PhD 16056

SYMMETRIC STRUCTURES

IN BANACH SPACES

W. T. Gowers

A dissertation submitted for the degree of

Doctor of Philosophy, December 1989

William Timothy
GOWERS
Trinity College

SYMMETRIC STRUCTURES IN BANACH SPACES

W. T. Gowers

The backbone of the dissertation is a series of results to do with finding almost symmetric block bases of sequences which satisfy certain commonly occurring conditions. This has been known to be possible for some years: in 1982 and 1985 Amir and Milman published two important papers containing various results on the subject. My contribution has been to find substantially larger block bases and to construct examples in many cases to show that one cannot improve my new results further. Probabilistic methods play an important part in calculating bounds in both directions.

These results were originally motivated by a theorem of Krivine, which can be regarded as the finite-dimensional analogue of the well-known distortion problem. Also in the dissertation is an analogue of the distortion problem in c_0, which strengthens considerably a result of James, and indicates that, contrary to what is generally believed, the answer to the distortion problem itself could very well be positive.

There is also a counterexample to a fairly long-standing question about norm-attaining operators. I show that ℓ_p does not have property B if $1 < p < \infty$. That is, I give an operator into ℓ_p which cannot be approximated in norm by a norm-attaining operator.

In the last chapter, I give an unusual method of constructing the ℓ_p-spaces. By a very natural geometric process, one can build up a symmetric polytope that approximates the sphere to within $\sqrt{2}$. The result generalizes to give polytopes that approximate the unit balls of the other ℓ_p-spaces. Given any function $\lambda: \mathbb{N} \longrightarrow \mathbb{R}$ that satisfies $\lambda(k) = \left\| \sum_1^k x_i \right\|$ for some symmetric basis $x_1, x_2, \ldots$, one can use the same process to give a natural example of such a basis. Some elementary properties of this class of spaces are investigated.

Declaration

No part of this dissertation is derived from any other source, except where it is explicitly stated otherwise.

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration.

W T Gowers

CONTENTS

INTRODUCTION

In 1971 Milman [38] published a new proof of Dvoretzky's theorem that was to be the starting point of a great deal of research in the local theory of Banach spaces. His proof was based on Lévy's isoperimetric inequality on the sphere, and it gave extremely sharp bounds. It was later realized that one could use results about measure concentration in certain discrete metric spaces to obtain other interesting results about Banach spaces. For the main theorems of this dissertation, we shall use such techniques to find large almost symmetric block bases of bases which satisfy various different natural conditions. The two conditions that will principally concern us are, first, that the original basis should be equivalent to the unit vector basis of ℓ_p^n, and, second, that it should satisfy a growth condition of the following form: if the basis $x_1, \ldots, x_n$ is normalized, one asks that $\mathbb{E}\,\|\sum \epsilon_i x_i\|$ should be large, where the expectation is taken over all possible choices of signs $\epsilon_1, \ldots, \epsilon_n$, distributed uniformly.

The first people to show that these conditions allow one to obtain estimates using measure concentration that are significantly better than those obtainable by Ramsey theoretical techniques were Amir and Milman. They published two important papers [3, 4] in 1982 and 1985, giving a series of results in which they obtained unconditional and symmetric block bases of bases with various natural assumptions on them. They used some of these results to obtain local versions of Krivine's theorem and the Maurey-Pisier theorem.

The two results of Amir and Milman that concern us most are the following. They showed that, for any $1 \leqslant p < \infty$ and $\epsilon > 0$, a basis which is equivalent to the unit vector basis of ℓ_p^n has a $(1+\epsilon)$-symmetric block basis of cardinality of order $n^{1/3}$. This was the first step in the proof of their local version of Krivine's theorem. They also considered bases $x_1, \ldots, x_n$ that satisfy the condition

$$\mathsf{E}\left\|\sum_1^n \epsilon_i x_i\right\| \geqslant n^{1/p},$$

where $1 \leqslant p < 2$ and the expectation is taken over all choices of signs $(\epsilon_i)_1^n$. They showed that such a basis has a $(1+\epsilon)$-symmetric block basis of cardinality of order $n^{(2-p)^2/3p^3}$. They used results due to Maurey and Schechtman concerning concentration of measure in various discrete metric spaces. This result was the first step in the proof of their local version of the Maurey-Pisier theorem.

Now, it seemed likely that these results were not best possible. Although the bounds obtained by Amir and Milman in the local versions of Krivine's theorem and the Maurey-Pisier theorem are not significantly affected by improvements in the first step, the problem of determining best possible bounds for the size of a $(1+\epsilon)$-symmetric basis when the original basis satisfies such natural conditions as the ones above is an interesting one in its own right. The main results of this dissertation give bounds in these two cases which are close to being best possible. We show that, if $1 \leqslant p < \infty$, $C > 1$ and $\epsilon > 0$, then any basis which is C-equivalent to the unit vector basis of ℓ_p^n has a $(1+\epsilon)$-symmetric block basis of cardinality at least $\alpha(\epsilon, p, C)n/\log n$, where $\alpha(\epsilon, p, C) > 0$ depends on ϵ, p and C only. Moreover, any basis satisfying the growth condition above has a $(1+\epsilon)$-symmetric block basis of cardinality $\beta(\epsilon)n^{2/p-1}/\log n$, where $\beta(\epsilon) > 0$ depends on ϵ only. We also give examples to show that these estimates are indeed close to the best possible.

We shall now give an overview of the results contained in each chapter of the dissertation. The first chapter introduces the basic definitions and outlines the probabilistic tools which lie at the heart of the later proofs, namely Azuma's inequality, and a deviation inequality due to Hoeffding. The relevance of Azuma's inequality to Banach space theory was discovered by Maurey [35], and Maurey's method was developed by Schechtman [46]. The second chapter consists of results which are of some interest on their own, but are mainly proved as tools for later results. This chapter is quite long and in places technical. The reader may prefer to read the proofs of the results in the chapter only after seeing the results used later.

In the third chapter, we start on our main results, considering bases with large average growth, and proving the result mentioned above concerning such bases. In the fourth chapter, we show how the same method can be adapted to prove the result we have stated above about bases equivalent to the standard basis of an ℓ_p-space.

There follow three chapters of upper bounds. In Chapter 5, we construct a basis with the growth condition above which does not have a $(1+\epsilon)$-symmetric block basis of cardinality greater than $\alpha(\epsilon)n^{2/p-1}(\log n)^{4/3}$. If $1 \leqslant p < 3/2$, then this basis can be chosen to satisfy a lower p-estimate. In Chapter 6 we construct, for $1 < p < \infty$, a basis which is equivalent to the unit vector basis of ℓ_p^n which has no $(1+\epsilon)$-symmetric block basis of cardinality exceeding $\alpha(\epsilon, p)n \log\log n/\log n$. In Chapter 7, we construct a sequence which is C-equivalent to the unit vector basis of ℓ_∞^n with no $(1+\epsilon)$-symmetric block basis of cardinality $n^{\beta(\epsilon,C)}$, where β tends to zero as $\log(1+\epsilon)/\log C$ tends to zero. This shows that the restriction $p < \infty$ is necessary in the main result of Chapter 4. We also obtain an upper bound of a rather different nature, since it concerns arbitrary subspaces and not just those generated by a block basis. Specifically, we show that, for any absolute constants γ and M, and for $n \in \mathbb{N}$ sufficiently large, there exists an n-dimensional normed space X such that no n^γ-dimensional subspace of X has an M-symmetric basis. Rather surprisingly, the proof of this result is an adaptation of the proof of a similar result for block bases. The result itself gives a negative answer to a question of Milman.

The remaining four chapters are different in character from Chapters 4 to

7, although they mostly concern symmetric sequences in one way or another. In Chapter 8 we discuss the natural infinite-dimensional analogue of the main result of Chapter 4. That is, if X is a Banach space that is isomorphic to ℓ_p, then must X contain a $(1+\epsilon)$-symmetric basic sequence? We do not solve this problem. This is not too surprising, as we show that it is in fact equivalent to the distortion problem, which it clearly resembles. We also give a positive answer to a "distortion problem for c_0", showing that, for any $\epsilon > 0$ and any Lipschitz function F on the unit sphere of c_0, there exists an infinite-dimensional subspace of c_0 on whose unit sphere F varies by at most ϵ. We do this by constructing an ultrafilter with special properties.

In Chapter 9 we discuss a problem which is quite well known and is related to one of the results of the fourth chapter. It concerns the relationship between the type constants of a space and the distance of that space from a Hilbert space of the same dimension. In Chapter 10 we provide a simple counterexample to a question about norm-attaining operators. In Chapter 11 we give an unusual isomorphic construction of ℓ_2^n which arose out of a surprising observation about a fairly natural geometric process. This suggested several natural questions, many of which we answer. In particular, the result generalizes to give an isomorphic construction of ℓ_p^n for arbitrary $1 \leqslant p \leqslant \infty$.

Most of the results of Chapters 2-5, Section 6.1 and Chapter 9 will appear in two papers accepted for publication by the Israel Journal of Mathematics [22, 23]. All the results in the dissertation are original except where it is explicitly stated otherwise.

Throughout the dissertation, except at the end of Section 11.5, all our scalars will be assumed to be real. The results, however, carry over without any difficulty to the complex case. We shall often assume that n, the dimension of the space under consideration, is sufficiently large, without actually mentioning this. This is of particular importance in Sections 2.2 and 2.3, and Chapters 6 and 7. Finally, there are several places where it would be hopelessly cumbersome to make sure that all integer quantities are actually given as integers, so in most cases we have not done so. It is easy to modify the proofs so that they are rigorous in this respect. We have not claimed any result that is not true when appropriate integer parts are taken.

I am indebted to various people, but by far the most important is Dr B. Bollobás, my research supervisor. He drew my attention to the papers of Amir and Milman and suggested that some of their bounds might be improved. He then gave me great encouragement and very valuable advice: I could not have asked for a better supervisor. The presentation of the results in this dissertation has benefited enormously from his criticisms. Any inadequacies that remain are, of course, entirely my own responsibility.

I would also like to thank Yoshiharu Kohayakawa, Imre Leader and Jamie Radcliffe for many stimulating conversations over the last three years. The atmosphere they have provided has made my task much easier. In particular, they have increased my general knowledge and sometimes provided me with key references. I am grateful to Yoshiharu Kohayakawa for pointing out various minor errors in earlier drafts of the dissertation.

Finally, I would like to thank my wife, Emily, for her support, and for putting up with distracted and unsociable behaviour when I have been working on problems. I hope that one day I will be able to explain to her what she has helped me to produce.

PART I

PRELIMINARIES

CHAPTER 1

BASIC DEFINITIONS AND STANDARD RESULTS

§1.1 Basic Definitions

The definitions which follow are almost all standard. The most important for our purposes will be that of a symmetric basis. Given a sequence of vectors $x_1, \ldots, x_n$ in a normed space X, we say that it is 1-*symmetric* or simply *symmetric* if, for any permutation $\pi \in S_n$, any choice of signs $\epsilon_1, \ldots, \epsilon_n$ and any sequence of scalars $a_1, \ldots, a_n$,

$$\left\|\sum_1^n \epsilon_i a_i x_{\pi(i)}\right\| = \left\|\sum_1^n a_i x_i\right\|.$$

If under the same conditions we have only

$$\left\|\sum_1^n \epsilon_i a_i x_{\pi(i)}\right\| \leqslant \alpha\left\|\sum_1^n a_i x_i\right\| \tag{1}$$

then we say that the basis $x_1, \ldots, x_n$ is α-*symmetric.* Sometimes, we shall say that a basis is *almost symmetric.* This is not precisely defined, and means that it is $(1+\epsilon)$-symmetric for a fairly small ϵ.

If (1) holds whenever the permutation π is just the identity permutation, then we say that the basis $x_1, \ldots, x_n$ is α-*unconditional.* We shall sometimes say that a basis is *almost unconditional* if it is $(1+\epsilon)$-unconditional for a small value of ϵ.

The next definition is not standard, but it is a natural one for our purposes. During the course of the proofs, we shall often fix a choice of scalars. We shall say that a basis $x_1, \ldots, x_n$ is α-*symmetric at* $a_1, \ldots, a_n$ if (1) holds for the sequence $a_1, \ldots, a_n$. If a is the vector $\sum_1^n a_i x_i$ we shall also say that $x_1, \ldots, x_n$ is α-*symmetric at* a. If the norm $\|.\|$ being considered is not clear from the context, we shall sometimes say that $x_1, \ldots, x_n$ is α-*symmetric at* $a_1, \ldots, a_n$ or a *under* $\|.\|$. We shall also speak loosely of a basis being *almost symmetric at* a vector or sequence.

For the main body of the dissertation, our definition of a block basis is not completely standard, but it follows that of Amir and Milman and is natural in a finite-dimensional context. Given a basis $x_1, \ldots, x_n$, a *block basis* is a sequence $y_1, \ldots, y_m$ where each element y_i is a vector of the form $\sum_{j \in A_i} \lambda_j x_j$ and the sets $A_1, \ldots, A_m$ are disjoint. The more common definition includes the additional condition that if $i_1 < i_2$ and $j_1 \in A_{i_1}$ and $j_2 \in A_{i_2}$ then $j_1 < j_2$. In Chapter 8, it will be convenient to revert to that definition.

The *p-type constant* of a Banach space X is the smallest constant C such that, given any N and any sequence $x_1, \ldots, x_N$ of vectors in X,

$$\mathsf{E}\left\|\sum_1^N \epsilon_i x_i\right\| \leqslant C\left(\sum_1^N \|x_i\|^p\right)^{1/p}, \tag{2}$$

where the expectation is taken over all choices of signs $\epsilon_1, \ldots, \epsilon_N$. We denote it by $T_p(X)$. The *q-cotype constant* is the smallest constant C such that, for any sequence as above,

$$\mathsf{E}\left\|\sum_1^N \epsilon_i x_i\right\| \geqslant C\left(\sum_1^N \|x_i\|^q\right)^{1/q}. \tag{3}$$

We denote this by $C_q(X)$. The smallest constants for which (2) and (3) hold whenever $N \leqslant n$ are denoted by $T_p(X,n)$ and $C_q(X,n)$ respectively. They are known as the *p-type constant* and *q-cotype constant on n vectors.*

Instead of using sums of the form $\sum_1^N \epsilon_i x_i$ in the above definitions, it is often convenient to look at sums of the form $\sum_1^N g_i x_i$ where $g_1, \ldots, g_N$ is a sequence of independent $N(0,1)$ variables. The *Gaussian p-type* and *q-cotype* constants $\alpha_p(X)$ and $\beta_q(X)$ are defined to be the smallest constants such that for any sequence $x_1, \ldots, x_N$ in X, (2) and (3) respectively hold when $\sum_1^N \epsilon_i x_i$ is replaced by $\sum_1^N g_i x_i$. Similarly, one defines the *Gaussian p-type* and *q-cotype constants on n vectors*, $\alpha_p(X,n)$ and $\beta_q(X,n)$.

Given a normed space X with a specified 1-unconditional basis $x_1, \ldots, x_n$ and given a vector $\mathbf{a} = \sum_{i=1}^n a_i x_i \in X$, then the vector $|\mathbf{a}|$ is defined by $\sum_{i=1}^n |a_i| x_i$, and, for $1 \leqslant p < \infty$, the vector $|\mathbf{a}|^p$ is defined by $\sum_{i=1}^n |a_i|^p x_i$. The space X is said to be *p-convex* if, for every choice of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_N$ in X,

$$\left\| \left(\sum_{i=1}^N |\mathbf{a}_i|^p \right)^{1/p} \right\| \leqslant \left(\sum_{i=1}^N \|\mathbf{a}_i\|^p \right)^{1/p} .$$

There is also a natural notion of ∞-convexity, but we shall not use this. If X is p-convex for some $1 \leqslant p < \infty$, then the *p-concavification* of X is the space Y defined by

$$\|\mathbf{a}\|_Y^p = \| \, |\mathbf{a}|^p \|_X \ .$$

The condition that X should be p-convex guarantees that Y is actually a normed space.

Since the norm of a vector in a 1-symmetric space is unchanged if the coordinates of the vector are permuted or changed in sign, it is often useful to look at the so-called decreasing rearrangement of the vector. Given a vector $\mathbf{a} = (a_i)_1^n \in \mathbb{R}^n$, there exists a permutation $\pi \in S_n$ and a sequence of signs $\epsilon_1, \ldots, \epsilon_n$ such that $\epsilon_1 a_{\pi(1)} \geqslant \ldots \geqslant \epsilon_n a_{\pi(n)} \geqslant 0$. Given such a permutation and sequence of signs, a_i^* is defined to be $\epsilon_i a_{\pi(i)}$. This is easily seen to be well-defined. The vector $\mathbf{a}^* = (a_i^*)_1^n$ is the *decreasing rearrangement* of $\mathbf{a}$. Given $t > 0$ and a function from the closed interval $[0, t]$ to $\mathbb{R}$, the *decreasing rearrangement* f^* of f is defined by

$$f^*(x) = \sup\Big\{ s > 0 : \big|\{ y : f(y) \geqslant s \}\big| \geqslant x \Big\} .$$

We shall use some standard notation for various sets and set systems. In particular, $[n]$ will stand for the set $\{1, 2, \ldots, n\}$, and $[n]^{(r)}$ will stand for the collection of subsets of $[n]$ of cardinality r. This should not be confused with $[n]^r$, which stands for the set of r-tuples $(i_1, \ldots, i_r)$ of elements of $[n]$.

We shall frequently talk about the *standard basis* or the *unit vector basis* of certain normed spaces. The meaning of this is obvious. For example, given a normed space $X = (\mathbb{R}^n, \|.\|)$, the standard basis of X is the basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$, where $\mathbf{e}_i$ is the vector with 1 in the i^{th} coordinate and zero everywhere else. Of course, the identification of X with $\mathbb{R}^n$ must be explicit. If $\mathbf{a} = \sum_{i=1}^n a_i \mathbf{e}_i \in \mathbb{R}^n$, and $A \subset [n]$ is a subset of $[n]$, then *the restriction of* $\mathbf{a}$ *to* A, written $\mathbf{a}\,\big|_A$, is the vector $\sum_{i \in A} a_i \mathbf{e}_i$.

§1.2 Standard Results

In this short section we shall state two deviation inequalities from probability theory, a simple estimate for the size of nets in the unit ball of a finite-dimensional normed space and some bounds for the hypergeometric distribution. Apart from these results, which all have easy proofs, this dissertation is more or less self-contained.

Suppose we choose a sequence $a_1, \ldots, a_n$ of real numbers and look at the sum $\sum_1^n \epsilon_i a_i$, where $\epsilon_1, \ldots, \epsilon_n$ is a randomly chosen sequence of signs. Khintchine's inequality tells us that the expected value of the modulus of this sum is about $\left(\sum_1^n a_i^2\right)^{1/2}$. An inequality of Azuma [5] states that the probability of the sum being substantially larger than this in modulus is extremely small. The proof of this fact is very quick, but the consequences are far-reaching. It will be our main tool for proving measure-concentration results.

Theorem 1·1. *Let $(d_i)_1^n$ be a martingale difference sequence, and let $(c_i)_1^n$ be non-negative real numbers such that $\|d_i\|_\infty \leqslant c_i$ for every $1 \leqslant i \leqslant n$. Then, for any $\lambda > 0$,*

$$\mathbb{P}\left[\left|\sum_{i=1}^n d_i\right| \geqslant \lambda\right] \leqslant \exp\left(-\lambda^2 \Big/ 2\sum_{i=1}^n c_i^2\right) . \qquad \square$$

In fact a similar inequality to Azuma's was proved in 1963 by Hoeffding [26], which can sometimes be used to prove stronger results. A "martingale version" of this result was recently proved by McDiarmid [37]. We shall use the following result of Hoeffding.

Theorem 1·2. (Hoeffding 1963) *Let $X_1, \ldots, X_n$ be a sequence of independent random variables satisfying $0 \leqslant X_i \leqslant 1$ for each i, and set $\overline{X} = n^{-1} \sum_1^n X_i$, $p = \mathsf{E}\overline{X}$ and $q = 1 - p$. Then if $0 \leqslant t < q$,*

$$\mathsf{P}\left[\overline{X} - p \geqslant t\right] \leqslant \left(\left(\frac{p}{p+t}\right)^{p+t}\left(\frac{q}{q-t}\right)^{q-t}\right)^n . \qquad \square$$

The proof of this is very similar in style to the proof of Azuma's inequality. We shall use this theorem in a situation where p is much smaller than 1, where Azuma's inequality gives only weak results, if any.

A δ-net of a metric space X is a subset $\Delta \subset X$ with the property that, for every $x \in X$, there exists $x' \in \Delta$ such that $d(x, x') \leqslant \delta$. We turn now to a very standard estimate for the size of a δ-net of the unit ball of an arbitrary n-dimensional normed space. An easy proof may be found in [18] or [39].

Lemma 1·3. *Let X be an n-dimensional normed space and let $0 < \delta < 1$. Then the unit ball of X contains a δ-net of cardinality $(1 + 2/\delta)^n$.* $\square$

We end this chapter with useful estimates for the hypergeometric distribution which can be found in [10].

Lemma 1·4. *Let $0 < n < N$ and $0 < R < N$ be integers, and, for any integer $0 < k < \min\{n, R\}$, let q_k be defined by*

$$q_k = \binom{R}{k}\binom{N-r}{n-k}\Big/\binom{N}{n} = \binom{n}{k}\binom{N-n}{R-k}\Big/\binom{N}{R} .$$

Then, setting $p = R/N$ and $q = 1 - p$, we have

$$\binom{n}{k}\left(p - \frac{k}{N}\right)^k \left(q - \frac{n-k}{N}\right)^{n-k} < q_k < \binom{n}{k} p^k q^{n-k} \left(1 - \frac{k}{N}\right)^{-(n-k)} .$$

CHAPTER 2

TECHNICAL RESULTS

The results of this chapter are proved mainly because they will be needed later in the dissertation. Some of them are quite interesting independently of their later use, especially those in Section 2·1 and some of Section 2·5. The results of Sections 2·2 and 2·3 give sort of converses to the main result of Section 2·1, but the form in which they are stated is very much dictated by their use later (in this case in Chapter 6).

§2.1 The Size of Nets in the Decreasing Part of the Unit Ball

Lemma 1.3 was an estimate for the size of a δ-net of the unit ball of an n-dimensional normed space. The main result of this section is not strictly about nets, but it plays the role of estimates of the size of nets in similar proofs, and we shall use Lemma 1.3 in its proof. The importance of a net in Banach space theory is principally that if a norm is known on a sufficiently fine net of the unit ball of a space, then it is essentially known everywhere. Clearly, therefore, if the space has a 1-symmetric basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$, all one needs to know is the norm on a sufficiently fine net of the set of vectors $\sum_1^n a_i \mathbf{e}_i$ in the unit ball for which $a_1 \geqslant \ldots \geqslant a_n \geqslant 0$. We shall call this the *decreasing part* of the unit ball. It turns out that it is more convenient to control the norm using a set that is not actually a net. Our first few lemmas will show that there is a set of vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ of fairly small cardinality with the property that any norm which is almost symmetric at each $\mathbf{a}_i$ must be almost symmetric on the whole space. We shall see that the norm on an n-dimensional space with a symmetric basis is almost determined by its restriction to a subspace of dimension roughly $\log n$.

In the other direction, we shall need to find a largish collection of vectors in the decreasing part of the unit ball of ℓ_p^n with the property that the norm of any one of the vectors is *not* controlled by the norms of the others. This we shall do in Sections 2·2 and 2·3. One can get some sort of estimate using a separated set and well known estimates for the modulus of convexity of ℓ_p^n. Our approach will be more direct and will give a better result.

The main result of this section is the following.

Proposition 2·1. *Let $\delta > 0$, let $(\mathbb{R}^m, \|.\|)$ be a normed space and set $N = m^{\delta^{-1}\log(3\delta^{-1})}$. There exist N vectors $\mathbf{a}_1, \ldots, \mathbf{a}_N$ such that if $\|.\|$ is $(1+\delta)$-symmetric at $\mathbf{a}_i$ for every i, then the standard basis of $\mathbb{R}^m$ is $(1+\delta)(1-6\delta)^{-1}$-symmetric.*

We shall prove Proposition 2·1 by splitting it into a number of further simple lemmas, but first we need some notation. Let Ψ be the set $\{-1,1\}^m \times S_m$, and, for any $\mathbf{a} = \sum_{j=1}^m a_j \mathbf{e}_j \in \mathbb{R}^m$ and any $(\eta, \sigma) = (\eta_1, \ldots, \eta_m, \sigma) \in \Psi$, let $\mathbf{a}_{\eta,\sigma}$ stand for the vector $\sum_{j=1}^m \eta_j a_j \mathbf{e}_{\sigma(j)}$. Let us also define two orders on $\mathbb{R}^m$ as follows. We shall say $\mathbf{a} \leqslant_1 \mathbf{b}$ if $\mathbf{a}$ can be written as a convex combination of vectors of the form $\mathbf{b}_{\eta,\sigma}$, and $\mathbf{a} \leqslant_2 \mathbf{b}$ if $\sum_{i=1}^k a_i^* \leqslant \sum_{i=1}^k b_i^*$ for all $1 \leqslant k \leqslant n$. It is easy to check that both these orders are transitive, and indeed, that they are partial orders on the set of positive decreasing vectors in $\mathbb{R}^n$. The order $\leqslant_1$ can be defined a little more naturally, but with this definition it is immediate that if $\|.\|$ is a 1-symmetric norm on $\mathbb{R}^m$ then $\mathbf{a} \leqslant_1 \mathbf{b}$ implies that $\|\mathbf{a}\| \leqslant \|\mathbf{b}\|$. This is why we shall use the order. The following easy fact, which was originally noted by Rado [44], is well known.

Lemma 2·2. *Let $\mathbf{a} = \sum_1^m a_i \mathbf{e}_i$ and $\mathbf{b} = \sum_1^m b_i \mathbf{e}_i$ be vectors in $\mathbb{R}^m$. Then $\mathbf{a} \leqslant_1 \mathbf{b}$ if and only if $\mathbf{a} \leqslant_2 \mathbf{b}$.* □

For given $\delta > 0$, let us define a subspace $U = U(\delta)$ of $\mathbb{R}^m$, as follows. Let $r = \lfloor \log_{1+\delta} m \rfloor$, and for $1 \leqslant i \leqslant r$ set $k_i = \lfloor (1+\delta)^i \rfloor$ and $\mathbf{u}_i = \sum_{j=1}^{k_i} \mathbf{e}_j$. Then our subspace U will be that generated by $\mathbf{u}_1, \ldots, \mathbf{u}_r$. For technical reasons, we shall define k_{r+1} to be $n+1$. Note that for $1 \leqslant i \leqslant r$, we have $k_{i+1}-1 \leqslant (1+\delta)k_i$. In the next lemma we show that any 1-symmetric norm is determined to within a factor $1+\delta$ by its restriction to U. We shall eventually need to loosen the conditions on the norm, and obtain a slightly weaker conclusion.

For the next three lemmas, let $\delta > 0$ and let $U = U(\delta)$ be as defined above.

Lemma 2·3. *Let $\|.\|$ be a norm which is defined on U. Then there exists a function $f : \mathbb{R}^m \to \mathbb{R}_+$ such that any 1-symmetric norm $|\!|\!|.|\!|\!|$ with the property that $|\!|\!|x|\!|\!| = \|x\|$ for every $x \in U$ satisfies also*

$$f(x) \leqslant |\!|\!|x|\!|\!| \leqslant (1+\delta)f(x)$$

for every $x \in \mathbb{R}^n$.

Proof. Let $\mathbf{a} = \sum_1^m a_i \mathbf{e}_i$ and let $a_1 \geqslant \ldots \geqslant a_m \geqslant 0$. Let us define two vectors $\mathbf{a}'$ and $\mathbf{a}''$ in U by

$$\mathbf{a}' = \sum_{i=1}^{r}(a_{k_i} - a_{k_{i+1}}) \sum_{j=1}^{k_{i+1}-1} \mathbf{e}_j = \sum_{i=1}^{r} a_{k_i} \sum_{j=k_i}^{k_{i+1}-1} \mathbf{e}_j$$

and

$$\mathbf{a}'' = \sum_{i=1}^{r}(a_{k_i} - a_{k_{i+1}}) \sum_{j=1}^{k_i} \mathbf{e}_j = \sum_{i=1}^{r} a_{k_i} \sum_{j=k_{i-1}+1}^{k_i} \mathbf{e}_j \,.$$

Our function f will be defined by $f(\mathbf{a}) \equiv \|\mathbf{a}''\|$. Note that $\mathbf{a}''$ is dominated pointwise by $\mathbf{a}$, which is itself dominated pointwise by $\mathbf{a}'$. Thus $|\!|\!|\mathbf{a}''|\!|\!| \leqslant |\!|\!|\mathbf{a}|\!|\!| \leqslant |\!|\!|\mathbf{a}'|\!|\!|$. We shall show that $\mathbf{a}' \leqslant_2 (1+\delta)\mathbf{a}''$ and hence (by Lemma 2·2) that $|\!|\!|\mathbf{a}'|\!|\!| \leqslant (1+\delta)|\!|\!|\mathbf{a}''|\!|\!|$, which will complete the proof. Let us write $(a'_j)_1^m$ and $(a''_j)_1^m$ for the coordinates of $\mathbf{a}'$ and $\mathbf{a}''$ respectively. Then for any $1 \leqslant s \leqslant m$

$$\sum_{j=1}^{s} a'_j = \sum_{i=1}^{r}(a_{k_i} - a_{k_{i+1}}) \min\{s, k_{i+1}-1\}$$

and

$$\sum_{j=1}^{s} a''_j = \sum_{i=1}^{r} (a_{k_i} - a_{k_{i+1}}) \min\{s, k_i\} .$$

But for each i, as remarked above, $k_{i+1} - 1 \leqslant (1+\delta)k_i$, so clearly $\min\{s, k_{i+1} - 1\} \leqslant (1+\delta)\min\{s, k_i\}$. Hence $\mathbf{a}' \leqslant_2 (1+\delta)\mathbf{a}''$ as stated. □

Lemma 2·4. *Let $\delta' > 0$ and let $\|.\|$ be any norm on $\mathbb{R}^m$ which is $(1-\delta')^{-1}$-symmetric at every $\mathbf{a} \in U(\delta)$. Then $\|.\|$ is $(1+\delta)(1-2\delta')^{-1}$-symmetric on $\mathbb{R}^m$.*

Proof. Define a norm $|\!|\!|.|\!|\!|$ on $\mathbb{R}^m$ by $|\!|\!|\mathbf{a}|\!|\!| = \max\{\|\mathbf{a}_{\eta,\sigma}\| : (\eta,\sigma) \in \Psi\}$. Then, if $\mathbf{a} \in U$, we have $\|\mathbf{a}\| \leqslant |\!|\!|\mathbf{a}|\!|\!| \leqslant (1+\delta)\|\mathbf{a}\|$ by assumption. Now, given any $\mathbf{a} \in \mathbb{R}^m$, define $\mathbf{a}''$ as in the proof of Lemma 2·3. We know that $\|\mathbf{a}\| \leqslant |\!|\!|\mathbf{a}|\!|\!|$. In the other direction, since $\mathbf{a}''$ is dominated pointwise by $\mathbf{a}$, $2\mathbf{a}'' - \mathbf{a}$ is dominated pointwise by $\mathbf{a}''$, so $|\!|\!|2\mathbf{a}'' - \mathbf{a}|\!|\!| \leqslant |\!|\!|\mathbf{a}''|\!|\!|$. But, since $\mathbf{a}'' \in U$, we have

$$\begin{aligned} \|\mathbf{a}\| &\geqslant 2\|\mathbf{a}''\| - \|2\mathbf{a}'' - \mathbf{a}\| \\ &\geqslant 2(1-\delta')|\!|\!|\mathbf{a}''|\!|\!| - |\!|\!|\mathbf{a}''|\!|\!| \\ &= (1-2\delta')|\!|\!|\mathbf{a}''|\!|\!| \geqslant (1-2\delta')(1+\delta)^{-1}|\!|\!|\mathbf{a}|\!|\!| \end{aligned}$$

where the last inequality follows from the proof of Lemma 2·3.

Hence $\|.\|$ is $(1+\delta)(1-2\delta')^{-1}$-equivalent to the 1-symmetric norm $|\!|\!|.|\!|\!|$, which proves the lemma. □

Lemma 2·5. *Let $\|.\|$ be a norm on $\mathbb{R}^m$, and define a 1-symmetric norm $|\!|\!|.|\!|\!|$ on $\mathbb{R}^m$ by $|\!|\!|\mathbf{a}|\!|\!| = \max\{\|\mathbf{a}_{\eta,\sigma}\| : (\eta,\sigma) \in \Psi\}$. Suppose that the set of vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ forms a δ-net in $|\!|\!|.|\!|\!|$ of the $|\!|\!|.|\!|\!|$-unit ball of U and that $\|.\|$ is $(1+\delta)$-symmetric at every $\mathbf{a}_i$. Then $\|.\|$ is $(1+\delta)(1-6\delta)^{-1}$-symmetric on $\mathbb{R}^m$.*

Proof. By hypothesis, given any $1 \leqslant i \leqslant N$ and any $(\eta,\sigma) \in \Psi$, we have $\|(\mathbf{a}_i)_{\eta,\sigma}\| \geqslant (1+\delta)^{-1}|\!|\!|\mathbf{a}_i|\!|\!|$. Let us pick $\mathbf{a} \in U$ with $|\!|\!|\mathbf{a}|\!|\!| = 1$. We shall show that $\|.\|$ is $(1-3\delta)^{-1}$-symmetric at $\mathbf{a}$. So pick any $(\eta,\sigma) \in \Psi$ and pick i such that

$|||\mathbf{a} - \mathbf{a}_i||| \leqslant \delta$. Write $\mathbf{b} = \mathbf{a}_{\eta,\sigma}$ and $\mathbf{b}' = (\mathbf{a}_i)_{\eta,\sigma}$. Then we have

$$\begin{aligned} |||\mathbf{b}||| - \|\mathbf{b}\| &\leqslant |||\mathbf{b} - \mathbf{b}'||| + \big|\,|||\mathbf{b}'||| - \|\mathbf{b}'\|\,\big| + \|\mathbf{b}' - \mathbf{b}\| \\ &\leqslant \delta + 1 - (1+\delta)^{-1} + \delta < 3\delta \end{aligned}$$

But $|||\mathbf{b}||| = |||\mathbf{a}||| = 1$, so the norm $\|.\|$ is indeed $(1-3\delta)^{-1}$-symmetric at every vector in U. But then, by Lemma 2·4, it is $(1+\delta)(1-6\delta)^{-1}$-symmetric on $\mathbb{R}^m$, as stated. □

Proof of Proposition 2·1. The dimension of $U(\delta)$ is at most $\log_{1+\delta} m$, so, by Lemma 1.3, there is a δ-net of the $|||.|||$-unit ball of U of cardinality at most $(1 + 2/\delta)^{\log_{1+\delta} m}$. When $\delta \leqslant 1/11$ one can easily check that this is at most $m^{\delta^{-1}\log(3\delta^{-1})}$. Proposition 2·1 now follows immediately from Lemma 2·5. □

Although it will not be strictly necessary for our purposes, it is of some interest to estimate the size of a δ-net of the decreasing part of the unit ball of an arbitrary n-dimensional space with a 1-symmetric basis. We have not done this in general, but when the space is ℓ_p^n, we obtain an estimate of $n^{(c/\delta)\log(c'/\delta)}$ for some absolute constants c and c'. In other words, the size of a net can be about the same as the size of the set obtained in Proposition 2.1. It seems very likely that the same is true for any space with a 1-symmetric basis.

Proposition 2·6. *Let $1 \leqslant p \leqslant \infty$, $0 < \delta < 1$ and let $K \subset \ell_p^n$ be the set*

$$\{\mathbf{a} \in \ell_p^n : \|\mathbf{a}\|_p \leqslant 1,\ a_1 \geqslant \ldots \geqslant a_n \geqslant 0\}.$$

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

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

Proof. Let $\theta = \delta/3$ and let $\mathbf{a} = (a_i)_1^n \in K$. If $\mathbf{a}' = (a_i')_1^n \in \ell_p^n$ is any vector such that $a_i \leqslant a_i' \leqslant (1+\theta)a_i$ for all $1 \leqslant i \leqslant n$, then $\|\mathbf{a} - \mathbf{a}'\|_p \leqslant \theta(\sum_1^n a_i^p)^{1/p} \leqslant \theta$. So, given $\mathbf{a}$, let us define $\mathbf{a}'$ to be the vector with

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

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

Note that $k \geqslant \log_{1+\theta}(\theta^{-1}n^{1/p})$, so $(1+\theta)^{-k} \leqslant \theta n^{-1/p}$. It follows easily that $\|\mathbf{a}'' - \mathbf{a}'\|_p \leqslant \theta$, and therefore that $\|\mathbf{a}'' - \mathbf{a}\|_p \leqslant 2\theta$. If, given any vector $\mathbf{a}$, we can find a vector $\mathbf{b}$ in Δ such that $\|\mathbf{b} - \mathbf{a}''\| \leqslant \theta$, then $\|\mathbf{b} - \mathbf{a}\| \leqslant 3\theta = \delta$, so then Δ will be a δ-net. In other words, it is enough to approximate to within θ vectors of the form $\mathbf{a} = \sum_1^k \alpha_i \mathbf{u}_i$, where $\mathbf{u}_i = \chi_{U_i}$, for some sequence of possibly empty sets $U_1, \ldots, U_k$ 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 $\mathbf{a} = \sum_1^k \alpha_i \mathbf{u}_i$ and $\mathbf{a}' = \sum_1^k \alpha_i \mathbf{u}'_i$, where $(\mathbf{u}_i)_1^k$ and $(\mathbf{u}'_i)_1^k$ are of the above form. Writing $\mathbf{v}_i = \sum_{j=1}^i \mathbf{u}_i$, $\mathbf{v}' = \sum_{j=1}^i \mathbf{u}'_i$, we have $\mathbf{a} = \sum_{i=1}^k (\alpha_i - \alpha_{i+1})\mathbf{v}_i$, $\mathbf{a}' = \sum_{i=1}^k (\alpha_i - \alpha_{i+1})\mathbf{v}'_i$ and $\mathbf{a} - \mathbf{a}' = \sum_{i=1}^k (\alpha_i - \alpha_{i+1})(\mathbf{v}_i - \mathbf{v}'_i)$.

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

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

Thus

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

But $\|\mathbf{a}\|_p^p = \sum_{j=1}^k \alpha_j^p |\mathrm{supp}(\mathbf{u}_j)| = \sum_{j=1}^k (\alpha_j^p - \alpha_{j+1}^p)|\mathrm{supp}(\mathbf{v}_j)| \leqslant 1 + 2\theta$ and $\|\mathbf{a}'\|_p^p = \sum_{j=1}^k (\alpha_j^p - \alpha_{j+1}^p)|\mathrm{supp}(\mathbf{v}'_j)| \leqslant 1 + 2\theta$, so N is at most the size of a θ^p-net of $(1+2\theta)B(\ell_1^k)$, i.e.

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

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

$$N \leqslant n^{(2/\log(1+\theta))\log(5/\theta)} . \qquad \square$$

§2.2 Well-Separated Classes of Vectors in the Unit Ball of ℓ_p^n $(1 < p < \infty)$

In the last section we estimated from above the smallest possible size of a set of vectors which controls an almost symmetric norm. Knowing an almost symmetric norm everywhere on such a set tells us, to considerable accuracy, the norm everywhere in the space. In Chapter 6, we shall construct, for $1 < p < \infty$, an equivalent norm on ℓ_p^n with the property that no block basis of cardinality greater than $\alpha(p, \epsilon)n \log\log n / \log n$ is $(1 + \epsilon)$-symmetric, for a constant α which depends only on ϵ and p. We shall do this by taking a large collection of subsets of the unit sphere of ℓ_p^n, each of which is invariant under the symmetries of ℓ_p^n induced by permutations and changes of sign of the standard basis, and then defining a norm randomly in such a way that it is very unlikely to be symmetric at any given vector in any one of the subsets.

It is important in our proof that our subsets should have two other properties. The first is that any sufficiently large block basis of the standard basis should generate at least one vector in each subset, and the second is that we should be able to define our random norm *independently* on each subset. This second property is the most important. We demand that the subsets should not interfere with each other in the following sense. If we define a norm on $\mathbb{R}^n$ by taking its unit ball to be the convex hull of all the subsets but one, then the norm of a vector in the subset that we leave out should be greater than $1 + \epsilon$ for some ϵ which does not depend on n. Actually, for technical reasons, we will ask for slightly more than this because of our particular method of proof in Chapter 6, but this is the basic idea. This last property is what enables us to define a norm randomly and independently on each subset. We shall then obtain an estimate for the probability of a block basis being almost symmetric at any particular vector it generates in one of the subsets. Since these probabilities are independent, we will be able to multiply them all together. If we have found a large number of classes, then the resulting probability is very small indeed and enables us to show that, with positive probability, no block basis of cardinality $\alpha(\epsilon,p)n\log\log n/\log n$ is $(1+\epsilon)$-symmetric. We shall speak loosely of the subsets being "well separated". In a way, finding as many well-separated vectors as possible is the opposite of finding as few vectors as possible which control the norm.

The numbers used in the construction are those that we shall need when we apply the result. As usual, q is the conjugate index for p. In our application, the order of magnitude of h will be $\log n/\log\log n$.

So, let n and h be given, let $0<\epsilon<1/3$, let $l=n^{1/2}$, $k=h^{1+p+q}\epsilon^{-p/q}$, $\delta=h^{-q}\epsilon^{1/q}=(h/k)^{1/p}$ and let $t=\log(l/2h)/\log k\geqslant\log n/2(1+p+q)\log\log n$, so that $h\sum_1^t k^i\leqslant l$.

Then, for $1\leqslant i\leqslant t$, let A_i be the set of norm-1 vectors in ℓ_p^n supported on at most hk^i points, whose coordinates are bounded above in modulus 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 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 $\{|\mathbf{a}|^{p-1}\mathrm{sign}\,\mathbf{a}\colon \mathbf{a}\in A_i\}$.

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

Now let $\mathcal{B}$ be a subset of $[t]^{(t/2)}$ such that whenever B, C are in $\mathcal{B}$ and $B\neq C$, then $|B\cap C|\leqslant t/3$, and let $\mathcal{B}$ have cardinality $N=(23/20)^t$. (We shall show that such a collection of sets exists in Lemma 2.13). For any $B\in\mathcal{B}$ we define classes

A_B, F_B and G_B as follows.

$$A_B = \left\{ \bigoplus_{i \in B} \mathbf{a}_i : \mathbf{a}_i \in A_i \ \forall i \in B \right\}$$

$$F_B = \left\{ \bigoplus_{i \in B} f_i : f_i \in F_i \ \forall i \in B \right\}$$

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

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

Note that if $\mathbf{a} \in A_B$, then there exists $f \in F_B$ such that $f(\mathbf{a}) = (t/2)^{1/q} \|\mathbf{a}\|_p$. Indeed, let $\mathbf{a} = \bigoplus_{i \in B} \mathbf{a}_i$, where $\mathbf{a}_i \in A_i$ for each $i \in B$, let $f_i = |\mathbf{a}_i| \mathrm{sign}(\mathbf{a}_i)$ for each such i and let $f = \bigoplus_{i \in B} f_i$. Then $f(\mathbf{a}) = t/2$ while $\|\mathbf{a}\|_p = (t/2)^{1/p}$. The next lemma shows that the classes $\{A_B : B \in \mathcal{B}\}$ are "well-separated".

Lemma 2·7. *Suppose B and C are distinct elements of $\mathcal{B}$, and suppose $f \in F_B$, $g \in G_B$ and $\mathbf{a} \in A_C$. Then*

$$f(\mathbf{a}) + g(\mathbf{a}) \leqslant (t/2)^{1/q} \|\mathbf{a}\|_p .$$

Proof. It is simple to show that if x and y are two vectors in $\mathbb{R}^n$, then $\langle x_{\epsilon,\pi}, y_{\epsilon',\pi'} \rangle$ is maximized when $x_{\epsilon,\pi}$ and $y_{\epsilon',\pi'}$ are both non-negative decreasing vectors. We shall therefore assume this of f, g and $\mathbf{a}$. Let us write $f = \sum_{i \in B} f_i$ with $f_i \in F_i$ for each i and the f_i being disjointly supported, and similarly write $g = \sum_{i \in B} g_i$ and $\mathbf{a} = \sum_{j \in C} \mathbf{a}_j$. 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 $f_i \left(\sum_{j \in C} \mathbf{a}_j \right)$ 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

$$f_i\Big(\sum_{j\in C} \mathbf{a}_j\Big) \leqslant f_i'\Big(\sum_{j\in C} \mathbf{a}_j\Big)$$
$$= f_i'\Big(\sum_{j\in C, j<i} \mathbf{a}_j\Big) + f_i'(\mathbf{a}_i)$$
$$\leqslant \|f_i'\|_\infty \Big\|\sum_{j\in C,\ j<i} \mathbf{a}_j\Big\|_1 + 1$$

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

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

and hence

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

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

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

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

We have already estimated the first term. Also

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

so

$$f_i\Big(\sum_{j\in C} \mathbf{a}_j\Big) \leqslant 2h^{1/q}k^{-1/q} + \delta\,. \tag{2}$$

It follows from (1) and (2) that

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

If $B \neq C$ then $|B\cap C| \leqslant t/3$, 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

$$g_i\Big(\sum_{j\in C} \mathbf{a}_j\Big) \leqslant \epsilon + 2h^{2/q}\epsilon^{-p/q}k^{-1/q} ,$$

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

$$g_i\Big(\sum_{j\in C} \mathbf{a}_j\Big) \leqslant 2h^{2/q}\epsilon^{-p/q}k^{-1/q} + \epsilon k^{-1/p} .$$

It follows that

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

If $B \neq C$ then this gives

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

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

$$\begin{aligned}(2/t)^{1/q}\|\mathbf{a}\|_p^{-1}(f+g)(\mathbf{a}) &= (2/t)(f+g)(\mathbf{a}) \\ &\leqslant 2/3 + \delta + 2h^{2/q}\epsilon^{-p/q}k^{-1/q} + 2\epsilon/3 + \epsilon k^{-1/p}/3.\end{aligned}$$

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

Before finishing this section, we remark that the functionals in the sets $(G_B : B \in \mathcal{B})$ are needed only for technical reasons. It would not be necessary at all if we were only interested in the dependence of our eventual estimate in Chapter 6 on n. By using these extra functionals, we will, it turns out, be able to construct a basis equivalent to the standard basis of ℓ_p^n whose largest $(1+\epsilon)$-symmetric block basis has cardinality at most $\epsilon^p c_1(p)c_2(n) = c_3(\epsilon, p)n \log\log n/\log n$, where c_1 depends only on p, c_2 only on n and c_3 only on p and ϵ. The correct dependence on ϵ is probably more like ϵ^{2p}. In the next section, when $p = 1$ is an absolute constant, we shall prove a similar result, but it will be slightly simpler since there will be nothing to gain from an extra class of functionals.

§2.3 Well-Separated Classes of Vectors in the Unit Ball of an Isomorph of ℓ_1^n

In the case when $p = 1$, Lemma 2·7 does not give any information, and a careful examination of the proof shows that even if $p > 1$, while remaining close enough to 1 for ℓ_p^n to be C-equivalent to ℓ_1^n for some fixed C, we do not obtain N classes of vectors for some N which tends to infinity with n. However, using a different construction, it is possible to find N "well-separated" classes of vectors in the unit ball of an n-dimensional normed space which is C-equivalent to ℓ_1^n, with $N = (\log n/\log\log n)^\alpha$, where $\alpha = \alpha(C, \epsilon) > 0$. It is to be expected that we should obtain fewer classes in this case, because the unit ball of ℓ_1^n cannot be renormed to be uniformly convex. We can obtain a very small amount of convexity by renorming, however: it will be just enough to enable us to construct a basis which is C-equivalent to the unit vector basis of ℓ_1^n, such that no block basis of cardinality exceeding $\beta(\epsilon, C)n/\log\log n$ is $(1+\epsilon)$-symmetric. As before, we shall define some classes of vectors and functionals, and then prove a lemma about them. Let $\epsilon > 0$, $C > 1$ and $h \in \mathbb{N}$ be given. In our application, h will be about $\log\log n$.

First let us define four more parameters. Let $r = -\log C/\log(1-4\epsilon)$, let $k = h^2$, let $\lambda = \lfloor(\log_k n)/2r\rfloor$ and let $N = (\lambda/2)^{r/2}$. Next, for $1 \leqslant i \leqslant r$, let $\beta_i = C(1-4\epsilon)^i$. Thus $C > \beta_1 > \ldots > \beta_r = 1$, and $\beta_{i+1}/\beta_i \leqslant 1 - 4\epsilon$ for each i.

Now, for any $1 \leqslant i \leqslant r$ and $1 \leqslant j \leqslant \lambda/2$, let $A_{i,j}$ be the class of vectors $\mathbf{a} \in \mathbb{R}^n$ satisfying $\|\mathbf{a}\|_1 = \beta_i^{-1}$, $\|\mathbf{a}\|_\infty \leqslant \beta_i^{-1}k^{-(i-1)\lambda-j}$ and $|\text{supp}(\mathbf{a})| \leqslant hk^{(i-1)\lambda+j}$. Note that it follows from these conditions that if $\mathbf{a} \in A_{i,j}$, then $|\text{supp}(\mathbf{a})| \geqslant k^{(i-1)\lambda+j}$.

We shall also define a class of functionals to go with $A_{i,j}$. It turns out in this case, as we mentioned at the end of the last section, that only one class of functionals is necessary. Let $F_{i,j}$ be the set of functionals f on $\mathbb{R}^n$ such that

$|\mathrm{supp}(f)| = hk^{(i-1)\lambda+j}$ and each non-zero coordinate of f is $\pm\beta_i$.

We shall now define, for any $\mathbf{s} = (s_1, \ldots, s_r) \in [\lambda/2]^r$, two classes $A_\mathbf{s}$ and $F_\mathbf{s}$ as follows.

$$A_\mathbf{s} = \left\{ \bigoplus_{i=1}^{r} \mathbf{a}_i \ : \ \mathbf{a}_i \in A_{i,s_i} \ i = 1, \ldots, r \right\}$$

and

$$F_\mathbf{s} = \left\{ \bigoplus_{i=1}^{r} f_i \ : \ f_i \in F_{i,s_i} \ i = 1, \ldots, r \right\} .$$

The next lemma is the analogue of Lemma 2·7 for the classes $A_\mathbf{s}$ and $F_\mathbf{s}$.

Lemma 2·8. *Let $\mathbf{s}, \mathbf{s}' \in [\lambda/2]^r$ and suppose that there exists a subset $K \subset [r]$ of cardinality at least $r/3$ such that $s_i < s'_i$ for every $i \in K$. Then, given any $f \in F_{\mathbf{s}'}$ and $\mathbf{a} \in A_\mathbf{s}$, we have*

$$f(\mathbf{a}) \leqslant r(1-\epsilon).$$

Proof. Let i be such that $s'_i < s_i$ and let $f = \bigoplus_{j=1}^{r} f_j$ with $f_j \in F_{j,s'_j}$ for each $1 \leqslant j \leqslant r$. We shall first estimate $\left(\bigoplus_{j=1}^{r} f_j\right)(\mathbf{a}_i)$. Let us temporarily write X and Y for supp $\left(\bigoplus_{j=1}^{i-1} f_j\right)$ and $\mathrm{supp}(f_i)$ respectively. We then have

$$\left(\bigoplus_{j=1}^{i-1} f_j\right)(\mathbf{a}_i) \leqslant \left\| \bigoplus_{j=1}^{i-1} f_j \right\|_\infty \left\| \mathbf{a}_i \left|_X \right. \right\|_1 \leqslant C\beta_i^{-1} hk^{-\lambda/2}.$$

We also have

$$f_i(\mathbf{a}_i) \leqslant \|f_i\|_\infty \left\| \mathbf{a}_i \left|_Y \right. \right\|_1 \leqslant hk^{-1}$$

and

$$\left(\bigoplus_{j=i+1}^{r} f_j\right)(\mathbf{a}_i) \leqslant \left\| \bigoplus_{j=i+1}^{r} f_j \right\|_\infty \|\mathbf{a}_i\|_1 \leqslant \beta_{i+1}\beta_i^{-1} = 1 - 4\epsilon.$$

Now let i be such that $s'_i \geqslant s_i$. Then

$$\left(\bigoplus_{j=1}^{i-1} f_j\right)(\mathbf{a}_i) \leqslant C\beta_i^{-1} hk^{-\lambda/2}$$

as before. We also have

$$\Big(\bigoplus_{j=i}^{r} f_j\Big)(\mathbf{a}_i) \leqslant \Big\|\bigoplus_{j=i}^{r} f_j\Big\|_\infty \|\mathbf{a}_i\|_1 = \beta_i\beta_i^{-1} = 1.$$

Since $s_i' < s_i$ for at least $r/3$ values of i, we have

$$\Big(\bigoplus_{j=1}^{r} f_j\Big)\Big(\bigoplus_{i=1}^{r} \mathbf{a}_i\Big) \leqslant (r/3)(1-4\epsilon+Chk^{-\lambda/2}+hk^{-1})+(2r/3)(1+Chk^{-\lambda/2})$$
$$\leqslant r(1-\epsilon). \qquad \square$$

§2·4 A Simple Consequence of Azuma's Inequality

In this short section, we shall give one very standard application of Azuma's inquality. In fact, it is a special case of a general result of Schechtman [46] (cf. also [39]). We shall use the set $\Omega = \{-1,1\}^n \times S_n$ a great deal in this dissertation, since each element of this set corresponds in an obvious way to a symmetry of an n-dimensional 1-symmetric normed space. (We shall be completely explicit about this in the next chapter.) We consider this set as a metric space in various ways, but in each case we use a weighted versions of a distance known as the Hamming distance. Roughly speaking, given two elements of our set above, the Hamming distance between them is the number of coordinates where they differ. To be precise, given (ϵ,π) and (ϵ',π') in the set, the distance $d[(\epsilon,\pi),(\epsilon',\pi')]$ between them is given by

$$d[(\epsilon,\pi),(\epsilon',\pi')] = \big|\{i\colon \epsilon_i \neq \epsilon_i' \text{ or } \pi(i)\neq\pi'(i)\}\big|.$$

A weighted version of the Hamming distance is defined as follows. Given a positive sequence $(b_i)_1^n$, let the corresponding distance between (ϵ,π) and (ϵ',π') be given by

$$d[(\epsilon,\pi),(\epsilon',\pi')] = \sum\{b_i\colon \epsilon_i \neq \epsilon_i' \text{ or } \pi(i)\neq\pi'(i)\}.$$

Thus, the ordinary Hamming distance is what one obtains when each b_i takes the value 1. For ease of notation, we shall assume for the rest of the section that the particular choice $(b_i)_1^n$ of weights has been fixed. Let (Ω, d) be the corresponding metric space. We turn (Ω, d) into a metric probability space (Ω, d, P) by letting P be the normalized counting measure on Ω. Our aim in this section is to show that, if f is a Lipschitz function on (Ω, d), then, if (ϵ, π) is chosen at random from Ω, the probability of $f((\epsilon, \pi))$ differing much from its expected value is small. That is, (Ω, d, P) exhibits the phenomenon of measure concentration.

Azuma's inequality (Theorem 1.1) is useful in this context because there is a natural sequence of sigma-fields on Ω. With these, we can use a Lipschitz function f on (Ω, d) to define a martingale. We define equivalence relations $\sim_0, \ldots, \sim_n$ on Ω by $(\epsilon, \pi) \sim_i (\epsilon', \pi')$ iff $\epsilon_j = \epsilon'_j$ and $\pi(j) = \pi'(j)$ for $1 \leqslant j \leqslant i$. For $1 \leqslant i \leqslant n$ let $\mathcal{F}_i$ be the sigma-field whose atoms are the equivalence classes of $\sim_i$. Finally, let f be a γ-Lipschitz function on Ω, and set $f_i = \mathsf{E}(f \mid \mathcal{F}_i) \quad (1 \leqslant i \leqslant n)$. We have the following corollary of Theorem 1.1.

Proposition 2·9. *Let (Ω, d, P) and $f_0, \ldots, f_n$ be defined as above. Then for all $s > t$ and $\delta > 0$,*

$$\mathsf{P}\Big[f_s - f_t \geqslant \delta\Big] \leqslant \exp\Big(-\frac{\delta^2}{8\gamma^2 \sum_{t+1}^{s} b_i^2}\Big)$$

and

$$\mathsf{P}\Big[f_s - f_t \leqslant -\delta\Big] \leqslant \exp\Big(-\frac{\delta^2}{8\gamma^2 \sum_{t+1}^{s} b_i^2}\Big).$$

Proof. We shall prove the first inequality above. The second can be deduced from it by looking at the function $-f$ instead of f. We restrict our attention to a single atom of $\mathcal{F}_r$. It is then obvious that without loss of generality $s = n$ and $r = 0$. By Theorem 1.1, we need only show that, for $1 \leqslant i \leqslant n$, $f_i - f_{i-1} \leqslant 2\gamma b_i$.

Suppose $A, B \in \mathcal{F}_i$, $A, B \subset C \in \mathcal{F}_{i-1}$, and let (η, σ) be an element of B.

Then let ϕ be the bijection from A to B given by $(\epsilon, \pi) \mapsto (\epsilon', \pi')$, where

$$\epsilon'_j = \begin{cases} \epsilon_j & j \neq i \\ \eta_j & j = i \end{cases}$$

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

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

Since f is γ-Lipschitz, f_i varies by at most $2\gamma b_i$ in any atom of $\mathcal{F}_{i-1}$. It follows easily that $f_i - f_{i-1} \leqslant 2\gamma b_i$ as was needed. □

§2·5 A Generalization of the Hypergeometric Distribution

This section concerns a weighted version of the hypergeometric distribution. Specifically, we ask the following question. Suppose $\mathbf{a} = (a_i)_1^n$ is a given positive vector in $\mathbb{R}^n$ and B is a randomly chosen set in $[n]^{(k)}$. Then, given $t \geqslant 0$, what is the probability that $\sum_{i\in B} a_i \geqslant t$? If $\mathbf{a}$ is just the characteristic function of a set of size l then this probability can be estimated from above and below by using known bounds for the hypergeometric distribution. We shall need estimates in both directions (in different contexts) but only weak ones. Our two main results will be, roughly speaking, that if the maximum value of the a_i is at most 1, and $\sum_1^n a_i$ is at most l, then the probability of $\sum_{i\in B} a_i$ deviating from its mean by a given proportion of the mean is largest when $a_1 = \ldots = a_l = 1$ and the rest of the a_i are zero, whereas if one restricts the size of the support of the sequence to l and tries to minimize this probability, one cannot do much better than to pick the same sequence. In short, the probability, as one might expect, gets smaller the more the sequence is spread out.

Our first lemma is a straightforward deduction from the theorem of Hoeffding stated in the last chapter (Theorem 1.2). In fact, a slightly stronger result can be found in Hoeffding's paper [26] but since we shall not need the extra strength, and the next argument is a simple one, we shall content ourselves with the following statement.

Lemma 2·10. *Let $\lambda > 2e$, let k, l and n be positive integers, let $\mathbf{a} = (a_i)_1^n$ be a non-negative vector in $\mathbb{R}^n$ satisfying $\|\mathbf{a}\|_\infty \leqslant 1$ and $\|\mathbf{a}\|_1 \leqslant l$ and let $A \in [n]^{(k)}$ be chosen at random. Then*

$$\mathbb{P}\left[\sum_{i\in A} a_i \geqslant \lambda kl/n\right] \leqslant 2(2e/\lambda)^{\lambda kl/n} .$$

Proof. Let $X_1', \ldots, X_n'$ be a sequence of independent Bernoulli random variables of mean $2k/n$, and set $X_i = a_i X_i'$ for each i. Let $\tau \subset [n]$ be the (variable) set $\{i : X_i = 1\}$.

Then

$$\begin{aligned}
\mathsf{P}\left[\sum_{i=1}^{n} X_i \geqslant \lambda k l\right] &= \sum_{j=0}^{n} \mathsf{P}\left[\sum_{i=1}^{n} X_i \geqslant \lambda k l \,\middle|\, |\tau| = j\right] \mathsf{P}\Big[|\tau| = j\Big] \\
&\geqslant \sum_{j=k}^{n} \mathsf{P}\left[\sum_{i=1}^{n} X_i \geqslant \lambda k l \,\middle|\, |\tau| = j\right] \mathsf{P}\Big[|\tau| = j\Big] \\
&\geqslant \mathsf{P}\left[\sum_{i=1}^{n} X_i \geqslant \lambda k l \,\middle|\, |\tau| = k\right] \mathsf{P}\Big[|\tau| \geqslant k\Big]
\end{aligned}$$

where the last inequality followed from the obvious fact that

$$\mathsf{P}\left[\sum_{i=1}^{n} X_i \geqslant \lambda k l \,\middle|\, |\tau| = j\right]$$

is an increasing function of j.

However,

$$\mathsf{P}\left[\sum_{i=1}^{n} X_i \geqslant \lambda k l \,\middle|\, |\tau| = k\right]$$

is exactly the quantity we wish to estimate. We shall therefore show that $\mathsf{P}\Big[|\tau| \geqslant k\Big]$ is at least $1/2$ and $\mathsf{P}\left[\sum_{i=1}^{n} X_i \geqslant \lambda k l\right]$ is at most $(2e/\lambda)^{\lambda k l/n}$. This will prove the lemma.

First, we have $|\tau| = \sum_1^n X_i'$ and $\mathsf{E}|\tau| = 2k$. By the second part of Theorem 1.2 we therefore have

$$\begin{aligned}
\mathsf{P}\Big[|\tau| \leqslant k\Big] &= \mathsf{P}\Big[|\tau| - 2k \leqslant -k\Big] \\
&\leqslant 2^k \left(\frac{1 - 2k/n}{1 - k/n}\right)^{n-k} \\
&\leqslant 2^k e^{-k} = (2/e)^k ,
\end{aligned}$$

which is at most $1/2$ when k is sufficiently large. The important estimate is of course the upper bound for $\mathsf{P}\left[\sum_{i=1}^{n} X_i \geqslant \lambda k l\right]$. Setting $l' = \|\mathbf{a}\|_1$, we apply the first part of Theorem 1.2, with $p = 2kl'/n^2$ and $t = (\lambda kl'/n^2) - p$, obtaining

$$\mathsf{P}\left[\sum_{i=1}^{n} X_i \geqslant \lambda kl\right] \leqslant \left(\frac{2}{\lambda}\right)^{\lambda kl/n} \cdot \left(\frac{n^2 - 2kl}{n^2 - \lambda kl}\right)^{n - \lambda kl/n}$$
$$\leqslant (2/\lambda)^{\lambda kl/n} \cdot e^{\lambda kl/n} = (2e/\lambda)^{\lambda kl/n} . \qquad \square$$

Our estimate from below will be needed in Chapter 6. The situation that will interest us is when the size of the support of the vector **a** is around $\log n$ and k is proportional to n. The statement of the next lemma is slightly unnatural because of the appearance of the parameter t. This is to give us a little elbow-room when we come to apply it.

Lemma 2·11. *Let $r = \epsilon n$, $h \leqslant r$ and let $\mathbf{a} = (a_i)_1^n \in \mathbb{R}^n$ be a vector such that $a_1 \geqslant \ldots \geqslant a_h \geqslant a_{h+1} = \ldots = a_n = 0$ and $\sum_1^n a_i = 1$. Then if $t < r/4$ and K is chosen randomly from $[n-t]^{(r-t)}$, we have*

$$\mathsf{P}\left[\sum_{i\in K} a_i \geqslant 2\epsilon\right] \geqslant (1/16)^{32\epsilon h} \cdot (1-2\epsilon)^{2h}$$

and

$$\mathsf{P}\left[\sum_{i\in K} a_i = 0\right] \geqslant (1-2\epsilon)^h .$$

Proof. Note that $\mathsf{E}\left(\sum_{i\in K} a_i\right) = \epsilon$. Clearly $\mathsf{P}\left[\sum_{i\in K} a_i = 0\right] \geqslant \binom{n-r}{h}/\binom{n-t}{h} \geqslant \left(\frac{n-r-h}{n-h-t}\right)^h \geqslant (1-2\epsilon)^h$, as stated.

For the first estimate, we use the lower bounds for the hypergeometric distribution given by Lemma 1.4. For $l \leqslant r/2$, $0 < \alpha < 1$ we obtain

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

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

Suppose $|B_j \cap K| \geqslant 8\epsilon|B_j|$ for $1 \leqslant j \leqslant s$. In this case

$$\sum_{i\in K} a_i \geqslant \sum_{j=1}^{s} \sum\{a_i : i \in (B_j \setminus B_{j-1}) \cap K\} \geqslant \sum_{j=1}^{s} 2^{-j}|(B_j \cap K) \setminus (B_{j-1} \cap K)|$$

$$\geqslant \sum_{j=1}^{s} |B_j \cap K| \left(2^{-j} - 2^{-(j+1)}\right) \geqslant 4\epsilon \sum_{j=1}^{s} 2^{-j}|B_j|$$

$$= 4\epsilon \sum_{j=1}^{s} \left(2^{-(j-1)} - 2^{-j}\right)|B_j| \geqslant 4\epsilon \sum_{j=1}^{s} 2^{-(j-1)}|B_j \setminus B_{j+1}| \geqslant 2\epsilon$$

since $\sum\{a_i : a_i \leqslant 2^{-s}\} \leqslant 1/2$.

But

$$\mathsf{P}\Big[|B_j \cap K| \geqslant 8\epsilon|B_j| \;\Big|\; |B_{j+1} \cap K| \geqslant 8\epsilon|B_{j+1}|\Big]$$
$$\geqslant \binom{|B_j|}{8\epsilon|B_j|}\binom{n-r}{(1-8\epsilon)|B_j|} \Big/ \binom{n-t}{|B_j|}$$
$$= (1/16)^{8\epsilon|B_j|}(1-2\epsilon)^{|B_j|} \qquad (1 \leqslant j < s)$$

and

$$\mathsf{P}\Big[|B_s \cap K| \geqslant 8\epsilon|B_s|\Big] \geqslant (1/16)^{8\epsilon|B_s|}(1-2\epsilon)^{|B_s|}.$$

Hence

$$\mathsf{P}\Big[|B_j \cap K| \geqslant 8\epsilon|B_j| \;\forall 1 \leqslant j \leqslant s\Big] \geqslant (1/16)^{8\epsilon\sum_1^s |B_j|} \cdot (1-2\epsilon)^{\sum_1^s |B_j|}.$$

But $|B_j| \leqslant 2^j$, so $\sum_1^s |B_j| \leqslant 4h$, and so

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

Note that, in the proof of Lemma 2·11, we actually gave a lower estimate for the smaller event $\Big[|B_j \cap K| \geqslant 8\epsilon|B_j| \text{ for every } 1 \leqslant j \leqslant s\Big]$. This gives us the next corollary.

Corollary 2·12. *Let $r = \epsilon n$, $h \leqslant r$ and let $\mathbf{a} = (a_i)_1^n \in \mathbb{R}^n$ be a vector such that $a_1 \geqslant \ldots \geqslant a_h \geqslant a_{h+1} = \ldots = a_n = 0$ and $\sum_{i=1}^n a_i = 1$. If $t, r/4$ and K is chosen randomly from $[n-t]^{(r-t)}$, then, with probability at least $(1/16)^{32\epsilon h}(1-2\epsilon)^{2h}$, there exists a subset $A \subset [h] \cap K$ such that $|A| \leqslant 8\epsilon h$ and $\sum_{i\in A} a_i \geqslant 2\epsilon$.* □

As our next result of this kind, we show that there exists a large collection of sets $\mathcal{B} \subset [t]^{(t/2)}$, such that if B, $C \in \mathcal{B}$ are distinct sets in the collection, then $|B \cap C| \leqslant t/3$. Instead of using the theorem of Hoeffding, however, we shall prove the result directly from Lemma 1.4. This gives us the set $\mathcal{B}$ that we used in our construction just before Lemma 2.7.

Lemma 2·13. *There is a subset $\mathcal{B} \subset [t]^{(t/2)}$ of cardinality $(23/20)^t$ such that given any two distinct sets B, $C \in \mathcal{B}$, $|B \cap C| \leqslant t/3$.*

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 $\sum_{r=0}^{t/6} \binom{t/2}{t/6-r}\binom{t/2}{t/3+r} \leqslant 3\binom{t/2}{t/3}\binom{t/2}{t/6}$. Hence, by picking sets one at a time, each one disjoint from the previous ones, we can find $\mathcal{B}$ with $|\mathcal{B}| \geqslant \frac{1}{3}\binom{t}{t/2}/\binom{t/2}{t/3}\binom{t/2}{t/6}$. But, by Lemma 1.4,

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

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

Our final result of the section is, in a sense, a generalization of Lemma 2.13. In conjunction with Lemma 2.8, it gives a lower estimate for the size of a well-separated collection of subsets of the unit ball of an isomorph of ℓ_1^n.

Lemma 2·14. *Let $n, k \in \mathbb{N}$ and let k be an even number. There exists a subset*

$\mathcal{S} \subset [n]^k$ of cardinality $N = (n/3e)^{k/12}$ such that, whenever $\mathbf{s} \neq \mathbf{s}' \in \mathcal{S}$,

$$\left|\{i \in [k] : s_i < s_i'\}\right| \geqslant k/3.$$

Proof. Let us pick $\mathbf{s}$ and $\mathbf{s}'$ randomly from $[n]^{k/2}$ and estimate

$$\mathsf{P}\left[\left|\{i \in [k/2] : s_i = s_i'\}\right| > k/6\right].$$

Now $\left|\{i \in [k/2] : s_i = s_i'\}\right|$ is simply a sum of $k/2$ independent Bernoulli variables of mean $1/n$. By standard estimates for the binomial distribution [10], or by a simple application of Theorem 1.2, we obtain that the above probability is at most $(3e/n)^{k/6}$. Let us pick a set $\mathcal{S}' \subset [n]^{k/2}$ of cardinality N randomly. The probability that any two elements of $\mathcal{S}$ coincide in more than $k/6$ coordinates is certainly at most $\binom{N}{2}(3e/n)^{k/6}$. This is less than one, since $N \leqslant (n/3e)^{k/12}$. Thus there exists $\mathcal{S}' \subset [n]^{k/2}$ such that, for any $\mathbf{s} \neq \mathbf{s}' \in \mathcal{S}'$, $\left|\{i \in [k/2] : s_i = s_i'\}\right| \leqslant k/6$. Given $\mathbf{s} \in \mathcal{S}'$, define $\phi(\mathbf{s}) \in [n]^k$ by

$$(\phi(\mathbf{s}))_i = \begin{cases} s_i & i \leqslant k/2 \\ k+1-s_{i-k/2} & i > k/2 \end{cases}$$

Let $\mathcal{S} = \{\phi(\mathbf{s}) : \mathbf{s} \in \mathcal{S}'\}$. Then if $\mathbf{s} \neq \mathbf{s}' \in \mathcal{S}$, we clearly have that

$$\left|\{i \in [k] : s_i < s_i'\}\right| = \left|\{i \in [k] : s_i' < s_i\}\right|$$

and the sum of the two sides is at least $2k/3$. It follows that $\mathcal{S}$ has the desired properties. □

PART II

THE SIZE OF SYMMETRIC BLOCK BASES:

LOWER BOUNDS

CHAPTER 3

BASES WITH LARGE AVERAGE GROWTH

The main theorem of this chapter asserts that if $x_1, \ldots, x_n$ is a sequence of unit vectors in a normed space, and if $\mathsf{E}\,\|\sum_{i=1}^n \epsilon_i x_i\|$ is large, then $x_1, \ldots, x_n$ must have a large almost symmetric block basis. Here the expectation is taken over all possible choices of signs $\epsilon_1, \ldots, \epsilon_n$, distributed uniformly. Before proving anything, we shall point out a few simple facts to show why one considers this particular problem.

There are important results, such as a well known result of Elton [16], which show that sequences of vectors in a normed space satisfying certain conditions must have subsequences of large cardinality satisfying stronger conditions. One might ask whether, if a sequence satisfies a growth condition such as the one we are about to consider, it must have a large almost symmetric subsequence. However, simple examples show that this is certainly not the case. For example, given $1 < p \leqslant 2$, one can take the space $X = (\mathbb{R}^n, \|.\|)$, where the norm is given by

$$\|\mathbf{a}\| = \max\left\{ \|\mathbf{a}\|_p , \left|\sum_1^n a_i\right| \right\}.$$

Then the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ has large average growth, but, given any i and j, we have $\|\mathbf{e}_i + \mathbf{e}_j\| = 2$, while $\|\mathbf{e}_i - \mathbf{e}_j\| = 2^{1/p}$. If a normed space X has a natural basis and one wants to find an almost symmetric basic sequence in the space (as one does when proving local versions of Krivine's theorem and the Maurey-Pisier theorem), it is sensible anyway to try a block basis, and even a block basis with ± 1-coefficients and blocks of equal length. For example, any such block basis can be transformed into any other by the linear map induced by a permutation and change of signs of the original basis, and second, if the original basis is 1-symmetric, then such a block basis will be as well. Furthermore, if a basis is equivalent to the standard basis of ℓ_p^n, then a normalized block basis of it will be equivalent to the standard basis of the ℓ_p-space of the appropriate dimension.

Theorem 3·1. *Let $1 \leqslant p < 2$, $0 < \epsilon \leqslant 1$, and let $(x_i)_1^n$ be a sequence of unit vectors in a Banach space X. Suppose that $\mathsf{E}\,\|\sum_1^n \epsilon_i x_i\| \geqslant n^{1/p}$. Then there is a block basis of $(x_i)_1^n$ with blocks of ± 1 coefficients and equal length, which is $(1+\epsilon)$-symmetric and has cardinality*

$$m \geqslant \gamma(\epsilon)(\log n)^{-1} n^{\frac{2}{p}-1}$$

where $\gamma(\epsilon) = (\epsilon^3/3,000,000)(\log(33/\epsilon))^{-1}$.

The dependence on n in this theorem improves on the previously known dependence of $n^{(2-p)^2/3p^3}$ obtained by Amir and Milman in [4, Section 5.2]. They have also shown that, if X has p-type constant C, one can obtain a bound of $\gamma(\epsilon, p, C)n^{(2-p)/3p(2+3p)}$ [3, Theorem 2.4]. Our exponent of n of $2/p-1$ therefore improves both these results, with or without an assumption about type-conditions. We shall say a little about the effect of type conditions at the end of this chapter and more at the end of the next one.

We shall begin with some notation.

Let Ψ be the group $\{-1,1\}^m \times S_m$ with multiplication given by $((\eta_i)_1^m, \sigma) \circ ((\eta_i')_1^m, \sigma') = ((\eta_i \eta_i')_1^m, \sigma \circ \sigma')$, acting on $\mathbf{R}^m$ as follows. If $\mathbf{a} = \sum_1^m a_i \mathbf{e}_i \in \mathbf{R}^m$, and $(\eta, \sigma) \in \Psi$, $\eta = (\eta_i)_1^m$, then

$$(\eta, \sigma) \colon \mathbf{a} \mapsto \psi_{\eta,\sigma}(\mathbf{a}) \equiv \mathbf{a}_{\eta,\sigma} = \sum_1^m \eta_i a_i \mathbf{e}_{\sigma(i)}.$$

Let Ω be the group $\{-1,1\}^n \times S_n$ acting on X as follows. If $\mathbf{b} \in X$, $\mathbf{b} = \sum_1^n b_i x_i$, and $(\epsilon, \pi) \in \Omega$, then

$$(\epsilon, \pi) \colon \mathbf{b} \mapsto \omega_{\epsilon,\pi}(\mathbf{b}) \equiv \mathbf{b}_{\epsilon,\pi} = \sum_1^n \epsilon_i b_i x_{\pi(i)}.$$

We shall sometimes relabel the indices of $(x_i)_1^n$. Let $x_{ij} \equiv x_{(i-1)h+j}$ $(i = 1, \ldots, m,\ j = 1, \ldots, h)$, where $hm = n$, and similarly let $\epsilon_{ij} \equiv \epsilon_{(i-1)h+j}$ and $\pi_{ij} \equiv \pi((i-1)h + j)$ for $(\epsilon, \pi) \in \Omega$.

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

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

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

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

When $1 \leqslant i \leqslant m$ we shall write $(\mathbf{u}_i)_{\epsilon,\pi}$ for $\phi_{\epsilon,\pi}(\mathbf{e}_i)$. The sequence $((\mathbf{u}_i)_{\epsilon,\pi})_1^m$ will be our random block basis.

Like Amir and Milman, we shall rely heavily on measure-concentration techniques when we prove Theorem 3.1. The additional strength in our proof comes from considering, for an arbitrary vector $\mathbf{a} \in \mathsf{R}^m$, the whole class

$$\left\{ \mathbf{a}_{\eta,\sigma} \ :\ (\eta, \sigma) \in \Psi \right\}$$

at once, and estimating the probability that the norms of the images of all the vectors in the class are about the same. One might expect that some sort of result about sub-Gaussian processes would be useful here: in fact, the only probabilistic tool we need is Azuma's inequality (Theorem 1.1).

The proof of Theorem 3.1 is based on two statements. The first is Proposition 2.1, and the second is the following.

Lemma 3·2. *(i) Let the sequence $x_1, \ldots, x_n \in X$ satisfy the conditions of Theorem 3.1, let q be the conjugate index of p and let $\phi_{\epsilon,\pi}\colon \mathsf{R}^m \mapsto X$ be as defined above. Then, for any $\mathbf{a} \in \mathsf{R}^m$,*

$$\mathsf{E}_\Omega \left\| \phi_{\epsilon,\pi}(\mathbf{a}) \right\| \geqslant n^{-1/q} \left\| \mathbf{a} \right\|_1 h .$$

(ii) Let ϵ and m be as in the statement of Theorem 3.1, let $h = n/m$ and let $\delta = \epsilon/11$. For any $\mathbf{a} \in \mathsf{R}^m$, let $E(\mathbf{a})$ stand for $\mathsf{E}_\Omega \|\phi_{\epsilon,\pi}(\mathbf{a})\|$. Then

$$\mathsf{P}_\Omega \left[\exists(\eta,\sigma) \text{ s.t. } \left|\|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\| - E(\mathbf{a})\right| > \frac{\delta}{3} n^{-1/q} \|\mathbf{a}\|_1 h\right] < m^{-\delta^{-1}\log(3\delta^{-1})}.$$

The main step in the argument is the second part of Lemma 3.2. This states, roughly, that for any vector $\mathbf{a} \in \mathsf{R}^m$, the deviation of any vector of the form $\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})$ from their common average is, with large probability, small. The first part provides a lower bound for the average itself, so that the deviation is in fact proportionately small. Let us see why Proposition 2.1 and Lemma 3.2 are enough to prove Theorem 3.1.

If they are both true, then we can choose vectors $\mathbf{a}_1, \dots, \mathbf{a}_N$ that satisfy the conclusion of Proposition 2.1. Having done this, let us consider a single vector $\mathbf{a} = \mathbf{a}_i$. Taking $\delta = \epsilon/11$, we have $\delta < 1/11$, so certainly $(1+\delta/3)(1-\delta/3)^{-1} \leqslant 1+\delta$. Hence, by the two parts of Lemma 3.2,

$$\mathsf{P}_\Omega \left[\max_\Psi \|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\| \,/ \min_\Psi \|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\| > 1+\delta\right] < m^{-\delta^{-1}\log(3\delta^{-1})} = N^{-1}$$

It follows that there exists some $(\epsilon,\pi) \in \Omega$ such that if we define a norm on R^m by $\|\mathbf{a}\| \equiv \|\phi_{\epsilon,\pi}(\mathbf{a})\|$, then this norm is $(1+\delta)$-symmetric at each of the vectors $\mathbf{a}_1, \dots, \mathbf{a}_N$. But then, by Proposition 2.1, this norm is $(1+\delta)(1-6\delta)^{-1}$-symmetric on R^m. Since $\delta = \epsilon/11 \leqslant 1/11$, we have $(1+\delta)(1-6\delta)^{-1} \leqslant 1+\epsilon$, so the block basis $((\mathbf{u}_i)_{\epsilon,\pi})_1^m$ is $(1+\epsilon)$-symmetric. Thus Theorem 3.1 does indeed follow from Proposition 2.1 and Lemma 3.2.

Proof of Lemma 3.2 (i). This part is a simple observation. Let $|||\mathbf{a}|||$ stand for $\mathsf{E}_\Omega \|\phi_{\epsilon,\pi}(\mathbf{a})\|$. Then $|||.|||$ is a 1-symmetric norm on R^m, and by hypothesis $|||\sum_1^m \mathbf{e}_i||| \geqslant n^{1/p}$. Now let $\mathbf{a} = \sum_1^m a_i\mathbf{e}_i$ be a vector in R^m with $a_i \geqslant 0$ for every i.

For $1 \leqslant k \leqslant m$ let $\mathbf{a}^{(k)}$ be the vector $\sum_1^m a_i \mathbf{e}_{i+k}$ (where $i+k$ is reduced modulo m). Then, by the triangle inequality,

$$|||\mathbf{a}||| \geqslant m^{-1} \left|\left|\left| \sum_{k=1}^{m} \mathbf{a}^{(k)} \right|\right|\right| = m^{-1} \left|\left|\left| \sum_{i=1}^{m} \Big(\sum_{j=1}^{m} a_j \Big) \mathbf{e}_i \right|\right|\right| = m^{-1} \|\mathbf{a}\|_1 n^{1/p} = n^{-1/q} \|\mathbf{a}\|_1 h.$$

Proof of Lemma 3.2 (ii) Without loss of generality we may assume that $\|\mathbf{a}\|_1 = 1$ and that the coordinates of $\mathbf{a}$ satisfy $a_1 \geqslant \ldots \geqslant a_m \geqslant 0$. This is for ease of notation. Let $B_1, \ldots, B_{k+1} \subset [m]$ be defined by

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

where $k = \log_2(200mn^{1/q}/\epsilon)$.

Let $\mathbf{b}_1, \ldots, \mathbf{b}_k$ be given by $\mathbf{b}_j = \mathbf{a}|_{B_j}$, $(1 \leqslant j \leqslant k)$. Given $(\eta, \sigma) \in \Psi$, let us define $\mathbf{b}^j_{\eta,\sigma}$ to be $(\mathbf{b}_j)_{\eta,\sigma}$. Clearly $\mathbf{b}^j_{\eta,\sigma} = \mathbf{a}_{\eta,\sigma}|_{\sigma(B_j)}$, and the absolute values of the coefficients of $\mathbf{b}^j_{\eta,\sigma}$ lie between 2^{-j} and $2^{-(j-1)}$ when $j \leqslant k$, and are at most $\epsilon/200mn^{1/q}$ when $j = k+1$.

For each $1 \leqslant r \leqslant k$ and $(\eta, \sigma) \in \Psi$, let the function $f^r_{\eta,\sigma} : \Omega \longrightarrow \mathbb{R}$ be defined by

$$f^r_{\eta,\sigma}((\epsilon, \pi)) = \mathsf{E}\Big[\|\phi_{\epsilon',\pi'}(\mathbf{a}_{\eta,\sigma})\| \;\Big|\; \phi_{\epsilon',\pi'}(\mathbf{b}^j_{\eta,\sigma}) = \phi_{\epsilon,\pi}(\mathbf{b}^j_{\eta,\sigma}), j = 1, \ldots, r \Big]$$

Now, for any fixed (η, σ), the sequence of functions $f^0_{\eta,\sigma}$ $(= \mathsf{E}(\|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|))$, $f^1_{\eta,\sigma}, \ldots, f^k_{\eta,\sigma}$ is a martingale. Note also that $f^k_{\eta,\sigma}(\epsilon, \pi) = \|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|$, although the expectation is not taken over a singleton subset of Ω. This is because if $\phi_{\epsilon',\pi'}(\mathbf{b}^j_{\eta,\sigma}) = \phi_{\epsilon,\pi}(\mathbf{b}^j_{\eta,\sigma})$ for $j = 1, \ldots, k$ then $\phi_{\epsilon',\pi'}(\mathbf{a}_{\eta,\sigma}) = \phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})$. As it happens the fact that $f^0_{\eta,\sigma}, \ldots, f^k_{\eta,\sigma}$ is a martingale will not concern us. Instead we are interested in upper bounds for the following two quantities:

(a) the number of *distinct* functions $f^r_{\eta,\sigma}$ for any given r;

(b) the probability, for given r and (η, σ), that $f^r_{\eta,\sigma}((\epsilon, \pi))$ differs substantially from $f^{r-1}_{\eta,\sigma}((\epsilon, \pi))$, if (ϵ, π) is chosen randomly from Ω.

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

We shall obtain an estimate for (b) by using Proposition 2.9. Indeed, let us define our weights $b_1, \ldots, b_n$ by $b_i = a_{\lceil (i/h) \rceil}$ for each $i \in [n]$ and set $s = \gamma_r h$ and $t = \gamma_{r-1} h$. Let f be the function given by $f((\epsilon, \pi)) = \|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|$. We have $2^{-r} \leqslant b_i \leqslant 2^{-(r-1)}$ for $\gamma_{r-1} h \leqslant i \leqslant \gamma_r h$.

Since

$$(\epsilon, \pi) \sim_s (\epsilon', \pi') \Rightarrow \phi_{\epsilon,\pi}(\mathbf{b}^j_{\eta,\sigma}) = \phi_{\epsilon',\pi'}(\mathbf{b}^j_{\eta,\sigma})$$

when $1 \leqslant j \leqslant r$, and

$$(\epsilon, \pi) \sim_t (\epsilon', \pi') \Rightarrow \phi_{\epsilon,\pi}(\mathbf{b}^j_{\eta,\sigma}) = \phi_{\epsilon',\pi'}(\mathbf{b}^j_{\eta,\sigma})$$

when $1 \leqslant j \leqslant r-1$, and f is 1-Lipschitz, we obtain from Proposition 2.9 that

$$\begin{aligned} \mathsf{P}\left[f^r_{\eta,\sigma}((\epsilon,\pi)) - f^{r-1}_{\eta,\sigma}((\epsilon,\pi)) > \delta_r h n^{-1/q}\right] &< \exp\left(-\frac{\delta_r^2 h^2 n^{-2/q}}{8(\gamma_r - \gamma_{r-1})h \cdot 2^{-2(r-1)}}\right) \\ &= \exp\left(-\frac{2^{2(r-1)}\delta_r^2 h}{8n^{2/q}\beta_r}\right) \end{aligned}$$

Similarly, we obtain

$$\mathsf{P}\left[f^r_{\eta,\sigma}((\epsilon,\pi)) - f^{r-1}_{\eta,\sigma}((\epsilon,\pi)) < -\delta_r h n^{-1/q}\right] \leqslant \exp\left(-\frac{2^{2(r-1)}\delta_r^2 h}{8n^{2/q}\beta_r}\right).$$

Note that the above probabilities are both zero in the case $\beta_r = 0$.

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

$$f^{k+1}_{\eta,\sigma}((\epsilon,\pi)) - f^k_{\eta,\sigma}((\epsilon,\pi)) \leqslant (\epsilon/200)hn^{-1/q} = (\epsilon/200)hn^{-1/q}\|\mathbf{a}\|_1$$

Recall that $E(\mathbf{a})$ stands for $\mathsf{E}_\Omega(\|\phi_{\epsilon,\pi}(\mathbf{a})\|)$, and note that this is the same as $\mathsf{E}_\Omega(\|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|)$ for any $(\eta,\sigma)\in\Psi$.

Now suppose that for some $(\epsilon,\pi)\in\Omega$ it is true that there exists some (η,σ) for which

$$\|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\| - E(\mathbf{a}) > \frac{\epsilon}{33}\|\mathbf{a}\|_1\, hn^{-1/q},$$

that is, for which

$$f^{k+1}_{\eta,\sigma}((\epsilon,\pi)) - f^{0}_{\eta,\sigma}((\epsilon,\pi)) > \frac{\epsilon}{33}\|\mathbf{a}\|_1\, hn^{-1/q}.$$

Then

$$f^{k}_{\eta,\sigma}((\epsilon,\pi)) - f^{0}_{\eta,\sigma}((\epsilon,\pi)) > \frac{\epsilon}{40}\|\mathbf{a}\|_1\, hn^{-1/q}\ ,$$

so, if $\delta_1+\ldots+\delta_k\leqslant\epsilon/40$, there will be some $1\leqslant r\leqslant k$ such that

$$f^{r}_{\eta,\sigma}((\epsilon,\pi)) - f^{r-1}_{\eta,\sigma}((\epsilon,\pi)) > \delta_r\|\mathbf{a}\|_1\, hn^{-1/q}.$$

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

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

where s is the smallest value of r for which $\gamma_r>0$. This is therefore an upper bound for the probability we wish to estimate.

It remains for us to choose appropriate $\delta_1,\ldots,\delta_k$ and to verify that this probability is at most $\frac{1}{2}\cdot m^{-\delta^{-1}\log(3\delta^{-1})}$. Since the inequality in the other direction is exactly similar, we will then be done.

Choosing $\delta_r = 2^{-r}\beta_r^{1/2}\gamma_r^{1/2}\cdot\epsilon/66$ will do.

First, using the Cauchy-Schwarz inequality and the fact that $\sum_1^k 2^{-r}\beta_r\leqslant$

$\sum_1^m a_i = 1$ and $\gamma_r = \sum_{j=1}^r \beta_j$, we obtain

$$\begin{aligned}
\sum_1^k \delta_r &= \frac{\epsilon}{66} \cdot \sum_1^k 2^{-r} \beta_r^{1/2} \gamma_r^{1/2} \\
&\leqslant \frac{\epsilon}{66} \left(\sum_1^k 2^{-r} \beta_r \right)^{1/2} \left(\sum_1^k 2^{-r} \gamma_r \right)^{1/2} \\
&\leqslant \frac{\epsilon}{66} \left(\sum_{r=1}^k 2^{-r} \sum_{j=1}^r \beta_j \right)^{1/2} \\
&= \frac{\epsilon}{66} \left(\sum_{j=1}^k \beta_j \sum_{r=j}^k 2^{-r} \right)^{1/2} \\
&< \frac{\epsilon\sqrt{2}}{66} \left(\sum_{j=1}^k 2^{-j} \beta_j \right)^{1/2} \leqslant \frac{\epsilon}{40}
\end{aligned}$$

Second,

$$\begin{aligned}
\sum_{r=s}^k (2m)^{\gamma_r} \exp\left(-\frac{2^{2(r-1)} \delta_r^2 h}{8 \beta_r n^{2/q}} \right) &= \sum_{r=s}^k \exp\left(\gamma_r \left(\log(2m) - \frac{\epsilon^2 h}{4 \times 8 \times 66^2 n^{2/q}} \right) \right) \\
&\leqslant k \exp\left(\log(2m) - \frac{\epsilon^2 h}{150,000 n^{2/q}} \right)
\end{aligned}$$

since $\gamma_r > 0$ for every $r \geqslant s$.

But since $h \geqslant (20/\epsilon) \cdot (\log(33/\epsilon)) \cdot (150,000/\epsilon^2) \cdot n^{2/q} \log n$, we have

$$\begin{aligned}
k \exp\left(\log(2m) - \frac{\epsilon^2 h}{150,000 n^{2/q}} \right) &\leqslant k \exp\Big(\big(1 - (20/\epsilon) \log(33/\epsilon) \big) \log n \Big) \\
&\leqslant \frac{1}{2} \exp\Big(-(11/\epsilon) \log(33/\epsilon) \log n \Big) \\
&\leqslant \frac{1}{2} m^{-(11/\epsilon)\log(33/\epsilon)} \\
&= \frac{1}{2} m^{-\delta \log(3\delta^{-1})}
\end{aligned}$$

which is what we needed. This completes the proof of Lemma 3.2 (ii), and with it the proof of Theorem 3.1. □

Let us finish this chapter with two simple corollaries.

Corollary 3·3. *Let $x_1, \ldots, x_n$ be a sequence of vectors in a normed space C-equivalent to the unit vector basis of ℓ_p^n. Then there is a $(1+\epsilon)$-symmetric block basis $(\mathbf{u}_j)_1^m$ of $(x_i)_1^n$ with blocks of ± 1 coefficients and equal length, of cardinality*

$$m \geqslant \gamma(\epsilon) C^{-2} (\log n)^{-1} n^{\frac{2}{p}-1} ,$$

where $\gamma(\epsilon) = (\epsilon^3/3{,}000{,}000)(\log(33/\epsilon))^{-1}$.

Proof. Without loss of generality, the norm of each x_i is at most 1 and

$$\mathsf{E}\left\|\sum_{i=1}^{n} \epsilon_i x_i\right\| \geqslant C^{-1} n^{1/p} .$$

If p' is chosen so that $n^{1/p'} = C^{-1} n^{1/p}$, then $n^{2/p'} = C^{-2} n^{2/p}$. Corollary 3.3 is thus an immediate consequence of Theorem 3.1. □

It turns out, however, that one can do much better than Corollary 3.3 when $p > 1$. In this chapter, we have used the fact that, for an appropriate distance on Ω, the function $\|a_{\epsilon,\pi}\|$ is a Lipschitz function on Ω. If the norm concerned is equivalent to the ℓ_p-norm, there are stonger continuity properties available. We shall show how to use these in the Chapter 4.

As our final result in this chapter, we show that if the original sequence in Theorem 3.1 is contained in a space with a type condition, then the almost symmetric block basis obtained satisfies a stronger growth condition.

Corollary 3·4. *Let $1 \leqslant p < 2$, $0 < \epsilon \leqslant 1$, and let $(x_i)_1^n$ be a sequence of unit vectors in a Banach space X. Suppose that $\mathsf{E}\,\|\sum_1^n \epsilon_i x_i\| \geqslant n^{1/p}$ and suppose that the p-type constant of X is C. Then there is a block basis $(\mathbf{u}_i)_1^m$ of $(x_i)_1^n$ with blocks of ± 1 coefficients and equal length, which is $(1+\epsilon)$-symmetric and has cardinality*

$$m \geqslant \gamma(\epsilon)(\log n)^{-1} n^{\frac{2}{p}-1}$$

where $\gamma(\epsilon) = (\epsilon^3/3,000,000)(\log(33/\epsilon))^{-1}$. Moreover, if $h = n/m$, $1 \leqslant k \leqslant m$ and $A \subset [m]$ is any subset of $[m]$ of cardinality k, then

$$(10C)^{-1}(hk)^{1/p} \leqslant \Big\| \sum_{i \in A} \mathbf{u}_i \Big\| \leqslant 2C(hk)^{1/p} .$$

Proof. The first part of the result is just a restatement of Theorem 3.1, so all we need to do is prove the second. For the right-hand inequality, observe that, with the help of Lemma 3.2(ii), we have proved the existence of a block basis such that the norm of any vector is not merely invariant under changes of sign and order of its coefficients, but is also close to its expected norm. Now, if the block basis $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is chosen randomly as in the proof of Theorem 3.1, then it is clear that

$$\mathsf{E}\Big\| \sum_{i \in A} \mathbf{u}_i \Big\| \leqslant C(hk)^{1/p}$$

because of the type condition on X. The inequality now follows from the fact that $C > 1$ and $\epsilon \leqslant 1$.

To obtain the second inequality, let us set $l = \lfloor n/k \rfloor$ and partition $[n]$ into l sets $A_1, \ldots, A_l$ of size k and one set A_{l+1} of size less than k. Then it follows easily from the type condition on X and the fact that $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is a 2-symmetric block basis that

$$\begin{aligned} \Big\| \sum_{i=1}^{m} \mathbf{u}_i \Big\| &\leqslant 2C \Big(\sum_{j=1}^{l+1} \Big\| \sum_{i \in A_j} \mathbf{u}_i \Big\|^p \Big)^{1/p} \\ &\leqslant 4C(l+1)^{1/p} \Big\| \sum_{i \in A} \mathbf{u}_i \Big\| . \end{aligned}$$

The result follows now from the growth assumption on the original basis. $\square$

CHAPTER 4

BASES EQUIVALENT TO THE STANDARD BASIS OF ℓ_p^n

In the last chapter, we showed that a normalized basis for which the average of all ± 1-sums is at least $n^{1/p}$ must have an almost symmetric block basis of cardinality proportional to $n^{2/p-1}/\log n$. In this chapter, we shall show that, if we assume the stronger condition that the basis is equivalent to the unit vector basis of an ℓ_p-space, for some $p < \infty$, then we get a much stronger result. In fact, there is an almost symmetric block basis of cardinality proportional to $n/\log n$. Note that, in the case $p = 1$, we have already proved this result under a weaker assumption. The extra strength when $p > 1$ comes from the fact that we can define a natural metric on the set of rearrangements of a vector in ℓ_p^n in such a way that any equivalent norm on ℓ_p^n is Hölder continuous with respect to this metric, with exponent $1/p < 1$. In the previous chapter, we only used a Lipschitz condition. Let us now state our main result.

Theorem 4·1. *Let $1 \leqslant p < \infty$, $0 < \epsilon < 1/2$, $C > 1$ and let $(x_i)_1^n$ be a basis for a normed space X. Suppose that, for any $\mathbf{a} = \sum_1^n a_i \mathbf{e}_i \in \mathbb{R}^n$,*

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

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

$$m = (1/64)(\epsilon/66C)^{2p} \cdot (\epsilon/20) \cdot (\log(33/\epsilon))^{-1} \, n/\log n.$$

This improves on the known bound of $\alpha(\epsilon, p, C)n^{1/3}$ due to Amir and Milman [4]. The result is useful as the first step in the proof of a local version of Krivine's theorem.

As in the last chapter, we must begin with some notation. This time it is a little more complicated. Let Ψ and Ω and their actions on $\mathbb{R}^m$ and X respectively be defined as in the last chapter. Relabelling the indices of $(x_i)_1^n$ as before, let us also define an action of Ψ on Ω as follows. Given $(\eta, \sigma) \in \Psi$, let it send (ϵ, π) to $\psi_{\eta,\sigma}((\epsilon, \pi)) = (\epsilon', \pi')$, where

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

The proof of Theorem 4.1 will be similar to that of Theorem 3.1, but it is not possible to work directly with the norm. Instead, for a fixed $\mathbf{a} \in \ell_p^m$, such that $\|\mathbf{a}\|_p = 1$ and $a_1 \geqslant \ldots \geqslant a_m \geqslant 0$, we define, for each $(\eta, \sigma) \in \Psi$, a function $g_{\eta,\sigma}: \Omega \longrightarrow \mathbb{R}$ as follows.

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

$$\left\{(\epsilon, \pi): \|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\| \leqslant \mathsf{M}\, \|\phi_{\epsilon',\pi'}(\mathbf{a}_{\eta,\sigma})\|\right\}.$$

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

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

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

Then

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

Thus $g_{\eta,\sigma}$ measures how far (ϵ, π) is from some (ϵ', π') for which $\|\phi_{\epsilon',\pi'}(\mathbf{a}_{\eta,\sigma})\|$ 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

$$\begin{aligned} d_{\eta,\sigma}\big((\epsilon, \pi), (\epsilon', \pi')\big) &= \sum_{i=1}^{m} |a_i|^p \big|\{j: \epsilon_{\sigma(i)\,j} \neq \epsilon'_{\sigma(i)\,j} \text{ or } \pi_{\sigma(i)\,j} \neq \pi'_{\sigma(i)\,j}\}\big| \\ &= d\big(\psi_{\eta,\sigma}(\epsilon, \pi), \psi_{\eta,\sigma}(\epsilon', \pi')\big). \end{aligned}$$

Also

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

and thus

$$\begin{aligned}\Gamma_{\eta,\sigma} &= \left\{(\epsilon,\pi) \colon \left\|\phi_{\psi_{\eta,\sigma}(\epsilon,\pi)}(\mathbf{a})\right\| \leqslant \mathsf{M}\left\|\phi_{\epsilon',\pi'}(\mathbf{a})\right\|\right\} \\ &= \left\{\psi_{\eta,\sigma}^{-1}(\epsilon,\pi) \colon \left\|\phi_{\epsilon,\pi}(\mathbf{a})\right\| \leqslant \mathsf{M}\left\|\phi_{\epsilon',\pi'}(\mathbf{a})\right\|\right\} \\ &= \psi_{\eta,\sigma}^{-1}(\Gamma).\end{aligned}$$

Hence

$$\begin{aligned}g_{\eta,\sigma}(\epsilon,\pi) = d_{\eta,\sigma}\big((\epsilon,\pi),\Gamma_{\eta,\sigma}\big) &= d\big(\psi_{\eta,\sigma}(\epsilon,\pi),\psi_{\eta,\sigma}(\Gamma_{\eta,\sigma})\big) \\ &= d\big(\psi_{\eta,\sigma}(\epsilon,\pi),\Gamma\big) = g\big(\psi_{\eta,\sigma}(\epsilon,\pi)\big)\ .\end{aligned}$$

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

$$\begin{aligned}&\Big|\left\|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\right\| - \left\|\phi_{\epsilon',\pi'}(\mathbf{a}_{\eta,\sigma})\right\|\Big| \\ &\qquad\leqslant \left\|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma}) - \phi_{\epsilon',\pi'}(\mathbf{a}_{\eta,\sigma})\right\| \\ &\qquad= \left\|\phi_{\psi_{\eta,\sigma}(\epsilon,\pi)}(\mathbf{a}) - \phi_{\psi_{\eta,\sigma}(\epsilon',\pi')}(\mathbf{a})\right\| \\ &\qquad\leqslant 2C\left(\sum_{i=1}^{m}|a_i|^p\left|\{j \colon \epsilon_{\sigma(i)\,j} \neq \epsilon'_{\sigma(i)\,j} \text{ or } \pi_{\sigma(i)\,j} \neq \pi'_{\sigma(i)\,j}\}\right|\right)^{1/p} \\ &\qquad= 2C\Big(d_{\eta,\sigma}\big((\epsilon,\pi),(\epsilon',\pi')\big)\Big)^{1/p}.\end{aligned}$$

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

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

We are now ready to state the main lemma on which the proof of Theorem 4.1 is based. It corresponds to the second part of Lemma 3.2. We shall also need Proposition 2.1, as we did in the previous chapter.

Lemma 4.2. *Let ϵ, m and C be as in the statement of Theorem 4.1, let $\delta = \epsilon/11$ and let $h = n/m$. Let $\mathbf{a} \in \mathbf{R}^m$ be a given vector with $\|\mathbf{a}\|_p = 1$, and for $(\eta, \sigma) \in \Psi$ let $g_{\eta,\sigma}: \Omega \longrightarrow \mathbf{R}^+$ be as defined above. Then*

$$\mathbf{P}_\Omega\left[\exists(\eta,\sigma) \text{ s.t. } g_{\eta,\sigma}((\epsilon,\pi)) > (\delta/6C)^p h\right] < (1/2)m^{-\delta^{-1}\log(3\delta^{-1})}.$$

Let us first see how we may deduce Theorem 4.1 essentially from Proposition 2.1 and Lemma 4.2. Suppose both of these are true, and pick vectors $\mathbf{a}_1, \ldots, \mathbf{a}_N$ which satisfy the conclusion of Proposition 2.1. Now apply Lemma 4.2 to each of the vectors $\mathbf{a}_i$. Since $N = m^{\delta^{-1}\log(3\delta^{-1})}$, we have, with probability greater than 1/2 if we pick (ϵ, π) at random from Ω, that $g_{\eta,\sigma}((\epsilon,\pi)) \leqslant (\delta/6C)^p h$ for every $(\eta,\sigma) \in \Psi$ and every $1 \leqslant i \leqslant N$. (The dependence on i is via $g_{\eta,\sigma}$, which is defined in terms of $\mathbf{a}_i$.) By the remarks preceding the statement of Lemma 4.2, we have, with probability 1/2, that

$$\|\phi_{\epsilon,\pi}((\mathbf{a}_i)_{\eta,\sigma})\| - \mathbf{M}\,\|\phi_{\epsilon',\pi'}((\mathbf{a}_i)_{\eta,\sigma})\| < (\delta/3)h^{1/p}$$

for every $i \in [N]$ and $(\eta,\sigma) \in \Psi$.

Lemma 4.2 is not quite sufficient for our purposes because it only gives a bound for the probability of some $\|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|$ being significantly *larger* than its median. However, one can prove a bound in the other direction with an almost identical proof. One obtains that, with probability greater than 1/2,

$$\|\phi_{\epsilon,\pi}((\mathbf{a}_i)_{\eta,\sigma})\| - \mathbf{M}\,\|\phi_{\epsilon',\pi'}((\mathbf{a}_i)_{\eta,\sigma})\| > -(\delta/3)h^{1/p}$$

for every $i \in [N]$ and $(\eta,\sigma) \in \Psi$.

Now the median above is certainly at least $h^{1/p}$, and, since $\delta < 1/11$, we also have $(1+\delta/3)(1-\delta/3)^{-1} \leqslant 1+\delta$. With these two facts and the estimates above, we obtain that there exists $(\epsilon,\pi) \in \Omega$ such that

$$\max_{\Psi} \|\phi_{\epsilon,\pi}((\mathbf{a}_i)_{\eta,\sigma})\| \Big/ \min_{\Psi} \|\phi_{\epsilon,\pi}((\mathbf{a}_i)_{\eta,\sigma})\| \leqslant 1+\delta$$

for every $i \in [N]$.

Hence, if we define a norm $\|.\|'$ on $\mathbb{R}^m$ by $\|\mathbf{a}\|' = \|\phi_{\epsilon,\pi}(\mathbf{a})\|$, we find that this norm is $(1+\delta)$-symmetric at $\mathbf{a}_1, \ldots, \mathbf{a}_N$. It follows from Proposition 2.1 that it is $(1+\epsilon)$-symmetric. This tells us that $\phi_{\epsilon,\pi}(\mathbf{e}_1), \ldots, \phi_{\epsilon,\pi}(\mathbf{e}_m)$ is a $(1+\epsilon)$-symmetric block basis of $x_1, \ldots, x_n$.

It therefore remains for us to prove Lemma 4.2. Before doing this, we shall prove one other lemma, which we shall then use in the proof of Lemma 4.2. It is a technicality which is found in the proofs of many results of this kind.

Lemma 4·3. *Let δ, p, C, m, h, $\mathbf{a}$ and $g_{\eta,\sigma}$ be as in Lemma 4.2. Then*

$$\mathbb{E}_\Omega g_{\eta,\sigma}((\epsilon,\pi)) < 1/4(\delta/6C)^p h.$$

Proof. This lemma states that the expectation and median of $\|\phi_{\epsilon,\pi}(\mathbf{a}_{\eta,\sigma})\|$ are close. It follows easily from Proposition 2.9. Indeed, let us set Φ to be Ω, and define our weights by $b_i = |a_{\lceil (i/h)\rceil}|^p$ for each $i \in [n]$. Then the function $f = g_{\eta,\sigma}$ is 1-Lipschitz. Applying Proposition 2.9 with $s = n$ and $t = 0$ we obtain

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

Hence

$$\begin{aligned}\mathbb{P}\left[f = 0\right] &< \exp\left(-\frac{(\mathbb{E}f)^2}{8h}\right)\\ \Rightarrow \frac{1}{2} &< \exp\left(-\frac{(\mathbb{E}f)^2}{8h}\right)\\ \Rightarrow \mathbb{E}f &< (8h\log 2)^{1/2}\\ &< (1/4)\cdot(\delta/6C)^p h.\end{aligned}$$

□

Proof of Lemma 4·2. The proof is similar to that of Lemma 3.2 (ii). We fix a vector $\mathbf{a}$ with $\|\mathbf{a}\|_p = 1$ and $a_1 \geqslant \ldots \geqslant a_n \geqslant 0$, and define

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

where $k = \log_2(4 \cdot (6C/\delta)^p m)$.

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

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

$$f^r_{\eta,\sigma}((\epsilon,\pi)) = \mathsf{E}_\Omega\left[g_{\eta,\sigma}((\epsilon',\pi')) \;\middle|\; \epsilon'_{ij} = \epsilon_{ij}, \pi'_{ij} = \pi_{ij}, \forall i \in \bigcup_{s=1}^{r} \sigma(B_s), j = 1,\ldots,h\right]$$

Note that if $\epsilon'_{ij} = \epsilon_{ij}$ and $\pi'_{ij} = \pi_{ij}$ for all $i \in \bigcup_{s=1}^r \sigma(B_s)$, $j = 1, \ldots, h$, then $\phi_{\epsilon,\pi}(\mathbf{b}^s_{\eta,\sigma}) = \phi_{\epsilon',\pi'}(\mathbf{b}^s_{\eta,\sigma})$ 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),

$$f^{k+1}_{\eta,\sigma}((\epsilon,\pi)) - f^k_{\eta,\sigma}((\epsilon,\pi)) \leqslant \frac{1}{4}\left(\frac{\delta}{6C}\right)^p h$$

We would like to prove two facts, which correspond to the estimates (a) and (b) in the last chapter. 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, $1 \leqslant r \leqslant k$ and $\delta_r > 0$,

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

(where $\beta_r = |B_r|$ as before). Let us set $C_r = \bigcup_{s=1}^r B_s$ for each $1 \leqslant r \leqslant k$.

To prove (a)′, we use the fact that $g_{\eta,\sigma}((\epsilon',\pi')) = g(\psi_{\eta,\sigma}(\epsilon',\pi'))$, and hence

$$\begin{aligned} f^r_{\eta,\sigma}((\epsilon,\pi)) &= \\ &= \mathsf{E}_\Omega\left[g((\epsilon',\pi')) \;\middle|\; \eta_i\epsilon'_{\sigma^{-1}(i)\,j} = \epsilon_{ij}, \pi'_{\sigma^{-1}(i)\,j} = \pi_{ij}\ \forall i \in \sigma(C_r), j = 1,\ldots,h\right] \\ &= \mathsf{E}_\Omega\left[g((\epsilon',\pi')) \;\middle|\; \epsilon'_{ij} = \eta_i\epsilon_{\sigma(i)\,j}, \pi'_{ij} = \pi_{\sigma(i)\,j}\ \forall i \in C_r, j = 1,\ldots,h\right]. \end{aligned}$$

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

To prove (b)$'$, we once again apply Proposition 2.9. Set $b_i = a^p_{\lceil(i/h)\rceil}$ for $1 \leqslant i \leqslant n$, set $s = \gamma_r h$, $t = \gamma_{r-1} h$ and $f((\epsilon,\pi)) = g_{\eta,\sigma}((\epsilon,\pi))$. Then f is 1-Lipschitz, $f_s = f^r_{\eta,\sigma}$ and $f_t = f^{r-1}_{\eta,\sigma}$. We obtain (b)$'$ immediately.

Again, just as before, we may conclude that

$$\mathbb{P}_\Omega\left[\exists(\eta,\sigma) \in \Psi \text{ s.t. } f^k_{\eta,\sigma}((\epsilon,\pi)) - f^0_{\eta,\sigma}((\epsilon,\pi)) > \frac{1}{2}\Big(\frac{\delta}{6C}\Big)^p h\right] \leqslant \sum_{r=1}^{k}(2m)^{\gamma_r}\exp\Big(-\frac{2^{2(r-1)}\delta_r^2 h}{8\beta_r}\Big) \qquad (1)$$

where $\delta_1, \ldots, \delta_r$ is any sequence satisfying $\sum_1^k \delta_r \leqslant (1/2)\cdot(\delta/6C)^p$. We shall choose δ_r to be $(1/4)\cdot 2^{-r}\beta_r^{1/2}\gamma_r^{1/2}\cdot(\delta/6C)^p$.. Then, just as before, $\sum_1^k \delta_r$ is indeed at most $(1/2)\cdot(\delta/6C)^p$, and the right hand side of (1) is at most

$$k\exp\left(\log(2m) - \Big(\frac{\delta}{6C}\Big)^{2p}\frac{h}{64}\right).$$

But if

$$f^k_{\eta,\sigma}((\epsilon,\pi)) - f^0_{\eta,\sigma}((\epsilon,\pi)) \leqslant \frac{1}{2}\Big(\frac{\delta}{6C}\Big)^p h$$

and, by Lemma 4.3,

$$f^0_{\eta,\sigma}((\epsilon,\pi)) < \frac{1}{4}\Big(\frac{\delta}{6C}\Big)^p h$$

and, as we remarked earlier,

$$f^{k+1}_{\eta,\sigma}((\epsilon,\pi)) - f^k_{\eta,\sigma}((\epsilon,\pi)) \leqslant \frac{1}{4}\left(\frac{\delta}{6C}\right)^p h,$$

then

$$f^{k+1}_{\eta,\sigma}((\epsilon,\pi)) \leqslant \Big(\frac{\delta}{6C}\Big)^p h.$$

Hence

$$\mathbb{P}_\Omega\left[\exists(\eta,\sigma) \text{ s.t. } g_{\eta,\sigma}((\epsilon,\pi)) > (\delta/6C)^p h\right] \leqslant k\exp\left(\log 2m - \Big(\frac{\delta}{6C}\Big)^{2p}\frac{h}{64}\right).$$

But since $h > 64 \cdot (6C/\delta)^{2p} \cdot (20/\epsilon) \cdot \log(33/\epsilon)\log n$, this is at most

$$k\exp\Big(\log n\Big(1-(20/\epsilon)\log(33/\epsilon)\Big)\Big)$$
$$\leqslant n^{-(11/\epsilon)\log(33/\epsilon)}$$
$$< (1/2)m^{-(11/\epsilon)\log(33/\epsilon)} = (1/2)m^{-\delta^{-1}\log(3\delta^{-1})}.$$

This proves Lemma 4.2, and hence Theorem 4.1. □

One result of Amir and Milman [3] gives an estimate for the size of a $(1+\epsilon)$-symmetric block basis when the original basis not only has large average growth, but is the basis of a space with p-type constant γ, and has an average growth that is within a constant of being the greatest possible consistent with this type condition. As it happens, they use the type condition to find a subsequence of the original sequence which satisfies an average lower p-estimate. That is, if $x_1, \dots, x_n$ is the original sequence, they find a subsequence $y_1, \dots, y_m$ and a constant c such that $\mathsf{E}\left\|\sum_1^k \epsilon_i y_i\right\| \geqslant cn^{1/p}$ for every $k \leqslant m$. This, as well as enabling them to improve their bound over the one they obtain without a type-condition, yields a growth condition on their symmetric block basis (cf. Corollary 3.4). In Chapter 5, we shall exhibit, for $1 \leqslant p \leqslant 3/2$, a sequence which satisfies a lower p-estimate, but which does not have an almost symmetric block sequence of cardinality significantly greater than that guaranteed by Theorem 3.1. However, it turns out that a type condition on the space containing the original sequence *can* sometimes enable one to find larger almost symmetric block bases. This is a corollary of Theorem 4.1: if p is sufficiently close to 2 and the p-type constant of a space is sufficiently close to 1, then one can obtain information about the distance of that space from a Hilbert space of the same dimension. We shall do this, and then apply Theorem 4.1.

First, let us consider the case $p = 2$. Here, any result at all that gives a power of n is interesting, since no such bound follows from Theorem 3.1 or from results of Amir and Milman. We shall obtain a block sequence that is $(1+\epsilon)$-symmetric and whose cardinality is a power of n that depends on the precise type constant, tending to $1/2$ as the constant tends to 1. We shall of course insist that the original basis is independent, since the 1-dimensional Banach space has 2-type constant 1. This in fact shows that a purely probabilistic method cannot possibly work. In the next simple lemma, we show how to replace the original basis by a block basis that looks much nicer. Lemma 4.5 shows that a space with a good 2-type constant cannot have too large a 2-cotype constant. The proof of Lemma 4.5 uses a fairly standard technique.

Lemma 4·4. *Let $x_1, \ldots, x_n$ be a linearly independent sequence of vectors in ℓ_2. Then there is an orthonormal block sequence $y_1, \ldots, y_m$ of $x_1, \ldots, x_n$ of cardinality $m = \sqrt{2n}$.*

Proof. Set $y_1 = x_1/\|x_1\|$. Now there is at least one unit vector in the space spanned by x_2 and x_3 which is orthogonal to y_1. Let y_2 be such a vector. There is at least one unit vector in the space spanned by x_4, x_5 and x_6 which is orthogonal to both y_1 and y_2. Let y_3 be such a vector. Continuing this process gives an orthonormal block basis of cardinality m, where m is the greatest integer such that $\frac{1}{2}m(m-1) \leqslant n$. In particular, $m \geqslant \sqrt{2n}$. □

Lemma 4·5. *There exists an absolute constant γ with the following property. Let $0 \leqslant \epsilon \leqslant 1$ and let X be any n-dimensional Banach space satisfying $T_2(X) \leqslant 1+\epsilon$. Then $C_2(X) \leqslant \gamma n^{\log(1+\epsilon)/\log 2} \log n \leqslant \gamma n^{\epsilon/\log 2} \log n$.*

Proof. In fact, we use a rather weaker hypothesis. For any $x_1, x_2 \in X$ we have

$$\frac{1}{2}(\|x_1+x_2\|^2 + \|x_1-x_2\|^2) \leqslant (1+\epsilon)^2(\|x_1\|^2 + \|x_2\|^2)$$

By a simple substitution this implies that, for any $y_1, y_2 \in X$

$$\frac{1}{2}(\|y_1+y_2\|^2 + \|y_1-y_2\|^2) \geqslant (1+\epsilon)^{-2}(\|y_1\|^2 + \|y_2\|^2) \qquad (*)$$

König and Tzafriri [29] have shown that there is an absolute constant γ_0 such that for any n-dimensional space Y,

$$C_2(Y) \leqslant \gamma_0 C_2(Y, n)(\log n)^{1/2}.$$

Let $N = 2^k$ be the smallest power of 2 that is greater than n, so $k \leqslant 1 + \log n/\log 2$, and suppose we have a sequence of vectors $x_1, \dots, x_N \in X$. By repeated applications of $(*)$ we find that

$$\begin{aligned} \mathsf{E}\Big\|\sum_1^N \epsilon_i x_i\Big\|^2 &\geqslant (1+\epsilon)^{-2}\Big(\mathsf{E}\Big\|\sum_1^{N/2} \epsilon_i x_i\Big\|^2 + \mathsf{E}\Big\|\sum_{N/2+1}^{N} \epsilon_i x_i\Big\|^2\Big) \\ &\geqslant \dots \\ &\geqslant (1+\epsilon)^{-2k}\sum_1^N \|x_i\|^2 \end{aligned}$$

We therefore have $C_2(X,n) \leqslant C_2(X,N) \leqslant (1+\epsilon)^k \leqslant (1+\epsilon)\cdot(1+\epsilon)^{\log n/\log 2} = (1+\epsilon)n^{\log(1+\epsilon)/\log 2} \leqslant 2n^{\epsilon/\log 2}$. Thus, if we set $\gamma = 2\gamma_0$, we have that $C_2(X) \leqslant \gamma n^{\epsilon/\log 2}(\log n)^{1/2}$. □

Theorem 4·6. *Suppose $x_1, \dots, x_n$ is a linearly independent sequence of vectors spanning a Banach space X with 2-type constant $1+\epsilon$. Then there is a $(1+\epsilon)$-symmetric block sequence $y_1, \dots, y_m$ of $x_1, \dots, x_n$ of cardinality at least $m = \alpha n^{f(\epsilon)}/(\log n)^5$, where $f(\epsilon) = \frac{1}{2} - 4\epsilon/\log 2$ and $\alpha = \alpha(\epsilon)$.*

Proof. By Lemma 4.5, X has 2-cotype constant at most $C' = \gamma n^{\epsilon/\log 2}\log n$. A well known result of Kwapien [31] states that, for any k-dimensional Banach space Y,

$$d(Y, \ell_2^k) \leqslant T_2(Y)C_2(Y).$$

Hence, X is actually $C = (1+\epsilon)C'$-equivalent to ℓ_2^n. Let $T: X \to \ell_2^n$ be a map such that $\|T\| = 1$ and $\|T^{-1}\| \leqslant C$. By Lemma 4.4 we can pick a block sequence $z_1, \dots, z_m$ with $m = \sqrt{2n}$ such that $Tz_1, \dots, Tz_m$ is an orthonormal sequence.

This implies that $z_1, \ldots, z_m$ is C-equivalent to the unit vector basis of ℓ_2^m. Hence, by Theorem 4.1 with $p = 2$ and C as above, we can find a $(1+\epsilon)$-symmetric block basis of $z_1, \ldots, z_m$ of cardinality at least $\alpha(\epsilon)C^{-4}m/\log m$, that is, at least $\alpha(\epsilon)n^{f(\epsilon)}/(\log n)^5$. □

We shall now look at what happens when $p < 2$. It turns out that the proof of Theorem 4.6 can be adapted easily to give a result here, provided p is sufficiently close to 2 and C is sufficiently close to 1. Let X be an n-dimensional space with p-type constant $C = T_p(X)$. Then we have

$$\begin{aligned}\frac{1}{2}(\|x_1 + x_2\|^2 + \|x_1 - x_2\|^2) &\leqslant C^2(\|x_1\|^p + \|x_2\|^p)^{2/p} \\ &\leqslant 2^{2/p-1}C^2(\|x_1\|^2 + \|x_2\|^2)\end{aligned}$$

Hence, for any $y_1, y_2 \in X$, we have

$$\frac{1}{2}(\|y_1 + y_2\|^2 + \|y_1 - y_2\|^2) \geqslant 2^{2-2/p}C^{-2}(\|y_1\|^2 + \|y_2\|^2) \qquad (**)$$

Theorem 4·7. *Let $0 < \epsilon \leqslant 1/2$, $1 \leqslant p < 2$, $C \geqslant 1$ and let $x_1, \ldots, x_n$ be any sequence of vectors spanning an n-dimensional Banach space X which satisfies $T_p(X) \leqslant C$. Then $(x_i)_1^n$ has a $(1+\epsilon)$-symmetric block basis of cardinality at least $\alpha(\epsilon)n^{f(C,p)}/(\log n)^5$ where $f(C,p) = (1/2) - 4((\log C/\log 2) + (2/p - 1))$.*

Proof. By using the inequality $(**)$ in place of $(*)$ in the proof of Lemma 4.5 we obtain that $C_2(X) \leqslant \gamma n^{\log(2^{1/p-1/2}C)/\log 2}\log n$. Now a result of Tomczak-Jaegermann [48] states that for any n-dimensional space X, and any $p \leqslant 2$ and $q \geqslant 2$,

$$d(X, \ell_2^n) \leqslant 4\alpha_p(X)\beta_q(X).$$

Moreover, it is well known that $\alpha_p(X)$ and $T_p(X)$ are the same, up to a constant, and also that $\beta_p(X) \leqslant C_p(X)$. So in this case we have

$$\begin{aligned}d(X, \ell_2^n) &\leqslant C\gamma n^{\log(2^{1/p-1/2}C)/\log 2 + 1/p - 1/2}\log n \\ &= C\gamma n^{(\log C/\log 2) + (2/p-1)}\log n.\end{aligned}$$

Just as in the proof of Theorem 4.6, we may now apply Lemma 4.4 and Theorem 4.1 to obtain that if $x_1, \dots, x_n$ is any linearly independent sequence of vectors in a space X with $T_p(X) = C$, then it has a $(1+\epsilon)$-symmetric block basis of cardinality at least $\alpha(\epsilon)n^{f(C,p)}/(\log n)^5$, as stated. This improves on the bound of $n^{2/p-1}$ obtained in Chapter 3 if $5(2/p-1) + 4\log C/\log 2 < 1/2$. □

There is in fact an elementary proof that if X is an n-dimensional space, then $C_2(X) \leqslant \gamma C_2(X,n)\log n$, using well-known results about the duality between type and cotype and a result of Tomczak-Jaegermann about the relationship between $\alpha_p(X)$ and $\alpha_p(X,n)$. It is known (cf. [39] Chapters 9 and 14) that for any n-dimensional normed space X, $C_2(X,n) \geqslant (\log n)^{-1}T_2(X^*,n)$, and Tomczak-Jaegermann's result [48] states that $\alpha_p(X,n)$ and $\alpha_p(X)$ are the same, to within an absolute constant. We obtain

$$\begin{aligned} C_2(X,n) &\geqslant (\log\, n)^{-1}T_2(X^*,n) \geqslant (\log\, n)^{-1}\alpha_2(X^*,n) \\ &\geqslant \gamma_1(\log\, n)^{-1}\alpha_2(X^*) \geqslant \gamma_2(\log\, n)^{-1}T_2(X^*) \\ &\geqslant \gamma_3(\log\, n)^{-1}C_2(X) \end{aligned}$$

where the remaining steps are standard (see e.g. [39 Chapter 9]).

It seems likely that if the sequence in Theorem 4.7 satisfies $\mathbb{E}\,\|\sum_1^n \epsilon_i x_i\| \geqslant cn^{1/p}$ for some constant c, then Lemma 4.4 is too weak. However, we have not found a way of exploiting the growth in this case to obtain a better bound.

There are various other questions suggested by the results of this chapter. One is to what extent Theorem 4.1 is best possible. There is a strong analogy between Theorem 4.1 and results of Figiel, Lindenstrauss and Milman [18] connected with Dvoretzky's theorem. One of their results (cf. also [39]) is that if X is an n-dimensional normed space that is C-equivalent to ℓ_2, then it has a subspace of dimension $k = \alpha(\epsilon, C)n$ that is $(1+\epsilon)$-equivalent to ℓ_2^k. One can restate this as follows. Suppose $\|.\|$ is a norm on $\mathbb{R}^n$ such that $\|x\| \leqslant C\,\|Ax\|$ for any $x \in R^n$

and $A \in O(n)$. Then there is a k-dimensional subspace $Y \subset R^n$ such that for any $y \in Y$ and any $A \in O(n)$ that leaves Y invariant, $\|y\| \leqslant (1+\epsilon)\|Ay\|$. Theorem 4.1 is very similar, but is concerned with the symmetry group $\{-1,1\}^n \times S_n$ instead of $O(n)$. This suggests two possibilities. One is that one might be able to obtain a bound in Theorem 4.1 that is actually proportional to n. The other is that one might be able to relax the condition on the original basis from being C-equivalent to the unit vector basis of ℓ_p^n to being C-symmetric.

It turns out that neither of these works. In Chapter 6, we shall show that Theorem 4.1 as stated is best possible, up to a factor of $\log\log n$, and in Chapter 7, we shall exhibit a basis that is C-equivalent to the unit vector basis of ℓ_∞^n which has no almost symmetric block basis of cardinality $n^{\beta(\epsilon,C)}$, where $\beta(\epsilon, C)$ depends on ϵ and C only, and tends to zero as $\log(1+\epsilon)/\log C$ tends to zero.

PART III

THE SIZE OF SYMMETRIC BLOCK BASES:

UPPER BOUNDS

CHAPTER 5

BASES WITH LARGE AVERAGE GROWTH

In Chapter 3 we showed that any sequence $x_1, \ldots, x_n$ of unit vectors in a normed space which satisfies the condition that $\mathsf{E}\,\|\sum_{i=1}^n \pm x_i\|$ is at least $n^{1/p}$ must have an almost symmetric block basis of cardinality at least $\gamma(\epsilon)n^{2/p-1}/\log n$. Discussion of an upper bound for the size of an almost symmetric block basis that is guaranteed by these conditions is somewhat complicated by the fact that there can often be block bases which are almost symmetric but which do not satisfy certain other very natural conditions. The following example shows that in one sense the result of Chapter 3 is best possible.

Let $1 \leqslant p < 2$, $1/q + 1/p = 1$, $m = 2n^{2/p-1}$ and $h = n^{2/q}/2$. Then let $\mathbf{e}_1, \ldots, \mathbf{e}_m$ be the standard basis of ℓ_1^m. For $1 \leqslant i \leqslant n$ we set $x_i = \mathbf{e}_{k(i)}$, where $k(i) = \lceil i/h \rceil$. Then

$$\begin{aligned}\mathsf{E}\left\|\sum_1^n \epsilon_i x_i\right\| &= \mathsf{E}\left(\sum_{i=1}^m \left|\sum_{j=1}^h \epsilon_{(h-1)i+j}\right|\right)\\ &= m\,\mathsf{E}\left|\sum_{j=1}^h \epsilon_j\right|\\ &\geqslant 2^{-1/2} m h^{1/2} = n^{1/p}\end{aligned}$$

It is immediate by linear algebra that the sequence $(x_i)_1^n$ has no almost symmetric block basis of cardinality greater than m.

Unfortunately, however, this sequence is not linearly independent. The natural way of dealing with this ought to be simply to embed ℓ_1^m into any Banach space of dimension n and perturb the sequence very slightly, but this does not necessarily work. Suppose, for example, that we were to embed ℓ_1^m into ℓ_1^n in the natural way and then perturb the unit vector basis of ℓ_1^m by adding multiples of vectors from the unit vector basis of ℓ_1^n. We would then be able to construct a proportional-sized block basis that was 1-symmetric by taking blocks of the form $x_i - x_{i+1}$ (where i should not be a multiple of h) and normalizing. Although this is a contrived example, it illustrates the point that to find a linearly independent sequence with average growth as given in the main theorem, but with no almost symmetric block sequence, either one must use a completely different example, or one must solve the rather different problem of constructing some independent sequence with more or less no almost symmetric block sequence of any size. As p approaches 2, these become the same problem, which we shall discuss later in Chapter 7.

The 1-symmetric block basis given in the example above is, however, an uninteresting one. The reason is that the normalization can involve multiplying by an arbitrarily large number, so that the block basis bears very little relation to the original basis. If we pick any real number M and stipulate first that the coefficients appearing in the vectors in the block basis should be bounded above by M, and second that the block basis should be normalized, then any sufficiently small perturbation of the linearly dependent sequence given above will clearly do. This is obviously a natural constraint to impose on the block basis we find.

We shall present a randomly constructed space which gives a bound close to that achieved in the other direction in Chapter 3, even without this extra condition, when $1 \leqslant p \leqslant 3/2$. In fact, our example satisfies a lower p-estimate. When $p > 3/2$ our construction fails because the basis it gives has a large block basis that is close to the unit vector basis of an appropriate ℓ_∞^k.

Theorem 5·1. *There exists an absolute constant C such that for any $1 \leqslant p \leqslant 3/2$ and any $n \in \mathbb{N}$, there is a norm $\|.\|$ on $\mathbb{R}^n$ satisfying the following three conditions:*

(i) the standard basis is normalized;

(ii) for any $\mathbf{a} \in \mathbb{R}^n$, $\|\mathbf{a}\| \geqslant \|\mathbf{a}\|_p$;

(iii) if $k \geqslant Cn^{2/p-1}(\log n)^{4/3}$ *then no block basis* $\mathbf{u}_1, \ldots, \mathbf{u}_k$ *of the standard basis is 2-symmetric.*

Let us begin by setting $m = (C/2)n^{2/p-1}(\log n)^{1/3}$, $h = (2/C)n^{2/q}(\log n)^{-4/3}$, $N = 600h^{2/p}\log n$ and $M = (5n^2)^{mh}$. We shall pick N functionals $f_1, \ldots, f_N$ independently at random from the set of all functionals on $\mathbb{R}^n$ with ± 1-coordinates, where this set is endowed with the uniform probability measure. Given a vector $\mathbf{a} \in \mathbb{R}^n$, we define a random norm by setting

$$\|\mathbf{a}\| = \max\{\|\mathbf{a}\|_p, |f_1(\mathbf{a})|, \ldots, |f_N(\mathbf{a})|\} .$$

Our aim is to show that there exists a choice of $f_1, \ldots, f_N$ such that no block basis of the standard basis of $\mathbb{R}^n$ is 2-symmetric under the corresponding norm. In fact, it is easy to deduce from our proof that, with high probability, a random choice will do. The other two conditions of Theorem 5.1 are trivially satisfied by any norm of this kind.

As in the previous two chapters, we shall make one or two definitions and then state two lemmas which will be enough to prove our main theorem.

Suppose we have a block basis $\mathbf{u}_1, \ldots, \mathbf{u}_m$ of the standard basis, and suppose that we can find some $1 \leqslant j \leqslant m$ and some sequence of signs $\eta_1, \ldots, \eta_m$ such that

$$\left\|\sum_1^m \eta_i \mathbf{u}_i\right\| \geqslant \max\left\{80 \log n \, \|\mathbf{u}_j\|_1 \, , 6\left\|\sum_1^n \eta_i \mathbf{u}_i\right\|_p\right\} \qquad (*)$$

~~is greater than $1 - M^{-1}$.~~ In this case we shall say that the block basis $(\mathbf{u}_i)_1^m$ *satisfies condition* $(*)$.

We shall say that it *satisfies condition* $(**)$ if there exist $j \in [m]$ and a sequence of signs $\eta_1, \ldots, \eta_m$ such that

$$\left\|\sum_1^m \eta_i \mathbf{u}_i\right\| \geqslant \max\left\{64 \log n \, \|\mathbf{u}_j\|_1 \, , 4\left\|\sum_1^n \eta_i \mathbf{u}_i\right\|_p\right\} \qquad (**)$$

The next definition enables us to restrict our attention to block bases that are slightly more regular than an arbitrary block basis. Suppose $k_0 \geqslant k$ and $\mathbf{v}_1, \ldots, \mathbf{v}_{k_0}$ is a block basis of the standard basis. Since the supports of the $\mathbf{v}_i$ are disjoint, we can certainly find $k/2$ of them such that each one is supported on at most h coordinates. Passing to a subsequence again, and relabelling, we can use the pigeonhole principle to find a subbasis $\mathbf{u}_1, \ldots, \mathbf{u}_m$ and $\lambda > 0$ such that, for each $1 \leqslant i \leqslant m$, $\lambda \leqslant \|\mathbf{v}_i\|_p \leqslant 2\lambda$. By multiplying each element of the basis by λ^{-1} we obtain a block basis $\mathbf{u}_1, \ldots, \mathbf{u}_m$ such that, for each $1 \leqslant i \leqslant m$, $1 \leqslant \|\mathbf{u}_i\|_p \leqslant 2$. Let us call such a block basis *proper*. We have shown that any block basis of cardinality at least k has a subbasis, a multiple of which is proper. Note that if the proper block basis obtained this way fails to be 2-symmetric, then the original block basis also fails.

Our lemmas are as follows.

Lemma 5·2. *(i) Let m, M be as above and let $\mathbf{u}_1, \ldots, \mathbf{u}_m$ be any proper block basis. Then the probability that $\mathbf{u}_1, \ldots, \mathbf{u}_m$ satisfies condition $(*)$ is greater than $1 - M^{-1}$.*

(ii) Let $\mathbf{u}_1, \ldots, \mathbf{u}_m$ be any proper block basis which satisfies condition $(**)$. Then $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is not 2-symmetric.

Lemma 5·3. *There exist M proper block bases with the following property. If each one of these block bases satisfies condition $(*)$, then every proper block basis satisfies condition $(**)$.*

Lemmas 5.2 and 5.3 imply Theorem 5.1 as follows. Pick M proper block bases to satisfy the conclusion of Lemma 5.3. By Lemma 5.2(i), the probability that every one of them satisfies condition $(*)$ is greater than zero. In other words, for some choice of $f_1, \ldots, f_N$, each of these block bases satisfies condition $(*)$. By Lemma 5.3, it follows that every proper block basis satisfies condition $(**)$. Hence, by Lemma 5.2(ii) it follows that no proper block basis is 2-symmetric. However, by the remarks preceding the statement of Lemma 5.2, every block basis of cardinality at least k has a subbasis, a multiple of which is proper. Thus, with the choice of $f_1, \ldots, f_N$ above, no block basis of cardinality k or greater is 2-symmetric.

Proof of Lemma 5·2. (i) Let us fix a particular $j \in [N]$ and a particular proper block basis $\mathbf{u}_1, \ldots, \mathbf{u}_m$ and examine the effect of the functional f on the space generated by $\mathbf{u}_1, \ldots, \mathbf{u}_m$.

There certainly exist signs $\eta_1, \ldots, \eta_m$ such that $f\left(\sum_1^m \eta_i \mathbf{u}_i\right) = \sum_1^m |f(\mathbf{u}_i)|$. Further, it is well known (and a straightforward application of Theorem 1.1) that there exists an absolute constant c such that for each $\mathbf{u}_i$, we have

$$\mathsf{P}\left[|f(\mathbf{u}_i)| \geqslant c\,\|\mathbf{u}_i\|_2\right] \geqslant 1/2\ .$$

We can therefore define some independent random variables $\gamma_1, \ldots, \gamma_m$ in such a way that, for each i, γ_i is dominated by $|f(\mathbf{u}_i)|$, and takes the values $c\,\|\mathbf{u}_i\|_2$ or 0, each with probability 1/2. Then $\mathsf{E}\left(\sum_1^m \gamma_i\right) = \frac{c}{2}\sum_1^m \|\mathbf{u}_i\|_2$. By an easy appplication of Theorem 1.1 and the fact that each $\mathbf{u}_i$ is supported on at most h coordinates, we have

$$\begin{aligned}
\mathsf{P}\left[\sum_1^m \gamma_i < \frac{c}{4}\sum_1^m \|\mathbf{u}_i\|_2\right] &\leqslant \exp\left(-\frac{c^2}{16}\left(\sum_1^m \|\mathbf{u}_i\|_2\right)^2 \Big/ c\sum_1^m \|\mathbf{u}_i\|_2^2\right) \\
&= \exp\left(-\frac{c}{16}\left(\sum_1^m \|\mathbf{u}_i\|_2\right)^2 \Big/ \sum_1^m \|\mathbf{u}_i\|_2^2\right) \\
&\leqslant \exp\left(-\frac{c}{16}h^{1-2/p}\left(\sum_1^m \|\mathbf{u}_i\|_p\right)^2 \Big/ \sum_1^m \|\mathbf{u}_i\|_p^2\right) \\
&\leqslant \exp\left(-\frac{c}{16}h^{1-2/p} \cdot \frac{\lambda^2 m^2}{4\lambda^2 m}\right) \\
&= \exp\left(-\frac{c}{64}mh^{1-2/p}\right)\ .
\end{aligned}$$

Let us set $\theta = \exp\left(-\frac{c}{64} m h^{1-2/p}\right)$.

Now, since the support of each $\mathbf{u}_i$ is at most h, and since the $\|\mathbf{u}_i\|_p$ differ by a factor of at most 2, we have

$$\begin{aligned}
\sum_1^m \|\mathbf{u}_i\|_2 &\geqslant h^{1/2-1/p} \sum_1^m \|\mathbf{u}_i\|_p \\
&\geqslant \frac{1}{2} h^{1/2-1/p} m^{1-1/p} \left\| \sum_1^m \mathbf{u}_i \right\|_p \\
&= \frac{C^{1/2}}{2^{3/2}} (\log n)^{1/p+1/3} \left\| \sum_1^m \mathbf{u}_i \right\|_p .
\end{aligned}$$

Using these facts again, and also that $p \leqslant 3/2$, and picking the index j for which $\|\mathbf{u}_j\|_p$ is minimal, we have

$$\begin{aligned}
\sum_1^m \|\mathbf{u}_i\|_2 &\geqslant h^{1/2-1/p} \sum_1^m \|\mathbf{u}_i\|_p \\
&\geqslant m h^{1/2-1/p} \|\mathbf{u}_j\|_p \\
&\geqslant m h^{-1/2} \|\mathbf{u}_j\|_1 \\
&\geqslant (C/2)^{3/2} \log n \, \|\mathbf{u}_j\|_1 .
\end{aligned}$$

It follows that, if $(c/4) \cdot (C/2)^{3/2} > 80$, that is, if $C > 2 \cdot (320/c)^{2/3}$, then

$$\mathsf{P}\left[\sum_1^m |f(\mathbf{u}_i)| < \max\left\{ 80 \log n \, \|\mathbf{u}_j\|_1 , 6 \left\| \sum_1^m \mathbf{u}_i \right\|_p \right\} \right] \leqslant \theta .$$

Moreover, since the functionals $f_1, \ldots, f_N$ were chosen independently, the probability that, for every j and every choice of signs $\eta_1, \ldots, \eta_m$,

$$\left\| \sum_1^m \eta_i \mathbf{u}_i \right\| < \max\left\{ 80 \log n \, \|\mathbf{u}_j\|_1 , 6 \left\| \sum_1^m \mathbf{u}_i \right\|_p \right\}$$

is at most $\theta^N = \exp\left(-\frac{c}{64} m h^{1-2/p} N\right)$. In other words, we have shown that a given proper block basis satisfies condition $(*)$ with probability at least $1 - \theta^N$. It is easy to check that $\theta^N < M^{-1}$, which completes the proof of the first part of Lemma 5.1.

Proof of Lemma 5·2. (ii) We must show that a proper block basis $(\mathbf{u}_i)_1^m$ which satisfies condition (**) cannot be 2-symmetric. So let us suppose that it is 2-symmetric. Then for *all* choices of sign $(\eta_i)_1^m$ and any $j \in [m]$, we certainly have

$$\left\|\sum_1^m \eta_i \mathbf{u}_i\right\| \geqslant \max\left\{16 \log n \left\|\mathbf{u}_j\right\|_1, 2\left\|\sum_1^m \eta_i \mathbf{u}_i\right\|_p\right\}.$$

It follows that the support functional at each vector of the form $\sum_1^m \eta_i \mathbf{u}_i$ is one of $f_1, \ldots, f_N$. We shall show that, on the contrary, no given f_j can support as many as $N^{-1}2^m$ of these vectors. Let us fix $1 \leqslant j \leqslant N$ and set $f = f_j$. If the block basis $(\mathbf{u}_i)_1^m$ is 2-symmetric, then there exists some $\mu > 0$ such that for every choice of signs $(\eta_i)_1^m$,

$$\mu \left\|\mathbf{u}_1\right\| \leqslant \left\|\sum_1^m \eta_i \mathbf{u}_i\right\| \leqslant 2\mu \left\|\mathbf{u}_1\right\| ,$$

from which it follows that $\sum_1^m |f(\mathbf{u}_i)| \leqslant 2\mu \left\|\mathbf{u}_1\right\|$. We know also that $\mu \geqslant 16 \log n$ and that, for each i, $|f(\mathbf{u}_i)| \leqslant 2\left\|\mathbf{u}_1\right\|$. Let Ω be the probability space of all sequences of signs $(\eta_i)_1^m$ uniformly distributed. A final application of Theorem 1.1 gives

$$\begin{aligned}\mathsf{P}_\Omega\left[\left|f\left(\sum_1^m \eta_i \mathbf{u}_i\right)\right| \geqslant \mu \left\|\mathbf{u}_1\right\|\right] &\leqslant 2\exp\left(-\mu^2 \left\|\mathbf{u}_1\right\|^2 \Big/ 2\sum_1^m |f(\mathbf{u}_i)|^2\right) \\ &\leqslant 2\exp\left(-\frac{1}{4}\left(\sum_1^m |f(\mathbf{u}_i)|\right)^2 \Big/ 2\sum_1^m |f(\mathbf{u}_i)|^2\right)\end{aligned}$$

But if $\sum_1^m |f(\mathbf{u}_i)| \geqslant \mu \left\|\mathbf{u}_1\right\| \geqslant 16 \log n \left\|\mathbf{u}_1\right\|$, then $\sum_1^m |f(\mathbf{u}_i)| \geqslant 8 \log n |f(\mathbf{u}_j)|$ for each j. It follows that the above expression is at most $2\exp(-\log n) < N^{-1}$, as stated. This completes the proof of Lemma 5.2. □

Proof of Lemma 5·3. Given $\delta > 0$, let us define two block bases $(\mathbf{u}_i)_1^m$ and $(\mathbf{u}_i')_1^m$ to be δ-close if they are related in the following way. First, for any $i \neq j$, $\operatorname{supp}(\mathbf{u}_i) \cap \operatorname{supp}(\mathbf{u}_j') = \emptyset$, and second, for each i, $\left\|\mathbf{u}_i - \mathbf{u}_i'\right\|_p \leqslant \delta$. Let us estimate, for given δ, the size required of a set of block bases for every proper block basis of cardinality k to be δ-close to at least one block basis in the set. Since the number of ways of choosing m disjoint sets of size h is certainly less than n^{mh}, and, by Lemma 1.3, the number of vectors in a δ-net of the 2-ball of ℓ_p^h is at most $(1+4/\delta)^h$, we find that the size needed is certainly no greater than $(5n/\delta)^{mh} = M$.

Now fix δ to be n^{-1}, let $(\mathbf{u}_i)_1^m$ be a proper block basis, let $(\mathbf{v}_i)_1^m$ be δ-close to $(\mathbf{u}_i)_1^m$ and suppose that $(\mathbf{v}_i)_1^m$ satisfies condition $(*)$. Since $(\mathbf{u}_i)_1^m$ is a proper block basis, $\|\sum_1^m \eta_i \mathbf{u}_i\| \geqslant m^{1/p}$ for every choice of signs $(\eta_i)_1^m$. We therefore have, for some j,

$$\begin{aligned}\left\|\sum_1^m \eta_i \mathbf{u}_i\right\| &\geqslant \left\|\sum_1^m \eta_i \mathbf{v}_i\right\| - n^{1/q}\delta m^{1/p} \\ &\geqslant 80 \log n \left\|\mathbf{v}_j\right\|_1 - n^{1/q}\delta m^{1/p} \\ &\geqslant 80 \log n(\left\|\mathbf{u}_j\right\|_1 - \delta h^{1/q}) - n^{1/q}\delta m^{1/p} \\ &\geqslant 64 \log n \left\|\mathbf{u}_j\right\| \ .\end{aligned}$$

In a similar way, using the fact that for every choice of signs, $(\eta_i)_1^m$, $\|\sum_1^m \eta_i \mathbf{u}_i\|_p \geqslant m^{1/p}$, one can show that $\|\sum_1^m \eta_i \mathbf{u}_i\| \geqslant 4\,\|\sum_1^m \eta_i \mathbf{u}_i\|_p$, and hence that $(\mathbf{u}_i)_1^m$ satisfies condition $(**)$. □

We have left open the following problem. If $3/2 < p \leqslant 2$ and $x_1, \ldots, x_n$ is a normalized basis satisfying a lower p-estimate, then how large an almost symmetric block basis of $x_1, \ldots, x_n$ must there be? Theorems 3.1 and 5.1 show that the answer lies between about $n^{2/p-1}$ and about $n^{1/3}$. In Chapter 7 we shall give a partial answer to this. In particular, when $p = 2$, which is perhaps the most interesting case, we shall show that, for every $\alpha > 0$, if n is sufficiently large then there exists a normalized basis $x_1, \ldots, x_n$ satisfying a lower 2-estimate such that no block basis of cardinality n^α is 2-symmetric. This will justify the rough idea that one ought not to be able to deduce very much merely from the fact that a basis satisfies a lower 2-estimate, as this is a very weak property. Indeed, it can be shown that any n-dimensional space contains a basis of cardinality $n/2$ which satisfies such an estimate up to some absolute constant.

CHAPTER 6

BASES EQUIVALENT TO THE STANDARD BASIS OF ℓ_p^n

In Chapter 4, we showed that, if $1 \leqslant p < \infty$, $\epsilon > 0$ and $C \geqslant 1$, then every basis C-equivalent to the unit vector basis of ℓ_p^n has a $(1+\epsilon)$-symmetric block basis C-equivalent to the unit vector basis of ℓ_p^n has a $(1+\epsilon)$-symmetric block basis of cardinality $\alpha(\epsilon, p, C)n/\log n$, where $\alpha(\epsilon, p, C) > 0$ depends on ϵ, p and C only. In this chapter, we shall show that this result, at least when $p > 1$, is close to being best possible. When $p = 1$, we shall obtain an upper bound which shows that one cannot always find an almost symmetric block basis of proportional size. The two cases will be dealt with separately.

§6·1 The Case $1 < p < \infty$

The aim of this section is to construct, for given $0<\epsilon<1/2$ and $1<p<\infty$, a 1-unconditional norm $\|.\|$ on $\mathbb{R}^n$ which is 2-equivalent to $\|.\|_p$, such that no block basis of the standard basis of $\mathbb{R}^n$ with cardinality exceeding $m_0 = 1000(1+p+q)\epsilon^p n \log\log n/\log n$ is $(1+4^{-1/p}\epsilon/3)$-symmetric (where $1/p+1/q=1$). Of great importance in our construction is Lemma 2.7, since we shall use the sets A_B, F_B and G_B which were defined in Section 2.3. We shall collect together the salient properties of these sets in Lemma 6.1, but first, as we did in order to prove Theorem 5.1, we must show that a sufficiently large block basis has a large subbasis that is slightly easier to handle. That is, we will be able to restrict our attention to block bases which we shall call proper block bases. However, the definition of a proper block basis will be different in this section from the definition in the last chapter. For the rest of the chapter, we shall mean by a *block basis* a block basis of the standard basis of $\mathbf{R}^n$. Also, by the word "normalized", we shall mean normalized in ℓ_p^n.

Suppose then that $\|.\|$ is a norm on $\mathbf{R}^n$ which is 2-equivalent to $\|.\|_p$, let m_0 be as above, suppose $m_1 \geqslant m_0$, and let $\mathbf{u}_1, \dots, \mathbf{u}_{m_1}$ be a 2-symmetric block basis. Then there exists $\lambda > 0$ such that $1 \leqslant \|\lambda \mathbf{u}_i\|_p \leqslant 4$ for each i. For such a λ, set $\mathbf{u}'_i = \lambda \mathbf{u}_i$, for each i. By a simple and standard averaging argument there exists a subset $A \subset [m_1]$ with $|A| = n^{7/8}$ such that $\mathbf{u}_i$ is supported on at most $h = 25n/16m_0 = \log n/640(1+p+q)\epsilon^p \log\log n$ coordinates, for every $i \in A$. (The averaging argument gives more than this, of course, but we shall not need the extra precision.) By applying the pigeonhole principle, we can find $\mu \in [1,4]$ and $A' \subset A$ such that $|A'| = n^{3/4}$ and $\mu \leqslant \|\mathbf{u}'_i\| \leqslant \mu + 3n^{-1/8}$ for every $i \in A'$. If we set $\mathbf{u}''_i = \mu^{-1}\mathbf{u}'_i$ for each $i \in A'$, we certainly have

$$1 \leqslant \|\mathbf{u}''_i\| \leqslant 1 + 3n^{-1/8}$$

for every $i \in A'$.

We shall say that a block basis $\mathbf{v}_1, \dots, \mathbf{v}_m$ is *proper* if $m = n^{3/4}$, and $|\mathrm{supp}(\mathbf{v}_i)| \leqslant h$ and $1 \leqslant \|\mathbf{v}_i\| \leqslant 1 + n^{-1/8}$ for each $1 \leqslant i \leqslant m$. We have just shown that every block basis of cardinality at least m_0 has a subbasis, a multiple of which is proper. We shall construct a norm in such a way that no proper block basis is $(1 + 4^{-1/p}\epsilon/3)$-symmetric. It will follow that no block basis of cardinality m_0 or greater is $(1 + 4^{-1/p}\epsilon/3)$-symmetric under that norm.

The next lemma gives, as promised, the main properties that we shall use of the sets A_B, F_B and G_B. The third part of the lemma is simply a repetition of Lemma 2.7.

Lemma 6·1. *(i) For any set $B \in [t]^{(t/2)}$ and any normalized proper block basis $\mathbf{u}_1, \dots, \mathbf{u}_m$ there exists a sequence of scalars $a_1, \dots, a_m$ such that $\sum_1^m a_i \mathbf{u}_i \in A_B$.*

(ii) Let $B \in [t]^{(t/2)}$, let $\mathbf{u}_1, \dots, \mathbf{u}_m$ be a normalized proper block basis and let

$a_1, \ldots, a_m$ *be a sequence for which* $\sum_1^m a_i\mathbf{u}_i \in A_B$. *Then, for any permutation* $\pi \in S_m$ *and any sequence* $\epsilon_1, \ldots, \epsilon_m$ *of signs,*

$$\sum_1^m \epsilon_i a_{\pi(i)} \mathbf{u}_i \in A_B \ .$$

(iii) Suppose $B, C \in \mathcal{B}$, $B \neq C$ *and suppose* $f \in F_B$, $G \in G_B$ *and* $\mathbf{a} \in A_C$. *Then* $f(\mathbf{a}) + g(\mathbf{a}) \leqslant (t/2)^{1/q} \|\mathbf{a}\|_p$.

Proof. Note that, for any $X_j \subset [m]$ with $|X_j| = k^j$, the vector $|X_j|^{-1/p} \sum_{i \in X_j} \mathbf{u}_i$ is an element of A_j. Hence, if the sets X_j $(j \in B)$ are all disjoint, and we set $a_i = \left(\bigoplus_{j \in B} |X_j|^{-1/p}\chi_{X_j}\right)_i$, we have $\mathbf{a} = \bigoplus_1^m a_i\mathbf{u}_i \in A_B$. This proves the first part of the lemma. The second part is trivial and the third part was proved in Section 2.1. □

Let us now define the norm which we shall use. Let $\mathcal{B}$ be a set of the kind guaranteed to exist by Lemma 2.13, that is, a subset of $[t]^{(t/2)}$ of cardinality $(23/20)^t$ with the property that, whenever $B, C \in \mathcal{B}$ are distinct elements of $\mathcal{B}$, we have $|B \cap C| \leqslant t/3$. Set $r = \epsilon^p n$ and let $\gamma = (K_B : B \in \mathcal{B})$ be an element of $\Gamma = \left([n]^{(r)}\right)^{\mathcal{B}}$. Then we can define a norm $\|.\|_\gamma$ on $\mathbb{R}^n$ as follows:

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

There is one very important fact about the norm $\|.\|_\gamma$, namely, that if $C \in \mathcal{B}$ and $\mathbf{a} \in A_C$, then the maximum above is attained at C. That is, for such an $\mathbf{a}$ we have

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

This follows from Lemma 6.1 (iii). Indeed, if $B \neq C$, $f \in F_B$ and $g \in G_B$, then, by Lemma 6.1 (iii), $(2/t)^{1/q}(f(\mathbf{a}) + g(\mathbf{a})) \leqslant (t/2)^{1/p}$. On the other hand, if $\mathbf{a} \in A_C$, then $f = |\mathbf{a}|^{p-1}\mathrm{sign}(\mathbf{a}) \in F_C$, and $(2/t)^{1/q} f(\mathbf{a}) = (t/2)^{1/p}$.

Let P be the uniform probability measure on Γ. We shall show that if $\gamma \in \Gamma$ is chosen randomly, then there is a high probability that no normalized proper block basis is $(1+4^{-1/p}\epsilon)$-symmetric under $\|.\|_\gamma$.

We shall introduce one further definition to simplify the statement of the next lemma and later discussion. Suppose $\mathbf{u}$ is a unit vector in $\mathbb{R}^n$ such that $|\mathrm{supp}(\mathbf{u})| \leqslant h$, and suppose that $K \in [n]^{(r)}$. We shall say that $\mathbf{u}$ is *large on* K if it can be restricted to a vector $\mathbf{u}'$ with $|\mathrm{supp}(\mathbf{u})| \leqslant r = \epsilon^p h$, $\mathrm{supp}(\mathbf{u}') \subset K$ and $\left\|\mathbf{u}' \,|_K\right\|_p \geqslant 4^{-1/p}\epsilon$.

Lemma 6·2. *Let $\gamma = (K_B : B \in \mathcal{B}) \in \Gamma$, and let $(\mathbf{u}_i)_1^m$ be a normalized proper block basis. Suppose that there exist two sequences $i_1,\ldots,i_l$ and $j_1,\ldots,j_l$ and a set $B \in \mathcal{B}$ such that $\mathbf{u}_{i_1},\ldots,\mathbf{u}_{i_l}$ are all large on K_B, while $\mathrm{supp}(\mathbf{u}_{j_s}) \cap K_B = \emptyset$ for every $1 \leqslant s \leqslant l$. Then $(\mathbf{u}_i)_1^m$ is not $(1+4^{-1/p}\epsilon)$-symmetric under $\|.\|_\gamma$.*

Proof. By Lemma 6.1 (i) there exists a sequence $a_1,\ldots,a_m$ such that $\sum_1^m a_i\mathbf{u}_i \in A_B$. We shall show that $(\mathbf{u}_i)_1^m$ is not $(1+4^{-1/p})$-symmetric at $(a_i)_1^m$ under $\|.\|_\gamma$. Note that, because $\sum_{j=1}^t k^j \leqslant l$, the construction in Lemma 6.1 (i) gives a sequence $(a_i)_1^m$, all but at most l of whose terms are zero. By Lemma 6.1 (ii), we may assume that $a_{l+1} = \ldots = a_m = 0$. Let us set $\mathbf{a}' = \sum_{s=1}^l a_s\mathbf{u}_{i_s}$ and $\mathbf{a}'' = \sum_{s=1}^l a_s\mathbf{u}_{j_s}$. We shall then estimate $\|\mathbf{a}'\|_\gamma$ and $\|\mathbf{a}''\|_\gamma$.

First, let us calculate $\max\left\{g\left(\sum_{s\in X_j} a_s\mathbf{u}_{i_s}\right) : g \in G_j, \mathrm{supp}(g) \subset K_B\right\}$ when $j \in B$. Write $\mathbf{b}_j = \sum_{s\in X_j} a_s\mathbf{u}_{i_s}$. Then $\mathbf{b}_j \in A_j$ and since $\mathbf{u}_{i_1},\ldots,\mathbf{u}_{i_l}$ are large on K_B and disjointly supported, we can restrict $\mathbf{b}_j$ to a vector $\mathbf{b}'_j$ satisfying $\mathrm{supp}(\mathbf{b}'_j) \subset K_B$, $|\mathrm{supp}(\mathbf{b}'_j)| \leqslant \epsilon^p h k^j$ and $\left\|\mathbf{b}'_j\right\|_p \geqslant 4^{-1/p}\epsilon \left\|\mathbf{b}_j\right\|_p = 4^{-1/p}\epsilon$. We can therefore find $g \in G_j$ such that $g(\mathbf{b}_j) \geqslant 4^{-1/p}\epsilon$. Since $j \in B$ was arbitrary, it follows that we can find $g \in G_B$ such that $g(\mathbf{a}') \geqslant 4^{-1/p}\epsilon t/2$. It is obvious that we can find $f \in F_B$ such that $f(\mathbf{a}') \geqslant t/2$, so $\|\mathbf{a}'\|_\gamma \geqslant (t/2)^{1/p}(1+4^{-1/p}\epsilon)$.

Now $\|\mathbf{a}''\|_p = (t/2)^{1/p}$, and $\mathrm{supp}(\mathbf{a}'') \cap K_B = \emptyset$. But, for any $f \in F_B$,

$\|f\| = (t/2)^{1/q}$, so, by (1), $\|\mathbf{a}''\|_\gamma \leqslant (t/2)^{1/p}$. Hence $(\mathbf{u}_i)_1^m$ fails to be $(1+4^{-1/p}\epsilon)$-symmetric at $\mathbf{a}$, which proves the lemma. □

We shall show that the sequences needed in the conditions of Lemma 6.2 exist with very high probability. For this we shall use Corollary 2.12.

Lemma 6·3. *Let $(\mathbf{u}_i)_1^m$ be a normalized proper block basis and let $B \in \mathcal{B}$. Then the probability that we can find indices $i_1, \ldots, i_l$ such that $\mathbf{u}_{i_1}, \ldots, \mathbf{u}_{i_l}$ are large on K_B is at least*

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

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

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

Proof. We would like to estimate

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

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

Setting $\mathbf{a} = |\mathbf{u}_i|^p$ and taking $\epsilon^p/8$ instead of ϵ in Corollary 2.12, we obtain

$$\mathsf{P}\left[\mathbf{u}_i \text{ is large on } K \,\middle|\, \operatorname{supp}(\mathbf{u}_j) \cap K = W_j \text{ for } 1 \leqslant j < i\right]$$

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

$$p_1 \leqslant \binom{m}{l}\left(1 - (1/16)^{4\epsilon^p h}(1-\epsilon^p/4)^{2h}\right)^{m-l}.$$

Similarly,

$$p_2 \equiv \mathsf{P}\left[\left\|\mathbf{u}_i\,|_K\right\|_p = 0 \text{ for at most } l \text{ values of } i\right]$$
$$\leqslant \binom{m}{l}(1-(1-\epsilon^p/4)^h)^{m-l}$$

□

Corollary 6·4. *Let $(\mathbf{u}_i)_1^m$ be a normalized proper block basis. Then the probability that $(\mathbf{u}_i)_1^m$ is $(1+4^{-1/p}\epsilon)$-symmetric under $\|.\|_\gamma$ is at most*

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

Proof. For each $B \in \mathcal{B}$, construct $\mathbf{a}_B \in A_B$ generated by $\mathbf{u}_1, \ldots, \mathbf{u}_m$ as in the proof of Lemma 6.1(i). By (1), $\|\mathbf{a}_B\|_\gamma$ depends only on K_B, (where $\gamma = (K_B : B \in \mathcal{B})$). Let p_B be the probability that $\|.\|_\gamma$ is $(1+4^{-1/p}\epsilon)$-symmetric at a_B. Then, by Lemmas 6.2 and 6.3,

$$\begin{aligned} p_B &\leqslant \binom{m}{l}\left(1-(1/16)^{4\epsilon^p h}(1-\epsilon^p/4)^{2h}\right)^{m-l} + \binom{m}{l}\left(1-(1-\epsilon^p/4)^{h}\right)^{m-l} \\ &\leqslant 2\binom{m}{l}\left(1-(1/16)^{4\epsilon^p h}(1-\epsilon^p/4)^{2h}\right)^{m-l} \end{aligned}$$

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

□

The next lemma resembles Lemma 5.3. We shall prove it at the end of the section. First, however, we shall assume it and use it to prove the main theorem of the section.

Lemma 6·5. *Let $1 \leqslant p < \infty$, $0 < \eta < 1$ and let $M = (20n/\eta)^{mh}$. Then there exists a collection of M normalized proper block bases $(\mathbf{u}_i^1)_{i=1}^m, \ldots, (\mathbf{u}_i^M)_{i=1}^m$ of the standard basis of $\mathbb{R}^n$ such that any norm 2-equivalent to $\|.\|_p$ which fails to be $(1+\eta)$-symmetric on any of $(\mathbf{u}_i^1)_{i=1}^m, \ldots, (\mathbf{u}_i^M)_{i=1}^m$ fails to be $(1+\eta/3)$-symmetric on any proper block basis.*

The proof of our main theorem is now a simple matter of verification.

Theorem 6·6. *Let $0 < \epsilon < 1/2$, let $1 < p < \infty$, let $n \in \mathbb{N}$ and let $m_0 = 1000(1+p+q)\epsilon^p n \log\log n/\log n$. Then there exists a norm $\|.\|$ on $\mathbb{R}^n$ such that for any $x \in \mathbb{R}^n$, $\|x\|_p \leqslant \|x\| \leqslant 2\|x\|_p$, but no block basis of cardinality exceeding m_0 is $(1+\delta)$-symmetric for any $\delta < 4^{-1/p}\epsilon/3$.*

Proof. By Corollary 6.4 and Lemma 6.5 (with $\eta = 4^{-1/p}\epsilon$) it remains only to show that

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

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 $\|.\|_\gamma$.

Now

$$\begin{aligned}
&\left[2\binom{m}{l}\left(1-(1/16)^{4\epsilon^p h}(1-\epsilon^p/4)^h\right)^{m-l}\right]^N \\
&\qquad \leqslant 2^N m^{lN} \exp\left(-(m-l)N(1/16)^{4\epsilon^p h}(1-\epsilon^p/4)^h\right) \\
&\qquad = \exp\left(N\left(\log 2 + l\log m - (m-l)(1/16)^{4\epsilon^p h}(1-\epsilon^p/4)^h\right)\right)
\end{aligned}$$

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

$$\begin{aligned}
&\exp\left(-(1/2)N(m-l)(1/16)^{4\epsilon^p h}(1-\epsilon^p/4)^h\right) \\
&\qquad \leqslant \exp\left(-(1/2)N(m-l)\exp(-\epsilon^p h(\log 16 + 1/4))\right)
\end{aligned}$$

Now it is easy to check that $h \leqslant \log N/40\epsilon^p$, so this is at most $\exp(-N^{1/2}(m-l))$ which is certainly at most $(20n/4^{-1/p}\epsilon)^{-mh}$. This completes the proof of Theorem 6.6, assuming Lemma 6.5. □

It remains to prove Lemma 6.5.

Proof of Lemma 6·5. Let us call two proper block bases $(\mathbf{u}_i)_1^m$ and $(\mathbf{v}_i)_1^m$ α-*close* if they satisfy $\mathrm{supp}(\mathbf{u}_i)\cap\mathrm{supp}(\mathbf{v}_j) = \emptyset$ whenever $i \neq j$ and if $\|\mathbf{u}_i - \mathbf{v}_i\|_p \leqslant \alpha$ for each i. Suppose also, without loss of generality, that for any $x \in \mathbb{R}^n$, $\|x\|_p \leqslant \|x\| \leqslant 2\|x\|_p$. Now, if $(\mathbf{u}_i)_1^m$ and $(\mathbf{v}_i)_1^m$ are α-close, then, given any sequence $(a_i)_1^m \in \mathbb{R}^m$,

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

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

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

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 normalized proper block bases $(\mathbf{u}_i^1)_{i=1}^m, \dots, (\mathbf{u}_i^M)_{i=1}^m$ such that any proper block basis is $\eta/8$-close to $(\mathbf{u}_i^r)_{i=1}^m$ for some $1 \leqslant r \leqslant M$. This proves the lemma. □

§6·2 The Case $p=1$

We now turn to the case $p=1$. The proof given in the case $p>1$ breaks down when $p=1$ and one does not obtain any result even by approximating ℓ_1^n by ℓ_p^n for p sufficiently close to 1 for ℓ_p^n to be uniformly equivalent to ℓ_1^n. Roughly speaking, this is because the unit ball of ℓ_1^n is not sufficiently convex to allow many classes of "well-separated" vectors. However, using the construction of Section 2.3 and Lemma 2.8, we shall obtain a basis which is C-equivalent to the unit vector basis of ℓ_1^n such that no block basis of cardinality exceeding $\alpha(\epsilon, C)n/\log\log n$ is $(1+\epsilon)$-symmetric. The proof is similar to the proof of Theorem 6.6: we shall prove a sequence of lemmas resembling Lemmas 6.1 to 6.4.

First, we need a definition of a proper block basis appropriate for this context. Let $m_0 = (240\epsilon/\log C)n/\log\log n$ and let $h=(\log C/200\epsilon)\log\log n$. We shall say that a block basis $\mathbf{u}_1, \dots, \mathbf{u}_m$ is *proper* if $m=n^{3/4}$, $|\mathrm{supp}(\mathbf{u}_i)| \leqslant h$ and $1 \leqslant \|\mathbf{u}_i\| \leqslant 1+3n^{-1/8}$ for each $1 \leqslant i \leqslant m$. As in the last section, one can easily show that any 2-symmetric block basis of cardinality exceeding m_0 has a subbasis, a multiple of which is proper.

We shall begin by defining a random norm on $\mathbb{R}^n$. Recall from Section 2.3 that we set $r = -\log C/\log(1-4\epsilon)$, $k = h^2$, $\lambda = \lfloor \log_k n/2r \rfloor$ and $N = (\lambda/2)^{r/2}$. Let $\leqslant$ be the usual partial order on $[\lambda/2]^r$, that is $(s_1, \ldots, s_r) \leqslant (t_1, \ldots, t_r)$ iff $s_i \leqslant t_i$ for each $1 \leqslant i \leqslant r$. Let $\mathcal{S}$ be a subset of $[\lambda/2]^r$ of the kind which is guaranteed to exist by Lemma 2.14. That is, $\mathcal{S}$ has cardinality $N = (\lambda/6\epsilon)^{r/12}$, and if $\mathbf{s}, \mathbf{s}' \in \mathcal{S}$ are distinct then there are at least $r/3$ values of $i \in [r]$ for which $s_i < s_i'$ and vice versa. In other words, $\mathcal{S}$ is an antichain in a very strong sense. Now let $\Gamma = ([n]^{(n/2)})^{\mathcal{S}}$. For any element $\gamma = (K_{\mathbf{s}} : \mathbf{s} \in \mathcal{S}) \in \Gamma$, define a norm $\|.\|_\gamma$ on $\mathbb{R}^n$ by

$$\|x\|_\gamma = \|x\|_1 \vee \max_{\mathbf{s}\in\mathcal{S}} \max\{f_1(x) + \epsilon f_2(x) : f_1, f_2 \in F_{\mathbf{s}}, \mathrm{supp}(f_2) \subset K_{\mathbf{s}}\}.$$

The next lemma collects together the useful properties of the sets $A_{\mathbf{s}}$ and $F_{\mathbf{s}}$.

Lemma 6·7. *(i) For any $\mathbf{s} \in [\lambda/2]^r$ and any normalized proper block basis $\mathbf{u}_1, \ldots, \mathbf{u}_m$, there exists a sequence of scalars $a_1, \ldots, a_m$ such that $\sum_1^m a_i \mathbf{u}_i \in A_{\mathbf{s}}$.*

(ii) Let $\mathbf{s} \in [\lambda/2]^r$, let $\mathbf{u}_1, \ldots, \mathbf{u}_m$ be a normalized proper block basis and let $a_1, \ldots, a_m$ be a sequence of scalars for which $\sum_1^m a_i \mathbf{u}_i \in A_{\mathbf{s}}$. Then, for any permutation $\pi \in S_m$ and any sequence $\epsilon_1, \ldots, \epsilon_m$ of signs,

$$\sum_1^m \epsilon_i a_i \mathbf{u}_{\pi(i)} \in A_{\mathbf{s}} \; .$$

(iii) Suppose $i \neq j$ and suppose that $f_1, f_2 \in F_i$ and $\mathbf{a} \in A_{\mathbf{s}_j}$. Then $f_1(\mathbf{a}) + \epsilon f_2(\mathbf{a}) \leqslant r$.

Proof. If $X_i \subset [m]$ has cardinality $k^{(i-1)\lambda + s_i}$, then $\beta_i^{-1} k^{-(i-1)\lambda - s_i} \sum_{j \in X_i} \mathbf{u}_j \in A_{i,s_i}$. If we pick disjoint sets $X_1, \ldots, X_r$ each satisfying the above, and if we set

$$a_i = \left(\bigoplus_{j=1}^r \beta_j^{-1} k^{-(j-1)m - s_j} \chi X_j \right)_i$$

then $\mathbf{a} = \sum_{i=1}^m a_i \mathbf{u}_i \in A_{\mathbf{s}}$. The second part is trivial, and the third follows immediately from Lemma 2.8. $\square$

There is an important fact about $\|.\|_\gamma$ analogous to the fact mentioned after the proof of Lemma 6.1, namely that if $\mathbf{s} \in \mathcal{S}$ and $\mathbf{a} \in A_{\mathbf{s}}$, then the outer maximum in the definition of $\|\mathbf{a}\|_\gamma$ is attained at $\mathbf{s}$. That is, for such an $\mathbf{a}$, we have

$$\|\mathbf{a}\|_\gamma = \|\mathbf{a}\|_1 \vee \max\{f_1(\mathbf{a}) + \epsilon f_2(\mathbf{a}) : f_1, f_2 \in F_{\mathbf{s}}, \operatorname{supp}(f_2) \subset K_{\mathbf{s}}\}. \tag{2}$$

This follows from Lemma 6.7 and the fact that $\mathcal{S}$ is an antichain in the strong sense of Lemma 2.14. Indeed, suppose $\mathbf{s}$ and $\mathbf{s}'$ are distinct members of $\mathcal{S}$ and suppose that $f_1, f_2 \in F_{\mathbf{s}'}$ and $\mathbf{a} \in A_{\mathbf{s}}$. Then, by Lemma 6.7 (iii), $f_1(\mathbf{a}) + f_2(\mathbf{a}) \leqslant r$. On the other hand, if $\mathbf{a} = \sum_1^n a_i \mathbf{e}_i = \bigoplus_{j=1}^r \mathbf{a}_j$ with $\mathbf{a}_j \in A_{j,s_j}$ for each $1 \leqslant j \leqslant r$, then one can define a functional f on $\mathbb{R}^n$ by

$$f_i = \begin{cases} \beta_j \operatorname{sign}(a_i) & i \in \operatorname{supp}(\mathbf{a}_j) \\ 0 & \text{otherwise} \end{cases}$$

It is easy to see that $f \in F_{\mathbf{s}}$ and $f(\mathbf{a}) \geqslant r$.

Given $\mathbf{u} \in \mathbb{R}^n$ and $K \in [n]^{(n/2)}$, let us define $\mathbf{u}$ to be *large on* K if $\operatorname{supp}(\mathbf{u}) \subset K$. Suppose that $\mathbf{u}_1, \ldots, \mathbf{u}_t$ is a sequence of disjointly supported vectors such that $\sum_1^t |\operatorname{supp}(\mathbf{u}_i)| \leqslant n/6$, and suppose that K is chosen randomly from $[n]^{(n/2)}$. Provided that $i \leqslant \sqrt{n}$, the following two estimates are easy, whatever the sets W_j which appear might be.

$$\mathbb{P}\left[\operatorname{supp}(\mathbf{u}_i) \subset K \,\middle|\, \operatorname{supp}(\mathbf{u}_j) \cap K = W_j \text{ for } 1 \leqslant j < i\right] \geqslant 3^{-h} \tag{3}$$

and

$$\mathbb{P}\left[\operatorname{supp}(\mathbf{u}_i) \cap K = \emptyset \,\middle|\, \operatorname{supp}(\mathbf{u}_j) \cap K = W_j \text{ for } 1 \leqslant j < i\right] \geqslant 3^{-h}. \tag{4}$$

The next two lemmas are analogous to Lemmas 6.2 and 6.3.

Lemma 6·8. *Let* $\gamma = (K_{\mathbf{s}} : \mathbf{s} \in \mathcal{S})$, *let* $l = n^{1/2}$ *and let* $\mathbf{u}_1, \ldots, \mathbf{u}_m$ *be a normalized proper block basis. Suppose that there exist two sequences* $i_1, \ldots, i_l$ *and* $j_1, \ldots, j_l$ *and* $\mathbf{s} \in \mathcal{S}$ *such that* $\mathbf{u}_{i_1}, \ldots, \mathbf{u}_{i_l}$ *are all large on* $K_{\mathbf{s}}$, *while* $\mathrm{supp}(\mathbf{u}_{j_t}) \cap K_{\mathbf{s}} = \emptyset$ *for every* $1 \leqslant t \leqslant l$. *Then* $\mathbf{u}_1, \ldots, \mathbf{u}_m$ *is not* $(1+\epsilon)$*-symmetric under* $\|.\|_\gamma$.

Proof. By Lemma 6.7 (i) there exists a sequence $a_1, \ldots, a_m$ such that $\sum_1^m a_i \mathbf{u}_i \in A_{\mathbf{s}}$. We shall show that $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is not $(1+\epsilon)$-symmetric at $a_1, \ldots, a_m$ under $\|.\|_\gamma$. The construction of Lemma 6.7 (i) gives a sequence $a_1, \ldots, a_m$ such that all but at most l of the a_i are zero, so without loss of generality $a_{l+1} = \ldots = a_m = 0$. Let us set $\mathbf{a}' = \sum_{t=1}^l a_t \mathbf{u}_{i_t}$ and $\mathbf{a}'' = \sum_{t=1}^l a_t \mathbf{u}_{j_t}$, and estimate $\|\mathbf{a}'\|_\gamma$ and $\|\mathbf{a}''\|_\gamma$.

We have already seen that if $\mathbf{a} \in A_{\mathbf{s}}$ for some $\mathbf{s}$, then there is a functional $f \in F_{\mathbf{s}}$ such that $f(\mathbf{a}) = r$. Now $\mathbf{a}'$ has the additional property of being supported inside K, so we can of course find $f \in F_{\mathbf{s}}$ such that $\mathrm{supp}(f) \subset K$ and $f(\mathbf{a}) = r$. By the definition of the norm it follows that $\|\mathbf{a}'\|_\gamma \geqslant r(1+\epsilon)$. On the other hand, $\mathbf{a}'' \in A_{\mathbf{s}}$ and $\mathrm{supp}(\mathbf{a}'') \cap K = \emptyset$. By the remarks following the proof of Lemma 6.7, it follows that $\|\mathbf{a}''\|_\gamma \leqslant r$. The result follows. □

Lemma 6·9. *Let* $\mathbf{u}_1, \ldots, \mathbf{u}_m$ *be a normalized proper block basis and let* $K \in [n]^{(n/2)}$. *Then the probability that we can find* $i_1, \ldots, i_l$ *such that* $\mathbf{u}_{i_1}, \ldots, \mathbf{u}_{i_l}$ *are large on* K *is at least*

$$1 - \binom{m}{l}(1 - 3^{-h})^{m-l}$$

and the probability that we can find $j_1, \ldots, j_l$ *such that* $\mathrm{supp}(\mathbf{u}_{j_t}) \cap K = \emptyset$ *for every* $1 \leqslant t \leqslant l$ *is also at least*

$$1 - \binom{m}{l}(1 - 3^{-h})^{m-l}.$$

Proof. Let

$$p_1 = \mathbf{P}\left[\mathbf{u}_i \text{ is large on } K \text{ for at most } l \text{ values of } i\right]$$

where K is chosen randomly from $[n]^{(n/2)}$. From (3) we can deduce that

$$p_1 \leqslant \binom{m}{l}(1-3^{-h})^{m-l}.$$

Similarly, using (4), we obtain

$$\begin{aligned} p_2 &= \mathbb{P}\left[\operatorname{supp}(\mathbf{u}_i)\cap K=\emptyset \text{ for fewer than } l \text{ values of } i\right] \\ &\leqslant \binom{m}{l}(1-3^{-h})^{m-l}. \end{aligned}$$

□

Corollary 6·10. *Let $\mathbf{u}_1,\ldots,\mathbf{u}_m$ be a normalized proper block basis and let γ be chosen randomly from Γ. Then the probability that $\mathbf{u}_1,\ldots,\mathbf{u}_m$ is $(1+\epsilon)$-symmetric under $\|.\|_\gamma$ is at most*

$$\left[2\binom{m}{l}(1-3^{-h})^{m-l}\right]^N.$$

Proof. Given any $\mathbf{s}\in\mathcal{S}$, let $\mathbf{a_s}$ be the vector constructed in Lemma 6.7 (i). Then, by (2), the norm of any rearrangement of $\mathbf{a_s}$ depends only on $K_\mathbf{s}$, (where $\gamma=(K_\mathbf{s}:\mathbf{s}\in\mathcal{S})$). Let $p_\mathbf{s}$ be the probability that $\mathbf{u}_1,\ldots,\mathbf{u}_m$ is $(1+\epsilon)$-symmetric at $\mathbf{a_s}$ under $\|.\|_\gamma$. By Lemmas 6.8 and 6.9,

$$p_\mathbf{s}\leqslant 2\binom{m}{l}(1-3^{-h})^{m-l}$$

and if $\mathbf{s}\neq\mathbf{s}'$ then $p_\mathbf{s}$ and $p_{\mathbf{s}'}$ are independent. The result follows. □

We are left, as in the proof of Theorem 6.6, with some numerical verification.

Theorem 6.11. *Let $\epsilon>0$ and $C>1$, let n be a sufficiently large positive integer and let $m_0=(240\epsilon/\log C)n/\log\log n$. There exists a norm $\|.\|$ on $\mathbb{R}^n$ such that, for any $x\in\mathbb{R}^n$, $\|x\|_1\leqslant\|x\|\leqslant C\|x\|_1$, but no block basis of the standard basis of cardinality exceeding m_0 is $(1+\epsilon)$-symmetric.*

Proof. It is easily seen that $\|x\|_1\leqslant\|x\|_\gamma\leqslant C\|x\|_1$ for any $x\in R^n$ and $\gamma\in\Gamma$. Indeed, $\|.\|_\gamma$ is defined as a supremum of functionals, each of which is of ℓ_∞-norm at most $(1+\epsilon)\beta_1 \leqslant C$. Recall that the cardinality of the "strong antichain" $\mathcal{S}$ is $N = (\lambda/6\epsilon)^{r/12}$, where $r = -\log C/\log(1-4\epsilon)$ and $\lambda = \lfloor(\log_k n)/2r\rfloor$. In particular, $N \geqslant \lambda^{\log C/60\epsilon}$. Recall also that $h = (\log C/200\epsilon)\log\log n$. By Corollary 6.10 and Lemma 6.5, it remains only to show that

$$\left[2\binom{m}{l}(1-3^{-h})^{m-l}\right]^N < (20n/\epsilon)^{-mh}.$$

Now

$$\begin{aligned}\left[2\binom{m}{l}(1-3^{-h})^{m-l}\right]^N &\leqslant 2^N m^{lN}\exp(-3^{-h}(m-l)N)\\ &= \exp\big(N(\log 2 + l\log m - 3^{-h}(m-l))\big)\end{aligned}$$

Since $h < \log_3(n^{1/5})$, this is at most

$$\exp\big(-(1/2)3^{-h}N(m-l)\big)$$

Now h is certainly less than $\log_3(N^{1/2}/4)$, so this is at most $\exp(-N^{1/2}m)$. This is easily checked to be less than $(20n/\epsilon)^{-mh}$. □

This, then, completes the proof of our upper bounds in this section. However, we have left open one or two interesting questions. For example, we have no information about the case $p=\infty$, and in the case $p=1$, there is quite a gap between the lower bound obtained in Theorem 4.1 and the upper bound given by Theorem 6.11. We shall discuss the case $p=\infty$ in detail in the next chapter. The gap when $p=1$ suggests the following question. Suppose $x_1,\dots,x_n$ is a basis which is C-equivalent to the unit vector basis of ℓ_p^n, and suppose $a_1,\dots,a_n$ is any sequence of scalars. Then how large a block basis must there be which is $(1+\epsilon)$-symmetric at $a_1,\dots,a_n$? The results of Chapter 4 and this section show only that, up to a constant depending on ϵ, p and C, the answer is between $n/\log n$ and n. If the answer were $n/\log n$, we would be able to prove Theorem 6.6 without using

Lemma 2.7, and the result would be valid when $p = 1$. It is not hard to find a sequence of scalars $a_1, \ldots, a_m$ and a norm such that the probability of a random block basis of cardinality significantly larger than $n/\log n$ being almost symmetric at $a_1, \ldots, a_m$ is very small. However, my hope is that this is an interesting case of an event with a small but necessarily positive probability. Probabilistic methods proving the existence of rare events have been used in combinatorics ever since a simple but powerful result of Erdős and Lovász [17] (cf. [1], [10]).

I think the dependence on ϵ in Theorem 6.6 can be improved from ϵ^p to ϵ^{2p} by a slight adaptation of the definition of the norm. On the other hand, the correct dependence on C in Theorem 4.1 is far from clear. It is not hard to see that the norm constructed in this section is actually $(1+2\epsilon)$-equivalent to $\|.\|_p$. Again, if we had an upper bound for the size of a block basis almost symmetric at a single vector $\mathbf{a}$ then this situation would probably be clarified. An almost negligible dependence on C can be obtained by choosing p' so close to p that the standard basis of $\ell_{p'}^n$ is $(1+2\epsilon)^{-1}C$-equivalent to the standard basis of ℓ_p^n. One then uses the construction of this section with p replaced by p', thereby improving the dependence on ϵ from ϵ^p to $\epsilon^{p'}$. The dependence on C in Theorem 6.11 is also a very weak one. For large enough values of C (depending on n), Theorem 5.1 gives a better bound.

Lemma 2.7, and the result would be valid when $p = 1$. It is not hard to find a sequence of scalars $a_1, \ldots, a_m$ and a norm such that the probability of a random block basis of cardinality significantly larger than $n/\log n$ being almost symmetric at $a_1, \ldots, a_m$ is very small. However, my hope is that this is an interesting case of an event with a small but necessarily positive probability. Probabilistic methods proving the existence of rare events have been used in combinatorics ever since a simple but powerful result of Erdős and Lovász [17] (cf. [1], [10]).

I think the dependence on ϵ in Theorem 6.6 can be improved from ϵ^p to ϵ^{2p} by a slight adaptation of the definition of the norm. On the other hand, the correct dependence on C in Theorem 4.1 is far from clear. It is not hard to see that the norm constructed in this section is actually $(1 + 2\epsilon)$-equivalent to $\|.\|_p$. Again, if we had an upper bound for the size of a block basis almost symmetric at a single vector $\mathbf{a}$ then this situation would probably be clarified. An almost negligible dependence on C can be obtained by choosing p' so close to p that the standard basis of $\ell_{p'}^n$ is $(1 + 2\epsilon)^{-1}C$-equivalent to the standard basis of ℓ_p^n. One then uses the construction of this section with p replaced by p', thereby improving the dependence on ϵ from ϵ^p to $\epsilon^{p'}$. The dependence on C in Theorem 6.11 is also a very weak one. For large enough values of C (depending on n), Theorem 5.1 gives a better bound.

CHAPTER 7

GENERAL 1-UNCONDITIONAL BASES

Let us begin by recalling the main result of Chapter 4, namely Theorem 4·1. This states that, if $n \in \mathbb{N}$, $1 \leqslant p < \infty$, $C > 1$ and $\epsilon > 0$, then any basis $x_1, \ldots, x_n$ which is C-equivalent to the standard basis of ℓ_p^n has a block basis of cardinality at least $m = \alpha(\epsilon, p, C)n/\log n$ which is $(1+\epsilon)$-symmetric. Moreover, the blocks have ± 1-coefficients and they are of equal length. We did not state this as part of the theorem, but the proof gives blocks of length $h = n/m$. In other words, we did not discard any of the vectors in the original basis.

The results of this chapter were motivated by the question of what one could say about the size of almost symmetric block bases of sequences which were equivalent to the unit vector basis of ℓ_∞^n. A natural approach to this question is to approximate ℓ_∞^n by ℓ_p^n for some suitably large p (possibly depending on n) and to apply Theorem 4.1. However, this fails, because either p or C has to be so large that $\alpha(\epsilon, p, C)$ turns out to be less than n^{-1}. Moreover, there is a good reason for this failure. Consider the standard basis of $\mathbb{R}^n$ with the norm on $\mathbb{R}^n$ given by

$$\left\| \sum_{i=1}^{n} a_i \mathbf{e}_i \right\| = \max\Big\{ C|a_1|, |a_2|, |a_3|, \ldots, |a_n| \Big\} \,.$$

It is clear that no block basis with ± 1-coefficients of cardinality 2 or more which uses all the vectors in the original basis can be C'-symmetric for any $C' < C$.

In general it seems that, given any natural class of block bases of the standard basis of $\mathbb{R}^n$ with a natural probability measure on that class, it is easy to construct a norm on $\mathbb{R}^n$ so that the standard basis of $\mathbb{R}^n$ is equivalent to the standard basis of ℓ_∞^n and so that the probability of a random block basis being almost symmetric is very small. This suggests that a purely probabilistic method is unlikely to give an interesting result here.

Our main result concerning this question is, not surprisingly then, a negative one. Given $C > 1$ and $\epsilon > 0$, we shall construct a basis which is C-equivalent to the standard basis of ℓ_∞^n and has no $(1+\epsilon)$-symmetric block basis of cardinality $n^{\beta(\epsilon,C)}$, where $\beta(\epsilon, C)$ is a function of ϵ and C only which tends to zero as $\log(1+\epsilon)/\log C$ tends to zero. The number $\log(1+\epsilon)/\log C$ will crop up throughout the section and we will denote it by $\alpha = \alpha(\epsilon, C)$. Again, it is not possible to obtain this result by approximating ℓ_∞^n by ℓ_p^n for some large p and applying earlier results. One would like to apply Theorem 6.6 with p about $\log n$, but the implicit condition that n should be sufficiently large turns out not to be satisfied if $n \leqslant e^p$.

Our result is, in a weak sense, best possible, since a very simple technique, due essentially to James [27] (cf. e.g. [36] for a finite-dimensional version), shows that any basis which is C-equivalent to the unit vector basis of ℓ_∞^n has a block basis of cardinality $k = n^{\alpha/2}$, where $\alpha = \log(1+\epsilon)/\log C$, which is $(1+\epsilon)$-equivalent to the unit vector basis of ℓ_∞^k and which is *a fortiori* $(1+\epsilon)$-symmetric. On the other hand, our function $\beta(\epsilon, C)$ is considerably larger than $\alpha(\epsilon, C)$, and it would be of interest to determine the best possible function. We obtain $\beta = c/\log(\alpha^{-1})$ for some absolute constant c.

In our result, C may actually depend on n. If we take $C = C(n)$ tending to infinity with n, however slowly, then we obtain the first example of a sequence $E_1, E_2, \ldots$ of bases with the cardinality of E_n being n, such that, for any $\gamma > 0$ there exists n_0 such that for every n greater than n_0, the basis E_n does not have a $(1+\epsilon)$-symmetric block basis of cardinality n^γ. To put this loosely, we obtain a basis of length n such that the size of its largest almost symmetric block basis is bounded above by a function which grows more slowly than any positive power of n. In particular, if we have no condition on C at all, then α may be proportional to $(\log n)^{-1}$, and the function is $n^{c/\log\log n}$, where c is again some absolute constant.

It turns out that, with very little further work, one can show that, when C is indeed a power of n, the space generated by the basis E_n above does not contain any subspace with a 2-symmetric basis of dimension greater than $n^{c/\log\log n}$. In particular, for any $\gamma > 0$, there exists $n_0 \in \mathbb{N}$ such that for every $n \geqslant n_0$ there is a normed space of dimension n containing no n^γ-dimensional subspace with a 2-symmetric basis. One can, of course, replace 2 by any absolute constant.

Milman has asked (private communication) whether every normed space of dimension n contains a subspace of dimension proportional to $\sqrt{n}$ with a 2-symmetric basis. Our result therefore gives a negative answer to this question. It is also related to a well known problem concerning the dependence on ϵ in Dvoretzky's theorem. This well known result states that, for given $k \in \mathbb{N}$ and $\epsilon > 0$, there exists $n = n(k, \epsilon)$ such that, if X is any n-dimensional normed space, then X has a k-dimensional subspace which is $(1+\epsilon)$-isomorphic to ℓ_2^k. It is not known whether n is bounded above by a polynomial function of ϵ^{-1} when k is fixed. Bourgain and Lindenstrauss [13] have shown that it is bounded above polynomially if the space X has a 1-symmetric basis, and Lindenstrauss has told me that this is also the case if X has a 2-symmetric basis. Hence, if there were some fixed $\gamma > 0$ and $c > 0$ such that any n-dimensional normed space contained a 2-symmetric basic sequence of cardinality at least cn^γ, then the dependence on ϵ in Dvoretzky's theorem would indeed be polynomial. Although our construction shows that this premise is false, it does not give an example of non-polynomial dependence for Dvoretzky's theorem.

It is worth mentioning that, given a basis which is equivalent to the standard basis of ℓ_∞^n, we do not know anything about how large a $(1+\epsilon)$-unconditional block basis it must have, except, of course, for the lower bound of $n^{\alpha/2}$ mentioned above. Our examples are all in fact 1-unconditional sequences. It is possible that some technique involving the Sauer-Shelah lemma might produce a significant lower bound. This would be interesting in view of the fact that there is no obvious measure-concentration argument. However, it is at present nothing more than speculation.

There are some similarities between our methods in this chapter and those of Chapter 5. For example, we will tend to fix a k and show that if a block basis $\mathbf{u}_1, \ldots, \mathbf{u}_t$ is large enough, then not all sums of the form $\sum_{i \in K} \pm \mathbf{u}_i$, where K is a set of cardinality k, can have approximately the same norm. However, in order to show that there are sets K for which the norm of $\sum_{i \in K} \pm \mathbf{u}_i$ is large, we shall use simple deductions from the pigeonhole principle rather than probabilistic methods. On the other hand, we do use probabilistic methods to show that there are combinations of the form $\sum_{i \in K} \pm \mathbf{u}_i$ whose norms are small.

Unlike the norms defined in Chapters 5 and 6, the norms in this chapter have simple explicit definitions. In fact, they are all built up using iterated ℓ_p-direct sums. Given any normed space $X = (\mathbb{R}^k, \|.\|)$ such that the standard basis of $\mathbb{R}^k$ is 1-unconditional, let $(X)_p^r$ denote the p-direct sum of r copies of X. This is sometimes written $\ell_p^r(X)$, or else $(X \oplus \ldots \oplus X)_p$ with the number of copies specified. We shall regard $\mathbb{R}^{rk}$ as the underlying vector space of X in the natural way, and by the *standard basis* of X, we shall mean the concatenation of r copies of the standard basis of $\mathbb{R}^k$. In other words, denoting the standard basis by $\mathbf{e}_1, \ldots, \mathbf{e}_{rk}$, the j^{th} copy of X is generated by $\mathbf{e}_{(j-1)k+1}, \ldots, \mathbf{e}_{jk}$. All the spaces we consider in this chapter will be of the form $(\ldots(\ell_{p_1}^{m_1})_{p_2}^{m_2} \ldots)_{p^k}^{m^k}$ for some $p_1, \ldots, p_k$ in the closed interval $[1, \infty]$ and integers $m_1, \ldots, m_k$. We shall call such spaces *compound ℓ_p-spaces.*

We shall not prove our main theorem immediately, since its proof is rather complicated and should be easier to understand after a weaker result. We consider first the space $X = (\ell_p^m)_\infty^m$ and show that, for a suitable choice of p, the standard basis of X is C-equivalent to the standard basis of ℓ_∞^n, and its largest almost symmetric block basis has cardinality not much greater than $\sqrt{n}$. We shall make use of Lemma 2.10 here and several times later in the chapter. Just before we begin, we emphasize here that most of the results of this chapter are valid only when n is sufficiently large, and that we shall often assume this in our calculations.

Theorem 7·1. *Let $n = m^2$, let $p = \log n/2\log C$, let $X = (\ell_p^m)_\infty^m$ and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis of X. Let us set $\alpha = \log(1+\epsilon)/\log C$ and $s = n^{(1+\alpha)/2}(\log_2 n)^2$. Then $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is C-equivalent to the standard basis of ℓ_∞^n, and no block basis $\mathbf{u}_1, \ldots, \mathbf{u}_s$ of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of cardinality s is $(1+\epsilon)$-symmetric.*

Proof. It is easy to check that $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is C-equivalent to the standard basis of ℓ_∞^n. So let us suppose that $\mathbf{u}_1, \ldots, \mathbf{u}_s$ is a $(1+\epsilon)$-symmetric block basis of cardinality s. For $1 \leqslant j \leqslant m$, let us write A_j for the set $\{(j-1)m+1, \ldots, jm\}$. By the pigeonhole principle, there is a subset $A \subset [s]$ of cardinality $s/\log_2 n$ and some $h > 0$ such that $h \leqslant \|\mathbf{u}_i\|_X^p \leqslant 2h$ for every $i \in A$. Let $t = s/\log_2 n = mn^{\alpha/2}\log_2 m$. Without loss of generality, $A = [t]$ and $h = 1$.

Recall (from Chapter 1) that, given any vector $\mathbf{a} \in \mathbb{R}^n$, we denote by $\mathbf{a}\,|_K$ the restriction of $\mathbf{a}$ to the subspace generated by $\{\mathbf{e}_i : i \in K\}$. By definition, then, $\|\mathbf{a}\|_X^p = \max_{1\leqslant j\leqslant m} \|\mathbf{a}|_{A_j}\|_p^p$ for any vector $\mathbf{a} \in \mathbb{R}^n$. By the pigeonhole principle, there is a set $B \subset [t]$ of cardinality t/m and $j \in [m]$ such that, for every $i \in B$,

$$\|\mathbf{u}_i\|_X^p = \|\mathbf{u}_i|_{A_j}\|_p^p .$$

Since the $\mathbf{u}_i$ are disjointly supported, we therefore have

$$\Big\|\sum_{i\in B}\mathbf{u}_i\Big\|_X^p \geqslant t/m.$$

Let us set $\delta = (1+\epsilon)^{-p} = n^{-\alpha/2}$ and $l = t/m$. Because $\mathbf{u}_1, \ldots, \mathbf{u}_t$ is $(1+\epsilon)$-symmetric, we must have

$$\delta l \leqslant \Big\|\sum_{i\in B'}\mathbf{u}_i\Big\|_X^p$$

for any set $B' \in [t]^{(l)}$. It follows that there exists $j \in [m]$ such that if B' is chosen randomly from $[t]^{(l)}$, then

$$\mathsf{P}\left[\left\|\sum_{i\in B'} \mathbf{u}_i \Big|_{A_j}\right\|_p^p \geqslant \delta l\right] \geqslant 1/m.$$

We shall use Lemma 2.10 to show that this cannot be the case. For such a j, set $a_i = \left\|\mathbf{u}_i|_{X_j}\right\|_p^p$, for each $i \in [t]$. Then $0 \leqslant a_i \leqslant 2$ for every i, and $\sum_{i=1}^t a_i \leqslant 2m$;

By Lemma 2.10 it follows that, if B' is chosen randomly from $[t]^{(l)}$,

$$\mathsf{P}\left[\left\|\sum_{i\in B'} \mathbf{u}_i \Big|_{A_j}\right\|_p^p \geqslant \delta l\right] \leqslant 2(4em/\delta t)^{\delta l}.$$

However, $t \geqslant m \log_2 m/\delta$, so $l = t/m \geqslant \delta^{-1} \log_2 m$. It follows easily that the probability above is less than $1/m$. This contradiction establishes the result. $\square$

We note here that if one picks $n^{\alpha/2}$ vectors from each set of the form $\{\mathbf{e}_i : i \in A_j\}$, then the subbasis consisting of those vectors is $(1+\epsilon)$-equivalent to the unit vector basis of ℓ_∞^n and has cardinality $n^{(1+\alpha)/2}$.

We shall now turn to the main theorem of this chapter. For given $1 < p < \infty$ and $k, m \in \mathbb{N}$, we shall set $n = m^k$ and consider the n-dimensional normed space

$$X = ((\ldots((\ell_1^m)_p^m)_{p^2}^m \ldots)_{p^{k-2}}^m)_\infty^m .$$

We shall obtain an upper bound for the size of the largest almost symmetric block basis of X, and we will later use our result to obtain an upper bound for a different compound ℓ_p-space Y whose standard basis is C-equivalent to the standard basis of ℓ_∞^n.

In order to avoid an excessively bulky proof, we shall develop some notation before stating and proving the theorem. To begin with, let $[n]$ be identified with $[m]^k$ in the natural way, i.e.

$$(i_1, i_2, \ldots, i_k) \equiv \sum_{j=1}^{k-1} (i_j - 1)m^{k-j} + i_k$$

for $1 \leqslant i_1, \ldots, i_k \leqslant m$. Then, given $1 \leqslant j \leqslant k$, and $\mathbf{i} = (i_1, \ldots, i_j) \in [m]^j$, let

$$A_{\mathbf{i}} = A_{i_1,\ldots,i_j} = \left\{(i'_1, \ldots, i'_k) \in [m]^k : i'_1 = i_1, \ldots, i'_j = i_j\right\}.$$

The collection of sets $\{A_{\mathbf{i}} : \mathbf{i} \in [m]^j\}$ is a partition $\mathcal{P}_j$ of $[n]$. Although it will not be relevant, let us note that the partitions $\mathcal{P}_0 \equiv \{[n]\}, \mathcal{P}_1, \ldots, \mathcal{P}_k$ form a filtration of $[n]$. The point about this filtration is that if we restrict X to the subspace generated by $\{\mathbf{e}_{i'} : i' \in A_{\mathbf{i}}\}$ for any $\mathbf{i} \in [m]^k$, then we obtain a copy of ℓ_1^m, and similarly, if we restrict to the subspace generated by $\{\mathbf{e}_{i'} : i' \in A_{\mathbf{i}}\}$ for any $1 \leqslant j \leqslant k$ and $\mathbf{i} \in [m]^j$, we obtain a copy of $(\ldots((\ell_1^m)_p^m)_{p^2}^m \ldots)_{p^{k-j}}^m$.

Now, given $1 \leqslant j \leqslant k-1$, let us define a space X_j by

$$X_j = ((\ldots((\ell_{p^j-1}^{m^j})_{p^j}^m)_{p^j+1}^m \ldots)_{p^k-2}^m)_\infty^m$$

and let us set $X_k = \ell_\infty^n$. Note that $X_1 = X$.

If $\mathbf{u}$ is a vector in X, we define a sequence $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(k)}$ of associated vectors as follows. Given $1 \leqslant j \leqslant k$ and $\mathbf{i} \in [m]^{k-j}$, let $r_0 = \min(\mathrm{supp}(\mathbf{u}) \cap A_{\mathbf{i}})$ and let the restriction of $\mathbf{u}^{(j)}$ to $A_{\mathbf{i}}$ be given by

$$(\mathbf{u}^{(j)})_r = \begin{cases} \left\|\mathbf{u}\mid_{A_{\mathbf{i}}}\right\| & r = r_0 \\ 0 & \text{otherwise} \end{cases}$$

Thus $\mathbf{u}^{(j)}$ is a vector such that $\left|\mathrm{supp}(\mathbf{u}^{(j)}) \cap A_{\mathbf{i}}\right| \leqslant 1$ and $\left\|\mathbf{u}^{(j)}\mid_{A_{\mathbf{i}}}\right\| = \left\|\mathbf{u}\mid_{A_{\mathbf{i}}}\right\|$ for every $\mathbf{i} \in [m]^{k-j}$. We have insisted that $\mathrm{supp}(\mathbf{u}^{(j)}) \cap A_{\mathbf{i}} = \min(\mathrm{supp}(\mathbf{u}) \cap A_{\mathbf{i}})$ (or $\emptyset$) solely in order to define $\mathbf{u}^{(j)}$ explicitly.

Note the following two easy facts, which will be assumed in the proof of Theorem 7.2. If $\mathbf{u}$ is any vector in X, then, first,

$$\left\|\mathbf{u}^{(j)}\right\|_X = \left\|\mathbf{u}^{(j)}\right\|_{X_j},$$

and, secondly,

$$\|\mathbf{u}\|_X \leqslant \left\|\mathbf{u}^{(j)}\right\|_\infty \cdot (m^{1/p} m^{1/p^2} \ldots).$$

In the proof that follows, we shall, not surprisingly, use induction on j. Our inductive hypothesis will be divided into two parts which we shall call Steps j and j'. We shall show that Step j implies Step j', and then that Steps j and j' imply Step $j+1$. Some of the numbers, particularly exponents of m, can clearly be chosen to be smaller than we have chosen them, giving a better result than we in fact obtain. We have chosen the larger values in order to be able to use most of the proof again later for Theorem 7·6.

Theorem 7·2. *Let $n = m^k$, let $X = ((\dots((\ell_1^m)_p^m)_{p^2}^m \dots)_{p^{k-2}}^m)_\infty^m$ and let $\mathbf{e}_1, \dots, \mathbf{e}_n$ be the standard basis of X. Let $M \geqslant 1$ be such that $(M \log n)^{p^k} \leqslant m^{1-6/p}$. Then there is no M-symmetric block basis of $\mathbf{e}_1, \dots, \mathbf{e}_n$ of cardinality $s = m^8(\log_2 n)^k$.*

Proof. The structure of the proof is as follows. We begin by assuming that $\mathbf{u}_1, \dots, \mathbf{u}_s$ is an M-symmetric block basis of $\mathbf{e}_1, \dots, \mathbf{e}_n$. We then use an induction argument to deduce from this assumption a statement which is easily seen to be false. This contradiction shows that there cannot have been an M-symmetric block basis of cardinality s in the first place.

So let us assume, then, that $\mathbf{u}_1, \dots, \mathbf{u}_s$ is an M-symmetric block basis of $\mathbf{e}_1, \dots, \mathbf{e}_n$. We must standardize our basis a little, as we did in the proof of Theorem 7·1. Since the proof is easy and we have more or less given it in the proof of Theorem 7·1, we shall merely state the conclusion.

Claim. There exists a set $A \subset [s]$ of cardinality $t = s/(\log_2 n)^k = m^8$ and positive real numbers $h_1, \dots, h_k$ such that, for every $i' \in A$ and every $1 \leqslant j \leqslant k$,

$$h_j/2 \leqslant \max_{\mathbf{i} \in [m]^{k-j}} \left\| \mathbf{u}_{i'} \Big|_{A_\mathbf{i}} \right\|_{p^{j-1}}^{p^{j-1}} \leqslant h_j.$$

We shall assume, without loss of generality, that the set A guaranteed to exist by our claim is actually the set $[t]$. Throughout the rest of the proof, let q be the conjugate index of p. We may now state our inductive hypothesis.

Inductive Hypothesis.

Step j If B is chosen randomly from $[t]^{(m)}$, then

$$\mathsf{P}\left[\left\|\sum_{i'\in B}\mathbf{u}_{i'}^{(j)}\right\|_{X_j} < (2(\log n)^{1/q})^{-2\beta}\left\|\sum_{i'\in B}\mathbf{u}_{i'}\right\|_X\right] \leqslant j/2k \ ,$$

where $\beta = 0$ if $j = 1$, and $p^{-1} + \ldots + p^{-(j-1)}$ otherwise.

Step j' For any $\mathbf{i} \in [m]^{k-j}$,

$$\left\|\sum_{i'=1}^{t}\mathbf{u}_{i'}^{(j)}\ \Big|_{A_\mathbf{i}}\right\|_{p^{j-1}}^{p^{j-1}} \leqslant m^5 h_j \ .$$

Note that if $j = 1$ then Step j is trivial, and Step j' follows, with room to spare, from the fact that $|A_\mathbf{i}| = m$ when $\mathbf{i} \in [m]^{k-1}$, the supports of the $\mathbf{u}_{i'}^{(1)}$ are disjoint and $\left\|\mathbf{u}_{i'}^{(1)}\ |_{A_\mathbf{i}}\right\| \leqslant h$. Let us now show how to deduce from our inductive hypothesis the corresponding pair of statements with j replaced by $j+1$. As we said earlier, we shall in fact show that Step j implies Step j' and that Steps j and j' imply Step $j+1$.

Proof that Step j implies Step j'. Suppose then that Step j is false and let $\mathbf{i} = (i_1, \ldots, i_{k-j}) \in [m]^{k-j}$ be such that

$$\left\|\sum_{i'=1}^{t}\mathbf{u}_{i'}^{(j)}\ \Big|_{A_\mathbf{i}}\right\|_{p^{j-1}}^{p^{j-1}} > m^5 h_j \ .$$

Then there is certainly a set C of cardinality $m^2/2$ such that

$$\max_{1\leqslant i_{k-j+1}\leqslant m}\left\|\mathbf{u}_{i'}^{(j-1)}\ \Big|_{A_{i_1,\ldots,i_{k-j},i_{k-j+1}}}\right\|_{p^{j-2}}^{p^{j-2}} > (2m)^{-4/p}h_j^{1/p}$$

for every $i' \in C$, since otherwise we would have

$$\begin{aligned}\left\|\sum_{i'=1}^{t}\mathbf{u}_{i'}^{(j)}\ \Big|_{A_\mathbf{i}}\right\|_{p^{j-1}}^{p^{j-1}} &\leqslant m^5 h_j/2 + (2m)^{-4} m h_j t \\ &\leqslant m^5 h_j \ .\end{aligned}$$

By the pigeonhole principle, there exists a subset $B \subset C$ of cardinality $m/2$ and $i_{k-j+1} \in [m]$ such that

$$\left\| \mathbf{u}_{i'}^{(j-1)} \big|_{A_{i_1,\ldots,i_{k-j},i_{k-j+1}}} \right\|_{p^{j-2}}^{p^{j-2}} > (2m)^{-4/p} h_j^{1/p}$$

for every $i' \in B$. It follows that

$$\left\| \sum_{i' \in B} \mathbf{u}_{i'}^{(j-1)} \big|_{A_{i_1,\ldots,i_{k-j+1}}} \right\|_{p^{j-2}}^{p^{j-2}} > (m/2)(2m)^{-4/p} h_j^{1/p}$$

and therefore that

$$\left\| \sum_{i' \in B} \mathbf{u}_{i'} \right\|_X^{p^{j-2}} > (m/2)(2m)^{-4/p} h_j^{1/p} .$$

Since $\mathbf{u}_1, \ldots, \mathbf{u}_t$ is a 1-unconditional and M-symmetric block basis, we may deduce that, for every $B \in [t]^{(m)}$,

$$\left\| \sum_{i' \in B} \mathbf{u}_{i'} \right\|_X^{p^{j-2}} > M^{-p^{j-2}} (m/2)(2m)^{-4/p} h_j^{1/p} .$$

However, we know that $\left\| \sum_{i' \in B} \mathbf{u}_{i'}^{(j)} \right\|_{X_j}^{p^{j-2}}$ cannot possibly be greater than $h_j^{1/p}(m^{1/p} m^{1/p^2} \ldots) \leqslant h_j^{1/p} m^{2/p}$. The condition that $m > m^{6/p}(M \log n)^{p^{j-2}}$ guarantees that this contradicts Step j.

Proof that Steps j and j' imply Step $j+1$. First, let us fix $\mathbf{i} = (i_1, \ldots, i_{k-j}) \in [m]^{k-j}$ and consider the behaviour of the sequence $\mathbf{u}_1^{(j)}, \ldots, \mathbf{u}_t^{(j)}$ in the space X_j restricted to the set $A_{\mathbf{i}}$. The restriction of the space X_j to $A_{\mathbf{i}}$ is a copy of $\ell_{p^{j-1}}^{m^j}$, and $|\mathrm{supp}(\mathbf{u}_{i'}^{(j)}) \cap A_{\mathbf{i}}| \leqslant 1$ for every $i' \in [t]$. Let us set $a_{i'} = \left\| \mathbf{u}_{i'}^{(j)} \big|_{A_{\mathbf{i}}} \right\|^{p^{j-1}}$ for each $i' \in [t]$. Then, for any $B \subset [t]$, we have

$$\left\| \sum_{i' \in B} \mathbf{u}_{i'}^{(j)} \big|_{A_{\mathbf{i}}} \right\|_{X_j}^{p^{j-1}} = \sum_{i' \in B} a_{i'} \qquad (1)$$

and also

$$\left\| \sum_{i'\in B} \mathbf{u}_{i'}^{(j)} \Big|_{A_{\mathbf{i}}} \right\|_{X_{j+1}}^{p^{j-1}} = \left(\sum_{i'\in B} a_{i'}^{p} \right)^{1/p} . \tag{2}$$

Let $A \subset [t]$ be the set $\{i' : a_{i'} > m^{-2}h_j\}$. Then, by (1) and Step j', the cardinality of A is at most m^7. If B is chosen randomly from $[t]^{(m)}$, then, by Lemma 2.10, or else by a direct estimate of the hypergeometric distribution from Lemma 1.4,

$$\mathsf{P}\Big[|A \cap B| > \log n\Big] \leqslant 2(4e/\log n)^{\log n} \leqslant n^{-2} .$$

Therefore, with probability at least n^{-2}, we can partition B into two sets $B_1 \cup B_2$ such that

$$\sum_{i\in B_1} a_i \leqslant (\log n)^{1/q} \left(\sum_{i\in B_1} a_i^p \right)^{1/p}$$

and

$$\sum_{i\in B_2} a_i \leqslant m^{-1} h_j .$$

Hence, with probability at least $1 - n^{-1}$, we can find such a partition for every $\mathbf{i} \in [m]^{k-j}$.

Suppose $B \in [t]^{(m)}$ is one of the sets for which such a partition does indeed exist for every $\mathbf{i} \in [m]^{k-j}$. In that case, for every $\mathbf{i}$, we can set

$$\mathbf{v}_{\mathbf{i}} = \sum_{i'\in B_1} \mathbf{u}_{i'}^{(j)} \Big|_{A_{\mathbf{i}}}$$

and

$$\mathbf{w}_{\mathbf{i}} = \sum_{i'\in B_2} \mathbf{u}_{i'}^{(j)} \Big|_{A_{\mathbf{i}}} ,$$

where of course the partition $B = B_1 \cup B_2$ depends on $\mathbf{i}$.

Now

$$\|\mathbf{v}_{\mathbf{i}}\|_{X_j}^{p^{j-1}} \leqslant (\log n)^{1/q} \left\| \sum_{i'\in B} \mathbf{u}_{i'}^{(j)} \Big|_{A_{\mathbf{i}}} \right\|_{X_{j+1}}^{p^{j-1}}$$

and

$$\|\mathbf{w_i}\|_{X_j}^{p^j-1} \leqslant m^{-1}h_j$$

and

$$\sum_{i'\in B}\mathbf{u}_{i'}^{(j)} = \sum_{\mathbf{i}\in[m]^{k-j}}\mathbf{v_i} + \sum_{\mathbf{i}\in[m]^{k-j}}\mathbf{w_i}\ .$$

It follows that

$$\left\|\sum_{\mathbf{i}\in[m]^{k-j}}\mathbf{v_i}\right\|_{X_j}^{p^j-1} \leqslant (\log n)^{1/q}\left\|\sum_{i'\in B}\mathbf{u}_{i'}^{(j)}\right\|_{X_{j+1}}^{p^j-1}$$

and that

$$\left\|\sum_{\mathbf{i}\in[m]^{k-j}}\mathbf{w_i}\right\|_{X_j}^{p^j-1} \leqslant m^{-1}h_j(m^{1/p}m^{1/p^2}\ldots) \leqslant h_j\ .$$

But

$$h_j/2 \leqslant \left\|\sum_{i'\in B}\mathbf{u}_{i'}^{(j)}\right\|_{X_{j+1}}^{p^j-1} = \left\|\sum_{i'\in B}\mathbf{u}_{i'}^{(j+1)}\right\|_{X_{j+1}}^{p^j-1}\ .$$

Therefore

$$\left\|\sum_{i'\in B}\mathbf{u}_{i'}^{(j)}\right\|_{X_j}^{p^j-1} \leqslant 2(\log n)^{1/q}\left\|\sum_{i'\in B}\mathbf{u}_{i'}^{(j+1)}\right\|_{X_{j+1}}^{p^j-1}\ .$$

It follows that this is the case with probability at least $1-n^{-1}$. Combining this fact with Step j we can clearly deduce Step $j+1$.

Proof that Step $(k-1)'$ is false. Step $(k-1)'$ states that, for any $\mathbf{i}\in[m]$,

$$\left\|\sum_{i'=1}^{t}\mathbf{u}_{i'}^{(k-1)}\Big|_{A_\mathbf{i}}\right\|_{p^{k-2}}^{p^{k-2}} \leqslant m^5h_{k-1}\ .$$

However, by the pigeonhole principle, there exists a set $C\subset[t]$ of cardinality at least t/m, and some $\mathbf{i}\in[m]$ such that

$$\left\|\mathbf{u}_{i'}^{(k-1)}\Big|_{A_\mathbf{i}}\right\|_{p^{k-2}}^{p^{k-2}} \geqslant h_{k-1}/2$$

for every $i' \in C$. This shows that

$$\left\| \sum_{i'=1}^{t} \mathbf{u}_{i'}^{(k-1)} \Big|_{A_i} \right\|_{p^{k-2}}^{p^{k-2}} \geqslant m^7 h_{k-1}/2 \,,$$

contradicting Step $(k-1)'$. However, by induction, Step $(k-1)'$ must be true if $\mathbf{u}_1, \ldots, \mathbf{u}_s$ was an M-symmetric block basis. The contradiction therefore shows that no block basis of cardinality s can be M-symmetric, completing the proof of Theorem 7.2. □

As a corollary of Theorem 7·2, we can now deduce that there exists a basis C-equivalent to the standard basis of ℓ_∞^n whose largest $(1+\epsilon)$-symmetric block basis is of cardinality a power of n which tends to zero as $\log(1+\epsilon)/\log C$ tends to zero.

Corollary 7·3. *Let $\epsilon > 0$, $C > 1$, let $n \in \mathbb{N}$ and set $\alpha = \log(1+\epsilon)/\log C$ and $k = \log_{12}(1/8\alpha)$. Then there exists a basis C-equivalent to the unit vector basis of ℓ_∞^n with no $(1+\epsilon)$-symmetric block basis of cardinality $n^{8/k}(\log_2 n)^k$.*

Proof. Let $p = 12$, let $m = n^{1/k}$ and let $r = 2\log n/k\log C$. We shall show that the standard basis of

$$Y \equiv ((\ldots((\ell_r^m)_{pr}^m)_{p^2r}^m \ldots)_{p^{k-2}r}^m)_\infty^m$$

has the required properties.

First, this basis is C'-equivalent to the unit vector basis of ℓ_∞^n, where

$$C' \leqslant m^{(k/2)\log C(\log n)^{-1}(1+p^{-1}+p^{-2}+\ldots+p^{-(k-2)})} \leqslant e^{\log C} = C \,.$$

Let us suppose that $\mathbf{u}_1, \ldots, \mathbf{u}_s$ is a $(1+\epsilon)$-symmetric block basis of this basis. We shall consider the r-concavification of Y. Indeed, let

$$X = ((\ldots((\ell_1^m)_p^m)_{p^2}^m \ldots)_{p^{k-2}}^m)_\infty^m$$

and let $\mathbf{v}_i = |\mathbf{u}_i|^r$ for each $1 \leqslant i \leqslant s$. Then X is the r-concavification of Y, since, for any sequence $\lambda_1, \ldots, \lambda_s$ of scalars,

$$\left\| \sum_{i=1}^{s} |\lambda_i|^r \mathbf{v}_i \right\|_X = \left\| \sum_{i=1}^{t} \lambda_i \mathbf{u}_i \right\|^r .$$

It follows that $\mathbf{u}_1, \ldots, \mathbf{u}_s$ is $(1+\epsilon)$-symmetric if and only if $\mathbf{v}_1, \ldots, \mathbf{v}_s$ is M-symmetric, where $M = (1+\epsilon)^r = n^{2\alpha/k}$. However, by Theorem 7·4, $\mathbf{v}_1, \ldots, \mathbf{v}_s$ is not M-symmetric if

$$n^{2\alpha k^{-1} p^k} (\log n)^{p^k} \leqslant m^{1-6/p} = n^{1/2k} ,$$

and $s = m^8 (\log_2 n)^k$. We have assumed the second condition, and, since $k = \log_{12}(1/8\alpha)$, we have

$$n^{2\alpha k^{-1} p^k} (\log n)^{p^k} \leqslant n^{4\alpha k^{-1} p^k} = n^{4\alpha k^{-1}/8\alpha} = n^{1/2k} .$$

This proves the corollary. □

The next corollary we deduce directly from Theorem 7·2, since this is more natural than checking that the proof of Corollary 7.3 is valid when $C = n$.

Corollary 7·4. *There exists an absolute constant c such that, for every $n \in \mathbb{N}$, there is a 1-unconditional basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ such that its largest 2-symmetric block basis has cardinality at most $n^{c/\log\log n}$.*

Proof. Let $k = \log\log n / 2 \log 12$, let $p = 12$ and let $m = n^{1/k}$. Then

$$(2 \log n)^{p^k} = (2 \log n)^{\sqrt{\log n}} \leqslant n^{\log 12 / \log\log n} = m^{1-6/p} .$$

Let X be as in Theorem 7·2 and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be its standard basis. By Theorem 7·2, the largest 2-symmetric block basis of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ has cardinality at most $m^8 (\log_2 n)^k$, which is at most $n^{40/\log\log n}$ provided n is large enough. The condition $\log\log n > 40$ is certainly enough to guarantee this. □

It is reasonable to ask whether there is any easy improvement of Corollary 7·4. One could, for example, consider a different compound ℓ_p-space, and it is not even obvious whether the proof of Theorem 7·2 gives the correct bound for the space we did consider. We shall therefore give a simple argument which shows that an improved estimate in Corollary 7·4 would have to come from a different construction altogether.

Proposition 7·5. *Let $k \in \mathbb{N}$, let $m_1, \ldots, m_k$ be integers and let $p_1, \ldots, p_k$ be numbers in the closed interval $[1, \infty]$. Let X be the space $(\ldots(\ell_{p_1}^{m_1})_{p_2}^{m_2} \ldots)_{p_k}^{m_k}$, let $n = \prod_{i=1}^k m_i$ and let $\epsilon > 0$. Then the standard basis of X has a $(1+\epsilon)$-symmetric block basis of cardinality $\theta_1(\epsilon) n^{c/\theta_2(\epsilon)+\log\log n}$, where c is an absolute constant, and $\theta_1(\epsilon)$ and $\theta_2(\epsilon)$ depend only on ϵ.*

Proof. Let $l = \lceil \log_{3/2}(\log n / \log(1+\epsilon)) \rceil + 1 = c' \log\log n + \theta_2$ for some absolute constant c' and some constant $\theta_2 = \theta_2(\epsilon)$ which depends only on ϵ. For $1 \leqslant j \leqslant l$, let I_j be the real interval given by

$$I_j = \begin{cases} [(3/2)^{j-1}, (3/2)^j) & 1 \leqslant j < l \\ [(3/2)^{l-1}, \infty] & j = l\,. \end{cases}$$

Then, by our choice of l, $\bigcup_{j=1}^l I_j = [1, \infty]$.

By an obvious averaging argument, we can find a subset $A \subset [k]$ and $j \in [l]$ such that $p_i \in I_j$ for every $i \in A$ and such that $\prod_{i \in A} m_i \geqslant n^{1/l}$. Let $A = \{i_1, \ldots, i_r\}$, let $Y = (\ldots(\ell_{p_{i_1}}^{m_{i_1}})_{p_{i_2}}^{m_{i_2}} \ldots)_{p_{i_r}}^{m_{i_r}}$ and set $m = \dim(Y) = \prod_{i \in A} m_i \geqslant n^{1/l}$. The standard basis of Y is 1-equivalent to a subbasis of the standard basis of X. If $j = l$, then we know in addition that this basis is $(1+\epsilon)$-equivalent to the unit vector basis of ℓ_∞^m, in which case we are done. If $1 \leqslant j < l$, then the space Y is $(3/2)^{j-1}$-convex. Let Y' be the $(3/2)^{j-1}$-concavification of Y and let $\mathrm{e}_1, \ldots, \mathrm{e}_m$ be the standard basis of Y'. Note that Y is isometric to $(\ldots(\ell_{q_{i_1}}^{m_{i_1}})_{q_{i_2}}^{m_{i_2}} \ldots)_{q_{i_r}}^{m_{i_r}}$, where $q_{i_s} = (3/2)^{-(j-1)} p_{i_s}$ for $1 \leqslant s \leqslant r$.

As in the proof of Corollary 7·3, there exists a $(1+\epsilon)^{(3/2)^{j-1}}$-symmetric block basis of $e_1, \ldots, e_m$ of a given cardinality if and only if there exists a $(1+\epsilon)$-symmetric block basis of the standard basis of Y of the same cardinality. It will therefore certainly be enough for us to find a reasonably large $(1+\epsilon)$-symmetric block basis of the standard basis of Y'. But $1 \leqslant q_{i_s} \leqslant 3/2$ for every $1 \leqslant s \leqslant r$, from which it follows that $\|\sum_{h=1}^{m} e_h\| \geqslant m^{2/3}$. By Theorem 3.1, it follows that $e_1, \ldots, e_m$ has a $(1+\epsilon)$-symmetric block basis of cardinality $\theta_1 m^{1/3}/\log m > \theta_1 m^{1/4} = \theta_1 n^{1/4l}$, where $\theta_1 = \theta_1(\epsilon)$ depends only on ϵ. This proves the proposition.

□

The next result is a strengthening of Corollary 7·4. In that corollary, we constructed a basis with no almost symmetric block basis of size any fixed power of n. We shall now show that a space of the kind we have been considering actually contains no subspace with a uniformly symmetric basis of dimension a fixed power of n. The proof of this result is very similar to the proof of Theorem 7·2. In fact, we shall prove two steps, and obtain a basis which satisfies Steps 1 and 1′ in the proof of Theorem 7·2, with a slightly different value of M. The rest of the proof will then be a corollary of the proof of Theorem 7·2. Even the method used to get to Steps 1 and 1′ will resemble our previous arguments. Note that the dependence of our estimate on M in the next result is via the condition that n should be sufficiently large.

Theorem 7·6. *Let $M \geqslant 1$ and $p = 12$. Let $n \in \mathbb{N}$, $k = \log\log n / 2\log 12$ and $m = n^{1/k}$. Then the space*

$$X = ((\ldots(\ell_{2p}^m)_{2p^2}^m \ldots)_{2p^{k-1}}^m)_\infty^m$$

has no subspace of dimension $s = m^8(\log_2 n)^k \leqslant n^{40/\log\log n}$ with an M-symmetric basis.

Proof. Suppose that $\mathbf{u}_1, \ldots, \mathbf{u}_s$ is an M-symmetric basic sequence, and, for each $1 \leqslant i \leqslant s$, let the coordinates of $\mathbf{u}_i$ (with respect to the standard basis of X) be $u_{i1}, \ldots, u_{in}$. Let us define a correspondence between $[m]^k \times [s]$ and $[sn]$ by setting

$$(i_1, i_2, \ldots, i_k, i_{k+1}) \equiv s \sum_{j=1}^{k} (i_j - 1) m^{k-j} + i_k$$

for $1 \leqslant i_1, \ldots, i_k \leqslant m$ and $1 \leqslant i_{k+1} \leqslant s$. Given $1 \leqslant j \leqslant k$, and $\mathbf{i} \in [m]^j$, let $A_{\mathbf{i}}$ be defined as before. That is, $A_{\mathbf{i}}$ is the set of sequences in $[m]^k \times [s]$ that begin with $\mathbf{i}$. Let $[m]^k$ be identified with $[n]$ as it was in the proof of Theorem 7·4. No ambiguity will arise between these two correspondences.

Let us define a new sequence of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_s$ in the sn-dimensional space

$$X' = ((\ldots((\ell_2^s)_{2p}^m)_{2p^2}^m \ldots)_{2p^{k-2}}^m)_\infty^m$$

as follows. If $\mathbf{i} = (i_1, \ldots, i_k) \in [m]^k$, then $(i_1, \ldots, i_k, i_{k+1})^{\text{th}}$ coordinate of $\mathbf{v}_{i'}$ is the $(i_1, \ldots, i_k)^{\text{th}}$ coordinate of $\mathbf{u}_{i'}$ if $i_{k+1} = i'$, and zero otherwise. The sequence $\mathbf{v}_1, \ldots, \mathbf{v}_s$ is a sort of disjointification of $\mathbf{u}_1, \ldots, \mathbf{u}_s$. The main idea of the proof is to replace $\mathbf{u}_1, \ldots, \mathbf{u}_s$ by $\mathbf{v}_1, \ldots, \mathbf{v}_s$ and then use the technique of Theorem 7·4. This will be carried out in two steps. First, as ever, we need to standardize our basis: in this case we standardize the sequence $\mathbf{v}_1, \ldots, \mathbf{v}_s$. We will not prove the next statement, since it is very simple and we have seen several similar statements already.

Claim. There exists a subset $A \subset [s]$ of cardinality $t = s/(\log_2 n)^k = m^8$ and a sequence of positive real numbers $h_0, h_1, \ldots, h_k$ such that, for every $i' \in A$ and every $1 \leqslant j \leqslant k$,

$$h_j/2 \leqslant \max_{\mathbf{i} \in [m]^{k-j}} \left\| \mathbf{v}_{i'} \Big|_{A_{\mathbf{i}}} \right\|_{2p^{j-1}}^{2p^{j-1}} \leqslant h_j \,.$$

As usual, we shall assume, without loss of generality, that $A = [t]$. We shall also assume, for the sake of convenience, that $h_0 = 1$.

Step A. Let $\lambda_1, \dots, \lambda_t$ be any sequence of scalars. Then

$$\left\|\sum_{i=1}^{t} \lambda_i \mathbf{v}_i\right\|_{X'}^2 \leqslant \mathsf{E}\left\|\sum_{i=1}^{t} \pm\lambda_i \mathbf{u}_i\right\|_{X}^2 \leqslant 20\log n \left\|\sum_{i=1}^{t} \lambda_i \mathbf{v}_i\right\|_{X'}^2 ,$$

where the expectation is taken over all possible choices of signs, each with equal probability.

Sketch Proof of Step A. Since we have had many similar arguments already, our sketch of this will be extremely rudimentary. Let us fix $k \in [n]$ and, for $1 \leqslant i \leqslant t$, let a_i be the k^{th} coordinate of $\mathbf{u}_i$. Then, if $(\epsilon_i)_1^t \in \{-1,1\}^t$ is chosen at random, we have, by an easy application of Azuma's inequality (Theorem 1.1), that

$$\mathsf{P}\left[\sum_{i=1}^{t} \epsilon_i \lambda_i a_i \geqslant 10(\log n)^{1/2}\left(\sum_{i=1}^{t} \lambda_i^2 a_i^2\right)^{1/2}\right] \leqslant n^{-2} .$$

From this it follows easily that

$$\mathsf{E}\left\|\sum_{i=1}^{t} \epsilon_i \lambda_i \mathbf{u}_i\right\|_{X}^2 \leqslant 200\log n \left\|\sum_{i=1}^{t} \lambda_i \mathbf{v}_i\right\|_{X'}^2 .$$

The other inequality is an easy consequence of Jensen's inequality.

Step B. For any $\mathbf{i} \in [m]^k$,

$$\left\|\sum_{i'=1}^{t} \mathbf{v}_{i'}\Big|_{A_{\mathbf{i}}}\right\|_2^2 \leqslant m^5 .$$

Proof of Step B. Suppose not, and let $\mathbf{i} \in [m]^{(k)}$ be such that $\left\|\sum_{i'=1}^t \mathbf{v}_{i'}\big|_{A_{\mathbf{i}}}\right\|_2^2 > m^5$. Then, by the definition of $\mathbf{v}_1, \dots, \mathbf{v}_t$, we have $\sum_{i'=1}^t |u_{i'k}|^2 > m^5$, where $k = \mathbf{i}$. Since $h_0 = 1$, we know that $|u_{i'k}| \leqslant 1$ for each i'. It follows that $\sum_{i'=1}^t |u_{i'k}| > m^5$, from which we may deduce that

$$\left\|\sum_{i=1}^{t} \operatorname{sign}(u_{ik})\mathbf{u}_i\right\| > m^5 .$$

Since $\mathbf{u}_1, \ldots, \mathbf{u}_t$ is an M-symmetric sequence, we obtain that

$$\mathsf{E}\left\|\sum_{i=1}^{t} \epsilon_i \mathbf{u}_i\right\| > M^{-1} m^5 .$$

However, we also know that

$$\left\|\sum_{i=1}^{t} \operatorname{sign}(u_{ik})\mathbf{v}_i\right\| = \left\|\sum_{i=1}^{t} \mathbf{v}_i\right\| \leqslant t^{1/2} m^{2/p} = m^{4+2/p} .$$

Since $(20M \log n) < m^{1-2/p}$, this is a contradiction of Step A.

Let us now see why this theorem follows from the proof of Theorem 7·2. Let X'' be the 2-concavification of X', and let $\mathbf{w}_i = |\mathbf{v}_i|^2$, for each $1 \leqslant i \leqslant t$. Then $X'' = ((\ldots((\ell_1^s)_p^m)_{p^2}^m \ldots)_{p^{k-1}}^m)_\infty^m$ and, for any sequence of scalars $\lambda_1, \ldots, \lambda_t$, we have

$$\left\|\sum_{i=1}^{t} \lambda_i^2 \mathbf{w}_i\right\|_{X''} = \left\|\sum_{i=1}^{t} \lambda_i \mathbf{v}_i\right\|_{X'}^2 .$$

Using Steps A and B where necessary, the following facts are now obvious:

(i) $\mathbf{w}_1, \ldots, \mathbf{w}_t$ is a $400M^2(\log n)^2$-symmetric block basis of the standard basis of X'';

(ii) for every $0 \leqslant j \leqslant k-1$ and $1 \leqslant i' \leqslant t$,

$$h_j/2 \leqslant \max_{\mathbf{i}\in[m]^{k-j}} \left\|\mathbf{w}_{i'}\Big|_{A_{\mathbf{i}}}\right\|_{p^{j-1}}^{p^{j-1}} \leqslant h_j ;$$

(iii) for every $\mathbf{i} \in [m]^{(k)}$,

$$\left\|\sum_{i'=1}^{t} \mathbf{w}_{i'}\Big|_{A_{\mathbf{i}}}\right\|_1 \leqslant m^5 = m^5 h_0 .$$

Strictly speaking, we are in a slightly different situation from the beginning of the inductive process in the proof of Theorem 7·2 because X'' is a direct sum of spaces of the form ℓ_1^s rather than ℓ_1^m. However, the only place where the restriction on the dimension was used was in proving Step 1′. This is the reason we proved Step B above. The rest of the proof can now be used word for word, with $400M^2(\log n)^2$ replacing M, to contradict the fact (i) above, provided that

$$\left(400M^2(\log n)^2\right)^{12^k} \leqslant m^{1/2} = n^{1/2k} .$$

By our choice of k, the left hand side is equal to $(400M^2(\log n)^2)^{\sqrt{\log n}}$ and the right hand side is equal to $12^{\log n/\log\log n}$, which is clearly larger. □

It seems unlikely that Theorem 7·6 is best possible. In the light of Proposition 7·5, the best way to obtain an improved bound might be to construct a space using iterated direct sums as before, but in such a way as to avoid large-dimensional subspaces which are p-convex for some large p.

The best bound in the opposite direction is due to Alon and Milman [2]. They showed that every n-dimensional space has a k-dimensional subspace which is $(1+\epsilon)$-isomorphic to either ℓ_2^k or ℓ_∞^k, where k is proportional to $\exp(\sqrt{\log n})$.

Theorem 7·2 is, of course, even less likely to be best possible. Given an arbitrary basis, or even a 1-unconditional basis, the only known result which gives an almost symmetric block basis seems to be Krivine's theorem. The proof of Krivine's theorem requires a condition that the basis be slightly "spread out". One could use a result of Amir and Milman [3] for this. They have shown that an arbitrary basis has an almost-unconditional block basis of cardinality proportional to $\sqrt{\log n}$. Since the proof of Krivine's theorem uses Ramsey's theorem, the lower bound one obtains is extremely small and not unduly affected by dropping down to $\sqrt{\log n}$ as a first step. It is very hard to believe that this bound could not be improved dramatically, in which case it would be more sensible to "spread the vectors out" by using Lemma 4.4 or something similar, with respect to the Fritz John ellipsoid of the space generated by the basis.

It seems likely that the construction of Theorem 7.6 can be modified to yield a number of corollaries of some interest. We shall give two simple ones. First, we solve a problem mentioned at the end of Chapter 5.

Corollary 7·7. *Let $M \geqslant 1$ and let n be a power of 2. Then there there exists a normalized basis $x_1, \ldots, x_n$ satisfying a lower 2-estimate with no M-symmetric block basis of cardinality $n^{40/\log\log n}$.*

Proof. Let $p = 12$, $k = \log\log n/2\log 12$ and $m = n^{1/k}$, and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis of the space $X = ((\ldots(\ell_{2p}^m)_{2p^2}^m \ldots)_{2p^{k-1}}^m)_\infty^m$ used in Theorem 7.6. Let $A = (a_{ij})_{i,j=1}^n$ be the $n \times n$-Walsh matrix (this is defined in Chapter 9) and, for $i = 1, \ldots, n$, let $x_i' = \sum_{j=1}^n a_{ij}\mathbf{e}_j$ and let $x_i = x_i'/\|x_i'\|$. It is easy to check that the basis $x_1, \ldots, x_n$ is normalized and satisfies a lower 2-estimate.

However, by Theorem 7.6, X contains no M-symmetric basic sequence of cardinality $n^{40/\log\log n}$. The corollary follows trivially from this. □

Let us remark here that one can get a partial result about bases with a lower p-estimate, for $p < 2$, by defining a norm $\|.\|$ on $\mathbb{R}^n$ as follows. If $\mathbf{a} = \sum_{i=1}^n a_i\mathbf{e}_i \in \mathbb{R}^n$, one sets $\|\mathbf{a}\| = \|\sum_{i=1}^n a_i x_i\|_X \vee \|\mathbf{a}\|_p$. Then one uses the fact that the standard basis of $\mathbb{R}^n$ under this norm is $n^{1/p-1/2}$-equivalent to the basis $x_1, \ldots, x_n$. Setting $\beta = 1/p - 1/2$, one can obtain this way a basis satisfying a lower p-estimate with no 2-symmetric block basis of cardinality $n^{c/\log(\beta^{-1})}$, for some absolute constant c. However, compared with the known lower bound of $n^{2\beta}$ obtained in Theorem 3.1, this is very large, and it gives no information at all unless p is rather close to 2.

The next corollary is very simple. To prove it, one simply considers the space $(\ldots(\ell_{2p}^m)_{2p^2}^m \ldots)_{2p^l}^m$ as a subspace of X, for $l = \lfloor \log_{12} q \rfloor$.

Corollary 7·8. *There exist absolute constants c, c' such that, for every $q > e^c$ and $M \geqslant 1$, and for n sufficiently large, there exists an n-dimensional space with*

q-cotype constant at most c', with no subspace of dimension $n^{c/\log q}$ with an M-symmetric basis. □

Note also that this space has uniformly bounded 2-type constant, so we have a partial converse to Theorem 4.6.

The results of this chapter suggest that the key problem concerning upper estimates for the sizes of almost symmetric block bases is to obtain a best possible improvement of Theorem 7.6. From this it is likely that one would be able to obtain best possible results to do with bases close to the standard basis of ℓ_∞^n, bases with lower p-estimates, spaces with type or cotype conditions and probably spaces with other natural properties as well.

This concludes Chapter 7, and indeed the part of this dissertation which deals with symmetric sequences in finite-dimensional normed spaces. In the next chapter, we shall look a little at infinite symmetric sequences, but before that we shall summarize the results of Chapters 3 to 7.

SUMMARY OF RESULTS

AND SOME OPEN PROBLEMS

In Chapters 3 to 7 we have considered various properties that a basis can have. For each property P, we have asked the following question. If a basis of cardinality n satisfies property P, then what is the largest $(1+\epsilon)$-symmetric block basis that it is guaranteed to have? That is, if a sequence satisfies P, what is the smallest possible size of its largest $(1+\epsilon)$-symmetric block basis? We have been interested in both lower bounds and upper bounds for this quantity. We shall now summarize what is known about various such problems, and also about some of the similar problems obtained by replacing "$(1+\epsilon)$-symmetric" with "$(1+\epsilon)$-unconditional" in the question above. The open questions we discuss are those particularly related to our work: there are several others which are natural and interesting in the light of the work of Amir and Milman.

Given a property P, then, and given $n \in \mathbb{N}$ and $\epsilon > 0$, let $P_S(n,\epsilon)$ (respectively $P_U(n,\epsilon)$) be the *smallest* possible size of the *largest* $(1+\epsilon)$-symmetric (respectively $(1+\epsilon)$-unconditional) block basis of a basis satisfying P. We shall deal with various properties in turn. For the whole of this summary, we shall assume that $x_1,\ldots,x_n$ is the basis under discussion.

1. Let us begin by taking AG to be the property that $x_1,\ldots,x_n$ is a sequence of unit vectors and

$$\mathsf{E}\left\|\sum_{i=1}^{n} \pm x_i\right\| \geqslant n^{1/p}$$

where $1 \leqslant p < 2$ and the expectation is over all possible choices of signs, each with equal probability. Also, let I be the property that $x_1,\ldots,x_n$ is a linearly independent sequence.

Then, by the results of Chapters 3 and 5, we have

$$c(\epsilon)n^{2/p-1}/\log n \leqslant AG_S(n,\epsilon) \leqslant AG_U(n,\epsilon) \leqslant 2n^{2/p-1} .$$

If $1 \leqslant p \leqslant 3/2$, then

$$c(\epsilon)n^{2/p-1}/\log n \leqslant (AG \cap I)_S(n,\epsilon) \leqslant (AG \cap I)_U(n,\epsilon) \leqslant c'(\epsilon)n^{2/p-1}(\log n)^{4/3} .$$

Although we have not done this, it is possible to use Theorem 7.2 to construct an independent sequence which shows that, when $3/2 < p \leqslant 2$, we still have that $(AG \cap I)_S \leqslant n^{2/p-1+o(n)}$. This suggests the problem of estimating $(AG \cap I)_U(n,\epsilon)$ from above. We cannot improve on $c(\epsilon)n^{1/3}(\log n)^{4/3}$. Another problem suggested by the results of Chapter 3 is the following. Suppose that $3/2 < p \leqslant 2$ and LE is the property that $x_1, \ldots, x_n$ is a sequence of unit vectors, and, for any sequence of scalars $a_1, \ldots, a_n$,

$$\left\| \sum_{i=1}^{n} a_i x_i \right\| \geqslant \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p} .$$

Then what are $LE_S(n,\epsilon)$ and $LE_U(n,\epsilon)$? The results of Chapters 3 and 5 show only that the exponent of n lies between $n^{2/p-1}$ and $n^{1/3}$. In Chapter 7, a partial answer is obtained for $LE_S(n,\epsilon)$. In particular, it is shown that $LE_S(n,\epsilon)$ is bounded above by $n^{c/\log\log n}$ for an absolute constant c, when $p = 2$.

2. Let LP be the property that $x_1, \ldots, x_n$ is C-equivalent to the unit vector basis of ℓ_p^n, where $1 \leqslant p < \infty$. We showed in Chapter 4 that

$$c(\epsilon,p,C)n/\log n \leqslant LP_S(\epsilon,n) \leqslant LP_U(n,\epsilon)$$

and in Chapter 6 that

$$LP_S(n,\epsilon) \leqslant \begin{cases} c'(\epsilon,p)n\log\log n/\log n & \text{if } p > 1 \\ c'(\epsilon)n/\log\log n & \text{if } p = 1 . \end{cases}$$

We have not checked the details, but it seems that Theorems 6.6 and 6.11 could be modified fairly easily to show that the upper bound for $LP_S(n, \epsilon)$ is also an upper bound for $LP_U(n, \epsilon)$. Therefore, the main question still open is what the correct bound is when $p = 1$. Less interesting is the question of whether the $\log \log n$-factor is necessary in the upper bound when $p > 1$. It was needed for technical reasons, and can be dropped if one only wants to rule out large bases with ± 1-coefficients. It seems almost certain that the lower bound is best possible when $p > 1$, up to the constant $c(\epsilon, p, C)$. One final question which is relevant to the case $p = 1$ is the following. Suppose we fix a sequence $\mathbf{a} = (a_1, \ldots, a_m) \in \mathbb{R}^m$. Then how large must n be so that any basis satisfying LP must have a block basis of cardinality m which is $(1+\epsilon)$-symmetric at $\mathbf{a}$? Our methods show only that, up to a constant depending on ϵ, p and C, the answer lies between m and $m \log m$.

3. Let LI be the property that $x_1, \ldots, x_n$ is C-equivalent to the standard basis of ℓ_∞^n. Let us fix $\epsilon > 0$ and set $\alpha = \log(1+\epsilon)/\log C$. Then it follows from the well known finite-dimensional version of a result of James that $LI_U(n, \epsilon) \geqslant LI_S(n, \epsilon) \geqslant n^{\alpha/2}$. We have shown in Chapter 7 that $LI_S(n, \epsilon) \leqslant n^\beta$, where $\beta = c/\log(\alpha^{-1})$ for some absolute constant c. We have no interesting information about $LI_U(n, \epsilon)$.

If we let U be the property that $x_1, \ldots, x_n$ is a 1-unconditional sequence, then we have $U_S(n, \epsilon) \leqslant n^{c/\log \log n}$ for some absolute constant c. We have no interesting lower bound in this case. By Krivine's theorem, $U_S(n, \epsilon)$ does at least tend to infinity with n, but this result uses Ramsey's theorem and therefore almost certainly gives a very poor estimate. It is quite likely that results of Amir and Milman could be used to obtain a better bound without too much further work, but we have not investigated this.

A related question is the following. Suppose X is an n-dimensional normed space. Then how large a subspace must it contain with a 2-symmetric basis? We have shown, again in Chapter 7, that, for some absolute constant c, there exists for every n an n-dimensional space with no subspace of $n^{c/\log\log n}$ dimensions with a 2-symmetric basis. The best known lower bound was obtained by Alon and Milman, who showed that every n-dimensional space contains a subspace of k dimensions, where k is about $\exp(\sqrt{\log n})$, which is $(1+\epsilon)$-isomorphic to either ℓ_2^k or ℓ_∞^k. Just because this would be tidy, my guess is that this is best possible, that $U_S(n,\epsilon)$ is also about $\exp(\sqrt{\log n})$ and that $LI_S(n,\epsilon)$ is about $n^{\sqrt{\alpha}}$.

Finally, we do not know much about $I_U(n,\epsilon)$, where I, as before, is simply the property that $x_1,\ldots,x_n$ is linearly independent. Amir and Milman [3] showed that it is at least $c(\epsilon)\sqrt{\log n}$, and in Chapter 5 we showed that it is at most $c(\epsilon)n^{1/3}(\log n)^{4/3}$. This is a ridiculously large gap which we hope to make narrower in the future.

PART IV

OTHER RESULTS

CHAPTER 8

INFINITE ALMOST SYMMETRIC SEQUENCES

In Chapter 4 we proved that if $1 \leqslant p < \infty$ then any basis equivalent to the unit vector basis of ℓ_p^n has a block basis of cardinality proportional to $n/\log n$ that is almost symmetric. There is a natural infinite-dimensional analogue of this statement, namely that any infinite-dimensional Banach space that is isomorphic to ℓ_p contains an infinite almost symmetric basic sequence. Now, a famous question in Banach space theory, known as the distortion problem, is the question of whether any space which is isomorphic to ℓ_p has an infinite-dimensional subspace that is almost isometric to ℓ_p. A positive answer to this would trivially imply that such a space contained an infinite almost symmetric basic sequence. The object of this chapter is to show that the reverse implication also holds. Specifically, we show that if *every* Banach space isomorphic to ℓ_p contains an infinite $(1+\delta)$-symmetric basic sequence for every $\delta > 0$ then every Banach space isomorphic to ℓ_p contains, for every $\epsilon > 0$, an infinite-dimensional subspace which is $(1+\epsilon)$-isomorphic to ℓ_p.

The proof will be divided into two parts. The first is an adaptation of one of the standard proofs of Hindman's theorem, which we shall discuss later in the chapter. The second is a corollary of the standard proof of Krivine's theorem in the case when the original basis is 1-symmetric and equivalent to the unit vector basis of ℓ_p. The first part of the proof is perhaps interesting on its own, but to explain it we shall need some notation and a standard lemma which reformulates the distortion problem. We shall give the lemma first.

Lemma 8·1. *The following are equivalent, for $1 < p < \infty$.*

(i) Any Banach space isomorphic to ℓ_p contains, for every $\epsilon > 0$, an infinite-dimensional subspace which is $(1+\epsilon)$-isomorphic to ℓ_p.

(ii) *Let* $\|.\|$ *be any norm equivalent to the standard norm of* ℓ_p. *For every* $\epsilon > 0$ *there exists an* ℓ_p*-normalized block basis of the unit vector basis of* ℓ_p *which is* $(1+\epsilon)$*-equivalent, under* $\|.\|$, *to the unit vector basis of* ℓ_p.

(iii) *Let* $\mathcal{C} \subset S(\ell_p)$ *be any subset of the unit sphere of* ℓ_p. *For any* $\epsilon > 0$ *let* $\mathcal{C}_\epsilon$ *denote*

$$\{x \in S(\ell_p) : \|x - y\|_p \leqslant \epsilon \text{ for some } y \in \mathcal{C}\}.$$

Then for any $\epsilon > 0$ *there is an infinite normalized block basis of the unit vector basis of* ℓ_p *which generates a subspace* Y *such that either* $S(Y) \subset \mathcal{C}_\epsilon$ *or* $S(Y) \cap \mathcal{C} = \emptyset$.

Proof. Trivially (ii) implies (i). It is also easy to see that (iii) implies (ii). Indeed, let $\|.\|$ be an equivalent norm on ℓ_p. Without loss of generality there exists $C > 1$ such that $\|x\|_p \leqslant \|x\| \leqslant C\|x\|_p$ for every $x \in \ell_p$. Let $\mathcal{C} = \{x \in S(\ell_p) : \|x\| \leqslant \sqrt{C}\}$. Then given any $\delta > 0$ we certainly have $\mathcal{C}_\delta \subset \{x \in S(\ell_p) : \|x\| \leqslant \sqrt{C} + \delta C\}$. Let $\mathbf{e}_1, \mathbf{e}_2, \ldots$ be the unit vector basis of ℓ_p. If (iii) holds, then we can find a normalized block basis $\mathbf{u}_1, \mathbf{u}_2, \ldots$ of $\mathbf{e}_1, \mathbf{e}_2, \ldots$ which generates a subspace Z of ℓ_p such that either $1 \leqslant \|z\| \leqslant \sqrt{C} + \delta C$ for every $z \in S(Z)$ or $\sqrt{C} \leqslant \|z\| \leqslant C$ for every $z \in S(Z)$. By an easy iteration of this argument, we can obtain (ii). We shall complete the proof by showing that (i) implies (ii) and then that (ii) implies (iii).

Suppose then that (i) is true and let $\|.\|$ be an equivalent norm on ℓ_p. Let X be ℓ_p with the equivalent norm, let $\epsilon > 0$ and let $\delta > 0$ be such that $(1+\delta)^3 \leqslant 1+\epsilon$. By (i), there is a sequence $y_1, y_2, \ldots$ in ℓ_p which is $(1+\delta)$-equivalent to the unit vector basis of ℓ_p. We now make use of the following standard lemma (cf. [34 Proposition 1.a.12]) which is a slight extension of an observation of Bessaga and Pelczynski [8].

Lemma 8·2. *Let* $\delta > 0$ *and let* $x_1, x_2, \ldots$ *be a basis of a Banach space* X, *and let*

$y_i = \sum_{j=1}^{\infty} a_{ij} x_j$ for $i = 1, 2, \ldots$ be such that $\inf_i \|y_i\| > 0$ and $\lim_{j\to\infty} a_{ij} = 0$ for every i. Then there is a subsequence $(y_{i_k})_{k=1}^{\infty}$ of $(y_i)_{i=1}^{\infty}$ which is $(1+\delta)$-equivalent to a block basis of $x_1, x_2, \ldots$. □

It is easy to check that, if X and $y_1, y_2, \ldots$ are as above and we set $x_i = \mathbf{e}_i$ for each i, then the conditions of Lemma 8.2 are satisfied. We therefore obtain a block basis of $x_1, x_2, \ldots$ which is $(1+\delta)^2$-equivalent under $\|.\|$ to the unit vector basis of ℓ_p. By passing to a subsequence and using the pigeonhole principle, we can find $\lambda > 0$ and a block basis $\mathbf{v}_1, \mathbf{v}_2, \ldots$ which is $(1+\delta)^2$-equivalent under $\|.\|$ to the unit vector basis of ℓ_p and which in addition satisfies $\lambda \leqslant \|\mathbf{v}_i\|_p \leqslant \lambda(1+\delta)$ for each i. It is easy to check that, setting $\mathbf{u}_i = \mathbf{v}_i / \|\mathbf{v}_i\|$ for each i, we obtain a block basis $\mathbf{u}_1, \mathbf{u}_2, \ldots$ which is ℓ_p-normalized and $(1+\delta)^3$-equivalent to the unit vector basis of ℓ_p under the norm $\|.\|$.

Finally, let us suppose that (ii) is true, let $\epsilon > 0$ and let $\mathcal{C} \subset S(\ell_p)$ be a subset of the unit sphere of ℓ_p. We define an equivalent norm on ℓ_p as follows. Given any vector $y \in S(\ell_p)$, let f_y stand for the support functional at y. Then set

$$\|x\| = \sup\{|f_y(x)| : y \in \mathcal{C}\} \vee (1/2)\|x\|_p .$$

Now, since $1 < p < \infty$, the space ℓ_p is uniformly convex. We can therefore find $\delta = \delta(\epsilon) > 0$ such that, whenever $x, y \in S(\ell_p)$ and $\|x - y\| > \epsilon$, we have $\|(x+y)/2\| \leqslant 1 - \delta$. Now suppose $x \notin \mathcal{C}_\epsilon$ and $y \in \mathcal{C}$ and let $\alpha = f_y(x)$. Suppose that $\alpha > 0$. Then certainly $\alpha^{-1} f_y(\frac{x+y}{2}) \geqslant 1$. But $\|\frac{x+y}{2}\| \leqslant (1-\delta)$ and $\|f_y\| = 1$, from which it follows that $\alpha \leqslant 1 - \delta$ and hence that $\|x\| \leqslant 1 - \delta$ whenever $x \in S(\ell_p) \setminus \mathcal{C}_\epsilon$. However, by (ii) we can find a normalized block basis of the unit vector basis of ℓ_p which is $(1+\delta)$-equivalent under $\|.\|$ to the unit vector basis of ℓ_p. Let Y be the subspace generated by this normalized block basis. If (iii) is false then we can find $x, y \in S(Y)$ such that $x \notin \mathcal{C}_\epsilon$ and $y \in \mathcal{C}$. The above argument shows that $\|x\| \leqslant 1 - \delta$, and clearly $\|y\| = 1$. This contradiction proves the result. □

Note that the third form of the equivalence is no longer easy when ℓ_p is replaced by c_0. Later in the chapter we shall show that this question has a positive answer.

We shall now introduce a fair amount of notation and terminology that will greatly shorten several statements, and will emphasize the formal similarity with Hindman's theorem. First, given two vectors x, $y \in \ell_p$, we shall write $x < y$ if x and y have disjoint finite supports on the standard basis of ℓ_p and $\max(\text{supp}(x)) < \min(\text{supp}(y))$. We shall use the word "subspace" to mean "unit sphere of subspace" throughout the paper from now on. If a subspace of ℓ_p is generated by a basis $(x_i)_1^\infty$ which satisfies $x_i < x_{i+1}$ for every i, that is, by a block basis, we shall call it a block subspace. If $(x_i)_1^n$ or $(x_i)_1^\infty$ is a block basis of the standard basis, then $\langle x_1, \ldots, x_n\rangle$ or $\langle x_1, x_2, \ldots\rangle$ will stand for the block subspace it generates. We shall say $x < Y$ if Y is a block subspace and $x < y$ for every $y \in Y$. If $x < Y$, then $\langle x, Y\rangle$ will stand for the subspace generated by x and Y. Letters such as A and B will be used to stand for finite-dimensional block subspaces. We shall say $A < Y$ if $x < y$ for every $x \in A$ and $y \in Y$, and then write $\langle A, Y\rangle$ for the subspace generated by A and Y.

We shall use the following partial order on the infinite-dimensional block subspaces of ℓ_p, similar to a partial order used in Baumgartner's simplified proof [6] of Hindman's theorem [25] (cf. also [24]). If $X = \langle x_1, x_2, \ldots\rangle$ and $Y = \langle y_1, y_2, \ldots\rangle$, then we say $X \prec Y$ if $(x_i)_1^\infty$ is a block basis of $(y_i)_1^\infty$. Also inspired by Baumgartner's proof of Hindman's theorem, we say that a subset C of the unit sphere of ℓ_p is *large* on a block subspace X if $C \cap Y \neq \emptyset$ whenever $Y \prec X$. Given a subset C of the unit sphere of ℓ_p, we shall write C_ϵ as before for the set of x in the unit sphere of ℓ_p for which there exists $y \in C$ such that $\|x - y\|_p < \epsilon$. Using these definitions, we can formulate the distortion problem in the following way. If $1 < p < \infty$, is there a subset C of the unit sphere of ℓ_p and $\epsilon > 0$ such that both C and the complement of C_ϵ are large on ℓ_p? Let us now state the result which, together with the proof of Krivine's theorem in the symmetric case, will imply the equivalence of the infinite-dimensional analogue of Theorem 4.1 and a positive answer to this question.

Theorem 8·3. *Let $1 < p < \infty$. Suppose there exist $\epsilon > 0$ and a subset C of the unit sphere of ℓ_p such that C and the complement of C_ϵ are both large on ℓ_p. Then there exist $\delta > 0$ and a subset $\mathcal{D}$ of the unit sphere of ℓ_p such that $\mathcal{D}$ is large on ℓ_p, but, given any two dimensional block subspace $A = \langle x_1, x_2 \rangle$ of ℓ_p, there exists $y \in A$ such that $d(y, \mathcal{D}) \geqslant \delta$.*

For the proof of this theorem, we shall need some simple lemmas, which are similar to lemmas used by Baumgartner. First, let us introduce one more piece of shorthand. If m, $n \in \mathbb{N} \cup \{\infty\}$ and $\epsilon > 0$, we shall write $P(m, n, \epsilon)$ for the following statement:

Given any set $C \in \ell_p$, either there exists an m-dimensional block subspace X of ℓ_p such that $X \subset C_\epsilon$ or there exists an n-dimensional block subspace Y with $Y \cap C = \emptyset$.

We shall write "$P(m, n)$" for the statement "$P(m, n, \epsilon)$ for all $\epsilon > 0$". Theorem 8.3 states that $P(2, \infty)$ implies $P(\infty, \infty)$.

Lemma 8·4. *Suppose $X \prec \ell_p$ and C is large on X. Suppose also that $C = \bigcup_1^n C_i$. Then there exist $1 \leqslant i \leqslant n$ and $Y \prec X$ such that C_i is large on Y.*

Proof. We use induction. If C_n is not large on X then there exists $X' \prec X$ such that $X' \cap C_n = \emptyset$. Since C is large on X', it follows that $\bigcup_1^{n-1} C_i$ is large on X'. By induction we can find $1 \leqslant i \leqslant n-1$ and $Y \prec X$ such that C_i is large on Y.

The case $n = 1$ is trivial. □

Lemma 8·5. *Suppose $X \prec \ell_p$, $\epsilon > 0$ and C is large on X. Suppose also $P(2, \infty, \epsilon)$. Then there exist a finite-dimensional block subspace $A \subset X$ and $Y \prec X$ such that $A < Y$ and*

$$\{y \in Y : \langle x, y\rangle \subset C_\epsilon \text{ for some } x \in A\}$$

is large on Y.

Proof. Let $X = \langle x_1, x_2, \ldots\rangle$. Set $y_1 = x_1$ and $Y_1 = \langle x_2, x_3, \ldots\rangle$. Then either $\{y \in Y_1 : \langle y_1, y\rangle \subset C_\epsilon\}$ is large on Y_1 or there exists $X_1 \prec Y_1$ such that whenever $y \in X_1$, $\langle x_1, y\rangle \cap (C_\epsilon)^c \neq \emptyset$. In the first case we may set $A = A_1 \equiv \langle y_1\rangle$ and $Y = Y_1$. In the second case, let $X_1 = \langle x_{1,1}, x_{1,2}, \ldots\rangle$ and set $y_2 = x_{1,1}$, $Y_2 = \langle x_{1,2}, x_{1,3}, \ldots\rangle$ and $A_2 = \langle y_1, y_2\rangle$. Then either $\{y \in Y_2 : \langle x, y\rangle \subset C_\epsilon \text{for some } x \in A_2\}$ is large on Y_2 or there exists $X_2 \prec Y_2$ such that whenever $x \in A_2$ and $y \in Y_2$, we have $\langle x, y\rangle \cap (C_\epsilon)^c \neq \emptyset$. Continue this process. Either it terminates at some stage, in which case we are done, or we generate a block subspace $Y = \langle y_1, y_2, \ldots\rangle$ such that for any n, any $x \in \langle y_1, \ldots, y_n\rangle$ and any $y \in \langle y_{n+1}, y_{n+2}, \ldots\rangle$, we have $\langle x, y\rangle \cap (C_\epsilon)^c \neq \emptyset$. In the second case, we therefore have $\langle x, y\rangle \cap (C_\epsilon)^c \neq \emptyset$ whenever $x, y \in Y$, $x < y$. But then, by $P(2, \infty, \epsilon)$, there exists $Z \prec Y$ such that $Z \subset ((C_\epsilon)^c)_\epsilon$, that is, $Z \cap C = \emptyset$. This contradicts the assumption that C was large on X. □

Lemma 8·6. *Suppose $X \prec \ell_p$, $\epsilon > 0$, $\delta > 0$ and C is large on X. Suppose also $P(2, \infty, \epsilon)$. Then there exist $x \in X$ and $Y \prec X$ such that $x < Y$ and*

$$\{y \in Y : \langle x, y\rangle \subset C_{\epsilon+\delta}\}$$

is large on Y.

Proof. By Lemma 8.5 there exist $A \subset X$ and $Y' \prec X$ such that $A < Y$ and $\{y \in Y' : \langle x, y\rangle \subset C_\epsilon \text{ for some } x \in A\}$ is large on Y'. Now A is finite-dimensional, and therefore is totally bounded. Let $x_1, \ldots, x_N$ be a δ-net of A. It is easy to verify that if $\|x - x_i\|_p \leqslant \delta$ and $\langle x, y\rangle \subset \mathcal{C}_\epsilon$ then $\langle x_i, y\rangle \subset \mathcal{C}_{\epsilon+\delta}$. Thus

$$\bigcup_{i=1}^{N} \{y \in Y' \colon \langle x_i, y\rangle \subset \mathcal{C}_{\epsilon+\delta}\}$$

is large on Y'. By Lemma 2, there is some i and some $Y \prec Y'$ such that

$$\{y \in Y' \colon \langle x_i, y\rangle \subset \mathcal{C}_{\epsilon+\delta}\}$$

is large on Y. □

Proof of Theorem 8·3. We shall use Lemma 8.6 repeatedly. Assume then that $P(2, \infty)$ is true and that $\mathcal{C}$ is large on ℓ_p. We shall find some $X \prec \ell_p$ such that $X \subset \mathcal{C}_\epsilon$. This will show that $(\mathcal{C}_\epsilon)^c$ is not large on ℓ_p, thus proving Theorem 8.3.

By Lemma 8.6, we can find x_1 and $Y_1 \prec \ell_p$ such that $x_1 < Y_1$ and

$$\mathcal{C}_1 \equiv \{y \in Y_1 \colon \langle x_1, y\rangle \subset \mathcal{C}_{\epsilon/2}\}$$

is large on Y_1. By Lemma 8.6 again, we can find x_2 and $Y_2 \prec Y_1$ such that

$$\mathcal{C}_2 \equiv \{y \in Y_2 \colon \langle x_2, y\rangle \subset (\mathcal{C}_1)_{\epsilon/4}\}$$

is large on Y_2. In general, if we have constructed x_i, $\mathcal{C}_i$ and Y_i, let x_{i+1} and Y_{i+1} be such that $x_{i+1} < Y_{i+1}$, $Y_{i+1} \prec Y_i$ and

$$\mathcal{C}_{i+1} \equiv \{y \in Y_{i+1} \colon \langle x_{i+1}, y\rangle \subset (\mathcal{C}_i)_{2^{-(i+1)}\epsilon}\}$$

is large on Y_{i+1}. It is easily checked that $X = \langle x_1, x_2, \ldots\rangle$ satisfies the condition $X \subset \mathcal{C}_\epsilon$. □

Theorem 8·7. *Suppose that, whenever $1 < p < \infty$ and $\epsilon > 0$, every space isomorphic to ℓ_p contains an infinite $(1+\epsilon)$-symmetric basic sequence. Then every space isomorphic to ℓ_p also contains a subspace $(1+\epsilon)$-isomorphic to ℓ_p.*

We shall need a simple lemma. We shall not give a proof as it is very similar to the proof of Lemma 8.1.

Lemma 8·8. *The following are equivalent, for any $1 < p < \infty$.*

(i) Any Banach space isomorphic to ℓ_p contains, for every $\epsilon > 0$, an infinite $(1+\epsilon)$-symmetric basic sequence.

(ii) Let $\|.\|$ be any norm on ℓ_p equivalent to the standard norm. Then there exists an ℓ_p-normalized block basis of the standard basis of ℓ_p which is $(1+\epsilon)$-symmetric under $\|.\|$.

(iii) Let $\epsilon > 0$ and let $C \subset S(\ell_p)$ be any subset of the unit sphere of ℓ_p. Then there exists a normalized block basis $\mathbf{u}_1, \mathbf{u}_2, \ldots$ of the standard basis of ℓ_p such that, for any sequence of scalars $(\lambda_i)_1^\infty$, any sequence of signs $(\epsilon_i)_1^\infty$ and any permutation π of $\mathbb{N}$,

$$\left| d\Big(\sum_{i=1}^{\infty} \epsilon_i \lambda_i \mathbf{u}_{\pi(i)}\ , C\Big) - d\Big(\sum_{i=1}^{\infty} \lambda_i \mathbf{u}_i\ , C\Big) \right| \leqslant \epsilon\,. \qquad \square$$

Proof of Theorem 8·7. By Theorem 8.3, it is enough to prove $P(2,\infty)$ under our assumptions. By Lemma 8.8, we may assume that the third of the three equivalent statements is true. Given $\epsilon > 0$ and $C \subset \ell_p$, let $\delta = \epsilon/3$ and let $X \prec \ell_p$ be the subspace of ℓ_p generated by the block basis given by the third part of Lemma 8.8, with ϵ replaced by δ. Without loss of generality $X = \ell_p$. It was observed by Rosenthal [45] that for any $\eta > 0$ and $k \in \mathbb{N}$ there exists a vector $x \in S(\ell_p)$ of finite support such that, if $x_1, \ldots, x_k$ are disjointly supported copies of x and $\sum_1^k |a_i|^p = 1$, then there is a rearrangement y of $\sum_1^k a_i x_i$ with $\mathrm{supp}(y) \subset \mathrm{supp}(x)$ such that $\|x - y\|_p \leqslant \eta$. (An upper bound for the number of non-zero coordinates needed for x was given in [3]). Let us take $k = 2$, $\eta = \delta$, and let x be the vector we obtain. Let $x_1 < x_2 < \ldots$ be disjoint copies of x and set $Y = \langle x_1, x_2, \ldots \rangle$.

Suppose now that C is large on ℓ_p. It follows that there exists $x \in Y$ of finite support with $x \in C_\delta$. Suppose that $x = \sum_1^n \lambda_i x_i$ and let $x' > x$ be a disjointly supported copy of x, with $x' = \sum_1^n \lambda_i x_i'$. Now let a and b be any two real numbers such that $|a|^p + |b|^p = 1$ and let $y = ax + bx'$. For each $1 \leqslant i \leqslant n$ let z_i be a rearrangement of $ax_i + bx'_i$ such that $\operatorname{supp}(z_i) \subset \operatorname{supp}(x_i)$ and $\|x_i - z_i\| \leqslant \delta$ and let $z = \sum_1^n \lambda_i z_i$. A quick calculation, using the fact that the supports of $z_i - x_i$ are disjoint, gives that $\|x - z\|_p \leqslant \delta$. It follows that $z \in C_{2\delta}$ and hence, by Lemma 8.8, that $ax + bx' \in C_\epsilon$. We have therefore shown that $\langle x, x' \rangle \subset C_\epsilon$ which proves $P(2, \infty, \epsilon)$. Since $\epsilon > 0$ was arbitrary, we have in fact proved $P(2, \infty)$ and hence $P(\infty, \infty)$. □

In 1975, Hindman, as we have mentioned, proved a famous and important theorem. It states that, given any finite colouring of $\mathbb{N}$, there exists an infinite sequence $n_1, n_2, \ldots$ such that, for any finite set $A \subset \mathbb{N}$, the colour of $\sum_{i \in A} n_i$ is the same. This theorem has an equivalent formulation in terms of colourings of $\mathbb{N}^{(<\omega)}$, the set of finite subsets of $\mathbb{N}$. Given any finite colouring of $\mathbb{N}^{(<\omega)}$, there exists an infinite sequence $X_1 < X_2 < \ldots$ of elements of $\mathbb{N}^{(<\omega)}$ such that, for any $A \in \mathbb{N}^{(<\omega)}$, the colour of $\bigcup_{i \in A} X_i$ is the same. In a sense, Hindman's theorem in this form is a natural discrete analogue of the distortion problem.

The "finite unions version" can be regarded as a theorem about finite words in an alphabet consisting of the two letters 0 and 1, and as such has been generalized by Carlson and Simpson [14a] and Furstenberg and Katznelson [21] to theorems concerning larger alphabets. The theorem of Carlson and Simpson is an infinite version of the Hales-Jewett theorem while that of Furstenberg and Katznelson is a refinement of the Carlson-Simpson theorem. In both these theorems, the alphabet concerned is a finite set with no order (although for Furstenberg and Katznelson it has preferred elements). In this paper we prove another natural generalization of Hindman's theorem, this time using a totally ordered alphabet and proving a result which respects the order. Our proof has certain obvious similarities with the methods of [14] and [21], and generalizes Glazer's remarkable proof of Hindman's theorem [cf. 24]. In particular, the induction step of our Lemma 8.11 is essentially contained in [21, Theorem 1.3]. Since our proof is quite short, we give it in full, except for Lemma 8.10, which is well known, and a number of elementary facts that need to be checked, and have been checked elsewhere [cf. 14,21].

As a consequence, we shall be able to deduce easily that a Lipschitz function on the unit sphere of c_0 which does not depend on the signs of the coordinates of any vector can be restricted to an infinite-dimensional subspace on which it is almost constant. The corresponding theorem for general Lipschitz functions is a little harder. We prove an "approximate Ramsey result" which states, roughly speaking, that if a certain discrete structure is coloured with finitely many colours, then it has an infinite substructure all of whose points are close to a point of one particular colour.

Before stating our next theorem, we shall introduce some notation. For any $k \in \mathbb{N}$, let the shift $T : \mathbb{N}^k \to \mathbb{N}^k$ be defined by

$$T : (n_1, n_2, \ldots, n_k) \mapsto (0, n_1, \ldots, n_{k-1})$$

and let $X_k = \mathbb{N}^k \setminus T\mathbb{N}^k = \{(n_1, \ldots, n_k) : n_1 \neq 0\}$. Given a subset $A = \{\mathbf{n}_i : i \in I\} \subset X_k$ indexed by a set I, we shall say that *the subspace generated by* A is the set of elements of X of the form

$$\sum_{j=1}^{k} \sum_{i \in B_j} T^{j-1} \mathbf{n}_i,$$

where $B_1, \ldots, B_k$ are disjoint subsets of I and B_1 is (necessarily) non-empty. The use of the word "subspace" will be justified later.

Theorem 8·9. *Let $k, r \in \mathbb{N}$ and let $X_k = c_1 \cup \ldots \cup c_r$ be a colouring of X_k with r colours. Then there exists a monochromatic subspace of X_k generated by an infinite set.*

The case $k = 1$ of Theorem 8.9 is simply Hindman's theorem. In order to prove the result in general, we rely heavily on the following lemma, which was also used by Glazer.

Lemma 8·10. *Let $(S, +)$ be a compact semigroup such that the function $y \mapsto y + x$ on S is continuous for every $x \in S$. Then there exists an idempotent, that is, an element $x \in S$ such that $x + x = x$.* □

The following notation will be very useful when we deal with ultrafilters. Given a set X and an ultrafilter α on X, we shall define a quantifier Λ_α as follows. If $P(x)$ is any proposition involving the elements of X, then when we write $(\Lambda_\alpha x)P(x)$ we shall mean $\{x \in X : P(x)\} \in \alpha$. (This can be read "for a lot_α of x, $P(x)$".) If X is a semigroup, then the set $U(X)$ of ultrafilters on X can be turned into a compact semigroup by setting

$$\alpha + \beta = \{A \subset X : (\Lambda_\alpha x)(\Lambda_\beta y)\ x + y \in A\} .$$

It is not hard to verify that this operation on $U(X)$ is right-continuous. By Lemma 8.10, it follows that $U(X)$ contains an idempotent.

Given $k \geqslant 2$ let a "shift" operator $S : U(X_k) \to U(X_{k-1})$ be defined as follows. For any $\alpha \in U(X_k)$ we define $S(\alpha) \in X_{k-1}$ to be the set

$$\{A \subset X_{k-1} : (\Lambda_\alpha x)\ Tx \in A\} .$$

It is not hard to check that the operator S is continuous.

Given any $j < k \in \mathbb{N}$ there is an obvious identification between X_j and $T^{k-j}X_k$. We may therefore define a map $+ : (X_j, X_k) \to X_k$ by

$$(m_1, \ldots, m_j) + (n_1, \ldots, n_k) = (n_1, \ldots, n_{k-j}, m_1 + n_{k-j+1}, \ldots, m_j + n_k) .$$

We also define the map $+ : (X_k, X_j) \to X_k$ in a similar way. We can define corresponding maps on the spaces of ultrafilters. For example, the map

$+ : (U(X_k), U(X_j)) \to U(X_k)$ is defined by setting

$$\alpha + \beta = \{A \subset X_k \,:\, (\Lambda_\alpha x \in X_k)(\Lambda_\beta y \in X_j)\;\; x + y \in A\} \;.$$

It is not hard to verify that these maps are all right-continuous. It is also not hard to verify that, if $j, k \geqslant 2$, then $S(\alpha+\beta) = S\alpha + S\beta$ for any $\alpha \in U(X_j)$, $\beta \in U(X_k)$, and, if $1 = j < k$, then $S(\alpha+\beta) = S(\beta+\alpha) = S\beta$ for any $\alpha \in U(X_j)$, $\beta \in U(X_k)$.

We may now state and prove the main lemma upon which Theorem 8.9 depends.

Lemma 8·11. *For every $k \in \mathbb{N}$ there exists an ultrafilter $\alpha \in U(X_k)$ such that $S^j\alpha + \alpha = \alpha + S^j\alpha = \alpha$ for each $0 \leqslant j \leqslant k-1$.*

Proof. We use induction on k. The case $k = 1$ follows from Lemma 8.10 and the remarks just after it. Suppose then that there exists $\beta \in U(X_{k-1})$ such that $S^j\beta + \beta = \beta + S^j\beta = \beta$ for every $0 \leqslant j \leqslant k-2$. Then, since S is continuous and $U(X_k)$ is compact, $S^{-1}\beta$ is a compact subset of $U(X_k)$. Since addition is right-continuous, the set $S^{-1}\beta + \beta$ is also compact. It is closed under addition, since, if $S\gamma_1 = S\gamma_2 = \beta$, then $S(\gamma_1 + \beta + \gamma_2) = S\gamma_1 + S\beta + S\gamma_2 = \beta + S\beta + \beta = \beta + \beta = \beta$. Therefore, by Lemma 8.10, there exists $\gamma \in S^{-1}\beta$ such that $(\gamma+\beta)+(\gamma+\beta) = \gamma+\beta$.

To complete the proof, set $\alpha = \beta + \gamma + \beta$. Then, if $1 \leqslant j \leqslant k-1$, we have

$$\alpha + S^j\alpha = \beta + \gamma + \beta + S^j\beta + S^{j-1}\beta + S^j\beta = \beta + \gamma + \beta = \alpha \;.$$

Similarly, $S^j\alpha + \alpha = \alpha$. Furthermore,

$$\alpha + \alpha = \beta + \gamma + \beta + \beta + \gamma + \beta = \beta + (\gamma + \beta) + (\gamma + \beta) = \beta + (\gamma + \beta) = \alpha \;.$$

Therefore, α will do. □

Proof of Theorem 8·9. Given a finite sequence $\mathrm{n}_1, \ldots, \mathrm{n}_r$ of elements of X_k, let $\langle \mathrm{n}_1, \ldots, \mathrm{n}_r \rangle$ denote the subspace generated by $\{\mathrm{n}_1, \ldots, \mathrm{n}_r\}$. Let $\alpha \in U(X_k)$ be an ultrafilter of the kind guaranteed to exist by Lemma 8.11 and let us write Λ for Λ_α. It follows immediately from Lemma 8.11 that, given any subset $A \subset X_k$, we have $(\Lambda\, x)\ x \in A$ iff $(\Lambda\, x)(\Lambda\, y)\ \langle x, y\rangle \subset A$. Given an r-colouring of X_k, let c_s be the unique colour for which $(\Lambda_\alpha x)\ x \in c_s$ and set $A_1 = c_s$. By the above remarks, we have $(\Lambda\, x)(\Lambda\, y)\ \langle x, y\rangle \subset A_1$. Pick $x_1 \in X_k$ such that $(\Lambda\, y)\ \langle x_1, y\rangle \subset A_1$ and set $A_2 = A_1 \cap \{y : \langle x_1, y\rangle \subset A_1\}$. Then, since α is a filter, $(\Lambda\, x)\ x \in A_2$, and this implies that $(\Lambda\, x)(\Lambda\, y)\ \langle x, y\rangle \subset A_2$. Pick x_2 such that $(\Lambda\, y)\ \langle x_2, y\rangle \subset A_2$ and note that this implies that $(\Lambda\, y)\ \langle x_1, x_2, y\rangle \subset A_1$. Continuing this process, we produce an infinite sequence $x_1, x_2, \ldots$ such that the subspace it generates is contained in A_1, that is, is monochromatic of colour c_s. □

We come now to a "finite unions version" of Theorem 8.9. Let us set Y_k to be the set of functions $f : \mathbb{N} \to \{0, 1, \ldots, k\}$ which are finitely supported and take the value k at least once. Define a shift operator $T : Y_k \to Y_{k-1}$ by setting $(Tf)(n) = (f(n) - 1) \vee 0$. Given a subset $A = \{f_i : i \in I\} \subset Y_k$ of disjointly supported functions, *the subspace generated by* A is the collection of functions of the form $\sum_{i \in I} T^{r_i} f_i$ where $r_i = k$ in all but finitely many cases and $r_i = 0$ at least once. There is an obvious isomorphism between any infinitely generated subspace and Y_k itself. Indeed, Y_k is the subspace generated by the functions $(k, 0, 0, \ldots), (0, k, 0, \ldots), \ldots$.

Theorem 8·12. *Let $k, n \in \mathbb{N}$, and let Y_k be coloured with r colours $c_1, \ldots, c_r$. Then Y_k contains a monochromatic subspace generated by an infinite set.*

Proof. Let $\tilde{\phi} : \mathbb{N} \to \mathbb{N}^{(<\omega)}$ be the usual binary correspondence and let $\phi : X_k \to Y_k$ be defined by

$$\phi : (n_1, \ldots, n_k) \mapsto \max\{k\tilde{\phi}(n_1), (k-1)\tilde{\phi}(n_2), \ldots, \tilde{\phi}(n_k)\}$$

where this is of course a pointwise maximum of functions.

Now it is easy to show that, given any infinite sequence $\mathbf{n}_1, \mathbf{n}_2, \ldots \subset X_k$ and any integer m, there exists a finite subset $A \subset \mathbb{N}$ such that each coordinate of $\sum_{i \in A} \mathbf{n}_i$ is divisible by m. It follows easily that, given such an infinite sequence, there exists a sequence of subsets $A_1 < A_2 < \ldots$ such that the images of the vectors $\phi\left(\sum_{i \in A_j} \mathbf{n}_i\right)$ are disjointly supported.

Let us therefore colour X_k by setting the colour of $\mathbf{n}$ to be the colour of $\phi(\mathbf{n})$ in Y_k. By Theorem 8.9, X_k contains a monochromatic subspace generated by an infinite sequence $\mathbf{n}_1, \mathbf{n}_2, \ldots$, and, by the above remark, the infinite sequence can be chosen so that $\phi(\mathbf{n}_1), \phi(\mathbf{n}_2), \ldots$ are disjointly supported. It is also easy to see that the image of the subspace of X_k generated by $\mathbf{n}_1, \mathbf{n}_2, \ldots$ is the subspace of Y_k generated by $\phi(\mathbf{n}_1), \phi(\mathbf{n}_2), \ldots$. The result follows. □

We come now to the first of our results about Lipschitz functions on c_0. Given any normed space X, let $S(X)$ denote its unit sphere. Let us say that a real-valued Lipschitz function F on $S(c_0)$ is *unconditional* if $F(x) = F(|x|)$ for every $x \in S(c_0)$.

Theorem 8·13. *Let F be an unconditional Lipschitz function on $S(c_0)$. Then, for any $\epsilon > 0$, there exists an infinite-dimensional subspace X of c_0 such that F varies by at most ϵ on X.*

Proof. Without loss of generality, F has Lipschitz constant 1. Let $\delta = \epsilon/2$. There is a natural δ-net of $|S(c_0)|$, namely, the collection of functions $f : \mathbb{N} \to \{1, (1+\delta)^{-1}, \ldots, (1+\delta)^{-(k-1)}\}$ which are finitely supported and take the value 1 at least once, where k is chosen so that $(1+\delta)^{-(k-1)} < \delta$. Let us write Δ for this collection of functions.

Since F is Lipschitz, there exists an interval $[a, b) \subset \mathbb{R}$ such that $F(S(c_0)) \subset [a, b)$. Let r be such that $a + r\delta \geqslant b$ and let the intervals $I_1, \ldots, I_r$ be defined by $I_j = [a + (j-1)\delta, a + j\delta)$ for each $1 \leqslant j \leqslant r$. Given $f \in \Delta$, let us colour f

according to the interval I_j in which $F(f)$ falls.

Now there is an obvious bijection between Δ and Y_k. Indeed, let us define a map $\psi : \Delta \to Y_k$ by

$$(\psi f)(n) = \begin{cases} k + \log_{1+\delta} f(n) & f(n) \neq 0 \\ 0 & f(n) = 0 \end{cases}$$

The colouring on Δ induces a colouring on Y_k. By Theorem 8.12, Y_k contains a monochromatic subspace in this colouring, generated by an infinite set. This set corresponds to a block basis of c_0, and it is not hard to see that the subspace generated by the set corresponds to a δ-net of the positive part of the unit sphere of the subspace generated by the block basis in c_0. Therefore, since F is unconditional and varies by at most δ on this set, it can vary by at most $2\delta = \epsilon$ on the whole of the unit sphere of the subspace. This completes the proof of Theorem 8.13. □

We shall now extend Theorem 8.13 to arbitrary Lipschitz functions. The proof of this is harder, because the most obvious combinatorial approach does not work, and must be replaced with an "approximate Ramsey result", as we mentioned earlier.

Given $k \in \mathbb{N}$ let Z_k be the set of functions $f : \mathbb{N} \to \{-k, -(k-1), \ldots, k\}$ which are finitely supported and take one of the values $\pm k$ at least once. Let the shift T be defined by

$$(Tf)(n) = \text{sign}(f(n))((|f(n)| - 1) \vee 0) .$$

If the functions $\{f_i :\in I\}$ are disjointly supported, then let *the subspace generated by* $\{f_i :\in I\}$ be defined to be the set of functions of the form

$$\sum_{j=1}^{k} \sum_{i \in A_j} T^{k-j} f_i - \sum_{j=1}^{k} \sum_{i \in B_j} T^{k-j} f_i$$

where $A_1, \ldots, A_k, B_1, \ldots, B_k$ are all disjoint, and at least one of A_1 and B_1 is non-empty.

If every colouring of Z_k yielded an infinite monochromatic subspace, then we would be done, by imitating the deduction of Theorem 8.13 from Theorem 8.12. However, it does not take long to see that this is not the case. For example, one can colour each function f by the sign of its first non-zero coordinate, or by whether or not the first and last non-zero coordinates have the same sign. This second colouring shows that we cannot get anywhere by restricting our attention to colourings for which the colour of f is always the same as the colour of $-f$.

These colourings, however, cannot be adapted to give counterexamples to the next result, which will be sufficient for our purposes.

Theorem 8·14. *Let Z_k be finitely coloured. Then there exists a colour $A \subset Z_k$ and an infinite subspace $W \subset Z_k$ such that $W \subset A'$, where*

$$A' = \{f \in Z_k : (\exists g \in A)\,|f| \leqslant |g|, \|f - g\|_\infty \leqslant 1\}\ .$$

In other words, given any finite colouring of Z_k, there exists a colour A and an infinite subspace every point of which is close to a point in A and dominated by it.

The next lemma is in a sense the finite-dimensional version of what we want. It is perhaps surprising that the finite-dimensional version should be useful for proving the infinite-dimensional version, but this seems to be the case.

Lemma 8·15. *Let $n \in \mathbb{N}$ and $\epsilon > 0$. If $N = N(n, \epsilon)$ is sufficiently large, then, given any decomposition $S(\ell_\infty^N) = A \cup B$, there exists a block subspace $X \subset \ell_\infty^N$ such that $\dim(X) = n$ and either $X \subset A_\epsilon$ or $X \subset B_\epsilon$.*

Proof. Let $S(c_0) = A \cup B$ and let a Lipschitz fuction F be defined on $S(c_0)$ by $F(x) = d(x, A)$. By an obvious adaptation from norms to Lipschitz functions of the methods of [3], one can find, for any $M \in \mathbb{N}$, a block subspace $X = \langle x_1, \ldots, x_M \rangle$ of c_0 such that, whenever $x = \sum_{i=1}^M a_i x_i, y = \sum_{i=1}^M b_i x_i$ are in $S(X)$ and $|a_i| = |b_i|$

for each i, we have $|F(x) - F(y)| \leqslant \epsilon/2$. Let us pick such a subspace X and let G be the unconditional Lipschitz function defined on $S(\ell_\infty^M)$ by

$$G(\mathbf{a}) = \max\left\{ F\left(\sum_{i=1}^{M} a_i' x_i\right) : |a_i'| = |a_i| \text{ for each } i \right\} .$$

By the finite version of Theorem 8.13, which follows from the well-known compactness principle (cf. e.g. [24]), if M was large enough then ℓ_∞^M has an n-dimensional block subspace on whose unit sphere G varies by at most $\epsilon/2$. It follows easily that X has a block subspace on which F varies by at most ϵ. Hence, we must either have $Y \subset A_\epsilon$ or $Y \subset A^c \subset B$. By the compactness principle once more we obtain the result. □

Let us say that a subset $A \subset S(c_0)$ is *n-large* if, for every n-dimensional block subspace X of c_0, $A \cap X \neq \emptyset$. A subset that is n-large for some n, we shall call *finitely large.*

Corollary 8·16. *Let $\beta = \{A_\epsilon \subset S(c_0) : \epsilon > 0, A$ is finitely large$\}$. Then β is a filter-base.*

Proof. We must show that if A and B are finitely large and $\epsilon > 0$, then there exist $C \subset S(c_0)$ and $\delta > 0$ such that C is finitely large and $C_\delta \subset A_\epsilon \cap B_\epsilon$. Pick n such that A and B are both n-large. Let $N = N(n, \epsilon/4)$ be as given by Lemma 8.15 and let X be any n-dimensional block subspace of c_0. Then X is isometric to ℓ_∞^N and $X \subset A_{\epsilon/2} \cup A^c_{\epsilon/2}$. By Lemma 8.15, X has an n-dimensional block subspace Y which is either contained in $A_{3\epsilon/4}$ or $A^c_{\epsilon/4}$. Since A is n-large, the former must be the case. But then, since B is also n-large, $B \cap Y \neq \emptyset$. However, X was arbitrary, so $A_{3\epsilon/4} \cap B$ is N-large. But $(A_{3\epsilon/4} \cap B)_{\epsilon/4} \subset A_\epsilon \cap B_\epsilon$, so we may set $C = A_{3\epsilon/4} \cap B$ and $\delta = \epsilon/4$. □

Given a filter α on $S(c_0)$, let us say that it is *cofinite* if, for every $n \in \mathbb{N}$, the unit sphere of $\langle \mathrm{e}_n, \mathrm{e}_{n+1}, \ldots \rangle$ is in α. Note that these sets are all finitely large.

Corollary 8·17. *There exists a cofinite ultrafilter α on $S(c_0)$ with the property that, whenever $A \in \alpha$ and $\epsilon > 0$, the set $-A_\epsilon$ is also in α.*

Proof. Let α be a maximal filter on $S(c_0)$ with the following two properties. First, α extends the filter generated by β (and is therefore cofinite), and second, $-A_\epsilon \in \alpha$ whenever $A \in \alpha$ and $\epsilon > 0$. That such a maximal filter exists follows easily from Zorn's lemma. We shall show that α is in fact an ultrafilter. Indeed, suppose that $A \cup B = S(c_0)$ and let $X \subset S(c_0)$ be any 2-dimensional block subspace. We claim that there exists $x \in X$ such that $x \in (A \cap -A_\epsilon) \cup (B \cap -B_\epsilon)$. This is very easy. If $X \subset A$ or $X \subset B$ then we are done. Otherwise, since X is connected, $X \cap \bar{A} \cap \bar{B} \neq \emptyset$, and therefore there exists $x \in X \cap A \cap B_\epsilon$. We are then obviously done, whether $-x$ is in A or B. Thus, the set $E = (A \cap -A_\epsilon) \cup (B \cap -B_\epsilon)$ is 2-large. It follows that $E \in \beta$ and hence that $E \in \alpha$.

Now suppose that neither A nor B is in α. Since α is maximal, there must be some $C \in \alpha$ such that $C \cap A \cap -A_\epsilon = \emptyset$, and some $D \in \alpha$ such that $D \cap B \cap -B_\epsilon = \emptyset$. It follows that $(C \cap E) \cap (D \cap E) = \emptyset$. This contradicts the fact that α is a filter.

□

Given a filter on Z_k, there is an obvious notion of cofiniteness corresponding to the case of filters on $S(c_0)$. We shall say that a filter α is *cofinite* if, for every $n \in \mathbb{N}$, the subspace generated by $ke_n, ke_{n+1}, \ldots$ is in α.

Corollary 8·18. *For every $k \in \mathbb{N}$ there exists a cofinite ultrafilter $\tilde{\alpha}$ on Z_k such that $-\bar{A} \in \tilde{\alpha}$ whenever $A \in \tilde{\alpha}$.*

Proof. Let $\phi : S(c_0) \to Z_k$ be defined by

$$(\phi f)(n) = \mathrm{sign}(f(n))\Big(\max\{j \in \mathbb{Z} : (1+\delta)^{-(k-j)} \leqslant |f(n)|\} \vee 0\Big) .$$

In other words, if one takes the obvious analogue of Δ in the whole of the unit sphere of c_0, the map ϕ "rounds down" to the nearest point in the net Δ and takes the corresponding point of Z_k.

The ultrafilter $\tilde{\alpha}$ is defined by taking $A \subset Z_k$ to be in $\tilde{\alpha}$ iff $\phi^{-1}(A) \in \alpha$, where α is the ultrafilter constructed in Corollary 8.17. Note that $A \cap B = \emptyset$ iff $\phi^{-1}(A) \cap \phi^{-1}(B) = \emptyset$. It follows easily that $\tilde{\alpha}$ is indeed an ultrafilter. It is also clear that $\tilde{\alpha}$ is cofinite.

Now, if $A \in \tilde{\alpha}$, then $-\phi^{-1}(A)_\epsilon \in \alpha$ for every $\epsilon > 0$. If ϵ is sufficiently small, it is clear that $\phi(-\phi^{-1}(A)_\epsilon) = -\bar{A}$, and certainly that $-\phi^{-1}(A)_\epsilon \subset \phi^{-1}(-\bar{A})$. It follows that $-\bar{A} \in \tilde{\alpha}$. □

Note some important facts about the ultrafilter constructed in Corollary 8.18. First, the set of ultrafilters on Z_k which satisfy the conditions in the corollary is easily seen to be closed. Let us denote it by $V(Z_k)$. Given $j, k \in \mathbb{N}$ it is easy to check, using cofiniteness, that one can define a map $+ : (V(Z_j), V(Z_k)) \to V(Z_{j \vee k})$ by setting

$$\alpha + \beta = \{A \subset V(Z_{j\vee k} : (\Lambda_\alpha x)(\Lambda_\beta y) \operatorname{supp}(x) \cap \operatorname{supp}(y) = \emptyset, x + y \in A\} .$$

In order to avoid writing too much, let us adopt the convention that the operation $+$ is only defined on elements of Z_j and Z_k when they are disjointly supported We shall verify that $-\bar{A} \in \alpha + \beta$ whenever $A \in \alpha + \beta$. Note first that, if $\alpha \in V(Z_k)$ and P is a sentence taking $x \in Z_k$ as a variable, then

$$(\Lambda_\alpha x)\ P(x)$$

implies that

$$(\Lambda_\alpha x)(\exists x')\ |x| \leqslant |x'|, \|x - x'\|_\infty \leqslant 1, P(-x') .$$

Hence, if

$$(\Lambda_\alpha x)(\Lambda_\beta y)\ x + y \in A ,$$

then

$$(\Lambda_\alpha x)(\exists x')(\Lambda_\beta y)(\exists y')$$

$$|x| \leqslant |x'|, |y| \leqslant |y'|, \|x' - x\|_\infty \leqslant 1, \|y' - y\|_\infty \leqslant 1, -x' - y' \in A$$

which implies that $x' + y' \in -A$ and hence that

$$(\Lambda_\alpha x)(\Lambda_\beta y)\ x + y \in -\bar{A}\ .$$

We define a continuous shift operator $S : V(Z_k) \to V(Z_{k-1})$ just as before, that is, by setting

$$S(\alpha) = \left\{A \in Z_{k-1} : (\Lambda_\alpha x)\ Tx \in A\right\}\ .$$

It is simple to check that $S(V(Z_k)) = V(Z_{k-1})$. The proof of the next lemma now mirrors directly the proof of Lemma 8.11.

Lemma 8·19. *For every $k \in \mathbb{N}$ there exists an ultrafilter $\alpha \in V(Z_k)$ such that $S^j\alpha + \alpha = \alpha + S^j\alpha = \alpha$ for each $0 \leqslant j \leqslant k-1$.* □

Proof of Theorem 8·14. Given $x, y \in Z_k$ and $A \subset Z_k$, let $\langle x', y\rangle \subset A$ stand for the statement that there exists x' such that $|x| \leqslant |x'|$, $\|x' - x\| \leqslant 1$ and, for every $0 \leqslant j \leqslant k$, $T^jx + y \in A$, $x + T^jy \in A$, $-T^jx' + y \in A$ and $-x' + T^jy \in A$. Our ultrafilter α in Lemma 8.19 is constructed so that whenever $(\Lambda_\alpha x)\ x \in A$ we have also $(\Lambda_\alpha x)(\Lambda_\alpha y)\ \langle x', y\rangle \subset A$. Note that $\langle x', y\rangle \subset A$ implies that $\langle x, y\rangle \subset \bar{A}$.

Let $\alpha \in V(Z_k)$ be as given by Lemma 8.19, and let $A_1 \subset Z_k$ be the colour which is in α. Then we have $(\Lambda_\alpha x)(\Lambda_\alpha y)\ \langle x', y\rangle \subset A_1$. Let x_1 be such that $(\Lambda_\alpha y)\ \langle x_1', y\rangle \subset A_1$ and let A_2 be the set of $y \in A_1$ for which $\langle x_1', y\rangle \subset A_1$, so that $A_2 \in \alpha$. It follows that $(\Lambda_\alpha x)(\Lambda_\alpha y)\ \langle x', y\rangle \subset A_2$. Pick x_2 such that $(\Lambda_\alpha y)\ \langle x_2', y\rangle \subset A_2$ and let A_3 be the set of $y \in A_2$ for which $\langle x_2', y\rangle \subset A_2$. Note that if $\langle x_2', y\rangle \subset A_2$ then $\langle x_1, x_2, y\rangle \subset \bar{A}$. Continuing this process, we produce an infinite sequence $x_1, x_2, \ldots$ which generates a subspace $X \subset \bar{A}$, as stated. □

We have essentially proved our second main theorem.

Theorem 8·20. *Let $\epsilon > 0$ and let F be any real-valued Lipschitz function on the unit sphere of c_0. Then there is an infinite-dimensional subspace $X \subset c_0$ on whose unit sphere F varies by at most ϵ.*

Proof. Let $\delta = \epsilon/4$ and let Δ_1 be the set of functions $f : \mathbb{N} \to \{\pm 1, \pm(1+\delta)^{-1}, \ldots, \pm(1+\delta)^{-(k-1)}\}$ which are finitely supported and take one of the values ± 1 at least once, where k is chosen such that $(1+\delta)^{-(k-1)} \leqslant \delta$. This is a 2δ-net of $S(c_0)$. Let $S(c_0)$ be coloured as in the proof of Theorem 8.13, and let ψ be the map defined there. Let ψ_1 be the natural bijection between Δ_1 and Z_k, defined by

$$(\psi_1 f)(n) = \mathrm{sign}(f(n))(\psi|f|)(n) \ .$$

The colouring on Δ_1 induces a colouring on Z_k. It is now easy to see that Theorem 8.20 follows straight from Theorem 8.14. □

CHAPTER 9

DISTANCE FROM THE SPHERE

For the proofs of Theorems 4.6 and 4.7 we used the fact that if the 2-type constant of an n-dimensional normed space X is close to 1, then $d(X, \ell_2^n)$ is not too large. The main result in this direction is a theorem of König and Tzafriri [29] which is similar to Lemma 4.5, but appropriate when the 2-type constant is large. They showed that if $T_2(X) \leqslant C$ then $d(X, \ell_2^n) \leqslant 16n^{1/2-\epsilon}$, where $\epsilon = 1/6C^2$. They mentioned the following problem, which seems to be well known. Suppose we know, for some $1 < p < 2$ and some $C \geqslant 1$, that $T_p(X) \leqslant C$. Does it follow that $d(X, \ell_2^n) \leqslant Kn^{1/2-\epsilon}$ for some K and ϵ that depend only on p and C? König and Tzafriri showed that the answer is yes if p is sufficiently close to 2 and C is sufficiently close to 1. Note that, if p' is the conjugate index of p and $q = C^{p'}$, then $T_p(\ell_q^n) \leqslant C$. This shows that ϵ cannot be greater than $1/C^{p'}$. It is not even known whether this estimate is in fact sharp. For the rest of the chapter, let $\epsilon_{\min}$ be a function of p, C and n defined as follows. We set $\epsilon_{\min}(p, C, n)$ to be the smallest possible value of ϵ such that there exists an n-dimensional space X with $T_p(X) \leqslant C$ and $d(X, \ell_2^n) \geqslant n^{1/2-\epsilon}$. The question we are considering is whether $\epsilon_{\min}$ is bounded below by a positive function that does not depend on n.

This chapter consists of work in progress. We have not answered the above question, but we shall show that, under a rather strong but nevertheless commonly occurring condition on the space X, the bound of $\epsilon = 1/C^{p'}$ is roughly correct, and we shall relate the problem to a question of Tomczak-Jaegermann. In her recent book [47] she asks the following. Given a normed space X, let us set

$$d_k = \sup\left\{d(Y, \ell_2^k) : \dim(Y) = k,\ Y \subset X\right\}.$$

Then if $\dim(X) = n = kl$, does it follow that $d(X, \ell_2^n) \leqslant d_k d_l$?

We shall first show that a positive answer to this question implies that $\epsilon_{\min}$ is indeed bounded below by a function independent of n, and moreover that a function of the form $c/(CK)^{p'}$, where c and K are absolute constants and p' is the conjugate index of p, will do. For this we need a result of Milman and Wolfson [40]. They showed that if $c > 0$ is an absolute constant, then there exist constants $c' > 0$ and C, depending only on c with the following property. If X is any n-dimensional normed space such that $d(X, \ell_2^n) \geqslant c\sqrt{n}$, then X contains a subspace Y with $\dim(Y) = k = c' \log\log\log n$ and $d(Y, \ell_1^k) \leqslant C$. Recently, the bound on the dimension of Y was improved by Bourgain and Tomczak-Jaegermann [48], following a general approach of Pisier [43].

Theorem 9·1. *Let $c > 0$, let X be an n-dimensional normed space and suppose that $d(X, \ell_2^n) \geqslant c\sqrt{n}$. Then X has a subspace Y of dimension $k = c' \log n$, such that $d(Y, \ell_1^k) \leqslant C$, where c' and C are constants that depend only on c.* □

This theorem already gives some information about the main question we are considering. It tells us that if $(X_n)_1^\infty$ is a sequence of normed spaces with $\dim(X_n) = n$ and $d(X_n, \ell_2^n)$ close enough to $\sqrt{n}$ for each n, then $T_p(X_n)$ cannot be independent of n. However, the dependence of c' on c is such that it gives no information at all if $d(X_n, \ell_2^n) \leqslant n^{1/2}/\log n$.

In order to show that a positive answer to Tomczak-Jaegermann's question implies that $\epsilon_{\min}$ is bounded below by $c/(CK)^{p'}$ for some absolute constants c and K, we shall use one other result due to Tomczak-Jaegermann, which is a partial answer to her question. Suppose X is an n-dimensional normed space with $d(X, \ell_2^n) \geqslant d$. Then, for any positive integer k, X has a subspace Y with $\dim(Y) = m = \lceil n/k \rceil$ such that $d(Y, \ell_2^m) \geqslant d/\sqrt{2k}$. In particular, setting $k = 4$, we find that X contains a subspace Y with $\dim(Y) = m = \lceil n/4 \rceil$ and $d(Y, \ell_2^m) \geqslant d/2\sqrt{2}$. Let r be an integer such that $n/4 < 2^{2r} \leqslant n$, and let Z be any 2^{2r}-dimensional subspace of X which contains Y. Then certainly $d(Z, \ell_2^{2^{2r}}) \geqslant d/2\sqrt{2}$.

Now let us assume that the question of Tomczak-Jaegermann has a positive answer. Then we find immediately that Z contains a subspace X_1 of dimension 2^r such that $d(X_1, \ell_2^{2^r}) \geqslant 2^{-3/4} d^{1/2} > (1/2) d^{1/2}$. By repeating this process, we obtain, for any integer s, a subspace X_s of X with $\dim(X_s) = k'_s \leqslant n^{1/2^s}$ such that $d(X, \ell_2^{k'_s}) \geqslant (1/2)^{1+2^{-1}+\ldots+2^{-(s-1)}} d^{1/2^s} \geqslant (1/4) d^{1/2^s}$.

Suppose now that $d = n^{1/2-\epsilon}$. Then, by the above argument, we can find, for any integer s, a subspace $X_s \subset X$ such that $\dim(X_s) = k_s = \lfloor n^{1/2^s} \rfloor$ and $d(X_s, \ell_2^{k_s}) \geqslant (1/4) n^{2^{-s}(1/2-\epsilon)} \geqslant (1/4) k_s^{1/2-\epsilon}$.

Let us pick s such that $2 < n^{\epsilon/2^s} \leqslant 4$. Then $\lfloor 2^{1/\epsilon} \rfloor \leqslant k_s \leqslant \lfloor 4^{1/\epsilon} \rfloor$ and $d(X_s, \ell_2^{k_s}) \geqslant (1/16) k_s^{1/2}$. By Theorem 9.1 there exist absolute constants $c > 0$ and $K \geqslant 1$ such that X_s contains a subspace Y with $\dim(Y) = t \geqslant c/\epsilon$ and $d(Y, \ell_1^t) \leqslant K$.

It follows that if $1 < p \leqslant 2$ then $T_p(X) \geqslant K^{-1}(c \log 2/\epsilon)^{p'}$, where p' is the conjugate index of p. Hence, if $T_p(X) \leqslant C$ then $\epsilon \geqslant c/(CK)^{p'}$. In other words, for this value of ϵ, we have that $T_p(X) \leqslant C$ implies $d(X, \ell_2^n) \leqslant n^{1/2-\epsilon}$.

We have shown that if distance from Hilbert space does behave submultiplicatively, then $\epsilon_{\min}$ is bounded below by $c/(CK)^{p'}$ for some absolute constants c and K. That is, it would follow that the easy upper bound for $\epsilon_{\min}$ mentioned at the beginning of the chapter was roughly correct.

In the remainder of this chapter, we shall use a similar technique to prove that a lower bound for $\epsilon_{\min}$ of this form is valid if we restrict our attention to a fairly wide class of spaces. First, we shall need a characterization of the distance of an n-dimensional normed space from ℓ_2^n. This is a special case of a result concerning 2-factorable operators due to Lindenstrauss and Pelczynski [33].

Theorem 9·2. *Let X be an n-dimensional normed space. Then $d(X, \ell_2^n) \geqslant d$ if* *and only if there exists a sequence of vectors $x_1, \ldots, x_N$ in X and an orthogonal $N \times N$ matrix (a_{ij}) such that*

$$\sum_{i=1}^{N} \left\| \sum_{j=1}^{N} a_{ij} x_j \right\|^2 \geqslant d^2 \sum_{j=1}^{N} \|x_j\|^2 .$$

Of course, the important implication in this theorem is that the conditions given are necessary. Even so, the simple fact that they are sufficient gives a useful method of estimating the distance of a given n-dimensional space from ℓ_2^n from below, and an easy way of constructing spaces with extremal distance from ℓ_2^n.

We shall be concerned with one particular orthogonal matrix – the normalized Walsh matrix. The Walsh matrix is defined inductively as follows. Let W_0 be the 1×1-matrix whose single entry is 1. Now for $k \geqslant 1$ set

$$W_k = \begin{pmatrix} W_{k-1} & W_{k-1} \\ W_{k-1} & -W_{k-1} \end{pmatrix} .$$

Thus, W_k is a $2^k \times 2^k$-matrix with ± 1-entries. It is easy to verify that the rows and columns of W_k are orthogonal. An important property of the Walsh matrices which we shall use is the following. Suppose $r = s + t$ and let us write $n = 2^r$, $k = 2^s$ and $l = 2^t$. Then for any i and $j \in [n]$ we have $(W_r)_{ij} = (W_s)_{i'j'}(W_t)_{i''j''}$, where i', j', i'' and j'' are chosen so that $i = (l-1)i' + i''$ and $j = (l-1)j' + j''$. To be concise, $W_r = W_s \otimes W_t$. Our main theorem in this chapter is the following.

Theorem 9·3. *Let $n = 2^r$ and suppose that X is an n-dimensional normed space. Let $\epsilon > 0$, let $A = n^{-1/2} W_r$, and suppose that there exist vectors $x_1, \ldots, x_n$ in X such that*

$$\sum_{i=1}^{n} \left\| \sum_{j=1}^{n} a_{ij} x_j \right\|^2 \geqslant n^{1-2\epsilon} \sum_{j=1}^{n} \|x_j\|^2 ,$$

where a_{ij} is the ij^{th} entry of A. Then, for any $1 < p \leqslant 2$, we have $T_p(X) \geqslant K^{-1}(c/\epsilon)^{1/p'}$, where c and K are absolute constants and p' is the conjugate index for p.

Proof. Let us write $r = s+t$, where $s = t = r/2$ if r is even and $s+1 = t = (r+1)/2$ if r is odd. Let us also set $k = 2^s$ and $l = 2^t$, as above. Let us relabel $x_1, \ldots, x_n$ by setting $x_{ij} = x_{(l-1)i+j}$ for every $i \in [k]$ and $j \in [l]$. Let us also write $a_{iji'j'}$ for $a_{(l-1)i+j,(l-1)i'+j'}$ for every $i, i' \in [k]$ and $j, j' \in [l]$. Let $B = k^{-1/2}W_s$ and $C = l^{-1/2}W_t$, and let the ij^{th} entries of B and C be b_{ij} and c_{ij} respectively. Then, by the remarks before the statement of the theorem, we have $a_{iji'j'} = b_{ii'}c_{jj'}$, for every $i, i' \in [k]$ and $j, j' \in [l]$. Using this relabelling, we may write

$$\sum_{i=1}^{n}\left\|\sum_{j=1}^{n} a_{ij}x_j\right\|^2 = \sum_{i=1}^{k}\sum_{j=1}^{l}\left\|\sum_{i'=1}^{k}\sum_{j'=1}^{l} b_{ii'}c_{jj'}x_{i'j'}\right\|^2$$

and

$$\sum_{j=1}^{n}\|x_j\|^2 = \sum_{i'=1}^{k}\sum_{j'=1}^{l}\|x_{i'j'}\|^2 .$$

For any $j \in [l]$ and $i' \in [k]$, let us set

$$\mathbf{u}_{i'}^{j} = \sum_{j'=1}^{l} c_{jj'}x_{i'j'} .$$

Then

$$\frac{\sum_{i=1}^{n}\left\|\sum_{j=1}^{n} a_{ij}x_j\right\|^2}{\sum_{j=1}^{n}\|x_j\|^2} = \frac{\sum_{j=1}^{l}\sum_{i=1}^{k}\left\|\sum_{i'=1}^{k} b_{ii'}\mathbf{u}_{i'}^{j}\right\|^2}{\sum_{j=1}^{l}\sum_{i'=1}^{k}\left\|\mathbf{u}_{i'}^{j}\right\|^2} \cdot \frac{\sum_{j=1}^{l}\sum_{i'=1}^{k}\left\|\sum_{j'=1}^{l} c_{jj'}x_{i'j'}\right\|^2}{\sum_{i'=1}^{k}\sum_{j'=1}^{l}\|x_{i'j'}\|^2} .$$

Now it follows immediately from our assumption that either the left half of the expression on the right hand side is at least $k^{1/2-\epsilon}$ or the right half is at least $l^{1/2-\epsilon}$. In the first case, there must exist $j \in [l]$ such that

$$\sum_{i=1}^{k}\left\|\sum_{i'=1}^{k} b_{ii'}\mathbf{u}_{i'}^{j}\right\|^2 \geqslant k^{1/2-\epsilon}\sum_{i'=1}^{k}\left\|\mathbf{u}_{i'}^{j}\right\|^2 .$$

In the second case, there must exist $i' \in [k]$ such that

$$\sum_{j=1}^{l}\left\|\sum_{j'=1}^{l} c_{jj'}x_{i'j'}\right\|^2 \geqslant l^{1/2-\epsilon}\sum_{j'=1}^{l}\|x_{i'j'}\|^2 .$$

In the first case, let X_1 be the space spanned by $\mathbf{u}_1^j, \ldots, \mathbf{u}_k^j$ and in the second case, let X_1 be the space spanned by $x_{i'1}, \ldots, x_{i'l}$. Then we have found a subspace X_1 of X which satisfies the same conditions as X, except that $n = 2^r$ has been replaced by $k = 2^s$ or $l = 2^t$. Let us repeat the argument, to obtain a nested sequence $X_0 = X, X_1, X_2, \ldots$ of subspaces of X, with $\dim(X_s) = n_s = 2^{k_s}$ for each s and $k_s = (k_{s-1} \pm 1)/2$ if k_{s-1} is odd and $k_{s-1}/2$ if k_{s-1} is even. We can certainly find s such that $\sqrt{2} < n_s^\epsilon \leqslant 4$. For such an s, Theorem 9.2 gives that $d(X_s, \ell_2^{n_s}) \geqslant (1/4)n_s^{1/2}$. By Theorem 9.1, there are absolute constants c and K such that X_s contains a subspace Y of dimension $k \geqslant c/\epsilon$ with $d(Y, \ell_1^k) \leqslant K$. It follows that $T_p(X) \geqslant K^{-1}(c/\epsilon)^{1/p'}$ as stated. □

The most obvious spaces in which the conditions of Theorem 9.3 are satisfied are the ℓ_p-spaces. If $p < 2$ then we may take $x_1, \ldots, x_n$ to be the standard basis, while if $p > 2$ we can use the easy fact that the matrix A is its own inverse, and take $x_i = \sum_{j=1}^n a_{ij}\mathbf{e}_j$ for each i, where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is of course the standard basis. Of course, Theorem 9.3 does not give any interesting information in this case, apart from an unnecessarily complicated way of estimating a lower bound for the type constants of ℓ_p when $p > 2$.

There is one other normed space to which Theorem 9.3 applies which has been considered in the literature. It is a space constructed by Bourgain [12] as a counterexample to a conjectured strengthening of a theorem of Tomczak-Jaegermann used in Chapter 4. She proved [48] that if X is an n-dimensional normed space and $1 \leqslant p \leqslant 2 \leqslant q \leqslant \infty$, then $d(X, \ell_2^n) \leqslant T_p(X)C_q(X)n^{1/p-1/q}$. (The case $p = q = 2$ was proved by Kwapien [31]). The conjecture was that the power of n could be improved to $\max\{1/p - 1/2, 1/2 - 1/q\}$. Bourgain constructed, for every p and q such that $1/p - 1/q < 1/2$, a space X such that $T_p(X)$ and $C_q(X)$ are bounded above by an absolute constant, and the correct estimate for the distace of X from

ℓ_2^n is indeed given, up to a constant, by $n^{1/p-1/q}$. We shall discuss his example briefly, and then apply Theorem 9.3 to it.

It is constructed as follows. Let $n = 2^k$, let $A = n^{-1/2}W_k$, let $1 \leqslant p \leqslant 2$, let p' be the conjugate index for p and let $\|.\|$ and $|\!|\!|.|\!|\!|$ be the following two norms on $\mathbb{R}^n$. The norm $\|.\|$ is just the standard $\ell_{p'}$-norm, while for any $x \in \mathbb{R}^n$, $|\!|\!|x|\!|\!| = \|Ax\|_p$. Let $X_p = (\mathbb{R}^n, \|.\|) = \ell_\infty^n$ and let $Y_p = (\mathbb{R}^n, |\!|\!|.|\!|\!|)$. Finally, let Z_p be the real interpolation space "half way between" X_p and Y_p, that is the space $(X_p, Y_p)_{1/2,2}$. It follows from elementary facts about interpolation that $T_q(Z_p) \leqslant (p')^{1/4}$, where $q = 4p/2 + p$, and also that Z_p is isometric to its dual. It follows that $C_{q'}(Z) \leqslant (p')^{1/4}$. When $p = 1$, one obtains $T_{4/3}(Z_1) \leqslant (\log n)^{1/4}$ and $C_4(Z_1) \leqslant (\log n)^{1/4}$. It is an easy deduction from the easy implication of Theorem 9.1 that $d(Z_p, \ell_2^n) = n^{1/p-1/2}$.

If we now consider the space Z_1 and let $x_1, \ldots, x_n$ be given by $x_i = Ae_i$ for each i, a direct calculation of the norms at the relevant points (using again the fact that A is self-inverse) yields that

$$\sum_{i=1}^{n} \left\| \sum_{j=1}^{n} a_{ij} x_j \right\|^2 = n \sum_{i=1}^{n} \|x_i\|^2 .$$

It follows easily from the proof of Theorem 9.3 that the estimate for the q-type constant of Z_p given above is correct.

This of course follows from Theorem 9.1. However, by the same method, we can show that if p is very slightly greater than 1, so that $d(Z_p, \ell_2^n) \geqslant n^{1/2}/\log n$, for example, we obtain almost the same estimate for the type constants of Z_p, while Theorem 9.1 tells us nothing. This may well be known. However, it shows that Bourgain's construction cannot give an example of a sequence of spaces for which $\epsilon_{\min}$ tends to zero with n, for some fixed p and C, while strongly suggesting a construction that might do so. The proof of Theorem 9.3 says, speaking very loosely, that if X is an n-dimensional space whose distance from ℓ_2^n can be estimated from below using a Walsh matrix and the sufficiency of the conditions in Theorem 9.2, then its distance from ℓ_2^n behaves submultiplicatively. The same result can easily be shown by the same method for any orthogonal matrix built up iteratively in a similar way. However, a typical orthogonal matrix cannot be decomposed as a tensor product of two smaller ones. This suggests that if one replaced the Walsh matrix by a random orthogonal matrix in the example of Bourgain, one might obtain an improved estimate for the type constants, at the expense of a slightly (but of course necessarily) worse estimate for the distance from ℓ_2^n.

I have tried this, and unfortunately, if some rather unpleasant calculations were correct, it does not work. Nevertheless, I am convinced that there is a counterexample. In other words, I believe that there exist absolute constants $C \geqslant 1$ and $1 < p < 2$ such that $\epsilon_{\min}(p, C, n)$ tends to zero as n tends to infinity. This would of course show that Tomczak-Jaegermann's question also has a negative answer, which seems highly likely anyway. On the other hand, it is generally thought that there might well be a positive result if X is assumed not only to have a type condition, but also to be uniformly convex or uniformly smooth. I have not investigated this question, but intend to do so.

CHAPTER 10

THE HILBERT SPACE DOES NOT HAVE PROPERTY B

The result of this chapter is not related to any other results of this dissertation, except in the very weak sense that we shall construct a norm which happens to be 1-symmetric; it is to do with norm-attaining operators. In 1961, Bishop and Phelps [9] showed that, given any Banach space X, the set of norm-attaining continuous linear functionals on X is dense in X^*, a norm-attaining functional being simply a functional $x^* \in X^*$ for which there exists $x \in X$ such that $|x^*(x)| = \|x\|_{X^*}$. They asked in their paper to what extent this result could be generalized. That is, for which pairs of Banach spaces X, Y is it true that the set of norm-attaining linear operators from X to Y is dense in the set of bounded linear operators from X to Y? (A bounded operator $T: X \longrightarrow Y$ is said to be *norm-attaining* if there exists $x \in X$ such that $\|x\| = 1$ and $\|T(x)\| = \|T\| \equiv \sup\{\|T(x')\| : \|x'\| = 1\}$). Over the years, this question has been considered by various authors.

Let us write $B(X, Y)$ for the set of bounded linear operators from X to Y, and $NB(X, Y)$ for the set of norm-attaining ones. The question asked by Bishop and Phelps appears to be too general to allow a reasonably complete answer, so Lindenstrauss [32] made the following definitions. A Banach space X is said to have property A if, for all spaces Y, $NB(X, Y)$ is dense in $B(X, Y)$, and property B if, for all spaces Y, $NB(Y, X)$ is dense in $B(Y, X)$. One can now ask which Banach spaces have property A and which have property B.

Bourgain [11] showed that any space with the Radon-Nikodym property also has property A, but the results for property B are less satisfactory. The Bishop-Phelps theorem states that $\mathbb{R}$ has property B. In the paper mentioned above, Lindenstrauss used the Bishop-Phelps theorem to prove the following. Suppose X is a Banach space such that there exists a set $\{x_\alpha, f_\alpha : \alpha \in A\}$ with $x_\alpha \in X$, $f_\alpha \in X^*$, and suppose that there exists $\lambda < 1$ such that

(i) $\|f_\alpha\| = 1$ for every $\alpha \in A$, and $\|x\| = \sup_{\alpha \in A} |f_\alpha(x)|$ for every $x \in X$;

(ii) $\|x_\alpha\| = f_\alpha(x_\alpha) = 1$ for every $\alpha \in A$ and $|f_\alpha(x_\beta)| < \lambda$ whenever $\alpha \neq \beta$.

Then in that case the space X has property B. This is the main positive result of a general nature about property B. Properties related to property B have also been considered by Finet and Schachermayer [20] and Finet [19].

Lindenstrauss asked [32] whether every reflexive space has property B. Johnson and Wolfe [28] asked the same question, and also asked whether there is any classical space which does not have property B. The result of Lindenstrauss above was used by Partington [41] to show that any Banach space can be equivalently renormed so as to have property B. He remarked that it was not known whether ℓ_1 or ℓ_2 have the property.

In this chapter it is shown that ℓ_2 does not have property B. In fact the proof, which is straightforward, applies equally to ℓ_p whenever $1 < p < \infty$, so the result is presented in that form.

Theorem 10·1. *When $1 < p < \infty$, the space ℓ_p does not have property B.*

Proof. Let X be the space of real sequences $(a_i)_1^\infty$ such that if $(a_i^*)_1^\infty$ is the positive decreasing rearrangement of $(a_i)_1^\infty$ (that is, the decreasing rearrangement of $(|a_i|)_1^\infty$) then

$$\lim_{N\to\infty} \left(\sum_{i=1}^{N} a_i^* \Big/ \sum_{i=1}^{N} i^{-1} \right) = 0.$$

For $\mathbf{a} = (a_i)_1^\infty \in X$, set

$$\|\mathbf{a}\| = \max_{N\to\infty} \left(\sum_{i=1}^{N} a_i^* \Big/ \sum_{i=1}^{N} i^{-1} \right)$$

It is easily seen that $\|.\|$ is a 1-symmetric norm on X.

Finally, let $T: X \to \ell_p$ be the formal identity map. We shall prove the following easy facts.

(i) X is a Banach space;

(ii) T is a bounded linear operator;

(iii) If $S: X \to \ell_p$ is any norm-attaining map, then there is some $n \in \mathbb{N}$ such that $S\mathbf{e}_n = 0$.

Since our main interest is in the map T, we shall prove the first fact last. We shall prove that T is bounded by showing that if $\mathbf{a} \in X$ and $\|\mathbf{a}\| = 1$, then $\|T\mathbf{a}\| \leqslant (\sum_1^\infty i^{-p})^{1/p}$. In fact the following inequality is obviously enough: suppose $(a_i)_1^n$ is a monotone decreasing sequence of positive numbers such that for all $m \leqslant n$, $\sum_1^m a_i \leqslant \sum_1^m i^{-1}$. Then $\sum_1^n a_i^p \leqslant \sum_1^n i^{-p}$.

To prove this, pick a sequence $(a_i)_1^n$ satisfying these conditions such that $\sum_1^n a_i^p$ is maximal. If $a_i = i^{-1}$ for all i then there is nothing to prove. Otherwise pick a minimal j such that $a_j < j^{-1}$, and a maximal k such that $\sum_{i=1}^k a_i = \sum_{i=1}^k i^{-1}$. There must exist such a k and it is greater than j, for if it were not, we could increase a_j to produce a new sequence satisfying the given conditions with $\sum_1^n a_i$ larger than before, contradicting our maximality assumption. Now it is obvious that $a_j < a_{j-1}$ and $a_k > a_{k+1}$. Hence, for sufficiently small ϵ, we may replace a_j by $a_j + \epsilon$ and a_k by $a_k - \epsilon$ to obtain a new sequence satisfying the above conditions. But this again contradicts the maximality of $\sum_1^n a_i^p$, since

$$\begin{aligned}(a_j+\epsilon)^p &+ (a_k-\epsilon)^p \\ &= a_j^p + a_k^p + p(a_j^{p-1} - a_k^{p-1})\epsilon + p(p-1)/2(a_j^{p-2} + a_k^{p-2})\epsilon^2 + o(\epsilon^2) \\ &> a_j^p + a_k^p\end{aligned}$$

when ϵ is sufficiently small, as $a_j \geqslant a_k$.

We shall now turn our attention to the third fact, which is of course the important one. We use the fact that, for $1 < p < \infty$, ℓ_p is strictly convex. Suppose then that $S: X \to \ell_p$ is a norm-attaining operator. Pick $\mathsf{a} = (a_i)_1^\infty \in X$ such that $\|\mathsf{a}\| = 1$ and $\|S\mathsf{a}\| = \|S\|$. Without loss of generality $(a_i)_1^\infty$ is a positive monotone decreasing sequence. Note that, for every n, $a_n \leqslant n^{-1}\sum_1^n i^{-1}$, and certainly $a_n \leqslant a_1 \leqslant 1$. If the sequence $(a_i)_1^\infty$ is eventually zero, then let n be the minimal index such that $a_n = 0$. Then for $\delta \leqslant n^{-1}$ we have that $\|\mathsf{a} \pm \delta \mathsf{e}_n\| = \|\mathsf{a}\| \leqslant 1$. Suppose on the other hand that a_n is never zero. Find m such that for any $n \geqslant m$, $\sum_1^n a_i + 1 \leqslant \sum_1^n i^{-1}$. There must certainly exist an $n \geqslant m$ such that $a_n < a_{n-1}$, because if there were not, the expressions $\sum_1^N a_i / \sum_1^N i^{-1}$ would not even be bounded. For such an n we have that if $0 < \delta < a_{n-1} - a_n \leqslant 1$, then $\|\mathsf{a} \pm \delta \mathsf{e}_n\| \leqslant 1$. (We have used the fact that $a_{n-1} - a_n \leqslant 1$). Now since ℓ_p is strictly convex, either $S\mathsf{e}_n = 0$ or

$$\|S(\mathsf{a} + \delta \mathsf{e}_n)\| + \|S(\mathsf{a} - \delta \mathsf{e}_n)\| > 2\,\|S\mathsf{a}\|\,.$$

The second possibility is ruled out because S was supposed to attain its norm at a.

We shall now check that X is indeed a Banach space. First, let Y be the space of real sequences $(a_i)_1^\infty$ such that $\sum_1^N a_i^* / \sum_1^N i^{-1}$ is bounded above, and let the norm on Y be given by the supremum of $\sum_1^N a_i^* / \sum_1^N i^{-1}$. Then X is a subspace of the normed space Y. The space Y has been used by several authors, and is the dual of a Lorentz sequence space. We shall give a direct proof that it is complete and that X is a closed subspace of it.

First, for $N \in \mathbb{N}$ let the norm $\|.\|_N$ on Y be defined by setting $\|\mathsf{a}\|_N = \sum_1^N a_i^* / \sum_1^N i^{-1}$. Now, suppose $(\mathsf{a}_n)_1^\infty$ is a Cauchy sequence in Y. Then since the norm $\|.\|$ dominates the ℓ_∞-norm, the sequence $(\mathsf{a}_n)_1^\infty$ is a Cauchy sequence in each coordinate, so it has a pointwise limit, a, say. Pick an m such that given any $n \geqslant m$, $\|\mathsf{a}_m - \mathsf{a}_n\| \leqslant 1$. Then, in particular, $\|\mathsf{a}_n\|_N \leqslant \|\mathsf{a}_m\|_N + 1 \leqslant \|\mathsf{a}_m\| + 1$ for any N and any $n \geqslant m$. Now suppose that a were not in Y. We would then be able to find N such that $\|\mathbf{a}\|_N > \|\mathbf{a}_m\| + 2$. Since $\mathbf{a}$ is the pointwise limit of the $\mathbf{a}_n$, we would then easily find n large enough for $\|\mathbf{a}_n\|_N$ to be greater than $\|\mathbf{a}_M\| + 1$. This contradiction establishes that $\mathbf{a}$ is in the space Y.

We must now check that $\mathbf{a}_n \longrightarrow \mathbf{a}$ in norm. By subtracting the vector $\mathbf{a}$ from all of the $\mathbf{a}_n$ we may assume without loss of generality that in fact $\mathbf{a} = 0$. So suppose $(\mathbf{a}_n)_1^\infty$ is a Cauchy sequence and tends pointwise to 0. Pick M such that for any $m, n \geqslant M$, $\|\mathbf{a}_m - \mathbf{a}_n\| \leqslant \epsilon$. Now suppose that for some N and some $m \geqslant M$, $\|\mathbf{a}_m\|_N \geqslant 2\epsilon$. For ease of notation we will take $\mathbf{a}_m$ to be a positive decreasing vector. We then have that for any $n \geqslant M$, $\|\mathbf{a}_m - \mathbf{a}_n\|_N \leqslant \|\mathbf{a}_m - \mathbf{a}_n\| \leqslant \epsilon$, from which it follows that the sum of the first N coordinates of $\mathbf{a}_n$ is at least $\epsilon \sum_1^N i^{-1}$. But then $\mathbf{a}_n$ does not tend to zero pointwise. This completes the proof that Y is a Banach space.

Our final task is to show that X is a closed subspace of Y. If this were not the case, then we would be able to find $\mathbf{a}_1, \mathbf{a}_2, \ldots \in X$ such that $\mathbf{a}_n \longrightarrow \mathbf{a} \in Y \setminus X$. Then since $\mathbf{a} \notin X$, we could find $\epsilon > 0$ and $N_1 < N_2 < \ldots$ such that $\|\mathbf{a}\|_{N_i} \geqslant \epsilon$. Pick n large enough for $\|\mathbf{a} - \mathbf{a}_n\|$ to be at most $\epsilon/2$. Then for all i, $\|\mathbf{a} - \mathbf{a}_n\|_{N_i} \leqslant \epsilon/2$, which implies that for all i, $\|\mathbf{a}_n\|_{N_i} \geqslant \epsilon/2$. But then $\mathbf{a}_n \notin X$, contradicting our assumption above. It follows that X is indeed a closed subspace of Y and is thus a Banach space. □

This shows then that a reflexive space need not have property B. Many interesting questions are still left unanswered, however. The most natural is perhaps whether ℓ_1 has the property. Another question, described by Johnson and Wolfe as "the most irritating problem" in the area, is whether ℓ_2^2 has property B. I think that it probably has, and that a proof of this could be adapted to show that any finite-dimensional space has it.

CHAPTER 11

AN ISOMORPHIC CONSTRUCTION OF THE CLASSICAL SPACES

There are two common ways of defining random finite-dimensional normed spaces. One is to choose a certain number of points at random from the unit sphere of ℓ_2^n and to take their convex hull. The other, dual, method is to take the convex polytope defined by the hyperplanes tangent to the sphere at the given points. However, these spaces have properties, for example small type or cotype constants, which general spaces cannot be expected to have. The first result of this chapter originated in an attempt to mix these two methods to obtain a natural random space without small type or cotype constants. The attempt failed completely because, far from finding a space with large type and cotype constants, we found that, if we started with $\{\pm \mathrm{e}_1, \ldots, \pm \mathrm{e}_n\}$ and mixed the two constructions above in a very natural way, we obtained spaces which were uniformly isomorphic to ℓ_2^n. This observation seemed interesting, however, and suggested further results which are the topic of this chapter.

§11.1 An Isomorphic Construction of ℓ_2^n

For $n \geqslant 1$ let $\mathcal{F}_n$ stand for the set of normed spaces of the form $(\mathbb{R}^n, \|.\|)$ for which the standard basis $\mathrm{e}_1, \ldots, \mathrm{e}_n$ is 1-symmetric. Let us also define two operators S and T which map $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ to $\bigcup_{n=1}^{\infty} \mathcal{F}_{n+1}$ as follows. Given $X \in \mathcal{F}_n$, we set $T(X)$ to be the unique normed space in $\mathcal{F}_{n+1}$ such that for any $\mathrm{a} = \sum_1^{n+1} a_i \mathrm{e}_i \in \mathbb{R}^{n+1}$ with $a_1 \geqslant \ldots \geqslant a_{n+1} \geqslant 0$, $\|\mathrm{a}\|_{T(X)} = \|\sum_1^n a_i \mathrm{e}_i\|_X$. The norm on $T(X)$ is as small as it can be given that it is 1-symmetric and that its restriction to any n coordinates yields the norm on X. The operator S is defined in the opposite way: the norm on $S(X)$ is as big as it can be under the same conditions. Thus $S(X) = (T(X^*))^*$. Alternatively, given the unit ball of X, embed it into $\mathbb{R}^{n+1}$ in the $n+1$ natural ways (up to symmetry) given by ignoring each of the coordinates in turn. The convex hull of these $n+1$ copies is the unit ball of $S(X)$.

Now, if we start with the unique normalized 1-dimensional space, which we shall call $\mathbb{R}$, and apply the operator S $n-1$ times, we clearly end up with ℓ_1^n. Similarly, if we apply T $n-1$ times we obtain ℓ_∞^n. What happens if we alternate S with T? The answer to this is the main result of this section and the motivation for the rest of the chapter.

Theorem 11·1. *Let X be a space in $\mathcal{F}_n$ obtained from the space $\mathbb{R}$ by applying the operators S and T alternately. Then $d(X, \ell_2^n) = \sqrt{2}$.*

The proof of this theorem is based on two lemmas: the first is trivial and the second also easy.

Lemma 11·2. *If $X, Y \in \mathcal{F}_n$ and $\|\mathbf{a}\|_X \leqslant \|\mathbf{a}\|_Y$ for any $\mathbf{a} \in \mathbb{R}^n$, then $\|\mathbf{a}'\|_{SX} \leqslant \|\mathbf{a}'\|_{SY}$ and $\|\mathbf{a}'\|_{TX} \leqslant \|\mathbf{a}'\|_{TY}$ for any $\mathbf{a}' \in \mathbb{R}^{n+1}$.* □

Lemma 11·3. *Suppose $X = \ell_2^n$. Then for any $\mathbf{a} \in \mathbb{R}^{n+2}$,*

$$\|\mathbf{a}\|_{\ell_2^{n+2}} \leqslant \|\mathbf{a}\|_{TSX} \,.$$

Before proving Lemma 11.3, let us see why the two lemmas are sufficient to prove the main result. The following unconventional notation will be extremely useful. Given $n \in \mathbb{N}$, two spaces $X, Y \in \mathcal{F}_n$ and positive constants c_1 and c_2, we shall write $c_1 X < c_2 Y$ if $c_1 \|\mathbf{a}\|_X \leqslant c_2 \|\mathbf{a}\|_Y$ for any vector $\mathbf{a} \in \mathbb{R}^n$.

First let us consider the space $(TS)^k(\mathbb{R})$ for some $k \in \mathbb{N}$. It is clear from the two lemmas that $(TS)^k(\mathbb{R}) > \ell_2^{2k+1}$. Moreover, since $T(\mathbb{R}) > 2^{-1/2}\ell_2^2$, we also have that $(TS)^k T(\mathbb{R}) > 2^{-1/2}\ell_2^{2k+2}$, and hence that $S(TS)^{k-1}T(\mathbb{R}) = (ST)^k(\mathbb{R}) >$

$2^{-1/2}\ell_2^{2k+1}$. However, $(TS)^k(\mathsf{R})$ and $(ST)^k(\mathsf{R})$ are dual to each other, so we have the relations

$$2^{-1/2}\ell_2^{2k+1} < (ST)^k(\mathsf{R}) < \ell_2^{2k+1} < (TS)^k(\mathsf{R}) < 2^{1/2}\ell_2^{2k+1}$$

It follows immediately also that

$$2^{-1/2}\ell_2^{2k} < T(ST)^{k-1}(\mathsf{R}) < \ell_2^{2k} < S(TS)^{k-1}(\mathsf{R}) < 2^{1/2}\ell_2^{2k} .$$

This establishes Theorem 11.1, as long as Lemma 11.3 is true.

Proof of Lemma 11·3. Throughout this proof, all sequences of real numbers are to be assumed to be positive and decreasing. Writing $X = \ell_2^n$ we have, for any $\mathbf{a} \in \mathsf{R}^{n+1}$ and $\mathbf{b} \in \mathsf{R}^{n+2}$,

$$\|\mathbf{a}\|_{SX} = \max\left\{\sum_1^{n+1} f_i a_i \,:\, \sum_1^n f_i^2 \leqslant 1\right\}$$

and

$$\|\mathbf{b}\|_{TSX} = \max\left\{\sum_1^{n+1} f_i b_i \,:\, \sum_1^n f_i^2 \leqslant 1\right\}$$

Hence it is enough to show that for any $a_1 \geqslant \ldots \geqslant a_{n+2} \geqslant 0$ with $\sum_1^{n+2} a_i^2 = 1$, we can find $f_1 \geqslant \ldots \geqslant f_{n+1} \geqslant 0$ such that $\sum_1^n f_i^2 \leqslant 1$ and $\sum_1^{n+1} f_i a_i \geqslant 1$. We do this by setting $f_i = \lambda a_i$ for $1 \leqslant i \leqslant n$, where $\lambda = (1 - a_{n+1}^2 - a_{n+2}^2)^{-1/2}$, and setting $f_{n+1} = f_n$. Then certainly $\sum_1^n f_i^2 = 1$, and moreover

$$\begin{aligned}\sum_1^{n+1} f_i a_i &= \lambda \sum_1^n a_i^2 + \lambda a_n a_{n+1} \\ &= \frac{1 - a_{n+1}^2 - a_{n+2}^2 + a_n a_{n+1}}{(1 - a_{n+1}^2 - a_{n+2}^2)^{1/2}} \\ &\geqslant \frac{1 - a_{n+2}^2}{(1 - 2a_{n+2}^2)^{1/2}} \geqslant 1.\end{aligned}$$

□

The proof of Theorem 11.1 can be generalized very easily to approximate ℓ_p^n in a similar way when p or its conjugate is an integer. However, with a little more work, one can approximate ℓ_p^n for an arbitrary p. We shall do this in Section 11.3. First, we show how to calculate the norm of a vector in a space of the form $U_1(U_2(\ldots U_{n-1}(\mathbb{R})\ldots))$ where each U_i is either S or T. We shall call such a space an ST-space.

§11.2 Calculating the Norm

Let $X \in \mathcal{F}_n$ and let $\mathbf{a} = \sum_1^{n+1} a_i\mathbf{e}_i$ be a vector in $\mathbb{R}^{n+1}$ such that $a_1 \geqslant \ldots \geqslant a_{n+1} \geqslant 0$. Then by definition $\|\mathbf{a}\|_{TX} = \|\sum_1^n a_i\mathbf{e}_i\|_X$. We shall now show how to calculate $\|\mathbf{a}\|_{SX}$. We shall use Lemma 2.2, writing $\mathbf{a} \prec \mathbf{b}$ instead of $\mathbf{a} \leqslant_2 \mathbf{b}$. As in Chapter 2, the important use of Lemma 2.2 will be to show that if $\|.\|$ is a 1-symmetric norm on $\mathbb{R}^n$, then $\mathbf{a} \prec \mathbf{b}$ implies that $\|\mathbf{a}\| \leqslant \|\mathbf{b}\|$.

Given $X \in \mathcal{F}_n$ and $\mathbf{a} \in \mathbb{R}^{n+1}$, let us define a function $\|.\|'$ on $\mathbb{R}^{n+1}$ by

$$\|\mathbf{a}\|' = \inf\{\|\mathbf{b}\|_X : \mathbf{b} \in X, \mathbf{a} \prec \mathbf{b}\}$$

where $\mathbf{a} \prec \mathbf{b}$ means of course that $\mathbf{a} \prec \mathbf{b}'$, where $\mathbf{b}'$ is the image of $\mathbf{b}$ in $\mathbb{R}^{n+1}$ under the formal identity map. Recall that, given a vector $\mathbf{b}$, we denote the decrasing rearrangement of $\mathbf{b}$ by $\mathbf{b}^*$.

Lemma 11·4. *Let $X \in \mathcal{F}_n$ and $\mathbf{a} \in \mathbb{R}^{n+1}$. Then $\|\mathbf{a}\|_{SX} = \|\mathbf{a}\|'$.*

Proof. Given two vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ in $\mathbb{R}^n$, let $\mathbf{b}_i \in \mathbb{R}^{n+1}$ be such that $\mathbf{a}_i \prec \mathbf{b}_i$ and $\|\mathbf{b}_i\|_X$ is minimal, for $i = 1, 2$. Then we certainly have $\mathbf{a}_1 + \mathbf{a}_2 \prec \mathbf{b}_1^* + \mathbf{b}_2^*$, and thus, by Lemma 2.2 and the fact that X is a 1-symmetric space,

$$\begin{aligned}\|\mathbf{a}_1 + \mathbf{a}_2\|' &\leqslant \|\mathbf{b}_1^* + \mathbf{b}_2^*\|_X \\ &\leqslant \|\mathbf{b}_1^*\|_X + \|\mathbf{b}_2^*\|_X \\ &= \|\mathbf{b}_1\|_X + \|\mathbf{b}_2\|_X \\ &= \|\mathbf{a}_1\|' + \|\mathbf{a}_2\|'\end{aligned}$$

It follows that $\|.\|'$ is a norm. Since its unit ball is clearly contained in that of SX, and contains the $n+1$ natural images of $B(X)$ in $\mathbb{R}^{n+1}$, we have that $\|.\|' = \|\cdot\|_{SX}$, as stated. □

By an easy compactness argument, the infimum of $\{\|\mathbf{b}\|_X : \mathbf{b} \in X, \mathbf{a} \prec \mathbf{b}\}$ is actually attained. In the next lemma we find a vector at which it is attained.

Lemma 11·5. *Let $X \in \mathcal{F}_n$, let $\mathbf{a} = \sum_1^{n+1} a_i \mathbf{e}_i$ be a vector in $\mathbb{R}^{n+1}$ for which $a_1 \geqslant \ldots \geqslant a_{n+1} > 0$ and let $\gamma \geqslant 0$ be the unique number for which*

$$\sum_1^n (a_i \vee \gamma) = \sum_1^{n+1} a_i$$

Then $\|\mathbf{a}\|_{SX} = \|\sum_1^n (a_i \vee \gamma)\mathbf{e}_i\|_X$.

Proof. We insist that $a_{n+1} > 0$ for convenience: it is obvious how to calculate $\|\mathbf{a}\|_{SX}$ if $a_{n+1} = 0$. This gives us the uniqueness of γ. Let us set $\mathbf{a}' = \sum_1^n (a_i \vee \gamma)\mathbf{e}_i$. We need to show that if $\mathbf{b} \in \mathbb{R}^n$ and $\mathbf{a} \prec \mathbf{b}$ then $\mathbf{a}' \prec \mathbf{b}$. It is clear that $\mathbf{a} \prec \mathbf{a}'$, and by Lemma 2.2 and the symmetry of X, we will also have $\|\mathbf{a}'\|_X \leqslant \|\mathbf{b}\|_X$. We may clearly suppose that $b_i = b_i^*$ for every i and that $\sum_1^n b_i = \sum_1^{n+1} a_i$. Since $\mathbf{a} \prec \mathbf{b}$ we then have, for every $1 \leqslant k \leqslant n$, that $\sum_k^n b_i \leqslant \sum_k^{n+1} a_i$. Let k be maximal such that $a_k > \gamma$. Clearly $k < n$, and $\gamma = (n-k)^{-1} \sum_{k+1}^{n+1} a_i$. For $l \leqslant k$ it is obvious that $\sum_1^l (a_i \vee \gamma) \leqslant \sum_1^l b_i$. When $l > k$ we have

$$\begin{aligned}
\sum_1^l (a_i \vee \gamma) &= \sum_1^{n+1} a_i - (n-l)\gamma \\
&= \sum_1^{n+1} a_i - \frac{n-l}{n-k} \sum_{k+1}^{n+1} a_i \\
&\leqslant \sum_1^{n+1} a_i - \frac{n-l}{n-k} \sum_{k+1}^{n} b_i \\
&\leqslant \sum_1^{n+1} a_i - \sum_{l+1}^{n} b_i = \sum_1^{l} b_i
\end{aligned}$$

This proves the lemma. □

To conclude this section we shall show how to calculate the norm of the vector $(3,3,3,2,2,2,1,1,1)$ in the spaces $(ST)^4(\mathbb{R})$ and $(TS)^4(\mathbb{R})$ by repeated application of Lemma 11.5 and the definition of the operator T. At each stage, we replace the vector we have by one of length one less, while preserving its norm. At a "T" stage, we simply remove the last coordinate. At an "S" stage, we apply Lemma 11.5.

$$
\begin{array}{llllllllll}
 & 3 & 3 & 3 & 2 & 2 & 2 & 1 & 1 & 1 \\
\xrightarrow{S} & 3 & 3 & 3 & 2 & 2 & 2 & 1\frac{1}{2} & 1\frac{1}{2} & \\
\xrightarrow{T} & 3 & 3 & 3 & 2 & 2 & 2 & 1\frac{1}{2} & & \\
\xrightarrow{S} & 3 & 3 & 3 & 2\frac{1}{2} & 2\frac{1}{2} & 2\frac{1}{2} & & & \\
\xrightarrow{T} & 3 & 3 & 3 & 2\frac{1}{2} & 2\frac{1}{2} & & & & \\
\xrightarrow{S} & 3\frac{1}{2} & 3\frac{1}{2} & 3\frac{1}{2} & 3\frac{1}{2} & & & & & \\
\xrightarrow{T} & 3\frac{1}{2} & 3\frac{1}{2} & 3\frac{1}{2} & & & & & & \\
\xrightarrow{S} & 5\frac{1}{4} & 5\frac{1}{4} & & & & & & & \\
\xrightarrow{T} & 5\frac{1}{4} & & & & & & & &
\end{array}
\qquad
\begin{array}{llllllllll}
 & 3 & 3 & 3 & 2 & 2 & 2 & 1 & 1 & 1 \\
\xrightarrow{T} & 3 & 3 & 3 & 2 & 2 & 2 & 1 & 1 & \\
\xrightarrow{S} & 3 & 3 & 3 & 2 & 2 & 2 & 2 & & \\
\xrightarrow{T} & 3 & 3 & 3 & 2 & 2 & 2 & & & \\
\xrightarrow{S} & 3 & 3 & 3 & 3 & 3 & & & & \\
\xrightarrow{T} & 3 & 3 & 3 & 3 & & & & & \\
\xrightarrow{S} & 4 & 4 & 4 & & & & & & \\
\xrightarrow{T} & 4 & 4 & & & & & & & \\
\xrightarrow{S} & 8 & & & & & & & &
\end{array}
$$

Notice that $\|(3,3,3,2,2,2,1,1,1)\|_2 = \sqrt{42}$, and, as must be the case by the proof of Theorem 11.1, $\sqrt{21} \leqslant 5\frac{1}{4} \leqslant \sqrt{42} \leqslant 8 \leqslant \sqrt{84}$.

§11.3 Approximating ℓ_p^n by an ST-Space

In order to show that ℓ_p^n can be approximated by an ST-space, we shall prove an inequality which is similar to, but a little more sophisticated than, Lemma 11.3. However, our main difficulty is a notational one: it makes our proof much tidier if we generalize the class of spaces we deal with, and also the operators S and T, and for this we need a number of definitions.

Given a real number $t > 0$, let $F[0,t]$ denote the vector space of step functions from the closed interval $[0,t]$ to $\mathbb{R}$, and let $\mathcal{G}_t$ denote the set of normed spaces

$(F[0,t], \|.\|)$ such that $\|f\| = \|f^*\|$ for any $f \in F[0,t]$, where f^* is the decreasing rearrangement of f. Note that we do not ask for these normed spaces to be complete. If $s > t$ and $f \in F[0,s]$ we shall write f_t for the restriction of the function f to the interval $[0,t]$. For $n \in \mathbb{N}$ let I_n denote the linear map from $\mathbb{R}^n$ to $F[0,n]$ determined by $e_j \mapsto \chi_{[j-1,j)}$. We shall prove a result about norms on $F[0,t]$ but our interest will eventually be in subspaces of the form $I_n(\mathbb{R}^n)$.

We now define operators which are similar to S and T, but which map $\mathcal{G}_{t_1}$ to $\mathcal{G}_{t_2}$, where $t_1 < t_2$. Given $\alpha > 1$, let us define an operator $T_\alpha : \bigcup_{t>0} \mathcal{G}_t \longrightarrow \bigcup_{t>0} \mathcal{G}_{\alpha t}$ as follows. If $X \in \mathcal{G}_t$ and $f \in F[0,\alpha t]$ then $\|f\|_{T_\alpha(X)} = \|(f^*)_t\|_X$. Thus, the norm on the space $T_\alpha(X)$ is as small as it can be given that $T_\alpha(X)$ is in the set $\mathcal{G}_{\alpha t}$ and that the norm on $T_\alpha(X)$ coincides with that on X for functions supported on the interval $[0,t]$. As in the discrete case, S_α is defined in the opposite way: the norm on $S_\alpha(X)$ is as large as it can be under the same conditions. The following lemma we state without proof, since the analogy with Lemma 11.5 is very close.

Lemma 11·6. *Let $t > 0$, $\alpha > 1$ and $X \in \mathcal{G}_t$. Let $f \in F[0,\alpha t]$ be a step function satisfying $f = f^*$ and $f(s) > 0$ for some $s > t$. Then if $\gamma > 0$ is the unique number for which*

$$\int_0^t (f(t) \vee \gamma)\, dt = \int_0^{\alpha t} f(t)\, dt\,,$$

we have $\|f\|_{S_\alpha(X)} = \|(f \vee \gamma)_t\|_X$. □

We shall also need the analogue of Lemma 11.2. Like Lemma 11.2 it is trivial, so we state it without proof. As in the first section, we use an unconventional but useful notation. If X and Y are spaces in $\mathcal{G}_t$ and c_1 and c_2 are positive constants, then we shall write $c_1 X < c_2 Y$ if $c_1 \|f\|_X \leqslant c_2 \|f\|_Y$ for every $f \in F[0,t]$. For $1 \leqslant p \leqslant \infty$ and $t \geqslant 1$ let ℓ_p^t denote the space $(F[0,t], \|.\|_p)$, where $\|.\|_p$ is the usual norm on $L_p[0,t]$ restricted to $F[0,t]$. If $X \in \mathcal{G}_t$ and $c > 0$, it will be convenient to denote by cX the space $Z \in \mathcal{G}_t$ defined by $\|f\|_Z = c\|f\|_X$ for every $f \in F[0,t]$.

Lemma 11·7. *Let $t > 0$, $\alpha > 1$ and let $X_1, X_2 \in \mathcal{G}_t$. Then if $X_1 < X_2$ we have $S_\alpha X_1 < S_\alpha X_2$ and $T_\alpha X_1 < T_\alpha X_2$.* □

The next lemma is the main lemma of the section and the only one in the chapter whose proof requires much work.

Lemma 11·8. *Let $1 < p < \infty$, $t > 0$, $\alpha > 1$ and $\beta = \alpha^{p-1}$. Then*

$$S_\alpha(T_\beta(\ell_p^t)) < \ell_p^{\alpha\beta t} < T_\beta(S_\alpha(\ell_p^t)) .$$

Proof. We shall prove the right-hand inequality only. The left-hand one can be proved by a similar argument, or else by using duality.

Suppose then that the right-hand inequality does not hold. In that case we can find $N \in \mathbb{N}$ and sequences $0 = x_0 < x_1 < \ldots < x_N = \alpha\beta t$ and $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_N \geqslant 0$ such that, setting $A_i = [x_{i-1}, x_i)$ and $f = \sum_1^N \lambda_i \chi_{A_i}$, we have

$$\int_0^{\alpha\beta t} f(x)^p \, dx > \int_0^t (f(x) \vee \gamma)^p \, dx$$

where

$$\int_0^t (f(x) \vee \gamma) \, dt = \int_0^{\alpha t} f(x) \, dx .$$

Let i_0 be the minimal index i for which $\lambda_i \leqslant \gamma$ and set $s = x_{i_0 - 1}$. Thus $f(x) \leqslant \gamma$ if and only if $x \geqslant s$. Without loss of generality, there exists i_1 such that $x_{i_1} = \alpha t$. We have

$$\gamma(t - s) = \int_s^{\alpha t} f(x) \, dx ,$$

and, by assumption,

$$\begin{aligned} &\int_0^{\alpha\beta t} f(x)^p \, dx - \int_0^t (f(x) \vee \gamma)^p \, dx \\ &\qquad = \int_s^{\alpha\beta t} f(x)^p \, dx - (t - s)^{-p} \Big(\int_s^{\alpha t} f(x) \, dx \Big)^p (t - s) > 0 . \end{aligned}$$

Let us write A for $(t - s)^{p-1} \int_s^{\alpha\beta t} f(x)^p \, dx$ and B for $\left(\int_s^{\alpha t} f(x) \, dx \right)^p$. Our hypothesis now reads that $A > B$. If we fix the sequence $x_0, \ldots, x_N$, we may assume, by a simple compactness argument, that $\lambda_1 \geqslant \ldots \geqslant \lambda_N \geqslant 0$ have been chosen so as to maximize the ratio A/B. Let us pick $i_0 \leqslant i < i_1$. We find

$$\frac{\partial}{\partial \lambda_i}\frac{A}{B} = \frac{1}{B^2}\Big(B(t-s)^{p-1}(x_i - x_{i-1})p\lambda_i^{p-1}$$
$$-Ap\Big(\int_s^{\alpha t} f(x)\,dx\Big)^{p-1}(x_i - x_{i-1})\Big) .$$

But since $\lambda_i \leqslant \gamma$, we have also $\left(\int_s^{\alpha t} f(x)\,dx\right)^{p-1} \geqslant (t-s)^{p-1}\lambda_i^{p-1}$. Since $A > B$ we obtain that $\frac{\partial}{\partial \lambda_i}\frac{A}{B} < 0$. A simple calculation shows that if we decrease λ_i very slightly, then the value of i_0 does not change. If the change to λ_i is small enough, and the resulting sequence is still a decreasing one, then it follows that A/B was not maximized by the original sequence. Hence, we obtain that $\lambda_{i_0} = \lambda_{i_0+1} = \ldots = \lambda_{i_1}$. It is obvious also that $\lambda_{i_1} = \ldots = \lambda_N$. Without loss of generality they all take the value 1.

It remains to show that

$$(\alpha\beta t - s)(t-s)^{p-1} \leqslant (\alpha t - s)^p$$

By applying the weighted arithmetic-geometric-mean inequality twice and using the fact that $\beta = \alpha^{p-1}$, we have, writing q for the conjugate index of p,

$$\begin{aligned}(\alpha\beta t - s)^{1/p}(t-s)^{1/q} &= \alpha t\left(1 - \frac{s}{\alpha\beta t}\right)^{1/p}\left(1 - \frac{s}{t}\right)^{1/q} \\ &\leqslant \alpha t\left(1 - \frac{s}{p\alpha\beta t} - \frac{s}{qt}\right) \\ &\leqslant \alpha t\left(1 - \frac{s}{\alpha t}\right) = \alpha t - s\end{aligned}$$

The result follows on raising both sides to the power p. □

The next lemma is similar to Lemma 11.8, but it is rather easier.

Lemma 11·9. *Let $t > 0$, $X \in \mathcal{G}_t$ and $\alpha, \beta > 1$. Then $S_\alpha T_\beta X < T_\beta S_\alpha X$.*

Proof. Let $f \in F[0, \alpha\beta t]$. Without loss of generality $f = f^*$. Then $\|f\|_{S_\alpha T_\beta X} = \|(f \vee \gamma)_t\|_X$, where γ satisfies

$$\int_0^{\beta t} (f(x) \vee \gamma)\, dx = \int_0^{\alpha\beta t} f(x)\, dx$$

and $\|f\|_{T_\beta S_\alpha X} = \|(f \vee \gamma')_t\|_X$, where γ' satisfies

$$\int_0^{t} (f(x) \vee \gamma')\, dx = \int_0^{\alpha t} f(x)\, dx.$$

It is therefore enough to show that $\int_0^t (f(x) \vee \gamma)\, dx \leqslant \int_0^{\alpha t} f(x)\, dx$, and thus that $\gamma' \geqslant \gamma$. Let $s = \inf\{x: f(x) \leqslant \gamma\}$. If $s \geqslant t$, the result is trivial. Otherwise, we wish to show that $(t-s)\gamma \leqslant \int_s^{\alpha t} f(x)\, dx$, given that $\gamma = (\beta t - s)^{-1} \int_s^{\alpha\beta t} f(x)\, dx$. But since f is a positive decreasing function and $\alpha, \beta > 1$, we have

$$\frac{\int_s^{\alpha\beta t} f(x)\, dx}{\int_s^{\alpha t} f(x)\, dx} \leqslant \frac{\alpha\beta t - s}{\alpha t - s} \leqslant \frac{\beta t - s}{t - s}$$

which proves the lemma. □

In fact, it is not hard to see that the only functions whose norms are equal in the two spaces $S_\alpha T_\beta X$ and $T_\beta S_\alpha X$ are those which are either supported on the interval $[0, t]$, or have constant modulus (except on a set of measure zero) on the interval $[0, \alpha\beta t]$.

Lemma 11·10. *Let $s, t > 0$, let $1 < p < \infty$, let q be the conjugate index of p and let $\alpha, \beta > 1$. Then $S_\alpha(\alpha^{-1/q}\ell_p^s) < \ell_p^{\alpha s}$ and $\ell_p^{\beta t} < T_\beta(\beta^{1/p}\ell_p^t)$.*

Proof. The second inequality states that, for any positive decreasing function f on the interval $[0, \beta t]$,

$$\int_0^{\beta t} (f(x))^p\, dx \leqslant \beta \int_0^t (f(x))^p\, dx, \tag{1}$$

which is obvious. To prove the first one, observe that by (1), with $t = \alpha^{-p/q}s$ and $\beta = \alpha^{p/q}$, we have

$$\alpha^{-1/q}\ell_p^s < T_\beta(\ell_p^t)\ ,$$

so, by Lemma 11.7, we have

$$S_\alpha(\alpha^{-1/q}\ell_p^s) < S_\alpha T_\beta(\ell_p^t) .$$

Now $\beta = \alpha^{p/q} = \alpha^{p-1}$ so, by Lemma 11.8,

$$S_\alpha T_\beta(\ell_p^t) < \ell_p^{\alpha\beta t} = \ell_p^s . \qquad \square$$

Lemmas 11.7 to 11.10 contain the essence of the proof that we can approximate ℓ_p^n by an ST-space. The remaining arguments are all very easy, but this may not be immediately apparent. The reader is strongly advised to consider, for any space $X \in \mathcal{G}_t$ under discussion, the function $\lambda(t) = \|\chi_{[0,t]}\|_X$, and especially a graph of $\log\lambda$ against $\log t$. Note that the slope of this graph, whenever X is an ST-space, will be either zero or one. The terminology that follows is needed in order to formalize arguments which from such a diagram are simple.

Let us define an *ST-sequence* to be a sequence of the form $U_0, U_1, \ldots, U_k$, where U_0 is a space of the form $c\ell_p^s$, for some real numbers $c > 0$, $1 \leqslant p \leqslant \infty$ and $s > 0$, and each U_i is an operator of the form S_α or T_α for some α. Let us define a *generalized ST-space* to be any space of the form $U_k(U_{k-1}(\ldots U_1(U_0)\ldots))$, for some ST-sequence $U_0, U_1, \ldots, U_k$. We shall often drop the word "generalized" if no ambiguity is likely to arise. We shall say that a function $\lambda : [0,t] \longrightarrow \mathbb{R}$ is an *ST-function* if it is a function of the form $\lambda(t) = \|\chi_{[0,t]}\|_X$ for some generalized ST-space $X \in \mathcal{G}_t$. It is easy to show that a function λ is an *ST-function* if and only if for some $s > 0$, $c > 0$ and $1 \leqslant p \leqslant \infty$ the restriction of λ to the interval $[0,s]$ is the function $cx^{1/p}$ (where $x^{1/\infty}$ is understood to be 1), while the restriction of λ to the interval $[s,t]$ is piecewise linear, and the right derivative of λ at a point $x \in [s,t]$ is either zero or $\lambda(x)/x$.

Given an ST-sequence $U_0, U_1, \ldots, U_k$, we shall define *the corresponding* or *associated* ST-space to be the space $U_k(U_{k-1}(\ldots U_1(U_0)\ldots))$. If the ST-space corresponding to $U_0, U_1, \ldots, U_k$ is not the ST-space corresponding to any shorter ST-sequence, we shall say that the sequence $U_0, U_1, \ldots, U_k$ is *reduced*. If an ST-space is the space corresponding to two ST-sequences, we shall say that the two sequences are *equivalent*. There is an obvious one-to-one correspondence between reduced ST-sequences, generalized ST-spaces and ST-functions. Whenever we use the words "corresponding" or "associated" relating reduced ST-sequences, ST-spaces and ST-functions, this is the correspondence we shall mean.

Lemma 11·11. *Let $U_0, U_1, \ldots, U_k$ and $V_0, V_1, \ldots, V_k$ be two ST-sequences such that $U_0 = V_0$. Let $Y_1, Y_2 \in \mathcal{G}_t$ be the corresponding ST-spaces, and let λ_1, λ_2 the corresponding ST-functions. Then if $\lambda_1(x) \leqslant \lambda_2(x)$ for every $x \in [0, t]$ it follows that $Y_1 < Y_2$.*

Proof. This is an easy consequence of Lemma 11.9. Let W be the set of real numbers of the form $\lambda_i(x)$, for $i = 1$ or 2, such that the left derivative of λ_i at x is 1 and the right derivative is 0, and let $w_1, \ldots, w_N$ be the elements of W in increasing order. Let $s > 0$ be maximal such that λ_1 and λ_2 are equal on the interval $[0, s]$. Then $\lambda_1(s) = \lambda_2(s) = w_i$, say. Let $U_1, \ldots, U_l$ be the reduced ST-sequence corresponding to the restriction of λ_1 (or λ_2) to the interval $[0, s]$. Since s was maximal, there must be numbers $\alpha_0, \alpha_2 > 1$ and $\alpha_1 = w_{i+1}/w_i$ and an ST-sequence of either of the form

$$U_0, U_1, \ldots, U_l, T_{\alpha_0}, S_{\alpha_1}, S_{\alpha_2}$$

or of the form

$$U_0, U_1, \ldots, U_l, T_{\alpha_0}, S_{\alpha_1}, T_{\alpha_2}$$

whose corresponding ST-function is λ_1. Either way, let λ_1' be the ST-function corresponding to the sequence obtained by exchanging T_{α_0} and S_{α_1} and let Y_1' be the corresponding ST-space. Then, for every $x \in [0, t]$ we have $\lambda_1(x) \leqslant \lambda_1'(x) \leqslant$

$\lambda_2(x)$, and also, for every $x \in [0, \lambda_2^{-1}(w_{i+1})]$, we have $\lambda_1'(x) = \lambda_2(x)$. By Lemma 11.9, we obtain that $Y_1 < Y_1'$.

Hence, by an easy induction argument on i, we can transform the function λ_1 into λ_2 in a finite number of steps, in such a way that the corresponding ST-space increases at each step. It follows that $Y_1 < Y_2$, as stated. □

We now come to the main theorem of this section.

Theorem 11·12. *Let $0 < c_1 < c_2$, let λ be an ST-function on the interval $[0, t]$ such that $c_1 x^{1/p} \leqslant \lambda(x) \leqslant c_2 x^{1/p}$ for every $x \in [0, t]$ and let Y be the ST-space corresponding to λ. Then*

$$c_1 \ell_p^t < Y < c_2 \ell_p^t \ .$$

Proof. We shall prove only the left-hand inequality: the other is similar. Without loss of generality, $c_1 = 1$.

Let the reduced ST-sequence corresponding to λ be $U_0, U_1, \ldots, U_k$. It is clear that U_0 must equal $c\ell_p^s$ for some $s > 0$ and $c \geqslant 1$. If $c > 1$, let us replace U_0 by the space $S_{c^q} \ell_p^{c^{-q}s}$ the norm of which, by Lemma 11.10, is dominated by the norm on U_0. Let Y' be the ST-space corresponding to this new sequence. By Lemma 11.7 we have that $Y' < Y$, and it is clear that the ST-function λ' corresponding to Y' dominates the function $x^{1/p}$. Clearly the first term of the reduced ST-sequence corresponding to Y' is ℓ_p^s.

By an induction argument similar in style to the proof of Lemma 11.11, we can find an ST-sequence $\ell_p^s, S_{\alpha_1}, T_{\beta_1}, S_{\alpha_2}, T_{\alpha_2}, \ldots, S_{\alpha_k}, T_{\beta_k}$ such that $\beta_i = \alpha_i^{p-1}$ for $1 \leqslant i \leqslant k$, and the associated ST-function of the sequence is dominated by λ. Let X be the ST-space corresponding to this sequence. Then, by Lemmas 11.7 and 11.8, we have $\ell_p^t < X$, and, by Lemma 11.12, we have $X < Y'$. This completes the proof of the theorem. □

It remains to show that, whatever the value of p, we can approximate ℓ_p^n by an ST-space in the sense of Section 11.1, that is, by a non-generalized ST-space. The importance of Theorem 11.12 is that we only have to check that the corresponding ST-function approximates the function $x^{1/p}$.

Theorem 11·13. *Let $1 \leqslant p \leqslant \infty$. Then, for any $n \in \mathbb{N}$, there exists a non-generalized ST-space X such that $d(X, \ell_p^n) < 3/2$.*

Proof. By duality, we may assume that $1 \leqslant p \leqslant 2$. Given a non-generalized ST-space $X = U_{n-1}(\ldots U_2(U_1(\mathbb{R}))\ldots) \in \mathcal{F}_n$, let us set $V_0 = \ell_p^1$, and set V_i to be $S_{1+1/i}$ or $T_{1+1/i}$, according to whether U_i is S or T, for $1 \leqslant i \leqslant n-1$. Let Y be the generalized ST-space corresponding to the ST-sequence $V_0, V_1, \ldots, V_{n-1}$. Then the linear map which takes $\mathrm{e}_i \in \mathbb{R}^n$ to $\chi_{[i-1,i)}$ is easily seen to be an isometric embedding of X into Y.

By Theorem 11.12, then, it is enough to find an ST-sequence $V_0, V_1, \ldots, V_{n-1}$ such that $V_0 = \ell_p^1$, such that for $i \geqslant 1$ each V_i is either $S_{1+1/i}$ or $T_{1+1/i}$ and such that the corresponding ST-function λ satisfies $x^{1/p} \leqslant \lambda(x) < (3/2)x^{1/p}$.

Clearly we only need to check this inequality when x is an integer, so let us suppose that we have chosen $V_1, \ldots, V_{k-1}$. Then if $\lambda(k) \geqslant (k+1)^{1/p}$, set $V_k = T_{1+1/k}$. Otherwise, let it be $S_{1+1/k}$. Then for each k, $\lambda(k)$ is either $\lambda(k-1)$ or it is less than $k^{1/p} \cdot k/(k-1)$. In the first case we are done, by induction. In the second, we are done, unless $k = 2$. But in this case, $\lambda(k) = 2 < (3/2) \cdot 2^{1/p}$. □

The next result is a very easy corollary of Lemma 11.11 and Theorem 11.12. We state it here for convenience. Given a non-generalized ST-space $X \in \mathcal{F}_n$, let the *associated ST-function* be the function from $[n]$ to $\mathbb{R}_+$ given by $i \mapsto \left\| \sum_{j=1}^{i} \mathrm{e}_j \right\|_X$.

Theorem 11·14. *(i) Let $X, Y \in \mathcal{F}_n$ be non-generalized ST-spaces, and let the associated ST-functions of X and Y be λ and μ respectively. Then, if c_1 and c_2 are constants such that $c_1\lambda(i) \leqslant \mu(i) \leqslant c_2\lambda(i)$ for every $i \in [n]$, we have also that $c_1 \|x\|_X \leqslant \|x\|_Y \leqslant c_2 \|x\|_X$ for every $x \in \mathbb{R}^n$.*

(ii) Let $X \in \mathcal{F}_n$ be a non-generalized ST-space with associated ST-function λ, and let c_1 and c_2 be constants such that $c_1\lambda(i) \leqslant i^{1/p} \leqslant c_2\lambda(i)$ for every $i \in [n]$. Then $c_1 \|x\|_X \leqslant \|x\|_p \leqslant c_2 \|x\|_X$ for every $x \in \mathbb{R}^n$. □

Note that Theorem 11.14 implies that if $k \in \mathbb{N}$ is even and $X \in \mathcal{F}_{2^k}$ is the space $T^{2^{k-1}}(S^{2^{k-2}}(\ldots(S^4(T^2(S(\mathbb{R}))))\ldots))$, then $d(X, \ell_2^{2^k}) = \sqrt{2}$. Indeed, if λ is the ST-function associated with this space, then $\lambda(2^l) = 2^{l/2}$ if l is even, and $2^{(l+1)/2}$ if l is odd, and in between λ is linear. It follows that $\sqrt{x} \leqslant \lambda(x) \leqslant \sqrt{2x}$ for every x, which, by Theorem 11.14, is enough to prove the above estimate. This can also be proved directly from Lemma 11.8, using a duality argument similar to that used in the proof of Theorem 11.1. In general, to approximate ℓ_p^n to within an absolute constant by an ST-space, one only needs about $\log n$ changes of direction of the corresponding ST-function λ. This shows that there is considerable elbow room in Theorem 11.1.

§11.4 An Alternative Definition of the Norm

In Section 11.2 we obtained an algorithm for calculating the norm of a vector in any given ST-space, and we commented in the next section that the same method could be used, with obvious adaptations, to calculate the norm of a vector in a generalized ST-space. This method has an analogue which applies to any function whose logarithmic gradient is always between 0 and 1 and not simply ST-functions. Let us say that a function $\lambda : [0, t] \longrightarrow \mathbb{R}$ is a *growth function* if it is a uniform limit of a sequence of ST-functions such that the first term of the reduced ST-sequence corresponding to each function is the same. It is not hard to show that λ is a growth function if and only if there exist $0 < s \leqslant t$ and constants $c > 0$ and $1 \leqslant p \leqslant \infty$ such that $\lambda(x) = cx^{1/p}$ for every $x \in [0, s]$, and $\lambda(x) \leqslant \lambda(y) \leqslant (y/x)\lambda(x)$ whenever $x \in [s, t]$. It is also an easy consequence of Theorem 11.11 that, given a growth function λ, the following element Y of $\mathcal{G}_t$ is well defined. Pick any sequence $\lambda_1, \lambda_2, \ldots$ of ST-functions tending uniformly to λ with $\lambda_i(1) = 1$ for every i, and let

$$\|x\|_Y = \lim_{n \to \infty} \|x\|_{Y_n}$$

where Y_n is the ST-space corresponding to λ_n. We shall call the space Y the *continuous ST-space corresponding to* λ. Any space obtained this way we shall call a *continuous ST*-space. We shall often drop the word "continuous" if no ambiguity arises.

So far, we have only defined our norms on step functions. It is easier for the rest of the chapter to work with a larger class of functions. Let $G[0, t]$ denote the space of bounded piecewise continuous functions on $[0, t]$. Given any norm on $F[0, t]$ we can by an obvious limiting argument extend it to a norm on $G[0, t]$. We shall continue to use our previous notation for the rest of the section to apply to these extended spaces.

In this section, we shall show how to calculate the norm of functions $f \in G[0, t]$ in continuous ST-spaces. We shall not give a detailed proof. Such a proof would be a natural extension of the methods of Section 11.2. To simplify matters, we shall assume that λ is differentiable and f is bounded, positive and differentiable with derivative less than zero on the interval $[0, t]$. Let us call a function f satisfying these conditions *standard*. The main idea, as in Section 11.2, is to replace f by a function g_y supported on $[0, y]$ for some $y < t$ so that for some $0 \leqslant x \leqslant y$ we have

$$g_y(s) = \begin{cases} f(s) & 0 \leqslant s \leqslant x \\ f(x) & x \leqslant s \leqslant y \\ 0 & s > y \end{cases}$$

and, most importantly, $\|g_y\|_\lambda = \|f\|_\lambda$.

One obtains the following differential equation relating y and x:

$$y\frac{\lambda'(y)}{\lambda(y)}\frac{dy}{dx} = -\frac{f'(x)}{f(x)}(y-x) \tag{1}$$

with the condition that $y(t) = t$. When $x = 0$, we have replaced f by $f(0)\chi_{[0,y(1)]}$, so we obtain that the norm of f is $f(0)\lambda(y(1))$. One can obtain the norm of a function which is not standard by approximating its decreasing rearrangement by standard functions.

Note that if $\lambda(y) = y^{1/p}$, then the solution of this equation is

$$y = f(x)^{-p}\left(C - \int_0^x f(s)^p\, ds\right) + x$$

for some constant C. Since $y(t) = t$, we obtain that $C = \int_0^t f(s)^p\, ds$. and hence, that $y(0) = f(0)^{-p}\int_0^t f(s)^p\, ds$. It follows that $f(0)(y(0))^{1/p} = \|f\|_p$.

One can adopt this as a definition of ST-spaces. Once one has proved rigorously that it agrees with the original definition, one can prove Theorem 11.12 using Lemma 11.11 but not Lemma 11.8. It seems that one can also obtain a rather strange proof of Hölder's inequality and Minkowski's inequality without assuming either inequality in the process, but I have not checked this, as the proofs would be much longer than the usual proofs. In the next section, we shall consider two natural questions about ST-spaces that are most easily dealt with by using the differential equation discussed in this section.

§11.5 Two Results about ST-Spaces

We shall be concerned with two natural questions in this section. First, what is the relationship, if anything, between ST-spaces and Orlicz spaces? Second, what is the result of interpolating between two ST-spaces? First, let us note that the restriction on growth functions that they should begin as $cx^{1/p}$ is not necessary in order to use the differential equation in the last section. For this section, a growth function will stand for any positive differentiable function whose logarithmic right derivative lies between 0 and 1. The *corresponding (continuous) ST-space* will be simply the space for which the norm is calculated using the differential equation with the growth function λ.

Our first question is inspired by the fact that both ST-spaces and Orlicz spaces are 1-symmetric spaces or rearrangement-invariant function spaces which generalize ℓ_p-spaces by replacing the function t^p by a more general function. Suppose X is an Orlicz function space restricted to $\mathcal{G}_t$ with the norm given by

$$\|f\|_X = \inf\left\{\mu > 0 : \int_0^t \phi(|f(t)|/\mu)\, dt \leqslant 1\right\}$$

where $f \in G[0,t]$ and ϕ is an Orlicz function. For this space we have $\|\chi_{[0,t]}\|_X = (\phi^{-1}(t^{-1}))^{-1}$. Let λ be the function defined on the interval $[0,t]$ by $\lambda(x) = (\phi^{-1}(x^{-1}))^{-1}$ and suppose that λ is differentiable with non-zero derivative. It is natural to ask whether the continuous ST-space corresponding to λ is the same as X. We shall show that the two spaces are isometric if and only if, for some constants $1 \leqslant p < \infty$ and some $C > 0$, $\phi(t) = t^p$ and therefore $\lambda(t) = Ct^{1/p}$. In other words, the set of continuous ST-spaces and the set of Orlicz spaces intersect only in the ℓ_p-spaces.

Our proof of this is slightly indirect. Suppose ϕ is an Orlicz function, X is the corresponding Orlicz space and λ is a growth function with the property given above. Suppose moreover that X is the ST-space corresponding to λ. and let f

be any standard function in $G[0,t]$. For $y < t$ sufficiently close to t let us replace f by a function g_y with the properties we had in the last section. That is, for some $0 \leqslant x \leqslant y$,

$$g_y(s) = \begin{cases} f(s) & 0 \leqslant s \leqslant x \\ f(x) & x \leqslant s \leqslant y \\ 0 & s > y \end{cases}$$

and $\|g_y\|_\lambda = \|f\|_\lambda$. It is easy to see that, for each y, g_y is unique. Let us suppose that $\|f\|_X = 1$. Using the definition of the norm in X one readily obtains that

$$\int_0^x \phi(f(s))\, ds + (y-x)\phi(f(x)) = \int_0^t \phi(f(s))\, ds \tag{2}$$

If we differentiate this equation with respect to x and rearrange, we obtain that

$$\frac{\phi(f(x))}{\phi'(f(x))f(x)} \frac{dy}{dx} = -\frac{f'(x)}{f(x)}(y-x).$$

On the other hand, if we substitute $\lambda(y) = (\phi^{-1}(y^{-1}))^{-1}$ into (1), we obtain the equation

$$\frac{1}{y\phi^{-1}(y^{-1})\phi'(\phi^{-1}(y^{-1}))} \frac{dy}{dx} = -\frac{f'(x)}{f(x)}(y-x).$$

It follows from the uniqueness of g_y that $\phi(f(x))/\phi'(f(x))f(x)$ does not depend on $f(x)$. In other words, the function $t \mapsto \phi(t)/t\phi'(t)$ is a constant function. Solving this equation gives $\phi(t) = Ct^p$ for some constants C and p. We have proved the next theorem.

Theorem 11·15. *Let $\tilde{X}$ be the Orlicz function space determined by a differentiable Orlicz function, let $X \in \mathcal{G}_t$ be the restriction of $\tilde{X}$ to $G[0,t]$ and let Y be the continuous ST-space corresponding to the growth function $\lambda(x) = (\phi^{-1}(t^{-1}))^{-1}$. Then X is isometric to Y under the identity map if and only if, for some constants $C > 0$ and $1 \leqslant p \leqslant \infty$, $\|f\|_X = \|f\|_Y = C\,\|f\|_p$.*

I suspect it would be possible to be more careful about the proof of Theorem 11.15 and to obtain an isomorphic version of the result, and also a result that was valid without the assumption that the Orlicz function was differentiable. It would also be interesting to know how much of the structure theory for Orlicz spaces is also true for ST-spaces. For example I think the following is probably true. Given any infinite-dimensional ST-space, then for any $\epsilon > 0$ there is a block basis of the natural basis which is $(1+\epsilon)$-equivalent to the unit vector basis of ℓ_p for some p. Moreover, if the growth function of the space is λ, then this ought to hold for any p in the closed interval from $\liminf \log(\lambda(n))/\log n$ to $\limsup \log(\lambda(n))/\log n$. A similar result for Orlicz spaces was proved by Lindenstrauss and Tzafriri (cf. [34]). The following observation at least indicates that the structure of subspaces of ST-spaces is interesting, and that there is a relationship between ST-spaces and Orlicz spaces.

Proposition 11·16. *Let $n = m^2$, and let $X = T^{n-m}S^{m-1}(\mathbf{R})$. Then, for any $1 \leqslant p \leqslant \infty$, the unit vector basis of ℓ_p^m is 2-isomorphic to a block basis of the standard basis of X. This block basis is generated by disjointly supported copies of a single vector.*

Proof. The norm of a vector $\mathbf{a} = (a_i)_1^n \in X$ is given by

$$\|\mathbf{a}\|_X = \max\left\{\sum_{i\in A} |a_i| : A \subset [n], |A| = m\right\}.$$

If $p > 1$, let q be the conjugate index of p, let $\mathbf{u}$ be the vector

$$\mathbf{u} = m^{-1/q}(1^{1/q}, 2^{1/q} - 1^{1/q}, \ldots, m^{1/q} - (m-1)^{1/q}, 0, \ldots, 0)$$

and let $\mathbf{u}_1, \ldots, \mathbf{u}_m$ be disjointly supported copies of $\mathbf{u}$. Note that the coordinates of $\mathbf{u}$ have been given in decreasing order. Let $a_1, \ldots, a_m$ be scalars with $a_i \geqslant 0$ for each i and $\sum_1^m a_i^p = 1$ (or $\max_{1\leqslant i\leqslant m} a_i = 1$ if $p = \infty$). We shall estimate the norm of $\sum_1^m a_i\mathbf{u}_i$. Suppose we choose a set $A \in [n]^{(m)}$ and set $k_i = |A \cap \operatorname{supp}(\mathbf{u}_i)|$. Suppose also that $A \cap \operatorname{supp}(\mathbf{u}_i)$ contains the k_i largest coordinates of $\mathbf{u}_i$. We obtain that $\|\sum_{i=1}^m a_i \mathbf{u}_i\| \geqslant m^{-1/q} \sum_{i=1}^m a_i k_i^{1/q}$ and in general it is clear that

$$\left\| \sum_{i=1}^m a_i \mathbf{u}_i \right\| = \sup \left\{ m^{-1/q} \sum_{i=1}^m a_i k_i^{1/q} : k_i \in \mathbb{N}, \sum_{i=1}^m k_i \leqslant m \right\}.$$

By Hölder's inequality we therefore have

$$\left\| \sum_{i=1}^m a_i \mathbf{u}_i \right\| \leqslant m^{-1/q} \left(\sum_{i=1}^m a_i^p \right)^{1/p} \left(\sum_{i=1}^m k_i \right)^{1/q} \leqslant 1.$$

On the other hand, if we pick

$$k_i = \begin{cases} \lceil a_i^p m \rceil & \sum_{j=1}^i \lceil a_j^p m \rceil \leqslant m \\ 0 & \text{otherwise} \end{cases}$$

we have $k_i \geqslant a_i^p m$ if $\sum_{j=1}^i a_j^p \leqslant 1/2$, and so $m^{-1/q} \sum_{i=1}^m a_i k_i^{1/q} \geqslant 1/2$ and $\sum_{i=1}^m k_i \leqslant m$. The case when $p = 1$ follows by continuity, or else by inspection. □

This method actually shows that the dual of any Orlicz sequence space of m dimensions embeds into X in a similar way. Given an Orlicz function ϕ, set $\psi = m\phi$ and $f = \psi^{-1}$. If $\mathbf{u}_1, \ldots, \mathbf{u}_m$ are disjointly supported copies of the vector

$$\mathbf{u} = (f(1), f(2) - f(1), \ldots, f(m) - f(m-1), 0, \ldots, 0),$$

then they generate a subspace uniformly isomorphic to the dual of ℓ_ϕ^m.

The next result shows that ST-spaces interpolate in the way one would expect. Since we prove an isometric result, we must use complex interpolation and therefore complex scalars. One can either define the norm of any vector to be the norm of its modulus, or follow the original approach making obvious modifications. These give the same result. For standard facts about the complex interpolation method, see e.g. [7]. We shall also make use of the obvious fact that if λ is a growth function, then the dual of the ST-space corresponding to λ is the ST-space corresponding to λ', where $\lambda' = n^{1/2}/\lambda$.

Theorem 11·17. *Let λ and μ be growth functions and let X_λ and X_μ be the ST-spaces corresponding to λ and μ respectively. Given $0 < \theta < 1$, let $\nu = \lambda^\theta \mu^{1-\theta}$ and let X_ν be ST-space corresponding to ν. Then*

$$(X_\lambda, X_\mu)_{[\theta]} = X_\nu.$$

Proof. The proof of this is very similar to the standard proof that L_p-spaces interpolate in the way one would expect. Let f be a standard function with $\|f\|_{X_\nu} = 1$ and let $y = y(x)$ be defined by the differential equation

$$y\frac{\nu'(y)}{\nu(y)}\frac{dy}{dx} = -\frac{f'(x)}{f(x)}(y - x)$$

and the initial condition $y(t) = t$. Let $D \subset \mathbb{C}$ be the set $\{z \in \mathbb{C} : 0 \leqslant \mathrm{Re}(z) \leqslant 1\}$. Fixing y, we may now define, for every $z \in D$, a function $g_z \in G[0,t]$ by the differential equation

$$y\left((1-z)\frac{\lambda'(y)}{\lambda(y)} + z\frac{\mu'(y)}{\mu(y)}\right)\frac{dy}{dx} = -\frac{g_z'(x)}{g_z(x)}(y - x)$$

and the initial condition

$$g_z(0) = \lambda(y(0))^{-(1-z)}\mu(y(0))^{-z}.$$

It is easy to check that

$$g_z(x) = g_0(x)^{1-z}g_1(x)^z. \tag{3}$$

Let us now set $\tilde{f}(z,x) = \exp(\epsilon z^2 - \epsilon\theta^2)g_z(x)$. Thus $\tilde{f} : D \times [0,t] \to \mathbb{C}$. We have the properties of $\tilde{f}$ necessary to estimate $\|f\|_{(X_\lambda,X_\mu)_{[\theta]}}$. First, for each x it is clear that $\tilde{f}(z,x)$ is analytic in z on the interior of D. Second, $\left\|\tilde{f}(it,\cdot)\right\|_{X_\lambda}$ and $\left\|\tilde{f}(1+it,\cdot)\right\|_{X_\mu}$ both tend to zero as $|t|$ tends to infinity. Moreover, we have $\left\|\tilde{f}(it,\cdot)\right\|_{X_\lambda} \leqslant 1$ and $\left\|\tilde{f}(1+it,\cdot)\right\|_{X_\mu} \leqslant \exp(\epsilon)$ for every $t \in \mathbb{R}$. It follows that $\|f\|_{(X_\lambda,X_\mu)_{[\theta]}} \leqslant \exp(\epsilon)$. Since $\epsilon > 0$ was arbitrary, we have shown that $\|f\|_{(X_\lambda,X_\mu)_{[\theta]}} \leqslant \|f\|_{X_\nu}$ for any function f.

Conversely, suppose that $\|f\|_{(X_\lambda, X_\mu)_{[\theta]}} = 1$. This tells us that for each $\epsilon > 0$ there exists a function $\tilde{f}$ with the above properties. We also know that

$$\|f\|_{X_\nu} = \sup\left\{|\langle f, h\rangle| : h \in G[0,t] \text{ is standard}, \|h\|_{X'_\nu} = 1\right\}$$

where $\nu' = n^{1/2}/\nu$.

Given a standard function $h \in G[0,t]$, let $\tilde{h} : D \times [0,t] \to \mathbb{C}$ be given by the method used to construct $\tilde{f}$ from f, replacing all the spaces in that construction by their duals. Let us then set

$$F(z) = \int_0^t \tilde{f}(z,x)\tilde{h}(z,x)\,dx$$

for every $z \in D$. Then F is analytic on the interior of D, continuous on D, and $F(it) \leqslant \exp(\epsilon)$ and $F(1+it) \leqslant \exp(2\epsilon)$ for every $t \in \mathbb{R}$. By the Hadamard 3-line theorem (cf. [7]), we obtain that

$$|\langle f, h\rangle| \leqslant |F(\theta)| \leqslant \exp(2\epsilon).$$

Since $\epsilon > 0$ was arbitrary, we have $\|f\|_{X_\nu} \leqslant 1$. □

There is probably a good deal more that one can say about ST-spaces. I hope to be able to use them to obtain upper bounds for some of the local versions of Krivine's theorem. For example, I think it may be possible to find an n-dimensional ST-space whose standard basis is C-equivalent to the standard basis of ℓ_p^n, but has no block basis of size $k = n^\alpha$ equivalent to the standard basis of ℓ_p^k, where α is around $(\epsilon/C)^p$. This would show that the lower bound due to Amir and Milman [3,4] was sharp, at least for block bases. The best known upper bound, noted by Amir and Milman [4], is obtained by taking the standard basis of ℓ_q^n, for some q suitably close to p. When $p \neq 2$, one obtains an upper bound of n^β for β about $\log(1+\epsilon)/\log C$, so the correct dependence of the exponent on p is not at all clear.

With reference to the very beginning of this chapter, it is possible to define non-generalized ST-spaces in a purely geometrical way. One begins with the standard basis of $\mathbb{R}^n$ and the antipodal points, and builds up first the edges of a polytope, then the 2-dimensional faces and so on. This method can then be used with other sets of points on the surface of the sphere. However, we have not been able to show that, under suitable conditions on the original set of points, one obtains a body that is close to the sphere, so we will not go into it. For the moment, this is where we shall finish our discussion.

BIBLIOGRAPHY

[1] N. Alon, *Probabilistic proofs of existence of rare events*, GAFA 87-88, Springer Lecture Notes **1376** (1989), 186-201.

[2] N. Alon and V. D. Milman, *Embedding of ℓ_∞^k in finite dimensional Banach spaces*, Israel J. Math. **45**, 265-280.

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

[4] D. Amir and V. D. Milman, *A quantitative finite dimensional Krivine theorem*, Israel J. Math. **50** (1985), 1-12.

[5] K. Azuma, *Weighted sums of certain dependent random variables*, Tôhoku Math. J. **19** (1967), 357-367.

[6] J. E. Baumgartner, *A short proof of Hindman's theorem*, J. Comb. Th. (A) **17** (1974), 384-386.

[7] J. Bergh and J. Löfström, *Interpolation Spaces, an Introduction*, Springer-Verlag (1976), x + 207 pp.

[8] C. Bessaga and A. Pelczynski, *On bases and unconditional convergence of series in Banach spaces*, Studia Math. **17** (1958), 165-174.

[9] E. Bishop and R. R. Phelps, *A proof that every Banach space is subreflexive*, Bull. Amer. Math. Soc. **67** (1961), 97-98.

[10] B. Bollobás, *Random Graphs*, Academic Press (1985), xvi + 447 pp.

[11] J. Bourgain, *On dentability and the Bishop-Phelps property*, Israel J. Math. **28** (1977), 265-271.

[12] J. Bourgain, *Subspaces of L_N^∞, arithmetical diameter of Sidon sets*, Probability in Banach Spaces V, Proc. Medford 84, Springer Lecture Notes **1153** (1985), 96-127.

[13] J. Bourgain and J. Lindenstrauss, *Almost Euclidean sections in spaces with a*

symmetric basis, GAFA 87-88, Springer Lecture Notes **1376** (1989), 278-288.

[14] T. J. Carlson, *Some unifying principles in Ramsey theory*, Discrete Math. **68** (1988), 117-169.

[14a] T. J. Carlson and S. G. Simpson, *A dual form of Ramsey's theorem*, Adv. in Math., **53** (1984), 265-290.

[15] A. Dvoretzky, *Some results on convex bodies and Banach spaces*, Proc. Symp. on Linear Spaces, Jerusalem (1961), 123-160.

[16] J. Elton, *Sign-embeddings of ℓ_1^n*, Trans. Amer. Math. Soc. **279** (1983), 113-124.

[17] P. Erdős and L. Lovász, *Problems and results on 3-chromatic hypergraphs and some related questions*, Infinite and Finite Sets, North Holland, Amsterdam (1975), 609-628.

[18] T. Figiel, J. Lindenstrauss and V. D. Milman, *The dimension of almost spherical sections of convex bodies*, Acta Math. **139** (1977), 53-94.

[19] C. Finet, *Renorming Banach spaces with many projections and smoothness and properties*, to appear.

[20] C. Finet and W. Schachermayer, *Equivalent norms on separable Asplund spaces*, Studia Math. **93.2** (1988), to appear.

[21] H. Furstenberg and Y. Katznelson, *Idempotents in compact semigroups and Ramsey theory*, preprint.

[22] W. T. Gowers, *Symmetric block bases in finite-dimensional normed spaces*, Israel J. Math., to appear.

[23] W. T. Gowers, *Symmetric block bases of sequences with large average growth*, Israel J. Math., to appear.

[24] R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey Theory*, Wiley, New York (1980), ix + 174 pp.

[25] N. Hindman, *Finite sums from sequences within cells of a partition of* $\mathbb{N}$,

J. Comb. Th. (A) **17** (1974), 1-11.

[26] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), 13-30.

[27] R. C. James, *Uniformly non-square Banach spaces*, Ann. of Math. **80** (1964), 542-550.

[28] J. Johnson and J. Wolfe, *Norm attaining operators*, Studia Math. **65** (1979), 7-19.

[29] H. König and L. Tzafriri, *Some estimates for type and cotype constants*, Math. Ann. **256** (1981), 85-94.

[30] J. L. Krivine, *Sous-espaces de dimension finie des espaces de Banach réticulés*, Ann. of Math. **104** (1976), 1-29.

[31] S. Kwapien, *Isomorphic characterizations of inner product spaces by orthogonal series with vector coefficients*, Studia Math. **44** (1972), 583-595.

[32] J. Lindenstrauss, *On operators which attain their norm*, Israel J. Math. **1** (1963), 139-148.

[33] J. Lindenstrauss and A. Pelczynski, *Absolutely summing operators in L_p-spaces and their applications*, Studia Math. **29** (1968), 275-326.

[34] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces I, Sequence Spaces*, Springer-Verlag, Berlin (1977), xiii + 190 pp.

[35] B. Maurey, *Construction de suites symétriques*, C. R. A. S., Paris, **288** (1979), 679-681.

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

[37] C. McDiarmid, *On the method of bounded differences*, Surveys in Combinatorics 1989, LMS Lecture Note Series **141**, CUP (1989), 149-188.

[38] V. D. Milman, *A new proof of the theorem of A. Dvoretzky on sections of*

convex bodies, Func. Anal. Appl., **5** (1971), 28-37 (translated from Russian).

[39] V. D. Milman and G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Springer Lecture Notes **1200** (1986), viii + 156 pp.

[40] V. D. Milman and H. Wolfson, *Minkowski spaces with extremal distance from Euclidean space*, Israel J. Math., **29** (1978), 113-130.

[41] J. R. Partington, *Norm attaining operators*, Israel J. Math. **43** (1982), 273-276.

[42] G. Pisier, *Factorization of Linear Operators and the Geometry of Banach Spaces*, CBMS Regional Conf. Series in Math. (1986).

[43] G. Pisier, *Sur les espaces de Banach de dimension finie a distance extremal d'un espace euclidien*, Séminaire d'Analyse Fonct., Exp. **16**, Ecole Polytechnique, Palaiseau.

[44] R. Rado, *An inequality*, J. Lond. Math. Soc. **27** (1952), 1-6.

[45] H. Rosenthal, *On a theorem of Krivine concerning block finite representability of ℓ_p in general Banach spaces*, J. Func. Anal. **28** (1978), 197-225.

[46] G. Schechtman, *Lévy type inequality for a class of metric spaces*, Martingale Theory in Harmonic Analysis and Banach Spaces, Springer-Verlag 1981, 211-215.

[47] N. Tomczak-Jaegermann, *Banach-Mazur Distances and Finite Dimensional Operator Ideals*, Pitman (1989), xii + 395 pp.

[48] N. Tomczak-Jaegermann, *Computing 2-summing norms with few vectors*, Ark. Math. **17** (1979), 273-277.