ISRAEL JOURNAL OF MATHEMATICS, Vol. 68, No. 2, 1989

SYMMETRIC BLOCK BASES IN FINITE-DIMENSIONAL NORMED SPACES

BY

W. T. GOWERS

Department of Pure Mathematics and Mathematical Statistics, 16 Mill Lane, Cambridge CB2 1SD, England

ABSTRACT

It is shown that for $1 \leqq p < \infty$, any basis C-equivalent to the unit vector basis of l_p^n has a $(1+\varepsilon)$-symmetric block basis of cardinality proportional to $n/\log n$. When $1 < p < \infty$, an upper bound proportional to $n \log\log n/\log n$ is also obtained. These results extend results of Amir and Milman in [2].

§1. Introduction

After Milman's very successful proof of Dvoretzky's theorem using the isoperimetric inequality, it was soon noticed that other concentrations of measure results were useful in the local theory of Banach spaces, particularly for almost isometric embeddings. Two pioneering papers in this respect were [1] and [2], by Amir and Milman. Some of their results used concentration of measure in the space $\{-1, 1\}^n \times S_n$ to find almost symmetric block bases when the original basis satisfied certain conditions.

Recall that a basis $(y_i)_1^n$ is $(1+\varepsilon)$-*symmetric* if for any $\eta, \eta' \in \{-1, 1\}^n$, $\pi, \pi' \in S_n$ and sequence of scalars $(a_i)_1^n \in \mathbb{R}^n$,

$$\left\| \sum_1^n \eta_i a_i y_{\pi(i)} \right\| \leqq (1+\varepsilon) \left\| \sum_1^n \eta_i' a_i y_{\pi'(i)} \right\| .$$

The definition of a block basis in this context is natural, but not standard. A basis $(\mathbf{u}_j)_1^m$ is said to be a *block basis* of another basis $(x_i)_1^n$ if for some sequence of scalars $(b_i)_1^n$ and some sequence of disjoint subsets $(A_j)_1^m$ of $[n]$, $\mathbf{u}_j =$

Received August 7, 1988 and in revised form March 27, 1989

$\Sigma_{i\in A_j} b_i x_i$. This is not standard because we do not also require every element of A_i to precede every element of A_j whenever $i<j$.

Amir and Milman proved in [2] that given a basis $(x_i)_1^n$ which satisfies, for some $1 \leqq p < \infty$ and for all sequences $(a_i)_1^n \in \mathbb{R}^n$, the inequality

$$\left(\sum_1^n |a_i|^p\right)^{1/p} \leqq \left\|\sum_1^n a_i x_i\right\| \leqq C\left(\sum_1^n |a_i|^p\right)^{1/p},$$

there is a $(1+\varepsilon)$-symmetric block basis of $(x_i)_1^n$ of cardinality proportional to $n^{1/3}$. This was an intermediate step in their proof of a local version of Krivine's theorem, but the result is interesting in its own right. However, measure concentration arguments, which give best possible results for Dvoretzky's theorem, did not appear to do so here. The aim of this paper is to present an improvement of Amir and Milman's result to one that is essentially best possible. Both lower and upper bounds for the cardinality of an almost symmetric block basis are estimated. The lower bound is proportional to $n/\log n$ and an upper bound proportional to $n \log\log n/\log n$ is obtained when $1<p<\infty$. At the end of the paper, various problems related to this one are discussed.

The inequalities in the paper do not necessarily apply when n is small. This is not mentioned again. For the sake of tidiness, and because in this context it is not important, the dimensions calculated are not given as integers. Finally, the scalars throughout are assumed to be real, but the results carry over easily to the complex case.

§2. The lower bound

The theorem we shall prove in this section is the following.

THEOREM 1. *Let $1 \leqq p < \infty$, $0<\varepsilon<1/2$, $C>1$ and let $(x_i)_1^n$ be a basis for a normed space X. Suppose that for any $\mathbf{a} = \Sigma_1^n a_i \mathbf{e}_i \in \mathbb{R}^n$,*

$$\|\mathbf{a}\|_p \leqq \left\|\sum_1^n a_i x_i\right\| \leqq C\|\mathbf{a}\|_p.$$

Then $(x_i)_1^n$ has a block basis with blocks of ± 1 coefficients and equal length, which is $(1+\varepsilon)$-symmetric and has cardinality

$$m = (1/64)(\varepsilon/24C)^{2p}\cdot(\varepsilon/100C)\cdot(\log(90C/\varepsilon))^{-1} n/\log n.$$

We shall begin with some notation.

Let Ψ be the group $\{-1, 1\}^m \times S_m$ with multiplication given by $((\eta_i)_1^m, \sigma) \circ ((\eta_i')_1^m, \sigma') = ((\eta_i \eta_i')_1^m, \sigma \circ \sigma')$. There is a natural action of Ψ on $\mathbb{R}^n$: namely, for $\mathbf{a} = \Sigma_1^m a_i \mathbf{e}_i \in \mathbb{R}^m$, and $(\eta, \sigma) \in \Psi$, $\eta = (\eta_i)_1^m$, let (η, σ) send $\mathbf{a}$ to $\mathbf{a}_{\eta,\sigma} = \Sigma_1^m \eta_i \mathbf{e}_{\sigma(i)}$.

Let Ω be the group $\{-1, 1\}^n \times S_n$ with the same multiplication, *mutatis mutandis*, and for $\mathbf{b} \in X$, $\mathbf{b} = \Sigma_1^n b_i x_i$, $(\varepsilon, \pi) \in \Omega$, let $\omega_{\varepsilon,\pi}(\mathbf{b}) \equiv \mathbf{b}_{\varepsilon,\pi} = \Sigma_1^n \varepsilon_i b_i x_{\pi(i)}$.

It will sometimes be convenient to relabel the indices of $(x_i)_1^n$. We shall set $x_{ij} = x_{(i-1)h+j}$ $(i = 1, \ldots, m, j = 1, \ldots, h)$, where $hm = n$, and similarly we shall set $\varepsilon_{ij} = \varepsilon_{(i-1)h+j}$ and $\pi_{ij} = \pi((i-1)h + j)$ for $(\varepsilon, \pi) \in \Omega$.

Using this relabelling, we can define an action of Ψ on Ω by $\psi_{\eta,\sigma}((\varepsilon, \pi)) = (\varepsilon', \pi')$, where

$$\left.\begin{aligned} \varepsilon_{ij}' &= \eta_i \varepsilon_{\sigma(i)j} \\ \pi_{ij}' &= \pi_{\sigma(i)j} \end{aligned}\right\} \quad i = 1, \ldots, m, \quad j = 1, \ldots, h.$$

We shall regard a block basis of $(x_i)_1^n$ as a random embedding of $\mathbb{R}^n$ into X. Let $\phi : \mathbb{R}^m \to X$ be the embedding defined by

$$\phi : \sum_{i=1}^m a_i \mathbf{e}_i \mapsto \sum_{i=1}^m \sum_{j=1}^h a_i x_{ij}$$

and write $\mathbf{u}_i = \Sigma_{j=1}^h x_{ij}$, for $i = 1, \ldots, m$. Then let $\phi_{\varepsilon,\pi} = \omega_{\varepsilon,\pi} \circ \phi$, i.e.

$$\phi_{\varepsilon,\pi} : \sum_{i=1}^m a_i \mathbf{e}_i \mapsto \sum_{i=1}^m \sum_{j=1}^h \varepsilon_{ij} a_i x_{\pi_{ij}}.$$

Then $(\mathbf{e}_i)_1^m$ maps under $\phi_{\varepsilon,\pi}$ to the block basis $((\mathbf{u}_i)_{\varepsilon,\pi})_1^m$.

The proof of Theorem 1 is based on the following three assertions.

(i) Let A be the set $\{\mathbf{a} \in l_p^m : \| \mathbf{a} \|_p \leqq 1, a \geqq \cdots \geqq a_m \geqq 0\}$ and let $\delta > 0$. Then A contains a δ-net of cardinality at most $m^{2/\log(1+\delta/3))\log(15/\delta)}$.

(ii) If, for a particular $(\varepsilon, \pi) \in \Omega$, we have that for all $\mathbf{a}$ in the above net, $\| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) \|$ is constant to within $\delta \| \mathbf{a} \|_p h^{1/p}$, as (η, σ) runs through Ψ, then the block basis $(\mathbf{u}_1)_{\varepsilon,\pi}, \ldots, (\mathbf{u}_m)_{\varepsilon,\pi}$ is $(1 + 6C\delta)$-symmetric.

(iii) For any $\mathbf{a} \in A$, $\exists M(\mathbf{a}) \in \mathbb{R}$ such that

$$\mathbb{P}_\Omega\left[\exists (\eta, \sigma) \text{ s.t. } | \| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) \| - M(\mathbf{a})| > \frac{\varepsilon}{12C} \| \mathbf{a} \|_p h^{1/p}\right] < m^{-(2/\log(1+\varepsilon/18C))\log(90C/\varepsilon)}.$$

Once we have shown (i), (ii) and (iii) the theorem can be proved. Set δ to be $\varepsilon/6C$ and $N = m^{(2/\log(1+\varepsilon/18C))\log(90C/\varepsilon)}$. Then by (i) the set A contains a δ-net Δ of cardinality N. It follows from (iii) that for any $\mathbf{a}$ in the net, $\| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) \|$ varies by at most $\delta \| \mathbf{a} \|_p h^{1/p}$ as (η, σ) runs through Ψ, with probability greater than

$1-N^{-1}$ in Ω. Therefore there is some $(\varepsilon,\pi)\in\Omega$ for which $\|\phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|$ is constant to within $\delta\|\mathbf{a}\|_p h^{1/p}$ for all $\mathbf{a}\in\Delta$. In other words there is an $(\varepsilon,\pi)\in\Omega$ which satisfies the conditions of (ii). But then it follows from (ii) that the block basis $(\mathbf{u}_1)_{\varepsilon,\pi},\dots,(\mathbf{u}_m)_{\varepsilon,\pi}$ is $(1+\varepsilon)$-symmetric. Thus Theorem 1 follows from the three assertions.

Of these steps, (i) is not standard, but not especially difficult. Because the norm $\|\cdot\|_{\varepsilon,\pi}$ defined on $\mathbb{R}^m$ by $\|\mathbf{a}\|_{\varepsilon,\pi}=\max\{\|\phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|:(\eta,\sigma)\in\Psi\}$ is a 1-symmetric norm, and $\{\mathbf{a}_{\eta,\sigma}:\mathbf{a}\in A,\ (\eta,\sigma)\in\Psi\}$ is a δ-net of $B(l_p^n)$, the statement in (ii) follows from a standard argument to be found in virtually all proofs of almost isometric embeddings, namely that the behaviour of a norm is controlled by its behaviour on a sufficiently fine net. We shall discuss steps (i) and (ii) later, but the important step of the proof is (iii). It turns out that examining whole Ψ-orbits in $\mathbb{R}^m$ at a time gives good enough estimates in step (iii) for it not to matter that our δ-net $\{\mathbf{a}_{\eta,\sigma}:\mathbf{a}\in A,(\eta,\sigma)\in\Psi\}$ is not as small as it could be.

We shall begin by restricting ourselves to the case $p=1$. For $p>1$ there is an extra technicality which we shall discuss afterwards.

Let us fix a vector $\mathbf{a}$ with $\|\mathbf{a}\|=1$, and $a_1\geqq\cdots\geqq a_m\geqq 0$. Let $B_1,\dots,B_{k+1}\subset[m]$ be defined by

$$B_j=\begin{cases}\{i\in[m]:2^{-j}<a_i\leqq 2^{-(j-1)}\}, & 1\leqq j\leqq k\\ \{i\in[m]:a_i\leqq 2^{-k}\}, & j=k+1\end{cases}$$

where $k=\log_2(60Cm/\varepsilon)$.

Let $\mathbf{b}_1,\dots,\mathbf{b}_{k+1}$ be given by $\mathbf{b}_j=\mathbf{a}|_{B_j}$ $(1\leqq j\leqq k+1)$. For $(\eta,\sigma)\in\Psi$, we define $\mathbf{b}^j_{\eta,\sigma}=(\mathbf{b}_j)_{\eta,\sigma}$. Clearly $\mathbf{b}^j_{\eta,\sigma}=\mathbf{a}_{\eta,\sigma}|_{\sigma(B_j)}$, and the absolute value of the coefficients of $\mathbf{b}^j_{\eta,\sigma}$ lie between 2^{-j} and $2^{-(j-1)}$ when $j\leqq k$, and are at most $\varepsilon/60Cm$ when $j=k+1$.

For each $1\leqq r\leqq k+1$, $(\eta,\sigma)\in\Psi$, define a function $f^r_{\eta,\sigma}:\Omega\to\mathbb{R}$ by

$$f^r_{\eta,\sigma}((\varepsilon,\pi))=\mathbb{E}[\|\phi_{\varepsilon',\pi'}(\mathbf{a}_{\eta,\sigma})\|\ \big|\ \phi_{\varepsilon',\pi'}(\mathbf{b}^j_{\eta,\sigma})=\phi_{\varepsilon,\pi}(\mathbf{b}^j_{\eta,\sigma}),j=1,\dots,r].$$

Now, for any fixed (η,σ), the sequence of functions $f^0_{\eta,\sigma}$ $(=\mathbb{E}(\|\phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|)$, $f^1_{\eta,\sigma},\dots,f^{k+1}_{\eta,\sigma}$ is a martingale. Note that $f^{k+1}_{\eta,\sigma}(\varepsilon,\pi)=\|\phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|$, although the expectation is not taken over a singleton subset of Ω. This is because if $\phi_{\varepsilon',\pi'}(\mathbf{b}^j_{\eta,\sigma})=\phi_{\varepsilon,\pi}(\mathbf{b}^j_{\eta,\sigma})$ for $j=1,\dots,k+1$ then $\phi_{\varepsilon',\pi'}(\mathbf{a}_{\eta,\sigma})=\phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma})$. Actually the fact that $f^0_{\eta,\sigma},\dots,f^{k+1}_{\eta,\sigma}$ is a martingale will not concern us. Instead we are interested in two facts.

(a) The number of *distinct* functions $f^r_{\eta,\sigma}$ can be controlled, and is small when r is small.

More important than this is

(b) The probability (in Ω) of $f^r_{\eta,\sigma}((\varepsilon,\pi))$ differing substantially from $f^{r-1}_{\eta,\sigma}((\varepsilon,\pi))$ is small.

The estimate in (a) is simple. If we have (η,σ), (η',σ') such that $\mathbf{b}^j_{\eta,\sigma} = \mathbf{b}^j_{\eta',\sigma'}$ for $j = 1,\dots,r$, then it is easy to see that $f^r_{\eta,\sigma} \equiv f^r_{\eta',\sigma'}$. But the number of distinct choices of $\mathbf{b}^1_{\eta,\sigma},\dots,\mathbf{b}^r_{\eta,\sigma}$ is certainly at most $m(m-1)\cdots(m-\Sigma_{j=1}^r |B_j|)\cdot 2^{\Sigma_{j=1}^r |B_j|}$. So, writing $\beta_j = |B_j|$ $(j = 1,\cdots,k)$ and $\gamma_j = \Sigma_{i=1}^j \beta_i$, we obtain that there are at most $(2m)^{\gamma_r}$ distinct functions $f^r_{\eta,\sigma}$.

We shall use well-known martingale techniques to get an estimate in (b). (See Lemma 4, Corollary 5 and the remarks that follow.) For now we quote a result and show why it is all that is needed to prove the statement in (iii).

The result we quote is that for any $(\eta,\sigma)\in\Psi$, $1 \leqq r \leqq k$ and $\delta_r > 0$,

$$\mathbb{P}_\Omega[f^r_{\eta,\sigma}((\varepsilon,\pi)) - f^{r-1}_{\eta,\sigma}((\varepsilon,\pi)) > \delta_r h] < \exp\left(-\frac{\delta_r^2 2^{2(r-1)}h}{8C^2\beta_r}\right)$$

and

$$\mathbb{P}_\Omega[f^r_{\eta,\sigma}((\varepsilon,\pi)) - f^{r-1}_{\eta,\sigma}((\varepsilon,\pi)) < -\delta_r h] < \exp\left(-\frac{\delta_r^2 2^{2(r-1)}h}{8C^2\beta_r}\right).$$

Because of the bound given earlier for $\| \mathbf{b}^{k+1}_{\eta,\sigma} \|_\infty$ we also have, for any (ε,π), that

$$f^{k+1}_{\eta,\sigma}((\varepsilon,\pi)) - f^k_{\eta,\sigma}((\varepsilon,\pi)) \leqq (\varepsilon/60)h = (\varepsilon/60)h \,\| \mathbf{a} \|_1.$$

We shall take the number $M(\mathbf{a})$ appearing in (iii) to be $\mathbb{E}_\Omega(\| \phi_{\varepsilon,\pi}(\mathbf{a}) \|)$. Note that this is the same as $\mathbb{E}_\Omega(\| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) \|)$ for any $(\eta,\sigma)\in\Psi$, as has already been mentioned. (The letter M is used because later it will stand for a median.) Note also that the above probabilities are both zero in the case $\beta_r = 0$.

Now suppose that for some $(\varepsilon,\pi)\in\Omega$ it is true that

$$\exists\,(\eta,\sigma) \text{ s.t. } \| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) \| - M(\mathbf{a}) > \frac{\varepsilon}{12} \| \mathbf{a} \|_1 h,$$

i.e.

$$f^{k+1}_{\eta,\sigma}((\varepsilon,\pi)) - f^0_{\eta,\sigma}((\varepsilon,\pi)) > \frac{\varepsilon}{12} \| \mathbf{a} \|_1 h.$$

Then

$$f^k_{\eta,\sigma}((\varepsilon,\pi)) - f^0_{\eta,\sigma}((\varepsilon,\pi)) > \frac{\varepsilon}{15} \| \mathbf{a} \|_1 h.$$

Hence, if $\delta_1 + \cdots + \delta_k \leqq \varepsilon/15$, there will be some $1 \leqq r \leqq k$ and $(\eta, \sigma) \in \Psi$ such that

$$f^r_{\eta,\sigma}((\varepsilon,\pi)) - f^{r-1}_{\eta,\sigma}((\varepsilon,\pi)) > \delta_r \| \mathbf{a} \|_1 h.$$

However, by the estimates in (a) and (b) and the normalization $\| \mathbf{a} \|_1 = 1$, the probability of such r and (η, σ) existing is at most

$$\sum_{r=s}^{k} (2m)^{\gamma_r} \exp\left(-\frac{2^{2(r-1)}\delta_r^2 h}{8C^2\beta_r} \right)$$

where s is the smallest value of r for which $\gamma_r > 0$.

It remains to choose appropriate $\delta_1, \ldots, \delta_k$ and to verify that this probability is at most $\frac{1}{2} \cdot N^{-1}$. Since the other inequality is exactly similar, we will then be done.

Choosing $\delta_r = 2^{-r}\beta_r^{1/2}\gamma_r^{1/2} \cdot \varepsilon/24$ will do.

First,

$$\sum_1^k \delta_r = \frac{\varepsilon}{24} \cdot \sum_1^k 2^{-r}\beta_r^{1/2}\gamma_r^{1/2}$$

$$\leqq \frac{\varepsilon}{24} \left(\sum_1^k 2^{-r}\beta_r \right)^{1/2} \left(\sum_1^k 2^{-r}\gamma_r \right)^{1/2} \quad \text{(by the Cauchy–Schwarz inequality)}$$

$$\leqq \frac{\varepsilon}{24} \left(\sum_{r=1}^k 2^{-r} \sum_{j=1}^r \beta_j \right)^{1/2} \quad \left(\text{since } \sum_1^k 2^{-r}\beta_r \leqq \sum_1^m a_i = 1 \text{ and } \gamma_r = \sum_{j=1}^r \beta_j \right)$$

$$= \frac{\varepsilon}{24} \left(\sum_{j=1}^k \beta_j \sum_{r=j}^k 2^{-r} \right)^{1/2}$$

$$< \frac{\varepsilon\sqrt{2}}{24} \left(\sum_{j=1}^k 2^{-j}\beta_j \right)^{1/2} < \frac{\varepsilon}{15}.$$

Second,

$$\sum_{r=s}^{k}(2m)^{\gamma_r}\exp\left(-\frac{2^{2(r-1)}\delta_r^2 h}{8\beta_r C^2}\right)=\sum_{r=s}^{k}\exp\left(\gamma_r\left(\log(2m)-\frac{\varepsilon^2 h}{18,432C^2}\right)\right)$$

$$\leqq k\exp\left(\log(2m)-\frac{\varepsilon^2 h}{18,432C^2}\right)$$

(since $\gamma_r>0\ \forall r\geqq s$).

But since $h>32(576C^2/\varepsilon^2)\cdot(100C/\varepsilon)\cdot\log(90C/\varepsilon)\cdot\log n$, we have

$$k\exp\left(\log(2m)-\frac{\varepsilon^2 h}{18{,}234C^2}\right)\leqq k\exp\left(\log n\left(1-\frac{100C}{\varepsilon}\cdot\log\left(\frac{90C}{\varepsilon}\right)\right)\right)$$

$$\leqq\frac{1}{2}n^{-(50C/\varepsilon)\log(90C/\varepsilon)}$$

$$\leqq\frac{1}{2}m^{-(50C/\varepsilon)\log(90C/\varepsilon)}<\frac{1}{2}N^{-1}$$

which is what was needed.

When $p>1$, the proof is very similar, but it is not possible to work directly with the norm. Instead, for fixed $a\in l_p^m$, $\|\mathbf{a}\|_p=1$, $a_1\geqq\cdots\geqq a_m\geqq 0$, we define, for each $(\eta,\sigma)\in\Psi$, a function $g_{\eta,\sigma}:\Omega\to\mathbb{R}$ as follows.

Let $\Gamma_{\eta,\sigma}\in\Omega$ be the set

$$\{(\varepsilon,\pi):\|\phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|\leqq\mathbf{M}\|\phi_{\varepsilon',\pi'}(\mathbf{a}_{\eta,\sigma})\|\}.$$

(The symbol **M** denotes here and for the rest of the section the median taken over Ω.)

Let $d_{\eta,\sigma}$ be a metric on Ω defined by

$$d_{\eta,\sigma}((\varepsilon,\pi),(\varepsilon',\pi'))=\sum_{i=1}^{m}|a_i|^p\{j:\varepsilon_{\sigma(i)j}\neq\varepsilon'_{\sigma(i)j}\text{ or }\pi_{\sigma(i)j}\neq\pi'_{\sigma(i)j}\}|.$$

Then

$$g_{\eta,\sigma}((\varepsilon,\pi))=d_{\eta,\sigma}((\varepsilon,\pi),\Gamma_{\eta,\sigma}).$$

Thus $g_{\eta,\sigma}$ measures how far (ε,π) is from some (ε',π') for which $\|\phi_{\varepsilon',\pi'}(\mathbf{a}_{\eta,\sigma})\|$ is below the median. Moreover, the distance is weighted according to $\mathbf{a}_{\eta,\sigma}$.

When $(\eta,\sigma)=1_\Psi$, let us write g for $g_{\eta,\sigma}$, d for $d_{\eta,\sigma}$ and Γ for $\Gamma_{\eta,\sigma}$. Recalling that $\psi_{\eta,\sigma}$ represents the action of (η,σ) on Ω, we have

$$d_{\eta,\sigma}((\varepsilon,\pi),(\varepsilon',\pi')) = \sum_{i=1}^{m} |a_i|^p |\{j : \varepsilon_{\sigma(i)j} \neq \varepsilon'_{\sigma(i)j} \text{ or } \pi_{\sigma(i)j} \neq \pi'_{\sigma(i)j}\}|$$

$$= d(\psi_{\eta,\sigma}(\varepsilon,\pi), \psi_{\eta,\sigma}(\varepsilon',\pi')).$$

Also

$$\| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) \| = \left\| \sum_{i=1}^{m} \sum_{j=1}^{h} \varepsilon_{\sigma(i)j} \eta_i a_i x_{\pi_{\sigma(i)j}} \right\| = \| \phi_{\psi_{\eta,\sigma}(\varepsilon,\pi)}(\mathbf{a}) \|$$

and thus

$$\Gamma_{\eta,\sigma} = \{(\varepsilon,\pi) : \| \phi_{\psi_{\eta,\sigma}(\varepsilon,\pi)}(\mathbf{a}) \| \leqq \mathbf{M} \| \phi_{\varepsilon',\pi'}(\mathbf{a}) \| \}$$

$$= \{\psi_{\eta,s}^{-1}(\varepsilon,\pi) : \| \phi_{\varepsilon,\pi}(\mathbf{a}) \| \leqq \mathbf{M} \| \phi_{\varepsilon',\pi'}(\mathbf{a}) \| \}$$

$$= \psi_{\eta,\sigma}^{-1}(\Gamma).$$

Hence

$$g_{\eta,\sigma}(\varepsilon,\pi) = d_{\eta,\sigma}((\varepsilon,\pi),\Gamma_{\eta,\sigma}) = d(\psi_{\eta,\sigma}(\varepsilon,\pi), \psi_{\eta,\sigma}(\Gamma_{\eta,\sigma}))$$

$$= d(\psi_{\eta,\sigma}(\varepsilon,\pi),\Gamma) = g(\psi_{\eta,\sigma}(\varepsilon,\pi)).$$

Now the main reason $g_{\eta,\sigma}$ is useful is that

$$| \| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) \| - \| \phi_{\varepsilon',\pi'}(\mathbf{a}_{\eta,\sigma}) \| |$$

$$\leqq \| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) - \phi_{\varepsilon',\pi'}(\mathbf{a}_{\eta,\sigma}) \|$$

$$= \| \phi_{\psi_{\eta,\sigma}(\varepsilon,\pi)}(\mathbf{a}) - \phi_{\psi_{\eta,\sigma}(\varepsilon',\pi')}(\mathbf{a}) \|$$

$$\leqq 2C \left(\sum_{i=1}^{m} |a_i|^p |\{j : \varepsilon_{\sigma(i)j} \neq \varepsilon'_{\sigma(i)j} \text{ or } \pi_{\sigma(i)j} \neq \pi'_{\sigma(i)j}\}| \right)^{1/p}$$

$$= 2C(d_{\eta,\sigma}((\varepsilon,\pi),(\varepsilon',\pi')))^{1/p}.$$

Hence, if $g_{\eta,\sigma}(\varepsilon,\pi) < \delta$, then

$$\| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) \| - \mathbf{M} \| \phi_{\varepsilon',\pi'}(\mathbf{a}) \| < 2C\delta^{1/p}.$$

From now on, the proof is virtually the same as before. We define

$$B_j = \begin{cases} \{i \in [m] : 2^{-j} < a_i^p \leqq 2^{-(j-1)}\} & 1 \leqq j \leqq k \\ \{i \in [m] : a_i^p \leqq 2^{-k}\} & j = k+1 \end{cases}$$

with $k = \log_2(4 \cdot (24C/\varepsilon)^p m)$.

As before, $\mathbf{b}_j = \mathbf{a}\big|_{B_j}$ $(1 \leqq j \leqq k)$, and $\mathbf{b}_{\eta,\sigma}^j = (\mathbf{b}_j)_{\eta,\sigma}$.

Then for $1 \leqq r \leqq k+1$ and $(\eta,\sigma) \in \Psi$, we set

$f^r_{\eta,\sigma}((\varepsilon,\pi))$

$$= \mathbb{E}_\Omega\left[g_{\eta,\sigma}((\varepsilon',\pi')) \,\middle|\, \varepsilon'_{ij} = \varepsilon_{ij}, \pi'_{ij} = \pi_{ij}, \forall i \in \bigcup_{s=1}^{r} \sigma(B_s), j = 1,\ldots,h\right].$$

Note that if $\varepsilon'_{ij} = \varepsilon_{ij}$ and $\pi'_{ij} = \pi_{ij}$ for all $i \in \bigcup_{s=1}^{r} \sigma(B_s)$, $j = 1,\ldots,h$, then $\phi_{\varepsilon,\pi}(\mathbf{b}^s_{\eta,\sigma}) = \phi_{\varepsilon',\pi'}(\mathbf{b}^s_{\eta,\sigma})$ for $s = 1,\ldots,r$, but more is true, since for example there cannot be some i and some $\tau \in S_h$ such that $\pi'_{ij} = \pi_{i\tau(j)}$. This is purely for the sake of convenience.

Note also that for any (ε,π),

$$f^{k+1}_{\eta,\sigma}((\varepsilon,\pi)) - f^k_{\eta,\sigma}((\varepsilon,\pi)) \leqq \frac{1}{4}\left(\frac{\varepsilon}{24C}\right)^p h.$$

We would like to prove three facts, of which the first two correspond to facts (a) and (b) in the proof when $p = 1$. These are

(a)′ If $\eta_i = \eta'_i$ and $\sigma(i) = \sigma'(i)$ for all i in $\bigcup_{s=1}^{r} B_s$, then $f^r_{\eta,\sigma} = f^r_{\eta',\sigma'}$.

(b)′ For all (η,σ) in Ψ, $1 \leqq r \leqq k$ and $\delta_r > 0$,

$$\mathbb{P}_\Omega[f^r_{\eta,\sigma}((\varepsilon,\pi)) - f^{r-1}_{\eta,\sigma}((\varepsilon,\pi)) > \delta_r h] < \exp\left(-\frac{\delta_r^2 2^{2(r-1)} h}{8\beta_r}\right)$$

(where $\beta_r = |B_r|$ as before).

(c)′ $f^0_{\eta,\sigma} < \frac{1}{4}(\varepsilon/24C)^p h$, i.e. $\mathbb{E}(d_{\eta,\sigma}((\varepsilon,\pi),\Gamma_{\eta,\sigma}))$ is small.

We shall prove (b)′ later, using martingale techiques. The fact (c)′ is a technicality common to many concentration of measure arguments, which will also be proved later. To prove (a)′, we use the fact that $g_{\eta,\sigma}((\varepsilon',\pi')) = g(\psi_{\eta,\sigma}(\varepsilon',\pi'))$, and hence

$f^r_{\eta,\sigma}((\varepsilon,\pi))$

$$= \mathbb{E}_\Omega\left[g((\varepsilon',\pi')) \,\middle|\, \eta_i \varepsilon'_{\sigma^{-1}(i)j} = \varepsilon_{ij}, \pi'_{\sigma^{-1}(i)j} = \pi_{ij} \;\forall i \in \bigcup_{s=1}^{r} \sigma(B_s), j = 1,\ldots,h\right]$$

$$= \mathbb{E}_\Omega\left[g((\varepsilon',\pi')) \,\middle|\, \varepsilon'_{ij} = \eta_i \varepsilon_{\sigma(i)j}, \pi'_{ij} = \pi_{\sigma(i)j} \;\forall i \in \bigcup_{s=1}^{r} B_s, j = 1,\ldots,h\right].$$

(a)′ follows immediately, and with $\gamma_r = \Sigma_{s=1}^r \beta_s$, the number of distinct $f^r_{\eta,\sigma}$ is at most $(2m)^{\gamma_r}$ as before.

Again just as before, we may now conclude that

$$\mathbb{P}_\Omega\left[\exists(\eta,\sigma)\in\Psi \text{ s.t.} f^k_{\eta,\sigma}((\varepsilon,\pi))-f^0_{\eta,\sigma}((\varepsilon,\pi))>\frac{1}{2}\left(\frac{\varepsilon}{24}\right)^p h\right]$$

$$\leqq \sum_{r=1}^{k}(2m)^{\gamma_r}\exp\left(-\frac{2^{2(r-1)}\delta_r^2 h}{8\beta_r}\right) \tag{1}$$

where $\delta_1,\ldots,\delta_r$ is any sequence satisfying $\Sigma_1^k\delta_r\leqq(1/2)\cdot(\varepsilon/24C)^p$. We shall choose δ_r to be $(1/4)\cdot 2^{-r}\beta_r^{1/2}\gamma_r^{1/2}\cdot(\varepsilon/24C)^p$. Then just as before, $\Sigma_1^k\delta_r$ is indeed at most $(1/2)\cdot(\varepsilon/24C)^p$, and the right hand side of (1) is at most

$$k\exp\left(\log(2m)-\left(\frac{\varepsilon}{24C}\right)^{2p}\frac{h}{64}\right).$$

But if

$$f^k_{\eta,\sigma}((\varepsilon,\pi))-f^0_{\eta,\sigma}((\varepsilon,\pi))\leqq\frac{1}{2}\left(\frac{\varepsilon}{24C}\right)^p h$$

and

$$f^0_{\eta,\sigma}((\varepsilon,\pi))<\frac{1}{4}\left(\frac{\varepsilon}{24C}\right)^p h \qquad \text{(from (c)')}$$

and

$$f^{k+1}_{\eta,\sigma}((\varepsilon,\pi))-f^k_{\eta,\sigma}((\varepsilon,\pi))\leqq\frac{1}{4}\left(\frac{\varepsilon}{24C}\right)^p h,$$

then

$$f^{k+1}_{\eta,\sigma}((\varepsilon,\pi))\leqq\left(\frac{\varepsilon}{24C}\right)^p h$$

and so

$$\|\phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|-\mathbf{M}\|\phi_{\varepsilon',\pi'}(\mathbf{a}_{\eta,\sigma})\|\leqq\frac{\varepsilon h^{1/p}}{12}=\frac{\varepsilon}{12}\|\mathbf{a}\|_p h^{1/p}.$$

Hence

$$\mathbb{P}_\Omega\left[\exists(\eta,\sigma)\text{ s.t. }|\,\|\phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|-\mathbf{M}(\mathbf{a})|>\frac{\varepsilon}{12}\|\mathbf{a}\|_p h^{1/p}\right]$$

$$\leqq 2k\exp\left(\log 2m-\left(\frac{\varepsilon}{24C}\right)^{2p}\frac{h}{64}\right).$$

But since $h>64\cdot(24C/\varepsilon)^{2p}\cdot(100C/\varepsilon)\log n$, this is at most

$$2k \exp(\log n(1 - (100C/\varepsilon)\log(90C/\varepsilon))) \leqq n^{-(50C/\varepsilon)\log(90C/\varepsilon)}$$

$$\leqq m^{-(50C/\varepsilon)\log(90C/\varepsilon)}$$

$$< N^{-1}.$$

We shall now turn to the details left out so far. First we prove (i).

LEMMA 2. *Let $1 \leqq p \leqq \infty$, $0 < \delta < 1$ and let $K \subset l_p^n$ be the set*

$$\{\mathbf{a} \in l_p^n : \| \mathbf{a} \|_p \leqq 1, a_1 \geqq \cdots \geqq a_n \geqq 0\}.$$

Then K contains a δ-net Δ of cardinality N, where

$$N \leqq n^{(2/\log(1+\delta/3))\log(15/\delta)}.$$

PROOF. Let $\theta = \delta/3$ and let $\mathbf{a} = (a_i)_1^n \in K$. If $\mathbf{a}' = (a_i')_1^n \in l_p^n$ is any vector such that $a_i \leqq a_i' \leqq (1+\theta)a_i$ for all $1 \leqq i \leqq n$, then $\| \mathbf{a} - \mathbf{a}' \|_p \leqq \theta(\Sigma_1^n a_i^p)^{1/p} \leqq \theta$. So given $\mathbf{a}$, let us define $\mathbf{a}'$ to be the vector with

$$a_i' = \min\{(1+\theta)^{-(j-1)} : j \geqq 1, (1+\theta)^{-(j-1)} \geqq a_i\}.$$

Let $\mathbf{a}'' \in l_p^n$ be defined by $a_i'' = \max\{a_i', (1+\theta)^{-k}\}$, where $k = 2\log_{1+\theta}(n^{1/p})$.

Note that $k \geqq \log_{1+\theta}(\theta^{-1}n^{1/p})$, so $(1+\theta)^{-k} \leqq \theta n^{-1/p}$. It follows easily that $\| \mathbf{a}'' - \mathbf{a}' \|_p \leqq \theta$, and therefore that $\| \mathbf{a}'' - \mathbf{a} \|_p \leqq 2\theta$. If, given any vector $\mathbf{a}$, we can find a vector $\mathbf{b}$ in Δ such that $\| \mathbf{b} - \mathbf{a}'' \| \leqq \theta$, then $\| \mathbf{b} - \mathbf{a} \| \leqq 3\theta = \delta$, so then Δ will be a δ-net. In other words, it is enough to approximate to within θ vectors of the form $\mathbf{a} = \Sigma_1^k \alpha_i \mathbf{u}_i$, where $\mathbf{u}_i = \chi_{U_i}$, for some sequence of possibly empty sets $U_1, \cdots, U_k$ satisfying $\bigcup_1^k U_i = [n]$, and $k_i < k_j$ whenever $i < j$, $k_i \in U_i$, $k_j \in U_j$.

Consider two vectors $\mathbf{a} = \Sigma_1^k \alpha_i \mathbf{u}_i$ and $\mathbf{a}' = \Sigma_1^k \alpha_i \mathbf{u}_i'$, where $(\mathbf{u}_i)_1^k$ and $(\mathbf{u}_i')_1^k$ are of the above form. Writing $\mathbf{v}_i = \Sigma_{j=1}^i \mathbf{u}_i$, $\mathbf{v}' = \Sigma_{j=1}^i \mathbf{u}_i'$, we have

$$\mathbf{a} = \Sigma_{i=1}^k (\alpha_i - \alpha_{i+1})\mathbf{v}_i, \quad \mathbf{a}' = \Sigma_{i=1}^k (\alpha_i - \alpha_{i+1})\mathbf{v}_i'$$

and

$$\mathbf{a} - \mathbf{a}' = \Sigma_{i=1}^k (\alpha_i - \alpha_{i+1})(\mathbf{v}_i - \mathbf{v}_i').$$

Now since $p \geqq 1$, $(\alpha_i - \alpha_{i+1} + x)^p - x^p$ is an increasing function of x (when $x \geqq 0$), so

$$\left\| \sum_{i=1}^{j} (\alpha_i - \alpha_{i+1})(\mathbf{v}_i - \mathbf{v}'_i) \right\|_p^p - \left\| \sum_{i=1}^{j-1} (\alpha_i - \alpha_{i+1})(\mathbf{v}_i - \mathbf{v}'_i) \right\|_p^p$$

$$\leqq (\alpha_j^p - \alpha_{j+1}^p)|\operatorname{supp}(\mathbf{v}_j - \mathbf{v}'_j)|.$$

Thus

$$\| \mathbf{a} - \mathbf{a}' \|_p^p \leqq \sum_{j=1}^{k} (\alpha_j^p - \alpha_{j+1}^p)|\operatorname{supp}(\mathbf{v}_j - \mathbf{v}'_j)|.$$

But

$$\| \mathbf{a} \|_p^p = \sum_{j=1}^{n} \alpha_p^p |\operatorname{supp}(\mathbf{u}_j)| = \sum_{j=1}^{k} (\alpha_j - \alpha_{j+1}^p)|\operatorname{supp}(\mathbf{v}_j)| \leqq 1 + 2\theta$$

and

$$\| \mathbf{a}' \|_p^p = \sum_{j=1}^{k} (\alpha_j^p - \alpha_{j+1}^p)|\operatorname{supp}(\mathbf{v}'_j)| \leqq 1 + 2\theta,$$

so N is at most the size of a θ^p-net of $(1 + 2\theta)B(l_1^k)$, i.e.

$$N \leqq (1 + 2(1 + 2\theta)/\theta^p)^k \leqq (5/\theta)^{pk}.$$

But since $k = 2 \log n/p \log(1 + \theta)$,

$$N \leqq n^{(2/\log(1+\theta))\log(5/\theta)}.$$ □

Assertion (ii) will be an easy consequence of the following lemma.

LEMMA 3. *Let $\delta > 0$, $2C\delta < 1$, let $\| \cdot \|$, $\| \cdot \|'$ and $|||\cdot|||$ be norms on $\mathbb{R}^m$ and let Δ be a δ-net of the unit ball of $|||\cdot|||$ such that for all $\mathbf{a}$ in Δ the following relations are satisfied*:

(a) $|||\mathbf{a}||| \leqq \| \mathbf{a} \| \leqq \| \mathbf{a} \|' \leqq C |||\mathbf{a}|||$;

(b) $\| \mathbf{a} \|' \leqq (1 + \delta) \| \mathbf{a} \|$.

Then for all $\mathbf{a}$ in $\mathbb{R}^m$, $\| \mathbf{a} \|' \leqq (1 + 6C\delta) \| \mathbf{a} \|$.

PROOF. Let $1 + \gamma = \sup\{ \| \mathbf{a} \|' : \| \mathbf{a} \| = 1 \}$. Then if $\| \mathbf{b} \| = 1$, choose $\mathbf{a}$ from Δ such that $|||\mathbf{a} - \mathbf{b}||| \leqq \delta$. Then $\| \mathbf{a} - \mathbf{b} \| \leqq C\delta$, so $\| \mathbf{a} \| \leqq 1 + C\delta$. Furthermore

$$\begin{aligned} \| \mathbf{b} \|' - 1 &= | \| \mathbf{b} \|' - \| \mathbf{b} \| | \\ &\leqq \| \mathbf{b} - \mathbf{a} \|' + | \| \mathbf{a} \|' - \| \mathbf{a} \| | + \| \mathbf{a} - \mathbf{b} \| \\ &\leqq (1 + \gamma)C\delta + \delta(1 + C\delta) + C\delta. \end{aligned}$$

Hence

$$\gamma \leqq C\delta\gamma + \delta(1 + 2C + C\delta)$$

and so, since $C\delta \leqq 1/2$,

$$\gamma \leqq \frac{3C\delta(1+\delta/3)}{1-C\delta} \leqq 6C\delta. \qquad \square$$

Assertion (ii) follows from Lemma 3 upon setting $|||\mathbf{a}||| \equiv \|\mathbf{a}\|_p$, $\|\mathbf{a}\| \equiv \|\phi_{\varepsilon,\pi}(\mathbf{a})\|$ and $\|\mathbf{a}\|' = \max\{\|\phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma})\| : \eta, \sigma \in \Psi\}$, since then $\|\mathbf{a}\|'$ is a 1-symmetric norm.

The probabilistic estimates are based on the following standard martingale inequality which can be found in many places in the literature, for example in [5].

LEMMA 4. *Let $f_0 = \mathbb{E}f, f_1, \ldots, f_n = f$ be a martingale with difference sequence $d_i = f_i - f_{i-1}$ satisfying $\|d_i\|_\infty \leqq c_i$ for $1 \leqq i \leqq n$. Then $\forall a > 0$,*

$$\mathbb{P}[f - \mathbb{E}f \geqq a] \leqq \exp\left(-\frac{a^2}{2\sum_{i=1}^n c_i^2}\right). \qquad \square$$

Let $(\Phi, d, \mathbb{P})$ be the metric probability space $\{-1, 1\}^n \times S_n$, where

$$d[(\varepsilon, \pi), (\varepsilon', \pi')] = \sum_{i=1}^n \{b_i : \varepsilon_i \neq \varepsilon_i' \text{ or } \pi(i) \neq \pi'(i)\}$$

for a sequence $b_1 \geqq \cdots \geqq b_n \geqq 0$, and the measure $\mathbb{P}$ is the normalized counting measure on Φ. Define equivalence relations $\sim_0, \ldots, \sim_n$ on Φ by $(\varepsilon, \pi) \sim_i (\varepsilon', \pi')$ iff $\varepsilon_j = \varepsilon_j'$ and $\pi(j) = \pi'(j)$ for $1 \leqq j \leqq i$. For $1 \leqq i \leqq n$ let $\mathscr{F}_i$ be the sigma-field whose atoms are the equivalence classes of $\sim_i$. Finally, let f be a γ-Lipschitz function on Φ, and set $f_i = \mathbb{E}(f \mid \mathscr{F}_i)$ $(1 \leqq i \leqq n)$. We have the following corollary of Lemma 4.

COROLLARY 5. *Let $(\Phi, d, \mathbb{P})$ and $f_0, \ldots, f_n$ be defined as above. Then for all $s > t$ and $\delta > 0$,*

$$\mathbb{P}[f_s - f_t \geqq \delta] \leqq \exp\left(-\frac{\delta^2}{8\gamma^2 \sum_{i=t+1}^s b_i^2}\right)$$

and

$$\mathbb{P}[f_s - f_t \leqq -\delta] \leqq \exp\left(-\frac{\delta^2}{8\gamma^2 \sum_{i=t+1}^s b_i^2}\right).$$

PROOF. We shall prove only the first inequality above, since the second is

similar (and indeed can easily be deduced from the first by looking at the function $-f$ instead of f). We restrict our attention to a single atom of $\mathscr{F}_r$. It is then obvious that without loss of generality $s = n$ and $r = 0$. By Lemma 4, we need only show that for $1 \leqq i \leqq n, f_i - f_{i-1} \leqq 2\gamma b_i$.

Suppose $A, B \in \mathscr{F}_i$, $A, B \subset C \in \mathscr{F}_{i-1}$, and let (η, σ) be an element of B. Then let ϕ be the bijection from A to B given by $(\varepsilon, \pi) \mapsto (\varepsilon', \pi')$, where

$$\varepsilon' = \begin{cases} \varepsilon_j, & j \neq i \\ \eta_j, & j = i \end{cases}$$

and $\pi' = \rho \circ \pi$, where ρ is the transposition $(\pi(i)\sigma(i))$.

Since $b_1 \geqq \cdots \geqq b_n \geqq 0$, and A and B are contained in the same atom of $\mathscr{F}_{r-1}$, (ε, π) and (ε', π') are equal except perhaps at i or $\pi^{-1}(\sigma(i))$, and $b_{\pi^{-1}(\sigma(i))} \leqq b_i$. Thus for any (ε, π) in A, $d((\varepsilon, \pi), \phi((\varepsilon, \pi))) \leqq 2b_r$.

Since f is γ-Lipschitz, f_i varies by at most $2\gamma b_i$ in any atom of $\mathscr{F}_{i-1}$, so $f_i - f_{i-1} \leqq 2\gamma b_i$ as was needed. □

Setting $\Phi = \Omega$, $b_i = a_{\lceil (i/h) \rceil}$ $(1 \leqq i \leqq n)$ and $s = \gamma_r h$, $t = \gamma_{r-1} h$, we have $2^{-r} \leqq b_i \leqq 2^{-(r-1)}$ for $\gamma_{r-1} \leqq i \leqq \gamma_r$. Set $f((\varepsilon, \pi)) = \| \phi_{\varepsilon,\pi}(\mathbf{a}_{\eta,\sigma}) \|$.

Since

$$(\varepsilon, \pi) \sim_s (\varepsilon', \pi') \Rightarrow \phi_{\varepsilon,\pi}(\mathbf{b}^j_{\eta,\sigma}) = \phi_{\varepsilon',\pi'}(\mathbf{b}^j_{\eta,\sigma}) \qquad (1 \leqq j \leqq r)$$

and

$$(\varepsilon, \pi) \sim_t (\varepsilon', \pi') \Rightarrow \phi_{\varepsilon,\pi}(\mathbf{b}^j_{\eta,\sigma}) = \phi_{\varepsilon',\pi'}(\mathbf{b}^j_{\eta,\sigma}) \qquad (1 \leqq j \leqq r-1)$$

and f is C-Lipschitz, we obtain from Corollary 5 that

$$\mathbb{P}[f^r_{\eta,\sigma}((\varepsilon, \pi)) - f^{r-1}_{\eta,\sigma}((\varepsilon, \pi)) > \delta_r h] < \exp\left(- \frac{\delta_r^2 h^2}{8C^2(\gamma_r - \gamma_{r-1})2^{-2(r-1)}h} \right)$$

$$= \exp\left(- \frac{2^{2(r-1)}\delta_r^2 h}{8C^2\beta_r} \right).$$

This establishes the result quoted for fact (b) above.

When $p > 1$, set $\Phi = \Omega$, $b_i = a^p_{\lceil (i/h) \rceil}$ $(1 \leqq i \leqq n)$, $s = \gamma_r h$, $t = \gamma_{r-1} h$ and $f((\varepsilon, \pi)) = g_{\eta,\sigma}((\varepsilon, \pi))$. This time f is 1-Lipschitz, $f_s = f^r_{\eta,\sigma}$ and $f_t = f^{r-1}_{\eta,\sigma}$. By Corollary 5, we therefore obtain (b)′.

It remains to prove assertion (c)′. Now setting $f \equiv g_{\eta,\sigma}$, we obtain

$$\mathbb{P}[f - \mathbb{E}f < -\delta_j] < \exp\left(- \frac{\delta^2 h}{8 \sum_1^m a_i^{2p}} \right).$$

Hence

$$\mathbb{P}[f=0] < \exp\left(-\frac{(\mathbb{E}f)^2}{8h}\right)$$

$$\Rightarrow \frac{1}{2} < \exp\left(-\frac{(\mathbb{E}f)^2}{8h}\right)$$

$$\Rightarrow \mathbb{E}f < (8h \log 2)^{1/2}$$

$$< (1/4)\cdot(\varepsilon/24C)^p h.$$

§3. The upper bound

The main interest of this section is negative: the upper bound presented shows that the results of the previous section cannot be substantially improved, but it is not particularly interesting in itself. Some of the details of the proof will therefore be omitted.

We shall begin with a simple but useful definition. Let $\varepsilon > 0$, let $(x_i)_1^n$ be a basis of a normed space $(X, \|\cdot\|)$, and let $(a_i)_1^n \in \mathbb{R}^n$. Then we shall say that $(x_i)_1^n$ is $(1+\varepsilon)$-*symmetric at* $(a_i)_1^n$ *under* $\|\cdot\|$ if for any (ε, π), (ε', π) in Ω,

$$\left\| \sum_1^n \varepsilon_i a_i x_{\pi(i)} \right\| \leqq (1+\varepsilon) \left\| \sum_1^n \varepsilon_i' a_i x_{\pi'(i)} \right\| .$$

We shall also say that $(x_i)_1^n$ is $(1+\varepsilon)$-*symmetric at* $\mathbf{a}$, where $\mathbf{a} = \Sigma_1^n a_i x_i$. (If either $(1+\varepsilon)$ or $\|\cdot\|$ is obvious from the context, it will sometimes be dropped. Thus we may say merely that $(x_i)_1^n$ is symmetric at $(a_i)_1^n$ or at $\mathbf{a}$.)

The aim of this section is to construct, for given $0 < \varepsilon < 1/2$ and $1 < p < \infty$, a 1-unconditional norm $\|\cdot\|$ on $\mathbb{R}^n$ which is 2-equivalent to $\|\cdot\|_p$, such that no block basis of the standard basis of $\mathbb{R}^n$ with cardinality exceeding $m_0 = 1000(1+p+q)\varepsilon^p n \log\log n/\log n$ is $(1 + 4^{-1/p}\varepsilon/3)$-symmetric, (where $1/p + 1/q = 1$). Now suppose $\|\cdot\|$ is any such norm, $m_1 \geqq m_0$ and $(\mathbf{u}_i)_1^{m_1}$ is a block basis of the standard basis of $\mathbb{R}^n$, which is 2-equivalent to the unit vector basis of $l_p^{m_1}$. Then it is easy to see that it has a sub-basis $(\mathbf{v}_j)_1^m$, where $m = n^{3/4}$, each vector $\mathbf{v}_j$ is supported on at most $h = \log n/640(1+p+q)\varepsilon^p \log\log n$ coordinates, and the norms of the $\mathbf{v}_j$ vary by at most $n^{-1/8}$. This fact will be used in the proof of Lemma 9 at the end of the section. Let us call a block basis $(\mathbf{u}_i)_1^m$ *proper* if $m = n^{3/4}$, no $\mathbf{u}_i$ is supported on more than h coordinates, and in addition if $\|\mathbf{u}_i\|_p = 1$ for each i.

The norm $\|\cdot\|$ is obtained in four steps as follows.

(i) Let $\mathscr{B}$ be a set system of cardinality to be defined later. Denote its

cardinality by N and let $r = \varepsilon^p n$. Let Γ be the collection $([n]^{(r)})^{\mathscr{B}}$ of all sequences of r-sets of $[n]$ of the form $(K_B : B \in \mathscr{B})$, and let $\mathbb{P}$ be the uniform distribution on Γ. For each $\gamma \in \Gamma$ define a norm $\|\cdot\|_\gamma$ 2-equivalent to $\|\cdot\|_p$, in a way to be explained.

(ii) For each proper block basis $(\mathbf{u}_i)_1^m$, define a sequence of vectors $(\mathbf{a}_B : B \in \mathscr{B})$ generated by the block basis, such that

$$p_B \equiv \mathbb{P}_\Gamma[(\mathbf{u}_i)_1^m \text{ is } (1 + 4^{-1/p}\varepsilon)\text{-symmetric at } \mathbf{a}_B \text{ under } \|\cdot\|_\gamma]$$

is at most $p = 2\binom{m}{l}(1 - (1/16)^{4\varepsilon^p h}(1 - \varepsilon^p/4)^{2h})^{m-l}$ for all $B \in \mathscr{B}$, and such that the probabilities p_B are independent.

(iii) Show that there are $M = (20n/4^{-1/p}\varepsilon)^{mh}$ proper block bases such that if for some γ they all fail to be $(1 + 4^{-1/p}\varepsilon)$-symmetric under $\|\cdot\|_\gamma$, then no block basis of cardinality exceeding m_0 is $(1 + 4^{-1/p}\varepsilon/3)$-symmetric under $\|\cdot\|_\gamma$.

(iv) Verify that $p^N < M^{-1}$.

Once we have completed these four steps, we easily obtain a basis 2-equivalent to the unit vector basis of l_p^n with no large $(1 + 4^{-1/p}\varepsilon/3)$-symmetric block basis. From step (ii) it follows that the probability that a given proper block basis is $(1 + 4^{-1/p}\varepsilon)$-symmetric is at most p^N, since in order to be symmetric, it must certainly be symmetric at all the $\mathbf{a}_B$. But since $Mp^N < 1$, there is some $\gamma \in \Gamma$ such that none of the block bases obtained in step (iii) is $(1 + 4^{-1/p}\varepsilon)$-symmetric under $\|\cdot\|_\gamma$. Our basis with no large symmetric block basis is then the standard basis of $\mathbb{R}^n$ with the norm $\|\cdot\|_\gamma$.

DEFINITION OF THE RANDOM NORM. Let $l = n^{1/2}$, $k = h^{1+p+q}\varepsilon^{-p/q}$, $\delta = h^{-q}\varepsilon^{1/q} = (h/k)^{1/p}$ and let $t = \log(1/2h)/\log k \geqq \log n/2(1 + p + q)\log\log n$, so that $h\sum_1^t k^i \leqq l$.

Then for $1 \leqq i \leqq t$, let A_i be the set of non-negative norm-1 vectors in l_p^n supported on at most hk^i points, whose coordinates are bounded above by $k^{-i/p}$. Note that all vectors in A_i are therefore supported on at least k^i coordinates.

Now let F_i be the set of support functionals for the vectors in A_i. That is, F_i is the set of non-negative norm-1 vectors in l_q^n supported on at most hk^i points, with all their coordinates bounded above by $k^{-i/q}$. So F_i is the set of vectors $\{|\mathbf{a}|^{p-1}\operatorname{sign}\mathbf{a} : \mathbf{a} \in A_i\}$.

We also define a second set of functionals G_i for each $1 \leqq i \leqq t$. It is the set of non-negative vectors of norm 1 in l_q^n supported on at most $\varepsilon^p hk^i$ points, whose coordinates are bounded above by $\varepsilon^{-p/q}k^{-i/q}$.

Now let $\mathscr{B}$ be a subset of $[t]^{(t/2)}$ such that whenever B, C are in $\mathscr{B}$ and $B \neq C$, then $|B \cap C| \leqq t/3$, and let $\mathscr{B}$ have cardinality $N = (23/20)^t$. (This will be shown to be possible in Lemma 12 later.) For any $B \in \mathscr{B}$ we define classes A_B, F_B and G_B as follows:

$$A_B = \left\{ \bigoplus_{i \in B} \mathbf{a}_i : \mathbf{a}_i \in A_i \ \forall i \in B \right\}, \quad F_B = \left\{ \bigoplus_{i \in B} f_i : f_i \in F_i \ \forall i \in B \right\},$$

$$G_B = \left\{ \bigoplus_{i \in B} g_i : g_i \in G_i \ \forall i \in B \right\}$$

where $\oplus$ denotes a sum with disjoint supports.

Finally, let $r = \varepsilon^p n$ and let $\gamma = (K_B : B \in \mathscr{B})$ be an element of $\Gamma = ([n]^{(r)})^{\mathscr{B}}$. We define $\| \cdot \|_\gamma$ on $\mathbb{R}^n$ as follows:

$$\| x \|_\gamma = \| x \|_p \vee \max_{B \in \mathscr{B}} [(2/t)^{1/q} \max\{ f(x) + g(x) : f \in F_B, g \in G_B, \operatorname{supp}(g) \subset K_B \}].$$

It is clear that $\| \cdot \|_\gamma$ is 2-equivalent to $\| \cdot \|_p$. The motivation for this definition of $\| \cdot \|_\gamma$ is as follows. The classes A_i, F_i and G_i $(1 \leqq i \leqq t)$ are defined so that a functional in F_i or G_i can only be large at a vector in A_j if $i = j$. Then if $B, C \in \mathscr{B}$, $B \neq C$, it follows from the separation of B and C that a functional in F_B or G_B cannot be made to fit a vector in A_C all that well, or in other words cannot be large at such a vector. This argument is made more precise in the proof of Lemma 11 later, where it is shown that if $C \in \mathscr{B}$ and $\mathbf{a} \in A_C$, then

$$\| \mathbf{a} \|_\gamma = \| \mathbf{a} \|_p \vee (2/t)^{1/q} \max\{ f(\mathbf{a}) + g(\mathbf{a}) : f \in F_C, g \in G_C, \operatorname{supp}(g) \subset K_C \}. \tag{2}$$

But since the subsets K_B are chosen independently, it follows from (2) (which for now we shall assume) that the probabilities p_B defined in (ii) are independent.

To make the statement of Lemma 6 easier, we introduce the following definition. Given a proper block basis $(\mathbf{u}_i)_1^m$ and a set $K \in [n]^{(r)}$ we shall say that a vector $\mathbf{u}_i$ from the block basis is *large on* K if it can be restricted to a vector $\mathbf{u}_i'$ with $\operatorname{supp}(\mathbf{u}_i') \leqq \varepsilon^p h$, $\operatorname{supp}(\mathbf{u}_i') \subset K$ and $\| \mathbf{u}_i' \|_K \|_p \geqq 4^{-1/p} \varepsilon$. The reason for this rather artificial seeming definition will become clear in the statement and proof of Lemma 6.

LEMMA 6. *Let* $\gamma = (K_B : B \in \mathscr{B}) \in \Gamma$, *and let* $(\mathbf{u}_i)_1^m$ *be a proper normalized*

block basis of the standard basis of $\mathbb{R}^n$ *such that there exist two sequences* $i_1, \dots, i_l$ *and* $j_1, \dots, j_l$ *and* $B \in \mathscr{B}$ *with* $\mathbf{u}_{i_1}, \dots, \mathbf{u}_{i_l}$ *all large on* K_B, *while* $\| \mathbf{u}_{j_s} |_{K_B} \|_p = 0$ *for* $1 \leqq s \leqq l$. *Then there is a sequence* $(a_i)_1^m$ *such that* $\Sigma_1^m a_i \mathbf{u}_i \in A_B$ *and* $(\mathbf{u}_i)_1^m$ *fails to be* $(1 + 4^{-1/p})$*-symmetric at* $(a_i)_1^m$ *under* $\| \cdot \|_\gamma$.

PROOF. Given the assumption (2) above, the proof is simple. First, note that for any $X_j \subset [m]$ with $|X_j| = k^j$ we have $|X_j|^{-1/p} \Sigma_{i \in X_j} \mathbf{u}_i \in A_j$. Hence, if the sets X_j $(j \in \mathscr{B})$ are all disjoint, and we set $a_i = (\Sigma_{j \in B} |X_j|^{-1/p} \chi_{X_j})_i$, we have $\mathbf{a} = \Sigma_1^m a_i \mathbf{u}_i \in A_B$. Furthermore, since $h \Sigma_1^l k^i \leqq l$, all but l of the a_i are zero, so without loss of generality $X_j \subset [l]$ for each $j \in B$ and thus $a_{l+1} = \cdots = a_m = 0$. Let us write $\mathbf{a}' = \Sigma_{s=1}^l a_s \mathbf{u}_{i_s}$ and $\mathbf{a}'' = \Sigma_{s=1}^l a_t \mathbf{u}_{j_s}$. We shall then estimate $\| \mathbf{a}' \|_\gamma$ and $\| \mathbf{a}'' \|_\gamma$.

First, let us calculate $\max\{g(\Sigma_{s \in X_j} a_s \mathbf{u}_{i_s}) : g \in G_j, \operatorname{supp}(g) \subset K_B\}$ when $j \in \mathscr{B}$. Write $\mathbf{b}_j = \Sigma_{s \in X_j} a_s \mathbf{u}_{i_s}$. Then $\mathbf{b}_j \in A_j$ and since $\mathbf{u}_{i_1}, \dots, \mathbf{u}_{i_l}, \dots, \mathbf{u}_{i_l}$ are large on K_B, we can restrict $\mathbf{b}_j$ to a vector $\mathbf{b}_j'$ satisfying $\operatorname{supp}(\mathbf{b}_j') \subset K_B$, $|\operatorname{supp}(\mathbf{b}_j')| \leqq \varepsilon^p h k^j$ and $\| \mathbf{b}_j' \|_p \geqq 4^{-1/p} \varepsilon \| \mathbf{b}_j \|_p = 4^{-1/p} \varepsilon$. We can therefore find $g \in G_j$ such that $g(\mathbf{b}_j) \geqq 4^{-1/p} \varepsilon$. But then we can find $g \in G_B$ such that $g(\mathbf{a}') \geqq 4^{-1/p} \varepsilon t/2$. It is obvious that we can find $f \in F_B$ such that $f(\mathbf{a}') \geqq t/2$, so $\| \mathbf{a}' \|_\gamma \geqq (t/2)^{1/p}(1 + 4^{-1/p}\varepsilon)$.

Now $\| \mathbf{a}'' \|_p = (t/2)^{1/p}$, and $\operatorname{supp}(\mathbf{a}'') \cap K_B = \varnothing$. But for any $f \in F_B$, $\| f \| = (t/2)^{1/q}$, so therefore $\| \mathbf{a}'' \|_\gamma \leqq (t/2)^{1/p}$. Hence $(\mathbf{u}_i)_1^m$ fails to be $(1 + 4^{-1/p}\varepsilon)$-symmetric at $\mathbf{a}$, which proves the lemma. □

We shall show that the sequences needed in the conditions of Lemma 6 exist with very high probability. We need a probabilistic estimate whose proof is rather technical and deferred to the end of the section.

LEMMA 7. *Let* $(\mathbf{u}_i)_1^m$ *be a proper block basis and let* $B \in \mathscr{B}$. *Then the probability that we can find* $i_1, \dots, i_l$ *such that* $\mathbf{u}_{i_1}, \dots, \mathbf{u}_{i_l}$ *are large on* K_B *is at least*

$$1 - \binom{m}{l} (1 - (1/16)^{4\varepsilon^p h}(1 - \varepsilon^p/4)^{2h})^{m-l}$$

and the probability that we can find $j_1, \dots, j_l$ *such that the restrictions of* $\mathbf{u}_{j_1}, \dots, \mathbf{u}_{j_l}$ *to* K_B *are zero is at least*

$$1 - \binom{m}{l} (1 - (1 - \varepsilon^p/4)^h)^{m-l}.$$

COROLLARY 8. *Let* $(\mathbf{u}_i)_1^m$ *be a proper block basis. Then the probability that* $(\mathbf{u}_i)_1^m$ *is* $(1 + 4^{-1/p}\varepsilon)$*-symmetric under* $\| \cdot \|_\gamma$ *is at most*

$$\left[2\binom{m}{l}(1-(1/16)^{4\varepsilon^p h}(1-\varepsilon^p/4)^{2h})^{m-l}\right]^N.$$

PROOF. For each $B\in\mathscr{B}$, construct $\mathbf{a}_B\in A_B$ as in the proof of Lemma 6. By (2), $\|\mathbf{a}_B\|_\gamma$ depends only on K_B (where $\gamma=(K_B: B\in\mathscr{B})$), so by Lemmas 6 and 7,

$$p_B\leqq\binom{m}{l}(1-(1/16)^{4\varepsilon^p h}(1-\varepsilon^p/4)^{2h})^{m-l}+\binom{m}{l}(1-(1-\varepsilon^p/4)^h)^{m-l}$$

$$\leqq 2\binom{m}{l}(1-(1/16)^{4\varepsilon^p h}(1-\varepsilon^p/4)^{2h})^{m-l}$$

and if $B\neq C$ then p_B and p_C are independent. The result follows immediately. □

The next lemma is a precise statement of step (iii) earlier. It is proved at the end of the section.

LEMMA 9. *Let $0<\eta<1$ and $M=(20n/\eta)^{mh}$. Then there are proper block bases $(\mathbf{u}_i^1)_{i=1}^m,\dots,(\mathbf{u}_i^M)_{i=1}^m$ of the standard basis of $\mathbb{R}^n$ such that any norm 2-equivalent to $\|\cdot\|_p$ which fails to be $(1+\eta)$-symmetric on any of $(\mathbf{u}_i^1)_{i=1}^m,\dots,(\mathbf{u}_i^M)_{i=1}^m$ fails to be $(1+\eta/3)$-symmetric on any block basis of cardinality exceeding m_0.*

The proof of the upper bound is now a simple matter of verification.

THEOREM 10. *Let $0<\varepsilon<1/2$. Then there exists a norm $\|\cdot\|$ on $\mathbb{R}^n$ such that for any $x\in\mathbb{R}^n$, $\|x\|_p\leqq\|x\|\leqq 2\|x\|_p$, but no $\|\cdot\|_p$-normalized block basis of cardinality exceeding m_0 is $(1+\delta)$-symmetric for any $\delta<4^{-1/p}\varepsilon/3$.*

Proof. By Corollary 8 and Lemma 9 (with $\eta=4^{-1/p}\varepsilon$) it remains only to show that

$$\left[2\binom{m}{l}(1-(1/16)^{4\varepsilon^p h}(1-\varepsilon^p/4)^h)^{m-l}\right]^N<\left(\frac{20n}{4^{-1/p}\varepsilon}\right)^{-mh}.$$

From this it will follow that there is at least one $\gamma\in\Gamma$ such that no block basis of cardinality exceeding m_0 is symmetric under $\|\cdot\|_\gamma$.

Now

$$\left[2\binom{m}{l}(1-(1/16)^{4\varepsilon^p h}(1-\varepsilon^p/4)^h)^{m-l}\right]^N$$

$$\leqq 2^N m^{lN}\exp(-(m-l)N(1/16)^{4\varepsilon^p h}(1-\varepsilon^p/4)^h)$$

$$=\exp(N(\log 2 + l\log m-(m-l)(1/16)^{4\varepsilon^p h}(1-\varepsilon^p/4)^h)).$$

But since $h \leqq \log n/80\,\varepsilon^p$, this is at most

$$\exp(-(1/2)N(m-l)(1/16)^{4\varepsilon^p h}(1-\varepsilon^p/4)^h)$$

$$\leqq \exp(-(1/2)N(m-l)\exp(-\varepsilon^p h(\log 16 + 1/4))).$$

Now it is easy to check that $h \leqq \log N/40\varepsilon^p$, so this is at most $\exp(-N^{1/2}(m-l))$ which is certainly at most $(20n/4^{-1/p}\varepsilon)^{-mh}$. □

We are left with the task of proving the various lemmas and assertions assumed without proof earlier.

LEMMA 11. *Let $C\in\mathscr{B}$ and let $\mathbf{a}$ be a vector in A_C. Then for all $\gamma\in\Gamma$, $\gamma=(K_B: B\in\mathscr{B})$,*

$$\|\mathbf{a}\|_\gamma = \|\mathbf{a}\|_p \vee (2/t)^{1/q}\max\{f(\mathbf{a})+g(\mathbf{a}): f\in F_C, g\in G_C, \operatorname{supp}(g)\subset K_C\}.$$

PROOF. It is enough to prove that if $B, C\in B$, $B\neq C$, $f\in F_B$, $g\in G_B$ and $\mathbf{a}\in A_C$, then $f(\mathbf{a})+g(\mathbf{a})\leqq (t/2)^{1/q}\|\mathbf{a}\|_p$. It is simple to show that if x and y are two vectors in $\mathbb{R}^n$, then $\{x_{\varepsilon,\pi}, y_{\varepsilon',\pi'}\rangle$ is maximized when $x_{\varepsilon,\pi}$ and $y_{\varepsilon',\pi'}$ are both non-negative decreasing vectors. We shall therefore assume this of f, g and $\mathbf{a}$. Let us write $f=\Sigma_{i\in B} f_i$ with $f_i\in F_i$ for each i, and similarly write $g=\Sigma_{i\in B} g_i$ and $\mathbf{a}=\Sigma_{j\in C}\mathbf{a}_j$. In order to estimate $(f+g)(\mathbf{a})$ we shall estimate $f_i(\mathbf{a})$ and $g_i(\mathbf{a})$ in the cases $i\in B\cap C$ and $i\in B\setminus C$.

First let us look at $f_i(\Sigma_{j\in C}\mathbf{a}_j)$ in the case $i\in B\cap C$. Writing f_i' for the decreasing rearrangement of f_i (i.e. for f_i without the string of zeros at the front) we have

$$f_i\left(\sum_{j\in C}\mathbf{a}_j\right)\leqq f_i'\left(\sum_{j\in C}\mathbf{a}_j\right)$$

$$=f_i'\left(\sum_{j\in C, j<i}\mathbf{a}_j\right)+f_i'(\mathbf{a}_i)$$

$$\leqq \|f_i'\|_\infty\left\|\sum_{j\in C, j<i}\mathbf{a}_j\right\|_1+1.$$

Now $\| f_i' \|_\infty = \| f_i \|_\infty \leqq k^{-i/q}$, and $\| \Sigma_{j\in C, j<i} \mathbf{a}_j \|_1 = \Sigma_{j\in C, j<i} \| \mathbf{a}_j \|_1$. Since $\mathrm{supp}(\mathbf{a}_j) \leqq hk^j$ and $\| \mathbf{a}_j \|_p = 1$, $\| \mathbf{a}_j \|_1 \leqq h^{1/q} k^{j/q}$. Thus

$$\sum_{j\in C, j<i} \| \mathbf{a}_j \|_1 \leqq \sum_{j=1}^{i-1} h^{1/q} k^{j/q} \leqq \frac{h^{1/q} k^{i/q}}{k^{1/q} - 1}$$

and hence

$$f_i\left(\sum_{j\in C, j<i} \mathbf{a}_j \right) \leqq \frac{h^{1/q}}{k^{1/q} - 1} + 1 \leqq 2h^{1/q} k^{-1/q} + 1. \tag{3}$$

Now suppose $i \in B \setminus C$. This time

$$f_i\left(\sum_{j\in C} \mathbf{a}_j \right) \leqq f_i'\left(\sum_{j\in C, j<i} \mathbf{a}_j \right) + f_i'(\mathbf{a}_k),$$

where k is minimal such that $k > i$, $k \in C$.

We have already estimated the first term. Also

$$\begin{aligned} f_i(\mathbf{a}_k) &\leqq \| f_i' \|_1 \| \mathbf{a}_k \|_\infty \leqq h^{1/p} k^{i/p} \cdot k^{-(i+1)/p} \\ &\leqq h^{1/p} k^{-1/p} = \delta, \end{aligned}$$

so

$$f_i\left(\sum_{j\in C} \mathbf{a}_j \right) \leqq 2h^{1/q} k^{-1/q} + \delta. \tag{4}$$

It follows from (3) and (4) that

$$\left(\sum_{i\in B} f_i \right)\left(\sum_{j\in C} \mathbf{a}_j \right) \leqq th^{1/q} k^{-1/q} + |B \cap C| + \delta |B \setminus C|.$$

If $B \neq C$ then $|B \cap C| \leqq t/3$, so $f(\mathbf{a})$ is at most $t/3 + \delta t/2$.

The calculations for $g(\mathbf{a})$ are very similar. When $i \in B \cap C$ we obtain

$$g_i\left(\sum_{j\in C} \mathbf{a}_j \right) \leqq \varepsilon + 2h^{2/q} \varepsilon^{-p/q} k^{-1/q},$$

and when $i \in B \setminus C$, then

$$g_i\left(\sum_{j\in C} \mathbf{a}_j \right) \leqq 2h^{2/q} \varepsilon^{-p/q} k^{-1/q} + \varepsilon k^{-1/p}.$$

It follows that

$$\left(\sum_{i\in B} g_i \right)\left(\sum_{j\in C} \mathbf{a}_j \right) \leqq th^{2/q} \varepsilon^{-p/q} k^{-1/q} + \varepsilon |B \cap C| + \varepsilon k^{-1/p} |B \setminus C|.$$

If $B \neq C$ then this gives

$$g(\mathbf{a}) \leqq th^{2/q}\varepsilon^{-p/q}k^{-1/q} + \varepsilon t/3 + \varepsilon k^{-1/p}t/6.$$

Now $\| \mathbf{a} \|_p = (t/2)^{1/p}$, so

$$(2/t)^{1/q} \| \mathbf{a} \|_p^{-1}(f+g)(\mathbf{a}) = (2/t)(f+g)(\mathbf{a})$$
$$\leqq 2/3 + \delta + 2h^{2/q}\varepsilon^{-p/q}k^{-1/q} + 2\varepsilon/3 + \varepsilon k^{-1/p}/3.$$

But since $\varepsilon < 1/3$ and $\delta \leqq h^{-q}$, this is at most 1, as required. □

In order to prove Lemma 7, we shall prove a subsidiary lemma which can be regarded as a rather weak generalization of a lower bound for the hypergeometric distribution. It is exactly such a bound when all the non-zero coordinates of the vector $\mathbf{a}$ are equal.

LEMMA 7a. *Let $r = \varepsilon n$, $h \leqq r$ and let $\mathbf{a} = (a_i)_1^n \in \mathbb{R}^n$ be a vector such that $a_1 \geqq \cdots \geqq a_h \geqq a_{h+1} = \cdots = a_n = 0$ and $\Sigma_1^n a_i = 1$. Then if $t < r/4$ and K is chosen randomly from $[n-t]^{(r-t)}$, then*

$$\mathbb{P}\left[\sum_{i\in K} a_i \geqq 2\varepsilon\right] \geqq (1/16)^{32\varepsilon h} \cdot (1-2\varepsilon)^{2h} \quad \text{and} \quad \mathbb{P}\left[\sum_{i\in K} a_i = 0\right] \geqq (1-2\varepsilon)^h.$$

PROOF. Note that $\mathbb{E}(\Sigma_{i\in K} a_i) = \varepsilon$. Clearly

$$\mathbb{P}\left[\sum_{i\in K} a_i = 0\right] \geqq \binom{n-r}{h} \Big/ \binom{n-t}{h} \geqq \left(\frac{n-r-h}{n-h-t}\right)^h \geqq (1-2\varepsilon)^h$$

as stated.

For the first estimate, we use the lower bounds for the hypergeometric distribution given in [3] p. 8. For $l \leqq r/2$, $0 < \alpha < 1$ we obtain

$$\binom{r-t}{\alpha l}\binom{n-r}{(1-\alpha)l} \Big/ \binom{n}{l} \geqq \binom{l}{\alpha l}\left(\frac{r-t-\alpha l}{n}\right)^{\alpha l}\left(\frac{n-r-(1-\alpha)l}{n}\right)^{(1-\alpha)l}$$
$$= \binom{l}{\alpha l}\left(\frac{r-t-\alpha l}{n-r-(1-\alpha)l}\right)^{\alpha l}\left(\frac{n-r-(1-\alpha)l}{n}\right)^{l}$$
$$\geqq (1/\alpha)^{\alpha l}(\varepsilon/2)^{\alpha l}(1-2\varepsilon)^l = (\varepsilon/2\alpha)^{\alpha l}(1-2\varepsilon)^l.$$

Now let $B_0 \subset B_1 \subset \cdots \subset B_s \subset \{1, \ldots, h\}$ be defined by $B_0 = \varnothing$ and $B_j = \{i \in [n] : a_i \geqq 2^{-j}\}$ $(1 \leqq j \leqq s)$, where $s = \log_2(2h)$.

Suppose $|B_j \cap K| \geqq 8\varepsilon |B_j|$ for $1 \leqq j \leqq s$. In this case

$$\sum_{i\in K} a_i \geqq \sum_{j=1}^{s} \sum \{a_i : i \in (B_j \setminus B_{j-1}) \cap K\}$$

$$\geqq \sum_{j=1}^{s} 2^{-j} |(B_j \cap K) \setminus (B_{j-1} \cap K)|$$

$$\geqq \sum_{j=1}^{s} |B_j \cap K| (2^{-j} - 2^{-(j+1)})$$

$$\geqq 4\varepsilon \sum_{j=1}^{s} 2^{-j} |B_j|$$

$$= 4\varepsilon \sum_{j=1}^{s} (2^{-(j-1)} - 2^{-j}) |B_j|$$

$$\geqq 4\varepsilon \sum_{j=1}^{s} 2^{-(j-1)} |B_j \setminus B_{j+1}| \geqq 2\varepsilon$$

(since $\Sigma\{a_i : a_i \leqq 2^{-s}\} \leqq 1/2$).

But

$$\mathbb{P}[|B_j \cap K| \geqq 8\varepsilon |B_j| \mid |B_{j+1} \cap K| \geqq 8\varepsilon |B_{j+1}|]$$

$$\geqq \binom{|B_j|}{8\varepsilon |B_j|} \binom{n-r}{(1-8\varepsilon)|B_j|} \Big/ \binom{n-t}{|B_j|}$$

$$= (1/16)^{8\varepsilon |B_j|} (1-2\varepsilon)^{|B_j|} \qquad (1 \leqq j < s)$$

and

$$\mathbb{P}[|B_s \cap K| \geqq 8\varepsilon |B_s|] \geqq (1/16)^{8\varepsilon |B_s|} (1-2\varepsilon)^{|B_s|}.$$

Hence

$$\mathbb{P}[|B_j \cap K| \geqq 8\varepsilon |B_j| \ \forall\, 1 \leqq j \leqq s] \geqq (1/16)^{8\varepsilon \Sigma_1^s |B_j|} \cdot (1-2\varepsilon)^{\Sigma_1^s |B_j|}.$$

But $|B_j| \leqq 2^j$, so $\Sigma_1^s |B_j| \leqq 4h$, and so

$$\mathbb{P}\left[\sum_{i\in K} a_i \geqq 2\varepsilon\right] \geqq (1/16)^{32\varepsilon h} \cdot (1-2\varepsilon)^{2h}. \qquad \square$$

In fact we obtain more from the proof of Lemma 7a. Notice that if $|B_j \cap K| \geqq 8\varepsilon |B_j|$, for $j = 1, \ldots, s$, then we can restrict $\mathbf{a}$ to a vector $\mathbf{a}'$ such that, with the obvious definitions of $B_1', \ldots, B_s'$, $|B_j' \cap K| = 8\varepsilon |B_j|$ for each j, and $B_1' \subset \cdots \subset B_s'$. Thus there is a vector $\mathbf{a}'$ with $|\mathrm{supp}(\mathbf{a}')| \leqq 8\varepsilon h$, $\mathrm{supp}(\mathbf{a}') \subset$

K and $\Sigma_{i\in K} a_i' \geqq 2\varepsilon$, with probability at least $(1/16)^{32\varepsilon h}\cdot(1-2\varepsilon)^{2h}$. This observation allows us to prove Lemma 7 with ease.

PROOF OF LEMMA 7. We would like to estimate

$$p_1 \equiv \mathbb{P}[\mathbf{u}_i \text{ is large on } K \text{ for at most } l \text{ values of } i],$$

when K is chosen randomly from $[n]^{(r)}$ and $r = \varepsilon^p n$.

Setting $\mathbf{a} = |\mathbf{u}_i|^p$ and taking $\varepsilon^p/8$ instead of ε in Lemma 7a and the remarks following it, we obtain

$$\mathbb{P}[\mathbf{u}_i \text{ is large on } K \mid \operatorname{supp}(\mathbf{u}_j) \cap K = W_j \text{ for } 1 \leqq j < i]$$

is at least $(1/16)^{4\varepsilon^p h}\cdot(1-\varepsilon^p/4)^{2h}$, and so

$$p_1 \leqq \binom{m}{l}(1-(1/16)^{4\varepsilon^p h}(1-\varepsilon^p/4)^{2h})^{m-l}.$$

Similarly,

$$p_2 \equiv \mathbb{P}[\,\|\,\mathbf{u}_i\,|_K\,\|_p = 0 \text{ for at most } l \text{ values of } i]$$

$$\leqq \binom{m}{l}(1-(1-\varepsilon^p/4)^h)^{m-l}. \qquad \square$$

Next we shall prove Lemma 9.

PROOF OF LEMMA 9. Let us call two block bases $(\mathbf{u}_i)_1^m$ and $(\mathbf{v}_i)_1^m$ *α-close* if they satisfy $\operatorname{supp}(\mathbf{u}_i)\cap\operatorname{supp}(\mathbf{v}_i)=\varnothing$ whenever $i\neq j$ and $\|\mathbf{u}_i-\mathbf{v}_i\|_p \leqq \alpha$ for each i. Suppose also, without loss of generality, that for any $x\in\mathbb{R}^n$, $\|x\|_p \leqq \|x\| \leqq 2\|x\|_p$. Now if $(\mathbf{u}_i)_1^m$ and $(\mathbf{v}_i)_1^m$ are α-close then given any sequence $(a_i)_1^m\in\mathbb{R}^m$,

$$\begin{aligned}\left|\left\|\sum_1^m a_i\mathbf{u}_i\right\| - \left\|\sum_1^m a_i\mathbf{v}_i\right\|\right| &\leqq 2\left|\left\|\sum_1^m a_i\mathbf{u}_i\right\|_p - \left\|\sum_1^m a_i\mathbf{v}_i\right\|_p\right| \\ &\leqq 2\left\|\sum_1^m a_i(\mathbf{u}_i-\mathbf{v}_i)\right\|_p \\ &= 2\left(\sum_1^m |a_i|^p\,\|\mathbf{u}_i-\mathbf{v}_i\|_p^p\right)^{1/p} \\ &\leqq 2\alpha\left(\sum_1^m |a_i|^p\right)^{1/p}.\end{aligned}$$

Since $0<\eta<1$, it follows that if $(\mathbf{u}_i)_1^m$ and $(\mathbf{v}_i)_1^m$ are $\eta/8$-close and $(\mathbf{u}_i)_1^m$ is

not $(1+\eta)$-symmetric under $\|\cdot\|$ then $(\mathbf{v}_i)_1^m$ is not $(1+2\eta/5)$-symmetric under $\|\cdot\|$.

Similarly, if $(\mathbf{v}_i)_1^m$ and $(\mathbf{w}_i)_1^m$ are $n^{-1/8}$-close and $(\mathbf{v}_i)_1^m$ is not $(1+2\eta/5)$-symmetric under $\|\cdot\|$, then $(\mathbf{w}_i)_1^m$ is not $(1+\eta/3)$-symmetric under $\|\cdot\|$.

Now the number of ways of choosing m disjoint sets of size h from $[n]$ is certainly no more than n^{mh}, and there is an $\eta/6$-net of the unit sphere of l_p^h of cardinality at most $(15/\eta)^h$. It is thus easy to see that with $M=(20n/\eta)^{mh}$, there are proper block bases $(\mathbf{u}_i^1)_{i=1}^m, \ldots, (\mathbf{u}_i^M)_{i=1}^m$ such that any proper block basis is $\eta/8$-close to $(\mathbf{u}_i^r)_{i=1}^m$ for some $1 \leqq r \leqq M$. But by the remarks at the beginning of the section, any block basis of cardinality exceeding m_0 has a sub-basis, a multiple of which is $n^{-1/8}$-close to a proper block basis. The result follows. □

Finally, we prove the simple fact that $\mathscr{B}$ may be taken to be $(23/20)^t$.

LEMMA 12. *There is a subset $\mathscr{B} \subset [t]^{(t/2)}$ of cardinality $(23/20)^t$ such that given any two distinct sets $B, C \in \mathscr{B}$, $|B \cap C| \leqq t/3$.*

PROOF. For any $B \in [t]^{(t/2)}$, the number of $C \in [t]^{(t/2)}$ such that $|B \cap C| > t/3$ is at most

$$\sum_{r=0}^{t/6} \binom{t/2}{t/6-r}\binom{t/2}{t/3+r} \leqq 3\binom{t/2}{t/3}\binom{t/2}{t/6}.$$

Hence, by picking sets one at a time, each one disjoint from the previous ones, we can find $\mathscr{B}$ with

$$|\mathscr{B}| \geqq \frac{1}{3}\binom{t}{t/2}\Big/\binom{t/2}{t/3}\binom{t/2}{t/6}.$$

But by the estimates for the hypergeometric distribution in [3],

$$\begin{aligned}\binom{t/2}{t/3}\binom{t/2}{t/6}\Big/\binom{t}{t/2} &\leqq \binom{t/2}{t/3}(1/2)^{t/3}(1/2)^{t/6}(2/3)^{-t/6} \\ &= \binom{t/2}{t/3}(1/2)^{t/3}(3/4)^{t/6} \\ &\leqq (3/2)^{t/3}(1/2)^{t/3}(3/4)^{t/6} = (3/4)^{t/2}\end{aligned}$$

so $|\mathscr{B}| \geqq \frac{1}{3}(4/3)^{t/2} \geqq (23/20)^t$. □

§4. Remarks

(1) When $1 < p < \infty$, the upper bound can be improved to $n/\log n$ if it is only required that no large block basis with ± 1-coefficients should be symmetric. This is because one then has more control over the vectors that a block basis must generate. By methods similar to those of Section 3, an upper bound proportional to $n/\log\log n$ can be obtained in the case $p = 1$. This case is different because the unit ball of l_1^n is not convex enough to make it possible to vary the norms of nearly so many vectors independently. It is not clear whether this fact could be the basis of an improved lower bound.

There is also no obvious way of exploiting significantly the original distance from l_p^n for the purposes of the upper bound. In fact, the norm constructed in Section 3 can be shown, by methods similar to the proof of Lemma 11, to be $(1 + 2\varepsilon)$-equivalent to the p-norm. This is a similar difficulty, since increasing the distance from l_p^n hardly helps to increase the number of vectors whose norms can be varied independently.

(2) It is natural to ask what the correct answer is when $p = \infty$. A simple argument, which can be found in [4] pp. 50–51, shows that if $(x_i)_1^n$ is a basis which is C-equivalent to the unit vector basis of l_∞^n, then for an absolute constant α, it has a block basis of cardinality $k = n^{\alpha \log(1+\varepsilon)/\log C}$ which is $(1 + \varepsilon)$-equivalent to the unit vector basis of l_∞^k, and which is *a fortiori* $(1 + \varepsilon)$-symmetric. By considering the space l_p^n for $p = \log n/\log C$, it is easy to show that this result is best possible. Work in progress strongly suggests that even if one only wants symmetry, the correct upper bound is still a power of n which depends on the parameters ε and C.

(3) Other results of Amir and Milman in [1] and [2] can be improved by methods very similar to those of this paper. For example, it can be shown that if $1 \leqq p < 2$ and $(x_i)_1^n$ is a basis for a normed space X such that the p-type constant of X is γ, and for some $c > 0$,

$$\mathbb{E}\left\|\sum_1^n \varepsilon_i x_i\right\| \geqq cn^{1/p},$$

then $(x_i)_1^n$ has a $(1 + \varepsilon)$-symmetric block basis of cardinality proportional to $n^{2/p-1}/\log n$. No interesting upper bound is known, other than the upper bounds in this paper for isomorphis of l_p^n. As was pointed out by Amir and Milman, the deterioration when p approaches 2 is necessary, since a constant sequence of length n in l_2^n satisfies the given conditions.

Acknowledgement

The author would like to thank his supervisor Dr B. Bollobás for his constant encouragement and for suggesting this problem.

References

1. D. Amir and V. D. Milman, *Unconditional and symmetric sets in n-dimensional normed spaces*, Isr. J. Math. **37** (1980), 3–20.

2. D. Amir and V. D. Milman, *A quantitative finite dimensional Krivine theorem*, Isr. J. Math. **50** (1985), 1–12.

3. B. Bollobás, *Random Graphs*, Academic Press, New York, 1985.

4. B. Maurey and G. Pisier, *Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach*, Studia Math. **58** (1976), 45–90.

5. V. D. Milman and G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Lecture Notes in Mathematics, Vol. 1200, Springer-Verlag, Berlin, 1986.