Embedding l_p^m into l_1^n

by

WILLIAM B. JOHNSON(1) and GIDEON SCHECHTMAN(1)

Ohio State University
Columbus, OH, U.S.A.

Texas A&M University
College Station, TX, U.S.A.

Weizmann Institute of Science
Rehovot, Israel

Ohio State University
Columbus, OH, U.S.A.

1. Introduction

It is by now well-known that for any $1<s<2$, $L_1(0,1)$ contains a subspace which is isometrically isomorphic to l_s. This of course implies that for any $m=1,2,\dots$ and any $\varepsilon>0$, l_s^m is $1+\varepsilon$-isomorphic to a subspace of l_1^n if $n=n(\varepsilon,s,m)$ is sufficiently large. Theorem 1, the result of this paper, states that n can be of order m; i.e., that n can be chosen smaller than $\beta^{-1}m$ for some constant $\beta=\beta(\varepsilon,s)>0$. This complements the theorem of Figiel, Lindenstrauss and Milman [4] (cf. also [2], [3] for a somewhat weaker result) which treated the case $s=2$.

Actually the proof of Theorem 1 yields more than the above-mentioned result. First, it shows for $0<s<2$ and $0<r<s$ with $r\leqslant 1$, that for every $\varepsilon>0$, l_s^m is $1+\varepsilon$-isomorphic to a subspace of l_r^n if $m\leqslant\beta n$, where $\beta=\beta(\varepsilon,s,r)>0$ is a constant independent of n. Secondly, the condition that the range of the isomorphism be l_r^n can be relaxed. What is needed is that the range be an r-normed space which possesses a basis $(e_i)_{i=1}^n$ so that for all scalars $(b_i)_{i=1}^n$,

$$\operatorname*{Av}_{\pm}\left\|\sum_{i=1}^n \pm b_i e_i\right\| \approx \left(\sum_{i=1}^n |b_i|^r\right)^{1/r}.$$

The proof of Theorem 1, like the earlier proof of the $s=2$ case in [4], [2], [3], [5] and [11], is probabilistic in nature. A schematic outline of the usual argument specialized to the case $1<s<2$ and $r=1$ goes like this: For appropriate m and n, one defines a probability space (Ω,P) and a random linear operator or matrix $A=A_\omega$ $(\omega\in\Omega)$ from l_s^m

(1) Supported in part by NSF MCS-7903042.

into L_1^n ($=l_1^n$ with the L_1-normalization). It is important that for each $x \in l_s^m$, $\|Ax\|_1$ is, on the average, close to $|x|_s$, the norm of x in l_s^m; say,

$$|E\|Ax\|_1 - |x|_s| \leq \varepsilon |x|_s.$$

Now it is a standard fact that for small $\varepsilon>0$, if $|\|Ax\|_1-|x|_s| \leq \varepsilon$ for all x in an ε-net of the unit sphere of l_s^m, then $|\|Ax\|_1-|x|_s| \leq 3\varepsilon|x|_s$ for all $x \in l_s^m$ (see Lemma 3). Equally standard is the fact that the unit sphere of an m-dimensional normed space contains an ε-net of cardinality at most $\exp(2m\varepsilon^{-1})$ (cf. Lemma 2). Thus, in order to conclude the proof of the embedding theorem, it is sufficient to prove (and this is the main step) a distributional inequality which guarantees that for $x \in l_s^m$,

$$P[|\|Ax\|_1 - E\|Ax\|_1| \geq \varepsilon|x|_s] < \exp(-2\varepsilon^{-1}m).$$

In [2], [3] and [11] the probability space is $\{-1,1\}^{n \cdot m}$ with equal mass assigned to each point, and A_ω is defined for $x=\sum_{i=1}^m b_i e_i \in l_2^m$ and $\omega=\{\varepsilon_{i,j}\}_{i=1,j=1}^{n,m}$ by

$$A_\omega x = \sum_{i=1}^n \sum_{j=1}^m \varepsilon_{i,j} b_j e_i.$$

Notice that the entries in this random matrix are independent. Although we use a more complicated probability space and the random matrix we use does not have independent entries, the approach in [11] which uses a distributional inequality for general martingales also works in the present situation.

For the most part we use standard Banach space theory notation, as may be found in [6]. Since it is convenient for us to use the L_r-normalization in the range of the random matrix and the l_s-normalization in the domain of the random matrix, we define for

$$x = \sum_{i=1}^n \alpha_i e_i \in \mathbf{R}^n$$

(where $(e_i)_{i=1}^n$ are the unit vectors in $\mathbf{R}^n$) and for $0<p<\infty$,

$$\|x\|_p = n^{-1/p}\left(\sum_{i=1}^n |\alpha_i|^p\right)^{1/p}$$

$$|x|_p = \left(\sum_{i=1}^n |\alpha_i|^p\right)^{1/p}$$

$(\mathbf{R}^n, \|\cdot\|_p)$ is denoted by L_p^n and $(\mathbf{R}^n, |\cdot|_p)$ by l_p^n.

We also need the "weak l_p" norm on $\mathbf{R}^n$, defined for $x=\Sigma_{i=1}^n \alpha_i e_i$ by

$$|x|_{p,\infty} = \max_{1\le i\le n} \alpha_i^* i^{1/p},$$

where $(\alpha_i^*)_{i=1}^n$ is the decreasing rearrangement of $(|\alpha_i|)_{i=1}^n$. The space $(\mathbf{R}^n, |\cdot|_{p,\infty})$ is denoted by $l_{p,\infty}$.

For $0<r\le 1$, an r-norm is a non-negatively valued function, $\|\cdot\|$, on a vector space which is 0 only at 0 and satisfies the axioms $\|\alpha x\|=|\alpha|\,\|x\|$, $\|x+y\|^r\le\|x\|^r+\|y\|^r$. The most basic example of an r-norm is, of course $\|\cdot\|_r$.

If f is a measurable function on a measure space, f^* is used to denote the decreasing rearrangement of $|f|$.

Finally, we would like to thank Gilles Pisier for a discussion which yielded the present version of Theorem 1. Originally we used only Azuma's inequality, Proposition 2 (i), which led to a proof that for $1<s<2$ and for all $\varepsilon>0$, l_s^m $1+\varepsilon$-embeds into $l_{s/2}^n$ (and hence uniformly embeds into l_1^n, by Maurey's theorem [8]) as long as $m\le\alpha n/\log n$ for a certain constant $\alpha(\varepsilon, s)>0$. Pisier pointed out to us that by substituting the inequality of Proposition 2 (ii) for Azuma's in our proof, we could uniformly embed l_s^m directly into l_1^n provided $m\le\beta n$ for a certain constant $\beta=\beta(s)>0$.

2. The random matrix

Given positive integers n and m with $m\le n$, define $\Omega=\Omega(n, m)$ to be the space

$$\{-1, 1\}^{nm}\times[S(n)]^m, \tag{2.1}$$

where $S(n)$ is the symmetric group on $\{1, \dots, n\}$, and endow Ω with the probability measure P which assigns equal mass to each atom.

Given a fixed sequence $a_1\ge a_2\ge\dots\ge a_n\ge 0$, the entries of the random $n\times m$-matrix $A=A_\omega$ are defined for $1\le i\le n$, $1\le j\le m$, and for

$$\omega = ((\varepsilon_{i,j})_{i=1,j=1}^{n\ \ m}, \pi_1, \pi_2, \dots, \pi_m)\in\Omega$$

by

$$A_\omega(i,j) = \varepsilon_{i,j}\, a_{\pi_j(i)}. \tag{2.2}$$

For fixed $0<r<s<2$ with $r\le 1$, we want A_ω to be, for some $\omega\in\Omega$, a good isomorphism from l_s^m into L_r^n. Of course, the main case is $r=1$ and the reader may want to make this substitution on first reading in order to clean up messy-looking exponents and constants. In fact, from Maurey's work [8] (which is based on Rosenthal's paper [10]), it follows formally that if l_s^m embeds uniformly into L_r^n for some fixed $r<s<2$ (where $m=m(n)$ for $n=1,2,\dots$), then l_s^m embeds uniformly into L_r^n for all $r<s$. So for $1<s<2$, $r=1$ is, essentially, the general case. On the other hand, the direct embedding gives a stronger result than that which can be obtained by applying Maurey's theorem in that the embedding into L_r^n can be taken to be a $1+\varepsilon$-isomorphism and the range space need only be a "random L_r^n" space.

Notice that the columns of the random matrix A defined by (2.2) are independent, symmetric, and identically distributed, but entries in the same column of the matrix are not independent. The sequence $(a_i)_{i=1}^n$ is chosen so that the columns of A have approximately s-stable distribution; what we need is that m-independent functions with the same distribution as Ae_1, are, in L_r, $1+\varepsilon$-equivalent to the unit vector basis of l_s^m. That such a sequence $(a_i)_{i=1}^n$ exists is the content of our first lemma.

LEMMA 1. *Let $0<r<s<2$, $\varepsilon>0$, and let g be a symmetric s-stable random variable on $[0,1]$ with $\|g\|_r=1$. Then there exists $\alpha=\alpha(\varepsilon,r,s)>0$ so that for all positive integers m and n with $m\leq\alpha n$, if $y_1,y_2,\dots,y_m$ is a sequence of independent, symmetric random variables such that each $|y_i|$ $(1\leq i\leq m)$ has the same distribution as that of*

$$y=\sum_{i=1}^n a_i 1_{[(i-1)/n,\,i/n]} \tag{2.3}$$

where

$$a_i=g^*\left(\frac{i}{n}\right) \quad (1\leq i<n),$$

then for all scalars $(b_j)_{j=1}^m$ we have

$$(1-\varepsilon)\left(\sum_{j=1}^m |b_j|^s\right)^{1/s}\leq\left(E\left|\sum_{j=1}^m b_j y_j\right|^r\right)^{1/r}\leq(1+\varepsilon)\left(\sum_{j=1}^m |b_j|^s\right)^{1/s} \tag{2.5}$$

Proof. We first show that if $r\geq 1$ or if $s\leq 1$ and $r>s/(s+1)$ then $E|g^*-y|^r\leq Kn^{r/s-1}$, where $K=K(r,s)$ depends on r and s only. For $1\leq r<s$ we use the fact (cf. [7] or [14]) that for all $t>0$

$$P(|g|\geq t)\leq Ct^{-s} \tag{2.6}$$

for some constant $C=C(r,s)$ to get

$$\begin{aligned}E|g^*-y|^r\leq E(g^{*r}-y^r)&\leq\int_0^{1/n} g^{*r}\\ &\leq C^{r/s}\int_0^{1/n} t^{-r/s}\,dt=C^{r/s}(1-r/s)^{-1}n^{r/s-1}.\end{aligned}$$

For $s\leqslant 1$, $s/(s+1)<r<s$ we use the fact (cf. [14], p. 54) that the density function p of a symmetric s-stable random variable normalized in L_r satisfies the inequality

$$p(t)\geqslant D^{-1}t^{-s-1} \quad \text{for } |t|\geqslant D$$

for some constant $D=D(r,s)$.

Let

$$F(u)=P(|g|\geqslant u),$$

then $g^*=F^{-1}$ and $g^{*\prime}(t)=-1/p(g^*(t))$, $0\leqslant t\leqslant 1$. Thus $|g^{*\prime}(t)|\leqslant D(g^*(t))^{s+1}\leqslant Dc^{(s+1)/s}t^{-(s+1)/s}$ as long as $g^*(t)\geqslant D$ and

$$\begin{aligned} E|g^*-y|^r &\leqslant \int_0^{1/n} g^*(t)^r+\int_{1/n}^{F(D)-1/n}\left(g^*(t)-g^*\left(t+\frac{1}{n}\right)\right)^r+\int_{F(D)-1/n}^{1}\left(g^*(t)-g^*\left(t+\frac{1}{n}\right)\right)^r \\ &\leqslant C^{r/s}\left(1-\frac{r}{s}\right)^{-1}n^{r/s-1}+n^{-r}D^rC^{r(s+1)/s}\int_{1/n}^{\infty}t^{-r(s+1)/s}+\left(\int_{F(D)-1/n}^{1}g^*(t)-g^*\left(t+\frac{1}{n}\right)\right)^r \\ &\leqslant C^{r/s}\left(1-\frac{r}{s}\right)^{-1}n^{r/s-1}+D^rC^{r(s+1)/s}(r(s+1)/s-1)^{-1}n^{r/s-1}+g^*\left(F(D)-\frac{1}{n}\right)n^{-r} \\ &\leqslant Kn^{r/s-1}. \end{aligned}$$

Let $g_1, g_2, \dots, g_m$ be independent, symmetric s-stables with $\|g_i\|_r=1$ and set for $1\leqslant j\leqslant m$

$$z_j=\sum_{i=1}^{n} a_i \operatorname{sign}(g_j)\, 1_{[a_i<|g_j|<a_{i-1}]}.$$

Now the z_j's have the same distribution as the y_j's and, by the fact that $(g_j-z_j)_{j=1}^m$ forms a 1-unconditional basic sequence in L_r and by Hölder's inequality, we get that if $1\leqslant r<s<2$ or $s/(s+1)<r<s\leqslant 1$,

$$\begin{aligned} E\left|\sum_{j=1}^m b_j(g_j-z_j)\right|^r &\leqslant \sum_{j=1}^m |b_j|^r E|g_j-z_j|^r \\ &=\sum_{j=1}^m |b_j|^r E|g^*-y|^r\leqslant Kn^{r/s-1}\sum_{j=1}^m |b_j|^r \\ &\leqslant K\left(\frac{m}{n}\right)^{(s-r)/s}\left(\sum_{j=1}^m |b_j|^s\right)^{r/s}. \end{aligned}$$

So for all $\delta>0$ there exists an $\alpha>0$ such that if $m/n<\alpha$

$$\left(E\left|\sum_{j=1}^m b_j(g_j-z_j)\right|^r\right)^{1/r}\leqslant \delta\left(\sum |b_j|^s\right)^{1/s}.$$

The monotonicity of the function $\varphi(r)=(E|f|^r)^{1/r}$ implies that the same conclusion holds for all $r, s, 0<r<s<2$. The conclusion of the lemma follows now from the fact that

$$E\left|\sum_{j=1}^{m} b_j g_j\right|^r = \left(\sum_{j=1}^{m} |b_j|^s\right)^{r/s}. \qquad \text{Q.E.D.}$$

Returning now to the random matrix A defined by (2.2) and (2.4), we check that $E\|Ax\|_r \approx |x|_s$ for all $x \in l_s^m$. So suppose $\alpha=\alpha(\varepsilon, r, s)>0$ satisfies the conclusion of Lemma 1 and $m \leqslant \alpha n$ and fix $x=\Sigma_{j=1}^m b_j e_j \in l_s^m$. Then

$$E\|Ax\|_r^r = En^{-1}\sum_{i=1}^{n}\left|\sum_{j=1}^{m} b_j \varepsilon_{i,j} a_{\pi_j(i)}\right|^r = E\left|\sum_{j=1}^{m} b_j \varepsilon_{1,j} a_{\pi_j(1)}\right|^r.$$

The random variables $\varepsilon_{1,j} a_{\pi_j(1)}$ $(1 \leqslant j \leqslant m)$ on Ω are independent and symmetric; moreover, the common distribution of their absolute values is the same as that of "y" in (2.3), thus from Lemma 1 we have

$$(1-\varepsilon)^r |x|_s^r \leqslant E\|Ax\|_r^r \leqslant (1+\varepsilon)^r |x|_s^r. \tag{2.7}$$

We now state a distributional inequality, to be proved in Section 2, which allows us to select an $\omega \in \Omega$ for which $\|A_\omega x\|_r \approx |x|_s$ for all $x \in l_s^m$. (In the notation of Proposition 1, set $x=(b_j)_{j=1}^m$ and $\|\cdot\|=\|\cdot\|_r$. Then for $\omega=(\varepsilon, \pi) \in \Omega$,

$$\|A_\omega x\|_r^r = \left\|\sum_{j=1}^{m} b_j \sum_{i=1}^{n} \varepsilon_{i,j} a_{\pi_j(i)} e_i\right\|_r^r,$$

which is not the same as $f(\omega)$. Of course, f and $\|Ax\|_r^r$ have the same distribution, which is all that we need. We state Proposition 1 for f rather than for $\|Ax\|_r^r$ in order to simplify notation in its proof.)

PROPOSITION 1. *Let $0<r\leqslant 1<p<2$, let m and n be positive integers, and let Ω be given by* (2.1). *Suppose that $\|\cdot\|$ is an r-norm on $\mathbf{R}^n$, $(b_j)_{j=1}^m \in \mathbf{R}^m$ and $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n \geqslant 0$. Define for $\omega=(\varepsilon, \pi) \in \Omega$*

$$f(\omega) = \left\|\sum_{j=1}^{m} b_j \sum_{i=1}^{n} \varepsilon_{i,j} a_i e_{\pi_j(i)}\right\|^r.$$

Then for all $t>0$

$$P[|f-Ef| \geqslant t] \leqslant 2\exp\left[-4^{-q}\delta_p t^q |(a_i b_j)_{i=1,j=1}^{n}{}^{m}|_{rp,\infty}^{-rq} \cdot \max_{1\leqslant k\leqslant n} \|e_k\|^{-rq}\right] \tag{2.8}$$

where $p^{-1}+q^{-1}=1$ and $\delta_p>0$ depends only on p.

To apply Proposition 1, we need two standard lemmas involving δ-nets of the unit sphere. By a δ-net of a subset, U, of an r-normed space $(X, \|\cdot\|)$ we mean a subset M of U such that for all $x\in U$,

$$\inf_{y\in M}\|x-y\|^r\leqslant\delta.$$

LEMMA 2. *Let $(X, \|\cdot\|)$ be an r-normed space $(0<r\leqslant 1)$ of dimension m and suppose that $0<\delta$. Then the unit sphere of X contains a δ-net of cardinality at most* $\exp(2r^{-1}\delta^{-1}m)$.

Proof. Let M be a subset of the unit sphere of X maximal with respect to "$\|x-y\|^r\geqslant\delta$ for all distinct points x, y in M". M is obviously a δ-net of the unit sphere and the open balls (relative to the metric $\|x-y\|^r$) of radius $\delta/2$ around the points of M are pairwise disjoint and contained in $B(1+\delta/2)$, the open ball around the origin of radius $1+\delta/2$. Consequently,

$$\operatorname{card} M\cdot\operatorname{vol} B(\delta/2)\leqslant\operatorname{vol} B(1+\delta/2).$$

Since $B(ts)=t^{1/r}B(s)$ for any $s, t>0$ and $\dim X=m$, we conclude

$$\operatorname{card} M\leqslant[2(1+\delta/2)/\delta]^{m/r}=(1+2/\delta)^{m/r}\leqslant\exp(2r^{-1}\delta^{-1}m).\qquad\text{Q.E.D.}$$

LEMMA 3. *Suppose that $(X, \|\cdot\|)$ is an s-normed space $(0<s\leqslant 1)$, $(Y, \|\cdot\|)$ is an r-normed space $(0<r\leqslant 1)$ and $T: X\to Y$ is a continuous linear operator. Suppose that $0<\varepsilon, \delta<1$ are such that for some $\delta^{s/r}$-net, M, of the unit sphere of X and all $x\in M$ we have*

$$1-\varepsilon\leqslant|Tx|^r\leqslant 1+\varepsilon.$$

Then for all x in the unit sphere of X we have

$$\frac{1-2\delta-\varepsilon}{1-\delta}\leqslant|Tx|^r\leqslant\frac{1+\delta}{1-\delta}(1+\varepsilon).\tag{2.9}$$

Proof. Given x in the unit sphere of X, write

$$x=x_0+\sum_{n=1}^{\infty}a_n x_n$$

with $(x_n)_{n=0}^{\infty}\subseteq M$ and $0\leqslant a_n^s\leqslant\delta^{sn/r}$ for $n=1,2,\ldots$. Then

$$\big||Tx|^r-|Tx_0|^r\big|\leqslant|Tx-Tx_0|^r=\left|\sum_{n=1}^{\infty}a_n Tx_n\right|^r\leqslant\sum_{n=1}^{\infty}|a_n|^r|Tx_n|^r\leqslant\frac{\delta}{1-\delta}(1+\varepsilon).$$

A trivial computation now yields the desired conclusion. Q.E.D.

We can now prove the main result.

THEOREM 1. *Let $\tau>0$, and suppose that $0<r<s<2$ with $r\leq 1$. Then there exists $\beta=\beta(\tau, r, s)>0$ so that if m and n are positive integers with $m\leq\beta n$, then l_s^m is $1+\tau$-isomorphic to a subspace of L_r^n.*

Proof. For a value of $\varepsilon=\varepsilon(\tau, r)>0$ to be specified later, we take $\alpha=\alpha(\varepsilon, r, s)$ from Lemma 1 and let $0<\beta\leq\alpha$ be such that β also satisfies another numerical inequality which comes up later. Now fix $m\leq\beta n$ and let g, $(a_i)_{i=1}^n$, Ω, and A be as in Lemma 1, so that for all $x\in l_s^m$,

$$(1-\varepsilon)\,|x|_s^r\leq E\|Ax\|_r^r\leq(1+\varepsilon)\,|x|_s^r. \tag{2.10}$$

Now fix any $x=\sum_{i=1}^m b_i e_i\in l_s^m$ with $|x|_s=1$. Recalling the distributional inequality for s-stable variables mentioned at the beginning of the proof of Lemma 1, we see that the a_i's defined by (2.4) satisfy

$$a_i\leq C^{1/s}n^{1/s}i^{-1/s}$$

For some constant $C=C(r, s)$. Using this and the easy observation that

$$\left|\sum x_j\right|_{s,\infty}^s\leq\sum|x_j|_{s,\infty}^s$$

if the x_j's are disjointly supported vectors in $l_{s,\infty}$, we get

$$|(a_i b_j)_{i=1,j=1}^{n\ \ m}|_{s,\infty}^s\leq\sum_{j=1}^m|b_j|_{s,\infty}^s|(a_i)_{i=1}^n|_{s,\infty}^s\leq Cn|(i^{-1/s})_{i=1}^n|_{s,\infty}^s=Cn.$$

Assume that $r>s/2$ and set $p=s/r$ so that $1<p<2$. Applying Proposition 1 we get for any $x\in l_s^m$, $|x|_s=1$,

$$\begin{aligned}P[|\|Ax\|_r^r-E\|Ax\|_r^r|\geq\varepsilon]&\leq 2\exp(-\delta_p\,\varepsilon^q C^{-rq/s}n^{-rq(1/s-1/r)})\\&=2\exp(-\delta_p\,\varepsilon^q C^{r/(r-s)}n)\end{aligned}$$

so that (since $m\leq\beta n\leq\alpha n$; α from Lemma 1)

$$P[1-2\varepsilon\leq\|Ax\|_r^r\leq 1+2\varepsilon]\geq 1-2\exp(-\delta_p\,\varepsilon^q C^{r/(r-s)}n).$$

Using Lemma 2, pick an $\varepsilon^{s/r}$-net of the unit sphere of l_s^m with card$M\leq\exp(2r^{-1}\varepsilon^{-s/r}m)$. Then

$$P[1-2\varepsilon \leq \|Ax\|_r^r \leq 1+2\varepsilon \quad \text{for all } x \in M] \geq 1-2\exp(2r^{-1}\varepsilon^{-s/r}m-\delta_p \varepsilon^q C^{r/(r-s)}n)$$

hence, by Lemma 3,

$$P\left[\frac{1-4\varepsilon}{1-\varepsilon}|x|_s^r \leq \|Ax\|_r^r \leq \frac{(1+\varepsilon)(1+2\varepsilon)}{1-\varepsilon}|x|_s^r \quad \text{for all } x \in l_s^m\right]$$
$$\geq 1-2\exp\left[(2r^{-1}\varepsilon^{-s/r}\beta-\delta_p\varepsilon^q C^{r/(r-s)})\,n\right].$$

Thus if we choose $\varepsilon=\varepsilon(\tau, r)>0$ and $\beta=\beta(\varepsilon, r, s)>0$ sufficiently small, we get for $m\leqslant\beta n$ that

$$P[(1-\tau)\,|x|_s \leq \|Ax\|_r \leqslant (1+\tau)\,|x|_s \quad \text{for all } x \in l_s^m] > \frac{1}{2}. \tag{2.11}$$

This completes the proof in the case $r>s/2$.

The general case follows formally from the case $r>s/2$ by iteration. A more elegant way to finish (which yields a better estimate for β) is to use (2.11) for two different values r_1, r_2 with $s/2<r_1, r_2<s$ to select an $\omega \in \Omega$ so that A_ω is simultaneously a good isomorphism form l_s^m into $L_{r_1}^n$ and into $L_{r_2}^n$ and use a standard extrapolation argument to conclude that A_ω is also a good isomorphism from l_s^m into L_r^n. Q.E.D.

Remarks. (1). It follows from a result of Maurey's [8] and Theorem 1 that for $0<r<s<2$ and $m\leqslant\beta n$, l_s^m is $K(r, s)$-isomorphic to a subspace of l_r^n, but we do not know whether $K(r, s)$ can be taken close to one when $r>1$.

(2) As is easily seen from the proof, the assumption in Theorem 1 that the range space is L_r^n can be relaxed a bit. It is enough to assume that the range is an r-normed space which contains vectors $(e_i)_{i=1}^n$ so that for all $(b_i)_{i=1}^n$,

$$\operatorname*{Av}_{\pm}\left\|\sum_{i=1}^n \pm b_i e_i\right\|^r = \sum_{i=1}^n |b_i|^r. \tag{2.12}$$

This perhaps explains why our proof breaks down when r approaches 2 (i.e. for r>1), because (2.12) is true for $r=2$ if the e_i's are all the same unit vector in any Banach space.

3. The distributional inequality

The main tool for proving Proposition 1 is a martingale inequality which, along with its proof, was communicated to the authors by Gilles Pisier (part (ii) of Proposition 2). This inequality is in turn a consequence of Azuma's martingale inequality (part (i) of

Proposition 2) [1], [13]. Versions of Azuma's inequality have previously been used in Banach space theory [9], [11], [12].

PROPOSITION 2. *Let* $(d_k)_{k=1}^n$ *be a uniformly bounded martingale difference sequence* (i.e., $(\Sigma_{i=1}^k d_i)_{k=1}^n$ *is an* L_∞*-bounded martingle which has mean zero*).

(i) (Azuma) *For all* $t>0$,

$$P\left[\left|\sum_{k=1}^n d_k\right| \geq t\right] \leq 2\exp\left[-t^2 \Big/ 4\sum_{k=1}^n \|d_k\|_\infty^2\right].$$

(ii) *For all* $1<p<2$ *and all* $t>0$,

$$P\left[\left|\sum_{k=1}^n d_k\right| \geq t\right] \leq 2\exp\left[-\delta_p t^q / |(\|d_k\|_\infty)_{k=1}^n|_{p,\infty}^q\right]$$

where $1/p+1/q=1$ *and* $\delta_p=(2-p)/8p(q+1)^q$.

Proof. (i) Let E_i $(1\leq i\leq n)$ be the conditional expectation with respect to the sigma field generated by $d_1, d_2, \ldots, d_i$, so that $E_i d_j=0$ for $1\leq i<j\leq n$. Given any real λ, we have

$$\begin{aligned} E\exp\left(\lambda\sum_{i=1}^n d_i\right) &= EE_{n-1}\exp\left(\lambda\sum_{i=1}^n d_i\right) \\ &= E\exp\left(\lambda\sum_{i=1}^{n-1} d_i\right)E_{n-1}\exp(\lambda d_n) \\ &\leq E\exp\left(\lambda\sum_{i=1}^{n-1} d_i\right)E_{n-1}[\lambda d_n+\exp(\lambda^2 d_n^2)] \quad (\text{since } e^x\leq x+e^{x^2}) \\ &= E\exp\left(\lambda\sum_{i=1}^{n-1} d_i\right)E_{n-1}\exp(\lambda^2 d_n^2) \\ &\leq \exp\lambda^2\|d_n\|_\infty^2\, E\exp\left(\lambda\sum_{i=1}^{n-1} d_i\right). \end{aligned}$$

By iterating the above we obtain

$$E\exp\left(\lambda\sum_{i=1}^{n-1} d_i\right) \leq \exp\left(\lambda^2\sum_{i=1}^n \|d_i\|_\infty^2\right).$$

Hence for all $t>0$,

$$P\left[\sum_{i=1}^n d_i \geqslant t\right] = P\left[\exp\left(\lambda \sum_{i=1}^n d_i\right) \geqslant e^{\lambda t}\right]$$

$$\leqslant e^{-\lambda t} E \exp\left(\lambda \sum_{i=1}^n d_i\right) \leqslant \exp\left(\lambda^2 \sum_{i=1}^n \|d_i\|_\infty^2 - \lambda t\right).$$

Setting $\lambda = t/(2\sum_{i=1}^n \|d_i\|_\infty^2)$ we get

$$P\left[\sum_{i=1}^n d_i \geqslant t\right] \leqslant \exp\left[-t^2 \Big/ \left(2\sum_{i=1}^n \|d_i\|_\infty^2\right)\right].$$

Since also

$$P\left[-\sum_{i=1}^n d_i \geqslant t\right] \leqslant \exp\left[-t^2 \Big/ \left(2\sum_{i=1}^n \|d_i\|_\infty^2\right)\right]$$

we get the desired result.

(ii) Assume, without loss of generality, that

$$|(\|d_k\|_\infty)_{k=1}^n|_{p,\infty} = 1$$

and choose a permutation π of $\{1,\dots,n\}$ so that

$$\|d_{\pi(k)}\|_\infty = \|d_k\|^* \quad (1 \leqslant k \leqslant n).$$

Thus we have for $k=1,2,\dots,n$,

$$\|d_{\pi(k)}\|_\infty \leqslant k^{-1/p}.$$

Given an integer $N \leqslant n$ we have

$$P\left[\left|\sum_{k=1}^n d_k\right| \geqslant (q+1)N^{1/q}\right] \leqslant P\left[\left|\sum_{k=1}^N d_{\pi(k)}\right| \geqslant qN^{1/q}\right] + P\left[\left|\sum_{k=N+1}^n d_{\pi(k)}\right| \geqslant N^{1/q}\right].$$

But

$$\left|\sum_{k=1}^N d_{\pi(k)}\right| \leqslant \sum_{k=1}^N \|d_{\pi(k)}\|_\infty \leqslant \sum_{k=1}^N k^{-1/p} < qN^{1/q},$$

so we get by Proposition 2 (i),

$$P\left[\left|\sum_{k=1}^{n} d_k\right| \geq (q+1)N^{1/q}\right] \leq 2\exp\left[-N^{2/q}\Big/\left(4\sum_{k=N+1}^{n}\|d_{\pi(k)}\|_\infty^2\right)\right]$$

$$\leq 2\exp[N^{2/q}(1-2/p)/(4N^{(1-2/p)}]$$

$$= 2\exp[-(2-p)N/4p].$$

If $t \geqslant q+1$, set

$$N = \left[\left(\frac{t}{q+1}\right)^q\right],$$

so that

$$1 \leqslant N \leqslant \left(\frac{t}{q+1}\right)^q \leqslant 2N.$$

Then

$$P\left[\left|\sum_{k=1}^{n} d_k\right| \geqslant t\right] \leqslant P\left[\left|\sum_{k=1}^{n} d_k\right| \geqslant (q+1)N^{1/q}\right] \leqslant 2\exp[-(2-p)N/4p]$$

$$\leqslant 2\exp[-(2-p)t^q/8p(q+1)^q].$$

If $t \leqslant q+1$, then

$$2\exp[-(2-p)t^q/8p(q+1)^q] \geqslant 2\exp[-(2-p)/8p] \geqslant 2e^{-1/8} > 1. \qquad \text{Q.E.D.}$$

We turn to

Proof of Proposition 1. For the convenience of the reader, we recall that

$$\Omega = \{-1, 1\}^{n\cdot m} \times (S(n))^m$$

and for $\omega = (\varepsilon, \pi) \in \Omega$, we define

$$f(\varepsilon, \pi) = \left\|\sum_{j=1}^{m} b_j \sum_{i=1}^{n} \varepsilon_{i,j}\, a_i\, e_{\pi_j(i)}\right\|^r,$$

where $\|\cdot\|$ is an r-norm on $\mathbf{R}^n$, the b_j's are reals,

$$a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n \geqslant 0, \quad 0 < r \leqslant 1 < p < 2, \quad \text{and} \quad 1/p + 1/q = 1.$$

In order to apply Proposition 2 we need to define a martingale difference sequence which sums to $f - Ef$.

Set $L=\{1,\dots,n\}\times\{1,\dots,m\}$ and linearly order $\{0\}\cup L$ by taking 0 as the first element and using the Hebrew dictionary order on L; i.e.,

$$0<(1,1)<(2,1)<\dots<(n,1)<(1,2)<\dots<(n,2)<(1,3)<\dots.$$

We let $\mathcal{F}_0=\{\varnothing,\Omega\}$ and for $(i,j)\in L$ we define a sigma field on Ω by saying that an atom of $\mathcal{F}_{(i,j)}$ is determined by specifying the values of $\varepsilon_{l,k}$ and $\pi_k(l)$ for all $(l,k)\leq(i,j)$. Then $\{\mathcal{F}_t: t\in\{0\}\cup L\}$ is an increasing sequence of sigma fields; the first field is trivial and the last is the collection of all subsets of Ω.

For $(i,j)\in L$ let $(i,j)'$ be the immediate predecessor of (i,j) in $\{0\}\cup L$ and define

$$d_{i,j}=E(f\mid\mathcal{F}_{(i,j)})-E(f\mid\mathcal{F}_{(i,j)'})$$

so that $(d_{(i,j)})_{(i,j)\in L}$ is a martingale difference sequence which sums to $f-Ef$. Thus the conclusion (1.8) of Proposition 1 is an immediate consequence of Proposition 2 (ii) and the following inequality, valid for all $(i,j)\in L$:

$$\|d_{(i,j)}\|_\infty\leq 4|a_i b_j|^r \max_{1\leq k\leq n}\|e_k\|^r. \tag{3.1}$$

For any fixed $(i,j)\in L$, fix any atom A in $\mathcal{F}_{(i,j)'}$ and let $\mathcal{A}$ be the collection of all atoms in $\mathcal{F}_{(i,j)}$ which are contained in A. On A, $E(f\mid\mathcal{F}_{(i,j)'})$ is the average value of f on A, and if B is an atom of $\mathcal{F}_{(i,j)}$, then $E(f\mid\mathcal{F}_{(i,j)})$ is on B the average value of f on B. Thus (3.1) will follow once we check that for all $B, C\in\mathcal{A}$

$$\left|\operatorname*{Av}_{\omega\in B} f(\omega)-\operatorname*{Av}_{\omega\in C} f(\omega)\right|\leq 4|a_i b_j|^r\max_{1\leq k\leq n}\|e_k\|^r.$$

So fix $B, C\in\mathcal{A}$. Since B and C are both contained in the same atom of $\mathcal{F}_{(i,j)'}$, we have that the values of $\varepsilon_{u,v}$ and $\pi_v(u)$ are specified and equal on B and C for all $(u,v)<(i,j)$. Let us say that on B, $\varepsilon_{i,j}$ and $\pi_j(i)$ are specified by

$$\varepsilon_{i,j}=\varepsilon_B,\quad \pi_j(i)=s$$

while on C, $\varepsilon_{i,j}$ and $\pi_j(i)$ are specified by

$$\varepsilon_{i,j}=\varepsilon_C,\quad \pi_j(i)=t.$$

We define a one to one correspondence from B onto C by defining $(\varepsilon,\pi)\to(\varepsilon^*,\pi^*)$, where

$$\varepsilon^*_{u,v} = \begin{cases} \varepsilon_{u,v}, & \text{if } (u,v) \neq (i,j) \\ \varepsilon_C, & \text{if } (u,v) = (i,j) \end{cases}$$

$$\pi^*_w(y) = \begin{cases} t, & \text{if } (y,w) = (i,j) \\ s, & \text{if } w = j \text{ and } \pi_w(y) = t \\ \pi_w(y), & \text{otherwise.} \end{cases}$$

Given $(\varepsilon, \pi) \in B$, let z be the unique number in $\{1, \ldots, n\}$ such that $\pi_j(z) = t$. If $t = s$ then of course $z = i$. If $t \neq s$ then $z > i$ because $\pi_j(y) = \pi^*_j(y)$ for all $y < i$ and $t = \pi^*_j(i)$. Thus $|a_i| \geq |a_z|$ since $a_1 \geq a_2 \geq \ldots \geq a_n \geq 0$ and we have by the triangle inequality,

$$|f(\varepsilon, \pi) - f(\varepsilon^*, \pi^*)| \leq \left\| \sum_{w=1}^{m} \sum_{y=1}^{n} b_w \varepsilon_{y,w} a_y e_{\pi_w(y)} - b_w \varepsilon^*_{y,w} a_y e_{\pi^*_w(y)} \right\|^r$$

$$= \|b_j \varepsilon_B a_i e_s - b_j \varepsilon_C a_i e_t + b_j \varepsilon_{z,j} a_z e_t - b_j \varepsilon_{z,j} a_z e_s\|^r$$

$$\leq 2|b_j|^r (|a_i| + |a_z|)^r \max_{1 \leq k \leq n} \|e_k\|^r$$

$$\leq 2^{1+r} |b_j \cdot a_i|^r \max_{1 \leq k \leq n} \|e_k\|^r.$$

The inequality (3.8) now follows by averaging over (ε, π) in B. Q.E.D.

References

[1] AZUMA, K., Weighted sums of certain dependent random variables. *Tôhoku Math J.*, 19 (1967), 357–367.

[2] BENNETT, G., DOR, L. E., GOODMAN, V., JOHNSON, W. B. & NEWMAN, C. M., On uncomplemented subspaces of L_p, $1<p<2$. *Israel J. Math.*, 26 (1977), 178–187.

[3] BENNETT, G., GOODMAN, V. & NEWMAN, C. M., Norms of random matrices. *Pacific J. Math.*, 59 (1975), 359–365.

[4] FIGIEL, T., LINDENSTRAUSS, J. & MILMAN, V., The dimension of almost spherical sections of convex sets. *Acta Math.*, 139 (1977), 53–94.

[5] KASHIN, B. S., Diameter of some finite dimensional sets and of some classes of smooth functions. *Izv. Akad. Nauk SSSR, Ser. Mat.*, 41 (1977), 334–351 (russian).

[6] LINDENSTRAUSS, J. & TZAFRIRI, L., *Classical Banach Spaces*, I: *Sequence Spaces*. Springer-Verlag, 1977.

[7] LOÉVE, M., *Probability Theory* I. Springer-Verlag, 1977.

[8] MAUREY, B., *Théorémes de factorisation pour les opérateurs linéaires á valeurs dans les espaces L^p*. Astérisque No. 11, Société Math. de France, 1974.

[9] — Construction de suites symetriques. *C.R. Acad. Sci. Paris*, 228 (1979), 679–681.

[10] ROSENTHAL, H. P., On subspaces of L^p. *Ann. of Math.*, 97 (1973), 344–373.

[11] SCHECHTMAN, G., Random embeddings of euclidean spaces in sequence spaces. *Israel J. Math.*, 40 (1981), 187–192.

[12] — Levy type inequality for a class of finite metric spaces. *Proceedings of the NSF CBMS Conference on Martingale Theory in Harmonic Analysis and Banach Spaces* 1981. To appear.
[13] STOUT, W. F., *Almost Sure Convergence*. Academic Press, 1974.
[14] IBRAGIMOV, I. A. & LINNIK, YU, V., *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing, Groningen, 1971.

Received December 14, 1981