Acta Math., 162 (1989), 73–141

Approximation of zonoids by zonotopes(¹)

by

J. BOURGAIN J. LINDENSTRAUSS and V. MILMAN

I.H.E.S., Bures-sur-Yvette France

The Hebrew University Jerusalem, Israel

Tel Aviv University Tel Aviv, Israel

1. Introduction

Consider the Euclidean ball B^n in $\mathbf{R}^n$. It is well known that B^n can be approximated in the Hausdorff metric by a sum of segments. Given $\varepsilon>0$, what is the number N needed so that the Hausdorff distance between B^n and a sum of segments $\Sigma_{j=1}^{N} I_j$ is less than ε? It is quite clear that $N=\exp(c(\varepsilon)n)$ for a suitable $c(\varepsilon)$ will suffice. The surprising fact is that actually $N=c(\varepsilon)n$ will do. This was proved in [F.L.M.]. In this paper we show that this fact is not a special property of B^n but that essentially the same holds for any convex body in $\mathbf{R}^n$ which is a limit of a sum of segments. Questions related to this topic have been studied in the literature till now mainly in the framework of Banach space theory. Also the main tools we use in this paper are taken from Banach space theory. In order to make the paper accessible to experts in convexity theory as well as those in Banach space theory we include in the paper somewhat more than the usual amount of background material.

The introduction is divided into two parts. We first explain the geometric problem and state the main results. We then pass on to the functional analytic formulation, survey the history of the problem and explain the contents of the various sections of this paper.

A zonotope in $\mathbf{R}^n$ is a polytope P which is a vector sum of segments $\{I_j\}_{j=1}^{N}$, i.e.,

$$P=\left\{x; x=\sum_{j=1}^{N} x_j,\ x_j\in I_j,\ j=1,\dots,N\right\}.$$

(¹) This work was supported in part by the U.S.–Israel Binational Science Foundation.

(By a segment we mean a compact one dimensional convex set.) A zonotope always has a center of symmetry, namely the sum of the centers of the segments I_j. It is easy to check that every face of a zonotope is again a zonotope and hence all faces of a zonotope have a center of symmetry. Conversely, if P is a polytope all of whose faces have a center of symmetry then P is a zonotope (it actually suffices to check just its two dimensional faces; for a proof of these facts and other facts mentioned below we refer to the expository paper [Bo]). Thus, it is easy to check which polytopes are zonotopes; in $\mathbf{R}^2$ all centrally symmetric polytopes (i.e., convex polygons) are zonotopes, in $\mathbf{R}^3$ the cube for example is a zonotope while the octahedron is not. A zonoid in $\mathbf{R}^n$ is defined to be a limit in the Hausdorff metric of a sequence of zonotopes, and thus in particular is a compact convex set with a center of symmetry. From what was said above, it follows that any compact convex symmetric set in $\mathbf{R}^2$ is a zonoid. For $n \geq 3$ the unit ball in l_p^n i.e.,

$$\left\{ x = (t_1, t_2, \dots, t_n);\ \|x\| = \left(\sum_{i=1}^{n} |t_i|^p \right)^{1/p} \leq 1 \right\}$$

is a zonoid if and only if $p \geq 2$. In particular, the usual Euclidean unit ball B^n in l_2^n is a zonoid. A polytope is a zonoid if and only if it is already a zonotope. The zonoids arise naturally in many contexts. For example the zonoids are exactly the ranges of atom free vector (i.e., $\mathbf{R}^n$ valued) measures.

Clearly, an affine image of a zonotope (zonoid) is itself a zonotope (resp. zonoid). In general the polar body of a zonoid fails to be a zonoid. It follows from a result of Grothendieck that there is an absolute constant γ so that if B is a convex body in $\mathbf{R}^n$ with the origin as its center of symmetry so that both it and its polar

$$B^0 = \{y; |\langle x, y \rangle| \leq 1 \text{ all } x \in B\}$$

are zonoids then there is an affine automorphism T of $\mathbf{R}^n$ so that

$$B \subset T(B^n) \subset \gamma B.$$

There are however, for every n, convex bodies B in $\mathbf{R}^n$ which are not affine images of balls so that both B and its polar are zonoids (cf. [Schn.] and also the recent survey paper [Schn.W]). The Euclidean ball B^n plays a central role in the theory of the zonoids; the result of Grothendieck mentioned above is just one indication of this fact.

The problem we study in the present paper is the following: Given a zonoid B in $\mathbf{R}^n$

and given $\varepsilon>0$ what is the minimal $N=N(B,\varepsilon)$ so that there is a sum P_N of N segments so that

$$B\subset P_N\subset(1+\varepsilon)B. \tag{1.1}$$

The ε measures the distance between B and P_N. It is related to the Hausdorff distance ϱ between B and P_N by the obvious relation

$$\varepsilon r<\varrho<\varepsilon R \tag{1.2}$$

where r (respectively R) is the radius of the largest Euclidean ball contained (resp. the smallest ball containing) B. This ε is more natural in our context than ϱ since it makes $N(B,\varepsilon)$ an affine invariant of B.

As mentioned already above for $B=B^n$ the Euclidean ball, it was proved in [F.L.M.] that (if $\varepsilon<1/2$ say)

$$N(B^n,\varepsilon)\leqslant c\varepsilon^{-2}\log\varepsilon^{-1}\cdot n \tag{1.3}$$

where c is an absolute constant. This estimate was slightly improved in [G] where it is shown that

$$N(B^n,\varepsilon)\leqslant c\varepsilon^{-2}n. \tag{1.4}$$

The main interest in (1.3) (or (1.4)) is in the fact that $N(B^n,\varepsilon)$ depends linearly on n. This dependence on n is certainly optimal. As far as the dependence on ε is concerned it is obviously not optimal for $n=2$ where one easily sees that $N(B^2,\varepsilon)\approx\varepsilon^{-1/2}$ as $\varepsilon\to 0$. It was pointed out in [Be.Mc] that also for $n=3,4$ (1.4) can be improved. It was proved recently that also for $n\geqslant 5$ the dependence on ε in (1.4) can be improved (cf. [Li]).([1]) We shall prove however below (in section 6) that for a given $\tau>0$ we have for all sufficiently large n that

$$N(B^n,\varepsilon)\geqslant c(n)\,\varepsilon^{-2+\tau}$$

for a suitable positive constant $c(n)$ depending on n.

The main results in this paper show that the estimate obtained previously for B^n and which, as we just remaked, is quite sharp in this case, is valid in a slightly weaker form for an arbitrary zonoid.

([1]) See also the note "added in proof".

THEOREM 1.1. *Let B be a zonoid in $\mathbf{R}^n$ which is the unit ball of a uniformly convex norm in $\mathbf{R}^n$ and let $\tau>0$. Then we have*

$$N(B,\varepsilon)\leqslant C\varepsilon^{-(2+\tau)}n \tag{1.5}$$

where the constant $C=C(\tau,\delta)$ depends only on r and the degree of uniform convexity, i.e., on the constant

$$\delta=\min\{2-\|x+y\|;\|x\|,\|y\|\leqslant 1,\|x-y\|\geqslant 1\} \tag{1.6}$$

Where $\|\ \|$ is the norm in $\mathbf{R}^n$ whose unit ball is B.

Again, the main point here is the linear dependence of $N(B,\varepsilon)$ on n. For general zonoids in $\mathbf{R}^n$ (i.e., without the additional uniform convexity assumption) we have the following result.

THEOREM 1.2. *For every $\tau>0$ there is a constant $c(\tau)$ so that for any zonoid B in $\mathbf{R}^n$, $n\geqslant 2$,*

$$N(B,\varepsilon)\leqslant c(\tau)\,\varepsilon^{-(2+\tau)}(\log n)^3 n. \tag{1.7}$$

The proofs of both theorems give the following additional information on the approximating zonotopes. If B is a zonoid and P a zonotope $\Sigma_{j=1}^N I_j$ which approximates B up to ε then there exist $\{j_i\}_{i=1}^{\tilde{N}}$ where $\tilde{N}$ does not exceed the right hand side of (1.5) (resp. (1.7)) and scalars $\{\lambda_i\}_{i=1}^{\tilde{N}}$ so that $\Sigma_{i=1}^{\tilde{N}}\lambda_i I_{j_i}$ approximates P up to 2ε.

Before leaving the geometric part of the introduction, we would like to point out another interpretation of our problem and results. Let B be a convex compact subset of $\mathbf{R}^n$ having the origin as an interior point and center of symmetry. For a unit vector x in $\mathbf{R}^n$ (taken in the usual Euclidean norm) we denote by $h(x,B)$ the $(n-1)$-dimensional volume of the projection of B on the hyperplane orthogonal to x. It is well known (cf. [Bon.F.]) that

$$h(x,K)=\frac{1}{2}\int_{S^{n-1}}|\langle x,u\rangle|\,d\sigma_B(u)$$

where $S^{n-1}=\partial(B^n)$ and $\sigma_B(u)$ is a certain positive measure on S^{n-1} which is derived from the surface measure of B. Hence $h(x,K)$ is the restiction to S^{n-1} of a norm on $\mathbf{R}^n$. The unit ball of the dual of this norm is a zonoid $Z(B)$ (cf. (1.8) below) which is called the projection body of B. The correspondence $B\leftrightarrow Z(B)$ turns out to be a one to one map from the set of all convex bodies onto the set of all zonoids (we consider only those sets in $\mathbf{R}^n$ which have the origin as an interior point and center of symmetry). This classic fact is due to Minkowski and A. D. Aleksandrov (see [Bo] and [Schn.W.] as well as [Bo.F.] for details). It is easy to see and well known that B is a polytope if and only if $Z(B)$ is a zonotope; the number of segments in the representation of $Z(B)$ as a zonotope is equal to half the number of $(n-1)$-dimensional faces of B. Thus the problem we consider here can be phrased as follows. Given a centered symmetric convex body B in $\mathbf{R}^n$ and $\varepsilon>0$, find the smallest possible N so that there is a polytope P with N faces in $\mathbf{R}^n$ for which

$$Z(P) \subset Z(B) \subset (1+\varepsilon) Z(P).$$

A recent study of projection bodies which is closely related to the present paper is [B.L.].

We pass now to functional analytic terminology. Note first that by definition the sum P of N segments is an affine image of the unit cube in $\mathbf{R}^N$. Hence, if the origin is the center and an interior point of P, the norm induced by P on $\mathbf{R}^n$ is a quotient norm of l_∞^N (whose unit ball is the N-dimensional cube). Consequently the polar P^0 of P is the unit ball of an n dimensional subspace of $l_1^N=(l_\infty^N)^*$. Similarly a convex body B in $\mathbf{R}^n$ with center at the origin is a zonoid if and only if the norm $\|\ \|_*$ induced by B^0 is given by

$$\|x\|_* = \int_{S^{n-1}} |\langle x, u\rangle| d\sigma(u) \tag{1.8}$$

where $\sigma(u)$ is a positive measure on the sphere S^{n-1} symmetric with respect to 0. (The zonoid is a zonotope if and only if the measure σ is purely atomic with a finite number of atoms.) In other words B is a zonoid if and only if B^0 is the unit ball of an n-dimensional subspace of $L_1(\Omega,\mu)$ for some measure space (Ω,μ) (as mentioned above we may always take $\Omega=S^{n-1}$, another possible universal choice is $\Omega=[0,1]$ and μ=the Lebesgue measure; all these elementary facts are explained in detail in [Bo]).

Let us recall the notion of the Banach Mazur distance between two finite dimensional spaces X and Y (of the same dimension)

$$d(X, Y) = \inf\{\|T\|\,\|T^{-1}\|; T \text{ linear isomorphism from } X \text{ to } Y\}. \tag{1.9}$$

It follows that $d(X, Y)\leq d$ if and only if there is an affine map T so that $B_X \subset TB_Y \leq dB_X$ where B_X (resp. B_Y) denotes the unit ball of X (resp. Y). We can now restate our problem as follows.

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

As mentioned above this question was considered in [F.L.M] and [G] for $X=l_2^n$. For $X=l_p^n$, $1\leq p<2$ it was proved by Johnson and Schechtman [J.Sche.1] that there is a function $c(\varepsilon,p)$ so that $N(l_p^n,\varepsilon)\leq c(p,\varepsilon)\,n$. The dependence of $c(p,\varepsilon)$ on ε was of the order $\varepsilon^{-p'}$ where p' is the conjugate exponent to p and thus the result is weaker than what can be deduced from (1.5) in this special case. The proof of [J.Sche.1] depends on a delicate analysis of truncations of p-stable variables. The result of [J.Sche.1] was generalized by Pisier [P.1] who considered economical embeddings of l_p^n spaces into general spaces with a known "stable type p" constant. Schechtman continued this line of research and obtained, using probability distributions related to stable variables some results on general subspaces of $L_1(0.1)$ (cf. [Sche.1] or [Sche.2]). In [Sche.3] Schechtman attacked the problem by a new and much simpler method using what is called in probability theory "the empirical distribution method". Using this method he proved that for every subspace X of $L_1(0,1)$ of dimension n and every $0<\varepsilon<1/2$

$$N(X,\varepsilon)\leq c\varepsilon^{-2}\log\varepsilon^{-1}\cdot n^2 \tag{1.10}$$

for some absolute constant c.

In section 2 below we recall the method of Schechtman [Sche.3]. We do this since our main results are obtained by analyzing and refining the same procedure.

In section 3 we prove a weak version of Theorem 1.1 above in which the dependence of $N(B,\varepsilon)$ in n is already linear but the dependence on ε is worse than that of (1.7). For $X=l_p^n$ the dependence on ε we get in section 3 is essentially the same as that in [J.Sche.1]. The main ingredient we use in section 3 besides the empirical distribution method is a factorization theorem of Pisier [P.2] (or equivalently earlier results of this nature due to Nikishin and Maurey) which shows that we may change the position of a uniformly convex subspace X of $L_1(0,1)$ so that in the new position X actually sits "nicely" in $L_p(0,1)$ for some $p>1$.

In section 4 we prove an estimate on the entropy of the unit ball of an n dimensional subspace of $L_1(\mu)$ for some probability measure μ in the $\|\ \|_\infty$ norm. In other words, we compute for a given $r>1$ and for a suitable "good positioning" of an n dimensional subspace X of $L_1(\mu)$ the number of balls of radius r in the norm of $L_\infty(\mu)$ (and also of $L_p(\mu)$ for large finite p) needed in order to cover the unit ball of X (in the norm induced from $L_1(\mu)$). The proof involves Sudakov's inequality and its dual as well as the notion of absolutely summing operators.

In section 5 the methods of section 3 and 4 are combined and used to prove Theorems 1.1 and 1.2.

Section 6 is devoted to the approximation of the usual Euclidean ball B^n. This section can be read directly after section 2. In the first part of this section we show that for every compact convex set K in $\mathbf{R}^n$ we can approximate up to ε the ball B^n by a Minkowski sum of N rotations of λK where λ is a suitable scalar and N is given by the right hand side of (1.3). The second part of section 6 is devoted to the examination of the dependence on ε of $N(B^n, \varepsilon)$. The main result here (which is obtained by using spherical harmonics) is the fact that $N(B^n, \varepsilon) \geq c(n)\,\varepsilon^{-2(n-1)/(n+2)}$ as $\varepsilon \to 0$ for fixed n. Section 6 concludes that part of the paper which is devoted directly to the geometric problem of approximating zonoids.

All the papers we quoted above in connection with problem (*) deal also with the problem of embedding n-dimensional subspaces X of $L_r(0,1)$ with $1<r$ into l_r^N for a suitable small N. The answer to this question in the special case $X=l_2^n$ is given in [F.L.M.]. In section 7 it is shown that, in analogy to the situation for $r=1$, the estimates obtained in [F.L.M.] for l_2^n are valid in a slightly weaker form for an arbitrary n-dimensional subspace of $L_r(0,1)$. These results strengthen the previous results in this direction (in [J.Sche.1], [P.1], [Sche.1] and [Sche.3]).

If X is an n-dimensional subspace of $L_r(0,1)$ which is nicely complemented we may wish to find Y in l_r^N with say $d(X, Y) \leq 2$ so that not only N is small but that Y is also nicely complemented in l_r^N. In section 8 we show, using the empirical distribution method, that this can be done for a general X and the estimate for N obtained is slightly weaker than the estimates obtained in [F.L.M.] for the special case of $X=l_2^n$. We also apply these results to the siuation studied in [B.Tz.1]. In particular we deduce a result on the Banach space structure of some spaces which arise in harmonic analysis.

The last section of the paper is again concerned with entropy estimates. We obtain a sharper form of estimates proved in section 4. These sharper estimates, in view of their connection with classical questions, are of some independent interest (though they do not yield better estimates for $N(X, \varepsilon)$). These estimates also have applications to questions in approximation theory related to Sobolev inequalities, especially the so-called Bernstein width (cf. [B.G.]).

The main results of this paper were announced in [B.L.M.1].

In this paper we consider only Banach spaces of dimension strictly larger than 1 over the real field. The results and their proofs are valid (with some minor modifications) also in the complex case. There are several universal constants which enter into the estimates below. These constants are denoted by letters like $\gamma, c, C, c_1, c_2, \ldots$. We did not distinguish carefully between the different constants, neither did we try to get good estimates for them (the proofs however, yield naturally some estimates). The same letter will be used to denote different universal constants in different parts of the paper. The notation such as $c(\tau,\delta)$ will mean a function depending only on the parameters τ and δ. The cardinality of a (finite) set A is denoted by $\bar{\bar{A}}$. The characteristic function of a subset A of Ω is denoted by χ_A (provided Ω is clear from the context).

2. The empirical distribution method

In this section we present the method and main result of Schechtman [Sche.3].

Let X be an n-dimensional subspace of $L_1(\Omega,\mu)$ where (Ω,μ) is some probability space. Let $t_1,t_2,\ldots,t_N$ be N randomly (and independently) chosen points in this measure space. Consider the linear map $T\colon L_1(\mu)\to l_1^N$ defined by

$$Tf=N^{-1}(f(t_1),f(t_2),\ldots,f(t_N)). \tag{2.1}$$

The basic idea of the method is that for N sufficiently large and some (usually most) choices of $(t_1,t_2,\ldots,t_N)$ the map T is almost an isometry on X, i.e.,

$$\sup\left\{\left|1-N^{-1}\sum_{j=1}^{N}|f(t_j)|\right|;\, f\in X,\ \|f\|=1\right\} \tag{2.2}$$

is small. The tool for verifying that this happens is the following standard classical estimate from probability theory, often called Bernstein's inequality.

LEMMA 2.1. *Let $\{g_j\}_{j=1}^N$ be independent random variables with mean* 0 *on some probability space (Ω,μ) which satisfy*

$$\|g_j\|_1\le 2,\quad \|g_j\|_\infty\le M,\quad 1\le j\le N, \tag{2.3}$$

for some constant M. Then for $0<\varepsilon<1$

$$\operatorname{Prob}\left\{\left|\sum_{j=1}^{N}g_j\right|\ge\varepsilon N\right\}<2\exp(-\varepsilon^2N/8M). \tag{2.4}$$

Proof. For $-\infty<x\le 1$ we have $e^x\le 1+x+x^2$. Hence, since $\int g_j\,d\mu=0$ and $|g_j(t)|\le M$ for every j and every $t\in\Omega$ we have if $0<\lambda M\le 1$

$$\int\exp(\lambda g_j)\,d\mu\le 1+\lambda^2\int g_j^2\,d\mu\le 1+\lambda^2\|g_j\|_1\,\|g_j\|_\infty<\exp(2\lambda^2M).$$

By the independence of $\{g_j\}_{j=1}^N$ we deduce that

$$\int \exp\left(\lambda \sum_{j=1}^N g_j\right) d\mu < \exp(2\lambda^2 NM)$$

and consequently

$$\exp(\lambda N\varepsilon)\,\mathrm{Prob}\left\{\sum_{j=1}^N g_j \geqslant \varepsilon N\right\} < \exp(2\lambda^2 NM).$$

Choosing $\lambda=\varepsilon/4M$ we certainly fulfil $0<\lambda M<1$ and get

$$\mathrm{Prob}\left\{\sum_{j=1}^N g_j \geqslant \varepsilon N\right\} < \exp(-\varepsilon^2 N/8M).$$

and this implies (2.4). □

COROLLARY 2.2. *Let (Ω,μ) be a probability space, let $\mathcal{F}$ be a finite set in $L_1(\Omega,\mu)$ so that*

$$\|f\|_1 \leqslant 1, \quad \|f\|_\infty \leqslant M, \quad f\in\mathcal{F} \tag{2.5}$$

for some constant M. Let $0<\varepsilon<1$ and let N be an integer so that

$$2\bar{\bar{\mathcal{F}}} \leqslant \exp(\varepsilon^2 N/8M) \tag{2.6}$$

then there exist $\{t_j\}_{j=1}^N$ in Ω so that

$$\left|\,\|f\|_1 - N^{-1}\sum_{j=1}^N |f(t_j)|\,\right| \leqslant \varepsilon, \quad f\in\mathcal{F}. \tag{2.7}$$

Proof. For every $f\in\mathcal{F}$ consider the random variables $\{g_j\}_{j=1}^N$ on (Ω^N,μ^N) defined by

$$g_j(t) = |f(t_j)| - \int |f|\,d\mu, \quad j=1,\dots,N; \quad t=(t_1,t_2,\dots,t_N)\in\Omega^N.$$

The $\{g_j\}_{j=1}^N$ satisfy the assumptions in Lemma 2.1 and hence by (2.4)

$$\mathrm{Prob}\left\{t\in\Omega^N,\ \left|\,\|f\|_1 - N^{-1}\sum_{j=1}^N |f(t_j)|\,\right| > \varepsilon\right\} < 2\exp(-\varepsilon^2 N/8M).$$

In view of (2.6) we deduce that there is a $t\in\Omega^N$ so that (2.7) holds. □

6–898282 *Acta Mathematica* 162. Imprimé le 8 mars 1989

For applying Lemma 2.1 and Corollary 2.2, a good estimate on $\| \|_\infty$ is needed for the subsets of L_1 we consider. Such an estimate is provided by the next lemma.

LEMMA 2.3. *Let X be an n-dimensional subspace of $L_1(\Omega,\mu)$ for some probability space (Ω,μ). Then there is a linear operator $U: L_1(\Omega,\mu)\to L_1(\Omega,\nu)$ for some probability measure ν on Ω so that U is an isometry on X and*

$$\|f\|_{L_\infty(\Omega,\nu)} \leq n\|f\|_{L_1(\Omega,\nu)} \quad \text{for every} \quad f\in UX. \tag{2.8}$$

Proof. Let $\{f_i\}_{i=1}^n$ be an Auerbach basis in X, i.e., an algebraic basis of X so that $\|f_i\|=1$ for all i and so that whenever $f=\Sigma_{i=1}^n a_i f_i$ is in the unit ball of X then $|a_i|\leq 1$ for all i (the existence of such a basis is easy to prove cf. e.g., [L.Tz.1. p.16]). Let $F(t)=\Sigma_{i=1}^n |f_i(t)|$. Then $\|F\|_1=n$; $|f(t)|\leq F(t)$ for every $t\in\Omega$ and every f in the unit ball of X. Let ν be defined by $d\nu=n^{-1}F\,d\mu$ and $U: L_1(\mu)\to L_1(\nu)$ by

$$Ug(t)=\begin{cases} ng(t)/F(t) & \text{if } F(t)\neq 0 \\ 0 & \text{if } F(t)=0 \end{cases}$$

With these definitions all the assertions are evident. □

The preceding lemma asserts that for an n-dimensional subspace X of $L_1(\mu)$ we can embed X in $L_1(\nu)$ for another measure so that in the new position we have $\|f\|_\infty\leq n$ for every f in the unit ball of X. The operation of passing from μ to ν is called a change of density operation. If $F(t)\neq 0$ for all t then the map U is an isometry from $L_1(\Omega,\mu)$ onto $L_1(\Omega,\nu)$.

A subset $\mathcal{F}$ of a set D is called an ε-net in D for some $\varepsilon>0$ if for every $x\in D$ there in a $y\in\mathcal{F}$ with $\|x-y\|\leq\varepsilon$. The following two lemmas are completely standard facts.

LEMMA 2.4. *The unit ball B of an n-dimensional space (and any subset of it) has for every $\varepsilon>0$ an ε net of cardinality $\leq(1+2/\varepsilon)^n$.*

Proof. Consider a maximal subset $\mathcal{F}$ of B with the property that $\|u-v\|\geq\varepsilon$ for every $u,v\in\mathcal{F}$, $u\neq v$. The open balls with centers in $\mathcal{F}$ and radius $\varepsilon/2$ are mutually disjoint and all are contained in $(1+\varepsilon/2)B$. By considering the volumes of these balls we get that $\bar{\bar{\mathcal{F}}}\leq(1+2/\varepsilon)^n$. □

LEMMA 2.5. *Let T be a bounded linear map from a Banach space X into a Banach space Y. Let $0<\varepsilon<1$ and assume that for an ε net $\mathcal{F}$ of the unit ball of X*

$$\big|\|Tx\|-\|x\|\big|\leq\varepsilon,\quad x\in\mathcal{F}.$$

Then

$$\big|\|Tx\|-\|x\|\big|\leq\Delta\|x\| \quad \textit{for every } x\in X \textit{ where } \quad \Delta=3\varepsilon/(1-\varepsilon). \tag{2.9}$$

Proof. Put $\Delta=\sup\{\big|\|Tx\|-\|x\|\big|,\ \|x\|\leq 1\}$. Let x be any point in the unit ball of X and let $y\in\mathcal{F}$ satisfy $\|x-y\|\leq\varepsilon$. By the triangle inequality

$$\big|\|Tx\|-\|x\|\big|\leq\big|\|Ty\|-\|y\|\big|+\big|\|T(x-y)\|-\|x-y\|\big|+2\|x-y\|.$$

Hence $\Delta\leq\varepsilon+\varepsilon\Delta+2\varepsilon$ and (2.9) follows. □

We now have all the tools needed to prove

THEOREM 2.6 (Schechtman [Sche.3]). *Let X be an n-dimensional subspace of $L_1[0,1]$ and let $0<\varepsilon<1/2$. Then there is a subspace Y of l_1^N with $d(X,Y)\leq 1+\varepsilon$ provided that*

$$N\geq c\varepsilon^{-2}\log\frac{1}{\varepsilon}\cdot n^2 \tag{2.10}$$

for some absolute constant c.

Proof. By Lemma 2.3 we may assume that $\|x\|_\infty\leq n$ for every x in the unit ball of X. Let $\mathcal{F}$ be an ε-net in the unit ball of X of cardinality $\leq(3/\varepsilon)^n$ (Lemma 2.4). By Corollary 2.2 if

$$\log 2+n\log(3/\varepsilon)\leq\varepsilon^2N/8n$$

i.e., if (2.10) holds for a suitable c, the map T from X into l_1^N defined by (2.1) satisfies $\big|\|Tx\|-\|x\|\big|<\varepsilon$ for every $x\in\mathcal{F}$. Hence by Lemma 2.5,

$$d(X,TX)\leq(1+\Delta)(1-\Delta)^{-1}\leq 1+8\varepsilon$$

if ε is small enough. □

Our goal in the coming three sections is to improve the estimate (2.10), i.e., to pass from n^2 to n (or almost n). The lemmas used above by themselves are sharp, i.e., cannot be improved. For example in Lemma 2.3 we cannot replace n by any smaller number. If $X=l_1^n$ then in any isometric embedding of X into $L_1(\Omega,\nu)$ the unit vectors of X have to map onto disjointly supported functions and thus since ν is a probability measure at least one of the unit vectors of X will have an image whose $\|\ \|_\infty$ norm is at least n. The point which makes the improvement possible, however, is the exploitation of the fact that the set $\mathcal{F}$ to which we applied Corollary 2.2 is not an arbitrary set of cardinality $(3/\varepsilon)^n$ in L_1 but a set sitting in a low dimensional subspace. Thus while we cannot avoid the situation that some members of $\mathcal{F}$ have an $\|\ \|_\infty$ norm equal to n, we can manage the data so that this happens only to relatively few elements or alternatively that a large value is attained only for a very small subset of Ω (substantially smaller than what follows just from the condition that $\|f\|_1=1$).

3. The uniformly convex case—the iteration procedure

In this section we shall treat uniformly convex spaces X and prove a weak version of Theorem 1.1. For the purpose of proving this result we have to measure quantitatively the degree of uniform convexity. In the introduction we defined the constant δ in (1.6). This δ is perhaps the easiest measure of uniform convexity to understand geometrically. However, for our purposes (and many other contexts as well) the notion of "type" is more useful.

The Banach space X is said to be of type $1\le p\le 2$ if there is a constant α so that for every choice of $\{x_i\}_{i=1}^m$, $m=1,2,\dots$ in X

$$\int_0^1 \left\| \sum_{i=1}^m r_i(t)\, x_i \right\| dt \le \alpha \left(\sum_{i=1}^m \|x_i\|^p \right)^{1/p}, \tag{3.1}$$

where $\{r_i\}_{i=1}^\infty$ denote the Rademacher functions. The quantity on the left hand side of (3.1) is just the average of $\|\Sigma_{i=1}^m \pm x_i\|$ over all the possible 2^m choices of signs. The smallest α for which (3.1) holds is called the type p constant of X and is denoted by $T_p(X)$. Evidently $T_1(X)=1$ for every X. Every finite dimensional X is type p for every $1\le p\le 2$ (this, as well as the fact that p can never be taken >2, follows immediately from the generalized parallelogram identity in Hilbert space, i.e., the identity

$$\int \left\| \sum_{i=1}^m r_i(t)\, x_i \right\|^2 dt = \sum_{i=1}^m \|x_i\|^2$$

valid in suchs spaces) but in general $T_p(X)$ may tend to infinity with $\dim X$ if $p>1$. It is known that if an infinite dimensional space X fails to be of type p for any $p>1$ then for every $\varepsilon>0$, X has a subspace Y_ε with $d(Y_\varepsilon, l_2^1)\leqslant 1+\varepsilon$. In other words such X contains for every $\varepsilon>0$ unit vectors $u_\varepsilon, v_\varepsilon$ so that $\|u_\varepsilon+v_\varepsilon\|\geqslant 2-\varepsilon$ and $\|u_\varepsilon-v_\varepsilon\|\geqslant 2-\varepsilon$ and thus the δ defined by (1.6) is 0 (cf. [L.Tz.2] or [M.Sche] for further references and other basic facts concerning the notion of type). Hence by a straightforward compactness argument we get that for a Banach space X (finite or infinite dimensional) with a given $\delta>0$ defined by (1.6) there is a $p(\delta)>1$ and a finite $C(\delta)$ so that $T_{p(\delta)}(X)\leqslant C(\delta)$. It is not hard to give a good explicit estimate of $p(\delta)$ and $C(\delta)$ in terms of δ but this is of no real relevance to the problems we consider in this paper. We shall from now on state our results in terms of p and $T_p(X)$ without restating them in terms of δ.

We exploit the uniform-convexity assumption via the next proposition which is a special case of a factorization theorem of Pisier [P.2, Theorem 1.2] (which in turn is based on earlier results of Maurey and Nikishin). In order to state the proposition we have to recall the notion of weak L_p spaces. Let (Ω,μ) be a probability measure space and let $p\geqslant 1$. We denote by $L_{p,\infty}(\mu)$ the space of real valued measurable functions f on Ω so that

$$\|f\|_{p,\infty} = \sup_{t>0} t(\mu\{\omega\in\Omega; |f(\omega)|>t\})^{1/p} < \infty. \tag{3.2}$$

PROPOSITION 3.1 (Pisier). *Let X be a subspace of $L_1(\Omega,\mu)$ of type p for some $p>1$. Then there is a non-negative function F in $L_1(\Omega,M)$ so that*

$$\|F\|_{L_1(\mu)}=1,\quad \{\omega; F(\omega)=0\}\subset \bigcap_{x\in X}\{\omega; x(\omega)=0\}$$

and so that for every $x\in X$

$$\|x(\omega)/F(\omega)\|_{L_{p,\infty}(\Omega,F\mu)} \leqslant eT_p(X)\|x\|_{L_1(\Omega,\mu)} \tag{3.3}$$

where e is the base of the natural logarithm.

In other words the proposition asserts that after a suitable change of density transformation (passing from μ to the measure ν defined by $d\nu=Fd\mu$) we can assume that the $L_{p,\infty}$ norm on the unit ball of X is bounded by a constant which is essentially $T_p(X)$.

It will be very convenient for our purposes to work not with general change of density operations but only with those for which $F(\omega)\geqslant 1/2$ for every $\omega\in\Omega$. This can always be achieved at the cost of an increase of the constant by an amount depending on p.

LEMMA 3.2. *Let X be a subspace of $L_1(\Omega,\mu)$ of type $p>1$. Then there is a function G in $L_1(\Omega,\mu)$ so that*

$$G(\omega)\geq 1/2 \quad \textit{for every} \quad \omega\in\Omega, \quad \|G(\omega)\|_{L_1(\mu)}=1 \tag{3.4}$$

and so that for every $x\in X$

$$\|x(\omega)/G(\omega)\|_{L_{p,\infty}(\Omega,G\mu)}\leq cT_p(X)\|x\|/(p-1) \tag{3.5}$$

where c is an absolute constant.

Proof. Let F be the function given by Proposition 3.1. The function $G=(1+F)/2$ clearly satisfies (3.4). We shall show that also (3.5) holds. Let ν be the measure on Ω defined by $d\nu=Fd\mu$ let $t>0$ and let $x\in X$ with $\|x\|=1$. Put

$$A_0=\{\omega;F(\omega)\geq 1\},\quad A_k=\{\omega;2^{-k}\leq F(\omega)<2^{-k+1}\},\quad k=1,2,\ldots$$
$$D=\{\omega;|x(\omega)|\geq t(1+F(\omega))/2\}.$$

Clearly

$$D\subset\{w;|x(\omega)|\geq tF(\omega)/2\}\cap\{\omega;|x(\omega)|\geq t/2\}$$

and hence by (3.3)

$$\nu(D)\leq\nu\{\omega;|x(\omega)|\geq tF(\omega)/2\}\leq(2eT_p(X)/t)^p$$

and also

$$\begin{aligned}\mu(D)&=\sum_{k=0}^{\infty}\mu(D\cap A_k)\leq\sum_{k=0}^{\infty}2^k\nu(D\cap A_k)\\&\leq\nu(D)+\sum_{k=1}^{\infty}2^k\nu\{\omega;|x(\omega)|>t2^{k-2}F(\omega)\}\\&\leq\left(\frac{eT_p(X)}{t}\right)^p\left(2^p+\sum_{k=1}^{\infty}2^k2^{p(2-k)}\right).\end{aligned}$$

Consequently, for some absolute c_1

$$\frac{\mu(D)+\nu(D)}{2}\leq\left(\frac{c_1T_p(X)}{(p-1)t}\right)^p$$

and this is equivalent to (3.5). □

The main technical result in this section is the following

LEMMA 3.3. *Let X be a subspace of l_1^N of dimension n. Let $1<p\leq 2$ and $0<\varepsilon<1/2$ and assume that $N\geq n\varepsilon^{-2}$. Then there is a subspace Y of $l_1^{\tilde N}$ with $d(X, Y)\leq 1+\varepsilon$ provided that*

$$\tilde N \geq cT_p(X)\, N^{1/p} n^{1-1/p}\varepsilon^{-2}(\log\varepsilon^{-1})^{1-1/p}\left(\log\frac{N}{n}\right)^{1/p}(p-1)^{-3} \tag{3.6}$$

where c is an absolute constant.

Proof. By Lemmas 2.3 and 3.2 we may assume that there is a probability measure μ on $\{1,2,\ldots,N\}$ so that $\mu\{i\}\geq 1/2N$ for every i and so that

$$\|x\|_{L_\infty(\mu)}\leq c_2\, n\|x\|;\quad \|x\|_{L_{p,\infty}(\mu)}\leq c_2\, T_p(X)\,\|x\|/(p-1),\quad x\in X. \tag{3.7}$$

By Lemma 2.4 there is an ε-net $\mathscr{F}$ in the unit ball of X so that

$$\log \bar{\bar{\mathscr{F}}}\leq c_3\, n\log\varepsilon^{-1}. \tag{3.8}$$

For every $x\in\mathscr{F}$ we put

$$A_1(x)=\{i;|x(i)|\leq 2\},\quad A_k(x)=\{i;2^{k-1}<|x(i)|\leq 2^k\},\quad k=2,3,\ldots \tag{3.9}$$

Note that by (3.7) $A_k(x)$ is empty for $k>[\log c_2 n]$. We also put for $x\in\mathscr{F}$, $x_k=x\cdot\chi_{A_k(x)}$, $k=1,2,\ldots$ and note that

$$x=\sum_{k=1}^{[\log c_2 n]} x_k,\quad |x_k|\wedge|x_h|=0\quad\text{if}\quad h\neq k,\quad x\in\mathscr{F}. \tag{3.10}$$

We shall specify below a choice of positive scalars ε_k, $k=1,\ldots$ so that

$$\sum_{k=1}^{[\log c_2 n]}\varepsilon_k\leq c_4\,\varepsilon/(p-1) \tag{3.11}$$

and prove, applying the empirical distribution method to the families $\{x_k\}_{x\in\mathscr{F}}$, $k=1,2\ldots$ that if

$$N_0\geq c_5\, T_p(X)\, N^{1/p}n^{1-1/p}\varepsilon^{-2}(\log\varepsilon^{-1})^{1-1/p}\left(\log\frac{N}{n}\right)^{1/p}(p-1)^{-1} \tag{3.12}$$

there is a choice of integers $\{i_j\}_{j=1}^{N}$ (not necessarily distinct) in $\{1,2,\dots,N\}$ so that

$$\left| \|x_k\| - N_0^{-1} \sum_{j=1}^{N_0} |x_k(i_j)| \right| \leq \varepsilon_k, \quad x \in \mathscr{F}, \quad k=1,2\dots \tag{3.13}$$

Once we verify (3.13) we are done. Indeed by (3.10) and (3.11) it will follow that

$$\left| \|x\| - N_0^{-1} \sum_{j=1}^{N_0} |x(i_j)| \right| \leq \sum_k \varepsilon_k \leq c_4 \varepsilon/(p-1), \quad x \in \mathscr{F}.$$

Consequently, by applying Lemma 2.5 to $T: X \to l_1^{N_0}$ defined by $Tx = N_0^{-1}(x(i_1), \dots, x(i_{N_0}))$, we get that $d(X, TX) \leq c_6 \varepsilon/(p-1)$. We have only now to replace ε by $c_6 \varepsilon/(p-1)$; this transforms the N_0 of (3.12) to the $\tilde{N}$ of (3.6) and will conclude the proof.

In order to verify (3.13) note first that by (3.7) and (3.9) we get for every $x \in \mathscr{F}$ and $k \geq 2$ that

$$\overline{\overline{A_k(x)}} \leq 2N\mu(A_k(x)) \leq [c_7 N (T_p(X)/(p-1)2^k)^p] = u_k. \tag{3.14}$$

Consequently, all the elements in $\{|x_k|\}_{k \in \mathscr{F}}$, $k=2,\dots$ are convex combinations of the family $\mathscr{D}_k$ of all the functions on $\{1,2,\dots,N\}$ which vanish outside a certain subset of cardinality u_k and on this set take one of the three values $0, 2^{k-1}$ or 2^k. In order to ensure that (3.13) holds for a certain k it suffices that a similar statement holds for all $y \in \mathscr{D}_k$. The cardinality of $\mathscr{D}_k$ is at most $\binom{N}{u_k} 3^{u_k}$ and hence

$$\log \bar{\bar{\mathscr{D}}}_k \leq c_8 u_k \log\left(1 + \frac{N}{u_k}\right) \leq c_9 k N \left(\frac{T_p(X)}{p-1}\right)^p 2^{-kp}. \tag{3.15}$$

Let k_0 be the smallest integer such that for $k \geq k_0$, $\bar{\bar{\mathscr{D}}}_k \leq \bar{\bar{\mathscr{F}}}$. By (3.8) and (3.15) we have

$$2^{k_0} \leq c_{10} \frac{T_p(X)}{p-1} \left(\frac{N}{\log \bar{\bar{\mathscr{F}}}}\right)^{1/p} \left(\log \frac{N}{\log \bar{\bar{\mathscr{F}}}}\right)^{1/p} \tag{3.16}$$

In view of Lemma 2.1 we can ensure that (3.13) holds for all k provided that

$$\bar{\bar{\mathscr{F}}} \sum_{k=1}^{k_0} \exp(-\varepsilon_k^2 N_0/8 \cdot 2^k) + \sum_{k=k_0+1}^{[\log c_2 n]} \bar{\bar{\mathscr{D}}}_k \exp(-\varepsilon_k^2 N_0/8 \cdot 2^k) < 1/2. \tag{3.17}$$

We choose first $\varepsilon_{k_0} = \varepsilon$ and $\varepsilon_{k_0-h} = (2/3)^{h/2}\varepsilon$ for $h=1,2,\dots,k_0-1$. Then $\Sigma_{k=1}^{k_0} \varepsilon_k \leq \varepsilon/(1-\sqrt{2/3})$ and so far we fulfil (3.11). If we put $\eta = \exp(-\varepsilon^2 N_0/8 \cdot 2^{k_0})$ then in view of (3.8), (3.12) and

(3.16) we get that $\bar{\bar{\mathscr{F}}}\eta<\frac{1}{8}$ (provided the constant c_5, which we are still free to choose, is large enough). Hence

$$\bar{\bar{\mathscr{F}}}\sum_{k=1}^{k_0}\exp(-\varepsilon_k^2 N_0/8\cdot 2^k)\leqslant \bar{\bar{\mathscr{F}}}(\eta+\eta^{4/3}+\eta^{(4/3)^2}+\ldots)$$

$$\leqslant \bar{\bar{\mathscr{F}}}\eta(1+\eta^{1/3}+\eta^{2/3}+\ldots)\leqslant \bar{\bar{\mathscr{F}}}\eta/(1-\eta^{1/3})\leqslant 2\bar{\bar{\mathscr{F}}}\eta<1/4.$$

We pass to the second sum in (3.17). We put $k_1=[\log c_2 n]$. If $k_1\leqslant k_0$ the sum does not appear and there is no problem. So we assume that $k_0<k_1$. To ensure that (3.17) holds it suffices to choose ε_k for $k_0<k\leqslant k_1$ so that

$$\bar{\bar{\mathscr{D}}}_k\leqslant(\exp(\varepsilon_k^2 N_0/8\cdot 2^k))^{1/2},\quad k_0<k\leqslant k_1, \tag{3.18}$$

$$\sum_{k=k_0+1}^{k_1}(\exp(-\varepsilon_k^2 N_0/8\cdot 2^k))^{1/2}<\frac{1}{4}. \tag{3.19}$$

The condition (3.18) is equivalent to

$$2c_9 kN\left(\frac{T_p(X)}{p-1}\right)^{p-1}2^{-kp}\leqslant \varepsilon_k^2 N_0\cdot 2^{-k+3},\quad k>k_0$$

but since N_0 was chosen so that this holds for $k=k_0$ we get that (3.18) is verified once

$$\varepsilon_{k_0+h}\geqslant\varepsilon\left(\frac{k_0+h}{k_0}2^{-h(p-1)}\right)^{1/2},\quad h\geqslant 1.$$

As for (3.19) we choose $\varepsilon_{k_1}=\varepsilon$. Then

$$(\exp(-\varepsilon^2 N_0/8\cdot 2^{k_1}))^{1/2}<1/8,$$

provided $N_0>c_{11}n\varepsilon^{-2}$. This requirement on N_0 is ensured by (3.12) since we assume from the start that $N\geqslant n\varepsilon^{-2}$. Now the same computation as that done for the first sum in (3.17) shows that if

$$\varepsilon_{k_1-h}\geqslant\varepsilon\left(\frac{2}{3}\right)^{h/2},\quad h>0 \tag{3.21}$$

then (3.19) holds. We take now ε_k for $k_0<k<k_1$ to be the maximum of the two numbers appearing in the right hand sides of (3.20) and (3.21). With this choice of the $\{\varepsilon_k\}_{k=1}^{k_1}$ we get that (3.17) holds and that $\Sigma_{k=1}^{k_1}\varepsilon_k$ is bounded by an absolute constant times $\varepsilon/(p-1)$. □

Lemma 3.3 is evidently a useful tool for an iteration procedure. Let X be an n-dimensional subspace of $L_1(0,1)$ and let $\delta_1>0$. We find a subspace Y_1 in $l_1^{N_1}$ with $d(X, Y_1)\leq 1+\delta_1$ so that N_1 is not too large (say by using Theorem 2.6). We then choose another number δ_2 and by applying Lemma 3.3 find a subspace Y_2 of $l_1^{N_2}$ (with N_2 the $\tilde{N}$ of Lemma 3.3) with $d(Y_1, Y_2)\leq 1+\delta_2$. We continue to apply Lemma 3.3 as long as we can and as long as the numbers N_k decrease. If we stop after s stages we get a space Y_s for which $d(X, Y_s)\leq\Pi_{k=1}^s(1+\delta_k)$.

To put this procedure into a precise framework we recall the quantity $N(X,\varepsilon)$ defined by the formulation of problem (*) in the introduction. Since

$$(1+\varepsilon)\left(1+\frac{\delta}{2}\right)\leq 1+\varepsilon+\delta \quad \text{if} \quad 0<\varepsilon, \delta\leq 1/2$$

we can restate Lemma 3.3 as follows.

Lemma 3.3′. *There is an absolute constant c so that for every choice of* $0<\varepsilon$, $\delta<1/2$, *every* $1<p\leq 2$ *and every n-dimensional subspace X of* $L_1(0,1)$

$$N(X,\varepsilon+\delta)\leq cT_p(X)\,N(X,\varepsilon)^{1/p}n^{1-1/p}\delta^{-2}(\log\delta^{-1})^{1-1/p'}\left(\log\frac{N(X,\varepsilon)}{n}\right)^{1/p}(p-1)^{-3}, \quad (3.22)$$

provided that $N(X,\varepsilon)\geq 4n\delta^{-2}$.

Corollary 3.4. *Let X be an n-dimensional subspace of $L_1(0,1)$, let $1<p\leq 2$, $0<\varepsilon<1$ and let $\varrho>2$. Then*

$$N(X,\varepsilon)\leq C(p, T_p(X),\varrho)\,\varepsilon^{-\varrho p/(p-1)}n \quad (3.23)$$

where, as the notation indicates, $C(p, T_p(X),\varrho)$ depends just on p,ϱ and the type p constant of X.

Proof. Let $\tau>1$ be such that $\varrho p/(p-1)>2\tau p/(p-\tau)$ and put $\varphi(\varepsilon)=N(X,\varepsilon)/n$. By (3.22) there is a $\gamma=\gamma(p, T_p(X),\tau)$ so that for every $0<\varepsilon<1$

$$\varphi(\varepsilon)\leq\gamma\cdot\varepsilon^{-2\tau}\varphi\left(\frac{\varepsilon}{2}\right)^{\tau/p} \quad (3.24)$$

provided $\varphi(\varepsilon/2)\geq(4/\varepsilon)^2$. By iterating (3.24) k times we get

$$\varphi(\varepsilon)\leq\gamma^{1+\tau/p+\ldots+(\tau/p)^k}\varepsilon^{-2\tau(1+\tau/p+\ldots+(\tau/p)^k)}2^{\tau/p+2\tau^2/p^2+\ldots+k(\tau/p)^k}\varphi\left(\frac{\varepsilon}{2^k}\right)^{\tau^k/p^k}. \quad (3.25)$$

By Theorem 2.6, $\varphi(\varepsilon/2^k)\leq n\varepsilon^{-2\tau}2^{2k\tau}$. Substituting this inequality in (3.25) with $k\sim\log n$ we get the desired result. (If the iteration stops earlier, i.e., $\varphi(\varepsilon/2^i)\leq(2^{i+1}/\varepsilon)^2$ for some $i<k$ then we have an even better inequality.) □

Corollary 3.4 is a weak version of Theorem 1.1 and as mentioned in the introduction it already generalizes the main result of [J.Sche.1]. In order to obtain Theorem 1.1 as stated we have to add to the method of this section an argument involving certain entropy numbers, which will be evaluated in the next section.

4. Entropy estimates

In this section we obtain an estimate on the number of balls of radius t in $\|\ \|_\infty$ norm needed to cover the unit ball of a subspace X of L_1 which is in a "good position". The estimate is based on several known results. We devote a large part of this section to the statement of those known (or essentially known) results and the explanation of the notions from Banach space theory which enter in their formulations.

We start with the notion of entropy. Let D and B be subsets of a linear space and let $t>0$. We put

$$E(D,B,t)=\min\left\{k;\ \exists\{x_i\}_{i=1}^k,\ D\subset\bigcup_{i=1}^k(x_i+tB)\right\}.$$

In the cases we consider here this number will always be finite. Sometimes it is convenient to use instead of $E(D,B,t)$ the number

$$\bar{E}(D,B,t)=\min\left\{k;\ \exists\{x_i\}_{i=1}^k,\ x_i\in D,\ D\subset\bigcup_{i=1}^k(x_i+tB)\right\}.$$

It follows directly from the definition that always (if B is convex and symmetric)

$$\bar{E}(D,B,2t)\leq E(D,B,t)\leq\bar{E}(D,B,t). \tag{4.1}$$

An important quantity which enters into entropy computations for subsets of $\mathbf{R}^n$ is the following: Let $|||\ |||$ denote the Euclidean norm in $\mathbf{R}^n$ (with B^n its unit ball and S^{n-1} its boundary) and let $\|\ \|$ be another norm on $\mathbf{R}^n$. We denote $X=(\mathbf{R}^n,\|\ \|)$ and $X^*=(\mathbf{R}^n,\|\ \|_*)$ where the duality is with respect to the inner product defined by $|||\ |||$. The average of $\|x\|$ on S^{n-1} is denoted by M_X, i.e.,

$$M_X=\int_{S^{n-1}}\|x\|\,d\sigma(x) \tag{4.2}$$

where σ is the normalized rotation invariant measure on S^{n-1}. By the homogeneity of $\|\cdot\|$ and by using polar coordinates we get easily other convenient formulas for M_X for example

$$M_X = \alpha_n (2\pi)^{-n/2} \int_{\mathbf{R}^n} \|x\| \exp(-|||x|||^2/2)\, dx, \quad \alpha_n \sim n^{-1/2}. \tag{4.3}$$

A probabilistic way to rewrite (4.3) is to consider n independent and normalized Gaussian variables $\{g_i(\omega)\}_{i=1}^n$ on some probability measure space (Ω, μ). Then

$$M_X = \alpha_n \int_\Omega \left\| \sum_{i=1}^n g_i(\omega)\, e_i \right\| d\mu(\omega) \tag{4.4}$$

where $\{e_i\}_{i=1}^n$ is an orthonormal basis in $\mathbf{R}^n$.

The following result is well known in probability theory.

PROPOSITION 4.1 (Sudakov [Su]). *Let $X=(\mathbf{R}^n, \|\ \|)$ and let $|||\ |||$ be the Euclidean norm on $\mathbf{R}^n$. Then*

$$\log E(B_X, B^n, t) \leq Cn \left(\frac{M_{X^*}}{t} \right)^2 \tag{4.5}$$

where B_X (resp. B^n) denote the unit balls of $\|\ \|$ (resp. $|||\ |||$) and C is a universal constant.

A dual version of this proposition was recently obtained by A. Pajor and N. Tomczak-Jaegermann [Pa.T-J].

PROPOSITION 4.2. *With the same notation as in Proposition 4.1 we have*

$$\log E(B^n, B_X, t) \leq cn \left(\frac{M_X}{t} \right)^2. \tag{4.6}$$

We shall present here a proof of (4.6) which is different from that indicated in [Pa.T-J]. This proof was kindly communicated to us by A. Pajor and is due to him and M. Talagrand.

Proof. We let μ be the probability measure on $\mathbf{R}^n$ defined by

$$d\mu = (2\pi)^{-n/2} \exp(-|||x|||^2/2)\, dx.$$

By (4.3)

$$\mu\{x;\|x\|\leqslant 2M_X\alpha_n^{-1}\}\geqslant 1/2. \tag{4.7}$$

Let $\{x_i\}_{i=1}^N$ be a maximal subset in B^n relative to the requirement that $\|x_i-x_j\|\geqslant t$ for $i\neq j$. The sets $\{x_i+\frac{1}{2}tB_X\}_{i=1}^N$ have mutually disjoint interiors. Hence,

$$1\geqslant\sum_{i=1}^N\mu\{y_i+2M_X\alpha_n^{-1}B_X\}\quad\text{where}\quad y_i=4M_X(t\alpha_n)^{-1}x_i. \tag{4.8}$$

Fix $1\leqslant i\leqslant N$. By (4.7), the convexity of the function e^{-u} and the symmetry of B_X with respect to the origin, we get

$$\begin{aligned}\mu\{y_i+2M_X\alpha_n^{-1}B_X\} &= (2\pi)^{-n/2}\int_{2M_X\alpha_n^{-1}B_X}\exp(-|||x-y_i|||^2/2)\,dx\\ &\geqslant(2\pi)^{-n/2}\int_{2M_X\alpha_n^{-1}B_X}\exp(-(|||x-y_i|||^2+|||x+y_i|||^2)/4)\,dx\\ &=(2\pi)^{-n/2}\int_{2M_X\alpha_n^{-1}B_X}\exp(-(|||x|||^2+|||y_i|||^2)/2)\,dx\\ &\geqslant\frac{1}{2}\exp(-|||y_i|||^2/2)\geqslant\frac{1}{2}\exp(-(4M_X/\alpha_n t)^2).\end{aligned}$$

Hence by (4.8), $N\leqslant 2\exp(4M_X/\alpha_n t)^2$ and since $\alpha_n\sim n^{-1/2}$ we deduce (4.6). □

We recall now the notion of the K-convexity constant $K(X)$ of the Banach Space X. Let Q be the natural projection from $L_2([0,1],X)$ onto the span of $\{x_ir_i(t)\}_{i=1}^\infty$ where $x_i\in X$ and $r_i(t)$ are the Radamacher functions i.e.,

$$Qf(s)=\sum_{i=1}^\infty\left(\int_0^1 f(t)\,r_i(t)\,dt\right)r_i(s),\quad f\in L_2([0,1],X). \tag{4.9}$$

Then

$$K(X)=\|Q\|\quad\text{(with the usual operator norm)}. \tag{4.10}$$

It is evident that $\|Q\|=1$ if X is a Hilbert space and thus in particular $K(X)<\infty$ whenever $\dim X<\infty$. The K convexity constant will enter our computations in view of the following inequality which is an immediate consequence of the definition. For every choice of $\{x_i^*\}_{i=1}^m$ in X^* we have

$$\begin{aligned}\left(\int_0^1 \left\|\sum_{i=1}^m r_i(t)x_i^*\right\|_{X^*}^2 dt\right)^{1/2} &= \sup\left\{\int_0^1 \left\langle \sum_{i=1}^m r_i(t)x_i^*, f(t)\right\rangle dt,\ \|f\|_{L_2(X)} \leq 1\right\} \\ &\leq K(X)\sup\left\{\int_0^1 \left\langle \sum_{i=1}^m r_i(t)x_i^*, \sum_{i=1}^m r_i(t)x_i\right\rangle dt,\ \left\|\sum_{i=1}^m r_i(t)x_i\right\|_{L_2(X)} \leq 1\right\} \qquad (4.11) \\ &= K(X)\sup\left\{\sum_{i=1}^m x_i^*(x_i);\ \left\|\sum_{i=1}^m r_i(t)x_i\right\|_{L_2(X)} \leq 1\right\}.\end{aligned}$$

As for estimates on $K(X)$; the following are due to Pisier [P.3] (see also [M.Sche], Chapter 14).

$$K(X) \leq C\log d(X, l_2^n) \leq C\log n, \quad n = \dim X \qquad (4.12)$$

$$K(X) \leq C(\log n)^{1/2}, \quad n = \dim X, \quad X \subset L_1(0,1) \qquad (4.13)$$

$$K(L_p(\Omega,\mu)) \sim (p-1)^{1/2}, \quad 1 < p \leq 2. \qquad (4.14)$$

An estimate which seems to be new is given in

LEMMA 4.3. *Let X be an n-dimensional subspace of $L_1(0,1)$ and let $1<p\leq 2$. Then*

$$K(X) \leq C(\log(1+T_p(X))/(p-1))^{1/2}, \qquad (4.15)$$

where C is an absolute constant and $T_p(X)$ the type p constant of X.

Proof. By Proposition 3.1 there is no loss of generality to assume that $X \subset L_1(\Omega,\mu)$ for some probability measure space (Ω,μ) and that we have for every $x \in X$

$$\|x\|_{L_{p,\infty}} \leq eT_p(X)\|x\|_{L_1}. \qquad (4.16)$$

Let $0<\theta<1$ and $1/s=(1-\theta)/1+\theta/p$. A standard computation (involving the sets $\{\omega, 2^k \leq |f(\omega)| \leq 2^{k+1}\}$ for every k) shows that for $f \in L_{p,\infty}(\Omega,\mu)$ we have for some absolute c_1

$$\|f\|_{L_s} \leq c_1\left(\frac{p-1}{p-s}\right)^{1/s} \|f\|_{L_1}^{1-\theta}\|f\|_{L_{p,\infty}}^{\theta}$$

and in particular if we restrict θ to be in $(0, \frac{1}{2}]$

$$\|f\|_{L_s} \leq c_2 \|f\|_{L_1}^{1-\theta} \|f\|_{L_{p,\infty}}^{\theta} \tag{4.17}$$

for some absolute c_2. Hence by (4.16) and (4.17) we get that if $\theta \leq \frac{1}{2}$ then the Banach Mazur distance of X from a subspace, of $L_s(\Omega, \mu)$ is $\leq c_3 T_p(X)^\theta$. Consequently by (4.14) $K(X) \leq c_4 (s-1)^{-1/2} T_p(X)^\theta$. By choosing θ so that $2(s-1)\log(1+T_p(X)) = p-1$ we get (4.15). □

Two other notions from Banach space theory will be useful for us: 2-absolutely summing norms and cotype q. The 2-absolutely summing norm $\pi_2(T)$ of a linear operator $T: X \to Y$ is defined to be the smallest constant γ so that

$$\sum_{i=1}^{m} \|Tx_i\|^2 \leq \gamma^2 \sup \left\{ \sum_{i=1}^{m} |x^*(x_i)|^2, \ \|x^*\|_{X^*} \leq 1 \right\} \tag{4.18}$$

holds for every choice of $\{x_i\}_{i=1}^m$, $m = 1, 2, \ldots$, in X (in our case $\dim X$ will be finite and this ensures in particular that $\pi_2(T) < \infty$ for every operator T). The notion of cotype is dual to that of type. For $2 \leq q < \infty$ the cotype q constant of the Banach space Y, denoted by $C_q(Y)$ is the smallest constant γ for which

$$\left(\sum_{i=1}^{m} \|y_i\|^q \right)^{1/q} \leq \gamma \int_0^1 \left\| \sum_{i=1}^{m} y_i r_i(t) \right\| dt \tag{4.19}$$

holds for every choice of $\{y_i\}_{i=1}^m$, $m = 1, 2, \ldots$, in Y. A simple consequence of the classical Khintchine inequality is that $C_2(L_1(0,1)) < \infty$, i.e., $L_1(0,1)$ is a space of cotype 2. With these notions we can state the next lemma which is just a reformulation of Lemma 1 in [D.M.T-J].

Lemma 4.4. *With X as in Proposition 4.1. we have*

$$M_{X^*} \leq c n^{-1/2} K(X)\, C_2(X)\, \pi_2(T) \tag{4.20}$$

where T is the formal identity map from X into $l_2^n = (\mathbf{R}^n, \||\ \||)$. In particular, if X is a subspace of $L_1(0,1)$ then

$$M_{X^*} \leq \tilde{c} n^{-1/2} K(X)\, \pi_2(T). \tag{4.21}$$

Proof. Evidently (4.21) follows from (4.20) with $\tilde{c}=cC_2(L_1(0,1))$. We prove (4.20). By the Pietch factorization theorem (cf. e.g., [L.Tz.1 p. 64]) there exists an $m\geq n$ so that T can be factored as $T=V\Delta U$

$$X\xrightarrow{U} l_\infty^m \xrightarrow{\Delta} l_2^m \xrightarrow{V} l_2^n$$

where $\|U\|\leq 1$, $\|V\|\leq 1$, $\Delta f_j=\lambda_j f_j$, the $\{f_j\}_{j=1}^m$ being the unit vectors in l_∞^m and l_2^m and

$$\left(\sum_{j=1}^m \lambda_j^2\right)^{1/2} = \|\Delta\| \leq 2\pi_2(T). \tag{4.22}$$

Since an orthogonal $m\times m$ matrix takes m independent normalized Gaussian variables $\{g_i\}_{i=1}^m$ (on some probability space (Ω,μ)) into a set of m variables having the same joint distribution and since every operator of norm ≤ 1 on l_2^n is a convex combination of orthogonal transformations we deduce (using (4.4) and the fact that $\|V^*\|\leq 1$) that

$$M_{X^*} = \alpha_n \int_\Omega \left\| \sum_{i=1}^n g_i(\omega)\, e_i \right\|_{X^*} d\mu(\omega) \leq \alpha_n \int_\Omega \left\| \sum_{j=1}^m g_j(\omega)\, U^*\Delta^* f_j \right\|_{X^*} d\mu(\omega).$$

Again, since for every t the functions $\{r_j(t)\, g_j(\omega)\}_{j=1}^m$ have the same joint distribution as $\{g_j(\omega)\}_{j=1}^m$ it follows that

$$\begin{aligned} M_{X^*} &\leq \alpha_n \int_\Omega \int_0^1 \left\| \sum_{j=1}^m r_j(t)\, g_j(\omega)\, U^*\Delta^* f_j \right\| dt\, d\mu(\omega) \\ &\leq \alpha_n \int_\Omega \left(\int_0^1 \left\| \sum_{j=1}^m g_j(\omega)\, r_j(t)\, U^*\Delta^* f_j \right\|^2 dt \right)^{1/2} d\mu(\omega). \end{aligned} \tag{4.23}$$

For every $\omega\in\Omega$ we get by (4.11)

$$\begin{aligned} &\left(\int_0^1 \left\| \sum_{j=1}^m g_j(\omega)\, r_j(t)\, U^*\Delta^* f_j \right\|^2 \right)^{1/2} \\ &\leq K(X) \sup\left\{ \sum_{j=1}^m g_j(\omega)\, U^*\Delta^* f_j(x_j);\ \left\| \sum_{j=1}^m r_j(t)\, x_j \right\|_{L_2(X)} \leq 1 \right\}. \end{aligned} \tag{4.24}$$

By (4.19) we have for all $\{x_j\}_{j=1}^m$ in X

$$\left(\sum_{j=1}^m \|x_j\|^2 \right)^{1/2} \leq C_2(X) \left\| \sum_{j=1}^m r_j(t)\, x_j \right\|_{L_1(X)} \leq C_2(X) \left\| \sum_{j=1}^m r_j(t)\, x_j \right\|_{L_2(X)}.$$

Hence $\|\Sigma_{j=1}^m r_j(t)\, x_j\|_{L_2(X)}\leq 1$ implies that

$$\sum_{j=1}^m g_j(\omega)\, U^*\Delta^* f_j(x_j)\leq \left(\sum_{j=1}^m |g_j(\omega)\,\lambda_j|^2\right)^{1/2}\left(\sum_{j=1}^m \|Ux_j\|^2\right)^{1/2}\leq C_2(X)\left(\sum_{j=1}^m |g_j(\omega)\,\lambda_j|^2\right)^{1/2}.$$

Combining this inequality with (4.22), (4.23) and (4.24) we get

$$\begin{aligned} M_{X^*} &\leq \alpha_n K(X)\, C_2(X)\int_\Omega \left(\sum_{j=1}^m |g_j(\omega)\,\lambda_j|^2\right)^{1/2} d\mu(\omega) \\ &\leq c_1\,\alpha_n K(X)\, C_2(X)\left(\sum_{j=1}^m \lambda_j^2\right)^{1/2} \\ &\leq cn^{-1/2} K(X)\, C_2(X)\,\pi_2(T). \end{aligned}$$

□

Our next lemma is another change of density result.

Lemma 4.5. *Let (Ω,μ) be a probability space and let X be an n dimensional subspace of $L_1(\Omega,\mu)$. Then there is another probability measure ν on Ω and a subspace $\tilde{X}$ of $L_1(\Omega,\nu)$ which is isometric to X so that*

$$\pi_2(T)\leq cn^{1/2} \tag{4.25}$$

where T is the identity map from $\tilde{X}$ onto $(\tilde{X}, |||\ |||)$ where $|||\ |||$ is the norm induced on $\tilde{X}$ from $L_2(\Omega,\nu)$ and c, as usual, denotes an absolute constant.

Proof. As shown by Lewis [Le] there is a basis $\{\varphi_i\}_{i=1}^n$ of X so that for all scalars $\{\lambda_j\}_{j=1}^n$

$$\sum_{j=1}^n \lambda_j^2/n = \int_\Omega \left(\left|\sum_{j=1}^n \lambda_j\varphi_j\right|^2 \Big/ \Phi\right) d\mu$$

where

$$\Phi = \left(\sum_{j=1}^n \varphi_j^2\right)^{1/2}, \quad \|\Phi\|_{L_1(\mu)} = 1.$$

Define now a measure ν on Ω by $d\nu=\Phi d\mu$ and let $f\to\Phi^{-1} f$ map X onto a subspace $\tilde{X}$ of $L_1(\Omega,\nu)$. Clearly this map is an isometry and $\tilde{X}$ has a basis $\psi_j=\varphi_j/\Phi$, $j=1,2,\dots,n$ satisfying

$$\sum_{j=1}^{n} \psi_j^2 \equiv 1, \quad \int_\Omega \psi_j^2 \, d\nu = n^{-1}, \quad \int_\Omega \psi_k \psi_j \, d\nu = 0 \quad \text{if} \quad k \neq j. \tag{4.26}$$

Consider now the functions h_t defined on $\Omega, 0 \leq t \leq 1$ by $(i^2 = -1)$

$$h_t(\omega) = \operatorname{Imag} \prod_{j=1}^{n} (1 + i r_j(t) \psi_j(\omega)) = \sum_{j=1}^{n} r_j(t) \psi_j(\omega) + \text{higher order Walsh function in } t.$$

By (4.26)

$$\|h_t\|_{L_\infty(\nu)} \leq \sup_\omega \prod_{j=1}^{n} (1 + \psi_j^2(\omega))^{1/2} \leq \sqrt{e} \tag{4.27}$$

and for every $x = \Sigma_{j=1}^n \lambda_j \psi_j$ in $\tilde{X}$

$$\int_\Omega x(\omega) \, h_t(\omega) \, d\nu = \sum_{j=1}^{n} \lambda_j r_j(t)/n + \text{higher order Walsh functions}$$

and thus

$$\int_0^1 \left(\int_\Omega x(\omega) \, h_t(\omega) \, d\nu \right)^2 dt \geq \sum_{j=1}^{n} \lambda_j^2 / n^2 = |||x|||^2_{L_2(\nu)}/n. \tag{4.28}$$

By (4.27) and (4.28) we get for any choice of $\{x_k\}_{k=1}^m$ in X

$$\sup\left\{ \sum_{k=1}^{m} |x^*(x_k)|^2; \ \|x^*\| \leq 1 \right\} \geq e^{-1} \sup_{0 \leq t \leq 1} \sum_{k=1}^{m} \left(\int_\Omega h_t(\omega) \, x_k(\omega) \, d\nu \right)^2$$

$$\geq e^{-1} \sum_{k=1}^{m} \int_0^1 \left(\int_\Omega h_t(\omega) \, x_k(\omega) \, d\nu \right)^2 dt \geq \sum_{k=1}^{m} |||x_k|||^2 \Big/ ne$$

and this is equivalent to (4.25). □

Remark. Lemma 4.5 is already the third change of density result mentioned in this paper. By forming averages of the densities obtained in the various results and allowing a possible change in the absolute constants involved, we can ensure that we work with densities which have all the desired properties at once. In all the situations appearing here it is easy to check that the averaging procedure is permissible (the least obvious case involves Proposition 3.1 where the checking is made in Lemma 3.2).

We combine now most of the preceding results to obtain the following entropy estimate.

PROPOSITION 4.6. *Let X be an n-dimensional subspace of $L_1(\Omega,\mu)$. Then there is a probability measure ν on Ω given by $d\nu=Fd\mu$ so that for the space $\tilde{X}=F^{-1}X$ in $L_1(\Omega,\nu)$ we have for every $t>0$ and every $q\geq 1$*

$$\log E\{B_{\tilde{X}},B_q,t\}\leq cK(X)\,nq^{1/2}/t \tag{4.29}$$

where c is an absolute constant and B_q is the unit ball of $\tilde{X}_q=\tilde{X}$ endowed with the norm induced by $L_q(\Omega,\nu)$.

Proof. Let ν be as in the proof of Lemma 4.5. From the definition of entropy it follows immediately that

$$E\{B_{\tilde{X}},B_q,t\}\leq E\{B_{\tilde{X}},B_2,\lambda\}\,E\{B_2,B_q,t/\lambda\} \tag{4.30}$$

for every $0<\lambda<t$. We take as $|||\ |||$ the norm in $\tilde{X}_2$ i.e., that induced by $L_2(\Omega,\nu)$. By Propositions 4.1, 4.2 and (4.30) we deduce that

$$\log E\{B_{\tilde{X}},B_q,t\}\leq c_1\,n\left(\frac{M_{X^*}^2}{\lambda^2}+\frac{\lambda^2}{t^2}M_{\tilde{X}_q}^2\right). \tag{4.31}$$

By Lemmas 4.4 and 4.5, $M_{X^*}\leq c_2K(X)$. Let $\{g_j(\omega')\}_{j=1}^n$ be normalized independent Gaussians on a probability space (Ω',μ'). We take as unit vectors in $\tilde{X}_2$ the functions $\{\sqrt{n}\,\psi_j\}_{j=1}^n$ with the ψ_j as in (4.26). By (4.4)

$$M_{\tilde{X}_q}=\alpha_n n^{1/2}\int_{\Omega}\left\|\sum_{j=1}^n g_j(\omega')\,\psi_j\right\|_{L_q(\Omega',\nu)}d\mu(\omega').$$

It follows from the Maurey-Khintchine inequality (cf. e.g., [L.Tz.2 p. 49]) (4.26) and the fact that $\alpha_n\sim n^{-1/2}$ that

$$M_{\tilde{X}_q}\leq c_3\,q^{1/2}\left\|\left(\sum_{j=1}^n\psi_j^2\right)^{1/2}\right\|_{L_q(\Omega,\nu)}=c_3\,q^{1/2}.$$

Substituting in (4.31) we get that

$$\log E\{B_{\tilde{X}},B_q,t\}\leq c_4\,n\left(\frac{K(X)^2}{\lambda^2}+\frac{q\lambda^2}{t^2}\right).$$

By taking minimum over λ, (4.29) follows. □

Remark. As in the remark above we may replace ν by $(\nu+\mu)/2$. Then we get a measure on Ω which is dominated from below by $\mu/2$ and for which (4.29) still holds (with c replaced by $2c$).

COROLLARY 4.7. *Let X be an n-dimensional subspace of l_1^N. Then there exists a probability measure ν on $\{1,\dots,N\}$ and a subspace $\tilde{X}$ of $L_1(N,\nu)$ isometric to X so that*

$$\log E\{B_{\tilde{X}}, B_\infty, t\} \leq c(\log n \log N)^{1/2} n/t \tag{4.32}$$

where B_∞ is the unit ball of $l_\infty^N=(N,\nu)$.

Proof. Let μ be the usual probability measure on $\{1,2,\dots,N\}$, i.e., $\mu\{i\}=N^{-1}$ for every i. We apply Proposition 4.6 to find $\tilde{X}$ and ν. By the remark after Proposition 4.6 we may assume that $\nu\{i\}>(2N)^{-1}$ for every i. For every $f\in L_q(N,\nu)$

$$\max_i |f_i|/(2N)^{1/q} \leq \|f\| = \left(\sum_{i=1}^N |f_i|^q \nu\{i\}\right)^{1/q} \leq \max_i |f_i|$$

and hence $d(L_q(N,\nu), L_\infty(N,\nu)) \leq (2N)^{1/q}$. We take now $q=\log N$ in (4.29), use the estimate (4.13) for $K(X)$ and (4.32) follows. □

COROLLARY 4.8. *Let X be an n-dimensional subspace of $L_1(0,1)$. Then for every $\varepsilon>0$ there is a probability space (Ω,ν) and a subspace $\tilde{X}$ of $L_1(\Omega,\nu)$ with $d(X,\tilde{X})\leq 1+\varepsilon$ so that*

$$\log E\{B_{\tilde{X}}, B_{L_\infty(\Omega,\nu)}, t\} \leq c(\log n(\log n\varepsilon^{-1}))^{1/2} n/t. \tag{4.33}$$

Proof. Combine (4.32) with Theorem 2.6. □

In the two corollaries above we may use the estimate (4.15) for $K(X)$ instead of (4.13). Then for example we get, instead of (4.33), the following estimate with $1<p\leq 2$,

$$\log E\{B_{\tilde{X}}, B_{L_\infty(\Omega,\nu)}, t\} \leq c[\log(T_p(X)+1)\log(n\varepsilon^{-1})(p-1)^{-1}]^{1/2} n/t. \tag{4.34}$$

In section 9 we shall prove a slightly stronger version of (4.33), which however will not be of use for estimating $N(X,\varepsilon)$.

5. Proof of the main theorems

Our purpose now is to combine the tools developed in the last three sections and to derive from them Theorems 1.1 and 1.2. We show first how the entropy numbers enter into the analysis of the empirical method.

PROPOSITION 5.1. *Let X be an n-dimensional subspace of $L_1(\Omega,\mu)$ for some probability measure μ. Let $\|\cdot\|_\infty$ denote the norm in $L_\infty(\Omega,\mu)$ and put*

$$B_X=\{x;\ x\in X,\ \|x\|\le 1\},\quad B_\infty=\{x;\ x\in X,\ \|x\|_\infty\le 1\}$$

and $M=\sup\{\|x\|_\infty;\ x\in X,\ \|x\|\le 1\}$ which we assume to be finite. Then for some absolute constant C and all $0<\varepsilon<1/2$,

$$N(X,\varepsilon)\le C\varepsilon^{-2}\left(\left(\int_1^M\left(\frac{\log E(B_X,B_\infty,t)}{t}\right)^{1/2}dt\right)^2+n\log\varepsilon^{-1}\right). \tag{5.1}$$

Proof. Let $l=[\log_2 M]+1$ and for $s=0,1,\dots,l$ let $\mathscr{A}_s$ be a subset of B_X so that

$$B_X\subset\bigcup_{y\in\mathscr{A}_s}\{y+2^sB_\infty\},\quad \bar{\bar{\mathscr{A}}}_s\le\tilde{E}(B_X,B_\infty,2^s). \tag{5.2}$$

For $s=l$ we choose $\mathscr{A}_l$ to consist just of the origin 0. Let $\mathscr{F}$ be an ε-net in B_X so that

$$\log\bar{\bar{\mathscr{F}}}\le c_1 n\log\varepsilon^{-1} \tag{5.3}$$

(cf. Lemma 2.4). For every $x\in\mathscr{F}$ and $0\le s\le l$ let $y_{s,x}\in\mathscr{A}_s$ satisfy $\|x-y_{s,x}\|_\infty\le 2^s$. We clearly have

$$|x(\omega)|=(|x(\omega)|-|y_{0,x}(\omega)|)+\sum_{s=1}^{l}(|y_{s,x}(\omega)|-|y_{s-1,x}(\omega)|),\quad x\in X,\quad \omega\in\Omega. \tag{5.4}$$

We are going to apply the empirical method to the following collection of subsets of $L_\infty(\Omega,\mu)$

$$\mathscr{E}_0=\{|x|-|y_{0,x}|;\ x\in\mathscr{F}\},\quad \mathscr{E}_s=\{|y_{s,x}|-|y_{s-1,x}|;\ x\in\mathscr{F}\},\quad 1\le s\le l.$$

Since

$$\|\,|y_{s,x}|-|y_{s-1,x}|\,\|_\infty\le\|y_{s,x}-x\|_\infty+\|y_{s-1,x}-x\|_\infty\le 2^s+2^{s-1}\le 2^{s+1}$$

we have

$$\|f\|_1 \leq 2, \quad \|f\|_\infty \leq 2^{s+1}, \quad f \in \mathscr{E}_s, \quad 0 \leq s \leq l. \tag{5.5}$$

Also by (4.1), (5.2) and (5.3)

$$\begin{gathered} \log \bar{\bar{\mathscr{E}}}_0 \leq c_1 n \log \varepsilon^{-1}, \\ \log \bar{\bar{\mathscr{E}}}_s \leq \log \bar{\bar{\mathscr{A}}}_s + \log \bar{\bar{\mathscr{A}}}_{s-1} \leq 2 \log E(B_X, B_\infty, 2^{s-2}), \quad 1 \leq s \leq l. \end{gathered} \tag{5.6}$$

Our aim it to find $\{\omega_j\}_{j=1}^N$ in Ω where N is given by the right hand side of (5.1) so that

$$\left| N^{-1} \sum_{j=1}^N f(\omega_j) - \int_\Omega f(\omega)\, d\mu(\omega) \right| \leq \varepsilon, \quad f \in \mathscr{E}_0 \tag{5.7}$$

$$\left| N^{-1} \sum_{j=1}^N f(\omega_j) - \int_\Omega f(\omega)\, d\mu(\omega) \right| \leq \varepsilon_s, \quad f \in \mathscr{E}_s, \quad 1 \leq s \leq l \quad \text{with} \quad \sum_{s=1}^l \varepsilon_s \leq \varepsilon. \tag{5.8}$$

Once (5.7) and (5.8) hold we get by (5.4) that

$$\left| N^{-1} \sum_{j=1}^N |x(\omega_j)| - \|x\| \right| \leq 2\varepsilon, \quad x \in \mathscr{F}$$

and an application of Lemma 2.5 will conclude the proof.

To ensure that (5.7) and (5.8) hold it is enough to verify that the measure of the set of $\{\omega_j\}_{j=1}^N$ in Ω^N for which at least one of the inequalities fails is less than 1. In view of Lemma 2.1 this will be the case in particular if

$$2\bar{\bar{\mathscr{E}}}_0 \exp(-\varepsilon^2 N/32) < 1/2 \tag{5.9}$$

$$2 \sum_{s=1}^l \bar{\bar{\mathscr{E}}}_s \exp(-\varepsilon_s^2 N/32 \cdot 2^{s+1}) < 1/2. \tag{5.10}$$

(We apply Lemma 2.1 to the functions $f - \int f\, d\mu$, $f \in \mathscr{E}_s$; these functions have an L_1 norm bounded by 4 and an L_∞ norm bounded by 2^{s+1}.) By (5.6) it follows that (5.9) is ensured once

$$N \geq c_2 n \varepsilon^{-2} \log \varepsilon^{-1}. \tag{5.11}$$

Note that since obviously $E(B_X, B_\infty, \lambda) \geq M/\lambda$, it follows from (5.6) that (5.10) is ensured once we have for a suitable $c_3>0$

$$\exp(-\varepsilon_s^2 N/32 \cdot 2^{s+1}) \geq c_3 E(B_X, B_\infty, 2^{s-2})^4, \quad 1 \leq s \leq l$$

i.e.,

$$\varepsilon_s N^{1/2} \geq c_4 (\log E(B_X, B_\infty, 2^{s-2}))^{1/2} 2^{s/2}, \quad 1 \leq s \leq l.$$

Since we also need to have that $\Sigma_{s=1}^{l} \varepsilon_s \leq \varepsilon$, and $E(B_X, B_\infty, t)$ is a monotone decreasing function of t, we get that what will ensure (5.10) is

$$\varepsilon N^{1/2} \geq c_4 \sum_{s=1}^{l} \log E(B_X, B_\infty, 2^{s-2}) 2^{s/2} \geq c_5 \int_1^M \left(\frac{\log E(B_X, B_\infty, t)}{t} \right)^{1/2} dt. \tag{5.12}$$

Since the right hand side of (5.1) verifies both (5.11) and (5.12) the proposition is proved. □

We are now ready to prove Theorem 1.2 in the following, more precise, formulation.

THEOREM 5.2. *Let X be an n dimensional subspace of $L_1(0,1)$. Then for some absolute constant c*

$$N(X, \varepsilon) \leq c\varepsilon^{-2} \log(n\varepsilon^{-1})(\log n)^2 n, \quad \varepsilon > 0. \tag{5.13}$$

Proof. By Lemma 2.3 and Corollary 4.8 there is for every $\varepsilon>0$ a probability space (Ω, ν) and a subspace $\tilde{X}$ of $L_1(\Omega, \nu)$ so that $d(X, \tilde{X}) \leq 1+\varepsilon$, $\sup\{\|x\|_\infty;\ x \in \tilde{X},\ \|x\|=1\} \leq 2n$ and

$$\log E(B_{\tilde{X}}, B_\infty, t) \leq c_1 [\log n \log(n\varepsilon^{-1})]^{1/2} n/t, \quad t > 1.$$

Hence by (5.1)

$$N(\tilde{X}, \varepsilon) \leq c_2 \varepsilon^{-2} n \left(\left(\int_1^{2n} \frac{dt}{t} \right)^2 \cdot [\log n \cdot \log(n\varepsilon^{-1})]^{1/2} + \log \varepsilon^{-1} \right)$$

and this implies (5.13). □

We pass now to estimates on $N(X, \varepsilon)$ for spaces for which we have information on $T_p(X)$ for some $p>1$. Our aim is to prove Theorem 1.1. We shall combine in the proof the weak L_p estimates given by Proposition 3.1 with the entropy estimates from section 4. The strategy is just to combine the proof of Lemma 3.3. with that of Proposition 5.1. While the proof involves no additional ideas to those already used above the task of keeping track of all the parameters leads to some quite lengthy (but elementary) computations.

LEMMA 5.3. *Let X be an n-dimensional subspace of l_1^N. Let $1<p\le 2$, $0<\varepsilon<1/2$ and assume that $N\ge n\varepsilon^{-2}$. Then there is a subspace Y of $l_1^{\tilde N}$ with $d(X,Y)\le 1+\varepsilon$ provided that*

$$\tilde N\ge cn\varepsilon^{-2}(\log T)^{1/2}(p-1)^{-3}(\log(NT/n(p-1))^{5/2} \tag{5.14}$$

where c is a universal constant and $T=T_p(X)+1$.

Proof. By Lemmas 2.3 and 3.2 we may assume that there is a probability measure μ on $\{1,2,\dots,N\}$ so that $\mu\{i\}>1/2N$ for every i and

$$\|x\|_{L_\infty(\mu)}\le c_2 n\|x\|,\quad \|x\|_{L_{p,\infty}(\mu)}\le c_2 T_p(X)(p-1)^{-1}\|x\|,\quad .\, x\in X. \tag{5.15}$$

In addition, in view of Lemma 4.3 and Proposition 4.6, we may also assume that for every $q>1$

$$\log E(B_X,B_q,t)\le c_3 K(X)\,nq^{1/2}/t\le c_4\left(\frac{\log(1+T_p(X))}{p-1}\right)^{1/2} q^{1/2}n/t. \tag{5.16}$$

Note that, by (5.15), $\|x\|_{L_q(\mu)}\le c_2 n^{1-1/q}\|x\|$ for every $x\in X$ and thus for $t>c_2 n^{1-1/q}$, $E(B_X,B_q,t)=1$. Let $\mathscr{F}$ be an ε-net in the unit ball of X satisfying (5.3). We fix now a $q\ge 2$ and put

$$\alpha=8c_2T_p(X)/(p-1),\quad \beta=8c_4\left(\frac{\log(1+T_p(X))}{p-1}\right)^{1/2}\cdot q^{1/2},\quad l=[\log c_2 n^{1-1/q}]. \tag{5.17}$$

For $0\le s\le l$ we choose now a set G_s, in B_X so that

$$B_X\subset\bigcup_{y\in G_s}\{y+2^sB_q\},\quad \bar{\bar{G}}_s\le \bar E(B_X,B_q,2^s). \tag{5.18}$$

The set G_l consists just of the origin. For $x\in\mathscr{F}$, $0\le s\le l$ let $y_{s,x}\in G_s$ satisfy $\|x-y_{s,x}\|_q\le 2^s$. Put next

$$\mathscr{E}_0=\{|x|-|y_{0,x}|;\ x\in\mathscr{F}\},\quad \mathscr{E}_s=\{|y_{s,x}|-|y_{s-1,x}|;\ x\in\mathscr{F}\},\quad 1\le s\le l.$$

By (4.1), (5.16), (5.17) and (5.18) we get that

$$\log \bar{\bar{\mathscr{E}}}_s \leq \beta n 2^{-s}, \quad 1 \leq s \leq l. \tag{5.19}$$

By (5.15) we get that for every $z \in \mathscr{E}_s$, $0 \leq s \leq l$

$$\|z\|_{L_{p,\infty}(\mu)} \leq 2, \quad \|z\|_{L_q(\mu)} \leq 2^{s+1}, \quad \|z\|_{L_\infty(\mu)} \leq \min((2N)^{1/q} 2^{s+1}, 2c_2 n). \tag{5.20}$$

As in the proof of Proposition 5.1 we shall apply the empirical method to all the sets $\mathscr{E}_s$, $0 \leq s \leq l$. We find below positive numbers $\{\theta_s\}_{s=1}^l$ and $\{\delta_s\}_{s=1}^l$ and certain expressions $N_{1,q}, N_{2,q}$ and $N_{3,q}$ so that

$$\sum_{s=1}^{l} \theta_s \leq \varepsilon, \quad \sum_{s=1}^{l} \delta_s < 1/2 \tag{5.21}$$

and so that for $\tilde{N} \geq \max(N_{1,q}, N_{2,q}, N_{3,q})$ we have

$$\mathrm{Prob}\left\{ \left| \tilde{N}^{-1} \sum_{j=1}^{\tilde{N}} z(i_j) - \int z\, d\mu \right| > \varepsilon \text{ for at least one } z \in \mathscr{E}_0 \right\} < 1/2 \tag{5.22}$$

$$\mathrm{Prob}\left\{ \left| \tilde{N}^{-1} \sum_{j=1}^{\tilde{N}} z(i_j) - \int z\, d\mu \right| > \theta_s \text{ for at least one } z \in \mathscr{E}_s \right\} < \delta_s, \quad 1 \leq s \leq l, \tag{5.23}$$

where the probability of $\{i_j\}_{j=1}^{\tilde{N}}$ is taken of course with respect to $\mu^{\tilde{N}}$. Once we verify this the proof will be concluded by showing that the right hand side of (5.14) exceeds $\max(N_{1,q}, N_{2,q}, N_{3,q})$ for a suitably chosen q.

We start with (5.22). As in the proof of Lemma 3.3 we cut each z in $\mathscr{E}_0$ into its level sets $A_k(z) = \{i;\ 2^{k-1} < |z(i)| \leq 2^k\}$ and consider instead of $\mathscr{E}_0$ the $\sim \log n$ sets $\mathscr{E}_{0,k}$ consisting of $z \cdot \chi_{A_k(z)}$. We apply the empirical method to all those sets using the reasoning of Lemma 3.3 but with L_q estimates (instead of $L_{p,\infty}$ estimated used there). Thus, for example, (3.16) becomes in our setting (note that $\bar{\bar{\mathscr{F}}} = \bar{\bar{\mathscr{E}}}_0$)

$$2^{k_0} \leq c_5 \left(\frac{N}{\log \bar{\bar{\mathscr{F}}}} \right)^{1/q} \left(\log \left(\frac{N}{\log \bar{\bar{\mathscr{F}}}} \right) \right)^{1/q}.$$

We get from this argument that (5.22) holds provided that

$$\tilde{N} \geq N_{1,q} = c_6 n \varepsilon^{-2} \log \frac{1}{\varepsilon} \left(\frac{N}{n} \log \frac{N}{n} \right)^{1/q}. \tag{5.24}$$

We pass now to sets $\mathscr{E}_s$ for $1 \leq s \leq l$. Again we define $\mathscr{E}_{s,k} = \{z\chi_{A_k(z)};\ z \in \mathscr{E}_s\}$ and apply the empirical method to these sets. By using the L_q and L_∞ estimates from (5.20) the proof of Lemma 3.3 shows that (5.23) is satisfied for a given s if

$$\tilde{N} \geq c_7 \theta_s^{-2} 2^s \log \bar{\bar{\mathscr{E}}}_s \left(\frac{N}{\log \bar{\bar{\mathscr{E}}}_s} \cdot \log \frac{N}{\log \bar{\bar{\mathscr{E}}}_s} \right)^{1/q} \tag{5.25}$$

and

$$\delta_s > \exp(-c_8 \theta_s^2 \tilde{N}/2^s N^{1/q}). \tag{5.26}$$

Similarly, by using the $L_{p,\infty}$ estimates instead of the L_q estimates the same argument gives that (5.23) holds for a given s if

$$\tilde{N} \geq c_7 \theta_s^{-2} \alpha (p-1)^{-2} \log \bar{\bar{\mathscr{E}}}_s \left(\frac{N}{\log \bar{\bar{\mathscr{E}}}_s} \log \frac{N}{\log \bar{\bar{\mathscr{E}}}_s} \right)^{1/p} \tag{5.27}$$

and

$$\delta_s > \exp(-c_8 \theta_s^2 (p-1)^2 \tilde{N}/2^s N^{1/q}). \tag{5.28}$$

Our choice of $\{\theta_s\}_{s=1}^l$ will be so that in particular $\theta_s \geq c_9 \varepsilon (3/2)^{s-l}$, $1 \leq s \leq l$, for some positive (small) absolute constant c_9. This requirement does not contradict the first inequality in (5.21) and allows us to choose the δ_s so that the second inequality in (5.21) holds as well as (5.28) (and thus also (5.26)) for every $1 \leq s \leq l$, provided that

$$\tilde{N} \geq N_{2,q} = c_{10} n \varepsilon^{-2} (p-1)^{-2} \left(\frac{N}{n} \right)^{1/q}. \tag{5.29}$$

We now let s_0 be defined by the equation

$$\frac{\alpha}{(p-1)^2} \left(\frac{N}{\log \bar{\bar{\mathscr{E}}}_{[s_0]}} \log \frac{N}{\log \bar{\bar{\mathscr{E}}}_{[s_0]}} \right)^{1/p} = 2^{s_0} \left(\frac{N}{\log \bar{\bar{\mathscr{E}}}_{[s_0]}} \log \frac{N}{\log \bar{\bar{\mathscr{E}}}_{[s_0]}} \right)^{1/q} \tag{5.30}$$

and define θ_s, so that for $h > 0$

$$\theta_{[s_0]+h} \geq \theta_{[s_0]} \left(\frac{3}{2} \right)^{-\frac{h}{2}(1-1/p)}, \quad \theta_{[s_0]-h} \geq \theta_{[s_0]} \left(\frac{3}{2} \right)^{-h/2q}. \tag{5.31}$$

Such a choice can be made (consistent with our previous requirement on θ_s) so that $\Sigma_{s=1}^l \theta_s \leq \varepsilon$ provided that $\theta_{[s_0]} \geq c_{11} \varepsilon \min((p-1), q^{-1})$. It is clear from (5.30) and (5.31) that if $\tilde{N}$ is such that (5.27) holds for $[s_0]$ then (5.27) holds for all $s \geq [s_0]$ and (5.25) holds for $s < [s_0]$.

In other words in order to ensure that for every $1 \leq s \leq l$ either (5.25) or (5.27) holds for our choice of $\{\theta_s\}_{s=1}^{l}$ it suffices that

$$\tilde{N} \geq c_{12}\varepsilon^{-2}\alpha(p-1)^{-2}(\max((p-1)^{-1}, q))^2 \log \bar{\bar{\mathscr{E}}}_{[s_0]} \left(\frac{N}{\log \bar{\bar{\mathscr{E}}}_{[s_0]}} \cdot \log \frac{N}{\bar{\bar{\mathscr{E}}}_{[s_0]}}\right)^{1/p}. \quad (5.32)$$

A direct computation using (5.17), (5.19) and (5.30) shows that (5.32) holds once

$$\tilde{N} \geq N_{3,q} = c_{13}\, n\varepsilon^{-2} \max\left(\frac{1}{(p-1)^2}, q^2\right) \left(\frac{q \log T}{p-1}\right)^{1/2} \left[\left(\log \frac{NT}{n(p-1)}\right)\left(\frac{NT}{n(p-1)^4}\right)\right]^r \quad (5.33)$$

where $T = T_p(X)+1$ and $r = p/(pq-q+p)$.

For $q = (p-1)^{-1}\log(NT/n(p-1))$ the right hand side of (5.14) satisfies all three conditions on $\tilde{N}$ namely (5.24), (5.29) and (5.33). □

As in section 3 we can now use iteration and get a good estimate for $N(X, \varepsilon)$ for every $X \subset L_1(0,1)$. In particular we get the following refined version of Theorem 1.1.

THEOREM 5.4. *Let X be an n-dimensional subspace of $L_1(0,1)$, let $0<\varepsilon<1/2$, let $1<p\leq 2$ and $\tau>0$. Then*

$$N(X,\varepsilon) \leq c(\tau)\, n\varepsilon^{-2}(\log(T_p(X)+1))^{1/2}(p-1)^{-3-\tau}(\log(T_p(X)/\varepsilon(p-1)))^{5/2+\tau} \quad (5.34)$$

where $c(\tau)$ is a function of τ only.

Proof. Put $\varphi(\varepsilon) = N(X,\varepsilon)/n$, $\alpha = (\log(T_p(X)+1))^{1/2}(p-1)^{-3}$, $\beta = \log((T_p(X)+1)/(p-1))$. By Lemma 5.3 we have for every integer k that whenever $\varphi(\varepsilon 2^{-k}) \geq 2^{2k}\varepsilon^{-2}$ then

$$\varphi(\varepsilon 2^{-(k-1)}) \leq c 2^{2k}\varepsilon^{-2}\alpha(\beta + \log \varphi(\varepsilon 2^{-k}))^{5/2}. \quad (5.35)$$

By a routine calculation (5.34) follows from (5.35). □

Remark. If we take $p-1 = (\log n)^{-1}$ in Theorem 5.4 we get a slightly weaker version of Theorem 5.2. Indeed for every n-dimensional subspace X of $L_1(0,1)$, $T_p(X) \leq c_1$ (for $p = 1+(\log n)^{-1}$). Hence the dependence of $N(X,\varepsilon)$ on n we obtain is $n(\log n)^{3+\tau}$ (we get $n(\log n)^3$ in Theorem 5.2).

6. The approximation of Euclidean balls

In the previous sections we treated the question of embedding arbitrary subspaces of $L_1(0,1)$ in l_1^N. In [F.L.M.] the question of embedding Euclidean spaces into arbitrary Banach spaces was considered. This section is devoted to some results concerning the case which is common to both of these studies, i.e., the embedding of l_2^n in l_1^N, or in geometric language, the approximation of Euclidean balls by zonotopes. While this case has been examined in several papers by now, and is certainly simpler than the general problems treated in the previous sections or in [F.L.M.], there are still (as we shall see) some open problems related to it.

In [Be.Mc.] the authors raise the question of how well B^n can be approximated by a sum of segments of equal length. The empirical distribution method can be easily applied also for treating this question. We start by proving a simple variant of Lemma 2.1. This variant involves Orlicz spaces so we recall the definition of such spaces. Let (Ω,μ) be a probability space and let $\Phi(t)$ be a convex increasing function on $[0,\infty)$ so that $\Phi(0)=0$ and $\lim_{t\to\infty}\Phi(t)=\infty$. We denote by $L_\Phi(\mu)$ the space of all measurable real-valued functions f on Ω so that $\int_\Omega \Phi(|f|/\lambda)\,d\mu<\infty$ for some $\lambda>0$ and put

$$\|f\|_{L_\Phi(\mu)}=\inf\left\{\lambda>0;\ \int_\Omega \Phi(|f|/\lambda)\,d\mu\leq 1\right\}. \tag{6.1}$$

We shall be concerned here only with the two functions

$$\psi_1(t)=e^t-1,\quad \psi_2(t)=e^{t^2}-1,\quad t\geq 0 \tag{6.2}$$

(besides of course the functions t^p which give rise to the $L_p(\mu)$ spaces).

LEMMA 6.1. *Let $\{g_j\}_{j=1}^N$ be independent random variables with mean 0 on some probability space (Ω,μ). Assume that $\|g_j\|_{L_{\psi_1}(\mu)}\leq M$ for some constant M and every $1\leq j\leq N$. Then for $0<\varepsilon<4M$*

$$\mathrm{Prob}\left\{\left|\sum_{j=1}^N g_j\right|\geq\varepsilon N\right\}\leq 2\exp(-\varepsilon^2N/16M^2). \tag{6.3}$$

Proof. We may clearly assume that $M=1$. For every g with $\|g_j\|_{L_{\psi_1}(\mu)}\leq 1$ we have $\int_\Omega \exp(|g|)\,d\mu\leq 2$ and hence

$$\int_\Omega |g|^k\,d\mu\leq k!\int_\Omega \exp(|g|)\,d\mu\leq 2\cdot k!,\quad 1\leq k<\infty.$$

Consequently, if in addition $\int_\Omega g\,d\mu=0$, we get for $0<\lambda\leq 1/2$

$$\int_\Omega \exp(\lambda g)\,d\mu \leq 1+\sum_{k=2}^{\infty}\int_\Omega |\lambda g|^k\,d\mu/k! \leq 1+2\sum_{k=2}^{\infty}\lambda^k \leq 1+4\lambda^2.$$

Hence

$$\operatorname{Prob}\left\{\sum_{j=1}^{N} g_j > \varepsilon N\right\}\exp(\varepsilon\lambda N) \leq \int \exp\left(\lambda\sum_{j=1}^{N} g_j\right)d\mu = \prod_{j=1}^{N}\int \exp(\lambda g_j)\,d\mu$$
$$\leq (1+4\lambda^2)^N \leq \exp(4\lambda^2 N).$$

By taking $\lambda=\varepsilon/8$ ($\leq 1/2$) we get

$$\operatorname{Prob}\left\{\sum_{j=1}^{N} g_j > \varepsilon N\right\} \leq \exp(-\varepsilon^2 N/16)$$

and this implies (6.3). □

Remark. The lemma obviously holds also if we take $M=\max\|g_j\|_{L_{\psi_2}(\mu)}$. In this form we shall actually apply it below. There are however instances which are of some geometric interest in which an L_{ψ_1} estimate on the functions involved is available while no L_{ψ_2} estimate holds. We shall discuss an example of such a situation in [B.L.M.2].

PROPOSITION 6.2. *Let B^n be the Euclidean ball and let $0<\varepsilon<1/2$. Then for*

$$N \geq cn\varepsilon^{-2}\log\varepsilon^{-1} \tag{6.4}$$

there exist N segments $\{I_j\}_{j=1}^N$ of equal length so that

$$(1-\varepsilon)B^n \subset \sum_{j=1}^{N} I_j \subset (1+\varepsilon)B^n. \tag{6.5}$$

Proof. Consider $S^{n-1}=\partial(B^n)$ with the normalized rotation invariant measure σ_n. Let X^n be the subspace of $L_1(\sigma_n)$ consisting of the (restrictions to S^{n-1} of the) linear functions. (We identify the elements of X^n with the points in $\mathbf{R}^n$.) Put

$$\beta_n = \int_{S^{n-1}} |\langle x,u\rangle|\,d\sigma_n(u), \quad x\in S^{n-1}. \tag{6.6}$$

Then $\beta_n \sim n^{-1/2}$. A simple computation also shows that for $x \in S^{n-1}$

$$\int \exp(n\langle x, u\rangle^2/4)\, d\sigma_n(u) = \int_0^{\pi/2} \cos^{n-2}\theta \exp(n \sin^2\theta/4)\, d\theta \Big/ \int_0^{\pi/2} \cos^{n-2}\theta\, d\theta \le c_1$$

for some absolute constant c_1. Hence

$$\|\beta_n^{-1}\langle x, u\rangle\|_{L_{\psi_2}(\sigma_n)} \le c_2, \quad x \in S^{n-1}. \tag{6.7}$$

We now apply Lemma 6.1 to functions of the form

$$\beta_n^{-1}|\langle x, u_j\rangle| - 1, \quad x,\ u_j \in S^{n-1}$$

and get by the reasoning of section 2 in view of (6.2), (6.3) and (6.7) that for a suitable choice of $\{u_j\}_{j=1}^N$ in S^{n-1}

$$\left|(N\beta_n)^{-1}\sum_{j=1}^N |\langle x, u_j\rangle| - 1\right| < \varepsilon, \quad x \in S^{n-1}. \tag{6.8}$$

It follows from the Hahn Banach theorem (or more precisely the separation theorem) that (6.8) implies (6.5) if we take $I_j = (N\beta_n)^{-1}[-u_j, u_j]$. All these intervals have a common length $2/N\beta_n$. $K(X)$ and (4.32) follows. □

The proof of Proposition 6.2 actually enables us to obtain a more general geometric result. Proposition 6.2 asserts that we can approximate B^n by a sum of $N = N(n, \varepsilon)$ sets each obtained from a fixed segment by a suitable rotation. It turns out that it is possible to replace the segment above by an arbitrary compact convex set.

THEOREM 6.3. *Let K be a compact convex set in $\mathbf{R}^n$ and let $0<\varepsilon<1/2$. There are a constant $r = r(K)$ and orthogonal transformations $\{U_j\}_{j=1}^N$ on $\mathbf{R}^n$, with $N \le cn\varepsilon^{-2}\log\varepsilon^{-1}$ so that*

$$(1-\varepsilon)\, rB^n \subset \frac{1}{N}\sum_{j=1}^N U_j K \subset (1+\varepsilon)\, rB^n. \tag{6.9}$$

Proof. There is no loss of generality to assume that K is symmetric with respect to the origin (otherwise replace K by $K-K$). There is also no loss of generality to assume that K has the origin as an interior point. Let $\|\ \|$ be the norm on $\mathbf{R}^n$ whose unit ball is the polar of K, i.e.,

$$\|x\| = \sup\{|\langle x, u\rangle|, u \in K\}, \quad x \in \mathbf{R}^n.$$

We shall now proceed as in the proof of Proposition 6.2. The only difference is that the L_{ψ_2} estimate we need now lies much deeper. This estimate is however available in the literature. Let O^n be the group of orthogonal transformations on $\mathbf{R}^n$ and let μ_n be the normalized Haar measure on O^n. For every $x \in \mathbf{R}^n$ we consider $\|Ux\|$ as a function on O^n. A result of Marcus and Pisier ([Ma.P.] section V2), which is based on the so called Landau-Shepp-Fernique theorem, states that there is a universal constant c_1 so that

$$\|Ux\|_{L_{\psi_2}(\mu_n)} \leq c_1 \|Ux\|_{L_1(\mu_n)}, \quad x \in \mathbf{R}^n. \tag{6.10}$$

Observe that

$$|||x||| = \int_{O^n} \|Ux\| \, d\mu_n(U) \tag{6.11}$$

is a rotation invariant norm on $\mathbf{R}^n$ and thus is a suitable multiple of the usual Euclidean norm. By applying the empirical distribution method to the functions $\|Ux\| - |||x|||$ on O^n for all x satisfying $|||x|||=1$ (to be precise for an ε net with respect to $|||\cdot|||$ of this set) we get as in section 2 (and in view of Lemma 6.1 and (6.10)) that for N as in the statement of the theorem and for suitable $\{U_j\}_{j=1}^N \subset O^n$

$$\left| \frac{1}{N} \sum_{j=1}^{N} \|U_j x\| - |||x||| \right| \leq \varepsilon |||x|||, \quad x \in \mathbf{R}^n. \tag{6.12}$$

Assertion (6.9) follows from (6.12) by a straightforward duality argument. □

Remark. For K a zonotope, Theorem 6.3 is an easy consequence of Proposition 6.2 and its proof. Thus in this case we do not have to use the result of Marcus and Pisier.

A more detailed study of questions related to Theorem 6.3 is presented in [B.L.M.2]. In this paper we determine in particular the number of Minkowski symmetrizations needed in order to pass from a general convex body in $\mathbf{R}^n$ to a body ε close to a sphere.

We turn next to the second topic of this section and this is the study whether (1.4) can be improved as far as dependence on ε is concerned. We have seen in the previous section that up to logarithmic factors the same estimate holds for general zonoids and again in (6.4) we got a similar estimate for the case of equal segments. These results make the question above more interesting from the geometric point of view.

Some facts concerning this question are contained in [Be.Mc.]. One of the observations in [Be.Mc] is the following. Assume that (6.5) holds with $I_j=[-y_j, y_j]$. Then for every choise of $(n-1)$ indices $\{j_i\}_{i=1}^{n-1}$ and every choice of signs $\{\theta_i\}_{i=1}^{n-1}$ there is on the boundary of $\Sigma_{j=1}^N I_j$ a segment of length $2\|\Sigma_{i=1}^{n-1}\theta_i y_{j_i}\|_2$. Hence by (6.5)

$$(1-\varepsilon)^2+\left\|\sum_{i=1}^{n-1}\theta_i y_{j_i}\right\|_2^2 \geq (1+\varepsilon)^2. \tag{6.13}$$

Also by (6.5) we have for every $x\in S^{n-1}$ that $\Sigma_{j=1}^N|\langle x, y_j\rangle|\geq 1-\varepsilon$ and hence we may choose the θ_i and j_i, $1\leq i\leq n$, so that

$$\left\|\sum_{i=1}^{n-1}\theta_i y_{j_i}\right\|_2 \geq (1-\varepsilon)(n-1)/N. \tag{6.14}$$

From (6.13) and (6.14) it follows that (for $\varepsilon<1/4$)

$$N\geq n/4\varepsilon^{1/2}. \tag{6.15}$$

An estimate which is better than (6.15) for a suitable range of ε ($\varepsilon\geq n^{-2}$, to be precise) is given in the next proposition.

PROPOSITION 6.4. *Assume that* (6.5) *holds for some choice of intervals* $\{I_j\}_{j=1}^N$ *not necessarily of equal length. Then for some absolute constant* $c>0$

$$N\geq cn^2/(1+n\varepsilon). \tag{6.16}$$

Proof. We assume that n is even; the case n odd follows from n even by projecting on a subspace with codimension 1. Consider again the numbers β_n defined in (6.6). The precise value of β_n is given by

$$\beta_n=\Gamma(n/2)/(\Gamma((n+1)/2)\,\Gamma(1/2)). \tag{6.17}$$

Hence, by Stirling's formula

$$\frac{\beta_{n/2}}{\beta_n}=\frac{\Gamma(n/4)\,\Gamma((n+1)/2)}{\Gamma(n/2)\,\Gamma((n+2)/4)}=\sqrt{2}\left(1+\frac{1}{4n}+O\left(\frac{1}{n^2}\right)\right). \tag{6.18}$$

We write (6.5) in the following form. For some $\{y_j\}_{j=1}^N$ in S^{n-1} and some scalars $\lambda_1\geq\lambda_2\geq\ldots\geq\lambda_N>0$,

$$\left|\beta_n^{-1}\sum_{j=1}^N\lambda_j|\langle x, y_j\rangle|-1\right|\leq\varepsilon,\quad x\in S^{n-1}. \tag{6.19}$$

By integrating (6.19) with respect to σ_n we get

$$\left|\sum_{j=1}^{N} \lambda_j - 1\right| < \varepsilon. \tag{6.20}$$

Hence (if $\varepsilon<1/2$)

$$\sum_{j=1}^{n/2} \lambda_j \geq n(1-\varepsilon)/2N \geq n/4N. \tag{6.21}$$

Let $E\subset\mathbf{R}^n$ be an $n/2$ dimensional subspace containing $\{y_j\}_{j=1}^{n/2}$ and let Q be the orthogonal projection on E.

Then for $x\in E$ with $\|x\|_2=1$ we get from (6.19)

$$\left|\beta_n^{-1}\sum_{j=1}^{N} \lambda_j|\langle x, Qy_j\rangle| - 1\right| < \varepsilon$$

and by integrating on the unit sphere of E with respect to its invariant measure it follows that

$$\beta_{n/2}\beta_n^{-1}\sum_{j=1}^{N} \lambda_j \|Qy_j\|_2 \geq 1-\varepsilon.$$

By the Cauchy Schwartz inequality

$$\left(\sum_{j=1}^{N} \lambda_j\right)\left(\sum_{j=1}^{N} \lambda_j \|Qy_j\|_2^2\right) \geq (1-\varepsilon)^2 (\beta_n/\beta_{n/2})^2. \tag{6.22}$$

Similarly, by replacing Q and E with $I-Q$ and $E^{\perp}$ respectively and noting that $(I-Q)\,y_j=0$ for $1\leq j\leq n/2$ we get

$$\left(\sum_{j=n/2+1}^{N} \lambda_j\right)\left(\sum_{j=n/2+1}^{N} \lambda_j \|(I-Q)\,y_j\|_2^2\right) \geq (1-\varepsilon)^2 (\beta_n/\beta_{n/2})^2 \tag{6.23}$$

and in particular (if $\varepsilon<1/5$ and $n\geq 3$)

$$\sum_{j=n/2+1}^{N} \lambda_j \|(I-Q)\,y_j\|_2^2 \geq 1/4. \tag{6.24}$$

By (6.20)–(6.24)

$$(1+\varepsilon)^2 \geq \left(\sum_{j=1}^{N} \lambda_j\right)^2 = \left(\sum_{j=1}^{N} \lambda_j\right)\left(\sum_{j=1}^{N} \lambda_j \|Qy_j\|_2^2 + \sum_{j=1}^{N} \lambda_j \|(I-Q)\,y_j\|_2^2\right)$$

$$\geq \left(\sum_{j=1}^{n/2} \lambda_j\right)\left(\sum_{j=1}^{N} \lambda_j \|(I-Q)y_j\|_2^2\right) + 2(1-\varepsilon)^2(\beta_n/\beta_{n/2})^2$$

$$\geq n/16N + 2(1-\varepsilon)^2(\beta_n/\beta_{n/2})^2.$$

Hence by (6.13),

$$n/16N \leq (1+\varepsilon)^2 - (1-\varepsilon)^2(1-1/2n+O(n^{-2})) = 4\varepsilon + 1/2n + O(\varepsilon+1/n)^2$$

and this implies (6.16). □

In the notation of section 1, (6.16) states that $N(B^n,\varepsilon) \geq cn^2/(1+n\varepsilon)$. We consider next the behavior of $N=N(B^n,\varepsilon)$ for a fixed n as a function of ε as $\varepsilon\to 0$. It is trivial that for $n=2$ the right value of N is given by (6.15), i.e., $N\sim c\varepsilon^{-1/2}$. For $n=3,4$ better upper bounds for N than (1.4) were obtained (using the notion of a projection body) in [Be.Mc.], namely

$$N(B^n,\varepsilon) \leq \gamma_n \varepsilon^{(1-n)/2}. \tag{6.25}$$

We establish here a lower bound.

THEOREM 6.5. *For every $n\geq 2$ there is a positive constant c_n so that*

$$N(B^n,\varepsilon) \geq c_n \varepsilon^{2(1-n)/(n+2)}. \tag{6.26}$$

Thus for example for $n=3$ we get from (6.25) and (6.26) that

$$c_3\varepsilon^{-4/5} \leq N(B^3,\varepsilon) \leq \gamma_3 \varepsilon^{-1}.$$

We do not know the exact order of magnitude of $N(B^3,\varepsilon)$.(1)

Theorem 6.5 is an immediate consequence of the following proposition.

PROPOSITION 6.6. *Let μ be a symmetric probability measure on S^{n-1} supported by N points. Put*

$$h(x) = \int_{S^{n-1}} |\langle x,y\rangle|\, d\mu(y). \tag{6.27}$$

(1) See also the note "added in proof".

Then

$$\delta \equiv \sup_{x\in S^{n-1}} h(x) - \inf_{x\in S^{n-1}} h(x) \geq \|h-\beta_n\|_{L_2(\sigma_n)} \geq c(n)\, N^{-(n+2)/2(n-1)} \tag{6.28}$$

for a suitable positive constant $c(n)$.

Here σ_n denotes, as before, the normalized rotation invariant measure on S^{n-1} and

$$\beta_n = \int_{S^{n-1}} |\langle x, y\rangle|\, d\sigma_n(x) = \int_{S^{n-1}} h(x)\, d\sigma_n(x).$$

Proof. We shall use some results on spherical harmonics. For each k let $\{Y_{k,j},\ 1\leq j\leq M(n,k)\}$ be an orthonormal basis of the spherical harmonics of degree k on S^{n-1}. By the addition theorem (see [Mu] p.10)

$$\sum_j Y_{k,j}(\xi)\, Y_{k,j}(\eta) = M(n,k)\, P_k(\langle \xi, \eta\rangle), \quad \xi, \eta \in S^{n-1} \tag{6.29}$$

where $P_k(t)$ is the Legendre polynomial of degree k and dimension n, given by Rodriques' formula (see [Mu] p. 17)

$$P_k(t) = \left(-\frac{1}{2}\right)^k \frac{\Gamma((n-1)/2)}{\Gamma(k+(n-1)/2)} (1-t^2)^{(3-n)/2} \left(\frac{d}{dt}\right)^k \{(1-t^2)^{k+(n-3)/2}\}. \tag{6.30}$$

By the Funk-Hecke formula (see [Mu] p. 20)

$$\begin{aligned}[h, Y_{k,j}] = \int_{S^{n-1}} h(x)\, Y_{k,j}(x)\, d\sigma_n(x) &= \int_{S^{n-1}} \left[\int_{S^{n-1}} |\langle x, y\rangle|\, Y_{k,j}(x)\, d\sigma_n(x)\right] d\mu(y) \\ &= \lambda_k \int_{S^{n-1}} Y_{k,j}(y)\, d\mu(y)\end{aligned} \tag{6.31}$$

where $\lambda_k=0$ for odd k and for even k (by (6.30))

$$\begin{aligned}\lambda_k &= \frac{\Gamma(n/2)}{\pi^{1/2}\Gamma((n-1)/2)} \int_{-1}^{1} |t|(1-t^2)^{(n-3)/2} P_k(t)\, dt \\ &= \frac{2^{-k+1}\Gamma(n/2)}{\pi^{1/2}\Gamma(k+(n-1)/2)} \int_0^1 t\left(\frac{d}{dt}\right)^k \{(1-t^2)^{k+(n-3)/2}\}\, dt \\ &= \frac{(-1)^{(k-2)/2}\Gamma(n/2)\,\Gamma(k-1)}{\pi^{1/2}2^{k-1}\Gamma(k/2)\,\Gamma((k+n+1)/2)} \approx k^{-n/2-1} \quad \text{for} \quad k\to\infty.\end{aligned} \tag{6.32}$$

Note that we denote the inner product in $L_2(\sigma_n)$ by $[\cdot,\cdot]$. Recall also the following identity ([Mu] p. 30, Lemma 17)

$$\frac{1-r^2}{(1+r^2-2rt)^{n/2}}=\sum_{k=0}^{\infty} M(n,k)\, r^k P_k(t), \quad |t|\leqslant 1, \quad 0\leqslant r<1. \tag{6.33}$$

By (6.29) we deduce that

$$\frac{1-r^2}{(1+r^2-2r\langle \xi,\eta\rangle)^{n/2}}=\sum_{k=0}^{\infty} r^k \sum_j Y_{k,j}(\xi)\, Y_{k,j}(\eta), \quad \xi,\eta\in S^{n-1}, \quad 0\leqslant r<1, \tag{6.34}$$

and by integrating with respect to μ we get (using also (6.31))

$$\int_{S^{n-1}} \frac{1-r^2}{(1+r^2-2r\langle \xi,\eta\rangle)^{n/2}}\, d\mu(\eta)=\sum_{k=0}^{\infty} r^k \lambda_k^{-1} \sum_j [h, Y_{k,j}]\, Y_{k,j}. \tag{6.35}$$

Since

$$h=\beta_n+\sum_{k=1}^{\infty}\sum_j [h, Y_{k,j}]\, Y_{k,j}$$

we get by the definition of δ in (6.28) and Parseval's identity that

$$\delta^2 \geqslant \sum_{k=1}^{\infty}\sum_j [h, Y_{k,j}]^2. \tag{6.36}$$

By our assumption $\mu=\Sigma_{i=1}^N a_i \delta_{\eta_i}$ where $\eta_i\in S^{n-1}$ and $a_i\geqslant 0$ with $\Sigma_{i=1}^N a_i=1$. Hence, by (6.35) and (6.36)

$$\left\| 1-\sum_{i=1}^{N} \frac{a_i(1-r^2)}{(1+r^2-2r\langle \xi,\eta_i\rangle)^{n/2}} \right\|_{L_2(\sigma_n)} \leqslant \delta \max_{k\geqslant 1} (r^k \lambda_k^{-1}). \tag{6.37}$$

Next observe that by (6.34) the square of the left hand side of (6.37) is equal to

$$\sum_{i=1}^{N}\sum_{i'=1}^{N} a_i a_{i'} \frac{1-r^4}{(1+r^4-2r^2\langle \eta_i,\eta_{i'}\rangle)^{n/2}}-1$$

which is at least

$$-1+\left(\sum_{i=1}^{N} a_i^2\right)\frac{1-r^4}{(1-r^2)^n} \geqslant N^{-1}(1-r)^{-n+1}-1.$$

We choose now r so that

$$(1-r^2)^{-n+1} = 2N. \tag{6.38}$$

Then by (6.32) and (6.37) there is a constant $c^1(n)$ so that

$$c^1(n)\,\delta \max_{k\,\text{even}}(r^k k^{1+n/2}) \geq 1. \tag{6.39}$$

By taking $k \sim (1-r^2)^{-1}$ in (6.39) and using (6.38) we deduce (6.28). □

It was pointed out recently by Linhart [Li] that it follows from results of J. Beck (presented in detail in [Be.C.]) on irregularities of distribution on the sphere that the upper estimate $N(B^n, \varepsilon) \leq c(n)\,\varepsilon^{-2}$ can be improved also for $n \geq 5$. The estimate given in [Li] is $N(B^n, \varepsilon) \leq c(n)\,\varepsilon^{-2+2/n} |\log \varepsilon|$.(1)

7. Embedding into l_p^n

In this section we treat the question of embedding subspaces X of $L_p(0,1)$ into l_p^N with small N in the case $1<p<\infty$ (the trivial case $p=2$ is excluded). As already mentioned in the introduction, the results we get show that the estimates obtained in [F.L.M.] for the case X an inner product space hold in a slightly weaker form for arbitrary X. In general the arguments here will resemble those used in sections 2–5 for the case $p=1$. Therefore we shall be somewhat more brief here and mainly emphasize those technical points which have to be treated differently for $p>1$.

We start with the definition of the subject of our study. For a finite-dimensional subspace X of $L_p(0,1)$ and $\varepsilon>0$ we put

$$N_p(X, \varepsilon) = \min\{N;\, d(X, Y) \leq 1+\varepsilon \text{ for some } Y \subset l_p^N\}. \tag{7.1}$$

The change of density lemmas we used above for $p=1$ generalize in a straightforward way to the case $p>1$. For example the lemma of Lewis (used in section 4 for $p=1$) reads for general p as follows.

(1) See also the note "added in proof".

LEMMA 7.1 [Le]. *Let X be an n-dimensional subspace of $L_p(\Omega,\mu)$, $1\leq p<\infty$. Then there is a basis $\{\varphi_i\}_{i=1}^n$ of X so that for all scalars $\{\lambda_i\}_{i=1}^n$*

$$\int_\Omega \left|\sum_{i=1}^n \lambda_i\varphi_i\right|^2 F^{p-2}\,d\mu = n^{-1}\left(\sum_{i=1}^n \lambda_i^2\right) \tag{7.2}$$

where

$$F=\left(\sum_{i=1}^n \varphi_i^2\right)^{1/2}, \quad \|F\|_p=1. \tag{7.3}$$

It follows from this lemma that if ν is the probability measure defined on Ω by $d\nu=F^p d\mu$ then the map $f\to F^{-1}f$ maps X into a subspace $\tilde X$ of $L_p(\Omega,\nu)$ which is isometric to X and which has a basis $\psi_i=\varphi_iF^{-1}$, $1\leq i\leq n$ satisfying

$$\sum_{i=1}^n \psi_i^2\equiv 1, \quad \left\|\sum_{i=1}^n \lambda_i\psi_i\right\|_2^2 = n^{-1}\sum_{i=1}^n \lambda_i^2. \tag{7.4}$$

An immediate consequence of (7.4) is the fact that for $f\in\tilde X$, $\|f\|_\infty\leq n^{1/2}\|f\|_2$. Consequently if $1\leq p<2$, $f\in\tilde X$

$$\|f\|_2^2\leq\|f\|_\infty^{2-p}\|f\|_p^p\leq n^{1-p/2}\|f\|_2^{2-p}\|f\|_p^p$$

or

$$\|f\|_2\leq n^{(2-p)/2p}\|f\|_p \tag{7.5}$$

and in conclusion

$$\|f\|_\infty\leq n^{1/p}\|f\|_p,\quad 1\leq p\leq 2,\quad \|f\|_\infty\leq n^{1/2}\|f\|_p,\quad 2\leq p,\quad f\in X. \tag{7.6}$$

As pointed out by Schechtman [Sche.3] the empirical method of section 2, applied to the expressions

$$|f(\omega)|^p-\int_\Omega |f|^p\,d\nu$$

where f ranges over an ε-net in the boundary of the unit ball of $\tilde X$ gives immediately, in view of (7.6), the following estimates.

$$N_p(X,\varepsilon)\leq cn^2\varepsilon^{-2}\log\varepsilon^{-1},\quad 1\leq p<2 \tag{7.7}$$

$$N_p(X,\varepsilon)\leq cn^{1+p/2}\varepsilon^{-2}\log\varepsilon^{-1},\quad 2<p. \tag{7.8}$$

Our aim is to improve on these estimates by entropy considerations or the iteration procedure.

The next proposition gives the relevant entropy estimates.

PROPOSITION 7.2. (i) *Let X be an n-dimensional subspace of $L_p(0,1)$ with $2<p<\infty$ and let $\varepsilon>0$. Then there is a probability space (Ω,ν) and a subspace $\tilde{X}$ of $L_p(\Omega,\nu)$ so that $d(X,\tilde{X})\leq 1+\varepsilon$ and*

$$\log E(\tilde{B}_p,\tilde{B}_\infty,t)\leq c\cdot p\log(n\varepsilon^{-1})\,nt^{-2},\quad 1\leq t\leq n^{1/2}. \tag{7.9}$$

(ii) *With X as above but for $1<p<2$ there are a ν and $\tilde{X}$ as above so that*

$$\log E(\tilde{B}_p,\tilde{B}_\infty,t)\leq c(p-1)^{-1}\log(n\varepsilon^{-1})\,nt^{-p},\quad 1\leq t\leq n^{1/p}. \tag{7.10}$$

Here $\tilde{B}_r$ denotes the unit ball of $\tilde{X}$ in the norm induced by $L_r(\Omega,\nu)$.

Proof. The proof of part (i) is by a trivial modification of arguments appearing already in section 4. If Y is a subspace of $L_p(\nu)$ for some probability space, with a basis satisfying (7.4) then by the proof of Proposition 4.6 we have for every $q>p$

$$\log((B_Y)_p,(B_Y)_q,t)\leq\log E((B_Y)_2,(B_Y)_q,t)\leq c_1qnt^{-2}. \tag{7.11}$$

By Lemma 7.1 there is a probability measure ν on $\{1,\dots,N_p(X,\varepsilon)\}$ with $\nu(\{i\})\geq 1/2N_p(X,\varepsilon)$ for every i and a subspace $\tilde{X}$ of $L_p(\nu)$ so that (7.4) holds and $d(\tilde{X},X)\leq 1+\varepsilon$. Taking $q=\log N_p(X,\varepsilon)$ in (7.11) and using (7.8) we get (7.9) (see the proof of Corollary 4.7).

The proof of part (ii) requires more work namely both a duality argument and interpolation. Let again Y be a subspace of $L_p(\nu)$ with a basis satisfying (7.4) and denote by B_r its unit ball in the $L_r(\nu)$ norm.

Fix $2<r$ and let $q>r$ and $0<\theta<1$ be such that

$$\frac{1}{r}=\frac{1-\theta}{2}+\frac{\theta}{q}.$$

Since for $f,g\in B_2$

$$\|f-g\|_r\leq\|f-g\|_2^{1-\theta}\|f-g\|_q^{\theta}\leq 2\|f-g\|_q^{\theta}$$

we get that

$$\log \bar{E}(B_2, B_r, t) \leqslant \log \bar{E}(B_2, B_q, (t/2)^{1/\theta}) \leqslant c_2\, qn(t/2)^{-2r(q-2)/q(r-2)}. \tag{7.12}$$

By taking $q = r\log n$ (and remembering that only $t \leqslant n$ count) we get that

$$\log \bar{E}(B_2, B_r, t) \leqslant c_3\, rn \log n (t/c_3)^{-2r/(r-2)}. \tag{7.13}$$

We shall use (7.13) for estimating $E(B_p, B_2, t)$. For $k=0,1,\ldots$, let $\mathscr{E}_k$ be a maximal subset of B_p with respect to the condition $\|f-g\|_2 > 8^k t$ for every $f \neq g$ in $\mathscr{E}_k$ (if $8^k t > n^{(2-p)/2p}$ we take $\mathscr{E}_k = \{0\}$). Clearly

$$\bar{\bar{\mathscr{E}}}_k \geqslant \bar{E}(B_p, B_2, 8^k t).$$

It is also obvious that for every k there is an $h_k \in B_p$ so that if

$$\mathscr{F}_k = \{f;\, f \in \mathscr{E}_k,\ \|f-h_k\|_2 \leqslant 8^{k+1} t\}$$

then

$$\bar{\bar{\mathscr{F}}}_k \geqslant \bar{\bar{\mathscr{E}}}_k / \bar{E}(B_p, B_2, 8^{k+1} t) \geqslant \bar{E}(B_p, B_2, 8^k t)/\bar{E}(B_p, B_2, 8^{k+1} t). \tag{7.14}$$

Let $\mathscr{G}_k = \{(f-h_k)/8^{k+1} t,\ f \in \mathscr{F}_k\}$. Then

$$\|g\|_2 \leqslant 1, \quad \|g\|_p \leqslant 2/8^{k+1} t, \quad g \in \mathscr{G}_k, \quad \|g-g'\|_2 > 1/8, \quad g \neq g' \in \mathscr{G}_k.$$

Hence for $g \neq g'$ in $\mathscr{G}_k$ we have, with p' denoting the conjugate exponent to p,

$$8^{-2} \leqslant \|g-g'\|_2^2 \leqslant \|g-g'\|_p \|g-g'\|_{p'} \leqslant 4\|g-g'\|_{p'}/8^{k+1} t$$

or $\|g-g'\|_{p'} \geqslant 2 \cdot 8^{k-2} t$. Consequently

$$\log \bar{E}(B_2, B_{p'}, 8^{k-2} t) \geqslant \log \bar{\bar{\mathscr{G}}}_k = \log \bar{\bar{\mathscr{F}}}_k \geqslant \log \bar{E}(B_p, B_2, 8^k t) - \log \bar{E}(B_p, B_2, 8^{k+1} t).$$

By summing this inequality over all k and using (7.13) we get that for a suitable universal c_4

$$\log \bar{E}(B_p, B_2, t) \leqslant c_4 (p-1)^{-1} n \log n \cdot (t/c_4)^{-2p/(2-p)}. \tag{7.15}$$

Since for all $1 < \lambda < t$

$$\log E(B_p, B_q, t) \leqslant \log E(B_p, B_2, \lambda) + \log E(B_2, B_q, t/\lambda) \tag{7.16}$$

we get by (7.7), (7.11) and (7.15) and by taking a suitable λ, $q=\log N_p(X,\varepsilon)$ and $\tilde{X}$ a suitable subspace of $L_1(\nu)$ where ν is probability measure on $\{1,2,\ldots,N_p(X,\varepsilon)\}$ that (7.10) holds. □

Remark. If we take $q=(4r/(r-2))\log s$ in (7.12) (which can be done for $s>\exp((r-2)/4)$) and then follow the argument above up to (7.16) we get that for $1<p<2<q<\infty$ and say $qp/(q-p)\geq 1$ that

$$\log E(B_p, B_q, t) \leq cn\log t\cdot t^{-pq/(q-p)}, \quad t>\exp(q+(p-1)^{-1}). \tag{7.17}$$

This is a very precise estimate for a general finite dimensional subspace of $L_2(\mu)$ in a "good position" (i.e., with a basis satisfying (7.4)). In fact, Schütt computed in [Schu] the left hand side of (7.17) for the special case of $X=l_2^n$ (i.e., μ the usual probability measure on $\{1,\ldots,n\}$ and X being the entire $L_2(\mu)$), and found that

$$\log E(B_p, B_q, t) \sim n(\log t)\, t^{-pq/(q-p)}.$$

We are now ready to prove sharper estimates on $N_p(X,\varepsilon)$ than those given by (7.7) and (7.8). We start with the case $p>2$.

THEOREM 7.3 *Let $2<p<\infty$ and let $0<\varepsilon<1/2$. Then there is a constant $c(p,\varepsilon)$ so that for every n-dimensional subspace X of $L_p(0,1)$*

$$N_p(X,\varepsilon) \leq c(p,\varepsilon)\, n^{p/2}\log n. \tag{7.18}$$

Proof. By Lemma 7.1 and Proposition 7.2 there is no loss of generality to assume that for some probability space (Ω,ν), $X\subset L_p(\Omega,\nu)$, $\|f\|_\infty\leq n^{1/2}\|f\|_p$, $f\in X$ and that (7.9) holds for $E(B_p, B_\infty, t)$.

Let $\mathcal{F}$ be an ε-net on the boundary of B_p (=the unit ball of X) with $\log\bar{\bar{\mathcal{F}}}\leq 4n\log\varepsilon^{-1}$. For $k=1,2,\ldots,l=[\log n^{1/2}/\log(1+\varepsilon)]+1$ let $\mathcal{A}_k\subset B_p$ be such that

$$B_p\subset \bigcup_{g\in\mathcal{A}_k}(g+\varepsilon(1+\varepsilon)^k/3\cdot B_\infty)$$

and

$$\log\bar{\bar{\mathcal{A}}}_k\leq c_1 p\log(n\varepsilon^{-1})\, n\varepsilon^{-2}(1+\varepsilon)^{-2k}. \tag{7.19}$$

For every $x\in\mathscr{F}$ and $1\leq k\leq l$ let $f_{k,x}\in\mathscr{A}_k$ satisfy $\|x-f_{k,x}\|_\infty\leq\varepsilon(1+\varepsilon)^k/3$. Put

$$C_{k,x}=\{\omega;|f_{k,x}(\omega)|\geq(1+\varepsilon)^{k-1}\}$$

$$D_{k,x}=C_{k,x}\sim\bigcup_{h>k}C_{h,x},\quad D_{0,x}=\Omega\sim\bigcup_{k\geq 1}C_{k,x}$$

and

$$\tilde{x}=x\cdot\chi_{D_{0,x}}+\sum_{k=1}^{l}(1+\varepsilon)^k\chi_{D_{k,x}}. \tag{7.20}$$

Note that if $\omega\in C_{k,x}$, $k\geq 1$, then

$$|x(\omega)|\geq(1+\varepsilon)^{k-1}-\varepsilon(1+\varepsilon)^k/3>(1+\varepsilon)^{k-2}$$

while if $\omega\notin C_{k,x}$

$$|x(\omega)|\leq(1+\varepsilon)^{k-1}+\varepsilon(1+\varepsilon)^k/3<(1+\varepsilon)^{k+1}.$$

Hence for every $\omega\in D_{k,x}$, $k\geq 1$,

$$(1+\varepsilon)^{k-2}\leq|x(\omega)|\leq(1+\varepsilon)^{k+2}$$

while for $\omega\in D_{0,x}$, $|x(\omega)|\leq(1+\varepsilon)^2$. It follows that

$$(1+\varepsilon)^{-2}\leq|\tilde{x}(\omega)|/|x(\omega)|\leq(1+\varepsilon)^2,\quad x\in\mathscr{F},\quad \omega\in\Omega. \tag{7.21}$$

As usual, we prove the theorem by using the empirical distribution method. We shall work with the functions

$$|\tilde{x}(\omega)|^p-\int_\Omega|\tilde{x}(\omega)|^p\,d\nu,\quad x\in\mathscr{F}.$$

In view of (7.21) it follows that if we show that for suitable $\{\omega_i\}_{i=1}^N$

$$\left|N^{-1}\sum_{i=1}^{N}|\tilde{x}(\omega_i)|^p-\int_\Omega|\tilde{x}(\omega)|^p\,d\nu\right|<\varepsilon,\quad x\in\mathscr{F}, \tag{7.22}$$

then a similar statements holds if $\tilde{x}$ is replaced by x provided we replace ε by $4p\varepsilon$. In order to prove (7.22) it suffices to verify that

$$\left|N^{-1}\sum_{i=1}^{N}|x(\omega_i)|^p\chi_{D_{0,x}}(\omega_i)-\int_{D_{0,x}}|x(\omega)|^p\,d\nu\right|<\varepsilon_0,\quad x\in\mathscr{F} \tag{7.23}$$

$$\left| N^{-1}(1+\varepsilon)^{pk}\left(\sum_{i=1}^{N}\chi_{D_{k,x}}(\omega_i)-N\nu(D_{k,x})\right)\right| < \varepsilon_k, \quad 1\le k\le l, \quad x\in \mathcal{F} \tag{7.24}$$

where

$$\sum_{k=0}^{l}\varepsilon_k \le \varepsilon. \tag{7.25}$$

Let $\mathcal{B}_k$, $1\le k\le l$ be the collection of all sets of form $D_{k,x}$. It follows from the definition that

$$\log \bar{\bar{\mathcal{B}}}_k \le \sum_{h=k}^{l} \log \bar{\bar{\mathcal{A}}}_h \le c_2 p \log(n\varepsilon^{-1})\, n\varepsilon^{-3}(1+\varepsilon)^{-2k}. \tag{7.26}$$

In view of Lemma 2.1, in order to ensure that (7.22) holds for some $\{\omega_i\}_{i=1}^N$, it is enough to have that

$$\bar{\bar{\mathcal{F}}}\exp(-\varepsilon_0^2 N/8(1+\varepsilon)^{2p})+\sum_{k=1}^{l}\bar{\bar{\mathcal{B}}}_k \exp(-\varepsilon_k^2 N/8(1+\varepsilon)^{pk}) < \tfrac{1}{2}. \tag{7.27}$$

In view of (7.26) it is readily checked that (7.25) and (7.27) hold with $N=c(p,\varepsilon)\, n^{p/2}\log n$ if we take $\varepsilon_0=c_3\varepsilon$, $\varepsilon_l=c_3\varepsilon/(p-2)$, $\varepsilon_k=(1+\varepsilon)^{(2-p)(l-k)/2}\cdot\varepsilon_l$, $1\le k<l$. □

Remark. The proof above gives that $c(p,\varepsilon)$ in (7.18) can be taken to be $c(p-2)^{-2}\varepsilon^{-5}|\log\varepsilon|\exp cp$. Some variants of the argument in the proof above give formulas for $N_p(X,\varepsilon)$ which improve on (7.8) as far as the dependence on n is concerned without changing the dependence on ε. We do not go into details since for $p>1$ the dependence on ε seems to be less interesting than in the case $p=1$. In the application presented in the next section only the isomorphic nature of the result (i.e., the dependence on n) is of interest.

In a completely similar manner we get, by using (7.10), the following

THEOREM 7.4. *Let $1<p<2$ and let $\varepsilon>0$. Then there is a constant $c(p,\varepsilon)$ so that for every n-dimensional subspace X of $L_p(0,1)$*

$$N_p(X,\varepsilon)\le c(p,\varepsilon)\, n(\log n)^3. \tag{7.28}$$

As in section 3 we can obtain also for $1<p<2$ a better estimate of $N_p(X,\varepsilon)$ (as far as dependence on $n=\dim X$ is concerned) if we make an extra assumption on the type of

X. The proof will be by iteration and the argument is identical to that used in section 3. We just formulate here the result.

PROPOSITION 7.5. *Assume that* $1<p<r\leq 2$ *and that* X *is an* n*-dimensional subspace of* l_p^N. *Let* $0<\varepsilon<1/2$ *and assume that* $N\geq n\varepsilon^{-2}$. *Then there is a subspace* $\tilde{X}$ *of* $l_p^{\tilde{N}}$ *with* $d(X,\tilde{X})\leq 1+\varepsilon$ *provided that*

$$\tilde{N}\geq cn^{1-p/r}N^{p/r}\left(\log\frac{N}{n}\right)^{p/r}\varepsilon^{-2}\log\varepsilon^{-1}(r-p)^{-(p+3)}T_r(X)^p. \tag{7.29}$$

COROLLARY 7.6. *For every* $\varrho>2$, $\lambda\geq 1$, $1<p<r\leq 2$ *there is a* $c(p,r,\varrho,\lambda)$ *so that whenever* X *is an* n*-dimensional subspace of* $L_p(0,1)$

$$N_p(X,\varepsilon)\leq c(p,r,\varrho,T_r(X))\,n\varepsilon^{-\varrho r/(r-p)}. \tag{7.30}$$

We conclude this section with a remark concerning the role of the l_p spaces in the discussion of this section. One may ask whether Theorems 7.3 and 7.4 or Corollary 7.6 are special cases of a general fact concerning symmetric structure. Assume that Z is a Banach space with a symmetric basis $\{e_i\}_{i=1}^{\infty}$ and denote $Z_n=\operatorname{span}\{e_i\}_{i=1}^{n}$. For every finite dimensional subspace X of Z and every $\varepsilon>0$ we define $N_Z(X,\varepsilon)$ in an obvious analogy to (7.1). Is it true that any estimate for $N_Z(l_2^n,\varepsilon)$ is valid (perhaps in a slightly weaker form) for an arbitrary n-dimensional subspace X of Z?

The answer to this question is negative. Consider for $0<\eta<1$ the space $Z(\eta)$ having a basis $\{e_i\}_{i=1}^{\infty}$ with

$$\left\|\sum_{i=1}^{\infty}\lambda_i e_i\right\|=\max\left\{\left(\sum_{i=1}^{\infty}\lambda_i^2\right)^{1/2},\ \eta\sum_{i=1}^{\infty}|\lambda_i|\right\}.$$

The space $Z(\eta)$ is isomorphic to l_1 and its cotype 2 constant is bounded by a constant independent of η. Hence by the results of [F.L.M.] there is a $c(\varepsilon)$ for all $\varepsilon>0$ so that $N_{Z(\eta)}(l_2^n,\varepsilon)\leq c(\varepsilon)\,n$. However $l_1^2\subset Z(\eta)$ for all $\eta>0$ and

$$\lim_{\eta\downarrow 0}N_{Z(\eta)}\left(l_1^2,\frac{1}{10}\right)=\infty.$$

It is also easy to construct a fixed space Z of cotype 2 with a symmetric basis and subspaces $X_n\subset Z$ with $\dim X_n=n$ so that, e.g., $N_Z(X_n,1)\geq\exp(n)$ in spite of the fact that, by [F.L.M.], $N_Z(l_2^n,\varepsilon)\leq c(\varepsilon)\,n$ for all $\varepsilon>0$.

8. Complemented subspaces

Let X be an n-dimensional subspace of $L_p(0,1)$ on which there is a linear projection P of norm λ. We are interested to embed X in l_p^N with N small so that on the image of X in l_p^N there will also be a projection with essentially the same norm. For $X=l_2^n$, $1<p<\infty$, it is easy to see and well known that we may take (ignoring ε for the moment) $N\approx n^{p^*/2}$ where $p^*=\max(p,p')$, p' being the conjugate exponent to p. This N is large enough to ensure that l_2^n embeds in both l_p^N and $l_{p'}^N$ (the complementation is obtained in this case automatically, any inner product space contained in $L_r(0,1)$ with $2\leq r<\infty$ is well complemented). In this section we prove, by using the empirical method, that a similar result holds for a general n-dimensional subspace of $L_p(0,1)$ (though, in the general case, complementation is no longer automatic). The first result in this direction was proved in [Sche.2]; it involved a special family of subspaces of $L_p(0,1)$ (the so called X_p^ω spaces). The estimate obtained in [Sche.2] does not, however, imply Theorem 8.1 below even in this special case.

In contrast to the situation for subspaces of l_p^n of which there are known many interesting examples (especially if $1\leq p<2$) the complemented subspaces of l_p^n are a rather restrictive class. In some (non-precise) sense they seem to be just "mixtures" of l_p^m spaces and inner product spaces. Therefore, while the main interest of the results of the preceding sections is in the existence of economic embeddings of concrete examples, the main application of the result in the complemented case seems to be in analyzing further the structure of complemented subspaces of l_p^n. In some instances it is possible to prove, using Theorem 8.1 below, that a complemented subspace of l_p^n has to be isomorphic to l_p^m for a suitable m. The second part of this section contains a specific result of this nature.

THEOREM 8.1. *Let X be an n-dimensional subspace of $L_p(\Omega,\mu)$ for some probability space (Ω,μ) and $1<p<\infty$. Assume that there is a bounded linear projection P from $L_p(\Omega,\mu)$ onto X with $\|P\|=\lambda$. Then for $0<\varepsilon<(c(p)\lambda)^{-1}$ and*

$$N=[c(p,\varepsilon)\, n^{p^*/2}(\log n)^{2p^*-1}] \tag{8.1}$$

where $p^*=\max(p,p')$, there is a subspace $\tilde{X}$ of l_p^N so that $d(X,\tilde{X})\leq 1+\varepsilon$ and so that there is a projection Q from l_p^N onto $\tilde{X}$ with

$$\|Q\|\leq\lambda(1+\varepsilon c(p)\lambda). \tag{8.2}$$

Proof. Our strategy for getting a complemented embedding of X in l_p^N is the following. Let X and P be as in the statement of the theorem and let $Y=P^*X^*\subset L_{p'}(\mu)$.

Assume that N_0 and the points $\{\omega_j\}_{j=1}^{N_0}$ in Ω are such that

$$\left| N_0^{-1} \sum_{j=1}^{N_0} |f(\omega_j)|^p - \int_\Omega |f|^p \, d\mu \right| < \varepsilon, \quad f \in X, \quad \|f\|_p = 1 \tag{8.3}$$

$$\left| N_0^{-1} \sum_{j=1}^{N_0} |g(\omega_j)|^{p'} - \int_\Omega |g|^{p'} \, d\mu \right| < \varepsilon, \quad g \in Y, \quad \|g\|_{p'} = 1 \tag{8.4}$$

$$\left| N_0^{-1} \sum_{j=1}^{N_0} f(\omega_j)\, g(\omega_j) - \int_\Omega fg \, d\mu \right| < \varepsilon, \quad f \in X, \quad g \in Y, \quad \|f\|_p = \|g\|_{p'} = 1. \tag{8.5}$$

Let $T: X \to l_p^{N_0}$ be defined by $Tf = (f(\omega_1), \ldots, f(\omega_{N_0}))$ (it is convenient in our context to consider $l_p^{N_0}$ as an L_p space over the set $\{1, 2, \ldots, N_0\}$ with each point having measure N_0^{-1}) and let $S: Y \to l_p^{N_0}$ be defined similarly. Let $i: X \to L_p(\mu)$ be the natural embedding. Then $Pi = I_X$=the identity operator of X, thus $i^*P^* = i_{|Y}^* \cdot P^* = I_{X^*}$ and $P^{**}(i_{|Y}^*)^* = I_X$ (we identify canonically X with X^{**}). The requirement (8.5) means that

$$|\langle Tf, Sg \rangle - \langle f, g \rangle| < \varepsilon, \quad f \in X, \quad g \in Y, \quad \|f\|_p = \|g\|_{p'} = 1$$

i.e., that $\|S^*T - (i_{|Y}^*)^*\| \leq \varepsilon$. Hence

$$\|P^{**}S^*T - I_X\| \leq \|P\| \cdot \varepsilon = \lambda \varepsilon.$$

It follows that

$$Q = T(P^{**}S^*T)^{-1}P^{**}S^*$$

is a projection of norm at most $\lambda(1+\varepsilon)^2/(1-\lambda\varepsilon)$ from $l_p^{N_0}$ on TX. Since $d(TX, X) \leq 1+\varepsilon$ this will prove the theorem once we verify that we can take as N_0 the number appearing in the right hand side of (8.1).

It is clear, in view of Lemma 2.5, that for the argument above it suffices to require that (8.3)–(8.5) hold just for all x and y belonging to an ε-net in the surface of the unit balls of X, respectively Y.

We shall use the procedure described above for obtaining a complemented embedding twice. In the first step we shall use the method of section 2 to embed X complementably in some $l_p^{N_1}$ where N_1 is a polynomial in n which is however larger than the right hand side of (8.1). This step is needed in order to insure that for X (resp. Y) we have at our disposal the sharp entropy estimates (7.9) and (7.10). Once we have these estimates we use the arguments of section 7 in order to obtain a complemented embedding into l_p^N for an N satisfying (8.1).

We examine the first step. In order to be specific, we assume that $1<p<2$ and then $p^*=p'>2$. For performing it we have first to make a change of density procedure so as to get good L_∞ estimates for the elements in the unit balls of X and $Y=P^*X^*$. By the Lemma 7.1 there is a function $F\geq 0$ on Ω with $\|F\|_{L_p(\mu)}=1$ so that if $d\nu=F^p\,d\mu$ and $J_F:X\to L_p(\nu)$ is defined by $J_Fx=F^{-1}x$ then J_F is an isometry and $\|f\|_\infty\leq n^{1/p}\|f\|_p$, $f\in J_FX$. As we have already done in previous sections, we replace F by $G_1=((F^p+1)/2)^{1/p}$. Then J_{G_1} is an isometry from all of $L_p(\mu)$ onto $L_p(\sigma_1)$ where $d\sigma_1=G_1^p\,d\mu$ and

$$G_1(\omega)\geq 2^{-1/p},\quad \omega\in\Omega;\quad \|f\|_\infty\leq(2n)^{1/p}\|f\|_p,\quad f\in J_{G_1}X. \tag{8.6}$$

Note that $P_1=J_{G_1}PJ_{G_1}^{-1}$ is a projection of norm λ from $L_p(\sigma_1)$ onto $J_{G_1}X$. Under this transformation the space Y is replaced by $P_1^*(J_{G_1}X)^*=J_{G_1^{p-1}}Y$. By applying Lemma 7.1 to the subspace $J_{G_1^{p-1}}Y$ of $L_{p'}(\sigma_1)$ we get in a similar way a function G_2 on Ω so that

$$\|G_2\|_{L_p(\sigma_1)}=\|G_2^{p-1}\|_{L_{p'}(\sigma_1)}=1 \tag{8.7}$$

and so that $J_{G_2}:L_p(\sigma_1)\to L_p(\sigma_2)$ is an isometry onto where $d\sigma_2=G_2^p\,d\sigma_1$. Moreover, we have

$$G_2(\omega)\geq 2^{-1/p},\quad \omega\in\Omega,\quad \|g\|_\infty\leq(2n)^{1/2}\|g\|_{p'}, \tag{8.8}$$

for $g\in J_{G_2^{p-1}}(J_{G_1^{p-1}}Y)=J_{(G_2G_1)^{p-1}}Y\subset L_{p'}(\sigma_2)$. Also for $f=J_{G_2G_1}x\in J_{G_2G_1}X$ we get

$$\begin{aligned}\|f\|_\infty&\leq\|G_2^{-1}\|_\infty\|J_{G_1}x\|_\infty\leq\|G_2^{-1}\|_\infty(2n)^{1/p}\|J_{G_1}x\|_{L_p(\sigma_1)}\\&=\|G_2^{-1}\|_\infty(2n)^{1/p}\|f\|_{L_p(\sigma_2)}\leq(4n)^{1/p}\|f\|_{L_p(\sigma_2)}.\end{aligned} \tag{8.9}$$

In conclusion, we see that by replacing X and Y respectively by $J_{G_2G_1}X$ and $J_{(G_2G_1)^{p-1}}Y$ there is no loss of generality to assume that for $f\in X$ with $\|f\|_p=1$ and $g\in Y$ with $\|g\|_{p'}=1$

$$\|f^p\|_\infty\leq 4n,\quad \|g^{p'}\|_\infty\leq(4n)^{p'/2},\quad \|fg\|_\infty\leq(4n)^{1/2+1/p}. \tag{8.10}$$

By (8.10) and the method of section 2 we get that (8.3)–(8.5) hold for suitable $\{\omega_j\}_{j=1}^{N_0}$ if

$$N_0\geq c_1\,n\varepsilon^{-2}\log\varepsilon^{-1}\cdot(4n)^{p'/2}.$$

We now apply Proposition 7.2 to obtain, via another change of density, good estimates on the entropy numbers. Since we work with densities which are bounded from below an argument similar to the one we did above (and using the definition of the entropy numbers) shows that we can make a change of density which is good for both X and Y. In other words there is no loss of generality to assume that for suitable constants c_2 and $c_3(p)$ we have

$$\|f\|_\infty \leqslant c_2 n^{1/p}\|f\|_p, \quad f \in X; \quad \|g\|_\infty \leqslant c_2 n^{1/2}\|g\|_{p'}, \quad g \in Y \tag{8.11}$$

$$\log E(B_X, B_\infty, t) \leqslant c_3(p) \log n\varepsilon^{-1} \cdot nt^{-p}, \quad 1 \leqslant t \leqslant c_2 n^{1/p} \tag{8.12}$$

$$\log E(B_Y, B_\infty, t) \leqslant c_3(p) \log n\varepsilon^{-1} \cdot nt^{-2}, \quad 1 \leqslant t \leqslant c_2 n^{1/2}. \tag{8.12}$$

From the proof of Theorem 7.3 it follows now that if we take as N_0 the right hand side of (8.1) then (8.3) and (8.4) hold for suitable $\{\omega_j\}_{j=1}^{N_0}$ in Ω and all x, resp. y, in a suitable ε-net $\mathcal{F}_X$, resp. $\mathcal{F}_Y$, in the boundary of the unit balls of X, resp. Y. By an argument which is similar to the proof of Theorem 7.3 we shall show next that also (8.5) holds for this N_0 and this will conclude the proof.

As in the proof of Theorem 7.3, we can replace the functions $x \in \mathcal{F}_X$ and $y \in \mathcal{F}_Y$ by the functions of the form

$$\tilde{x} = x\chi_{D_{0,x}} + \sum_{k=1}^{l_1} (1+\varepsilon)^k (\chi_{D'_{k,x}} - \chi_{D''_{k,x}}) = \sum_{k=0}^{l_1} \varphi_{k,x}$$

$$\tilde{y} = y\chi_{D_{0,y}} + \sum_{h=1}^{l_2} (1+\varepsilon)^h (\chi_{D'_{h,y}} - \chi_{D''_{h,y}}) = \sum_{h=0}^{l_2} \psi_{h,y}$$

where

$$l_1 = [\log(c_2 n^{1/p})/\log(1+\varepsilon)], \quad l_2 = [\log(c_2 n^{1/2})/\log(1+\varepsilon)]$$

and for each $x \in \mathcal{F}_X$, the sets $D_{0,x}, D'_{k,x}, D''_{k,x}$, $1 \leqslant k \leqslant l_1$ form a decomposition of Ω into disjoint sets, a similar statement holding for $y \in \mathcal{F}_Y$. In the present context we have to take into consideration also the signs of x and y and this is the reason for the appearance of $D'_{k,x}$ and $D''_{k,x}$ rather than just $D_{k,x}$ as in the proof of 7.3.

As in the proof of 7.3 (see (7.26)) we get from the construction and (8.12) and (8.13) that if we put $\mathcal{B}_k = \{\varphi_{k,x}\}_{x \in \mathcal{F}_X}$ and $\mathcal{C}_h = \{\psi_{h,y}\}_{y \in \mathcal{F}_Y}$ then

$$\log \bar{\bar{\mathcal{B}}}_k \leqslant c_4(p, \varepsilon)\, n \log n/(1+\varepsilon)^{kp}, \quad \log \bar{\bar{\mathcal{C}}}_h \leqslant c_4(p, \varepsilon)\, n \log n/(1+\varepsilon)^{2h}. \tag{8.14}$$

We consider the maps T and S defined after (8.5) with N_0 being the N given in (8.1). (Obviously T and S are defined not just on X and Y but on all the functions on Ω.) As mentioned above the argument of the proof of Theorem 7.3 and (8.1) ensure that we can assume that

$$\sum_{k=0}^{l_1} \big| \|\varphi_{k,x}\|_{L_p(\Omega)}^p - \|T\varphi_{k,x}\|_{l_p^N}^p \big| \leq \varepsilon, \quad x \in \mathcal{F}_X$$

and that a similar statements holds for $y \in \mathcal{F}_Y$. Our goal is to show that we can ensure also that

$$|\langle T\tilde{x}, S\tilde{y}\rangle - \langle \tilde{x}, \tilde{y}\rangle| \leq \sum_{k=0}^{l_1}\sum_{h=0}^{l_2} |\langle T\varphi_{k,x}, S\psi_{h,y}\rangle - \langle \varphi_{k,x}, \psi_{h,y}\rangle| < \varepsilon \tag{8.15}$$

for all $x \in \mathcal{F}_X$ and $y \in \mathcal{F}_Y$.

Note that for $k \geq 1$, $|\varphi_{k,x}|$ is $(1+\varepsilon)^k$ times the characteristic function of a set. Because of the special form of T, $|T\varphi_{k,x}|$ is also $(1+\varepsilon)^k$ times the characteristic function of a set with essentially the same measure. The same holds for $|\psi_{h,y}|$ and $|S\psi_{h,y}|$. Our strategy for verifying (8.15) is to show that the terms in the sum of (8.15) for which kp is far from hp' are small because $|\langle \varphi_{k,x}, \psi_{h,y}\rangle|$ are small for these indices, while the rest of the terms can be handled by the estimates provided by the empirical distribution theory (i.e., Lemma 2.1). In view of (2.4) and (8.14) the h and k for which we can apply the empirical distribution estimate are those for which

$$N/(1+\varepsilon)^{h+k} > c_5(p,\varepsilon)\, n \log n \max((1+\varepsilon)^{-kp}, (1+\varepsilon)^{-2h}). \tag{8.16}$$

Let α_n be such that

$$(1+\varepsilon)^{\alpha_n} = (\log n)^{2p/(p-1)},$$

and consider h and k for which

$$2h < kp < hp' + \alpha_n. \tag{8.17}$$

In this case (8.16) becomes

$$N > c_5(p,\varepsilon)\, n \log n (1+\varepsilon)^{k-h}$$

but

$$(1+\varepsilon)^{k-h} < (1+\varepsilon)^{h(p'/p-1)}(1+\varepsilon)^{\alpha_n/p} \leq (c_2 n^{1/2})^{p'/p-1} \log n^{2/(p-1)}$$

and thus since $1+\frac{1}{2}(p'/p-1)=p'/2$, (8.1) ensures that (8.16) holds for indices satisfying (8.17). Another case for which (8.16) is easily seen to hold is

$$kp<2h, \quad h\leq \log\log n. \tag{8.18}$$

For the h, k for which (8.17) and (8.18) fail we will show that the terms $|\langle \varphi_{k,x}, \psi_{h,y}\rangle|$ contribute only a negligible amount to (8.15) (and similarly $|\langle T\varphi_{k,x}, S\psi_{h,y}\rangle|$). Recall that $\int_\Omega |\varphi_{k,x}|^p\,dv\leq 2$ and $\int_\Omega |\psi_{h,y}|^{p'}\,dv\leq 2$, hence for all x, y, k and h

$$|\langle \varphi_{k,x}, \psi_{h,y}\rangle|\leq 2(1+\varepsilon)^{h+k}\min((1+\varepsilon)^{-kp},(1+\varepsilon)^{-hp'}). \tag{8.19}$$

Hence

$$\begin{aligned}\sum_{kp>hp'+\alpha_n}\sum |\langle \varphi_{k,x}, \psi_{h,y}\rangle| &\leq c_6(p,\varepsilon)\sum_{h=0}^{l_1}(1+\varepsilon)^{h+(hp'+\alpha_n)(1-p)/p}\\ &= c_6(p,\varepsilon)\sum_{h=0}^{l_1}(1+\varepsilon)^{\alpha_n(1-p)/p}\\ &= c_6(p,\varepsilon)\, l_1/(\log n)^2\leq c_7(p,\varepsilon)/\log n.\end{aligned} \tag{8.20}$$

We also have

$$\begin{aligned}\sum_{\substack{kp<2h\\ h>\log\log n}}\sum |\langle \varphi_{k,x}, \psi_{h,y}\rangle| &\leq 2\varepsilon^{-1}\sum_{h=\log\log n}^{l_1}(1+\varepsilon)^{h(2-p')/p}\\ &\leq 2\varepsilon^{-1}(\log n)^{-c_8(p,\varepsilon)}\end{aligned} \tag{8.21}$$

for some $c_8(p,\varepsilon)>0$.

(8.20) and (8.21) complete the verification of (8.15) and the proof is finally concluded. □

We apply Theorem 8.1. to prove a result in the spirit of [B.Tz.1].

PROPOSITION 8.2. *Assume that* $l_p^n=X\oplus Y$ *with* $m=\dim Y<n^{2/p^*-\varepsilon}$ *for some* $1<p<\infty$ *and* $\varepsilon>0$ *(as always* $p^*=\max(p,p')$*). Then*

$$d(X, l_p^{n-m})\leq c(\varepsilon,\lambda,M,p) \tag{8.22}$$

where λ is norm of the projection of l_p^n on Y which maps X to 0 and M is the basis constant of Y (i.e., the inf *of the basis constants of all possible bases of Y).*

Before we prove Proposition 8.2. we state an immediate corollary of it for spaces arising in harmonic analysis. Recall that for a subset Λ of the integers $L_{p,\Lambda}$ is the subspace of $L_p[0, 2\pi]$ of all the functions whose Fourier coefficients vanish outside Λ.

PROPOSITION 8.3. *Let $1<p<\infty, \varepsilon>0$ and $\Lambda\subset\{1, \ldots, n\}$ satisfy $\bar{\bar{\Lambda}}<n^{2/p^*-\varepsilon}$. Then*

$$d(L_{p,\{1,\ldots,n\}\setminus\Lambda}, l_p^{n-\bar{\bar{\Lambda}}}) \leq c(\varepsilon, \lambda, p), \tag{8.23}$$

where λ is the norm of the orthogonal projection from $L_p(0, 2\pi)$ onto $L_{p,\Lambda}$.

Proof. By a classical result $d(L_{p,\{1,\ldots,n\}}, l_p^n)\leq c_1(p)$ (see [Zy], Chapter X, Theorem (7.5)). Another classical fact, a theorem of M. Riesz, asserts that the characters in their natural order form a Schauder basis of $L_p(0, 2\pi)$. Hence the basis constant of $L_{p,\Lambda}$ is bounded by a number depending only on p. These two classical facts show that (8.23) is a consequence of (8.22). □

We turn to the proof of Proposition 8.2 which follows some arguments of [B.Tz.1]. The proof consists of two steps.

Step 1. For fixed Y with $\dim Y=m<n^{2/p^*-\varepsilon}$ the relation $l_p^n=Y\oplus X$ has a "unique" solution. In other words, if $\eta_i=d(l_p^n, Y\oplus X_i)$, $i=1,2$, then $d(X_1, X_2)\leq C(\varepsilon, p, \lambda_1, \lambda_2, \eta_1, \eta_2)$ where λ_1, resp. λ_2, is the norm of the projection onto Y taking X_1, resp. X_2, to 0. In this step the basis constant of Y plays no role.

Step 2. With Y as in the statement of Proposition 8.2.

$$d(l_p^n, l_p^{n-m}\oplus Y)\leq C_2(\varepsilon, \lambda, M, p)$$

where the direct sum in $l_p^{n-m}\oplus Y$ is taken, e.g., in the l_p sense.

Proof of Step 1. In the sequel, all the isomorphism and complementation constants are assumed to be controlled by the parameters of the given data. To prove that the solution X of $l_p^n\approx X\oplus Y$ is unique, it suffices by [B.Tz.1] to prove that X contains a subspace Z complemented in l_p^n, so that $\dim Z=n'>cn$, that both Z and its complement are isomorphic to l_p spaces of the appropriate dimension and that moreover Y embeds complementably in $l_p^{n'}$. The latter requirement follows from Theorem 8.1 since $n'>m^{p^*/2}(\log m)^{2p^*-1}$. The existence of Z follows from [J.Sche.2] or [B.Tz.2]. Indeed, let

P be the projection on Y whose kernel is X. By [B.Tz.2] there is a subset $A\subset\{1,2,\ldots,n\}$ with $\bar{\bar{A}}\geq cn$ so that

$$\|R_A P R_A\|_{p\to p}<\frac{1}{10} \tag{8.24}$$

where R^A is the natural projection from l_p^n onto $\operatorname{span}\{e_i\}_{i\in A}$. Let

$$Z=(I-P)\operatorname{span}\{e_i\}_{i\in A}\subset X.$$

Clearly by (8.24), $l_p^n=Z\oplus\operatorname{span}\{e_i\}_{i\notin A}$.

Proof of Step 2. We shall use the finite-dimensional version of the decomposition method introduced in [B.D.G.J.N.]. Assume that for some m_0 and every subspace Y of dimension m_0 of l_p so that Y is complemented in l_p and has a good basis, we know that for some n_0

$$Y\oplus l_p^{n_0}\approx l_p^{n_0+m_0} \tag{8.25}$$

with a good control of the constants. We shall verify that then for every m_1 a similar statement is valid if m_0 and n_0 are replaced by $m_0 m_1$ and $m_1((m_0 m_1)^{p^*/2+\varepsilon}+n_0)$ respectively, again with a good control of the constants. This argument will show that if ϱ is defined by $n_0=m_0^\varrho$ and if $\varrho>p^*/2$ then (by an appropriate choice of m_1) we can replace ϱ by any $\tilde{\varrho}$ satisfying

$$\tilde{\varrho}>\frac{p^*}{2}\left(1-\frac{1}{\varrho}+\frac{2}{p^*}\right).$$

In other words we will get that (8.25) is valid for every m_0 and $n_0=m_0^\varrho$ provided $\varrho>p^*/2$ and this is the assertion in Step 2.

Let now $m=m_0 m_1$ and $n=m^{p^*/2+\varepsilon}$. Let Y be a complemented subspace of l_p with a Schauder basis $\{u_i\}_{i=1}^m$. Write $\{1,\ldots,m\}$ as a union of m_1 consecutive intervals A_j each of size m_0 and put

$$W_j=\operatorname{span}\{u_i\}_{i\in A_j},\quad j=1,\ldots,m_1$$

By the assumption (8.25) we have for every j

$$W_j\oplus l_p^{n_0}\approx l_p^{n_0+m_0}. \tag{8.26}$$

Let

$$U_j=\operatorname{span}\{u_i\}_{i\in\cup_{k\geq j}A_k},\quad j=1,\ldots,m_1.$$

Note that $U_1 = Y$ and that $U_{m_1} = W_{m_1}$. Since $\dim U_j \leq m$ for every j it follows by our choice of n and Theorem 8.1 that for every j and suitable Z_j

$$l_p^n \approx U_j \oplus Z_j. \tag{8.27}$$

By (8.26) and (8.27) we get

$$l_p^{n+n_0} \approx U_j \oplus Z_j \oplus l_p^{n_0} \approx U_{j+1} \oplus W_j \oplus Z_j \oplus l_p^{n_0} \approx U_{j+1} \oplus Z_j \oplus l_p^{n_0+m_0}. \tag{8.28}$$

Write

$$l_p^{(m_1-1)n+m_1(m_0+n_0)} = l_p^{n+m_0+n_0} \oplus \ldots \oplus l_p^{n+m_0+n_0} \oplus l_p^{n_0+m_0}$$

with m_1 summands (the direct sums are in the l_p sense). By (8.26), (8.27) and (8.28) we get

$$\begin{aligned} l^{(m_1-1)n+m_1(m_0+n_0)} &\approx (U_1 \oplus Z_1 \oplus l_p^{n_0+m_0}) \oplus \ldots \oplus (U_{m_1-1} \oplus Z_{m_1-1} \oplus l_p^{n_0+m_0}) \oplus (W_{m_1} \oplus l_p^{n_0}) \\ &\approx U_1 \oplus (U_2 \oplus Z_1 \oplus l_p^{n_0+m_0}) \oplus \ldots \oplus (U_{m_1} \oplus Z_{m_1-1} \oplus l_p^{n_0+m_0}) \oplus l_p^{n_0} \\ &\approx U_1 \oplus l_p^{n+n_0} \oplus \ldots \oplus l_p^{n+m_0} \oplus l_p^{n+m_0} \oplus l_p^{n_0} \\ &= Y \oplus l_p^{(m_1-1)n+m_1 n_0} \end{aligned}$$

and this concludes the proof. □

9. Entropy estimates, sharper results

The number of balls of radius k in l_∞^n needed to cover the ball of radius n in l_1^n (=the unit ball of $L_1(\mu)$ where μ is the natural probability measure on $\{1, 2, \ldots, n\}$) is clearly equivalent to the number of solutions to the equation

$$|x_1| + |x_2| + \ldots + |x_n| \leq n$$

where each x_i is an integer (positive, negative or 0) divisible by k. This number can be computed explicitely and as observed in [Schu.] one gets from this computation that

$$\log E(nB_{l_1^n}, B_{l_\infty^n}, t) \approx n(\log t)/t, \quad 2 \leq t \leq n. \tag{9.1}$$

In this section we shall prove a generalization of (9.1) to the case where l_1^n is replaced by an n-dimensional subspace of l_1^N in a "good position". The result we obtain will be sharp and strengthen the results obtained in section 4. Like (9.1) the result has a natural interpretation as an estimate for the cardinality of a suitable set of lattice points. The counting of the number of lattice points in the ball of an arbitrary n-dimensional subspace X of l_1^N is naturally more complicated than in the concrete case of l_1^n. The results in this section are obtained by refining the Banach space techniques of section 4. Hence one may view the results of this section as an application of Banach space theory to a problem of counting lattice points.

As in section 4 the estimate of the L_∞ entropy of the L_1 ball will be done via L_2 estimates. Thus our problem naturally divides into two parts-estimating the L_2 entropy of the L_1 ball and the L_∞ entropy of the L_2 ball. We start with the second part.

PROPOSITION 9.1. *Let X be an n-dimensional subspace of $L_1(\mu)$ where μ is a probability measure on $\{1,\dots,N\}$ satisfying $\mu\{i\}\geq 1/2N$ for every i. Denote by X_2 resp. X_∞ the space X endowed with the norm induced by $L_2(\mu)$ resp. $L_\infty(\mu)$. Let $\{\varphi_j\}_{j=1}^n$ be an orthonormal basis of X_2 and define ϱ by*

$$\left\|\left(\sum_{j=1}^n \varphi_j^2\right)^{1/2}\right\|_\infty = \varrho n^{1/2}. \tag{9.2}$$

Then

$$\log E(B_{X_2}, B_{X_\infty}, t) \leq c\varrho^2 n t^{-2}\log(tN/n), \quad 2\leq t\leq \varrho n^{1/2}, \tag{9.3}$$

where c is an absolute constant.

We prove (9.3) by using another intermediate space, namely the Orlicz space $L_{\psi_2}(\mu)$ where (as in section 6) $\psi_2(s)=\exp(s^2)-1$.

LEMMA 9.2. *With X as in Proposition 9.1 we have for a suitable absolute constant c_1*

$$\log E(B_{X_2}, B_{X_{\psi_2}}, t) \leq c_2\varrho^2 n t^{-2}, \quad 1\leq t\leq \varrho n^{1/2}. \tag{9.4}$$

Proof. By Proposition 4.2 we have

$$\log E(B_{X_2}, B_{X_{\psi_2}}, t) \leq c_2 n t^{-2} M^2 \tag{9.5}$$

where

$$M = n^{-1/2} \int_{\Omega} \left\| \sum_{j=1}^{n} g_j(\omega) \varphi_j \right\|_{L_{\psi_2}(\mu)} d\sigma(\omega)$$

the $\{g_j\}_{j=1}^n$ being normalized independent Gaussians on the probability space (Ω, σ).

Observe that $\|f\|_{L_{\psi_2}(\mu)} \leq c_3 \int \exp(f^2)\, d\mu$ for every function f and some universal c_3. Hence for every $c_4 > 0$ there is a $c_5 < \infty$ so that

$$\varrho^{-1} n^{-1/2} \int_{\Omega} \left\| \sum_{j=1}^{n} g_j(\omega) \varphi_j \right\|_{L_{\psi_2}(\mu)} d\sigma(\omega) \leq c_5 \int_{\Omega} \left(\int \exp\left(c_4 \left| \sum_{j=1}^{n} g_j(\omega) \varphi_j \right|^2 \Big/ n\varrho^2 \right) d\mu \right) d\sigma(\omega)$$

$$= c_5 \int \left(\int_{\Omega} \exp\left(c_4 \left| \sum_{j=1}^{n} g_j(\omega) \varphi_j \right|^2 \Big/ n\varrho^2 \right) d\sigma(\omega) \right) d\mu.$$

By the Laudau-Schepp-Fernique theorem (see [Ma.P.]) we have that for suitable c_4 and c_6

$$\int_{\Omega} \exp\left(c_4 \left| \sum_{j=1}^{n} g_j(\omega)\, a_j \right|^2 \right) d\sigma(\omega) \leq c_6$$

whenever $\Sigma_{j=1}^n a_j^2 \leq 1$. Consequently, we deduce that $M \leq \varrho c_5 c_6$ and by substituting this estimate in (9.5) we get (9.4). □

LEMMA 9.3 *With μ as in the statement of Proposition 9.1 we have*

$$\log E(B_{L_{\psi_2}(\mu)}, B_{L_\infty(\mu)}, t) \leq c_7 \exp(-c_8 t^2) N, \quad 1 \leq t \leq (\log N)^{1/2}. \tag{9.6}$$

Proof. Let $f \in B_{L_{\psi_2}(\mu)}$ and put $A = \{i;\ |f(i)| \leq 2t\}$, $f_0 = f\chi_A$, $f_1 = f - f_0$. A direct computation shows that $\|f_1\|_{L_2(\mu)} \leq c_9 \exp(-t^2)$. Hence

$$B_{L_{\psi_2}(\mu)} \subset 2t B_{L_\infty(\mu)} + c_9 \exp(-t^2) B_{L_2(\mu)}. \tag{9.7}$$

Since $\mu\{i\} > 1/2N$ for every i we get from (9.1) (cf. [Schu.]) that

$$\log E(B_{L_2(\mu)}, B_{L_\infty(\mu)}, s) \leq c_{10} N s^{-2} \log s \leq c_{11} N s^{-1}$$

and hence, by (9.7),

$$\log E(B_{L_{\psi_2}(\mu)}, B_{L_\infty(\mu)}, c_{12}t) \leq c_{13} N \exp(-t^2)$$

and this implies (9.6).

Proof of Proposition 9.1. For $1<s<t$ we have by Lemmas 9.2 and 9.3

$$\log E(B_{X_2}, B_{X_\infty}, t) \leq \log E(B_{X_2}, B_{X_{\psi_2}}, ts^{-1}) + \log E(B_{L_{\psi_2}(\mu)}, B_{L_\infty(\mu)}, s)$$
$$\leq c_1 \varrho^2 n s^2 t^{-2} + c_7 N \exp(-c_8 s^2).$$

By taking $s \approx (\log Nt/n)^{1/2}$, (9.3) follows. □

COROLLARY 9.4. *Let X be as in Proposition* 9.1. *and let Y be an m-dimensional subspace of X. Then there is an $f \in Y$ satisfying*

$$\|f\|_{L_2(\mu)} = 1, \quad \|f\|_{L_\infty(\mu)} \leq c_{14} \varrho (nm^{-1} \log(N\varrho m^{-1}))^{1/2}. \tag{9.8}$$

Proof. Assume that for some $t>1$, $tB_{Y_\infty} \subset \frac{1}{2} B_{Y_2}$. Then by (9.3)

$$2m \leq \log E(B_{Y_2}, B_{Y_\infty}, t) \leq \log E(B_{X_2}, B_{X_\infty}, t)$$
$$\leq cp^2 nt^{-2} \log(tN/n),$$

and hence $2t < c_{14} \varrho (nm^{-1} \log(N\varrho/m))^{1/2}$.

COROLLARY 9.5. *Let X be an n-dimensional subspace of l_∞^N. Then*

$$d(X, l_2^n) \geq c_{15} n^{1/2} (1 + \log N/n)^{-1/2}. \tag{9.9}$$

Proof. Let μ be any probability measure on $\{1, \ldots, N\}$ satisfying $\mu\{i\} > 1/2N$ for every i and let ϱ be given by (9.2). By Corollary 9.4 there is an orthogonal system $f_1, \ldots, f_m$ $(m=n/2)$ in X_2 satisfying $\|f_j\|_{L_\infty(\mu)} \leq c_{16} \varrho (\log(N\varrho/n))^{1/2}$ for every j. By the orthonormality of the f_j

$$\int \left\| \sum_{j=1}^m f_j(i) f_j \right\|_X d\mu(i) \leq \left(\int \left\| \sum_{j=1}^m f_j(i) f_j \right\|_X^2 d\mu(i) \right)^{1/2} \leq m^{1/2} d(X, l_2^n)\, c_{16} \varrho (\log(N\varrho/n))^{1/2}.$$

On the other hand since the norm in X is the L_∞ norm

$$\int \left\| \sum_{j=1}^{m} f_j(i) f_j \right\|_X d\mu(i) \geq \int \left(\sum_{j=1}^{m} f_j(i)^2 \right) d\mu(i) = m.$$

Hence

$$d(X, l_2^m) \geq m^{1/2}(c_{16}\varrho(\log(N\varrho/n))^{1/2})^{-1}.$$

We are free to choose the measure μ. We show that μ can be chosen so that ϱ will be bounded by an absolute constant and this will prove (9.9). The 2-absolutely summing norm of the identity of X is $n^{1/2}$ (cf. e.g., [P4], Theorem 1.11) and hence by the Pietch factorization theorem there is a probability measure ν on $\{1,\ldots,N\}$ so that $\|x\|_\infty \leq n^{1/2}\|x\|_{L_2(\nu)}$ for every $x \in X$. Define μ by

$$\mu\{i\} = \frac{1}{2N} + \frac{1}{2}\nu\{i\}, \quad 1 \leq i \leq N.$$

For every orthonormal system $\{\varphi_j\}_{j=1}^n$ of X (in the norm induced by $L_2(\mu)$ and any choice of scalars $\{\alpha_j\}_{j=1}^n$ we have

$$\left\| \sum_{j=1}^{n} \alpha_j \varphi_j \right\|_\infty \leq 2n^{1/2} \left\| \sum_{j=1}^{n} \alpha_j \varphi_j \right\|_{L_2(\mu)} = 2n^{1/2} \left(\sum_{j=1}^{N} \alpha_j^2 \right)^{1/2}.$$

Consequently $\|(\Sigma_{j=1}^n \varphi_j^2)^{1/2}\|_\infty \leq 2n^{1/2}$. □

Remarks. (1) S. Kislyakov informed us that E. Gluskin recently obtained (9.9).

(2) A simple example shows how precise the estimate (9.9) is (and for that matter also the preceding estimates). Let $X(k,m) = (\Sigma \oplus l_2^k)_\infty$ with m summands in the direct sum. Then $n = \dim X(k,m) = km$, $X(k,m)$ is 2 isomorphic to a subspace of l_∞^N with $N = m2^{ck}$ and $d(X(k,m), l_2^n) \leq m^{1/2}$. It follows that for all choices of k and m, (9.9) gives an equivalence.

(3) In [F.L.M.] it is shown that $d(X, l_2^n) \geq cn^{1/2}(\log N)^{-1/2}$ and in [F.J.] it is proved that $d(X, l_2^n) \geq \frac{1}{2}n^{1/2}(n/N)^{1/2}$. Thus only if N is close to n but not too close (e.g., $N = n\log n$) is (9.9) an improvement on previously published results.

(4) B. Carl and A. Pajor [C.Pa.] obtained independently an estimate which is essentially the dual of Proposition 9.1 (and thus by [T-J.] also essentially equivalent to Proposition 9.1).

We pass to the estimate of the L_2 entropy of the unit ball of a subspace X of L_1. Let X be an n-dimensional subspace l_1^N. As we noted in Lemma 4.5 it follows from the lemma of Lewis that there is a probability measure μ on $\{1, \dots, N\}$ with $\mu\{i\} \geq 1/2N$ for every i, a subspace $\tilde{X}$ of $L_1(\mu)$ which is isometric to X and an orthonormal system $\{\varphi_j\}_{j=1}^n$ in $\tilde{X}$ so that

$$\sum_{j=1}^{n} \varphi_j^2(i) \leq 4n, \quad 1 \leq i \leq N. \tag{9.10}$$

The entropy estimate we are going to prove will be for the unit ball of $\tilde{X}$.

For the proof below it is convenient (though not really necessary) to note that if B is the unit ball of an inner product space and D is any compact convex set then $E(D, B, t) = \tilde{E}(D, B, t)$. Indeed, if a compact convex set K is in the unit ball it is also contained in a ball with radius 1 centered at K (take as center the nearest point to 0 in K).

Proposition 9.6. *Let $\tilde{X}$ be an n-dimensional subspace of $L_1(\mu)$ as above. Then*

$$\log E(B_{\tilde{X}_1}, B_{\tilde{X}_2}, t) \leq c \min(K(\tilde{X})^2, \log t)\, nt^{-2}, \quad 2 \leq t \leq n^{1/2} \tag{9.11}$$

where $K(\tilde{X})$ is K-convexity constant of $\tilde{X}$.

Proof. We have

$$\log E(B_{\tilde{X}_1}, B_{\tilde{X}_2}, t) \leq \log E(B_{\tilde{X}_1}, B_{\tilde{X}_1} \cap tB_{\tilde{X}_2}, 2) + \log E\left(B_{\tilde{X}_1} \cap tB_{\tilde{X}_2}, B_{\tilde{X}_2}, \frac{t}{2}\right). \tag{9.12}$$

Note that

$$E(B_{\tilde{X}_1}, B_{\tilde{X}_1} \cap tB_{\tilde{X}_2}, 2) \leq E(B_{\tilde{X}_1}, B_{\tilde{X}_2}, 2t). \tag{9.13}$$

Indeed if $B_{\tilde{X}_1} \subset \bigcup_{k=1}^s (u_k + 2tB_{\tilde{X}_2})$ with $u_k \in B_{\tilde{X}_1}$ for every k, then for every $x \in B_{\tilde{X}_1}$ there is a k such that $x - u_k \in 2tB_{\tilde{X}_2}$ and hence $x - u_k \in 2(B_{\tilde{X}_1} \cap tB_{\tilde{X}_2})$. We estimate $E(B_{\tilde{X}_1} \cap tB_{\tilde{X}_2}, B_{\tilde{X}_2}, t/2)$ by using (4.5) and (4.20). Note that the norm induced by $B_{\tilde{X}_1} \cap tB_{\tilde{X}_2}$ on $\tilde{X}$ is $\max\{\|x\|_{L_1(\mu)}, t^{-1}\|x\|_{L_2(\mu)}\}$. Denote by $\tilde{X}(t)$ the space $\tilde{X}$ endowed with this norm. Clearly $d(\tilde{X}(t), l_2^n) \leq t$ and hence by (4.13)

$$K(\tilde{X}(t)) \leq c_1 \min(K(\tilde{X}), (\log t)^{1/2}).$$

Also the π_2 norm of the identity map from $\tilde{X}(t)$ to $\tilde{X}_2$ is at most that of the identity map from $\tilde{X}_1$ into $\tilde{X}_2$, i.e. by (4.25), at most $c_2 n^{1/2}$. Thus, we get that

$$\log E(B_{\tilde{X}_1} \cap tB_{\tilde{X}_2}, B_{\tilde{X}_2}, t/2) \leqslant c_3 nt^{-2} \min(K(\tilde{X})^2, \log t).$$

Substituting into (9.12) we get that

$$\log E(B_{\tilde{X}_1}, B_{\tilde{X}_2}, t) \leqslant \log E(B_{\tilde{X}_1}, B_{\tilde{X}_2}, 2t) + c_3 nt^{-2} \min(K(\tilde{X})^2, \log t). \tag{9.14}$$

Assertion (9.11) follows directly from (9.14) by iteration. □

We can now easily deduce the main result of this section, which generalizes (9.1) to a general n-dimensional subspace of l_1^N in a "good position".

THEOREM 9.7. *Let X be an n-dimensional subspace of l_1^N. Then there is a probability measure μ on $\{1, \ldots, N\}$ and a subspace $\tilde{X}$ of $L_1(\mu)$ which is isometric to X so that*

$$\log E(B_{\tilde{X}_1}, B_{\tilde{X}_\infty}, t) \leqslant cnt^{-1}(\log t \cdot \log(Nt/n))^{1/2}, \quad 2 \leqslant t \leqslant n. \tag{9.15}$$

Proof. Choose μ and $\tilde{X}$ as in Proposition 9.6. Then by (9.10) we may apply Proposition 9.1 to $\tilde{X}$ with $\varrho=2$. By (9.3) and (9.11) we get for $2<s<t$

$$\log E(B_{\tilde{X}_1}, B_{\tilde{X}_\infty}, t) \leqslant \log E(B_{\tilde{X}_1}, B_{\tilde{X}_2}, t/s) + \log E(B_{\tilde{X}_2}, B_{\tilde{X}_\infty}, s)$$
$$\leqslant cn(t^{-2}s^2 \log t + s^{-2} \log sN/n).$$

By taking s so that $s^2 \approx t(\log(tN/n) \cdot (\log t)^{-1})^{1/2}$, (9.15) follows. □

Added in proof. It is now known that, up to a possible logarithmic factor, the estimate given in Theorem 6.5 is the best possible. More precisely; there is (for $n \geqslant 2$) a constant C_n so that for every $0<\varepsilon<1/2$ there is a zonotope $P(n, \varepsilon)$ with $B^n \subset P(n, \varepsilon) \subset (1+\varepsilon)B^n$ and so that the number $N(B^n, \varepsilon)$ of summands of $P(n, \varepsilon)$ satisfies

$$N(B^n, \varepsilon) \leqslant C_n(\varepsilon^{-2} \log|\varepsilon|)^{(n-1)/(n+2)}$$

The proof (as well as results for other zonoids) is presented in: J. Bourgain and J. Lindenstrauss, Distributions of points on spheres and approximation by zonotopes, *Israel Journal of Mathematics* (to appear).

References

[B.D.G.J.N.] 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.

[B.G.] BOURGAIN, J. & GROMOV, M., Estimates of Bernstein widths of Sobolev spaces. *GAFA Seminar 87/88,* Springer Lecture Notes. To appear.

[B.L.] BOURGAIN, J. & LINDENSTRAUSS, J., Projection bodies. *GAFA Seminar* 86/87, Springer Lecture Notes no. 1317 (1988), 250–270.

[B.L.M.1] BOURGAIN, J., LINDENSTRAUSS, J. & MILMAN, V., Sur l'approximation de zonoides par des zonotopes, *C.R. Acad. Sci. Paris*, 303 (1986), 987–988.

[B.L.M.2.] — Minkowski sums and symmetrizations. *GAFA Seminar* 86/87, Springer Lecture Notes no. 1317 (1988), 44–66.

[B.Tz.1] BOURGAIN, J. & TZAFRIRI, L., Complements of subspaces of l_p^m, $p\geq 1$, which are uniquely determined. *GAFA Seminar* 85/86, Springer Lecture Notes no. 1267 (1987), 39–52.

[B.Tz.2] — Invertibility of large submatrices with applications to the geometry of Banach spaces and harmonic analysis. *Israel J. Math.*, 57 (1987), 137–224.

[Be.C.] BECK, J. & CHEN, W., *Irregularities of distribution.* Cambridge Tracts in Mathematics, vol. 89, 1987.

[Be.Mc.] BETKE, U. & MCMULLEN, P., Estimating the sizes of convex bodies from projections. *J. London Math. Soc.*, 27 (1983), 525–538.

[Bo] BOLKER, E. D., A class of convex bodies. *Trans. Amer. Math. Soc.*, 145 (1969), 323–346.

[Bon.F.] BONNESEN, T. & FENCHEL, W., *Theorie der konvexen Körper.* Ergebnisse der Mathematik, vol. 3, Springer Verlag 1934.

[C.Pa.] CARL, B. & PAJOR, A., Gelfand numbers of operators with values in a Hilbert space. To appear.

[D.M.T-J.] DAVIS, W. J., MILMAN, V. & TOMCZAK-JAEGERMANN, N., The distance between certain n-dimensional spaces. *Israel J. Math.*, 39 (1981), 1–15.

[F.J.] FIGIEL, T. & JOHNSON, W. B., Large subspaces of l_∞^n and estimates of the Gordon Lewis constant, *Israel J. Math.*, 37 (1980), 92–112.

[F.L.M.] FIGIEL, T., LINDENSTRAUSS, J. & MILMAN, V., The dimension of almost spherical sections of convex bodies. *Acta Math.*, 129 (1977), 53–94.

[G] GORDON, Y., Some inequalities for Gaussian processes and applications. *Israel J. Math.*, 50 (1985), 265–289.

[J.Sche.1] JOHNSON, W. B. & SCHECHTMAN, G., Embedding l_p^m into l_1^n. *Acta Math.*, 149 (1982), 71–85.

[J.Sche.2] — On subspaces of L_1 with maximal distances to Euclidean space. *Proc. Research workshop on Banach space theory.* University of Iowa 1981, 83–96.

[L.Tz.1] LINDENSTRAUSS, J. & TZAFRIRI, L., *Classical Banach spaces, vol.* I, *Sequence spaces.* Ergebnisse der Mathematik, vol. 92, Springer Verlag 1977.

[L.Tz.2] — *Classical Banach spaces, vol.* II. *Function spaces.* Ergebnisse der Mathematik, vol. 97, Springer Verlag 1979.

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

[Li] LINHART, J., Approximation of a ball by zonotopes using uniform distribution on the sphere. To appear.

[Mu] MÜLLER, C., *Spherical harmonics.* Springer Lecture Notes no. 17 (1966).

[M.Sche.] MILMAN, V. & SCHECHTMAN, G., *Asymptotic theory of finite dimensional normed spaces.* Springer Lecture Notes no. 1200 (1986).

[Ma.P.] MARCUS, M. B. & PISIER, G., *Random Fourier series with applications to harmonic analysis.* Ann. Math. Studies, vol. 101, Princeton 1981.

[P.1] PISIER, G., On the dimension of the l_p^n subspaces of Banach spaces for $1 \leq p < 2$. *Trans. Amer. Math. Soc.*, 276 (1983), 201–211.

[P.2] — Factorization of operators through $L_{p,\infty}$ or $L_{p,1}$ and non-commutative generalization. *Math. Ann.*, 276 (1986), 105–136.

[P.3] — Un Théorème de factorization pour les operateurs lineaires entre espaces de Banach. *Ann. Sci. Ecole Norm. Sup.*, 13 (1980), 23–43.

[P.4] — *Factorization of linear operators and geometry of Banach spaces*. Regional Conferences Series no. 60, Amer. Math. Soc. 1986.

[Pa. T-J] PAJOR, A. & TOMCZAK-JAEGERMANN, N., Subspaces of small codimension of finite dimensional Banach spaces. *Proc. Amer. Math. Soc.*, 97 (1986), 637–642.

[Sche.1] SCHECHTMAN, G., Fine embeddings of finite dimensional subspaces of L_p, $1 \leq p < 2$, into l_1^m. *Proc. Amer. Math. Soc.*, 94 (1985), 617–623.

[Sche.2] — Embeddings X_p^m spaces into l_r^n. *GAFA Seminar* 85/86, Springer Lecture Notes no. 1267 (1987), 53–75.

[Sche.3] — More on embeddings subspaces of L_p in l_r^n. *Compositio Math.*, 61 (1987), 159–170.

[Schn.] SCHNEIDER, R., Zonoids whose polars are zonoids. *Proc. Amer. Math. Soc.*, 50 (1975), 365–368.

[Schn.W] SCHNEIDER, R. & WEIL, W., Zonoids and related topics. *Convexity and its Applications*, Birkhauser Verlag (1983), 296–317.

[Schu.] SCHÜTT, C., Entropy numbers of diagonal operators between symmetric Banach spaces. *J. Approx. Theory*, 40 (1984), 121–128.

[Su] SUDAKOV, V. N., Gaussian random processes and measures of solid angles in Hilbert spaces. *Soviet Math. Dokl.*, 12 (1971), 412–415.

[T-J.] TOMCZAK-JAEGERMANN, N., Dualite des nombres d'entropie pour des opérateurs à valeur dans un espace de Hilbert. *C.R. Acad. Sci. Paris*, 305 (1987), 299–301.

[Zy] ZYGMUND, A., *Trigonometric series*. Cambridge Univ. Press, 1977.

Received August 23, 1987