in W(E,d)$ and therefore $f \in W^+(\mathbb{R},d)$, it follows from the definition of G_m and the proof of Proposition 2.1 that

$$\sum_{m=1}^{\infty} G_m(x) = F(x) \quad \text{uniformly for } x \in \mathbb{R}^d \tag{2.13}$$

and

$$\sum_{m=1}^{\infty} \|G_m\|_1 = \sum_{m=1}^{\infty} \|g_m\|_1 = \|f\|_1 = \|F\|_1 . \tag{2.14}$$

Now to complete the proof of the theorem it is enough to show that

$$\sum_{m=1}^{\infty} \|\widetilde{\nabla} G_m\|_1 \le 3\|\nabla F\|_1 . \tag{2.15}$$

Indeed, (2.14) and (2.15) imply that $\sum_m G_m$ converges absolutely in $L^1_{(1)}(\mathbb{R}^d,E)$. Thus the completeness of $L^1_{(1)}(\mathbb{R}^d,E)$ and (2.13) yield that $F=\sum_m G_m$. Furthermore, (2.15) implies that $\sum_m \widetilde{\nabla} G_m(x)$ converges a.e. on $\mathbb{R}^d$. Hence $\nabla F(x) = \sum_m \widetilde{\nabla} G_m(x)$ for λ_d-a.e. $x \in \mathbb{R}^d$.

To prove (2.15) put as before $C_{m-1} = \{f > c_{m-1}\}$ and $A_m = C_{m-1}\setminus C_m$ for $m=1,2,\ldots$ Note that

$$\|\widetilde{\nabla} G_m\|_1 = \int_{C_{m-1}} \|\widetilde{\nabla} G_m(x)\|\,dx \quad (m=1,2,\ldots).$$

Moreover, for λ_d-a.e. $x \in C_{m-1}$, we have

$$\widetilde{\nabla} G_m(x) = (f^{-1}(\widetilde{\nabla} g)F)(x) - (g_m f^{-2}(\nabla f)F)(x) + (g_m f^{-1}\nabla F)(x).$$

Taking into account that $g_m(x)=0$ for $x \notin C_{m-1}$ and $\widetilde{\nabla} g_m(x) = 0$ for $x \notin A_m$ we get

$$\sum_{m=1}^{\infty} \|\widetilde{\nabla} G_m\|_1 \le I_1 + I_2 + I_3 ,$$

where

$$I_1 = \sum_{m=1}^{\infty} \int_{A_m} |\widetilde{\nabla} g_m(x)|\,dx ,$$

$$I_2 = \sum_{m=1}^{\infty} \int_{C_{m-1}} g_m f^{-2}|\nabla f|\|F(\cdot)\|_E\,dx ,$$

$$I_3 = \sum_{m=1}^{\infty} \int_{C_{m-1}} g_m f^{-1}\|\nabla F\|\,dx .$$

By (2.9), in view of Theorem 1.1, one gets

$$I_1 = \sum_{m=1}^{\infty} \|\widetilde{\nabla} g_m\|_1 = \|\nabla f\|_1 \le \|\nabla F\|_1 .$$

Next using the identity $C_{m-1} = \bigcup_{k=m}^{\infty} A_k$ for $m=1,2,\ldots$ and remembering that the A_k's are mutually disjoint we get

$$\begin{aligned}
I_2 &\le \sum_{m=1}^{\infty} \int_{C_{m-1}} g_m f^{-1}|\nabla f|\,dx = \sum_{m=1}^{\infty}\sum_{k=m}^{\infty} \int_{A_k} g_m f^{-1}|\nabla f|\,dx \\
&= \sum_{k=1}^{\infty}\sum_{m=1}^{k} \int_{A_k} g_m f^{-1}|\nabla f|\,dx = \sum_{k=1}^{\infty} \int_{A_k} \sum_{m=1}^{k} g_m f^{-1}|\nabla f|\,dx \\
&\le \sum_{k=1}^{\infty} \int_{A_k} |\nabla f(x)|\,dx \quad \text{(by (2.8))} \\
&= \int_{C_0} |\nabla f(x)|\,dx = \|\nabla f\|_1 \le \|\nabla F\|_1 \quad \text{(by Theorem 1.1)}.
\end{aligned}$$

Similarly

$$\begin{aligned}
I_3 &= \sum_{m=1}^{\infty}\sum_{k=m}^{\infty} \int_{A_k} g_m f^{-1}\|\nabla F(\cdot)\|\,dx = \sum_{k=1}^{\infty}\sum_{m=1}^{k} \int_{A_k} g_m f^{-1}\|\nabla F(\cdot)\|\,dx \\
&= \sum_{k=1}^{\infty} \int_{A_k} \|\nabla F(\cdot)\|\,dx = \|\nabla F\|_1 .
\end{aligned}$$

This completes the proof of (2.15). Thus we have proved Theorem 2.1 with the constants $a=3$ and $b=\beta(d)$. ∎

Remarks. 1. The constant $a=3$ obtained in the proof of Theorem 2.1 does not seem to be the best possible. For complex scalars, i.e. for the space $L^1_{(1)}(\mathbb{R}^d,\mathbb{C})$ one gets $a \le \sqrt{2}$. To see this, write $f = \sum_{j=0}^{3} i^j f_j$, where

$f_0 = (\Re f)_+$, $f_1 = (\Im f)_+$, $f_3 = (\Re f)_-$, $f_4 = (\Im f)_-$, and apply the same argument as at the end of the proof of Proposition 2.1.

2. For the space $L^1_{(1)}(\mathbb{R}^d, l^1)$ (here l^1 denotes the space of real absolutely convergent series) one gets $a = 1$ and $b = \beta(d)$ by decomposing a function in $L^1_{(1)}(\mathbb{R}^d, l^1)$ into induced coordinate functions and using the special property of the norm in l^1. By the localization technique we get from the result for l^1 the same constants for $L^1_{(1)}(\mathbb{R}^d, E)$ where E is an arbitrary abstract L^1-space.

3. For every $f \in L^1_{(1)}(\mathbb{R}^d, E)$ and every $\varepsilon > 0$ there exists a molecular decomposition consisting of functions belonging to $W(E, d)$.

Proof of Remark 3. Let (g_j) be any molecular decomposition of $f \neq 0$ satisfying (2.1)–(2.3) with given $\varepsilon > 0$. For $k = 1, 2, \ldots$ let $h_k = (g_{\varrho_k, \delta_k})$ be a regularization of $g = g_k$, defined as in the proof of Lemma 2.1, such that $h_k \in W(E, d)$, $\|h_k - g_k\|_{(1),1} < \varepsilon 2^{-k} \min(\|f\|_{(1),1}, \|g_k\|_{(1),1})$, $\|h_k\|_\infty \le \|g_k\|_\infty$, $\|h_k\|_1 \le \|g_k\|_1$. Put $h = \sum h_k$. Then (h_k) is a molecular decomposition of h, precisely (h_k) and h satisfy (2.1)–(2.3) with $\tilde{\varepsilon}_1 < 2\varepsilon + \varepsilon^2$. Moreover, $\|h - f\|_{(1),1} < (1+\varepsilon)\|f\|_{(1),1}$. Since $\varepsilon > 0$ can be chosen arbitrarily small, one can slightly modify the proof of Lemma 2.3 to get the desired molecular decomposition of f. ∎

Theorem 2.1 extends to Sobolev spaces on $\mathbb{T}^d$ as follows.

THEOREM 2.1_π. *Let E be a Banach space and let $d = 1, 2, \ldots$ There exist positive constants $a_\pi = a_\pi(d)$ and $b_\pi = b_\pi(d)$ such that for every Banach space E and for every $\varepsilon > 0$, given a function $f \in L^1_{(1)}(\mathbb{T}^d, E)$ there exists a sequence $(g_m) \subset L^1_{(1)}(\mathbb{T}^d, E)$ such that*

$$(2.1)_\pi \qquad \sum_m g_m(x) = f(x), \quad \sum_m \widetilde{\nabla} g_m(x) = \widetilde{\nabla} f(x)$$

for a.e. x with respect to the Haar measure of $\mathbb{T}^d$,

$$(2.2)_\pi \qquad \sum_m \|g_m\|_1 \le (1+\varepsilon) a_\pi \|f\|_1, \quad \sum_m \|\widetilde{\nabla} g_m\|_1 \le (1+\varepsilon) a_\pi \|\widetilde{\nabla} f\|_1,$$

$$(2.3)_\pi \qquad \|g_m\|_\infty^{1/d} \|g_m\|_1^{(d-1)/d} \le (1+\varepsilon) b_\pi (\|\widetilde{\nabla} g_m\|_1 + \|g_m\|_1)$$

for $m = 1, 2, \ldots$

Outline of the proof. Define $v : \mathbb{R} \to \mathbb{R}$ and $v_d : \mathbb{R}^d \to \mathbb{R}$ for $d = 1, 2, \ldots$ by

$$v(t) = \begin{cases} 1 & \text{for } |t| \le 1/2, \\ 3/2 - |t| & \text{for } 1/2 < |t| \le 3/2, \\ 0 & \text{for } |t| > 3/2, \end{cases}$$

$$v_d(x) = \prod_{j=1}^d v(x(j)) \quad \text{for } x = (x(j)) \in \mathbb{R}^d.$$

For $f \in L^1_{(1)}(\mathbb{T}^d, E)$ regarded as an E-valued function on $\mathbb{R}^d$, 1-periodic with respect to each coordinate, put $f^\# = v_d f$. Then $f^\# \in L^1_{(1)}(\mathbb{T}^d, E)$; moreover, if f is λ_d-nonflat so is $f^\#$. Next for $f^\#$ and for $\varepsilon > 0$ construct in the same way as in the proof of Theorem 2.1 a sequence $(g_m^0) \subset L^1_{(1)}(\mathbb{T}^d, E)$ satisfying (2.1)–(2.3). The analysis of the proof of Theorem 2.1 shows that for $m = 1, 2, \ldots$ the restriction of g_m^0 to the cube $\mathbb{I}^d = [-1/2, 1/2]^d$ has a unique extension to an E-valued function on $\mathbb{R}^d$, say g_m, which is 1-periodic with respect to each coordinate. Moreover, $g_m \in L^1_{(1)}(\mathbb{T}^d, E)$. The sequence (g_m) satisfies $(2.1)_\pi$ and $(2.2)_\pi$ with the same constant a which appears in the proof of Theorem 2.1. Indeed, note that to establish (2.1) and (2.2) in the proof of Theorem 2.1 one uses pointwise estimates only, so we can work with functions restricted to $\mathbb{I}^d$. To establish $(2.3)_\pi$ we use the inequalities

$$\|g_m\|_\infty = \|g_m^0\|_\infty, \quad \|g_m\|_1 \le \|g_m^0\|_1,$$
$$\|\widetilde{\nabla} g_m^0\|_1 \le A(d)(\|\widetilde{\nabla} g_m\|_1 + \|g_m\|_1) \quad (m = 1, 2, \ldots),$$

where $A(d)$ is a numerical constant depending only on d but independent of f, the g_m's and ε. The first two inequalities are trivial; the third follows by analyzing the construction of $f^\#$ and the g_m's. Thus, by (2.3), for $m = 1, 2, \ldots$ we have

$$\|g_m\|_\infty^{1/d} \|g_m\|_1^{(d-1)/d} \le \|g_m^0\|_\infty^{1/d} \|g_m^0\|_1^{(d-1)/d} \le (b(d)+\varepsilon) \|\widetilde{\nabla} g_m^0\|_1$$
$$\le (b(d)+\varepsilon) A(d) \|\widetilde{\nabla} g_m\|_1 . ∎$$

Remark. A similar argument shows that Theorem 2.1 on molecular decompositions extends to Sobolev spaces $L^1_{(1)}(\Omega, E)$ where Ω is a compact domain in $\mathbb{R}^d$ such that there exists a linear extension operator of Whitney type $\mathcal{E} : C^1(\Omega, \mathbb{R}) \to C^1(\mathbb{R}^d, \mathbb{R})$ which extends to continuous operators $L^1(\Omega, \mathbb{R}) \to L^1(\mathbb{R}^d, \mathbb{R})$, $L^\infty(\Omega, \mathbb{R}) \to L^\infty(\mathbb{R}^d, \mathbb{R})$ and $L^1_{(1)}(\Omega, \mathbb{R}) \to L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$. For the existence of such operators cf. [St], Chapt. VI, §2; [Hö], Theorem 2.3.6, and [Jo]. To extend the result to Sobolev spaces on a compact d-dimensional manifold we use a partition of the manifold into charts which are nice domains in $\mathbb{R}^d$ and apply a version of the theorem to each chart separately.

3. Refinements of the Sobolev embedding theorem for $L^1_{(1)}(\mathbb{R}^d, E)$.

In this section we apply the molecular decompositions to obtain improvements of the classical Gagliardo–Nirenberg embedding $L^1_{(1)}(\mathbb{R}^d, \mathbb{C}) \hookrightarrow L^{d/(d-1)}(\mathbb{R}^d, \mathbb{C})$. Most of the results presented here have been discovered in the case of scalar-valued functions by Bourgain [Br1], Poornima [Po], and Kolyada [K].

If X and Y are Banach spaces of functions with a common domain then we say that X embeds into Y, in symbols $X \hookrightarrow Y$, provided that X is a subset of Y and the identity inclusion map is continuous.

We begin with introducing Lorentz and Besov spaces of vector-valued functions.

Vector-valued Lorentz spaces. Given a measure space (Ω, Σ, μ) and a scalar-valued Σ-measurable function $h : \Omega \to \mathbb{C}$ we put

$$\mu_h(y) = \mu\{|h| > y\} \quad \text{for } y > 0,$$
$$h^*(t) = \inf\{y > 0 : \mu_h(y) \le t\} \quad \text{for } t > 0,$$
$$\|h\|_{p,q} = \begin{cases} \left(\frac{q}{p}\int\limits_0^\infty (t^{1/p}h^*(t))^q \frac{dt}{t}\right)^{1/q} & (0 < p < \infty,\ 0 < q < \infty), \\ \sup\limits_{t>0} t^{1/p}h^*(t) & (0 < p \le \infty,\ q = \infty) \end{cases}$$

(cf. [BS], p. 216; [H]). Given a Banach space E, we denote by $L_{p,q}(\Omega, \mu, E)$ the space of μ-equivalence classes of Σ-measurable functions $f : \Omega \to E$ such that $\|\|f(\cdot)\|_E\|_{p,q} < \infty$ $(0 < p \le \infty,\ 0 < q \le \infty)$. If $1 \le p \le \infty$ and $1 \le q \le \infty$ then $L_{p,q}(\Omega, \mu, E)$ is a Banach space under the norm $\|f\|_{p,q} = \|\|f(\cdot)\|_E\|_{p,q}$. For μ being the Lebesgue measure on $\mathbb{R}^d$ we write $L_{p,q}(\mathbb{R}^d, E)$ instead of $L_{p,q}(\Omega, \mu, E)$.

One has $L_{p,p}(\mathbb{R}^d, E) = L^p(\mathbb{R}^d, E)$; moreover, $\|f\|_p = \|f\|_{p,p}$ for $0 < p \le \infty$.

We shall need the following interpolation inequality:

$$\|f\|_{d/(d-1),1} \le \|f\|_\infty^{1/d}\|f\|_1^{(d-1)/d} \quad \text{for } f \in L_{d/(d-1),1}(\mathbb{R}^d, E). \tag{3.1}$$

The inequality (3.1) is well known in the scalar-valued case (cf. [BS], Chapt. 4, §4, Proposition 4.2). The vector-valued case follows immediately from the scalar one.

Besov spaces of vector-valued functions. Most of the definitions of Besov spaces of scalar-valued functions extend trivially to the case of functions taking values in a fixed Banach space. This is due to the observation that estimates of L^p-norms of products and convolutions of scalar-valued functions are usually preserved if exactly one of the functions is replaced by a vector-valued function. A typical example is:

LEMMA 3.1. *If $1 \le p \le \infty$, $1 \le p_1 \le \infty$, $1/r = 1/p + 1/p_1 - 1 > 0$, $\varphi \in L^p(\mathbb{R}^d, \mathbb{C})$, and $f \in L^{p_1}(\mathbb{R}^d, E)$ where E is a complex Banach space then $\varphi * f \in L^r(\mathbb{R}^d, E)$ and $\|\varphi * f\|_r \le \|\varphi\|_p\|f\|_{p_1}$.*

Proof. Cf. [Hö], Theorem 4.5.1 and Corollary 4.5.2. The proofs presented there for scalar-valued functions can be easily adapted to our case. ∎

To define vector-valued Besov spaces we first introduce an auxiliary family of partitions of unity "on the level of the Fourier Transform".

Let $\Psi(d)$ be the family of all sequences $\Psi = (\psi_k)_{k=0}^\infty$ of continuous scalar-valued functions on $\mathbb{R}^d$ such that:

(i) $\operatorname{supp}\psi_0 \subset \{|\xi| \le 2\}$, $\operatorname{supp}\psi_k \subset \{2^{k-1} \le |\xi| \le 2^{k+1}\}$ for $k = 1, 2, \ldots$,

(ii) $\sup_k \|\widehat{\psi}_k\|_1 = a_\Psi < \infty$,

(iii) $\sum_{k=0}^\infty \psi_k(\xi) = 1$ for $\xi \in \mathbb{R}^d$.

Now let E be a complex Banach space. Let $1 \le p \le \infty$, $1 \le q \le \infty$, $0 \le \theta < 1$.

For a fixed $\Psi \in \Psi(d)$ we define on $\mathcal{D}(\mathbb{R}^d, \mathbb{C})$ the norms

$$\|f\|_{B^\theta_{p,q}(\cdot;\Psi)} = \|f\|_p + B^\theta_{p,q}(f;\Psi),$$
$$B^\theta_{p,q}(f;\Psi) = \Big[\sum_{k=0}^\infty (2^{k\theta}\|\widehat{\psi}_k * f\|_p)^q\Big]^{1/q}. \tag{3.2}$$

Our notation differs from most of the books where the functions of Ψ are replaced by their Fourier Transforms.

LEMMA 3.2. *For every $\Psi, \Phi \in \Psi(d)$ the norms $B^\theta_{p,q}(\cdot;\Psi)$ and $B^\theta_{p,q}(\cdot;\Phi)$ are equivalent; precisely, there exists a $C = C(\Psi, \Phi) \ge 1$ such that*

$$C^{-1}B^\theta_{p,q}(f;\Psi) \le B^\theta_{p,q}(f;\Phi) \le CB^\theta_{p,q}(f;\Psi) \quad \text{for } f \in \mathcal{D}(\mathbb{R}^d, E). \tag{3.3}$$

Proof. Consider the nontrivial case $q < \infty$. Fix Ψ and Φ in $\Psi(d)$ and $f \in \mathcal{D}(\mathbb{R}^d, E)$. Put for convenience $\psi_{-1} = \phi_{-1} = 0$. Using (i) and (iii) for Φ and (i) for Ψ we get, for $k = 0, 1, \ldots$,

$$\widehat{\psi}_k * f = [\psi_k(\phi_{k-1} + \phi_k + \phi_{k+1})]^\wedge * f = \sum_{m=k-1}^{k+1} \widehat{\psi}_k * \widehat{\phi}_m * f.$$

Hence by (ii), Lemma 3.1 and the Hölder inequality, for $k = 0, 1, \ldots$,

$$\|\widehat{\psi}_k * f\|_p^q \le [a_\Psi]^q 3^{q-1} \sum_{m=k-1}^{k+1} \|\widehat{\phi}_m * f\|_p^q.$$

Thus summing over k with weights $2^{k\theta q}$ and using the triangle inequality we get

$$B^\theta_{p,q}(f;\Psi) \le 6a_\Psi B^\theta_{p,q}(f;\Phi).$$

Hence by symmetry we get (3.3) with $C = 6\max(a_\Psi, a_\Phi)$. ∎

DEFINITION 3.1. Let $1 \le p \le \infty$, $1 \le q \le \infty$, $0 \le \theta < 1$. The *Besov space* $B^{\theta}_{p,q}(\mathbb{R}^d, E)$ is the completion of $\mathcal{D}(\mathbb{R}^d, E)$ in any of the equivalent norms (3.2) for $\Psi \in \Psi(d)$.

Remarks. 1. Let E be a real Banach space. Let $E_{\mathbb{C}}$ denote the complexification of E. Then the Besov space $B^{\theta}_{p,q}(\mathbb{R}^d, E)$ is defined to be the real subspace of $B^{\theta}_{p,q}(\mathbb{R}^d, E_{\mathbb{C}})$ of functions taking values in E identified with its natural embedding in $E_{\mathbb{C}}$.

2. The definition of Besov spaces on $\mathbb{T}^d$ is similar. The only difference is that the periodic analog of the family $\Psi(d)$ consists of all sequences of functions (= sequences) on $\mathbb{Z}^d$ satisfying obviously modified conditions (i)–(iii).

3. Our definition of Besov spaces slightly differs from that of [Pee], p. 48, where it is assumed that $(\psi_k) \subset \mathcal{D}(\mathbb{R}^d, \mathbb{C})$ and our condition (ii) is replaced by

(ii)′ for every multiindex α of length d there exists $c_\alpha > 0$ such that

$$|D^\alpha \psi_k(x)| \le c_\alpha 2^{-|\alpha| k} \quad (x \in \mathbb{R}^d,\ k = 0, 1, \ldots).$$

One can show that (ii)′ implies (ii) (cf. [Pee], p. 48) by a similar technique to [Hö], Chapt. 7, §1. Now it follows from Lemma 3.2 that our definition of Besov spaces is equivalent to the original definition of Peetre.

Next we recall

LEMMA 3.3. *If* $0 \le 1/p - 1/r < 1/d$, $p \ge 1$, $r \ge 1$ *and* $d = 1, 2, \ldots,$ *then*

$$B^{d(1/p-1/r)}_{p,q}(\mathbb{R}^d, E) \hookrightarrow B^0_{r,q}(\mathbb{R}^d, E) \quad (1 \le q \le \infty).$$

Moreover, there exists $A(d) > 0$ *such that for every* $\Psi \in \Psi(d)$,

$$B^0_{r,q}(f; \Psi) \le A(d) B^{d(1/p-1/r)}_{p,q}(f; \Psi) \quad \text{for } f \in B^0_{r,q}(\mathbb{R}^d, E).$$

In particular,

$$B^{\theta(p,d)}_{p,1}(\mathbb{R}^d, E) \hookrightarrow B^0_{d/(d-1),1}(\mathbb{R}^d, E) \quad \left(1 < p \le \frac{d}{d-1},\ d = 1, 2, \ldots\right),$$

where $\theta(p, d) = d(1/p + 1/d - 1)$.

Proof. For $k = 0, 1, \ldots$ let $V_{k,d} : \mathbb{R}^d \to \mathbb{R}$ be the Fourier Transform of the d-dimensional de la Vallée-Poussin kernel; precisely,

$$V_{k,d}(x) = \prod_{j=1}^{d} v_{k,d}(x(j)) \quad \text{for } x = (x(j)) \in \mathbb{R}^d,$$

where $v_{k,d} : \mathbb{R} \to \mathbb{R}$ is defined by

$$v_{k,d}(t) = \begin{cases} 1 & \text{for } |t| \le 2^{k+1}, \\ \dfrac{2^{k+2} - |t|}{2^{k+2} - 2^{k+1}} & \text{for } 2^{k+1} < |t| \le 2^{k+2}, \\ 0 & \text{for } |t| > 2^{k+2}. \end{cases}$$

It is well known and easy to check that there exists a constant $A(d) > 0$ independent of k such that $\|\widehat{V}_{k,d}\|_1 \le A(d)$ and $\|\widehat{V}_{k,d}\|_\infty \le A(d) 2^{kd}$. Hence

$$\|\widehat{V}_{k,d}\|_{p_1} \le A(d) 2^{kd(1-1/p_1)} \quad \text{for } 1 \le p_1 \le d.$$

Next fix $\Psi \in \Psi(d)$. Then for $f \in \mathcal{D}(\mathbb{R}^d, E)$,

$$\|\widehat{V}_{k,d} * \widehat{\psi}_k * f\|_r = \|(V_{k,d} \cdot \psi_k)^\wedge * f\|_r = \|\widehat{\psi}_k * f\|_r .$$

Thus by Lemma 3.1, for $p_1 = (1/r - 1/p + 1)^{-1}$ we get

$$\|\widehat{\psi}_k * f\|_r \le \|\widehat{V}_{k,d}\|_{p_1} \|\widehat{\psi}_k * f\|_p \le A(d) 2^{kd(1/p-1/r)} .$$

Hence $B^0_{r,q}(f; \Psi) \le A(d) B^{d(1/p-1/r)}_{p,q}(f; \Psi)$ for every $f \in \mathcal{D}(\mathbb{R}^d, E)$, which is equivalent to the desired embedding. ∎

For $\theta > 0$ we introduce two other norms on Besov spaces. They are equivalent to the norm defined by (3.2).

The first involves the concept of pth modulus of smoothness. For $f \in L^p(\mathbb{R}^d, E)$ $(1 \le p < \infty)$, we put

$$\omega_p(f, t) = \sup_{|h| \le t} \|f_h - f\|_p$$

where $f_h(x) = f(x + h)$ for $x \in \mathbb{R}^d$. Next for $\theta > 0$ and $1 \le q < \infty$ we put

$$B^{\theta}_{p,q}(f) = \left[\int_0^\infty [t^{-\theta} \omega_p(f, t)]^q \frac{dt}{t} \right]^{1/q},$$

and

$$\|f\|_{B^{\theta}_{p,q}} = B^{\theta}_{p,q}(f) + \|f\|_p . \tag{3.4}$$

The second norm involves the degree of approximation of a function by entire functions of finite order.

For $f \in L^p(\mathbb{R}^d, E)$ we put

$$E_p(f, t) = \inf\{\|f - g\|_p : \operatorname{supp} \widehat{g} \subset \{|\xi| \le t\}\} \quad (t > 0,\ 1 \le p < \infty),$$

$$\overline{B}^{\theta}_{p,q}(f) = \left[\int_1^\infty [t^{\theta} E_p(f, t)]^q \frac{dt}{t} \right]^{1/q} \quad (\theta > 0,\ 1 \le q < \infty),$$

and

$$\|f\|_{\overline{B}^{\theta}_{p,q}} = \|f\|_p + \overline{B}^{\theta}_{p,q}(f). \tag{3.5}$$

PROPOSITION 3.1. *Let $1 \le p < \infty$, $1 \le q < \infty$, $0 < \theta < 1$. Then for every Banach space E the norms defined by (3.2), (3.4) and (3.5) are equivalent on $\mathcal{D}(\mathbb{R}^d, E)$. Moreover, the Besov space $B^\theta_{p,q}(\mathbb{R}^d, E)$ can be identified with the set*

$$\{f \in L^p(\mathbb{R}^d, E) : B^\theta_{p,q}(f) < \infty\} = \{f \in L^p(\mathbb{R}^d, E) : \overline{B}^\theta_{p,q}(f) < \infty\}.$$

Proof. The equivalence of the norms (3.2) and (3.5) can be established by repeating the argument presented in [Pee], pp. 72–74 (Chapt. 3, proof of Theorem 11). The only ingredients used there are the triangle inequality and the scalar version of Lemma 3.1.

To prove the equivalence of the norms (3.4) and (3.5) we first show that for some $C > 0$,

$$E_p(f, r) \le C\omega_p(f, r^{-1}), \quad r > 0,\ f \in \mathcal{D}(\mathbb{R}^d, E). \tag{3.6}$$

Fix $\varphi : \mathbb{R}^d \to \mathbb{C}$ such that

(j) $\int_{\mathbb{R}^d} (|x| + 1)|\varphi(x)|\, dx = C < \infty$,
(jj) $\int_{\mathbb{R}^d} \varphi(x)\, dx = 1$,
(jjj) $\operatorname{supp} \widehat{\varphi} \subset \{\xi \in \mathbb{R}^d : |\xi| \le 1\}$.

For instance define φ by $\widehat{\varphi}(\xi) = a\eta(\xi)\exp(-|\xi|^2)$ where $\eta \in \mathcal{D}(\mathbb{R}^d, \mathbb{C})$ satisfies $\eta(0) = 1$, $\operatorname{supp} \eta \subset \{\xi \in \mathbb{R}^d : |\xi| \le 1\}$ and a is determined by (jj).

Next for $r > 0$ put $\varphi_r(x) = r^d\varphi(rx)$ for $x \in \mathbb{R}^d$. Then, by (jjj), $\operatorname{supp}[\varphi_r * f]^\wedge = \operatorname{supp}(\widehat{\varphi}_r\widehat{f}) \subset \{|\xi| \le r\}$. Hence

$$E_p(f, r) \le \|f - \varphi_r * f\|_p \quad \text{for } r > 0.$$

By (jj),

$$(f - \varphi_r * f)(x) = \int_{\mathbb{R}^d} \varphi_r(h)[f(x) - f(x - h)]\, dh \quad \text{for } x \in \mathbb{R}^d.$$

Hence

$$\begin{aligned}\|f - \varphi_r * f\|_p &\le \int_{\mathbb{R}^d} |\varphi_r(h)| \|f(\cdot) - f(\cdot - h)\|_p\, dh \\ &\le \int_{\mathbb{R}^d} |\varphi_r(h)|\omega_p(f, |h|)\, dh = I_r.\end{aligned}$$

To estimate I_r we need the inequality

$$\omega_p(f, |h|) \le (r|h| + 1)\omega_p(f, r^{-1}), \quad r > 0,\ h \in \mathbb{R}^d. \tag{$*$}$$

Since $\omega_p(f, t)$ is increasing in t, $(*)$ is obvious for $|h| \le r^{-1}$. Let $|h| > r^{-1}$. Let the integer k satisfy $|h|r - 1 < k \le |h|r$. Then

$$\begin{aligned}&f(x + h) - f(x) \\ &= \sum_{j=1}^{k} \left[f\left(x + \frac{j}{|h|r}h\right) - f\left(x + \frac{j-1}{|h|r}h\right)\right] + f(x + h) - f\left(x + \frac{k}{|h|r}h\right).\end{aligned}$$

Thus, by the triangle inequality,

$$\begin{aligned}&\|f(\cdot + h) - f(\cdot)\|_p \\ &\le \sum_{j=1}^{k} \left\|f\left(\cdot + \frac{j}{|h|r}h\right) - f\left(\cdot + \frac{j-1}{|h|r}h\right)\right\|_p + \left\|f(\cdot + h) - f\left(\cdot + \frac{k}{|h|r}h\right)\right\|_p.\end{aligned}$$

Since no term on the right hand side exceeds $\omega_p(f, r^{-1})$, we get

$$\|f(\cdot + h) - f(\cdot)\|_p \le (k + 1)\omega_p(f, r^{-1}).$$

Thus, taking into account that $k \le r|h|$, we get $(*)$. Clearly $(*)$ yields

$$\begin{aligned}I_r &\le \omega_p(f, r^{-1}) \int_{\mathbb{R}^d} |\varphi_r(h)|(r|h| + 1)\, dh \\ &= \omega_p(f, r^{-1}) \int_{\mathbb{R}^d} (|u| + 1)|\varphi(u)|\, du \qquad \text{(substituting } u = hr) \\ &= C\omega_p(f, r^{-1}) \qquad \text{(by (j))}.\end{aligned}$$

This proves (3.6).

Clearly (3.6) yields

$$\begin{aligned}\overline{B}^\theta_{p,q}(f) &\le C \int_1^\infty r^\theta \omega_p(f, r^{-1})\, \frac{dr}{r} \le C \int_0^\infty r^\theta \omega_p(f, r^{-1})\, \frac{dr}{r} \\ &= CB^\theta_{p,q}(f) \quad \text{(substituting } r = t^{-1}).\end{aligned}$$

Thus $\|f\|_{\overline{B}^\theta_{p,q}} \le C\|f\|_{B^\theta_{p,q}}$.

To prove the reverse estimate put

$$\Delta(g, h) = g(\cdot + h) - g(\cdot) \quad (g \in L^p(\mathbb{R}^d, E),\ h \in \mathbb{R}^d).$$

Next recall an analogue of the classical Bernstein inequality

(BI) *$\exists C = C(d)$ such that if $g \in L^p(\mathbb{R}^d, E) \cap C^{(1)}(\mathbb{R}^d, E)$ and $\operatorname{supp} \widehat{g} \subset \{|\xi| \le r\}$ for some $r > 0$ then*

$$\|\nabla g\|_p \le Cr\|g\|_p.$$

Thus, for every $h \in \mathbb{R}^d$,

$$\|\Delta(g, h)\|_p = \Big\| \int_0^1 \nabla g(\cdot + th)(h)\, dt \Big\|_p \le Cr|h|\|g\|_p.$$

The proof of (BI) mimics the argument in [Pee], p. 51, proof of formula (13). Again we employ Lemma 3.1 in full generality.

Next fix $f \in L^p(\mathbb{R}^d, E)$ with $\overline{B}^\theta_{p,q}(f) < \infty$. For $m = 0, 1, \ldots$ pick $g_m \in L^p(\mathbb{R}^d, E) \cap C^{(1)}(\mathbb{R}^d, E)$ so that $\operatorname{supp} \widehat{g}_m \subset \{|\xi| \le 2^m\}$ and $\|f - g_m\|_p \le 2E_p(f, 2^m)$.

Put $w_0 = g_0$, $w_m = g_m - g_{m-1}$ for $m = 1, 2, \ldots$ Then $f = \sum_{m=0}^\infty w_m$ (the convergence of the series $\sum_{m=0}^\infty w_m$ is a routine consequence of the finiteness of the integral $\overline{B}^\theta_{p,q}(f)$ and the fact that the function $t \to E_p(f,t)$ for $t > 0$ is decreasing). Our choice of the w_m's implies $\|w_0\|_p \le \|f\|_p$ and for $m = 1, 2, \ldots,$

$$\|w_m\|_p \le \|f - g_m\|_p + \|f - g_{m-1}\|_p \le 4E_p(f, 2^{m-1}).$$

Thus for $h \in \mathbb{R}^d$ with $|h| \le 2^{-n}$ for some $n = 0, 1, \ldots,$ we have

$$\|\Delta(w_m, h)\|_p \le 2\|w_m\|_p \le 4E_p(f, 2^{m-1}) \quad \text{for } m > n$$

and, by (BI),

$$\begin{aligned}\|\Delta(w_m, h)\|_p &\le C|h|2^m\|w_m\|_p \le 4C2^{m-n}E_p(f, 2^{m-1}) \quad \text{for } m \le n,\\ \|\Delta(w_0, h)\|_p &\le C|h|\|w_0\|_p \le C2^{-n}\|f\|_p.\end{aligned}$$

Thus, for $n = 0, 1, \ldots,$

$$\begin{aligned}\omega_p(f, 2^{-n}) &\le \sum_{m=0}^\infty \sup_{|h| \le 2^{-n}} \|\Delta(w_m, h)\|_p\\ &\le b(0,n)\|f\|_p + \sum_{m=1}^\infty b(m,n)E_p(f, 2^{m-1})\end{aligned}$$

where

$$b(m,n) = \begin{cases} C_1 2^{m-n} & \text{for } m \le n,\\ C_1 & \text{for } m > n,\end{cases}$$

and $C_1 = C_1(d)$ is an absolute constant.

Let

$$I = \Big(\int_0^1 [t^{-\theta}\omega_p(f,t)]^q \frac{dt}{t}\Big)^{1/q}.$$

The crucial part of the proof is to estimate I from above. We have

$$\begin{aligned}I &= \Big(\sum_{n=0}^\infty \int_{2^{-n-1}}^{2^{-n}} [t^{-\theta}\omega_p(f,t)]^q \frac{dt}{t}\Big)^{1/q}\\ &\le 2^\theta\Big(\sum_{n=0}^\infty [2^{n\theta}\omega_p(f, 2^{-n})]^q\Big)^{1/q} \le 2^\theta \sum_{n=0}^\infty \lambda_n 2^{n\theta}\omega_p(f, 2^{-n}),\end{aligned}$$

where $(\lambda_n)_{n=0}^\infty$ is an appropriate positive sequence with $\sum_{n=0}^\infty \lambda_n^{q^*} = 1$ where $(q^*)^{-1} + q^{-1} = 1$.

Hence

$$\begin{aligned}2^{-\theta} I &\le \sum_{n=0}^\infty \lambda_n 2^{n\theta}\Big[b(0,n)\|f\|_p + \sum_{m=1}^\infty b(m,n)E_p(f, 2^{m-1})\Big]\\ &\le \|f\|_p \sum_{n=0}^\infty \lambda_n a(0,n) + \sum_{m=1}^\infty 2^{m\theta}E_p(f, 2^{m-1}) \sum_{n=0}^\infty \lambda_n a(m,n)\end{aligned}$$

where

$$a(m,n) = \begin{cases} C_2 2^{(m-n)(1-\theta)} & \text{for } m \le n,\\ C_2 2^{\theta(n-m)} & \text{for } m > n.\end{cases}$$

Let A be the operator determined by the matrix $(a_{m,n})_{m,n=0,1,\ldots}$ A direct calculation shows that A is bounded as an operator from l^∞ into l^∞ and from l^1 into l^1. Hence it is bounded from l^{q^*} into l^{q^*}. Let $C^* = \|A : l^{q^*} \to l^{q^*}\|$. Then by the Hölder inequality we get

$$\begin{aligned}I &\le 2^\theta C^*\Big(\|f\|_p^q + \sum_{m=1}^\infty [2^{m\theta}E_p(f, 2^{m-1})]^q\Big)^{1/q}\\ &\le 2^\theta C^*\Big[\|f\|_p + \Big(\sum_{m=1}^\infty [2^{m\theta}E_p(f, 2^{m-1})]^q\Big)^{1/q}\Big].\end{aligned}$$

Since $E_p(f,t)$ is decreasing in t, we easily obtain, for some constant C_2 independent of f,

$$\Big(\sum_{m=1}^\infty [2^{m\theta}E_p(f, 2^{m-1})]^q\Big)^{1/q} \le C_2\overline{B}^\theta_{p,q}(f).$$

Thus $I \le 2^\theta C^*\|f\|_p + C^*C_2\overline{B}^\theta_{p,q}(f)$. Also

$$\begin{aligned}\Big(\int_1^\infty t^{-\theta q}[\omega_p(f,t)]^q \frac{dt}{t}\Big)^{1/q} &\le 2\|f\|_p\Big(\int_1^\infty t^{-\theta q-1}\,dt\Big)^{1/q}\\ &= 2(\theta q)^{-1/q}\|f\|_p.\end{aligned}$$

Hence there exists a positive $C_3 = C_3(p,q,\theta)$ (independent of f!) such that

$$\|f\|_{B^\theta_{p,q}} \le C_3\|f\|_{\overline{B}^\theta_{p,q}}.$$

This completes the proof of the equivalence of the norms (3.4) and (3.5). ∎

Remarks. 1. Proposition 3.1 is well known in the scalar case (cf., e.g. [Tr], Theorems 2.3.2 and 2.5.4). However, some proofs presented in the literature do not carry over to the vector-valued case. In the vector-valued case Proposition 3.1 is a folklore. We give the proof for the sake of completeness.

2. Only cosmetic changes are needed to adapt our proof of Proposition 3.1 to the case of vector-valued Besov spaces on tori.

Some embedding theorems. We begin with the Gagliardo–Nirenberg embedding theorem (cf. [G] and [N]) in the vector-valued setting.

THEOREM 3.1. *Let E be a Banach space and let $d = 1, 2, \ldots$ Then*

$$L^1_{(1)}(\mathbb{R}^d, E) \hookrightarrow L^{d/(d-1)}(\mathbb{R}^d, E).$$

Moreover,

$$\|f\|_{d/(d-1)} \le \beta(d)\|\nabla f\|_1 \quad \text{for } f \in L^1_{(1)}(\mathbb{R}^d, E). \tag{3.7}$$

Proof. Assume that the assertion of the theorem has been established in the real case. Then one has

$$\begin{aligned}\|f\|_{d/(d-1)} = \|\|f(\cdot)\|_E\|_{d/(d-1)} &\le \beta(d)\|\nabla\|f(\cdot)\|_E\|_1 \quad &&\text{(by the assumption)}\\ &\le \beta(d)\|\nabla f\|_1 &&\text{(by Theorem 1.1)}.\end{aligned}$$

Now let $f \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$. By Proposition 2.1, given $\varepsilon > 0$ there exists a sequence of molecules (g_m) satisfying (2.1)–(2.3) with $a = 1$ and $b = \beta(d)$. For each molecule we have

$$\begin{aligned}\|g_m\|_{d/(d-1)} &\le \|g_m\|_\infty^{1/d}\|g_m\|_1^{(d-1)/d} \quad &&\text{(by the Hölder inequality)}\\ &\le \beta(d)(1+\varepsilon)\|\nabla g_m\|_1 &&\text{(by (2.3))}.\end{aligned}$$

Hence using the completeness of $L^{d/(d-1)}(\mathbb{R}^d, \mathbb{R})$, (2.1) and (2.2) we infer that $f \in L^{d/(d-1)}(\mathbb{R}^d, \mathbb{R})$ and

$$\begin{aligned}\|f\|_{d/(d-1)} &\le (1+\varepsilon)\sum_m \|g_m\|_{d/(d-1)}\\ &\le (1+\varepsilon)^2\beta(d)\sum_m \|\nabla g_m\|_1 \le (1+\varepsilon)^3\beta(d)\|\nabla f\|_1.\end{aligned}$$

Letting $\varepsilon \to 0$ we get (3.7). ∎

The next result is more subtle.

THEOREM 3.2. *Let E be a Banach space and let $d = 1, 2, \ldots$ Then*

$$L^1_{(1)}(\mathbb{R}^d, E) \hookrightarrow L_{d/(d-1),1}(\mathbb{R}^d, E).$$

Moreover,

$$\|f\|_{d/(d-1),1} \le \beta(d)\|\nabla f\|_1 \quad \text{for } f \in L^1_{(1)}(\mathbb{R}^d, E). \tag{3.8}$$

Proof. Repeat the proof of Theorem 3.1 with only one change: Use (3.1) instead of the Hölder inequality. ∎

In the scalar case, Theorem 3.2 has been proved by Poornima [Po] with another constant.

Remarks. 1. Theorems 3.1 and 3.2 fit the following abstract scheme:

PROPOSITION 3.2. *Let E be a Banach space. Let $X(\mathbb{R}^d, E)$ be a Banach space of λ_d-equivalence classes of measurable E-valued functions on $\mathbb{R}^d$.*

Assume that $X(\mathbb{R}^d, E)$ is equipped with a norm $\|\cdot\|_{X(\mathbb{R}^d,E)}$ such that for some $C > 0$,

$$\|f\|_{X(\mathbb{R}^d,E)} \le C\|f\|_1^{d/(d-1)}\|f\|_\infty^{1/d} \quad \text{for } f \in X(\mathbb{R}^d, E).$$

Then $L^1_{(1)}(\mathbb{R}^d, E) \hookrightarrow X(\mathbb{R}^d, E)$ and

$$\|f\|_{X(\mathbb{R}^d,E)} \le 3C\beta(d)\|\nabla f\|_1 \quad \text{for } f \in L^1_{(1)}(\mathbb{R}^d, E).$$

Moreover, if there exists a Banach space $X(\mathbb{R}^d, \mathbb{R})$ of λ_d-equivalence classes of scalar-valued functions on $\mathbb{R}^d$ such that for every $f \in X(\mathbb{R}^d, E)$ the function $\|f(\cdot)\|_E \in X(\mathbb{R}^d, \mathbb{R})$ and $\|\|f(\cdot)\|_E\|_{X(\mathbb{R}^d,\mathbb{R})} = \|f\|_{X(\mathbb{R}^d,E)}$ then

$$\|f\|_{X(\mathbb{R}^d,E)} \le C\beta(d)\|\nabla f\|_1 \quad \text{for } f \in L^1_{(1)}(\mathbb{R}^d, E).$$

The proof of the first part of the proposition uses Theorem 2.1 in its full strength; the proof of the "moreover" part just repeats the argument of the proof of Theorem 3.1.

2. It has been discovered by Federer and Fleming [FF] and Maz'ya [M1] (cf. [M2], Chapt. 2, §2.3, p. 103) that in the scalar case $\beta(d)$ is the best constant in the inequality (3.7) (by evaluating both sides of (3.7) at smooth approximations of the characteristic function of the Euclidean unit ball of $\mathbb{R}^d$). Hence $\beta(d)$ is the best constant in (3.7) for every Banach space E.

The same is true for (3.8).

3. Aubin [A] and Talenti [T] have found the best constants in the Sobolev inequality

$$\|f\|_{dp/(d-p)} \le \beta(p,d)\|\nabla f\|_p \quad \text{for } f \in L^p_{(1)}(\mathbb{R}^d, \mathbb{R})\ (1 \le p < d). \tag{3.9}$$

The argument given in the second part of the proof of Theorem 3.1 shows that (3.9) extends to vector-valued functions with the same best constants.

Now we pass to embeddings of $L^1_{(1)}(\mathbb{R}^d, E)$ into Besov spaces. The next theorem extends to Banach space-valued functions the result of Kolyada [K] for scalar-valued functions. Our proof is different from that of [K].

THEOREM 3.3. *Let $d = 1, 2, \ldots$ and let E be a Banach space. Then*

$$L^1_{(1)}(\mathbb{R}^d, E) \hookrightarrow B^{\theta(p,d)}_{p,1}(\mathbb{R}^d, E),$$

where $\theta(p,d) = d(1/p + 1/d - 1)$ and $1 < p < d/(d-1)$.

Moreover, there exists $c = c(p)$ such that

$$B^{\theta(p,d)}_{p,1}(f) \le 3c\|\nabla f\|_1 \quad \text{for } f \in L^1_{(1)}(\mathbb{R}^d, E).$$

We begin with two observations:

LEMMA 3.4. *One has*

$$\omega_p(f,t) \le (\omega_1(f,t))^{1/p}(2\|f\|_\infty)^{(p-1)/p}.$$

Proof. Integrate against λ_d over $\mathbb{R}^d$ both sides of the inequality

$$\|f(x+h)-f(x)\|_E^p \le \|f(x+h)-f(x)\|_E (2\|f\|_\infty)^{p-1}. \ \blacksquare$$

LEMMA 3.5. *If $f \in \mathcal{D}(\mathbb{R}^d, E)$ then*

$$\omega_1(f,t) \le t\|\nabla f\|_1 \quad \text{for } t>0.$$

Proof. The identity

$$f(x+h)-f(x) = |h| \int_0^1 \nabla f(x+th)\left(\frac{h}{|h|}\right) dt$$

combined with the Cauchy–Schwarz inequality and the Fubini theorem yields

$$\begin{aligned}\int_{\mathbb{R}^d} \|f(x+h)-f(x)\|_E\,dx &= |h| \int_{\mathbb{R}^d} \left\| \int_0^1 \nabla f(x+th)\left(\frac{h}{|h|}\right) dt \right\|_E dx \\ &\le |h| \int_{\mathbb{R}^d} \int_0^1 \|\nabla f(x+th)\|\,dt\,dx \\ &= |h| \int_0^1 \int_{\mathbb{R}^d} \|\nabla f(x+th)\|\,dx\,dt \\ &= |h| \int_0^1 \|\nabla f(\cdot+th)\|_1\,dt = |h| \int_0^1 \|\nabla f\|_1\,dt \\ &= |h|\|\nabla f\|_1. \ \blacksquare\end{aligned}$$

Proof of Theorem 3.3. The crucial point is to establish the appropriate inequality for molecules. Precisely, we shall show that there exists $c = c(p,d) > 0$ such that if $f \in \mathcal{D}(\mathbb{R}^d, E)$ satisfies $\|f\|_1^{(d-1)/d}\|f\|_\infty^{1/d} \le (\beta(d)+\varepsilon)\|\nabla f\|_1$ for some $\varepsilon > 0$ then

$$(3.10) \qquad \int_0^\infty t^{-\theta(p,d)}\omega_p(f,t)\,\frac{dt}{t} \le (c+\varrho(\varepsilon))\|\nabla f\|_1,$$

where $\varrho(\varepsilon) \to 0$ as $\varepsilon \to 0$.

For $f \ne 0$ we break the integral $\int_0^\infty$ into $\int_0^\alpha$ and $\int_\alpha^\infty$, where

$$\alpha = (\|f\|_1/\|f\|_\infty)^{1/d}.$$

Using Lemmas 3.4 and 3.5 and putting

$$C(p,d) = \frac{2^{1-1/p}}{(d-1)(1-1/p)}$$

we have

$$\begin{aligned}\int_0^\alpha t^{d-2-d/p}\omega_p(f,t)\,dt &\le \int_0^\alpha t^{d-2-d/p+1/p}(2\|f\|_\infty)^{(p-1)/p}\|\nabla f\|_1^{1/p}\,dt \\ &= C(p,d)\alpha^{(d-1)(1-1/p)}\|f\|_\infty^{(p-1)/p}\|\nabla f\|_1^{1/p} \\ &= C(p,d)(\|f\|_1^{(d-1)/d}\|f\|_\infty^{1/d})^{1-1/p}\|\nabla f\|_1^{1/p} \\ &\le C(p,d)\beta(d)^{(p-1)/p}(1+\varrho(\varepsilon))\|\nabla f\|_1.\end{aligned}$$

To estimate the second integral we need the trivial inequality $\omega_p(f,t) \le 2\|f\|_p$ and the inequality

$$\|f\|_p \le \beta(d)^{d(1-1/p)}\|f\|_1^{1-d(1-1/p)}\|\nabla f\|_1^{d(1-1/p)}.$$

The latter follows immediately from the Hölder inequality

$$\|f\|_p \le \|f\|_1^{1-d(1-1/p)}\|f\|_{d/(d-1)}^{d(1-1/p)}$$

and (3.7). Now putting $C'(p,d) = 2b^{-1}$ where

$$b = 1 + d/p - d$$

we get

$$\begin{aligned}\int_\alpha^\infty t^{d-2-d/p}\omega_p(f,t)\,dt &\le 2\beta(d)^{1-b}\|\nabla f\|_1^{1-b}\|f\|_1^b \int_\alpha^\infty t^{d-2-d/p}\,dt \\ &= C'(p,d)\beta(d)^{1-b}\|\nabla f\|_1^{1-b}\|f\|_1^b\alpha^{-b} \\ &= C'(p,d)\beta(d)^{1-b}\|\nabla f\|_1^{1-b}\|f\|_1^b(\|f\|_1/\|f\|_\infty)^{-b/d} \\ &= C'(p,d)\beta(d)^{1-b}\|\nabla f\|_1^{1-b}[\|f\|_1^{1-1/d}\|f\|_\infty^{1/d}]^b \\ &\le C'(p,d)\beta(d)^{1-b}\|\nabla f\|_1^{1-b}(\beta(d)+\varepsilon)^b\|\nabla f\|_1^b \\ &= C'(p,d)\beta(d)(1+\varrho(\varepsilon))\|\nabla f\|_1.\end{aligned}$$

Thus letting $\varepsilon \to 0$ we get (3.10) with $c = C(p,d)\beta(d)^{p/(p-1)} + C'(p,d)\beta(d)$.

Now using Theorem 2.1 together with Remark 3 after the proof of Theorem 2.1, and letting $\varepsilon \to 0$ we get $B_{p,1}^{\theta(p,d)}(f) \le 3c\|\nabla f\|_1$ for all $f \in L^1_{(1)}(\mathbb{R}^d, E)$, which proves the "moreover" part of the theorem.

Hence, by (3.7), for all $f \in L^1_{(1)}(\mathbb{R}^d, E)$,

$$\begin{aligned}\|f\|_{B_{p,1}^{\theta(p,d)}} &\le 3c\|\nabla f\|_1 + \|f\|_p \le 3c\|\nabla f\|_1 + \|f\|_{d/(d-1)}^{d(1-1/p)}\|f\|_1^{1-d(1-1/p)} \\ &\le 3c\|\nabla f\|_1 + \beta(d)^{d(1-1/p)}\|\nabla f\|_1^{d(1-1/p)}\|f\|_1^{1-d(1-1/p)} \\ &\le A(d,p)(\|\nabla f\|_1 + \|f\|_1)\end{aligned}$$

where $A(d,p)$ is a positive constant independent of f. $\blacksquare$

Theorem 3.3 extends to the limit case $p = d/(d-1)$. One has

COROLLARY 3.1. *For every Banach space E and for $d = 1, 2, \ldots$ one has*
$$L^1_{(1)}(\mathbb{R}^d, E) \hookrightarrow B^0_{d/(d-1),1}(\mathbb{R}^d, E).$$
Moreover, for every $\Psi \in \Psi(d)$ there exists $C_\Psi > 0$ such that
$$B^0_{d/(d-1),1}(f;\Psi) = \sum_{k=0}^{\infty} \|f * \widehat{\psi}_k\|_{d/(d-1)} \le C_\Psi \|\nabla f\|_1 \quad \text{for } f \in L^1_{(1)}(\mathbb{R}^d, E).$$

Proof. Combine Theorem 3.3 with Lemma 3.3. ∎

As a consequence of Corollary 3.1 and the Hausdorff-Young inequality ([Hö], 7.1.13) we get Bourgain's analogue of the Hardy inequality for analytic functions.

THEOREM 3.4. *Let $d = 2, 3, \ldots$ Then there exists a $C > 0$ such that*

$$\int_{\mathbb{R}^d} |\widehat{f}(\xi)|(1+|\xi|)^{1-d}\, d\xi \le C\|\nabla f\|_1 \quad \text{for } f \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R}). \tag{3.11}$$

Proof. Fix $\Psi \in \Psi(d)$ and $f \in L^1_{(1)}(\mathbb{R}^d, \mathbb{R})$. Taking into account that the Fourier Transform of the product of functions equals the convolution of their Fourier Transforms, the inversion formula, and the Hausdorff-Young inequality (note that $d \ge 2$) we infer that

$$\|\psi_k \widehat{f}\|_d \le C(d)\|\widehat{\psi}_k * f\|_{d/(d-1)} \quad (k = 0, 1, \ldots), \tag{3.12}$$

where the constant $C(d) > 0$ is independent of f and Ψ.

Put for convenience $\psi_{-1} = 0$ and consider the partition of $\mathbb{R}^d$ given by
$$A_0 = \{|\xi| < 1\}, \quad A_k = \{2^{k-1} \le |\xi| < 2^k\} \quad (k = 1, 2, \ldots).$$

Note that

$$\int_{A_m} (1+|\xi|)^{-d}\, d\xi \le 2^d \lambda_d(B_d) \quad (m, k = 0, 1, \ldots). \tag{3.13}$$

It follows from (i) that

$$(\operatorname{supp}\psi_k) \cap A_m \ne \emptyset \Rightarrow |m-k| \le 1 \ (m = 0, 1, \ldots). \tag{3.14}$$

Thus, by (iii), for $m = 0, 1, \ldots,$ we have
$$\int_{A_m} |\widehat{f}(\xi)|(1+|\xi|)^{1-d}\, d\xi$$
$$= \int_{A_m} \sum_{k=0}^{\infty} \psi_k(\xi)|\widehat{f}(\xi)|(1+|\xi|)^{1-d}\, d\xi$$

$$= \sum_{k=m-1}^{m+1} \int_{A_m} \psi_k(\xi)|\widehat{f}(\xi)|(1+|\xi|)^{1-d}\, d\xi \qquad \text{(by (3.14))}$$
$$\le \sum_{k=m-1}^{m+1} \int_{A_m} |\psi_k(\xi)\widehat{f}(\xi)|(1+|\xi|)^{1-d}\, d\xi$$
$$\le \sum_{k=m-1}^{m+1} \Big(\int_{A_m} |\psi_k(\xi)\widehat{f}(\xi)|^d\, d\xi\Big)^{1/d} \Big(\int_{A_m} (1+|\xi|)^{-d}\, d\xi\Big)^{(d-1)/d} \qquad \text{(by Hölder)}$$
$$\le 2^{d-1}\lambda_d(B_d)^{(d-1)/d} \sum_{k=m-1}^{m+1} \|\psi_k \widehat{f}\|_d \qquad \text{(by (3.13))}$$
$$\le 2^{d-1}\lambda_d(B_d)^{(d-1)/d} C(d) \sum_{k=m-1}^{m+1} \|\widehat{\psi}_k * f\|_{d/(d-1)} \qquad \text{(by (3.12))}.$$

Thus summing over m, and applying Corollary 3.1 we get
$$\int_{\mathbb{R}^d} |\widehat{f}(\xi)|(1+|\xi|)^{1-d}\, d\xi = \sum_{m=0}^{\infty} \int_{A_m} |\widehat{f}(\xi)|(1+|\xi|)^{1-d}\, d\xi$$
$$\le \frac{3}{2} C(d) 2^d \lambda_d(B_d)^{(d-1)/d} \sum_{k=0}^{\infty} \|\widehat{\psi}_k * f\|_{d/(d-1)}$$
$$\le \frac{3}{2} C(d) 2^d \lambda_d(B_d)^{(d-1)/d} C_\Psi (\|\nabla f\|_1 + \|f\|_1).$$

Thus we get (3.11) with
$$C = 3C(d)2^{d-1}\lambda_d(B_d)^{(d-1)/d} \inf_{\Psi \in \Psi(d)} C_\Psi. \ ∎$$

Remark. In fact we have shown that the function $(1+|\xi|)^{1-d}$ determines a functional on the Fourier Transforms of $B^0_{d/(d-1),1}(\mathbb{R}^d, \mathbb{C})$ for $d \ge 2$, while Poornima used the fact that $(1+|\xi|)^{1-d}$ determines a functional on the Fourier Transforms of $L_{d/(d-1),1}(\mathbb{R}^d, \mathbb{C})$ for $d \ge 3$.

COROLLARY 3.2. *Let $d = 2, 3, \ldots$ and $n = 1, 2, \ldots$ If $f \in L^1_{(n)}(\mathbb{R}^d, \mathbb{C})$ then*

$$\int_{\mathbb{R}^d} |\widehat{f}(\xi)|(1+|\xi|)^{n-d}\, d\xi < \infty. \tag{3.15}$$

Proof. Fix $d \ge 2$, $n \ge 2$ and $f \in L^1_{(n)}(\mathbb{R}^d, \mathbb{C})$. Note that

$$\frac{\partial^{(n-1)}}{\partial x(j)^{(n-1)}} f \in L^1_{(1)}(\mathbb{R}^d, \mathbb{C}), \quad j = 1, \ldots, d. \tag{3.16}$$

Thus, by Theorem 3.1,

$$\infty > \int_{\mathbb{R}^d} \left| \left[\frac{\partial^{(n-1)}}{\partial x(j)^{(n-1)}} f \right]^\wedge (\xi) \right| (1+|\xi|)^{1-d}\, d\xi$$
$$= \int_{\mathbb{R}^d} |\widehat{f}(\xi)|\, |\xi(j)|^{n-1} (1+|\xi|)^{1-d}\, d\xi .$$

Since $\sum_{j=1}^n |\xi(j)|^{n-1} \geq a(d,n-1)|\xi|^{n-1}$ where $a(d,n-1)=1$ for $n \leq 3$ and $a(d,n-1) = d^{(3-n)/2}$ for $n > 3$, we get

$$\int_{\mathbb{R}^d} |\xi|^{n-1} |\widehat{f}(\xi)| (1+|\xi|)^{1-d}\, d\xi < \infty .$$

Clearly $f \in L^1_{(1)}(\mathbb{R}^d, \mathbb{C})$, hence by Theorem 3.4,

$$\int_{\mathbb{R}^d} |\widehat{f}(\xi)| (1+|\xi|)^{1-d}\, d\xi < \infty .$$

Since $1 + |\xi|^{n-1} \geq 2^{2-n}(1+|\xi|)^{n-1}$, we get (3.15). ∎

The periodic analogue of Corollary 3.2 was discovered by Bourgain [Br1] in 1981. In [Po] our Corollary 3.2 was proved for $d \geq 3$.

We do not know whether (3.15) is already a consequence of the assumption that f and all its distributional pure partial derivatives of order $\leq n$ belong to $L^1(\mathbb{R}^d, \mathbb{C})$. In particular, is it true that f, $\widetilde{D}_{xx} f$, $\widetilde{D}_{yy} f$ in $L^1(\mathbb{R}^2, \mathbb{C})$ implies $\widehat{f} \in L^1(\mathbb{R}^2, \mathbb{C})$?

Another consequence of Corollary 3.1 in the case of scalar-valued functions is

COROLLARY 3.3. *Let $d = 2, 3, \ldots$ There exists a constant $C = C(d)$ such that if $f \in L^1_{(1)}(\mathbb{R}^d, \mathbb{C})$ then*

$$\sum_{k=0}^{\infty} \Big(\int_{D_k} |\widehat{f}(\xi)|^d \, d\xi \Big)^{1/d} < C(\|\nabla f\|_1 + \|f\|_1),$$

where

$$D_0 = \{\xi = (\xi(j)) \in \mathbb{R}^d : \max_j |\xi(j)| \leq 1\},$$
$$D_k = \{\xi \in \mathbb{R}^d : 2^{k-1} < \max_j |\xi(j)| \leq 2^k\} \quad (k = 1, 2, \ldots).$$

Proof. Combine Corollary 3.1 with the Hausdorff–Young inequality and the fact that the characteristic functions of the frames D_k's are bounded (uniformly in k) multipliers on $L^p(\mathbb{R}^d, \mathbb{C})$ for $1 < p < \infty$ (cf. e.g. [St], Chapt. IV, §4). ∎

Remarks. 1. The periodic analogue of Corollary 3.3 was discovered by Bourgain [Br2] around 1985.

2. Theorem 3.4 and Corollaries 3.2 and 3.3 carry over with the same proofs to the case of Hilbert space-valued functions.

3. It seems to be unknown for what Banach spaces E Theorem 3.4 remains valid for $L^1_{(1)}(\mathbb{R}^d, E)$. It still holds for E being isomorphic to a subspace of L^1 while it fails for c_0.

Embedding theorems in the periodic case. Theorems 3.1–3.3 and Corollary 3.1 extend to the case of spaces of Banach space-valued functions on tori. There are very little changes in the proofs. We have to use Theorem 2.1_π instead of Theorem 2.1. $\|\nabla f\|_1$ is replaced by $\|\nabla f\|_1 + \|f\|_1$. In Theorems 3.1 and 3.2 we loose control of the best constant. The Fourier Transform of a function on $\mathbb{T}^d$ is a function (a sequence) on the lattice $\mathbb{Z}^d$ of vectors with integer coordinates. Thus the counterpart of Theorem 3.4 and Corollary 3.2 is the following result due to Bourgain [Br1].

COROLLARY 3.2_π. *If $f \in L^1_{(1)}(\mathbb{T}^d, \mathbb{C})$ then*

$$\sum_{m \in \mathbb{Z}^d} |\widehat{f}(m)| (1+|m|)^{n-d} < \infty \quad (n = 1, 2, \ldots, \ d = 2, 3, \ldots).$$

The proofs of Corollary 3.2_π and of the periodic analogue of Corollary 3.3 require only cosmetic modifications of the proofs of their $\mathbb{R}^d$ counterparts.

References

[A] T. Aubin, *Problèmes isopérimétriques et espaces de Sobolev*, C. R. Acad. Sci. Paris 280 (1975), 279–281.

[BS] C. Bennett and R. Sharpley, *Interpolation of Operators*, Academic Press, London, 1988.

[Br1] J. Bourgain, *A Hardy inequality in Sobolev spaces*, Vrije University, Brussels, 1981.

[Br2] —, *Some examples of multipliers in Sobolev spaces*, IHES, 1985.

[BZ] Yu. D. Burago and V. A. Zalgaller, *Geometric Inequalities*, Nauka, Leningrad, 1980 (in Russian); English transl.: Springer, 1988.

[CSV] T. Coulhon, L. Saloff-Coste and N. Varopoulos, *Analysis and Geometry on Groups*, Cambridge University Press, Cambridge, 1992.

[DS] N. Dunford and J. T. Schwartz, *Linear Operators I*, Interscience, New York, 1958.

[F] H. Federer, *Geometric Measure Theory*, Springer, Berlin, 1969.

[FF] H. Federer and W. H. Fleming, *Normal and integral currents*, Ann. of Math. 72 (1960), 458–520.

[G] E. Gagliardo, *Proprietà di alcune classi di funzioni in più variabili*, Ricerche Mat. 7 (1958), 102–137.

[Hö] L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, Springer, Berlin, 1983.

[H] R. Hunt, *On $L(p,q)$ spaces*, Enseign. Math. (2) 12 (1966), 249–275.

[Jo] P. W. Jones, *Quasiconformal mappings and extendability of functions in Sobolev spaces*, Acta Math. 147 (1981), 71–88.

[K] V. I. Kolyada, *On relations between moduli of continuity in different metrics*, Trudy Mat. Inst. Steklov. 181 (1988), 117–136 (in Russian); English transl.: Proc. Steklov Inst. Math. 4 (1989), 127–148.

[Kr] A. S. Kronrod, *On functions of two variables*, Uspekhi Mat. Nauk 5 (1) (1950), 24–134 (in Russian).

[Le] M. Ledoux, *Semigroup proofs of the isoperimetric inequality in euclidean and Gauss space*, Bull. Sci. Math., to appear.

[LW] L. H. Loomis and H. Whitney, *An inequality related to the isoperimetric inequality*, Bull. Amer. Math. Soc. 55 (1949), 961–962.

[M1] V. G. Maz'ya, *Classes of sets and embedding theorems for function spaces*, Dokl. Akad. Nauk SSSR 133 (1960), 527–530 (in Russian).

[M2] —, *S. L. Sobolev's Spaces*, Leningrad University Publishing House, Leningrad, 1985 (in Russian).

[N] L. Nirenberg, *On elliptic partial differential equations*, Ann. Scuola Norm. Sup. Pisa (3) 13 (1959), 116–162.

[Pee] J. Peetre, *New Thoughts on Besov Spaces*, Duke Univ. Math. Ser. 1, Durham, N.C., 1976.

[Po] S. Poornima, *An embedding theorem for the Sobolev space $W^{1,1}$*, Bull. Sci. Math. (2) 107 (1983), 253–259.

[St] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, N.J., 1970.

[T] G. Talenti, *Best constant in Sobolev inequality*, Ann. Mat. Pura Appl. (4) 110 (1976), 353–372.

[Tr] H. Triebel, *Theory of Function Spaces*, Birkhäuser, Basel 1983.

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
ŚNIADECKICH 8
00-950 WARSZAWA, POLAND
E-mail: OLEK@IMPAN.IMPAN.GOV.PL
MIWOJ@IMPAN.IMPAN.GOV.PL

Received February 24, 1993 (3066)

INFORMATION FOR AUTHORS

Manuscripts should be typed on one side only, with double or triple spacing and wide margins, and submitted in duplicate, including the original typewritten copy. Poor quality copies will not be accepted.

An **abstract** of not more than 200 words and the AMS Mathematics Subject Classification are required.

Formulas must be typewritten. A complete **list of all handwritten symbols** with indications for the printer should be enclosed.

Figures should be drawn accurately on separate sheets, preferably twice the size in which they are required to appear. The author should indicate in the margin of the manuscript where figures are to be inserted.

References should be arranged in alphabetical order, typed with double spacing, and styled and punctuated according to the examples given below. Abbreviations of journal names should follow Mathematical Reviews. Titles of papers in Russian should be translated into English.

Examples:

[6] D. Beck, *Introduction to Dynamical Systems*, Vol. 2, Progr. Math. 54, Birkhäuser, Basel, 1978.

[7] R. Hill and A. James, *An index formula*, J. Differential Equations 15 (1982), 197–211.

[8] J. Kowalski, *Some remarks on $J(X)$*, in: Algebra and Analysis, Proc. Conf. Edmonton 1973, E. Brook (ed.), Lecture Notes in Math. 867, Springer, Berlin, 1974, 115–124.

[Nov] A. S. Novikov, *An existence theorem for planar graphs*, preprint, Moscow University, 1980 (in Russian).

Authors' **affiliation** should be given at the end of the manuscript.

Authors receive only page **proofs** (one copy). If the proofs are not returned promptly, the article will be printed in a later issue.

Authors receive 50 **reprints** of their articles. Additional reprints can be ordered.

The publisher would like to encourage submission of manuscripts written in TeX. **On acceptance of the paper**, authors should send discs (preferably PC) plus relevant details to the Editorial Committee, or transmit the paper by electronic mail to:

STUDIA@IMPAN.IMPAN.GOV.PL