Preprint (1999),

Embedding subspaces of L_p into ℓ_p^N, $0 < p < 1$

By Gideon Schechtman of Rehovot and Artem Zvavitch of Rehovot

( )

Abstract. It is shown that if X is n-dimensional subspace of L_p, $0 < p < 1$, then there exists a subspace Y of ℓ_p^N such that $d(X, Y) \leq 1 + \varepsilon$ and $N \leq C(\varepsilon, p)n(\log n)^3$.

1. Introduction

Given an n-dimensional subspace X of $L_p([0,1], dt)$ and, $\varepsilon > 0$ what is the smallest integer $N = N(n, \varepsilon)$ such that there is a subspace Y of ℓ_p^N with $d(X, Y) \leq 1 + \varepsilon$, where $d(X, Y)$ is the Banach-Mazur distance? For $p \geq 1$ the dependence of N on n is known, except for possibly redundant log factors. (It is, forgetting log factors, n, for $1 \leq p < 2$, and $n^{p/2}$, for $2 < p < \infty$. See Bourgain, Lindenstrauss and Milman [1], Talagrand [11], [12] for the best known results). In the case of $0 < p < 1$ the known result is only $N(n, \varepsilon) \leq c(\varepsilon, p)n^2$ (see Schechtman [10]). Recently Peña [9] considered the problem we address in the present paper, i.e. extending the result of [1] to the range $0 < p < 1$, and observed that any X as above admits a $Y \subseteq \ell_p^N$ with $N \leq C(p)n$ and $d(X, Y) \leq C(p)(\log n)^{1/p}$. The problem of reducing the $\log n$ factor to $1 + \varepsilon$ was left open in [9] and was the trigger to the present note. For a more extensive review of the background material see the forthcoming [4].

In this paper we extend the result of [1] to the range $0 < p < 1$ by proving the following result.

Theorem 1.1. *Let $0 < p < 1$ then*

$$N(n, \varepsilon) \leq C(p, \varepsilon)n(\log n)^3.$$

Throughout this paper C, c, c_1 etc. denote absolute constants whose value may change from line to line. Similarly $C(p), c(p), C(p, \varepsilon)$ denote constants depending on the arguments, p or p and ε, only.

1991 *Mathematics Subject Classification.* Primary 46E30, 46B07; Secondary 46B09, 60G15.
Keywords and phrases. L_p, ℓ_p^N, Lewis' change of density, entropy estimates.

The proof is by the method initiated in [10] and refined in [1] of random selection of coordinates. We may assume that $X \subseteq \ell_p^M$ for some finite (but large) M and we would like to choose randomly a small set A of coordinates for which the operator of restriction to A is an isomorphism on X. It is easy to see that this is impossible in general (for example if X is essentially supported on a small subset of the coordinates). So we need to first change the "location" of X in ℓ_p^M. This is done (as in all the previous references) via change of density. In the existing literature referred to above one uses the so called Lewis' change of density. This was available until now only in the range $p \geq 1$, indeed, the proof in [6] uses duality in an essential way. In section 2. below we extend Lewis' result to the range $0 < p < 1$. The proof is new and simpler even in the known range and we believe it is of independent interest.

Rather than choosing the set A of coordinates in one step we prefer do to it iteratively reducing the size of M by a factor of approximately $1/2$ each time. (This is the approach of [11].) This amounts to estimating

$$\sup_{x \in X, \|x\| \leq 1} \Big| \sum_{i=1}^{M} \varepsilon_i |x_i|^p \Big| \tag{1.1}$$

over a typical choice of signs ε_i, $i = 1, 2, \ldots, M$. Indeed, if for a particular choice of signs this quantity is smaller than δ then it is easy to see that the restriction operator to both $A = \{i;\ \varepsilon_i = 1\}$ and $A^c = \{i;\ \varepsilon_i = -1\}$ is an $\left(\frac{1+\delta}{1-\delta}\right)^{1/p}$-isomorphism.

The expected value of (1.1) over all choices of signs is dominated by a similar expression with independent Gaussian variables replacing the ε_i -s:

$$\mathbb{E} \sup_{x \in X, \|x\| \leq 1} \Big| \sum_{i=1}^{M} g_i |x_i|^p \Big|.$$

To estimate this quantity we use, as is done implicitly in [1], a theorem of Dudley. This is done in section 4.. The method requires fine entropy estimates and these are developed in section 3.. The proof of the entropy estimates is based on a proof of a similar estimate in [1]. The main difficulty for us was to realize that in spite of the way it is written there these estimates do not use duality in an essential way. In the last section of this paper we gather all the ingredients into a proof of Theorem 1.1

2. Lewis' theorem in the non-convex setting

We first state and prove a theorem (due to Lewis [6] in the range $1 \leq p < \infty$) concerning the existence of special bases in finite dimensional subspaces of L_p. Even for $p \geq 1$ the proof is somewhat different and more direct than that of Lewis and in particular does not use explicitly Lewis' lemma [7] nor the notion of p-summing operators. [8] (see also [13]) contains another proof, simpler than the original, for the case $1 < p < \infty$.

Theorem 2.1. *Let X be an n-dimensional subspace of $L_p(\Omega,\mu)$, $0<p<\infty$. Then, there is a basis $f_1,\dots,f_n$ of X satisfying*

$$\int\Big(\sum_{i=1}^n f_i^2\Big)^{(p-2)/2} f_k f_l d\mu=\delta_{kl},\ 1\le k,l\le n. \tag{2.1}$$

In particular, one can find a probability measure ν on Ω and a subspace $\tilde X$ of $L_p(\Omega,\nu)$ isometric to X which admits a basis $h_1,\dots,h_n$ orthonormal in $L_2(\Omega,\nu)$, such that $\sum_{i=1}^n h_i^2\equiv n$.

Proof. We first assume that Ω is finite. Assume without loss of generality that X has maximal support, i.e., for every $t\in\Omega$, there is a $x\in X$ with $x(t)\ne 0$.

Let $\{x_i\}_{i=1}^n$ be a basis for X. Given an $n\times n$ matrix $B=(b_{ij})_{i,j=1}^n$, we consider it as an operator from $\mathbb{R}^n$ into X sending the basis vector e_i to $\sum_{j=1}^n b_{ij}x_j$. Let $A=(a_{ij})_{i,j=1}^n$ be an $n\times n$ matrix which maximizes $\det(B)$ subject to

$$\Big\|\Big(\sum_{i=1}^n (Be_i)^2\Big)^{1/2}\Big\|_p\le 1.$$

Note that necessarily $\{Ae_i\}_{i=1}^n$ is a basis for X and in particular $\sum_{i=1}^n (Ae_i)^2(t)>0$ for all $t\in\Omega$. Consequently, for every t the function $\Big(\sum_{i=1}^n (Be_i)(t)^2\Big)^{p/2}$ is continuously differentiable at A as a function of (b_{ij}) and

$$\begin{aligned}\frac{\partial}{\partial b_{kl}}\Big\|\Big(\textstyle\sum_{i=1}^n (Be_i)^2\Big)^{1/2}\Big\|_p^p &= \frac{\partial}{\partial b_{kl}}\int\Big(\textstyle\sum_{i=1}^n\Big(\sum_{j=1}^n b_{ij}x_j(t)\Big)^2\Big)^{p/2}d\mu(t)\\ &= p\int\Big(\textstyle\sum_{i=1}^n\Big(\sum_{j=1}^n b_{ij}x_j(t)\Big)^2\Big)^{(p-2)/2}\Big(\sum_{s=1}^n b_{ks}x_s(t)\Big)x_l(t)d\mu(t).\end{aligned} \tag{2.2}$$

Now, as is well known (expand $\det(B)$ by the k^{th} row), $\frac{\partial}{\partial b_{kl}}\det(B)$ is the lk element of B^{-1} multiplied by $\det(B)$. By the method of Lagrange multipliers, we get that $C=\Big(\frac{\partial}{\partial a_{kl}}\Big\|\Big(\sum_{i=1}^n (Ae_i)^2\Big)^{1/2}\Big\|_p^p\Big)_{i,j}^n$ is a multiple of $\Big(\frac{\partial}{\partial a_{kl}}\det(A)\Big)_{i,j}^n$ which, as we have just remarked, is a multiple of A^{*-1}, i.e., A^*C is a multiple of the identity matrix. Put $f_k=\sum_{s=1}^n a_{ks}x_s$, then it follows from (2.2) that for some $\lambda\in\mathbb{R}$,

$$\int\Big(\sum_{i=1}^n f_i^2(t)\Big)^{(p-2)/2} f_k(t)f_u(t)d\mu(t)=\lambda\delta_{uk}$$

which, after normalization, gives a basis satisfying (2.1). Put

$$f(t)=\Big(\sum_{i=1}^n f_i^2(t)\Big)^{1/2},\quad \nu(\{t\})=\mu(\{t\})f^p(t)/\int f^p(s)d\mu(s),$$

$$\tilde X=\{x/f\ ;\ x\in X\}\ \text{ and }\ h_i=n^{1/2}f_i/f$$

then $h_1,\dots,h_n$ is the required basis in the 'In particular' part of the theorem.

This finishes the proof in the case that μ is finitely supported. This case is all we need for the rest of this paper so we only sketch the easy generalization. First, by changing the density, i.e. dividing the elements of X by some positive L_p function φ and changing the measure to $\varphi^p d\mu$, we may assume that for some constant K (depending on n) every norm one function in X is bounded by K. Now, for every ε there is a subspace X_ε of the same L_p space which consists of simple functions and which is ε close to X in the sense that for every norm one function f in X (resp. X_ε) there is a function g in X_ε (resp. X) with $\|f-g\|<\varepsilon$. By the first part of the proof there is a basis $f_1^\varepsilon,\dots,f_n^\varepsilon$ in X_ε satisfying (2.1). Any limit $f_1,\dots,f_n$ of these bases as $\varepsilon\to 0$ along some subsequence, will satisfy (2.1) as well and be in X. To get the basis $h_1\dots,h_n$ we change the density again as in the discrete case above. □

Remarks: 1. Note that if X is a subspace of ℓ_p^M and μ is the probability measure on $\{1,\dots,M\}$ given by the theorem above then, as observed in [3] we can split each atom of ν of mass large than $4/M$ into pieces each of size between $2/M$ and $4/M$. This will enlarge the number of atoms by at most $M/2$. The new measure, which we still denote by ν, is now supported on $\{1,\dots,M'\}$, where $M'\le 3M/2$. $\mu(\{i\})\le 4/M$ for all i, X is still isometric to a subspace (which we continue to denote by X) of L_p of the new measure space and the new X admits an orthonormal basis $h_1,\dots,h_n$ whose sum of squares is a constant.

2. Let $\tilde X$ be as given in Theorem 2.1. If $x\in\tilde X$ then, for $0<p<2$,

$$\|x\|_2\le n^{1/p-1/2}\|x\|_p \text{ and } \|x\|_\infty\le n^{1/p}\|x\|_p.$$

Indeed, If $x\in\tilde X$ and $x=\sum_{i=1}^n\alpha_i h_i$ then

$$\|x\|_\infty\le(\sum_{i=1}^n\alpha_i^2)^{1/2}\|(\sum_{i=1}^n h_i^2)^{1/2}\|_\infty\le n^{1/2}\|x\|_2$$

and

$$\|x\|_2^2\le\|x\|_\infty^{2-p}\|x\|_p^p\le n^{(2-p)/2}\|x\|_2^{2-p}\|x\|_p^p.$$

Thus $\|x\|_2\le n^{1/p-1/2}\|x\|_p$.
Note that this implies in particular that $d(X,\ell_2^n)\le n^{1/p-1/2}$ for every n dimensional subspace of L_p. This extends a result of Lewis [6] to the range $0<p<1$ (but it is easy to get this extension also by other means).

3. The basis with property (2.1) is unique up to orthogonal transformation: *Let X and $f_1,\dots,f_n$ be as in Theorem 2.1 and let $g_1,\dots,g_n$ be another basis satisfying (2.1), i.e.*

$$\int F^{p-2}f_kf_l d\mu=\int G^{p-2}g_kg_l d\mu=\delta_{kl},\ 1\le k,l\le n \tag{2.3}$$

where $F=\left(\sum_{i=1}^n f_i^2\right)^{1/2}$ and $G=\left(\sum_{i=1}^n g_i^2\right)^{1/2}$. Then there exists an orthonormal matrix $U=(a_{ij})$ such that $f_i=\sum_{j=1}^n a_{ji}g_j,\ \ i=1,\dots,n.$ We are going to show that $F=G$ from which the assertion easily follows. Denote by X_F (resp. X_G) X

considered as a subspace of $L_2(F^{p-2}d\mu)$ (resp. $L_2(G^{p-2}d\mu)$. Define three Hilbert space isometries, $u : \ell_2^n \to X_F$, $v : \ell_2^n \to X_G$, and $T : X_F \to X_G$ by

$$u(e_i) = f_i,\ v(e_i) = g_i,\ i = 1, \dots, n \text{ and } Tx = xG^{-(p-2)/2}F^{(p-2)/2}.$$

$U = v^{-1}Tu$ is a unitary operator on ℓ_2^n represented by the orthogonal matrix $U = (a_{ij})$. Then on one hand $Tf_i = \sum_{j=1}^n a_{ji}g_j$ and on the other hand $Tf_i = f_i G^{-(p-2)/2}F^{(p-2)/2}$ squaring, summing over i and using the fact that $\sum_{i=1}^n (\sum_{j=1}^n a_{ji}g_j)^2 \equiv \sum_{i=1}^n g_i^2$ we get that $F = G$.

4. As we remarked above Lewis' proof of Theorem 2.1 is based on another theorem of Lewis [7] which states that if α is a norm on $\mathbb{R}^{n^2}$ regarded as the space of $n \times n$ matrices and α^* is the dual norm where duality is given by the trace (traceBA $= \langle \mathrm{B}^*, \mathrm{A} \rangle$, where $\langle \cdot, \cdot \rangle$ is the usual scalar product in $\mathbb{R}^{n^2}$) then there is an invertible matrix A with $\alpha(A) = 1$ and $\alpha^*(A^{-1}) = n$. The proof (first used by F. John for α being an operator norm) is by a variational argument and A is obtained as a point at which $\det(A)$ is maximal under the constraint: $\alpha(A) \le 1$. The proof is quite simple but it may be instructive to notice that, although this is not the way it is commonly presented, it is really just a solution of a maximization problem under constraints, using the method of Lagrange multipliers. Indeed, assuming the norm α is smooth, the gradient of $\alpha(A)$ considered as a function on $\mathbb{R}^{n^2} \setminus \{0\}$ is (like the gradient of any norm on $\mathbb{R}^m$) the norming functional of A, i.e., a matrix $B = B_A$ such that $\alpha(B) = 1$ and traceBA $= \alpha(\mathrm{A})$. Since the gradient of $\det(A)$ is a multiple of A^{*-1} we get that B_A is a multiple of A^{-1}. (The reason for removing the adjoint is the relation between trace duality and the usual duality in $\mathbb{R}^{n^2}$.) It is now easy to see that the multiple must be $1/n$.

3. Entropy estimates

Let X be an n dimensional subspace of $L_p(\Omega, \nu)$ for some probability measure ν and assume X admits an orthonormal basis $f_1, \dots, f_n$ with $\sum f_i^2 \equiv n$. For $0 < q \le \infty$ denote by $B_q = B_q(X)$ the unit ball of X in the $L_q(\Omega, \nu)$ norm. By $E(B_r, B_s, t)$ we denote the metric entropy of B_p with respect to tB_s, i.e., the minimal number of translates of tB_s needed to cover B_r. The case $p \ge 1$ of the next proposition has been implicitly proved in [1] (see the proof of Proposition 7.2 there). A superficial glance at the proof there seems to imply that duality is used. As we shall see below this is not really the case and a small variation of the proof works also in the range $0 < p < 1$.

Proposition 3.1. *Let $0 < p < 2$ then, for any $t > 1$,*

$$\log E(B_p, B_2, t) \le c(p)n(\log n)t^{-2p/(2-p)}.$$

Proof. For $k = 0, 1, \dots$, let $\mathcal{E}_k$ be a maximal subset of B_p with respect to the condition: $||f - g||_2 > 8^k t$ for every $f \ne g$ in $\mathcal{E}_k$. (For k such that $8^k t > n^{(2-p)/2p}$ we take $\mathcal{E}_k = \{0\}$). Then

$$|\mathcal{E}_k| \ge E(B_p, B_2, 8^k t).$$

Since, for every k there are $x_i \in B_p$, $i = 1, 2, \ldots, E = E(B_p, B_2, 8^{k+1}t)$, satisfying $\cup_{i=1}^{E}\{x_i + 8^{k+1}tB_2\} \supseteq B_p$, we can find for every k a $y_k \in B_p$ such that

$$|\{y_k + 8^{k+1}tB_2\} \cap \mathcal{E}_k| \geq \frac{|\mathcal{E}_k|}{E(B_p, B_2, 8^{k+1}t)} \geq \frac{E(B_p, B_2, 8^k t)}{E(B_p, B_2, 8^{k+1}t)}. \tag{3.1}$$

Let $\mathcal{L}_k = \{(f - y_k)/8^{k+1}t, f \in \{y_k + 8^{k+1}tB_2\}\} \cap \mathcal{E}_k$. Then,

$$||g||_2 \leq 1, ||g||_p \leq 2^{1/p\vee 1}/8^{k+1}t, \text{ for any } g \in \mathcal{L}_k$$

and

$$||g - g'||_2 > 1/8, \text{for any } g \neq g', g, g' \in \mathcal{L}_k.$$

For $g \neq g'$ and $0 < \theta < p$,

$$\begin{aligned} ||g - g'||_2^2 &= \int_\Omega |g - g'|^\theta |g - g'|^{2-\theta} d\nu \\ &\leq \left(\int_\Omega |g - g'|^p d\nu\right)^{\theta/p} \left(\int_\Omega |g - g'|^{(2-\theta)p/(p-\theta)} d\nu\right)^{(p-\theta)/p}. \end{aligned}$$

Let $r = (2 - \theta)p/(p - \theta)$, then $r > 2$, and for any $g \neq g', g, g' \in \mathcal{L}_k$,

$$\begin{aligned} 8^{-2} &< ||g - g'||_2^2 = \int |g - g'|^\theta |g - g'|^{2-\theta} \\ &\leq ||g - g'||_p^\theta ||g - g'||_r^{2-\theta} \leq (2^{1/p\vee 1}/8^{k+1}t)^\theta ||g - g'||_r^{2-\theta} \end{aligned}$$

or

$$||g - g'||_r > 8^{-3/(2-\theta)}(8^{k+1}t)^{\theta/(2-\theta)} \text{ for all } \; g \neq g', g, g' \in \mathcal{L}_k.$$

It particular,

$$E(B_2, B_r, 8^{-3/(2-\theta)}(8^{k+1}t)^{\theta/(2-\theta)}) \geq |\mathcal{L}_k|. \tag{3.2}$$

We now use (7.13) in the proof of Proposition 7.2 of [1] according to which

$$\log E(B_2, B_r, s) \leq crn(\log n)(s/c)^{-2\frac{(2-\theta)p}{\theta(2-p)}}.$$

It follows that

$$\log E(B_2, B_r, 8^{-3/(2-\theta)}(8^{k+1}t)^{\theta/(2-\theta)}) \leq c^{1+2\frac{(2-\theta)p}{\theta(2-p)}} rn(\log n)(8^{k+1}t)^{-2p/(2-p)}$$

and picking $\theta = p/2$ we get

$$\log E(B_2, B_r, 8^{-3/(2-\theta)}(8^{k+1}t)^{\theta/(2-\theta)}) \leq c^{1/(2-p)} n(\log n)(8^{k+1}t)^{-2p/(2-p)}.$$

Now, (3.1) and (3.2) imply that

$$\begin{aligned} c^{1/(2-p)} n(\log n)(8^{k+1}t)^{-2p/(2-p)} &\geq \log|\mathcal{L}_k| = \log|\{y_k + 8^{k+1}tB_2\} \cap \mathcal{E}_k| \\ &\geq \log E(B_p, B_2, 8^k t) - \log E(B_p, B_2, 8^{k+1}t). \end{aligned}$$

By summing these inequalities over all k we obtain

$$\log E(B_p, B_2, t) \leq c(p)n(\log n)t^{-2p/(2-p)}.$$

The constant $c(p)$ is bounded by $c^{1/(2-p)}/p$ for some universal c. $\square$

Remark: As in [1] one can get also that for every n dimensional subspace X of L_p and for every ε there is another subspace $\tilde{X}$ of L_p whose Banach-Mazur distance from X is at most $1+\varepsilon$ and, letting $\tilde{B}_q$ denote the unit ball of $\tilde{X}$ with respect to the L_q norm,

$$\log E(\tilde{B}_p, \tilde{B}_\infty, t) \le c(p) n \log(n\varepsilon^{-1}) t^{-p}$$

for $t > 1$.

4. A Bound on a Gaussian process

Let X be an n dimensional subspace of $L_p(\Omega, \nu)$ for a set Ω with K points and some probability measure ν satisfying $\nu(\{i\}) \le 6/K$ and assume X admits an orthonormal basis $f_1, \ldots, f_n$ with $\sum f_i^2 \equiv n$.

Proposition 4.1. *Let $0 < p < 1$ and let g_i, $i = 1, \ldots, K$ be independent standard Gaussian random variables. Then*

$$\mathbb{E} \sup_{x \in B_p} \left| \sum_{i=1}^{K} \nu(\{i\}) g_i |x(i)|^p \right| \le c(p) \left(\frac{n}{K} (\log n)^3 \right)^{\frac{1}{2}}.$$

Proof. Let δ be the natural distance associated with the Gaussian process $\{G_x = \sum_{i=1}^{K} \nu(\{i\}) g_i |x(i)|^p\}_{x \in B_p}$, i.e.,

$$\delta(y,z) = \Big(\sum_{i=1}^{K} [\nu(\{i\})(|y_i|^p - |z_i|^p)]^2 \Big)^{1/2}, \quad y, z \in B_p.$$

Let $\theta = (1-p)/(2-p)$ then

$$\begin{aligned} \delta(y,z)^2 &\le \sum_{i=1}^{K} \nu(\{i\})^2 |y_i - z_i|^{2p} \\ &\le 6K^{-1} \|y - z\|_p^{\theta 2p} \|y - z\|_2^{(1-\theta)2p} \\ &\le 12K^{-1} \|y - z\|_2^{2p/(2-p)}. \end{aligned} \tag{4.1}$$

By the second remark following Theorem 2.1 $\|x\|_2 \le n^{(2-p)/2p} \|x\|_p$ for all $x \in X$ It follows that the δ diameter of B_p is at most $(12n)^{1/2} K^{-1/2}$.

Proposition 3.1 and (4.1) also imply

$$\begin{aligned} \log E(B_p, \delta, t) &\le \log E(B_p, (12K^{-1})^{1/2} \| \cdot \|_2^{p/(2-p)}, t) \\ &\le \log E(B_p, \| \cdot \|_2, 12^{-(2-p)/2p} K^{(2-p)/2p} t^{(2-p)/p}) \\ &\le c(p) n (\log n) K^{-1} t^{-2}. \end{aligned}$$

The last inequality in the later display requires that $t \geq 12^{1/2}K^{-1/2}$. For $0 < t < 12^{1/2}K^{-1/2}$ we first obtain that

$$\begin{aligned}\delta(y,z)^2 &\leq \sum_{i=1}^{K} \nu(\{i\})^2 |y_i - z_i|^{2p} \\ &\leq \Big(\sum_{i=1}^{K} (\nu\{i\}) |y_i - z_i|^p\Big)^2 \leq \|y-z\|_p^{2p}.\end{aligned}$$

Then we use standard volume considerations in the n-dimensional space X to obtain

$$\begin{aligned}\log E(B_p, \delta, t) &\leq \log E(B_p, \|\cdot\|_p^p, t) \\ &\leq \log E(B_p, \|\cdot\|_p, t^{1/p}) \leq cn \log t^{-1}.\end{aligned}$$

Finally, by Dudley's theorem ([2], [5]),

$$\begin{aligned}\mathbb{E} \sup_{x\in B_p} \Big| \sum_{i=1}^{K} \nu(\{i\}) g_i |x(i)|^p \Big| &\leq c \int_0^{(12n)^{1/2}K^{-1/2}} (\log E(B_p, \delta, t))^{1/2} \\ &\leq \int_{12^{1/2}K^{-1/2}}^{(12n)^{1/2}K^{-1/2}} c(p) n^{1/2} (\log n)^{1/2} K^{-1/2} t^{-1} dt \\ &\quad + c \int_0^{12^{1/2}K^{-1/2}} n^{1/2} (\log t^{-1})^{1/2} dt \\ &\leq c(p) \left(\frac{n}{K} (\log n)^3\right)^{\frac{1}{2}}.\end{aligned}$$

□

5. The proof of Theorem 1.1

By a simple approximation we may assume that X is a subspace of ℓ_p^M for some finite M. Using Theorem 2.1 and the first remark following it we may then assume that $X \subseteq L_p(\Omega, \nu)$ where $|\Omega| \leq 3M/2$, $\nu_i = \nu(\{i\}) \leq 4/M$ for all $i \in \Omega$ and X admits an orthonormal basis $f_1, \ldots, f_n$ with $\sum f_i^2 \equiv n$. We would like to find a subset $A \subset \Omega$ for which

$$\sup_{x\in X, \|x\|\leq 1} \Big| 2 \sum_{i\in A} \nu_i |x_i|^p - \sum_{i\in\omega} \nu_i |x_i|^p \Big| \leq (c(p) n M^{-1} (\log n)^3)^{1/2}. \tag{5.1}$$

Then the restriction operators to both A and the complement of A are $\left(\frac{1+(c(p)nM^{-1}(\log n)^3)^{1/2}}{1-(c(p)nM^{-1}(\log n)^3)^{1/2}}\right)^{1/p}$-isomorphisms and of course one of A or A^c is of cardinality at most $3M/4$. This reduces the dimension of the containing space from M to $3M/4$. Iterating, we get that as long as

$$\prod_{i=1}^{l} \left(\frac{1 + (c(p) n 3^i (4M)^{-i} (\log n)^3)^{1/2}}{1 - (c(p) n 3^i (4M)^{-i} (\log n)^3)^{1/2}}\right)^{1/p} \leq 1 + \varepsilon, \tag{5.2}$$

X $(1+\varepsilon)$-embeds into ℓ_p^m for some $m \le M(3/4)^{-l-1}$. Now (5.2) is satisfied for l such that $M(3/4)^{-l-1} \le C(p,\varepsilon)n(\log n)^3$ (for $C(p,\varepsilon)$ of the form $C(p)/\varepsilon^2$) which would conclude the proof if we only could establish (5.1).

To find a set A satisfying (5.1) it is enough to prove that

$$\begin{aligned}\text{Ave}_A \quad & \sup\nolimits_{x\in X,\|x\|\le 1}\left|2\sum\nolimits_{i\in A}\nu_i|x_i|^p - \sum\nolimits_{i\in\Omega}^m \nu_i|x_i|^p\right| \\ & = \mathbb{E}\sup\nolimits_{x\in X,\|x\|\le 1}\left|\sum\nolimits_{i\in\Omega}\varepsilon_i\nu_i|x_i|^p\right| \le (c(p)nM^{-1}(\log n)^3)^{1/2}.\end{aligned} \tag{5.3}$$

Here Ave_A denotes the average over all subsets of Ω while $\mathbb{E}$ is the expectation with respect to the natural product measure on $\{-1,1\}^\Omega$. The contraction principle (which is simple in this case) implies

$$\mathbb{E}\sup_{x\in X,\|x\|\le 1}\left|\sum_{i\in\Omega}\varepsilon_i\nu_i|x_i|^p\right| \le \sqrt{\frac{\pi}{2}}\,\mathbb{E}\sup_{x\in X,\|x\|\le 1}\left|\sum_{i\in\Omega}g_i\nu_i|x_i|^p\right|$$

which together with Proposition 4.1 gives (5.3).

Acknowledgements

Supported by ISF.

References

[1] J. Bourgain, J. Lindenstrauss and V. Milman, *Approximation of zonoids by zonotopes*, Acta Math. **162** (1989), 73-141.

[2] R. M. Dudley, *The sizes of compact subsets of Hilbert space and continuity of Gaussian processes*, J. Functional Analysis **1** (1967), 290–330.

[3] W. B. Johnson and G. Schechtman, *Computing p-Summing Norms With Few Vectors*, Israel Journal of Math. **87** (1994), 19-31.

[4] W. B. Johnson and G. Schechtman, *Finite dimensional subspaces of L_p*, in preparation.

[5] M. Ledoux and M. Talagrand, *Probability in Banach spaces* Ergeb. Math. Grenzgeb. **vol 23**, Springer, Berlin (1991).

[6] D. R. Lewis, *Finite dimensional subspaces of L_p*, Studia Math. **63** (1978), 207–212.

[7] D. R. Lewis, *Ellipsoids defined by Banach ideal norms*, Mathematika **26** (1979), 18–29.

[8] G. G. Lorentz and N. Tomczak-Jaegermann, *Projections of minimal norm*, Texas functional analysis seminar 1983–1984 (Austin, Tex.), 167–176, Longhorn Notes, Univ. Texas Press, Austin, TX, (1984).

[9] Ana Peña, *A note on the embedding of subspaces of L_p into $l_r^N, 0 < r < 1, r \le p < 2$*, Math. Nachr. **189** (1998), 195–207.

[10] G. Schechtman, *More on embedding subspaces of L_p in l_r^n*, Compositio Math. **61** (1987), 159-170.

[11] M. Talagrand, *Embedding subspaces of L_1 into L_1^N*, Proc. Amer. Math. Soc. **108** (1990), 363-369.

[12] M. Talagrand, *Embedding subspaces of L_p in l_p^N*, Geometric aspects of functional analysis (Israel, 1992–1994), Oper. Theory Adv. Appl. **77** (1995), 311–325.

[13] P. Wojtaszczyk, *Banach spaces for analysts*, Cambridge Studies in Advanced Mathematics **25**, Cambridge University Press, Cambridge (1991).

Department of Theoretical Mathematics
Weizmann Institute of Science
Rehovot
Israel
e-mail: gideon@wisdom.weizmann.ac.il
e-mail: tim@wisdom.weizmann.ac.il