PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 108, Number 2, February 1990

EMBEDDING SUBSPACES OF L_1 INTO l_1^N

MICHEL TALAGRAND

(Communicated by William J. Davis)

ABSTRACT. We simplify techniques of Schechtman, Bourgain, Lindenstrauss, Milman, to prove the following. If X is an n-dimensional subspace of L_1, there exists a subspace Y of l_1^N such that $d(X, Y) \leq 1+\varepsilon$ whenever $N \geq CK(X)^2\varepsilon^{-2}n$, where $K(X)$ is the K-convexity constant of X, and where C is a universal constant.

1. INTRODUCTION

Consider the following problem.

(*) Given an n-dimensional subspace X of $L_1 = L_1(0, 1)$ and $\varepsilon > 0$, what is the smallest $N = N(X, \varepsilon)$ such that there is a subspace Y of l_1^N with $d(X, Y) \leq 1+\varepsilon$?

This question was considered in [3] and [4] for $X = l_2^n$, and in [6] for $X = l_p^n$. A breakthrough was made in [13] by G. Schechtman, who, using empirical distributions proved that for every subspace X of $L_1(0, 1)$ of dimension n, and every $0 < \varepsilon < \frac{1}{2}$,

$$N(X, \varepsilon) \leq C\varepsilon^{-2} \log(\varepsilon^{-1})n^2. \tag{1}$$

Here, as in the sequel, C denotes a universal constant, that may vary at each occurrence. Schechtman's method was refined, and combined with facts from Banach space theory by Bourgain, Lindenstrauss, Milman [1] to yield a bound of $N(X, \varepsilon)$ that is nearly linear in n. In particular, they proved that for any n-dimensional subspace X of L_1,

$$N(X, \varepsilon) \leq C\varepsilon^{-2} \log(n\varepsilon^{-1})(\log n)^2 n. \tag{2}$$

Moreover, they showed that if $1 < p \leq 2$ and if $T_p(X)$ denote the type p constant of X, we have, for any $\tau > 0$

(3)

$$N(X, \varepsilon) \leq c(\tau)n\varepsilon^{-2}(\log(T_p(X)+1))^{1/2}(p-1)^{-3-\tau}(\log(T_p(X)(\varepsilon(p-1)))^{5/2+\tau}.$$

Received by the editors September 12, 1988.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 46B25, 46E30; Secondary 60G15.
Work partially supported by an National Science Foundation Grant DMS 880180.

We denote by $K(X)$ the K-convexity constant of X (see [11]), and we prove the following, where ε_0 is a universal constant.

Theorem. *For any n-dimensional subspace X of L^1, and $0 < \varepsilon \leq \varepsilon_0$, we have*

$$N(X, \varepsilon) \leq CK(X)^2 \varepsilon^{-2} n. \tag{4}$$

It has been proved by Pisier [12] that for

$$X \subset L^1, \quad n = \dim X,$$

we have

$$K(X) \leq C(\log n)^{1/2}.$$

Thus (4) improves on (2). It is proved in [1] that

$$K(X)^2 \leq C \log(T_p(X) + 1)/(p-1),$$

and thus (4) improves on(3). We do not know if the factor $K(X)$ is necessary in (4).

2. The random choice procedure

Central to our approach is a random choice argument, in spirit close to the empirical distribution method, but which avoids many of the technical complications of this method. Consider a subspace X of dimension n of l_1^M. To embed X into $l_1^{M'}$ where M' is of order $M/2$ we will, for each coordinate, flip a coin and disregard the coordinate if "head" comes up. More formally, consider a sequence $\varepsilon = (\varepsilon_i)_{i \leq M}$ of Bernoulli (or Rademacher) random variables. That is, the sequence is independent identically distributed (i.i.d.) and $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = \frac{1}{2}$. Consider the random diagonal operator $U_\varepsilon \colon l_1^M \to l_1^M$ given by $U_\varepsilon((x_i)_{i \leq M}) = ((1 + \varepsilon_i)x_i)_{i \leq M}$. We will give conditions under which U_ε is, with large probability, almost an isometry when restricted to X. We note that $U_\varepsilon(l_1^N)$ is isometric to $l_i^{M'}$, where $M' = \operatorname{card}\{i \leq M ; \varepsilon_i = 1\}$, and, with probability $\geq \frac{1}{2}$, we have $M' \leq M/2$.

For $x \in l_1^M$, consider the random variable $Z_x = \|U_\varepsilon(x)\| - \|x\|$, and let $A = \sup_{x \in X, \|x\|_1 \leq 1} |Z_x|$. The restriction T_ε of U_ε to X satisfies

$$\|T_\varepsilon\| \leq 1 + A, \quad \|T_\varepsilon^{-1}\| \leq 1/(1-A),$$

so that when $A \leq \frac{1}{2}$ we have $d(X, U_\varepsilon(X)) \leq 1 + 3A$.

For $x = (x_i)_{i \leq M}$, we have

$$\begin{aligned} Z_x &= \sum_{i \leq M} |(1 + \varepsilon_i)x_i| - \sum_{i \leq M} |x_i| \\ &= \sum_{i \leq M} (1 + \varepsilon_i)|x_i| - \sum_{i \leq M} |x_i| = \sum_{i \leq M} \varepsilon_i |x_i| \end{aligned}$$

so that

$$A = \sup_{x \in X, \|x\|_1 \le 1} \left| \sum_{i \le M} \varepsilon_i |x_i| \right| .$$

It follows by the comparison theorem for Rademacher processes ([7], Proposition 1) that

$$EA \le 2E \sup_{x \in X, \|x\|_1 \le 1} \left| \sum_{i \le M} \varepsilon_i x_i \right| . \tag{5}$$

Consider a sequence $(g_i)_{i \le N}$ of i.i.d. $N(0,1)$ random variables. Then by the "contraction principle" as e.g. in [5], we have

$$EA \le CE \sup_{x \in X, \|x\|_1 \le 1} \left| \sum_{i \le M} g_i x_i \right| . \tag{6}$$

Since $P(A \le 3EA) \ge \frac{2}{3}$ and

$$P(\operatorname{card} M' \le \tfrac{1}{2} \operatorname{card} M) \ge \tfrac{1}{2},$$

we have shown the following.

Proposition 1. *There exists two universal constants α, C, with the following property. If X is a subspace of l_1^M and*

$$H := E \sup_{x \in X, \|x\|_1 \le 1} \left| \sum_{i \le M} g_i x_i \right| < \alpha$$

then there exists $M' \le M/2$ and a subspace Y of $l_1^{M'}$ such that $d(X, Y) \le 1 + CH$.

Remark. Instead of using the comparison theorem for Rademacher processes (see [8] for a very simple proof) one can use above the usual comparison results for Gaussian processes, which, however, lie quite deeper.

3. Proof of the theorem

In order to successfully use Proposition 1, one must ensure that

$$H = E \sup_{x \in X, \|x\|_1 \le 1} \left| \sum_{i \le M} g_i x_i \right|$$

is small. Let us denote by $\langle \cdot , \cdot \rangle$ the scalar product of l_2^M. We have

$$H = \int \sup_{x \in X, \|x\|_1 \le 1} \langle x, y \rangle \, d\gamma_M(y)$$

where γ_M is the law of $\sum_{i\leq M} g_i e_i$, and where $(e_i)_{i\leq M}$ is the canonical basis of l_2^M. Since γ_M is the law of $\sum_{i\leq M} g_i y_i$ for any orthonormal basis l_2^M, we see that if $y_1, \dots, y_n$ is an orthonormal basis of $X \subset l_2^M$ we have

$$H = E \sup_{x\in X, \|x\|_1\leq 1} \left\langle x, \sum_{i\leq n} g_i y_i \right\rangle .$$

The following proposition is essentially a juxtaposition of a lemma of Lewis [9] and a lemma of Davis–Milman–Tomczak [2] which were also used in [1].

Proposition 2. *Consider a subspace X of l_1^M of dimension n. Then there exists $M' \leq 3M/2$ and a subspace Y of $l_1^{M'}$ isometric to X such that*

$$E \sup_{x\in Y, \|x\|_1\leq 1} \left| \sum_{i\leq M} g_i x_i \right| \leq CK(X) \left(\frac{n}{M}\right)^{1/2} .$$

Proof. Consider $D = \{-1, 1\}^{\mathbf{N}}$, provided with the Haar measure μ. For a Banach space X, the K-convexity constant $K(X)$ is the norm of the natural projection from $L_2(X) = L_2(D, \mu, X)$ onto the span of the functions $\sum \varepsilon_i x_i$, where ε_i is the ith coordinate function on D. Thus, if $f \in L_2(X)$, we have

$$\left\| \sum_{i\geq 1} \varepsilon_i E(\varepsilon_i f) \right\|_{L_2(X)} \leq K(X) \|f\|_{L^2(X)} .$$

It has been observed by Tomczak–Jaegermann [14] that for any probability space Ω, if $f \in L_2(X) = L_2(\Omega, X)$, we have

$$\left\| \sum_{i\geq 1} g_i E(g_i f) \right\|_{L_2(X)} \leq K(X) \|f\|_{L_2(X)} \tag{4.1}$$

where g_i is i.i.d. $N(0, 1)$ on Ω.

As observed in [1], it is an immediate consequence of a result of Lewis [9] that for each n-dimensional subspace X of l_1^M, there exists a probability measure ν on $\{1, \dots, M\}$, a subspace Y of $L_1(\nu)$, isometric to X, and a basis $(\psi_j)_{j\leq n}$ of Y, that is orthogonal in $L_2(\nu)$, and that satisfies

$$\sum_{j=1}^{n} \psi_j^2 = 1, \quad \|\psi_j\|_2 = n^{-1/2} .$$

If we split each atom of ν of mass $a \geq 2/M$ in $[aM/2] + 1$ equal pieces, we can assume that each atom of ν has mass $\leq 2/M$, and that ν is now supported by $\{1, \dots, M'\}$ where $M' \leq 3M/2$. Also, we can assume that $\nu(\{k\}) > 0$ for $k \leq M'$.

We denote by $\langle \cdot, \cdot \rangle$ the scalar product in $L_2(\nu)$. There exists

$$f \in L_\infty(Y), \quad \|f\|_{L_\infty(Y)} \leq 1 ,$$

such that

$$\sup_{x\in Y,\|x\|_1\le 1}\left\langle \sum_{j\le n} g_j\psi_j , x\right\rangle = \left\langle \sum_{j\le n} g_j\psi_j , f\right\rangle = \sum_{j\le n}\langle \psi_j , g_j f\rangle .$$

Thus

$$E\sup_{x\in Y,\|x\|_1\le 1}\left\langle \sum_{j\le n} g_j\psi_j , x\right\rangle = \sum_{j\le n}\langle \psi_j , E(g_j f)\rangle .$$

By (4.1), setting $x_j = E(g_j f)$, we have $\|\sum g_j x_j\|_{L_2(X)} \le K(Y)\|f\|_{L^2(Y)} = K(X)$. We have

$$\left\|\sum_{j\le n} g_j x_j\right\|^2_{L_2(X)} = E\int\left(\sum_{j\le n} g_j x_j(t)\right)^2 d\nu(t) = \int\left(\sum_{j\le n} x_j^2(t)\right) d\nu(t).$$

Now, since $\sum_{j\le n}\psi_j^2 = 1$

$$\begin{aligned}\sum_{j\le n}\langle \psi_j , x_j\rangle &= \int\sum_{j\le n}\psi_j(t)x_j(t)\,d\nu(t)\\ &\le \int\left(\sum_{j\le n} x_j^2(t)\right)^{1/2} d\nu(t)\\ &\le \left\|\sum_{j\le n} g_j x_i\right\|_{L_2(X)} \le K(X).\end{aligned}$$

For $k\le M'$, let

$$a_k = \nu(\{k\}),\quad v_k = 1_{\{k\}},$$

so that $(a_k^{-1/2}v_k)_{k\le M'}$ is an orthonormal basis of $L_2(\nu)$. As observed earlier, we have

$$E\sup_{x\in Y,\|x\|_1\le 1}\left\langle \sum_{k\le M'} g_k a_k^{-1/2}v_k , x\right\rangle = E\sup_{x\in Y,\|x\|_1\le 1}\left\langle \sum_{j\le n} g_j n^{1/2}\psi_j , x\right\rangle \le n^{1/2}K(X).$$

Consider now the isometry $T\colon L_1(\nu)\to l_1^{M'}$ given by $T(x)=\sum_{k\le M'} e_k\langle v_k , x\rangle$, where (e_k) denotes the canonical basis of $l_1^{M'}$. We have

$$\begin{aligned}E\sup_{x\in T(Y),\|x\|_1\le 1}\left\langle \sum_{k\le M'} g_k e_k , x\right\rangle &= E\sup_{x\in Y,\|x\|_1\le 1}\left\langle \sum_{k\le M'} g_k v_k , x\right\rangle\\ &= E\sup_{x\in Y,\|x\|_1\le 1}\left\langle \sum_{k\le M'} g_k \frac{v_k}{\sqrt{a_k}}\sqrt{a_k} , x\right\rangle\\ &\le (\operatorname{Max} a_k^{1/2})E\sup_{x\in Y,\|x\|_1\le 1}\left\langle \sum_{k\le M'} g_k a_k^{-1/2}v_k , x\right\rangle\\ &\le C(n/N)^{1/2}K(X)\end{aligned}$$

by contraction and since $a_k\le 2/N$ for $k\le M'$. The proof is complete.

Successive applications of Propositions 2 and 1 yield the following:

Proposition 3. *There exists two universal constants* α, C, *such that if* X *is a subspace of* l_1^M *of dimension* n, *for which* $K(X)(n/M)^{1/2} \leq \alpha$, *there exists* $M' \leq 3M/4$ *and a subspace* Y *of* $l_1^{M'}$ *for which* $d(X, Y) \leq 1+CK(X)(n/M)^{1/2}$.

To prove the theorem, we observe that given $\varepsilon > 0$, a subspace X of L_1 of finite dimension is always at distance $< 1+\varepsilon$ of a subspace l_1^M for some M (however large). This can be seen in a number of ways, e.g. using the fact that $\|f - E^n(f)\|_1 \to 0$ where E^n denotes the conditional expectation with respect to the n th dyadic subalgebra. Thus, it suffices to prove the theorem for subspaces of l_1^N. In that case it follows from repeated applications of Proposition 3, once we observe that $K(Y) \leq d(X, Y)K(X)$.

4. Further comments

Also of interest is the problem, studied in particular in [1], of embedding a subspace X of $L_p(p > 1)$ of dimension n in l_p^N for small N. Using the random choice procedure described in §2, one is led to consider the random variable

$$Z_x = \|U_\varepsilon(x)\|^p - \|x\|^p = \sum_{i \leq M} \varepsilon_i |x_i|^p$$

and $A = \sup_{x \in X, \|x\|_p \leq 1} |Z_x|$. To apply the method of Proposition 1, we need to have information on the Gaussian process $V_x = \sum_{i \leq M} g_i |x_i|^p$. If one tries to deduce this information from information on the process $W_x = \sum_{i \leq M} g_i x_i$ using the standard comparison theorems one does not get the correct result, as these comparison theorems do not take into account that X is a subspace of small dimension. Thus we seem to have no choice other than evaluating the expectation of A using the canonical distance $\delta(x, y) = \sum_{i \leq M} (|x_i|^p - |y_i|^p)^2)^{1/2}$ and entropy methods, i.e. Dudley's theorem as e.g. in [10]. For that purpose we observe e.g. that for

$$\|x\|_p \leq 1, \quad \|y\|_p \leq 1,$$

we have

$$\delta(x, y) \leq C\|x - y\|_{p/(2-p)} \leq C\|x - y\|_\infty^{p/2} \qquad \text{if } 1 \leq p \leq 2$$
$$\delta(x, y) \leq Cp(\mathrm{Max}(\|x\|_\infty, \|y\|_\infty))^{(p-2)/2} \|x - y\|_\infty \qquad \text{if } p \geq 2$$

and we use the entropy evaluations of [1].

While our approach in that case replaces (once the hard work on entropy evaluation has been done!) the computations of [1], Theorem 7.3 and Theorem 7.4 by an application of Dudley's theorem, it does not even decrease the power of the $\log n$ term there, and thus it does not yield any real improvement. A similar comment applies to the results of [1], §8.

Acknowledgment

Thanks are due for useful comments to M. Ledoux, and also to W. Johnson and G. Schechtman who pointed out and repaired an error (by splitting the large atoms) in an earlier formulation of Proposition 2.

References

1. J. Bourgain, J. Lindenstrauss and V. Milman, *Approximation of zonoids by zonotopes*, Acta Math. **162** (1989), 73–141.
2. W. Davis, V. Milman and N. Tomczak-Jaegermann, *The distance between certain n-dimensional spaces*, Israel J. Math. **39** (1981), 1–15.
3. T. Figiel, J. Lindenstrauss, and V. Milman, *The dimension of almost spherical sections of convex bodies*, Acta Math. **129** (1977), 53–94.
4. Y. Gordon, *Some inequalities for Gaussian processes and applications*, Israel J. Math. **50** (1985), 265–289.
5. J. Hoffmann-Jørgensen, *Sums of independent Banach space valued random variable*, Studia Math. LII (1984), 159–186.
6. W. B. Johnson and G. Schechtman, *Embedding l_p^m into l_1^n*, Acta. Math. **194** (1982), 71–85.
7. M. Ledoux and M. Talagrand, *Comparison theorems, random geometry and some limit theorems for empirical processes*, Ann. Probab. **17** (1989), 596–631.
8. ——, *Isoperimetry and processes in probability in a Banach space*, (to be published by Springer-Verlag in the "Ergebrisse" series).
9. D. Lewis, *Finite dimensional subspaces of L_p*, Studia Math. **63** (1978), 207–211.
10. M. B. Marcus and G. Pisier, *Random Fourier series with applications to harmonic analysis*, Ann. of Math. Stud. **101** (1981).
11. V. Milman and G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Lecture Notes in Math. **1200** (1986).
12. G. Pisier, *Un théorème de factorization pour les opérateurs linéaires entre espaces de Banach*, Ann. Ecole Norm. Sup. **13** (1980), 23–43.
13. G. Schechtman, *More on embedding subspaces of L_p in l_r^n*, Compositio Math. **61** (1987), 159–170.
14. N. Tomczak-Jaegermann, *Banach-Mazur distances and finite dimensional operation ideals*, Pitman, Marshfield, Massachusetts, 1988.

Equipe d'Analyse-Tour 46, E.R.A. associée au C.N.R.S. $n°$ 754, Université Paris VI, 4 Pl. Jussieu, 75230 Paris Cedex 05, France

and

The Ohio State University, Department of Mathematics, 231 West 18th Avenue, Columbus, Ohio 43210