Mixed Norms and Rearrangements: Sobolev's Inequality and Littlewood's Inequality (*) (**).

JOHN J. F. FOURNIER

Summary. – *In the 1930's, J. E. Littlewood and S. L. Sobolev each found useful estimates for L^p-norms. These results are usually not regarded as similar, because one of them is set in a discrete context and the other in a continuous setting. We show, however, that certain basic facts about mixed norms can be used to simplify proofs of both of these estimates. The same method yields a proof of a form of the isoperimetric inequality. We consider the effect of measure-preserving rearrangement on certain sums of permuted mixed norms of functions on R^K, and show these sums are minimal when the rearranged function, $f^\sim$ say, has the property that, for each positive real number λ, the set on which $|f^\sim| > \lambda$ is a cube with edges parallel to the coordinate axes. Finally, we use the fact about rearrangements to prove sharper forms of the estimates of Littlewood and Sobolev.*

1. – Introduction.

Given an integrable function f on R^K and an integer k between 1 and K, let $N_k(f)$ be the extended-real number obtained by first computing the essential supremum of $|f(x_1, x_2, \ldots, x_K)|$ with respect to the k-th variable, and then integrating this quantity with respect to the remaining variables. Let $N(f) = \sum_{k=1}^{K} N_k(f)$. Also, given f, let $f^\sim$ be a measure-preserving rearrangement of $|f|$ with the property that for each number $\lambda > 0$ the set where $f^\sim > \lambda$ is a K-cube with edges parallel to the coordinate axes. The central result in this paper is that $N(f^\sim) \leqslant N(f)$.

The stimulus for this paper was the recent proof given by S. POORNIMA [21] that, when $K > 1$ and $r = K/(K-1)$, the Sobolev space $W^{1,1}(R^K)$ imbeds into the Lorentz space $L(r, 1)$; the latter space is strictly smaller than L^r, which is the target space in the usual statement [1] of the imbedding theorem for $W^{1,1}(R^K)$. The first proof of the Sobolev imbedding theorem [24] did not apply to the case of $W^{1,1}$, but later, E. GAGLIARDO [13] and L. NIRENBERG [19] found a method of proof which worked in that exceptional case. Their idea was to observe that if $f \in W^{1,1}$, then the quantities $N_k(f)$ are finite for all k, and to deduce from this that $f \in L^r$. It is natural

(*) Entrata in Redazione il 27 novembre 1985; versione riveduta il 14 marzo 1986.
(**) Research partially supported by N.S.E.R.C. operating grant number 4822.
Indirizzo dell'A.: Department of Mathematics, University of British Columbia, Vancouver, Canada, V6T 1Y4.

to ask if these mixed-norm conditions, namely that $N_k(f) < \infty$ for all k, together imply that $f \in L(r, 1)$. We show here that this implication does indeed hold; we prove it in two ways, first by using our theorem about the effect of measure-preserving rearrangement of the quantity $N(f) = \sum_{k=1}^{K} N_k(f)$, and second by combining the method of Gagliardo and Nirenberg with Poornima's approach.

There is a striking similarity between the use of mixed norms in the context described above and the use of such norms in the study of bounded multilinear forms on l^∞. In [16] J. E. LITTLEWOOD showed that if a matrix $(a_{m,n})$ defines a bounded bilinear form on l^∞, with norm $\|a\|$ say, then

$$\sum_n \left\{\sum_m |a_{m,n}|^2\right\}^{1/2} \leqslant \varkappa \|a\|, \text{ and} \tag{1.1}$$

$$\sum_m \left\{\sum_n |a_{m,n}|^2\right\}^{1/2} \leqslant \varkappa \|a\|; \tag{1.2}$$

Littlewood deduced from these mixed-norm estimates that $\|a\|_{4/3} \leqslant \varkappa' \|a\|$. In Section 2, we show how the latter implication and the Gagliardo-Nirenberg proof of the Sobolev imbedding theorem can both be based on the same elementary properties of mixed norms. In Section 3, we use this approach to prove a form of the isoperimetric inequality due to L. H. LOOMIS and H. WHITNEY [17], and we present our analysis of the effect of measure-preserving rearranement on $N(f)$. In Section 4, we consider applications of our result on rearrangements. The embedding $W^{1,1} \subset L(r, 1)$ follows easily from this result. By using duality twice, we show that if a matrix $(a_{m,n})$ satisfies inequalities (1.1) and (1.2), then the matrix belongs to the Lorentz space $l(4/3, 1)$; this improves on Littlewood's conclusion that $a \in l^{4/3}$ in this case. We apply this improvement to the known [11] examples of 4/3-Sidon sets. Finally, we present alternate proofs of two of our main results in an appendix to the paper.

The idea of using measure-preserving rearrangements to prove imbedding theorems for Sobolev spaces goes back to the first proof [24], by S. L. Sobolev, of such imbeddings. More recently, several authors [4, 10, 18, 29] have used properties of measure-preserving rearrangements that are constant on *spheres* to obtain the best constants in various cases of Sobolev's inequality. This approach was used by W. G. FARIS [12] to prove a result that is equivalent by duality to the inclusion $W^{1,1} \subset L(r, 1)$. To get the dual formulation of this inclusion, first note that the standard imbedding of $W^{1,1}$ into L^r is equivalent by duality to the statement that

(i) if $f \in W^{1,1}$ then $\int |f \cdot g| < \infty$ for all functions g in the space $L^{r'}$,

where r' is the index conjugate to r. Let weak-$L^{r'}$ be the space of all measurable functions g on R^K for which there is a constant c so that, for each number $\lambda > 0$, the inequality $|g| > \lambda$ holds on a set of measure at most $c/\lambda^{r'}$; then $L^{r'}$ is *strictly* included in weak-$L^{r'}$. The inclusion $W^{1,1} \subset L(r, 1)$ is equivalent by duality to the statement that

(i′) if $f \in W^{1,1}$ then $\int |f \cdot g| < \infty$ for all functions g in the space weak-$L^{r'}$.

Assertion (i′) follows easily from the main result in [12].

The papers by FARIS and by POORNIMA and the present paper provide a variety of methods for proving the imbedding $W^{1,1} \subset L(r, 1)$. We will comment further on these methods in Section 4. We note here that our result that if $N(f) < \infty$ then $f \in L(r, 1)$ is equivalent by duality to an inclusion of the space weak-$L^{r'}$ in a sum of certain mixed-norm spaces. RON BLEI and the author have found a direct proof of the dual inclusion, and will present that proof in a joint paper. The direct proof of the dual inclusion provides an alternate approach to the applications in Section 4 of the present paper.

2. – Rearrangement of indices.

The functionals N_k described in the introduction are examples of what we will call *permuted mixed-norms*. Our goal in this section is to discuss some basic properties of these norms, and to show how certain important estimates follow easily from these properties.

Let X be a cartesian product of sigma-finite measure spaces, X_k say. Denote a typical element of X by $\boldsymbol{x} = (x_1, x_2, \ldots, x_k)$, and the product measure on X by $d\boldsymbol{x}$. Given a measurable function f on X, and an index vector $\boldsymbol{p} = (p_1, p_2, \ldots, p_K)$ with $0 < p_k \leqslant \infty$ for all k, consider the quantity obtained by first computing the norm in $L^{p_1}(X_1)$ of the function $x_1 \to f(x_1, x_2, \ldots, x_K)$ for each value of the tail vector $(x_2, .., x_K)$, and denoting this partial norm by $\|f(\cdot, x_2, \ldots, x_K)\|_{p_1}$, then computing the norm in $L^{p_2}(X_2)$ of the function $x_2 \to \|f(\cdot, x_2, \ldots, x_K)\|_{p_1}$, Denote the ultimate result of this computation by $\|f\|_{\boldsymbol{p}}$; for instance, in this notation, the quantity $N_1(f)$ is just $\|f\|_{\boldsymbol{p}}$, where $\boldsymbol{p} = (\infty, 1, 1, \ldots, 1)$. Denote the set of all functions f for which $\|f\|_{\boldsymbol{p}} < \infty$ by $L_{\boldsymbol{p}}(X)$.

The standard reference concerning these *mixed-norm spaces* $L^{\boldsymbol{p}}(X)$ is [5]. We need two elementary facts about mixed norms; the first fact is a version of Hölder's inequality, and the second is an easy consequence of the integral form of Minkowski's inequality. We will indicate how these facts are proved, partly because of a lack of accessible references for them in the forms that we need, and partly to convince the reader that these facts really are elementary. Given index sequences $\boldsymbol{p}$, $\boldsymbol{q}$ and $\boldsymbol{r}$, we write $1/\boldsymbol{r} = 1/\boldsymbol{p} + 1/\boldsymbol{q}$ if $1/r_k = 1/p_k + 1/q_k$ for all k.

THEOREM 2.1. – Let $f \in L^{\boldsymbol{p}}(X)$ and $g \in L^{\boldsymbol{q}}(X)$, and let $1/\boldsymbol{r} = 1/\boldsymbol{p} + 1/\boldsymbol{q}$. Then $f \cdot g \in L^{\boldsymbol{r}}(X)$, and

$$\|f \cdot g\|_{\boldsymbol{r}} \leqslant \|f\|_{\boldsymbol{p}} \cdot \|g\|_{\boldsymbol{q}} \,. \tag{2.1}$$

We will discuss the proof of this version of Hölder's inequality after we state our version of the integral form of Minkowski's inequality. To prepare for that statement we note that when $k > 1$ the functional N_k is, strictly speaking, not a mixed norm, because N_k must be computed by dealing first with the variable x_k, whereas the definition of mixed norm specifies that x_1 must come first. These functionals are examples, however, of *permuted* mixed norms. Any permutation, σ say, of the set $\{1, 2, \ldots, K\}$ induces an adjoint action on variables $\boldsymbol{x}$ and indices $\boldsymbol{p}$. Thus, we define $\sigma(\boldsymbol{x})$ to be the element $(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(K)})$ of the Cartesian product space $\sigma(X) = X_{\sigma(1)} \times X_{\sigma(2)} \times \ldots \times X_{\sigma(K)}$, and we define $\sigma(\boldsymbol{p})$ similarly. It would be more precise to denote these adjoint actions by σ^* or the like, but this distinction does not matter here. On the other hand, it is important in what we do to distinguish between product spaces like $X_1 \times X_2$ and $X_2 \times X_1$, even when the factors X_1 and X_2 are copies of the same space, R for instance. Given a function f on X, we continue our abuse of notation by letting σf be the function on $\sigma(X)$ given by $\sigma f(\sigma(\boldsymbol{x})) = f(\boldsymbol{x})$ for all elements $\sigma(\boldsymbol{x})$ of the space $\sigma(X)$; equivalently, $\sigma f(\boldsymbol{y}) = f(\sigma^{-1}(\boldsymbol{y}))$ for all $\boldsymbol{y}$ in $\sigma(X)$. For instance, the quantity $N_K(f)$ is, in this notation, $\|\sigma f\|_{\sigma(\boldsymbol{p})}$, where $p_K = \infty$ and $p_j = 1$ for all $j \neq K$, and σ is any permutation with the property that $\sigma(1) = K$; indeed, we get $N_K(f)$ by first taking the L^∞-norm with respect to the variable x_K, and then taking the L^1-norms with the respect to the other variables in any order. Our version of Minkowski's inequality for integrals concerns the effects, on mixed norms, of permutations that act on the variables and the indices *in the same way*.

THEOREM 2.2. – Fix a measurable function f and an index vector $\boldsymbol{p}$. Among the various quantities $\|\sigma f\|_{\sigma(\boldsymbol{p})}$, the one for which the sequence $\{p_{\sigma(k)}\}_{k=1}^K$ is nondecreasing is the smallest, and the one for which the sequence $\{p_{\sigma(k)}\}_{k=1}^K$ is nonincreasing is the largest.

We now indicate how these two theorems may be proved. Consider first the one-variable case of inequality (2.1), that is the estimate

$$\|f \cdot g\|_r \leq \|f\|_p \cdot \|g\|_q , \tag{2.2}$$

where f and g are measurable functions on the same measure space, and $1/r = 1/p + 1/q$. If r is infinite, then so are p and q, and the inequality above is obvious. If r is finite, then the indices p/r and q/r are conjugate, and inequality (2.2) follows from an application of Hölder's inequality, for these conjugate indices, to the integral of $|f|^r$. Finally, the multivariate case of inequality (2.1) follows by interating the single-variable case.

We will also apply Theorem 2.1 to products of K functions, where $K > 2$; if the k-th factor belongs to the mixed-norm space with index $\boldsymbol{p}^{(k)}$, then iterating the theorem yields that the product function belongs to the space $L^{\boldsymbol{r}}(X)$, where

$$\frac{1}{\boldsymbol{r}} = \sum_{k=1}^{K} \frac{1}{\boldsymbol{p}^{(k)}} ,$$

and that the appropriate product estimate holds for the L^r-norm of the product function. The case where each factor is a power of the same function is presented in [28]. Note that there is no requirement in the statement of the theorem or in its proof that the indices p_k, q_k, and r_k be at least 1; in fact, it will be convenient in the applications of the theorem to allow indices in the interval $(0, 1)$.

We begin our analysis of Theorem 2.2 by considering the case where there are only two variables. In this context, we drop the use of subscripts on variables and indices, denoting the product measure space by $X \times Y$, a typical element in it by (x, y), and a typical index pair by (p, q). We also suppose for definiteness that the indices are finite; our conclusions are also valid, with similar proofs, when one or more index is infinite. Theorem 2.2 addresses the question: Is

$$\left(\int_Y \left\{\left[\int_X |f(x, y)|^p \, dx\right]^{1/p}\right\}^q dy\right)^{1/q} \leqslant \left(\int_X \left\{\left[\int_Y |f(x, y)|^q \, dy\right]^{1/q}\right\}^p dx\right)^{1/p}, \tag{2.3}$$

or is the reverse inequality true? The theorem asserts that the larger of the two expressions above is the one for which the larger index is associated with the first variable to be integrated; that is, inequality (2.3) holds if $q \geqslant p$, while the reverse inequality holds if $p \geqslant q$. To see why this is so, we specialize further to the case where $(p, q) = (1, r)$ with $r \geqslant 1$; the assertion that

$$\left(\int_Y \left\{\int_X |f(x, y)| dx\right\}^r dy\right)^{1/r} \leqslant \int_X \left[\int_Y |f(x, y)|^r \, dx\right]^{1/r} dy \tag{2.4}$$

is known as Minkowski's inequality for integrals. To get a completely familiar inequality, we consider the case where Y has only two points, y and z say, each with mass 1. Let g and h be the functions in $L^r(X)$ given by $g(x) = f(x, y)$ and $h(x) = f(x, z)$ for all x; then inequality (2.4) states that

$$\| |g| + |h| \|_r \leqslant \|g\|_r + \|h\|_r , \tag{2.5}$$

which is Minkowski's inequality for sums. The corresponding inequality for integrals can be proved in the same way as the familiar inequality for sums [14, #202]; alternatively, the general case of inequality (2.4) can be deduced from the special case (2.5) by reagarding the integrals over Y as limits of sums. The fact that inequality (2.3) holds when $p \leqslant q$ follows by applying inequality (2.4) with $r = q/p$ and f replaced by $|f|^p$.

Now we return to the general setting for Theorem 2.2, and to the notation with subscripts on the names of variables, measure spaces, etc. Let τ be a permutation that transposes adjacent indices, k and $k+1$ say, and fixes all other indices. We say that the transposition τ *raises* the index sequence $\boldsymbol{p}$ if $p_{k+1} \geqslant p_k$. In the computation of the quantities $\|f\|_{\boldsymbol{p}}$ and $\|\tau f\|_{\tau(\boldsymbol{p})}$ the variables x_k and x_{k+1} come together, the only difference being that x_k and the corresponding index p_k come first for $\|f\|_{\boldsymbol{p}}$, whereas x_{k+1} and p_{k+1} come first for $\|\pi f\|_{\tau(\boldsymbol{p})}$. So, our analysis of inequaltity (2.3) yields that

$$\|f\|_{\boldsymbol{p}} \leqslant \|\tau f\|_{\tau(\boldsymbol{p})} \quad \text{if } \tau \text{ raises } \boldsymbol{p}\,. \tag{2.6}$$

If second transposition τ', of adjacent indices, raises $\tau(\boldsymbol{p})$, then $\|f\|_{\tau(\boldsymbol{p})} \leqslant \|\tau'(\tau f)\|_{\tau'(\tau(\boldsymbol{p}))}$ and so on. Now let σ be any permutation of the set $\{1, 2, \ldots, K\}$, and let σ' be a permutation for which $p_{\sigma'(1)} \geqslant p_{\sigma'(2)} \geqslant \ldots \geqslant p_{\sigma'(K)}$. Then we can pass from $\sigma(\boldsymbol{p})$ to $\sigma'(\boldsymbol{p})$ by a sequence of transpositions of adjacent indices so that each of these transpositions raises the index sequence to which it is applied. Specifically, we find a maximal index p_k, and then successively transpose this index with each of its neighbours to the left until it sits in the first position; then we find an index $p_{k'}$ that is maximal among the indices other than the one now in position 1, and we move $p_{k'}$ to left until it sits in position 2, etc. By iterating inequality (2.6) we get that

$$\|\sigma f\|_{\sigma(\boldsymbol{p})} \leqslant \|\sigma' f\|_{\sigma'(\boldsymbol{p})}\,,$$

as asserted in the theorem. Similarly, the smallest of the quantities $\|\sigma f\|_{\sigma(\boldsymbol{p})}$ occurs when $p_{\sigma(1)} \leqslant p_{\sigma(2)} \leqslant \ldots \leqslant p_{\sigma(K)}$. This completes our discussion of the proofs of Theorems 2.1 and 2.2.

Our first application of these theorems is to Littlewood's inequality. Suppose that a doubly-infinite matrix $(a_{m,n})$ has the property that for all elements b and c of $l^\infty(Z)$ the iterated series $\sum_{n=-\infty}^{\infty} \left\{ \sum_{m=-\infty}^{\infty} a_{m,n} b_n \right\} c_n$ converges to a number, $a(b, c)$ say, with

$$|a(b, c)| \leqslant C\|b\|_\infty \|c\|_\infty\,,$$

where C is some constant determined by a; define the norm $\|a\|$ of this bilinear form to be the infimum of the constants C for which the inequality above holds for all b and c in $l^\infty(Z)$. As noted in the introduction, Littlewood proved the mixed-norm inequalities (1.1) and (1.2). In the notation of this section, these inequalities become the estimates

$$\|a\|_{\boldsymbol{p}} \leqslant \varkappa \|a\|\,, \quad \text{and} \quad \|\tau a\|_{\tau(\boldsymbol{q})} \leqslant \varkappa \|a\|\,, \tag{2.7}$$

where $\boldsymbol{p} = (2, 1)$, while $\boldsymbol{q} = (1, 2)$, and τ is the transposition of the underlying two-element index set. Since $\tau(\boldsymbol{q}) = (2, 1)$, the permutation τ arranges the indices in $\boldsymbol{q}$ in decreasing order, and Theorem 2.2 yields that $\|a\|_{\boldsymbol{q}} \leqslant \|\tau a\|_{\tau(\boldsymbol{q})}$; a fortiori, $\|a\|_{\boldsymbol{q}} \leqslant \varkappa \|a\|$. With these estimates for $\|a\|_{\boldsymbol{p}}$ and $\|a\|_{\boldsymbol{q}}$ in hand, we apply Theorem 2.1 to deduce that $a^2 = l^{\boldsymbol{r}}$, where $1/\boldsymbol{r} = 1/\boldsymbol{p} + 1/\boldsymbol{q}$; moreover,

$$\|a^2\|_{\boldsymbol{r}} \leqslant \|a\|_{\boldsymbol{p}} \|a\|_{\boldsymbol{q}} \leqslant (\varkappa \|a\|)^2\,. \tag{2.8}$$

But $1/\mathbf{r} = (1/2, 1/1) + (1/1, 1/2) = (3/2, 3/2)$, so that the mixed-norm space $l^{\mathbf{r}}$ is isometric with $l^{2/3}$. Thus, $\|a^2\|_{2/3} \leqslant (\varkappa\|a\|)^2$, or equivalently, $\|a\|_{4/3} \leqslant \varkappa\|a\|$, which is the desired estimate.

Littlewood's proof yielded the estimate $\|a\|_{4/3} \leqslant 2\varkappa\|a\|$. The fact that our method gives a better constant is not important in this context, but in work of BLEI [8] on multilinear and fractional versions of Littlewood's inequality, it was desirable to get values for the constants in the inequalities that grew at an optimal rate as the dimension increased, and this required the use of a method like the one used here. For instance, it was shown, in [8, Lemma 5.3] that if a function a on Z^K defines a bounded K-linear form on $l^\infty(Z)$, then $\|a\|_{2K/(K+1)} \leqslant \varkappa^{K-1}\|a\|$, where the constant $\varkappa$ is independent of K. To prove this inequality, one first proves that $\|a\|_{\mathbf{p}} \leqslant \varkappa^{K-1}\|a\|$, where $\mathbf{p} = (2, 2, \ldots, 2, 1)$, and that the same estimate holds for $\|\sigma a\|_{\mathbf{p}}$ for all cyclic permutations σ. The desired estimate for $\|a\|_{2K/(2K+1)}$ then follows by our method, or the one used in [8].

Next, we use Theorems 2.1 and 2.2 to prove an endpoint case of the Sobolev imbedding theorem. We recall that the Sobolev space $W^{1,p}(R^K)$ consists of all functions in $L^1(R^K)$ whose first-order distributional partial derivatives also belong to $L^p(R^K)$; this space is complete with respect to the norm $\|\cdot\|_{1,p}$ given by

$$\|f\|_{1,p} = \left\{(\|f\|_p)^p + \sum_{k=1}^{K}(\|\nabla_k f\|_p)^p\right\}^{p/1},$$

where $\nabla_k f$ denotes the first partial derivative of f with respect to the k-th variable. See [1] or [25] for more information about Sobolev spaces. The imbedding theorem states that if $K > 1$ and if $1 \leqslant p < K$, then $W^{1,1}(R^K) \subset L^{p^*}(R^K)$, where $p^* = Kp/(K-p)$. Sobolev [24] proved this imbedding for the cases where $1 < p < K$, but his method did not work when $p = 1$. That case was settled affirmatively by GAGLIARDO [13] and NIRENBERG [19]; they also showed how the other cases follow easily from the one where $p = 1$. To deal with this fundamental case, first use the fact the set $C_c^1(R^K)$, of all compactly-supported C^1-functions on R^K, is dense in $W^{1,1}(R^K)$ to reduce matters to proving the estimate

$$\|f\|_{K/(K-1)} \leqslant \varkappa\|f\|_{1,1} \tag{2.10}$$

for all such functions f. Next, note that if g is a compactly-supported C^1-function on R, then

$$|g(t)| \leqslant \min\left\{\left|\int_{-\infty}^{t} g'(u)\,du\right|, \left|\int^{\infty} g'(u)\,du\right|\right\} \quad \text{for all } t;$$

hence, $\|g\|_\infty \leqslant (\|g'\|_1)/2$. Let $f \in C_c^1(R^K)$; then, by our estimate for $\|g\|_\infty$,

$$\sup_{x_1} |f(x_1, x_2, \ldots, x_K)| \leqslant \tfrac{1}{2}\int_R |\nabla_1 f(x_1, x_2, \ldots, x_K)|\,dx_1$$

for all fixed values of the tail vector $(x_2, \ldots, x_K)$. Integrating this inequality with respect to the variables in the tail vector yields that $N_1(f) \leqslant (\|\nabla_1 f\|_1)/2$. When $k > 1$, the same procedure starting with the k-th variable rather than the first one yields that $N_k(f) \leqslant (\|\nabla_k f\|_1)/2$.

As in the case of Littlewood's inequality, we now use Theorem 2.2 to get mixed-norm estimates from the estimates above for the permuted mixed norms $N_k(f)$. Let $\boldsymbol{p}^{(k)}$ be the index sequence whose k-th entry is ∞ and whose other entries are all 1's; let σ_k be any permutation of the set $\{1, 2, \ldots, K\}$ with the property that $\sigma_k(1) = k$. Then $N_k(f)$ is equal to the norm of the functions $\sigma_k f$ in the mixed-norm space with index $\sigma_k(\boldsymbol{p}^{(k)})$ on the measure space $\sigma_k(X)$. Since the indices in the sequence $\sigma_k(\boldsymbol{p}^{(k)})$ are arranged in decreasing order, Theorem 2.2 yields that $\|f\|_{\boldsymbol{p}^{(k)}} \leqslant N_k(f)$, and hence that $\|f\|_{\boldsymbol{p}^{(k)}} \leqslant (\|\nabla_k f\|_1)/2$ for all k. Let

$$\frac{1}{\boldsymbol{r}} = \sum_{k=1}^{K} \frac{1}{\boldsymbol{p}^{(k)}}.$$

It follows from Theorem 2.1 that $f^K \in L^{\boldsymbol{r}}(R^K)$ and that

$$\|f^K\|_{\boldsymbol{r}} \leqslant \prod_{k=1}^{K} [(\|\nabla_k f\|_1)/2]. \tag{2.11}$$

Each component of the index vector $1/\boldsymbol{r}$ is a sum of $K-1$ copies of $1/1$ and one copy of $1/\infty$, and is therefore equal to $K-1$. So, a more conventional way to write inequality (2.11) is to replace the quantity on the left by $\|f^K\|_{1/(K-1)}$. The resulting inequality is equivalent to the estimate

$$\|f\|_{K/(K-1)} \leqslant \tfrac{1}{2} \prod_{k=1}^{K} (\|\nabla_k f\|_1)^{1/K}. \tag{2.12}$$

This inequality implies the desired estimate (2.10), so that the proof of the fundamental case of the imbedding theorem is complete.

The reader may wish to compare the presentation here with the arguments in [13], [19], [9], [8] and [22]. These arguments are inductions on the dimension K, with the conventional form of Hölder's inequality and the integral form of Minkowski's inequality used in each step of the induction. In our method, Minkowski's inequality is used repeatedly at the beginning, to get from estimates for permuted mixed norms to estimates for mixed norms; then Theorem 2.1, which is just an iterated form of Hölder's inequality, is used to finish the proof. One advantage of this arrangement of the proof is that it becomes clear that full use was not made of the initial estimates on the quantities $N_k(f)$. First, we replaced these quantities by the quantities $\|f\|_{\boldsymbol{p}^{(k)}}$, which may be much smaller when $k > 1$. Then we factored f^K as a product of K copies of the same function and used Theorem 2.2, which also applies to products of distinct functions. This suggests that it may be possible to derive stronger conclusions from the initial estimates, and that is what we do in the rest of this paper.

3. – Measure-preserving rearrangements.

Our goal in this section is to prove the assertion, made in the introduction, about rearrangements and the sum of the permuted mixed norms N_k. We defer the discussion of the applications of this result to Section 4. In our proof of the rearrangement theorem, we use a variant of the isoperimetric inequality, which we prove by the method of the previous section. In the appendix to this paper, we will outline a second proof of the rearrangement theorem.

Let f be a measurable function on a sigma-finite measure space X; denote the measure of a given set S in X by $|S|$. For each number $\lambda > 0$, let $m_f(\lambda)$ be the extended-real number given by

$$m_f(\lambda) = \left|\{x \in X: |f(x)| > \lambda\}\right| , \tag{3.1}$$

and call the function m_f *the distribution function of* f. Call the function f *rearrangeable* if $m_f(\lambda) < \infty$ for all $\lambda > 0$. Given a rearrangeable function f, call a measurable function g, on some measure space that may differ from the space X, a *measure-preserving rearrangement of* f if the distribution functions m_f and m_g coincide.

See [26] for more about rearrangements of functions. All of the rearrangements that we will consider in this section will arise from measure-preserving transformations of the underlying measure space, which will always be R^K for some positive integer K. Let φ be a one-to-one, measurable map of R^K onto itself with the property that $|\varphi(S)| = |S|$ for all measurable subsets S if R^K. Then for each rearrangeable function f on R^K the function $\boldsymbol{x} \to f(\varphi(\boldsymbol{x}))$ is a measure-preserving rearrangement of f. In the next section, we will encounter pairs of functions, denoted there by a and A, for which $m_a = m_A$, although the functions a and A are not related by composition with any measure-preserving isomorphism between the underlying measure spaces, because these measure spaces do not have the same cardinality, and because the sets where a and A are equal to 0 do not have the same measure.

The term « rearrangement » will always mean a measure-preserving rearrangement. Replacing a given function by any such rearrangement of it does not change its L^p-norm; indeed [26, § V.3], there is a formula for computing $\|f\|_p$ from m_f. On the other hand the mixed norms of a function can be changed by passing to a rearrangement of the function. For example, let f be the indicator function of the rectangle $[0, 2) \times [0, 1/2)$ in R^2, and let g be the indicator function of the square $[0, 1) \times [0, 1)$; then $m_f = m_g$. Let N_1 and N_2 be the functionals defined in the abstract. Clearly, $N_1(f) = 1/2$, and $N_2(f) = 2$, while $N_1(g) = N_2(g) = 1$. Recall that the functional N is defined to the sum of the various functionals N_k, and note that $N(g) < N(f)$ in this example.

THEOREM 3.1. – Let K be an integer that is greater than or equal to 2. Let f be a function on R^K with the property that $N(f) < \infty$. Then f is rearrangeable. Let g

be a measure-preserving rearrangement of f with the property that for each number $\lambda > 0$ the set where $|g| > \lambda$ is essentially a K-cube with edges parallel to the coordinate axes. Then $N(g) \leqslant N(f)$.

PROOF. – We say that a set is *essentially a K-cube* if it differs from some K-cube by a set of measure 0. The hypothesis that $N(f) < \infty$ implies, as in the previous section that $f \in L^{K/(K-1)}(R^K)$; hence f is rearrangeable. The latter conclusion can also be proved directly. Since

$$N(f) = \sup\{N(F) \colon F \text{ is simple with bounded support, and } 0 \leqslant F \leqslant |f|\},$$

we may assume in the rest of the proof that f is a nonnegative simple function with bounded support.

Denote the nonzero values of f, in *decreasing* order, by $a_1, a_2, \ldots, a_M$, and let A_M be the set where $f \geqslant a_m$. Let

$$F_1(x_2, x_3, \ldots, x_K) = \operatorname*{ess\,sup}_{x_1} f(x_1, x_2, \ldots, x_K),$$

and let A_m^1 be the subset of R^{K-1} where $F_1 \geqslant a_m$. Similarly define functions F_k and sets A_m^k for integers k with $2 \leqslant k \leqslant K$. Denote the measure in R^K of A_m by $|A_m|$, and the measure in R^{K-1} of A_m^k by $|A_m^k|$. Then

$$N_k(f) = \sum_{m=1}^{M} a_m\big(|A_m^K| - |A_{m-1}^k|\big), \tag{3.3}$$

where $|A_0^k| = 0$ by convention. Summing this formula by parts for each index k and adding, we get that

$$N(f) = \sum_{m=1}^{M} (a_m - a_{m+1})\left\{\sum_{k=1}^{K} |A_m^k|\right\}, \tag{3.4}$$

where $a_{M+1} = 0$ by convention.

If we replace f by any rearrangement of it, the measures $|A_m^k|$ may change, but the numbers a_m and $|A_m|$ will not change. Moreover, by Lemma 3.2 below,

$$\prod_{k=1}^{K} |A_m^k| \geqslant |A_m|^{K-1}. \tag{3.5}$$

In the m-th term in formula (3.4), the factor $(a_m - a_{m+1})$ is positive and fixed, while

$$\sum_{k=1}^{K} |A_m^k| \geqslant K|A_m|^{(K-1)/K},$$

by the inequality between arithmetic and geometric means, and the constraint

(3.5). So,

$$N(f) \geqslant \sum_{m=1}^{M} (a_m - a_{m+1}) \cdot K |A_m|^{(K-1)/K} . \tag{3.6}$$

Let g be any rearrangement of f for which the sets where $|g| \geqslant a_m$ are essentially K-cubes with edges parallel to the coordinate axes. Then formula (3.4) yields that $N(g)$ is equal to the right hand side of formula (3.6). This completes the proof of the theorem, modulo the lemma below.

Given a measurable set A in R^K, denote its indicator function by 1_A, and define the *essential projection* of A into the K-th coordinate hyperplane to be the subset A_K of R^{K-1} with indicator function given by

$$1_{A_K}(x_1, x_2, \ldots, x_{K-1}) = \operatorname*{ess\,sup}_{x_K} 1_A(x_1, x_2, \ldots, x_K) .$$

Define the essential projection A_k of A into the k-th coordinate hyperplane in a similar way. Again denote the measure, in R^K, of A by $|A|$, and the measure, in R^{K-1}, of A_k by $|A_k|$. Observe that if A is a solid box with edges parallel to the coordinate axes, then $\prod_{k=1}^{K} |A_k| = |A|^{K-1}$.

The following lemma goes back at least as far as [17] and has been rediscovered at least three times, in [3], [8, inequality (2.4)], and [23]. I am grateful to Amram Meir and Ron Blei for bringing the reference [17] to my attention. The lemma is easy to prove by the methods of the previous section.

LEMMA 3.2. – For any measurable set A in R^K, the measures of A and of its essential projections must satisfy the condition that

$$|A|^{K-1} \leqslant \prod_{k=1}^{K} |A_k| . \tag{3.7}$$

PROOF. – Assume without loss of generality that $|A_k| < \infty$ for all k. Observe that $N_k(1_A) = |A_k|$ for all k. As in our proof of the Sobolev imbedding theorem,

$$\|1_A\|_{K/(K-1)} \leqslant \prod_{k=1}^{K} N_k(1_A)^{1/K}.$$

Taking K-th powers in this inequality yields inequality (3.7), thereby completing the proof of the lemma.

The methods used in the proof of Theorem 3.1 given above and in the alternate proof given in the appendix can also be used to prove similar statements about functionals where suprema are first taken with respect to more than one variable. Given a measurable function f on R^K, and a subset α of the underlying index set $\{1, 2, \ldots, K\}$, let $N_\alpha(f)$ be the extended-real number obtained by first taking the essential supremum of $|f(x_1, x_2, \ldots, x_K)|$ with respect to the variables x_k as k runs through the set α, and then integrating this quantity with respect to the remaining variables. Denote the number of elements in the set α by $|\alpha|$. Given an integer n with $1 < n < K$, let

$$N^{(n)}(f) = \sum_{|\alpha|=n} N_\alpha(f) .$$

THEOREM 3.3. – If $N^{(n)}(f) < \infty$, then f is rearrangeable. Let g be a rearrangement of f for which, for each number $\beta > 0$, the set where $|g| > \lambda$ is essentially a K-cube with edges parallel to the coordinate axes. Then $N^{(n)}(g) \leqslant N^{(n)}(f)$.

As noted above, this statement can be proved by the methods used to prove Theorem 3.1. We omit the details.

4. – Applications.

Before proving that certain functions must belong to certain Lorentz spaces, we briefly recall some basic facts about these spaces. We refer to the books [6] and [26] for more details; our notation is a compromise between the notations used in these sources. As in the previous section, let f be a rearrangeable function on a sigma-finite measure space X carrying the measure dx. Then there is a unique function f^* on the interval $(0, \infty)$ with the following properties:

(i) $f^* \geqslant 0$.

(ii) f^* is nonincreasing.

(iii) f^* is right-continuous.

(iv) f^* has the same distribution function as f.

Given indices p and q with $1 < p < \infty$ and $1 \leqslant q \leqslant \infty$, say that f belongs to the *Lorentz space* $L(p, q)$ if the function $t \to t^{1/p} f^*(t)$ belongs to L^1 with respect to the measure dt/t. It is thus appropriate to consider the quantity

$$\|f\|_{L(p,q)} = \left\{\frac{q}{p}\int_0^\infty [t^{1/p} f^*(t)]^q \frac{dt}{t}\right\}^{1/q}, \tag{4.1}$$

when $q < \infty$, and the quantity $\|f\|_{L(p,\infty)} = \text{ess sup}\,\{t^{1/p} f^*(t) : t > 0\}$, because $f \in L(p, q)$- if and only if $\|f\|_{L(p,q)} < c$. The factor q/p is inserted in the definition of $\|\cdot\|_{L(p,q)}$ when $q < +$ to guarantee that $\|f\|_{L(p,q)} = \|f\|_p$ whenever f is the indicator function of a set. Note also that $\|f\|_{L(p,p)} = \|f\|_p$ for all f, so that $L(p, p) = L^p$. For fixed p, the spaces $L(p, q)$ become larger as q increases; this can be seen by verifying that $f \in L(p, q)$ if an only if the sequence $\{2^{n/p} f^*(2^n)\}_{n=-\infty}^{\infty}$ belongs to l^q. The inclusions $L(p, 1) \subset L^p \subset L(p, \infty)$ are strict, except in trivial cases.

THEOREM 4.1. – Given an integer K with $K \geqslant 2$, let $r = K/(K-1)$. Let f be a measurable function on R^K for which $N(f) < \infty$. Then $f \in L(r, 1)$, and $\|f\|_{L(r,1)} \leqslant N(f)/K$.

PROOF. – Adopt the l^∞-norm for vectors in R^K; that is, let $|\boldsymbol{x}| = \max_k \{|x_k|\}$ for all such vectors $\boldsymbol{x}$. Define a function $f^\sim$ by letting $f^\sim(\boldsymbol{x}) = f^*(|\boldsymbol{x}|^K)$ in the open first orthant in R^K, and extending $f^\sim$ to be 0 on the rest of R^K. Then

$$|\{\boldsymbol{x}: f^\sim(\boldsymbol{x}) > \lambda\}| = |\{t: f^*(t) > \beta\}| = |\{\boldsymbol{x}: f(\boldsymbol{x}) > \lambda\}|$$

for all numbers $\beta > 0$. It follows that $(f^\sim)^* = f^*$, and hence that $\|f^\sim\|_{L(r,1)} = \|f\|_{L(r,1)}$. Moreover, the sets where $f^\sim > \lambda$ are open K-cubes; so Theorem 3.1 applies to yield that $N(f^\sim) \leqslant N(f)$.

We will complete the proof of the present theorem with a computation showing that $N(f^\sim) = K\|f\|_{L(r,1)}$, from which it follows that $\|f\|_{L(r,1)} \leqslant N(f)/K$. Given a vector $\boldsymbol{y}$ in R^{K-1} and a real number x, denote the vector $(y_1, \ldots, y_{K-1}, x)$ by $(\boldsymbol{y}, x)$. For each vector $\boldsymbol{y}$ in R^{K-1}, let $g(\boldsymbol{y}) = \operatorname{ess\,sup}_{x \in R} f^\sim(\boldsymbol{y}, x)$; clearly, this function vanishes off the positive orthant in R^{K-1}. On the other hand, if $\boldsymbol{y}$ lies in the positive orthant, then

$$g(\boldsymbol{y}) = \operatorname*{ess\,sup}_{x>0} f^\sim(\boldsymbol{y}, x) = \operatorname*{ess\,sup}_{x>0} f^*(|(\boldsymbol{y}, x)|^K) = f^*(|\boldsymbol{y}|^K),$$

because f^* is right-continuous and nonincreasing, In particular, in the positive orthant, $g(\boldsymbol{y})$ depends only on $|\boldsymbol{y}|$, and we also denote $g(\boldsymbol{y})$ by $G(|\boldsymbol{y}|)$ in this case. Using polar coordinates adapted to the norm $|\cdot|$, we have that

$$\int_{R^{K-1}} g(\boldsymbol{y})\, d\boldsymbol{y} = (K-1)\int_0^\infty G(s)\, s^{K-2}\, ds .$$

Now the left side above is just $N_K(f^\sim)$, and the right side is equal to

$$(K-1)\int_0^\infty f^*(s^K)\, s^{K-2}\, ds = \frac{K-1}{K}\int_0^\infty f^*(t)\, t^{-1/K}\, dt ,$$

by the change of variable $t = s^K$. But the latter integral is equal to

$$\frac{K-1}{K}\int_0^\infty t^{(K-1)/K} f^*(t)\, \frac{dt}{t} = \|f\|_{L(r,1)} .$$

By the symmetry of $f^\sim$, we have that $N(f^\sim) = K\cdot N_K(f^\sim) = K\cdot\|f\|_{L(r,1)}$, is claimed.

COROLLARY 4.2. – Under the same hypotheses,

$$\|f\|_{L(r,1)} \leqslant \prod_{k=1}^{K} N_k(f)^{1/K}.$$

PROOF. – Let $a_k = N_k(f)$ for each index k. If $a_k = 0$ for some k, then $f = 0$ almost everywhere, and the inequality above is trivial; similarly, there is nothing to prove if $a_k = \infty$ for some k. In the remaining cases, define a new function F by dilating f by the factor a_k in the k-th coordinate direction for each index k, that is by letting

$$F(\boldsymbol{x}) = f\left(\frac{x_1}{a_1}, \ldots, \frac{x_K}{a_K}\right) \quad \text{for all } \boldsymbol{x}.$$

Then $N_k(F) = \prod_{j=1}^{K} a_j$ for all k. By the theorem,

$$\|F\|_{L(r,1)} \leqslant \frac{N(F)}{K} = \prod_{k=1}^{K} a_k.$$

But $\|F\|_{L(r,1)} = \left[\prod_{k=1}^{K} a_k\right]^{1/r} \|f\|_{L(r,1)}$. So,

$$\|f\|_{L(r,1)} \leqslant \left[\prod_{k=1}^{K} a_k\right]^{1/r'},$$

where r' is the index conjugate to r. Since $r = K/(K-1)$ here, $r' = K$, and the inequality above is just the assertion of the corollary.

THEOREM 4.3. – Let K be an integer greater than 1, and let $r = K/(K-1)$. Then $W^{1,1}(R^K) \subset L(r, 1)$, and

$$\|f\|_{L(r,1)} \leqslant \tfrac{1}{2} \prod_{k=1}^{K} (\|\nabla_k f\|_1)^{1/K} \tag{4.2}$$

for all functions f in $W^{1,1}(R^K)$.

PROOF. – We saw in Section 2 that $N_k(f) \leqslant (1/2)\|w_k f\|_1$ for all indices k, and all C^1-functions f with compact support. The corollary above then yields inequality (4.2) for all such functions f; since $C^1_c(R^K)$ is dense in $W^{1,1}(R^K)$, the inequality above holds for all f in $W^{1,1}(R^K)$, and the inclusion $W^{1,1}(R^K) \subset L(r, 1)$ must hold. This completes the proof of the theorem.

We pause to compare the various methods for proving that $W^{1,1}(R^K) \subset L(r, 1)$. This inclusion was essentially proved by FARIS [12], who considered the inequality

$$\int_{R^K} |f \cdot g| \leqslant C_K \|\nabla f\|_1 \|g\|_{L(r',\infty)}, \tag{4.3}$$

for all C_c^1-functions f and all measurable functions g. He used rearrangements of f for which all the sets $\{x: |f(x)| > \lambda\}$ are balls centre at 0 to prove this inequality with the best value for the constant c_K. Inequality (4.3) is equivalent by a duality argument to the inequality

$$\|f\|_{L(r,1)} \leqslant \frac{c_K}{r} \|\nabla f\|_1 ,$$

although this is not stated explicitly in [12]. Subsequently, POORNIMA [21] studied the effect of certain composition operators on the space $W^{1,1}(R^K)$, and showed that $W^{1,1}(R^K) \subset L^r(R^K)$. Her method works on any domain $\boldsymbol{R}$ for which the imbedding $W^{1,1}(\boldsymbol{R}) \subset L^r(\boldsymbol{R})$ holds. If the boundary of $\boldsymbol{R}$ is sufficiently regular, then this imbedding can be deduced, by an extension technique, from the one for $W^{1,1}(R^K)$, but [2] there are domains $\boldsymbol{R}$ for which $W^{1,1}(\boldsymbol{R}) \subset L^r(\boldsymbol{R})$ although there is no bounded extennsion operator mapping $W^{1,1}(\boldsymbol{R})$ into $W^{1,1}(R^K)$. All such domains known to the author have the property that each function, f say, in $W^{1,1}$ can be expressed as a convex combination of functions F for each of which $N(F)$ is bounded by $c\|f\|_{1,1}$ in a suitable coordinate system; the imbedding of $W^{1,1}(\boldsymbol{R})$ into $L(r, 1)$ then follows from Theorem 4.1. In the appendix to the present paper, we combine one of the ideas in [21] with Lemma 3.2 to give another proof of Theorem 4.1.

There is an interesting connection between the method used by Faris and a point that arises in the theory of Lorentz spaces. In most cases, the functional $\|\cdot\|_{L(p,q)}$ fails to be subadditive, and is therefore not a norm. This defect can be remedied by using the *averaged rearrangement* f^{**} given by

$$f^{**}(t) = \frac{1}{t}\int_0^t f^*(s)\, ds ,$$

for all $t > 0$. Let $\|f\|^*_{L(p,q)}$ be the quantity obtained by replacing the function f^* by f^{**} in formula (4.1); then the functional $\|\cdot\|^*_{L(p,q)}$ is a norm, because, for each fixed positive number t, the map $f \to f^{**}(t)$ defines a norm. Moreover,

$$\|f\|_{L(p,q)} \leqslant \|f\|^*_{L(p,q)} \leqslant c_{p,q}\|f\|_{L(p,q)} .$$

for all indices p and q satisfying the conditions specified at the beginning of this section. Given a smooth, rearrangeable function f on R^K, let F be the unique nonnegative function on R^K with the property that for each number $\lambda > 0$ the set $\{x: F(x) > \lambda\}$ is an open ball centred at 0 and having the same measure as the set $\{x: |f(x) > 0\}$; call F the *radial rearrangement of f*. Faris proved inequality (4.3) by showing that $\|F\|_1 \leqslant \|\nabla f\|_1$ in general, and then using the special properties of F to prove inequality (4.3) in the case where $f = F$. Several other authors [4, 10, 29] have also proved inequalities of the form $\varphi(\nabla F) \leqslant \varphi(\nabla f)$ for various functionals φ.

These inequalities would all follow immediately if it were the case that $(\nabla F)^*(t) \leqslant (\nabla f)^*(t)$ for all $t > 0$, but there are easy counter examples to the latter inequality. The argument in [4, Chapter 2, § 6] does show, however, that

$$(4.4) \qquad (\nabla F)^{**}(t) \leqslant (\nabla f)^{**}(t) \quad \text{for all } t > 0\,,$$

from which it follows that $\|\nabla F\|^*_{L(p,q)} \leqslant \|\nabla f\|^*_{L(p,q)}$ for all indices p and q satisfying the conditions specified at the beginning of this section. By a classical theorem [15] of Hardy, Littlewood and Polya, many inequalities of the form $\varphi(\nabla F) \leqslant \varphi(\nabla f)$ also follow from inequality (4.4).

The referee points out that the imbedding of $W^{1,1}$ in $L(K/(K-1), 1)$ follows very easily from known facts about spherical rearrangements. If f belongs to $C^\infty(R^K)$ and has compact support, then f has a spherical rearrangement, g say, that also belongs to $C^\infty(R^K$; moreover, $\|\nabla g\|_1 \leqslant \|\nabla f\|_1$. The function f^* will also be smooth, and

$$\|\nabla g\|_1 = c_K \int_0^\infty s^{1-1/K}\left(-\frac{df^*}{ds}(s)\right) ds\,.$$

On the other hand, integrating by parts in the integral above yields the quantity $c_K(1-1/K)\int_0^\infty s^{-1/K} f^*(s)\, ds$, which, as in our proof of Theorem 4.1, is equal to $c_K\|f\|_{L(K/(K-1),1)}$.

The method used by FARIS [12] also yields that $W^{1,p}(R^K) \subset L(Kp/(K-p), p)$ when $1 < p < K$. This inclusion had been proved earlier, by other methods, by STRICHARTZ [27]. As the referee points out, it also follows easily a theorem of O'Neil [20] concerning convolution with the kernel used by Sobolev in his original proof of the imbedding theorem for the case where $1 < p < K$. It is not clear to what extent the methods of the present paper can also be applied in this case.

We now deal with the other application of Theorem 4.1. The following result does not seem to have appeared in print before, although it follows from unpublished work of Gilles Pisier. Our method of proof is new; another new proof will be presented in a joint paper with Ron Blei.

THEOREM 4.4. – Let S be a discrete set carrying a counting measure, and let a be a function on S^2 that defines a bounded bilinear form on $l^\infty(S)$, with norm $\|a\|$. Then the function a belongs to the Lorentz space $l(s, 1)$, where $s = 4/3$; moreover, there is an absolute constant C so that

$$(4.5) \qquad \|a\|_{l(s,1)} \leqslant C\|a\|\,.$$

PROOF. – We use the symbol $l(s, 1)$ here rather than $L(s, 1)$ as a reminder that the function a is defined on the discrete measure space S^2. It actually suffices for the conclusion of the theorem that the function a define a bounded bilinear form on $c_0(S)$. If a has the latter property, then for each finite subset S' of S, the restriction of a to $S' \times S'$ defines a bounded bilinear form on $l^\infty(S')$, with norm at most $\|a\|$, the norm of a as a bilinear form on $c_0(S)$. Moreover, $\|a\|_{l(s,1)}$ is the supremum of the corresponding norms of restrictions of a to the various finite subsets $S' \times S'$. So, we may suppose that the set S is finite. We identify it with the set $\{1, 2, \ldots, N\}$, and represent a by a matrix $\{a_{m,n}\}_{m,n=1}^{N}$. It can be shown [16] that

$$\sum_n \left\{ \sum_m |a_{m,n}|^2 \right\}^{1/2} \leqslant \varkappa \|a\| , \tag{4.6}$$

and that the same is true for the corresponding permuted mixed-norm obtained by first taking l^2-norm with respect to n, and then the l^1-norm with respect to m. The constant $\varkappa$ comes from an application of the version of Khintchine's inequality asserting that $\|f\|_2 \leqslant \varkappa \|f\|_1$ if f is a sum of Rademacher functions.

Let τ be the transposition mapping the ordered pair $\{m, n\}$ to $\{n, m\}$. It follows from the family of inequalities (4.6) that

$$\|a\|_{\boldsymbol{p}} + \|\tau a\|_{\boldsymbol{p}} \leqslant 2\varkappa \|a\| , \tag{4.7}$$

where $\boldsymbol{p} = (2, 1)$. Transfer the function a from the discrete set S^2 to the measure space R^2 by defining a function A on R^2 that vanishes outside the set $[0, N)^2$, and is equal to $a_{m,n}$ on each set $(m, n) + [0, 1]^2$. It is easy to check that inequality (4.7) also holds with the function a replaced in the left side of the inequality by A; moreover, $\|a\|_{l(s,1)} = \|a\|_{L(s,1)}$, because the functions a and A have the same distribution function. Denote the spaces of functions, f say, on R^2 for which the quantities $\|f\|_{\boldsymbol{p}'}$ and $\|\tau f\|_{\boldsymbol{p}'}$ are respectively finite by D, and E; then $D \cap E$ is a Banach space with the norm given by

$$\|f\|_{D \cap E} = \|f\|_{\boldsymbol{p}} + \|\tau f\|_{\boldsymbol{p}}$$

for all f. Transferring inequality (4.7) yields that $\|A\|_{D \cap E} \leqslant 2\varkappa \|a\|$.

We claim that there is an absolute constant c so that

$$\|f\|_{L(s,1)} \leqslant c \|f\|_{D \cap E} \quad \text{for all measurable functions } f . \tag{4.8}$$

Once this claim is proved, we will have that $\|A\|_{L(s,1)} \leqslant 2ck\|a\|$, as required. To prove the claim, we pass to a dual version of it. We will see below that the dual spaces $(D \cap E)'$ and $L(s, 1)'$ can be identified with spaces of measurable functions on R^2. It follows that statement (4.8) holds if and only if

$$\|f\|_{(D \cap E)'} \leqslant c \|f\|_{L(s,1)'} \quad \text{for all measurable functions } f .$$

Recall [6, § 2.7] that $(D \cap E)'$, the dual space of $D \cap E$, is the algebraic sum of the dual spaces D', and E', with the norm given by

$$\|g\|_{(D \cap E)'} = \inf \{\max (\|h\|_{D'}, \|k\|_{E'}) : g = h + k\} .$$

It is easy to verify that the dual spaces D', and E' are permuted mixed-norm spaces based on the index sequence $\boldsymbol{q} = (2, \infty)$. Indeed, $\|g\|_{D'} = \|g\|_{\boldsymbol{q}}$ and $\|g\|_{E'} = \|\tau g\|_{\boldsymbol{q}}$ for all measurable functions g. Also, $L(s, 1)' = L(s', \infty) =$ weak-$L^{s'}$, with equivalence of norms. Recall that $s = 4/3$, so that $s' = 4$ here. Hence, our claim (4.8) is equivalent to the following statement.

(I) Every function g in the space $L(4, \infty)$ can be split as a sum of two functions h, and k with

$$\max \{\|h\|_{\boldsymbol{q}'} \|\tau k\|_{\boldsymbol{q}'}\} \leq C\|g\|_{L(4,\infty)} \tag{4.9}$$

for some absolute constant C.

We also claim that a similar statement holds for the space $L(2, \infty)$.

(II) Every function g in the space $L(2, \infty)$ can be split as a sum of two functions h and k with

$$\max \{\|h\|_{\boldsymbol{r}}, \|\tau k_{\boldsymbol{r}}\|\} \leq C' \|g\|_{L(2,\infty)} , \tag{4.10}$$

where $\boldsymbol{r}$ is the index sequence $(1, \infty)$, and C' is an absolute constant.

We arrived at claim (I) by a duality argument starting with the desired inequality (4.8). Similarly, claim (II) follows by duality from the case of Theorem 4.1 where $K = 2$. Since the latter theorem has already been proved, it suffices to derive claim (I) from claim (II). Suppose that $g \in L(4, \infty)$; then $|g|^2 \in L(2, \infty)$, and

$$\| |g|^2 \|_{L(2,\infty)} = (\|g\|_{L(4,\infty)})^2 .$$

Applying claim (II) yields a pair of functions H and K with $|g|^2 = H + K$, and with $\|H\|_{\boldsymbol{r}} \leq C'(\|g\|_{L(4,\infty)})^2$ and $\|\tau K\|_{\boldsymbol{r}} \leq C'(\|g\|_{L(4,\infty)})^2$. Suppose that H and K have supports. Then the splitting

$$g = (\operatorname{sgn}(g)) \cdot H^{1/2} + (\operatorname{sgn}(g)) \cdot K^{1/2}$$

has the properties specified in claim (I).

Matters therefore reduce to showing that claim (II) holds with disjointly-supported pieces. This is an easy consequence of the fact that the claim holds with pieces that are allowed to have overlapping supports. Given a function g in $L(2, \infty)$, split it as specified in the claim. Let U be the set where $|h| > |g|/2$. Then define functions H and K by letting $H = g$ on the set U, while $H = 0$ off U, and letting $K = g$ off U, while $K = 0$ on U. These functions have disjoint support, and their sum is g. Moreover, $|H| \leqslant 2|h|$, whence $\|H\|_r \leqslant 2C'\|g\|_{L(2,\infty)}$; similarly, $\|\tau K\|_r \leqslant 2C'\|g\|_{L(2,\infty)}$. This completes the proof of the theorem.

Our improvement on Littlewood's inequality has an interesting consequence for the known examples of p-Sidon sets in harmonic analysis. For example, if E is the set of integers of the form 3^m, then every continuous function, f say, on the unit circle whose Fourier coefficients $\hat{f}(n)$ vanish outside the set $E + E$, has the property that $f \in l^{4/3}$; for this reason, the set $E + E$ is called a $4/3$-*Sidon set.* The proof that $E + E$ has this property [11, 7] uses Littlewood's inequality, and it follows from Theorem 4.4 that in fact $\hat{f} \in l(4/3, 1)$ for all such functions f. This suggests the question: If F is a p-Sidon set for some index p in the interval $(1, 2)$, does it follow that $\hat{f} \in l(p, 1)$ for all continuous functions f whose transform vanishes off F? It follows from the results in this paper that the answer is « yes » for every set that is known to be a p-Sidon set with $p \in (1, 2)$.

Littlewood showed that the index $4/3$ in his theorem is best possible, by exhibiting matrices a with all entries equal to 1 or -1, and of arbitrarily large dimension, for which $\|a\|_{4/3} \geqslant c\|a\|$. By taking direct sums of such matrices, we can show for each sequence b in the space $l(4/3, 1)$ that there is a doubly-infinite matrix a, that defines a bounded bilinear form on l^∞, so that a majorizes b in the sense that $a^*(t) \geqslant b^*(t)$ for all $t > 0$. Therefore, $l(4/3, 1)$ is the smallest rearrangement-invariant sequence space that contains the sequence of matrix entries for every bounded bilinear form on l^∞. It is also known that the index p^* is best possible in the Sobolev imbedding theorem; the standard examples to this effect can be combined to show that for each function g in $L(K/(K-1), 1)$, and each positive number B, there is a function f in $W^{1,1}(R^K)$ so that $f^*(t) \geqslant g^*(t)$ for all t in the interval $(0, B)$.

In this paper, we have concentrated on Theorem 3.1 and its applications. We end the main part of the paper by briefly considering the corresponding applications of Theorem 3.3. Fix an integer n with $1 < n < K$; let $r(n) = K/(K - n)$. The method used above to prove Theorem 4.1 also yields that if $N^n(f) < \infty$, then $f \in L(r(n), 1)$ and $\|f\|_{L(r(n),1)} \leqslant cN^{(n)}(f)$; moreover, there is a multiplicative estimate similar to Corollary 4.2. It follows that the Sobolev space $W^{n,1}(R^K)$ imbeds into $L(r(n), 1)$; as noted in [21] this can also be deduced from Theorem 4.3 and the mapping properties of Riesz potentials.

We saw in Section 2 that there are versions of Littlewood's theorem for K-linear forms on l^∞, yielding estimates for $\|a\|_{2K/(K+1)}$ in terms of $\|a\|$. Using the case of Theorem 3.3 where $n = K - 1$, the corresponding generalization of Theorem 4.1, and duality as in the proof above of Theorem 4.4, we can show that $\|a\|_{L(2K/(K+1),1)} \leqslant c\|a\|$. The dual statements can also be proved directly; the direct proofs and various applications in the setting of fractional cartesian products will be presented in a joint paper with Ron Blei.

Appendix.

Each of our two main theorems has an alternate proof that is of interest in its own right. We first present a second proof of Theorem 4.1. Since

$$\|f\|_{L(r,1)} = \sup\{\|F\|_{L(r,1)} : F \text{ is simple with } 0 \leqslant F \leqslant |f|\},$$

matters reduce to proving the estimate $\|f\|_{L(r,1)} \leqslant cN(f)$ when f is a nonnegative simple function. The last step in Poornima's proof [21] that $W^{1,1} \subset L(r, 1)$ is a lemma stating that

$$\|f\|_{L(r,1)} = \int_0^\infty m_f(t)^{1/r}\,dt\,. \tag{A.1}$$

As Poornima points out, this equation is easy to verify when f is a simple function. The integral defining the norm and the integral on the right above become sums, which can be seen to be equal by a summation by parts.

To estimate $m_f(t)$, introduce the functions F_k and the sets A_m^k as in Section 3. It is easy to verify that

$$N_k(f) = \int_0^\infty m_{F_k}(t)\,dt\,. \tag{A.2}$$

In Section 3, we used Lemma 3.2 to show that $|A_m|^{K-1} \leqslant \prod_{k=1}^{K} |A_m^k|$ for all m. Equivalently,

$$m_f(t)^{K-1} \leqslant \prod_{k=1}^{K} m_{F_k}(t) \tag{A.3}$$

for all t. Recall that $r = K/(K-1)$. Taking K-th roots in inequality (A.3) yields the estimate

$$m_f(t)^{1/r} \leqslant \prod_{k=1}^{K} m_{F_k}(t)^{1/K}\,.$$

Using this in the right-hand side of inequality (A.1) gives that

$$\|f\|_{L(r,1)} \leqslant \int_0^\infty \prod_{k=1}^{K} m_{f_K}(t)^{1/K}\,dt\,. \tag{A.4}$$

Then by Hölder's inequality and equation (A.2)

$$\|f\|_{L(r,1)} \leqslant \prod_{k=1}^{K} N_k(f)^{1/K}\,.$$

This completes our alternate proof of Theorem 4.1.

To prove Theorem 3.1 in another way, we imitate an argument in Sobolev's classic paper [24], where it was shown that convolution with the kernel $x \to |x|^{1-k}$ maps $L^p(R^K)$ into $L^q(R^K)$, where $1 < p < K$ and $q = Kp/(K - P)$. This kernel is a radial function, and Sobolev proved the desired norm estimate by reducing matters to the case where the other functions in the estimate are radial too. He did this in two stages, first dealing with the case where $K = 2$, and then using induction and spherical symmetrization in $(K-1)$-dimensional hyperplanes. We will reprove our theorem first in the case where $K = 2$, and then use induction and *cubical* symmetrization in $(K-1)$-dimensional hyperplanes.

We suppose again that f is a nonnegative simple function with bounded support, and we denote the nonzero values of f, in decreasing order, by $a_1, a_2, \ldots, a_M$. We define the sets A_m and A_m^k as before, and we recall formula (3.4) of Section 3, which asserts that

$$N(f) = \sum_{m=1}^{M} \left[(a_m - a_{m+1})\Big(\sum_{k=1}^{K} |A_m^k|\Big)\right], \tag{A.5}$$

where again $a_{M+1} = 0$ by convention. Passing to a rearrangement of f does not change the numbers a_m and $|A_m|$, so that matters reduce to showing for each m that, given $|A_m|$, the quantity

$$\sum_{k=1}^{K} |A_m^k| \tag{A.6}$$

is minimal when A_m is essentially a cube with edges parallel to the coordinate axes.

In Section 3, we used Lemma 3.2 to prove the minimality of this sum when A_m is such a cube. When $K=2$, however, we can prove this in an even more elementary way. Choose a measure-preserving isomorphism σ mapping R onto itself so that the set $\sigma(A_m^1)$ is an interval of the form $[0, b_m)$. Similarly, choose a measure-isomorphism τ of R onto itself so that $\tau(A_m^2)$ has the form $[0, c_m)$. Let φ be the measure-preserving transformation of R^2 that maps each point (x, y) to $(\sigma(x), \tau(y))$. Replace A_m by the set $\varphi(A_m)$. This change has no effect on the quantities $|A_m^1|$ and $|A_m^2|$, and it reduces matters to the case where the sets A_m^1 and A_m^2 are intervals $[0, b_m)$ and $[0, c_m)$ respectively. In particular, A_m is essentially included in the rectangle $[0, b_m) \times [0, c_m)$, so that $|A_m| \leqslant b_m \cdot c_m$. Also, the quantity (A.6) is equal to $b_m + c_m$ in this case. This sum is minimal when $b_m = c_m = |A_m|^{1/2}$, that is when A_m is essentially a square with edges parallel to the coordinate axes.

In dealing with the corresponding question when $K > 2$, it is convenient to revert to the notation used in Lemma 3.2. Thus, let A be a bounded, measurable set in R^K, and denote its essential projection into the k-th coordinate hyperplane by A_k. Our task is to show, given $|A|$, that the quantity

$$N(1_A) = \sum_{k=1}^{K} |A_k|$$

is minimal when A is essentially a cube with edges parallel to the coordinate axes. Since translation has no effect on $N(1_A)$, we may suppose that A is a bounded, measurable subset of the first orthant in R^K.

Suppose without loss of generality that A is measurable with respect to the *uncompleted* product measure on R^K, and form a sequence of sets $A^{(n)}$ by the following procedure. Let $A^{(0)}=A$. Given a real number b, denote the hyperplane with equation $x_1=b$ by R_b^{K-1}. Let $A^{(1)}$ be the set whose intersection with each such hyperplane is a $(K-1)$-cube of the form $\{b\}\times\{0, b'\}^{K-1}$ with the same measure in R^{K-1} as the slice $A\cap R_b^{K-1}$. Then let $A^{(2)}$ be the set obtained from $A^{(1)}$ by replacing the intersection of the latter set with each hyperplane, Π_c say, where $x_2=c$ by the $(K-1)$-cube of the form $[0, b')\times\{c\}\times[0, b')^{K-2}$ with the same measure in R^{K-1} as the slice $A^{(1)}\cap\Pi_c$. Given $A^{(2)}$, construct $A^{(3)}$ by rearranging each slice of $A^{(2)}$ perpendicular to the x_1-axis into a $(K-1)$-cube of the type used in forming $A^{(1)}$, and continue in this fashion, using alternate rearrangements in hyperplanes perpendicular to the x_1 and x_2-axes.

Let $b=|A|^{1/K}$, and let B be the K-cube $[0, b)^K$. Our goal is to show that $N(1_B)\leqslant N(1_A)$. This inequality follows from two properties of the sequence $\{A^{(n)}\}$, namely that

$$N(1_{A^{(n+1)}})\leqslant N(1_{A^{(n)}}) \quad \text{for all } n, \tag{A.7}$$

and

$$N(1_{A^{(n)}})\to N(1_B) \quad \text{as } n\to\infty. \tag{A.8}$$

To see why the inequalities (A.7) hold, consider the special but typical case when $n=1$. For each positive number c, denote the intersection of $A^{(0)}$ with the hyperplane $x_1=c$ by $A^{(0)}(c)$, and define $A^{(1)}(c)$ similarly. Denote the R^{K-1}-measures of these sets by $|A^{(0)}(c)|$ and $|A^{(1)}(c)|$ respectively. The construction of $A^{(1)}$ from $A^{(0)}$ guarantees that $|A^{(1)}(c)|=|A^{(0)}(c)|$ for all c. Moreover, $N(1_{A^{(1)}})=\operatorname{ess\,sup}_c |A^{(1)}(c)|$, while $N(1_{A^{(0)}})\geqslant\operatorname{ess\,sup}_c |A^{(0)}(c)|$. Hence

$$N(1_{A^{(1)}})\leqslant N(1_{A^{(0)}}). \tag{A.9}$$

Identify the sets $A^{(0)}(c)$ and $A^{(1)}(c)$ with subsets of R^{K-1} in the obvious way, and make the inductive assumption that the inequalities (A.7) hold for subsets of R^{K-1}. Thus $N(1_{A^{(2)}(c)})\leqslant N(1_{A^{(1)}(c)})$ for all c. Integrating this inequality with respect to c yields that

$$\sum_{k=2}^{K} N_k(1_{A^{(2)}})\leqslant\sum_{k=2}^{K} N_k(1_{A^{(1)}}). \tag{A.10}$$

Adding inequalities (A.9) and (A.10) yields that $N(1_{A^{(2)}})\leqslant N(1_{A^{(1)}})$, as claimed above.

Finally, we outline the proof of assertion (A.8). Each of the sets $A^{(n)}$, for $n\geqslant 1$, is a union of $(K-1)$-cubes, all perpendicular to the x_1-axis if n is odd, and all perpendicular to the x_2-axis if n is even. Define the *breadth of* $A^{(n)}$ to be the number $b(n)$ given by

$$b(n) = \sup\{x_3 : \boldsymbol{x} \in A^{(n)}\},$$

and the *length of* $A^{(n)}$ to be the number $c(n)$ given by

$$c(n) = \sup\{x_1 : \boldsymbol{x} \in A^{(n)}\} \quad \text{if } n \text{ is odd},$$

and

$$c(n) = \sup\{x_2 : \boldsymbol{x} \in A^{(n)}\} \quad \text{if } n \text{ is even}.$$

Then $A^{(n)}$ is included in a box with edges parallel to the coordinate axes, and with one dimension equal to $c(n)$ and the other $K-1$ dimensions equal to $b(n)$.

We claim that the sequences $\{b(n)\}$ and $\{c(n)\}$ both converge to the number $b = |A|^{1/K}$. Suppose, for the moment that this claim is true and let $d(n) = \max\{b(n), c(n)\}$. Then

$$\limsup_{n\to\infty} |A_1^{(n)}| \leqslant \lim_{n\to\infty} d(n)^{K-1} = b^{K-1}.$$

On the other hand, the set $A^{(n)}$, which has volume b^K, is included in the cylinder $[0, d(n)) \times A_1^{(n)}$, so that

$$b^K \leqslant \limsup_{n\to\infty} d(n) \cdot |A_1^{(n)}| = b \cdot \limsup_{n\to\infty} |A|_1^{(n)}.$$

Hence $|A_1^{(n)}| \to b^{K-1}$ as $n \to \infty$. The quantities $|A_k^{(n)}|$, where $K > 1$, also converge to b^{K-1} as $n \to \infty$, and assertion (A.8) holds.

In proving our claim about the limiting behaviour of $b(n)$ and $c(n)$, we first use a uniform change of scale to reduce matters to the case where $|A| = 1$; then $b = 1$ also. We now estimate $b(n+1)$ and $c(n+1)$ in terms of $c(n)$ and $b(n)$. Suppose for definiteness, that n is odd; if n is even, the only modification needed in our argument is an exchange of the roles of x_1 and x_2. The set $A^{(n)}$ is a union of $(K-1)$-cubes of the form $\{b\} \times [0, \varphi(b))^K$, with $\varphi(b) \leqslant b(n)$. So, if $c \geqslant b(n)$, then the hyperplane with equation $x = c$ does not intersect the set $A^{(n)}$. Therefore,

$$c(n+1) \leqslant b(n). \tag{A.11}$$

On the other hand, the hyperplane with equation $x_2 = 0$ intersects each of the nonempty $(K-1)$-cubes, perpendicular to the x_1-axis, that make up the set $A^{(n)}$. The intersection of this hyperplane with the $(K-1)$-cube where $x_1 = b$ is a $(K-2)$-cube with edge-length $\varphi(b)$; any other hyperplane perpendicular to the x_2-axis either misses the $(K-1)$-cube where $x_1 = b$, or also intersects it in a

(K-2)-cube with edge-length $\varphi(b)$. So, among the hyperplanes perpendicular to the x_2-axis, the one where $x_2 = 0$ intersects $A^{(n)}$ in a set of maximal R^{K-1}-measure. It follows that

$$b(n+1)^{K-1} = \int_0^{c(n)} \varphi(b)^{K-2}\, db \, .$$

Apply Hölder's inequality with the conjugate indices $(K-1)/(K-2)$ and $K-1$ to get that

$$b(n+1)^{K-1} \leqslant \Big\{ \int_0^{c(n)} \varphi(b)^{K-1}\, db \Big\}^{(K-2)/(K-1)} \cdot c(n)^{1/(K-1)} \, .$$

The integral above is equal to $|A^{(n)}|$, which has been normalized to be 1. Hence,

$$b(n+1) \leqslant c(n)^{1/(K-1)^2} \, . \tag{A.12}$$

By iterating inequalities (A.11) and (A.12) we get that

$$\max\{b(n+2), c(n+2)\} \leqslant \max\{b(n), c(n)\}^{1/(K-1)^2} \quad \text{for all } n \, .$$

So, $\limsup_{n\to\infty} b(n)$ and $\limsup_{n\to\infty} c(n)$ are both at most 1. On the other hand, $b(n)^{K-1} \cdot c(n) \geqslant |A^{(n)}| = 1$. for all n. Therefore,

$$\lim_{n\to\infty} b(n) = \lim_{n\to\infty} c(n) = 1 = b \, ,$$

as required. This completes our alternate proof of Theorem 3.1.

As we said at the beginning of the proof, the process used above is the analogue for cubes of a spherical symmetrization process used [24] by Sobolev. We briefly consider the analogue, for balls, of Theorem 3.1. Let U be the group of orthogonal transformations of R^K; denote a generic element of U by g, and the Haar measure on U bu $d\sigma$. For each measurable function f on R^K, let

$$N_U(f) = \int_U N_1(\sigma f)\, d\sigma \, ,$$

where again $\sigma f(\boldsymbol{x}) = f(\sigma(\boldsymbol{x}))$ for all vectors $\boldsymbol{x}$ in R^K. It is plausible that the rearrangements of f that minimize the quantity $N_U(f)$ are the ones for which the sets where $|f| > \lambda$ are balls centred at the origin. It also seems plausible that if $N(\sigma f) < \infty$ some transformation σ in the group U, then $N_U(f) < \infty$.

REFERENCES

[1] R. A. ADAMS, *Sobolev Spaces*, Academic Press, New York (1975).

[2] R. A. ADAMS - J. FOURNIER, *Cone conditions and properties of Sobolev spaces*, J. Math. Anal. Appl., **61** (1977), pp. 713-734.

[3] G. R. ALLEN, *An inequality involving product measures*, Radical Banach Algebras and Automatic Continuity (Long Beach, Calif., 1981), pp. 277-279; Lecture Notes in Math., **975**, Springer, Berlin - New York (1983).

[4] T. AUBIN, *Nonlinear Analysis on Manifolds. Monge-Ampère Equations*, Grundlehren der mathematischen Wissenschaften, **252**, Springer-Verlag, New York - Heidelberg - Berlin (1982).

[5] A. BENEDEK - R. PANZONE, *The spaces L^p with mixed norms*, Duke Math. J., **28** (1961), pp. 301-324.

[6] J. BERGH - J. LÖFSTROM, *Interpolation Spaces. An Introduction*, Grundlehren der mathematischen Wissenschaften, **223**, Springer-Verlag, Berlin - Heidelberg - New York (1976).

[7] R. C. BLEI, *Sidon partitions and p-Sidon sets*, Pacific J. Math., **65** (1976), pp. 307-313,

[8] R. C. BLEI, *Fractional cartesian products of sets*, Annales Institut Fourier (Grenoble). **29** (1979), pp. 79-105.

[9] A. P. CALDERÓN, *An inequality for integrals*, Studia Math., **57** (1976), pp. 275-277.

[10] G. F. D. DUFF, *A general integral inequality for the derivative of an equimeasurable rearrangement*, Canadian J. Math., **28** (1976), pp. 793-804.

[11] R. E. EDWARDS - K. A. ROSS, *p-Sidon sets*, J. Functional Anal., **15** (1974), pp. 404-427.

[12] W. J. FARIS, *Weak Lebesgue spaces and quantum mechanical binding*, Duke Math. J., **43** (1976), pp. 365-373.

[13] E. GAGLIARDO, *Proprietà di alcune classi di funzioni in più variabili*, Ricerche Mat., **7** (1958), pp. 102-137.

[14] G. H. HARDY - J. E. LITTLEWOOD - G. PÓLYA, *Inequalities*, Cambridge University Press, Cambridge (1934).

[15] G. H. HARDY - J. E. LITTLEWOOD - G. PÓLYA, *Some simple inequalities satisfied by convex functions*, Messenger of Math., **58** (1929), pp. 145-152.

[16] J. E. LITTLEWOOD, *On bounded bilinear forms in an infinite number of variables*, Quarterly J. Math., **1** (1930), pp. 164-174.

[17] L. H. LOOMIS - H. WHITNEY, *An inequality related to the isoperimetric inequality*, Bull. Amer. Math. Soc., **55** (1949), pp. 961-962.

[18] J. MOSER, *A sharp form of an inequality by N. Trudinger*, Indiana Univ. Math. J., **20** (1971), pp. 473-484.

[19] L. NIRENBERG, *On elliptic partial differential operators*, Annali della Scuola Normale Sup. Pisa, **13** (1959), pp. 116-162.

[20] R. O'NEIL, *Convolution operators and $L(p, q)$ spaces*, Duke Math. J., **30** (1963), pp. 129-142.

[21] S. A. POORNIMA, *An embedding theorem for the Sobolev space $W^{1,1}$*, Bull. Sci. Math., **107** (1983), pp. 253-259.

[22] T. PRACIANO-PEREIRA, *On bounded multilinear forms on a class of l^p spaces*, J. Math. Anal. Appl., **81** (1981), pp. 561-568.

[23] A. J. SCHWENK - J. I. MUNRO, *How small can the mean shadow of a set be?*, American Math. Monthly, **90** (1983), pp. 325-329.

[24] S. L. SOBOLEV, *On a theorem of functional analysis*, Mat. Sb., **46** (1938), pp. 471-496; translated in Amer. Math. Soc. Transl., **34** (1963), pp. 39-68.

[25] E. M. STEIN, *Singular Integrals and Differentiability Properties of Functions*, Princeton Mathematical Series # 30, Princeton University Press, Princeton (1970).

[26] E. M. STEIN - G. WEISS, *Introduction to Fourier Analysis an Euclidean Spaces*, Princeton Mathematical Series # 32, Princeton University Press, Princeton (1971).

[27] R. S. STRICHARTZ, *Multipliers on fractional Sobolev spaces*, J. Math. Mech., **16** (1967), pp. 1031-1060.

[28] J. SZELMECZKA, *On a property of a function belonging to various spaces with mixed norms*, Functiones et Approximatio, **9** (1980), pp. 25-28.

[29] G. TALENTI, *Best constant in Sobolev's inequality*, Ann. Mat. Pura Appl., **110** (1976), pp. 353-372.