

## ON TYPE OF METRIC SPACES

BY

J. BOURGAIN, V. MILMAN AND H. WOLFSON<sup>1</sup>

**ABSTRACT.** Families of finite metric spaces are investigated. A notion of metric type is introduced and it is shown that for Banach spaces it is consistent with the standard notion of type. A theorem parallel to the Maurey-Pisier Theorem in Local Theory is proved. Embeddings of  $l_p$ -cubes into the  $l_1$ -cube (Hamming cube) are discussed.

### 1. Introduction.

*Terminology and notations.*

1.1. DEFINITION. Two finite metric spaces  $(X, \rho)$ ,  $(Y, d)$  are called  $c$ -isomorphic if there is a one-to-one map  $\psi: X \rightarrow Y$  such that  $\|\psi\|_{\text{Lip}} \|\psi^{-1}\|_{\text{Lip}} \leq c$ . (We recall that  $\|\psi\|_{\text{Lip}} = \sup_{x \neq y} (d(\psi(x), \psi(y)) / \rho(x, y))$ ).

In  $[\mathbf{G}]$   $\|\psi\| \|\psi^{-1}\|$  is called the distortion of  $\psi$ .

Analogous to the Banach-Mazur distance between normed spaces, we define the Lipschitz distance between finite metric spaces as

$$d(X, Y) = \inf_{\psi: X \rightarrow Y} \|\psi\|_{\text{Lip}} \|\psi^{-1}\|_{\text{Lip}},$$

where the infimum is taken over all one-to-one and onto maps  $\psi: X \rightarrow Y$ .

1.2. We recall that the Banach-Mazur distance between two  $n$ -dimensional normed spaces  $X, Y$  is defined as  $d(X, Y) = \inf\{\|T\| \|T^{-1}\| : T: X \rightarrow Y \text{ is an isomorphism}\}$ . It is known that for the finite-dimensional real Banach spaces the Lipschitz distance coincides with the Banach-Mazur distance. (In the nontrivial direction use the fact that for any Lipschitz map  $\psi: X \rightarrow Y$  ( $\dim X = \dim Y < \infty$ ) there is a point  $x \in X$  such that  $\psi$  and  $\psi^{-1}$  have derivatives at  $x$  and  $y = \psi(x)$  respectively. Hence, there is a linear map  $T: X \rightarrow Y$  satisfying  $\|T\| \|T\|^{-1} \leq \|\psi\|_{\text{Lip}} \|\psi^{-1}\|_{\text{Lip}}$ ).

1.3. Let  $C_2^n = \{(\varepsilon_1, \dots, \varepsilon_n) | \varepsilon_i \in \{0, 1\}\} = \{0, 1\}^n$ . (Sometimes it is more convenient to use  $C_2^n = \{-1, 1\}^n$ . It will be clear from the context what representation is used.)

For every pair  $\varepsilon = (\varepsilon_j)_{j=1}^n$ ,  $\varepsilon' = (\varepsilon'_j)_{j=1}^n$  in  $C_2^n$  the Hamming metric is defined as  $h(\varepsilon, \varepsilon') = \#\{i | \varepsilon_i \neq \varepsilon'_i\}$ . In the case of  $C_2^n = \{0, 1\}^n$  this metric coincides with the standard  $l_1$  metric  $(\sum_{i=1}^n |\varepsilon_i - \varepsilon'_i|)$ .

DEFINITION. Let  $1 \leq p \leq \infty$ . The metric space  $(C_2^n, \rho_p)$ , where  $C_2^n = \{0, 1\}^n$  and  $\rho_p(\varepsilon, \varepsilon') = [\sum_{i=1}^n |\varepsilon_i - \varepsilon'_i|^p]^{1/p} = [\sum_{i=1}^n |\varepsilon_i - \varepsilon'_i|]^{1/p}$  for any pair  $\varepsilon, \varepsilon' \in C_2^n$  is called the  $l_p$   $n$ -cube (or  $l_p^n$ -cube).

---

Received by the editors May 25, 1984.

1980 *Mathematics Subject Classification*. Primary 46B20, 47H99, 51E99, 54E40.

<sup>1</sup>This paper includes part of the Ph.D. thesis of the third author.

REMARK. The  $l_1$   $n$ -cube will be often called the Hamming cube. We recall that for any  $1 \leq p \leq \infty$  the  $l_p^n$  space is the real (or complex)  $n$ -dimensional space equipped with the norm  $\|(a_i)_{i=1}^n\|_p = (\sum_{i=1}^n |a_i|^p)^{1/p}$ .

1.4. We say that an infinite-dimensional Banach space  $X$  contains  $l_p^n$ 's uniformly if for every  $\varepsilon > 0$  there exists a sequence of subspaces of  $X$ ,  $\{X_n; \dim X_n = n\}$ , such that for every  $n$ ,  $d(X_n, l_p^n) \leq 1 + \varepsilon$ .

An infinite metric space  $(X, \rho)$  is said to contain  $l_p$   $n$ -cubes uniformly if for any  $\varepsilon > 0$  there is a sequence of subsets of  $X$ ,  $\{X_n \subset X; \text{card } X_n = 2^n\}$ , such that for every  $n$ ,  $d\{(X_n, \rho), (C_2^n, \rho_p)\} \leq 1 + \varepsilon$ .

1.5. For any metric space  $(X, \rho)$  with cardinality  $2^n$  ( $|X| = 2^n$ ) consider a one-to-one map  $\psi: C_2^n \rightarrow X$ , which gives some way of ordering the elements of  $X$  by  $n$ -dimensional binary vectors.

Consider a point  $\varepsilon = (\varepsilon_j)_{j=1}^n$  in  $C_2^n$ . Take the opposite point on the cube  $\varepsilon^c = (1 - \varepsilon_j)_{j=1}^n$ . We define a *diagonal* in  $X$  as the unordered pair  $(\psi(\varepsilon); \psi(\varepsilon^c))$  and its length is

$$\text{diag}_\psi(\varepsilon) = \rho(\psi(\varepsilon), \psi(\varepsilon^c)).$$

(For the sake of convenience, we often use the same notation for a diagonal and its length. It is clear from the context what definition is meant.)

The set of all diagonals is denoted by  $D$ . Clearly,  $|D| = 2^{n-1}$ . (Sometimes, it is convenient to count every diagonal twice in order to get a one-to-one correspondence between the diagonals and the vertices.)

An *edge* in  $X$  is defined for every unordered pair of points  $(\varepsilon; \varepsilon')$  in  $C_2^n$  differing in one binary coordinate only, and its length (which we usually will call just "edge") is

$$\text{edge}_\psi(\varepsilon, \varepsilon') = \rho(\psi(\varepsilon), \psi(\varepsilon')).$$

The set of all edges is denoted by  $E$ .  $|E| = n2^{n-1}$ .

Two edges  $(\psi(\varepsilon_1); \psi(\varepsilon_2))$  and  $(\psi(\varepsilon'_1); \psi(\varepsilon'_2))$  are called *connected* if they share one point. (For example  $\psi(\varepsilon_2) = \psi(\varepsilon'_1)$ .) In the sequel, when it is obvious, we do not mention the specific ordering (map  $\psi$ ), which is chosen.

A *path* belonging to  $\text{diag}_\psi(\varepsilon)$  is a set of  $n$  connected edges, starting at  $\psi(\varepsilon)$  and terminating at  $\psi(\varepsilon^c)$ .

There are, of course, different paths belonging to the same diagonal. We may consider a path as the set of edges obtained by  $n$  consecutive changes of the coordinates of  $\varepsilon$  to obtain  $\varepsilon^c$ . The  $n$ -permutation  $\pi$ , by which these coordinates are changed, fully describes the path. Hence, we denote a path by  $E_\varepsilon(\pi)$ .

1.6. An infinite-dimensional Banach space  $X$  is called of type  $p$  if there is a constant  $C$  such that for any  $n$  and any sequence  $(x_i)_{i=1}^n$  in  $X$

$$(1.6.1) \quad \left[ \frac{1}{2^n} \sum_{\varepsilon} \left\| \sum_{i=1}^n \varepsilon_i x_i \right\|^p \right]^{1/p} \leq C \left( \sum_{i=1}^n \|x_i\|^p \right)^{1/p}$$

where the summing is over all  $\varepsilon \in \{-1, 1\}^n$ . (See [M-P].)

The definition makes sense only for  $1 \leq p \leq 2$  and it is clear that if  $X$  is of type  $p$  then there is a constant  $C'$  such that

$$(1.6.2) \quad \left[ \frac{1}{2^n} \sum_{\varepsilon} \left\| \sum_{i=1}^n \varepsilon_i x_i \right\|^2 \right]^{1/2} \leq C' n^{1/p-1/2} \left( \sum_{i=1}^n \|x_i\|^2 \right)^{1/2}.$$

Conversely, it is well known that (1.6.2) implies  $X$  is of type  $r$  whenever  $r < p$ .

Every Banach space  $X$  is of type 1 and the supremum of all types of  $X$  is denoted by  $p_X$ .

The Maurey-Pisier Theorem [M-P] says that  $X$  contains  $l_{p_X}^n$ 's uniformly.

The metric analog of type  $p$  (for any  $p \geq 1$ ) is defined as follows:

DEFINITION. Let  $p \geq 1$ . An (infinite) metric space  $(X, \rho)$  has *metric type  $p$*  if there is a constant  $\alpha$ , such that for every  $k$  and any  $k$ -cube (defined by any map  $\psi: C_2^k \rightarrow X$ ) the following inequality holds:

$$(1.6.3) \quad \left( \sum_D \text{diag}^2 \right)^{1/2} \leq \alpha k^{1/p-1/2} \left( \sum_E \text{edge}^2 \right)^{1/2},$$

where the summing is over all the diagonals and all the edges of the  $k$ -cube.

The smallest constant  $\alpha$  satisfying (1.6.3) is called the metric  $p$ -type constant of  $(X, \rho)$  and is denoted  $\alpha_p(X)$ . The smallest constant for a given  $k$  is denoted  $\alpha_p(k; X)$ . (When it causes no confusion we call the metric type just type.)

As usual, the metric type constant can be defined in the same way for a family of finite metric spaces  $(X_N, \rho_N)$ , where  $|X_N| = N$  ( $N \uparrow \infty$ ), as the best constant obtained simultaneously for all the spaces.

Another definition of a metric type is given in §3 (see 3.14).

*Summary of results.*

1.7. Recently, there has been a growing interest in understanding the similarity between asymptotic properties of finite-dimensional normed spaces and finite metric spaces. In a few known examples [Mar-P, G, J-L] there is a close analogy between the results obtained for metric spaces and the previously-known results from the Local Theory of Banach spaces, where the role of the dimension of a Banach space is played by the logarithm of the cardinality of the finite metric space. Below, we give more examples of these kind of results.

1.8. In §2, we show that if a metric space  $(X, \rho)$  has metric type no bigger than 1, then it contains  $l_1$   $n$ -cubes (Hamming cubes) uniformly. This is an analog of Pisier's result (see [P1]) for Banach spaces.

1.9. In §3 a result similar to that of §2 is shown; however the method is entirely different (the so-called "measure concentration" phenomenon is used). The estimates on the cardinality of the Hamming subcubes are improved; however, some restriction on the original space is imposed. Another definition of metric type influenced by Enflo's paper [E] (the so-called  $E$ -type) is introduced and discussed.

1.10. In §4 metric spaces of type no bigger than  $p$  ( $1 < p < 2$ ) are discussed. The connection of this property with the embeddability of  $l_p$   $n$ -cubes is investigated. In a sense, the results of §§2–4 may be viewed as a metric analog of the Maurey-Pisier Theorem (for type).

1.11. In §5, it is shown that for Banach spaces the supremum of the metric types equals  $p_X$ , the supremum of the standard types. This enables us not only to look for results in metric spaces which are extensions of well-known results in Banach spaces, but also gives us the opportunity to apply to Banach space theory some nontrivial results which follow from the existence of metric type.

For example, this implies that if a Banach space  $X$  contains a sequence of subsets  $\{X_n; |X_n| = 2^n\}$  such that  $d(X_n, (C_2^n, \rho_p)) \leq a$  for some fixed  $a \geq 1$  and for some  $1 \leq p \leq 2$ , then for any  $\varepsilon > 0$ , it contains a sequence of subspaces  $\{Y_n\}_{n=1}^\infty$  such

that  $d(Y_n, l_p^n) \leq 1 + \varepsilon$ . Another consequence is that if  $X$  is an infinite-dimensional Banach space of type  $p$ , then for any  $1 \leq p' < p$  there is a constant  $\alpha_{p'}(X)$  such that for any  $n$  and any  $\psi: C_2^n \rightarrow X$  the following inequalities hold:

$$\left\{ \sum_{\varepsilon \in C_2^n} \|\psi(\varepsilon) - \psi(-\varepsilon)\|^2 \right\}^{1/2} \\ \leq \alpha_{p'}(X) n^{1/p' - 1/2} \left\{ \sum_{j=1}^n \sum_{\varepsilon \in C_2^n} \|\psi(\varepsilon_1, \dots, \varepsilon_j, \dots, \varepsilon_n) \right. \\ \left. - \psi(\varepsilon_1, \dots, -\varepsilon_j, \dots, \varepsilon_n)\|^2 \right\}^{1/2}.$$

(Here  $C_2^n = \{-1, 1\}^n$ .)

1.12. §6 discusses  $(1 + \varepsilon)$ -isomorphic embeddings of  $l_p$   $k$ -cubes into the  $l_1$   $n$ -cube (Hamming cube) for a suitable function  $n = n(k, \varepsilon)$  or more generally the embedding of  $l_p$   $k$ -cubes into  $l_s$   $n$ -cubes with  $s < p$ . A nonconstructive method with a power type estimate ( $n \sim k^3$ ) and a constructive method with a logarithmic estimate ( $k \sim \log n$ ) are shown. Here, however, the similarity with the normed space theory is not precise, because the embedding may be constructed for any  $1 \leq s < p \leq \infty$ . (In normed spaces it is true only for  $p \leq 2$ .)

## 2. Type 1 case.

2.1. Let  $(X_N, \rho_N)$ ,  $|X| = N$  ( $N \uparrow \infty$ ), be a family of finite metric spaces. It is easily verified that every such family is at least of type 1 with 1-type constant 1; namely, for every  $k$ -cube

$$(2.1.1) \quad \left[ \sum_D \text{diag}^2 \right]^{1/2} \leq k^{1/2} \left[ \sum_E \text{edge}^2 \right]^{1/2}.$$

Indeed, let  $\text{diag}(\varepsilon)$  be any diagonal and take the “identity path”  $E_\varepsilon(\text{Id})$ . By the triangle and Cauchy-Schwartz inequalities

$$(2.1.2) \quad \text{diag}^2(\varepsilon) \leq \left( \sum_{E_\varepsilon(\text{Id})} \text{edge} \right)^2 \leq k \left( \sum_{E_\varepsilon(\text{Id})} \text{edge}^2 \right).$$

Summing over all the diagonals we get (2.1.1), since the edges, belonging to identity paths of different diagonals, are all distinct.

2.2. Consider the trivial case when equality in (2.1.1) is obtained.

**PROPOSITION.** *Let  $(X, \rho)$  be a metric space of cardinality  $2^k$ , such that there exists a map  $\psi: C_2^k \rightarrow X$ , for which equality in (2.1.1) holds. Then  $X$  is isometric to the  $C_2^k$  cube with the usual  $l_1$  (Hamming) metric.*

The proof is obvious. Indeed, equality in (2.1.1) means equality in (2.1.2) for every diagonal.

For any  $\varepsilon$  take the diagonal  $\text{diag}(\varepsilon)$ . Equality in the Cauchy-Schwartz inequality implies that all the edges in  $E_\varepsilon(\text{Id})$  are mutually equal.

The same argument shows that for a given permutation  $\pi$  and every  $\varepsilon$  all the edges of  $E_\varepsilon(\pi)$  are mutually equal. Hence, all  $k2^{k-1}$  edges are equal. Denote this length by  $a$ .

Equality in the triangle inequality in (2.1.2) implies that the length of every diagonal is  $ka$ . Take any  $\varepsilon, \varepsilon'$  in  $C_2^k$  such that the Hamming distance  $h(\varepsilon, \varepsilon') = h$ . Clearly

$$ha = \text{diag}(\varepsilon) - (k - h)a \leq \rho(\psi(\varepsilon), \psi(\varepsilon')) \leq ha.$$

2.3. Consider a family  $(X_N, \rho_N)$  of type not better than 1. We will show that in this case there is a subset of  $X_N$  (with cardinality growing to infinity as  $N \uparrow \infty$ ) which is “almost” isometric to a Hamming cube. The proofs in 2.3–2.6 are similar to G. Pisier’s proof in lineary theory (see [P1]).

First we consider a more general situation and show that for any  $p \geq 1$  the  $p$ -type constant is submultiplicative, namely:

LEMMA. For any metric space  $(X, \rho)$  and any  $p \geq 1$ ,

$$\alpha_p(nk; X) \leq \alpha_p(k; X)\alpha_p(n; X).$$

PROOF. Assume  $|X| = 2^{nk}$  and let  $\psi: C_2^{nk} \rightarrow X$  be an ordering of  $X$ . Divide  $\{1, \dots, nk\}$  into  $k$  consecutive  $n$ -blocks  $I_i = \{(i-1)n+1, \dots, in\}$ ,  $i = 1, \dots, k$ . For a given configuration  $\varepsilon = (\varepsilon^j)_{j=1}^{nk}$  consider only those points obtained by changing simultaneously values of  $n$ -blocks  $I_i$  in  $\varepsilon$ :

$$\varepsilon_{I_{i_1}, \dots, I_{i_s}} = \begin{cases} 1 - \varepsilon^j, & j \in I_{i_1} \cup \dots \cup I_{i_s}, \\ \varepsilon^j, & \text{otherwise,} \end{cases} \quad 1 \leq i_1 < i_2 < \dots < i_s \leq k.$$

Clearly, we get a  $k$ -cube  $Y(\varepsilon; k)$  ( $|Y| = 2^k$ ), where an edge is a pair of points  $(\psi(\varepsilon), \psi(\varepsilon'))$ , such that  $\varepsilon$  and  $\varepsilon'$  differ on one block  $I_i$  only and a diagonal is a pair of points differing on all  $k$   $n$ -blocks.

By the definition of  $\alpha_p(k; X)$  we obtain in the cube  $Y(\varepsilon; k)$

$$(2.3.1) \quad \sum \text{diag}^2(Y) \leq \alpha_p^2(k; X)k^{2/p-1} \sum \text{edge}^2(Y).$$

Consider all the different  $k$ -cubes  $Y(\varepsilon; k)$  defined above. (There are  $2^{n(k-k)}$  such cubes.)

The edges of these cubes may be considered as diagonals of  $n$ -cubes, obtained by changing all the coordinates on one of the  $n$ -blocks  $\{I_i\}$  and keeping the coordinates of all other  $n$ -blocks constant. It is easily seen that summing (2.3.1) over all different  $k$  cubes and applying the  $p$ -type condition on the new  $n$ -cubes results in

$$\sum_D \text{diag}^2 \leq \alpha_p^2(k; X)k^{2/p-1} \alpha_p^2(n; X)n^{2/p-1} \sum_E \text{edge}^2$$

where the sums are over all the diagonals and all the edges of the original  $nk$ -cube. Hence,  $\alpha_p(nk; X) \leq \alpha_p(k; X)\alpha_p(n; X)$ .  $\square$

2.4. COROLLARY. Let  $0 < \alpha < 1$ ,  $0 < \eta < 1$ ,  $p \geq 1$ . If  $\alpha_p(n; X) \geq \alpha$ , then there is a constant  $C(\alpha, \eta)$  and an integer  $k = C(\alpha; n) \log \log n$  such that  $\alpha_p(k; X) \geq 1 - \eta/k^2 2^{k+12}$ .

PROOF. By the previous lemma  $k^t = n$  for some integer  $t$  implies  $\alpha_p(n; X) \leq \alpha_p^t(k; X)$ . Hence, if  $\alpha_p(k; X) = 1 - \delta_k$ , then  $(1 - \delta_k)^t \geq \alpha$ ,  $\delta_k \lesssim -\log \alpha/t$ . Take

$t \sim -(\log \alpha / \eta) k^2 2^{k+12}$ ; then  $\delta_k \sim \eta / k^2 2^{k+12}$  and  $n = k^{-(\log \alpha / \eta) k^2 2^{k+12}}$ . Hence  $k \sim C(\alpha; \eta) \log \log n$ .  $\square$

We will also need the following trivial fact.

2.5. FACT. Let  $0 < \delta < 1$ ,  $a_i \geq 0$  ( $i = 1, \dots, n$ ) and

$$(1 - \delta) \sqrt{n} \left( \sum_{i=1}^n a_i^2 \right)^{1/2} \leq \sum_{i=1}^n a_i.$$

Then, for every  $i = 1, \dots, n$ ,

$$\left| \frac{a_i}{(\sum a_i^2)^{1/2}} - \frac{1}{\sqrt{n}} \right| \leq \sqrt{2\delta}.$$

PROOF. The result follows immediately from

$$\sum_{i=1}^n \left[ \frac{a_i}{(\sum a_i^2)^{1/2}} - \frac{1}{\sqrt{n}} \right]^2 = \sum_{i=1}^n \left[ \frac{a_i^2}{\sum a_i^2} - 2 \frac{\sum a_i}{\sqrt{n} (\sum a_i^2)^{1/2}} + \frac{1}{n} \right] \leq 2 - 2(1 - \delta) = 2\delta.$$

2.6. THEOREM. Let  $\{(X_n, \rho_n)\}_n$  be a family of finite metric spaces ( $|X_n| = 2^n$ ) with type no better than 1. Then, for every  $\delta > 0$ , there is a  $k(n)$ -cube in  $(X_n, \rho_n)$  which is  $(1 + \delta)$ -isometric to  $(C_2^k, h)$ , the  $l_1$   $k$ -cube and  $k(n) \uparrow \infty$  with  $n \uparrow \infty$ .

PROOF. Since  $(X, \rho_n)$  has type no better than 1, then by definition (1.6.3) we get

$$\lim \frac{|\log \alpha_1(n; X_n)|}{\log n} = 0.$$

Hence, for every  $0 < \delta < 1$  there is a subsequence  $\{n_k\}_{k=1}^\infty$  such that

$$0 \leq -\frac{\log \alpha_1(n_k; X_{n_k})}{\log n_k} \leq \frac{\delta^2}{k^2 2^{k+12} \log k}.$$

Then we have  $n_k \geq k^t$ , where  $t \sim -\delta^{-2} k^2 2^{k+12} \log \alpha_1(n_k; X_{n_k})$ . By 2.4

$$\alpha_1(k; X_{n_k}) \geq [\alpha_1(n_k; X_{n_k})]^{1/t} \sim \exp(-\delta^2 / k^2 2^{k+12}) \sim 1 - \theta,$$

denoting  $\theta = \delta^2 / k^2 2^{k+12}$ . Hence, there is a  $k$ -cube (namely, a map  $\psi: C_2^k \rightarrow X_{n_k}$ ) satisfying

$$\left( \sum_D \text{diag}^2 \right)^{1/2} \geq (1 - \theta) \sqrt{k} \left( \sum_E \text{edge}^2 \right)^{1/2}.$$

We may normalize by taking  $\sum_E \text{edge}^2 = 1$ . As in 2.2 assign to each diagonal  $(\text{diag}(\varepsilon))$  the  $k$  edges of its "identity path" ( $E_\varepsilon(\text{Id})$ ). Clearly, there is some diagonal  $\text{diag}(\varepsilon_0)$  satisfying

$$k \sum_{E_0(\text{Id})} \text{edge}^2 \geq \text{diag}^2(\varepsilon_0) \geq 2^{-k+1} (1 - \theta)^2 k \sum_E \text{edge}^2 = (1 - \theta)^2 k 2^{-k+1}.$$

We may assume  $\varepsilon_0 = \bar{0} = (0, \dots, 0)$ . Then

$$\begin{aligned}
 (2.6.1) \quad \left( \sum_{E_0(\text{Id})} \text{edge} \right)^2 &\geq \text{diag}^2(\bar{0}) \geq k \left[ (1 - \theta)^2 \sum_E \text{edge}^2 - \sum_{E \setminus E_0(\text{Id})} \text{edge}^2 \right] \\
 &= k \left[ \sum_{E_0(\text{Id})} \text{edge}^2 - (2\theta - \theta^2) \sum_E \text{edge}^2 \right] > k(1 - 2^{k+0.1}\theta) \sum_{E_0(\text{Id})} \text{edge}^2.
 \end{aligned}$$

By 2.5 every edge in  $E_0(\text{Id})$  satisfies

$$(2.6.2) \quad \left| \frac{\text{edge}}{(\sum_{E_0(\text{Id})} \text{edge}^2)^{1/2}} - \frac{1}{\sqrt{k}} \right| < (2^{k+0.1}\theta)^{1/2} < \frac{\delta}{48k}.$$

Hence, the  $k$  edges of  $E_0(\text{Id})$  are almost equal. This is the first step in evaluating all the edges of the  $k$ -cube.

The second step is to take all other paths  $E_0(\pi)$  ( $\pi \neq \text{Id}$ ) having at least one common edge with  $E_0(\text{Id})$ . For each such path

$$\sum_{E_0(\pi)} \text{edge}^2 \geq \frac{1}{k} \text{diag}^2(\bar{0}) \geq (1 - \theta)^2 2^{-k+1}.$$

As before, we obtain that an inequality similar to (2.6.2) holds for the edges of  $E_0(\pi)$ .

Since  $E_0(\pi)$  and  $E_0(\text{Id})$  share at least one edge, it follows that the edges of  $E_0(\text{Id})$  and  $E_0(\pi)$  (for all such  $\pi$ ) are “almost equal”.

In the next step paths having at least one edge, which were evaluated at a previous step, are considered. By this process all edges are evaluated in at most four steps.

(i) Get all the edges of  $E_0(\text{Id})$ , namely edges  $(\psi(\varepsilon_1), \psi(\varepsilon_2))$ , where

$$\varepsilon_1 = \begin{cases} \varepsilon_1^j = 0, & j > i, \\ \varepsilon_1^j = 1, & j \leq i, \end{cases} \quad \varepsilon_2 = \begin{cases} \varepsilon_2^j = 0, & j > i + 1, \\ \varepsilon_2^j = 1, & j \leq i + 1. \end{cases}$$

(ii) Get all the edges  $(\psi(1, 0, \dots, 0), \psi(\varepsilon_i))_{i>1}$ , where  $\varepsilon_i$  has 1 in the first and  $i$ th coordinate ( $i \neq 1$ ). Such edge is obtained by a path  $E_0(\pi)$  passing through this edge and sharing the edge  $(\psi(0, 0, \dots, 0), \psi(1, 0, \dots, 0))$  with  $E_0(\text{Id})$  ( $\pi = (1, i, \dots)$ ).

(iii) Get all the edges  $(\psi(0, \dots, 0), \psi(\varepsilon_i))_{i>1}$ , where  $\varepsilon_i$  is the  $i$ th unit vector. This is achieved by a path starting with the required edge and proceeding by an edge built in (ii). The required permutation is  $\pi = (i, 1, \dots)$ .

(iv) Any edge can be obtained by a path beginning with an edge built in (iii).

Hence, all the edges are “almost equal” and for any two edges of length  $a$  and  $a'$  we have  $|a - a'| < \delta/6k$ .

Define  $a$  to be the length of some edge. Then for any diagonal (by (2.6.1))

$$\begin{aligned}
 k^2 \left( 1 - \frac{\delta}{6k} \right)^2 a^2 - k(2\theta - \theta^2) &< \text{diag}^2 < k^2 \left( 1 + \frac{\delta}{6k} \right)^2 a^2, \\
 1 - \frac{\delta}{5k} &\leq \frac{\text{diag}}{ka} \leq 1 + \frac{\delta}{6k}.
 \end{aligned}$$

A standard use of the triangle inequality shows, now, that for any  $\varepsilon_1, \varepsilon_2 \in C_2^k$ ,  $|1 - \rho(\psi(\varepsilon_1), \psi(\varepsilon_2))/ah(\varepsilon_1, \varepsilon_2)| < \delta/2$ . Hence,  $\|\psi\|_{\text{Lip}}\|\psi^{-1}\|_{\text{Lip}} < 1 + \delta$ .  $\square$

2.7. From the proof of 2.6 it is obvious that the following proposition holds.

**PROPOSITION.** *Let  $0 < \alpha < 1$  and  $\delta < 0$ . Then there exists a constant  $C(\alpha; \delta) > 0$  depending on  $\alpha > 0$  and  $\delta > 0$  only, such that every metric space  $(X, \rho)$  with  $\alpha_1(n; X) \geq \alpha$  contains a  $k$ -cube which is  $(1 + \delta)$ -isometric to the  $l_1$   $k$ -cube for some  $k \geq C(\alpha; \delta) \log \log n$ . (Of course, this has nontrivial sense only for  $n$  large enough.)*

**3. Improved estimates; “concentration of measure phenomena” approach.** In this section a different approach is used to investigate the situation of Proposition 2.7. Much better (power-type) estimates are obtained on the cardinality of a Hamming subcube. However, a restriction on the distribution of edges is imposed.

3.1. **THEOREM.** *Let  $(X, \rho)$  be a finite metric space of cardinality  $2^n$ . Let  $\psi: C_2^n \rightarrow X$  be some ordering of  $X$  satisfying the following conditions:*

(i) *there is a constant  $0 < \gamma \leq 1$  such that*

$$\left( \sum_D \text{diag}_{\psi}^2 \right)^{1/2} \geq \gamma n^{1/2} \left( \sum_E \text{edge}_{\psi}^2 \right)^{1/2},$$

(ii) *there are constants  $0 \leq \beta < \frac{1}{2}$ ,  $\alpha > 0$  such that  $\max_E \text{edge}_{\psi} \leq \alpha n^{\beta}$ , where  $A$  is the  $l_2$ -average of the edges ( $A^2 = (1/n2^n) \sum_E \text{edge}_{\psi}^2$ ).*

*Then, there are a constant  $C(\alpha, \gamma)$  depending on  $\alpha$  and  $\gamma$  only and a subset  $Y \subset X$  of cardinality  $2^k$ ,  $k = C(\alpha, \gamma)n^{1/2-\beta}(\log n)^{-1/2}$  such that  $(Y, \rho)$  is  $3\gamma^{-1}$  isomorphic to  $(C_2^k, h)$ .*

The proof of Theorem 3.1 uses the so-called “concentration of measure phenomena” on some discrete spaces (see also [A-M, M-S]). We need some definitions and known results, which are stated without proof.

3.2. In a metric space  $(X, \rho)$  a  $\delta$ -inflation of a subset  $\mathcal{A} \subset X$  is defined to be the set  $\mathcal{A}_{\delta} = \{x \in X | \exists a \in \mathcal{A}, \rho(x, a) \leq \delta\}$ .

Given a metric probability space  $(X, \rho, \mu)$  and a real-valued function on it, we denote by  $L_f$  a Levy mean (median) of  $f: X \rightarrow \mathbb{R}$ ; namely,  $L_f$  is a number such that both  $\mu(\mathcal{A}^+) \geq \frac{1}{2}$  and  $\mu(\mathcal{A}^-) \geq \frac{1}{2}$ , where  $\mathcal{A}^+ = \{x \in X: f(x) \geq L_f\}$ ,  $\mathcal{A}^- = \{x \in X: f(x) \leq L_f\}$ .

3.3. **DEFINITION.** Let  $(X_n, \rho_n, \mu_n)$ ,  $n \nearrow \infty$ , be a family of normalized metric probability spaces (i.e.  $\text{diam } X_n = 1$ ,  $\mu_n(X_n) = 1$ ). It is called a “normal Levy family” if for every  $\delta > 0$  and  $\mathcal{A}_n \subset X_n$  with  $\mu_n(\mathcal{A}_n) \geq \frac{1}{2}$  we get  $\mu_n(\mathcal{A}_n)_{\delta} \geq 1 - c_1 \exp(-c_2 \delta^2 n)$ , where  $c_1, c_2$  are some absolute constants independent of  $n$ .

We will need the fact that the family of the Hamming cubes and the family of the  $n$ -permutations are normal Levy families.

3.4. **LEMMA [H, A-M].**  $(C_2^n, \bar{\rho}_n, \bar{\mu}_n)$ , where  $C_2^n$  is the Hamming  $n$ -cube of cardinality  $2^n$ ,  $\bar{\rho}_n = n^{-1}h_n$  ( $h_n$ -Hamming metric), and  $\bar{\mu}_n$  the standard normalized counting measure, is a normal Levy family with  $c_1 = \frac{1}{2}$ ,  $c_2 = 2$ .

3.5. **LEMMA [M].**  $(S_n, \tilde{\rho}_n, \tilde{\mu}_n)$ , where  $S_n$  is the permutation group on  $n$  elements,  $\tilde{\rho}_n(\sigma, \tau) = \frac{1}{n} \text{card}\{i: \sigma(i) \neq \tau(i)\}$  for every pair of permutations  $\sigma, \tau \in S_n$  and  $\tilde{\mu}_n(A \subset S_n) = \frac{1}{n!} \text{card } A$ , is a normal Levy family with  $c_1 = \frac{1}{2}$ ,  $c_2 = \frac{1}{16}$ .



3.6. COROLLARY. Let  $X_n = C_2^n \times S_n$  with the product measure  $\mu_n = \bar{\mu}_n \times \tilde{\mu}_n$  and metric  $\rho_n((\varepsilon, \sigma), (\varepsilon', \sigma')) = \#\{i: \varepsilon(i) \neq \varepsilon'(i) \text{ or } \sigma(i) \neq \sigma'(i)\}$ , where  $\bar{\mu}_n, \tilde{\mu}_n$  are as in Lemmas 3.4 and 3.5 respectively. Then  $(X_n, \rho_n, \mu_n)$  is a normal Levy family with  $c_1 = \frac{1}{2}$ ,  $c_2 = \frac{1}{16}$ .

We need two more results.

3.7. LEMMA [A-M]. Let  $(X_n, \rho_n, \mu_n)$  be a normal Levy family with constants  $c_1, c_2$  and let  $f: X_n \rightarrow R$  be a measurable function with modulus of continuity  $\omega_f$ . Then for every  $\delta > 0$

$$\left| L_f - \int_{X_n} f d\mu_n \right| \leq 2c_1 \exp(-c_2 \delta^2 n) \|f\|_\infty + \omega_f(\delta).$$

Here  $\|f\|_\infty = \max_{x \in X_n} |f(x)|$ .

3.8. LEMMA (SEE THEOREM 7.8 IN [M-S]). Let  $(X_n, \rho_n, \mu_n)$  be as in Corollary 3.6 and let  $f_n: X_n \rightarrow R$  satisfy for every  $x, y \in X_n$ ,  $|f_n(x) - f_n(y)| \leq \alpha \rho_n(x, y)$  with some constant  $\alpha > 0$ . Then

$$\mu_n(|f_n - Ef_n| \geq c) \leq 2 \exp(-c^2 n / 8\alpha^2).$$

$Ef_n$  is the expectation  $Ef_n = \int_{X_n} f_n d\mu_n$ .

3.9. PROOF OF THEOREM 3.1. Define a function  $\varphi: C_2^n \rightarrow R^+$  by  $\varphi(\varepsilon) = \text{diag}_\psi(\varepsilon)$  for every  $\varepsilon \in C_2^n$ .

The conditions of the theorem imply

$$\gamma An \leq \left[ \int_{C_2^n} \varphi^2(\varepsilon) \right]^{1/2} \leq An.$$

First evaluate

$$\left| L_\varphi - \left( \int_{C_2^n} \varphi^2(\varepsilon) \right)^{1/2} \right| \leq \left( \int_{C_2^n} |L_\varphi - \varphi(\varepsilon)|^2 \right)^{1/2}.$$

By Lemmas 3.4 and 3.7 for every  $\delta > 0$

$$\int_{C_2^n} |L_\varphi - \varphi(\varepsilon)|^2 \leq \exp(-2\delta^2 n) \left( n \max_E \text{edge} \right)^2 + \left( \delta n \max_E \text{edge} \right)^2.$$

Since  $\max_E \text{edge} \leq \alpha An^\beta$  ( $\beta < \frac{1}{2}$ ), we get (by calculating for  $\delta \leq \gamma/8\alpha n^\beta$ )

$$\left| L_\varphi - \left( \int \varphi^2(\varepsilon) \right)^{1/2} \right| < \frac{1}{4} \gamma An.$$

Hence

$$(3.9.1) \quad \frac{3}{4} \gamma An \leq L_\varphi \leq An \left( 1 + \frac{1}{4} \gamma \right).$$

By the same consideration for every  $\delta > 0$

$$\int |L_\varphi - \varphi(\varepsilon)| \leq \alpha An^\beta [\delta n + n \exp(-2\delta^2 n)],$$

and it is easily seen that

$$(3.9.2) \quad \frac{1}{2}\gamma An \leq \int_{C_2^n} \varphi \leq \left( \int_{C_2^n} \varphi^2 \right)^{1/2} \leq An.$$

Define  $\gamma_0 = \frac{1}{2}\gamma$ .

3.10. Now, we construct the subset which is isomorphic to the  $l_1$ -cube. Every edge of the new cube will be built on  $s$ -blocks of edges of the original cube.

Let  $s$  be an integer such that  $s|n$ , and divide  $\{1, \dots, n\}$  into  $k \equiv n/s$  blocks of length  $s$  each:

$$I_i = \{(i-1)s + 1, \dots, is\}, \quad i = 1, 2, \dots, k.$$

Define  $k$  functions  $\varphi_i: C_2^n \times S_n \rightarrow R^+$  by  $\varphi_i(\varepsilon, \sigma) = \rho(\psi(\sigma\varepsilon), \psi(\sigma\varepsilon_{I_i}))$ , where  $\varepsilon = (\varepsilon_i)_{i=1}^n \in C_2^n$ ,  $\sigma \in S_n$ ,  $\sigma\varepsilon = (\varepsilon_{\sigma(j)})_{j=1}^n$ ,  $\varepsilon_{I_i} = (\varepsilon_{I_i}^j)_{j=1}^n$  and

$$\varepsilon_{I_i}^j = \begin{cases} \varepsilon_j, & j \notin I_i \\ 1 - \varepsilon_j, & j \in I_i \end{cases}.$$

Clearly

$$(3.10.1) \quad \int_{S_n} \int_{C_2^n} \varphi_i(\varepsilon, \sigma) = \int_{S_n} \int_{C_2^n} \varphi_j(\varepsilon, \sigma)$$

for every  $i, j = 1, \dots, k$ . Define  $\gamma_1$  by  $\gamma_1 = (As)^{-1} \int_{S_n} \int_{C_2^n} \varphi_1(\varepsilon, \sigma)$ . By (3.9.2) and (3.10.1)

$$(3.10.2) \quad \gamma_0 An \leq \int_{C_2^n} \varphi \leq \sum_{i=1}^k \int_{S_n} \int_{C_2^n} \varphi_i(\varepsilon, \sigma) = \gamma_1 An.$$

Hence  $\gamma_1 \geq \gamma_0$ . Define  $E_1 = \int_{S_n} \int_{C_2^n} \varphi_1(\varepsilon, \sigma)$  and

$$B_i(\delta) = \{(\varepsilon, \sigma) : |\varphi_i(\varepsilon, \sigma) - E_1| \leq \delta As\} \quad \text{for some } \delta > 0.$$

By Lemma 3.8

$$\begin{aligned} \mu(B_i(\delta)) &\geq 1 - 2 \exp \left( -\frac{(\delta As)^2 n}{8(n \max_E \text{edge})^2} \right) \\ &\geq 1 - 2 \exp \left( -\frac{\delta^2 s^2}{8\alpha^2 n^{1+2\beta}} \right) \end{aligned}$$

and

$$\mu \left( \bigcap_{i=1}^k B_i(\delta) \right) \geq 1 - 2k \exp \left( -\frac{\delta^2 s^2}{8\alpha^2 n^{1+2\beta}} \right).$$

Notice that for  $i = 1, \dots, k$  and for every  $(\varepsilon, \sigma) \in \bigcap_{i=1}^k B_i(\delta)$  we have

$$(3.10.3) \quad A(\gamma_1 - \delta)s \leq \varphi_i(\varepsilon, \sigma) \leq A(\gamma_1 + \delta)s.$$

3.11. Now, we apply the same procedure to pairs of blocks  $I_i, I_j$  ( $i \neq j$ ). Consider the functions  $\varphi_{i,j}(\varepsilon, \sigma) = \rho(\psi(\sigma\varepsilon), \psi(\sigma\varepsilon_{I_i \cup I_j}))$  ( $i, j = 1, \dots, k; i \neq j$ ) and define  $E_2, \gamma_2$  by  $E_2 = \int_{S_n} \int_{C_2^n} \varphi_{1,2}(\varepsilon, \sigma)$ ,  $\gamma_2 = (2As)^{-1} E_2$ .

Clearly,  $E_2 = \int_{S_n} \int_{C_2^n} \varphi_{i,j}(\varepsilon, \sigma)$  for every  $i \neq j$ . Hence

$$\gamma_0 An \leq \int_{S_n} \int_{C_2^n} \varphi(\varepsilon, \sigma) \leq \frac{n}{2s} E_2 = \gamma_2 An$$

and  $\gamma_2 \geq \gamma_0$ . On the other hand  $\gamma_1 \geq \gamma_2$ , since  $2E_1 \geq E_2$ .

Define for  $i \neq j$  the sets

$$B_{i,j}(\delta) = \{(\varepsilon, \sigma) : |\varphi_{i,j}(\varepsilon, \sigma) - E_2| \leq 2\delta As\}.$$

By Lemma 3.8

$$\begin{aligned} \mu(B_{i,j}(\delta)) &\geq 1 - 2 \exp\left(-\frac{(2\delta s)^2}{8\alpha^2 n^{1+2\beta}}\right), \\ \mu\left(\bigcap_{i \neq j} B_{i,j}(\delta)\right) &\geq 1 - 2 \binom{k}{2} \exp\left(-\frac{(2\delta s)^2}{8\alpha^2 n^{1+2\beta}}\right) \end{aligned}$$

and for every  $(\varepsilon, \sigma) \in \bigcap_{i \neq j} B_{i,j}(\delta)$  we have

$$2A(\gamma_2 - \delta)s \leq \varphi_{i,j}(\varepsilon, \sigma) \leq 2A(\gamma_2 + \delta)s.$$

3.12. We proceed in a similar way and for every  $t$  ( $t = 1, \dots, k$ ) we define functions

$$\varphi_{i_1, \dots, i_t}(\varepsilon, \sigma) = \rho((\sigma\varepsilon), \psi(\sigma\varepsilon_{I_1 \cup \dots \cup I_t})), \quad (1 \leq i_1 < i_2 < \dots < i_t \leq k),$$

constants  $\gamma_t$ , expectations  $E_t$  and “almost constancy” sets  $B_{i_1, \dots, i_t}$  such that

$$(3.12.1) \quad E_t = \int_{S_n} \int_{C_2^n} \varphi_{i_1, \dots, i_t}(\varepsilon, \sigma) = \gamma_t t As.$$

For every  $(\varepsilon, \sigma) \in \bigcap_{1 \leq i_1 < \dots < i_t \leq k} B_{i_1, \dots, i_t}$  we have

$$(3.12.2) \quad tA(\gamma_t - \delta)s \leq \varphi_{i_1, \dots, i_t}(\varepsilon, \sigma) \leq tA(\gamma_t + \delta)s$$

and

$$(3.12.3) \quad \mu\left(\bigcap B_{i_1, \dots, i_t}\right) \geq 1 - 2 \binom{k}{t} \exp\left(-\frac{(ts)^2}{8\alpha^2 n^{1+2\beta}}\right)$$

As in 3.11, it is easy to show that

$$(3.12.4) \quad \gamma_0 \leq \gamma_t \leq \gamma_{t-1} \leq \dots \leq \gamma_2 \leq \gamma_1 \leq 1.$$

Assume that we have found an  $s$  such that  $B = \bigcap_{t=1}^k [\bigcap B_{i_1, \dots, i_t}]$ ,  $k = [n/s]$ , is not empty.

Then there exists a pair  $(\varepsilon, \sigma) \in B$ , which means that there are a binary vector  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$  and an  $n$ -permutation  $\sigma \in S_n$  such that taking  $k$  blocks

$$I_i = (\varepsilon_{\sigma[(i-1)s+1]}, \dots, \varepsilon_{\sigma(is)}), \quad (i = 1, \dots, k),$$

and defining “new edges” to be pairs of binary vectors, which are different exactly on one of the blocks  $I_i$ , gives us an  $k$ -cube, which is  $\tilde{\gamma} = (\gamma_1 + \delta)(\gamma_k - \delta)^{-1}$  isomorphic to the  $l_1^k$  (Hamming) cube.

Hence, in order to finish the proof we have to calculate  $\tilde{\gamma}$  and  $s$ . By (3.12.4) and by the definition of  $\gamma_0$

$$\tilde{\gamma} \leq (1 + \delta)(\frac{1}{2}\gamma - \delta)^{-1} \leq 3\gamma^{-1} \quad \text{if } 0 \leq \delta \leq \gamma(6 + 2\gamma)^{-1}.$$

In order to get  $\mu(B) > 0$  it is enough to assume (see (3.12.3))

$$(3.12.5) \quad 2 \sum_{t=1}^k \binom{k}{t} \exp \left( -\frac{(t\delta s)^2}{8\alpha^2 n^{1+2\beta}} \right) < 1.$$

Define  $a = \exp(-(\delta s)^2 / 8\alpha^2 n^{1+2\beta})$ . Then (3.12.5) implies

$$(3.12.6) \quad \sum_{t=1}^k \binom{k}{t} a^{t^2} < \frac{1}{2}, \quad (a < 1),$$

and (3.12.6) is trivially satisfied if  $ka < \frac{1}{4}$ , which implies

$$\left( \frac{n}{s} \right) \exp \left( -\frac{(\delta s)^2}{8\alpha^2 n^{1+2\beta}} \right) < \frac{1}{4}, \quad s \sim C'(\alpha, \gamma) n^{1/2+\beta} (\log n)^{1/2}$$

and  $k = [n/s] = C(\alpha, \gamma) n^{1/2-\beta} / \sqrt{\log n}$ .  $\square$

3.13. An obvious consequence of Theorem 3.1 is the following

**THEOREM.** *Let  $(X, \rho)$  satisfy the conditions of Theorem 3.1. Then for every  $\eta > 0$  there is a constant  $C(\alpha, \gamma, \eta)$  and a subset  $Y \subset X$  of cardinality  $2^k$ , where  $k = C(\alpha, \gamma, \eta)[n^{1/2-\beta}(\log n)^{-1/2}]^{\log(1+\eta/2)/\log 3\gamma^{-1}}$  such that  $(Y, \rho)$  is  $(1+\eta)$ -isometric to the  $l_1^k$  Hamming cube.*

**PROOF.** By 3.12  $(X, \rho)$  is  $C_0$ -isomorphic to the  $l_1$   $m$ -cube, where  $C_0 = (\gamma_1 + \delta)(\gamma_m - \delta)^{-1} \leq 3\gamma^{-1}$  and  $m = C(\alpha, \gamma) n^{1/2-\beta} \log n^{-1/2}$ . Recall that  $\gamma_i$  is an increasing sequence ( $\gamma_0 \leq \gamma_m < \dots \leq \gamma_1 \leq 1$ ). Now

$$(\gamma_1 + \delta)(\gamma_m - \delta)^{-1} = (\gamma_1 + \delta)(\gamma_{\sqrt{m}} + \delta)^{-1}(\gamma_{\sqrt{m}} + \delta)(\gamma_m - \delta)^{-1}.$$

Hence,

(i)  $(\gamma_{\sqrt{m}} + \delta)(\gamma_m - \delta)^{-1} \leq \sqrt{C_0}$  and then there is an  $\sqrt{m}$  subset of  $(X, \rho)$  which is  $\sqrt{C_0}$ -isomorphic to the  $l_1$   $m$ -cube, or

(ii)  $(\gamma_1 + \delta)(\gamma_{\sqrt{m}} + \delta)^{-1} \leq \sqrt{C_0}$ . In this case define  $C_1 = (\gamma_1 + \delta)(\gamma_{\sqrt{m}} - \delta)^{-1}$  and  $\xi = (\gamma_1 + \delta)(\gamma_0^2 - \delta^2)^{-1}$ . There is a  $\sqrt{m}$  subset of  $(X, \rho)$  which is  $C_1$ -isomorphic to the  $l_1$   $\sqrt{m}$ -cube and a simple calculation shows  $C_1 \leq \sqrt{C_0} + 2\xi\delta$ .

Applying the last argument  $t$  times we get a  $k = m^{2^{-t}}$  subcube, which is  $C_t$ -isomorphic to the  $l_1$   $k$ -cube. (The sequence  $(C_j)_{j=1}^t$  satisfies  $1 \leq C_{j+1} \leq \sqrt{C_j} + 2\xi\delta$  and hence  $C_t \leq C_0^{2^{-t}} + 4\xi\delta$ .)

In order to get the required result it is enough to have  $4\xi\delta < \eta/2$  (which can be obtained in proof of Theorem 3.1) and  $C_0^{2^{-t}} \leq 1 + \eta/2$  which implies  $2^{-t} \sim \log(1 + \eta/2)(\log C_0)^{-1}$ . Hence, there is a constant  $C(\alpha, \gamma, \eta)$  such that there is a  $k = C(\alpha, \gamma, \eta)[n^{1/2-\beta} \log n^{-1/2}]^{\log(1+\eta/2)(\log 3\gamma^{-1})^{-1}}$  subcube of  $(X, \rho)$  which is  $(1+\eta)$ -isometric to the  $l_1$   $k$ -cube.  $\square$

3.14. In [E] P. Enflo investigated a property of metric spaces which he called  $p$ -roundness and which suggests the following definition of metric type, in analogy with (1.6.1).

**DEFINITION.** Let  $p \geq 1$ . An (infinite) metric space  $(X, \rho)$  has  $E$ -type  $p$  if there is a constant  $a$  such that for every  $k$  and every  $k$ -cube (defined by any map  $\psi: C_2^k \rightarrow X$ )

$$(3.14.1) \quad \left( \sum_D \text{diag}_\psi^p \right)^{1/p} \leq a \left( \sum_E \text{edge}^p \right)^{1/p},$$

where the summing is over all the diagonals and all the edges of the  $k$ -cube.

The smallest  $a$  satisfying (3.14.1) is called the  $p$   $E$ -type constant of  $(X, \rho)$  and denoted  $a_p(X)$ . The smallest constant for a given  $k$  is denoted  $a_p(k; X)$ .

From a proof similar to that of 2.3 it is obvious that  $a_p(nk; X) \leq a_p(k; X)a_p(n; X)$ .

3.15. Following the lines of the proof of Theorem 3.1, where  $A$  is defined as  $A = (1/n2^n) \sum_E \text{edge}$  (which only simplifies the proof), we get the following

**THEOREM.** *Let  $(X, \rho)$  be a finite metric space  $(|X|) = 2^n$ . Let  $\psi: C_2^n \rightarrow X$  be some ordering of  $X$  satisfying the following conditions:*

(i) *there is a constant  $0 < \gamma \leq 1$  such that  $\sum_D \text{diag}_\psi \geq \gamma \sum_E \text{edge}_\psi$ ,*

(ii) *there are constants  $0 \leq \beta < \frac{1}{2}$ ,  $\alpha > 0$  such that  $\max_E \text{edge}_\psi \leq \alpha A n^\beta$ , where  $A$  is the  $l_1$ -average of the edges ( $A = (1/n2^n) \sum_E \text{edge}_\psi$ ).*

*Then there is a constant  $C(\alpha, \gamma)$  and a subset  $Y \subset X$  of cardinality  $k = C(\alpha, \gamma)n^{1/2-\beta}$  such that  $(Y, \rho)$  is  $3\gamma^{-1}$ -isomorphic to  $(C_2^k, h)$ .*

Of course, an analog of 3.13 is also true. It can be shown that in this situation the restriction on the distributions of edges (see (ii)) is essential. Take for example the case where  $\gamma = 1$  and equality in (i) holds. Then it is easily shown that  $X$  is a cube satisfying

(a) for any diagonal  $\text{diag}_\psi(\varepsilon)$  and any path  $E_\varepsilon(\pi)$   $\text{diag}(\varepsilon) = \sum_{E_\varepsilon(\pi)} \text{edge}$ ,

(b) all the edges obtained by changing the  $i$ th coordinate of any  $\varepsilon$  are equal; namely, there are only  $n$  lengths  $e_1, \dots, e_n$  and  $\text{edge}_\psi(\varepsilon, \varepsilon') = \sum_{i=1}^n e_i |\varepsilon_i - \varepsilon'_i|$  for any  $\varepsilon = (\varepsilon_i)$ ,  $\varepsilon' = (\varepsilon'_i)$  of Hamming distance 1. (We may say that "opposite" edges are equal and  $X$  is a "brick".)

In this example, if  $e_i = q^i$  ( $q > 1$ ) we cannot find a subset isomorphic to the Hamming cube, hence some kind of restriction on the distribution of edges is crucial.

In Theorem 3.1 this restriction seems to us nonessential and Theorem 2.6 supports this assumption.

**4. The type- $p$  case.** The main result of this section is the type- $p$  analog of Theorem 3.1.

4.1. **THEOREM.** *Let  $(X, \rho)$  be a finite metric space of cardinality  $2^n$  and of type  $p$  with type- $p$  constant  $a_p$  for some  $1 \leq p < 2$ . Let  $\psi: C_2^n \rightarrow X$  be a map satisfying the following two conditions:*

(i) *there is a constant  $0 < \gamma \leq 1$  such that*

$$\left( \sum_D \text{diag}_\psi^2 \right)^{1/2} \geq \gamma a_p n^{1/p-1/2} \left( \sum_E \text{edge}_\psi^2 \right)^{1/2},$$

(ii) *there are constants  $0 \leq \beta < \frac{1}{p} - \frac{1}{2}$ ,  $\alpha > 0$  such that*

$$\max_E \text{edge}_\psi \leq \alpha A n^\beta, \quad \text{where } A^2 = \frac{1}{n2^n} \sum_E \text{edge}_\psi^2.$$

*Then there is a constant  $C(a_p, \alpha, \gamma)$  and a subset  $Y \subset X$  of cardinality  $2^k$ ,  $k = C(a_p, \alpha, \gamma)n^{1-1/2-\beta p}(\log n)^{-p/2}$  such that  $(Y, \rho)$  is  $4a_p^2\gamma^{-1}$ -isomorphic to the  $l_p$   $k$ -cube.*

**PROOF.** The proof is analogous to the proof of Theorem 3.1.

Define a function  $\varphi: C_2^n \rightarrow R^+$  by  $\varphi(\varepsilon) = \text{diag}_\psi(\varepsilon)$  for every  $\varepsilon \in C_2^n$ . By the conditions of the theorem

$$\gamma a_p A n^{1/p} \leq \left[ \int_{C_2^n} \varphi^2(\varepsilon) \right]^{1/2} \leq a_p A n^{1/p}.$$

As in 3.9 we can show that

$$(4.1.1) \quad \frac{1}{2} \gamma a_p A n^{1/p} \leq \int_{C_2^n} \varphi \leq \left( \int_{C_2^n} \varphi^2 \right)^{1/2} \leq a_p A n^{1/p}.$$

Now, as in 3.10, divide  $\{1, \dots, n\}$  into  $s$  blocks  $I_i$  and define the  $k \equiv [n/s]$  functions  $\varphi_i: C_2^n \times S_n \rightarrow R^+$ .

This time the constant  $\gamma_1$  is defined by

$$\gamma_1 = (a_p A s^{1/p})^{-1} \int_{S_n} \int_{C_2^n} \varphi_1(\varepsilon, \sigma).$$

If  $n^{1/2} \max_E \text{edge} \leq n^{1/2+\beta} \alpha A < \frac{1}{4} \gamma_1 a_p A s^{1/p}$ , then we can show as in 3.9 that

$$(4.1.2) \quad \left| \int_{S_n} \int_{C_2^n} \varphi_1(\varepsilon, \sigma) - \left[ \int_{S_n} \int_{C_2^n} \varphi_1^2(\varepsilon, \sigma) \right]^{1/2} \right| < \frac{1}{2} \gamma_1 a_p A s^{1/p}.$$

By applying the  $p$ -type condition on  $s$ -blocks we get

$$(4.1.3) \quad \begin{aligned} (a_p \gamma A)^2 n^{2/p} &\leq \int_{S_n} \int_{C_2^n} \varphi^2(\varepsilon, \sigma) \leq a_p^2 k^{2/p-1} \int_{S_n} \int_{C_2^n} \sum_{i=1}^k \varphi_i^2(\varepsilon, \sigma) \\ &= a_p^2 k^{2/p} \int_{S_n} \int_{C_2^n} \varphi_1^2(\varepsilon, \sigma) \leq \left( \frac{3}{2} \gamma_1 a_p A s^{1/p} \right)^2 a_p^2 k^{2/p}, \end{aligned}$$

where the last inequality follows from (4.1.2). Hence  $\gamma_1 \geq 2\gamma/3a_p$ .

Define  $E_1 = \int \varphi_1(\varepsilon, \sigma)$  and sets

$$B_i(\delta) = \{(\varepsilon, \sigma) : |\varphi_i(\varepsilon, \sigma) - E_1| \leq \delta a_p A s^{1/p}\} \quad \text{for some } \delta > 0.$$

As in 3.10  $\mu(B_i(\delta)) \geq 1 - 2 \exp(-\delta^2 a_p^2 s^{2/p} / 8\alpha^2 n^{1+2\beta})$  and

$$\mu \left( \bigcap_{i=1}^k B_i(\delta) \right) \geq 1 - 2k \exp \left( -\frac{\delta^2 a_p^2 s^{2/p}}{8\alpha^2 n^{1+2\beta}} \right).$$

For every  $(\varepsilon, \sigma) \in \bigcap_{i=1}^k B_i(\delta)$  we have

$$(4.1.4) \quad (\gamma_1 - \delta) a_p A s^{1/p} \leq \varphi_i(\varepsilon, \sigma) \leq (\gamma_1 + \delta) a_p A s^{1/p}.$$

Now proceed as in 3.11 and define for pairs of blocks  $I_i, I_j (i \neq j)$  the functions  $\varphi_{i,j}(\varepsilon, \delta)$ ,  $E_2 = \int_{S_n} \int_{C_2^n} \varphi_{1,2}(\varepsilon, \sigma) = \gamma_2 a_p A (2s)^{1/p}$ . As in (4.1.3) we get  $\gamma_2 \geq 2\gamma/3a_p$ . On the other hand

$$\begin{aligned} \left[ \frac{1}{2} \gamma_2 a_p A (2s)^{1/p} \right]^2 &\leq \int_{S_n} \int_{C_2^n} \varphi_{1,2}^2(\varepsilon, \sigma) \leq \int_{S_n} \int_{C_2^n} 2^{2/p-1} a_p^2 [\varphi_1^2(\varepsilon, \sigma) + \varphi_2^2(\varepsilon, \sigma)] \\ &= 2^{2/p} a_p^2 \int_{S_n} \int_{C_2^n} \varphi_1^2(\varepsilon, \sigma) \leq 2^{2/p} a_p^2 \left( \frac{3}{2} \gamma_1 a_p A s^{1/p} \right)^2. \end{aligned}$$

Hence  $\gamma_1 \geq 2\gamma_2/3a_p$ .

The sets

$$B_{i,j}(\delta) = \{(\varepsilon, \sigma) : |\varphi_{i,j}(\varepsilon, \sigma) - E_2| \leq \delta a_p A(2s)^{1/p}\}$$

satisfy

$$\mu \left( \bigcap_{i \neq j} B_{i,j}(\delta) \right) \geq 1 - 2 \binom{k}{2} \exp \left( -\frac{\delta^2 a_p^2 (2s)^{2/p}}{8\alpha^2 n^{1+2\beta}} \right)$$

and for every  $(\varepsilon, \sigma) \in \bigcap_{i \neq j} B_{i,j}(\delta)$  we get

$$(\gamma_2 - \delta) A a_p (2s)^{1/p} \leq \varphi_{i,j}(\varepsilon, \sigma) \leq (\gamma_2 + \delta) A a_p (2s)^{1/p}.$$

Proceeding in a similar way we define for every  $t$ ,  $t = 1, \dots, k$ , functions  $\varphi_{i_1, \dots, i_t}$ , constants  $\gamma_t$ , “almost constancy” sets  $B_{i_1, \dots, i_t}$  such that the following hold:

- (i)  $E_t = \int_{S_n} \int_{C_2^n} \varphi_{i_1, \dots, i_t}(\varepsilon, \sigma) = \gamma_t a_p A(ts)^{1/p}$ ,
- (ii)  $1 \geq \gamma_t \geq 2\gamma/3a_p$  for every  $t$  and  $\gamma_{t_1} \geq (2/3a_p)\gamma_{t_2}$  for every  $t_2 > t_1$ .
- (iii) for every  $1 \leq i_1 < i_2 < \dots < i_t \leq k$ ,

$$(\gamma_t - \delta) A a_p (ts)^{1/p} \leq \varphi_{i_1, \dots, i_t}(\varepsilon, \sigma) \leq (\gamma_t + \delta) A a_p (ts)^{1/p},$$

- (iv) for every  $(\varepsilon, \sigma) \in \bigcap_{(i_1, \dots, i_t)} B_{i_1, \dots, i_t}$ ,

$$\mu \left( \bigcap B_{i_1, \dots, i_t} \right) \geq 1 - 2 \binom{k}{t} \exp \left( -\frac{\delta^2 a_p^2 (ts)^{2/p}}{8\alpha^2 n^{1+2\beta}} \right).$$

Now we want to find an  $s$  such that  $B = \bigcap_{t=1}^k (\bigcap_{(i_1, \dots, i_t)} B_{i_1, \dots, i_t})$  is not empty and the condition in (4.1.2) is satisfied.

Define  $a = \exp(-\delta^2 a_p^2 s^{2/p}/8\alpha^2 n^{1+2\beta})$ . If  $s^{1/p} > (4\alpha/\gamma a_p) n^{1/2+\beta}$  and

$$2 \sum_{t=1}^k \binom{k}{t} a^{t^{2/p}} < 1,$$

then there is a pair  $(\varepsilon, \sigma) \in B$  such that the  $k$ -cube based on this pair is  $\tilde{\gamma} \leq (\max_{1 \leq i \leq k} \gamma_i + \delta) (\min_{1 \leq i \leq k} \gamma_i - \delta)^{-1}$ -isomorphic to the standard  $l_p$   $k$ -cube. By (ii)

$$\tilde{\gamma} \leq \frac{\frac{3}{2} a_p \gamma_1 + \delta}{\frac{2\gamma}{3a_p} - \delta} \leq \frac{9a_p^2 + 4\gamma}{\gamma} \quad \text{if } \delta \leq \frac{\gamma}{6a_p}.$$

Hence, there is a constant  $\tilde{C}(a_p, \alpha, \gamma)$  such that  $s = [\tilde{C}(a_p, \alpha, \gamma)(n^{1+2\beta} \log n)^{p/2}]$  satisfies all the conditions above and

$$k \sim \frac{n}{s} \sim C(a_p, \alpha, \gamma) n^{1-p/2-\beta p} (\log n)^{-p/2}. \quad \square$$

## 5. Relation between standard and metric type.

5.1. In this section, we show that the notion of metric type is essentially an extension to metric spaces of the standard notion of type on Banach spaces.

It is obvious that a Banach space which has metric type  $1 \leq p \leq 2$  has also type  $p$ . Here we show that the supremum of types of a Banach space cannot exceed the supremum of metric types (Theorem 5.2). A number of consequences follow from this conclusion. (See Corollaries 5.9–5.10.)

The methods used in the proof of Theorem 5.2 are based on [P2].

5.2. In this section, it is more convenient for us to consider the binary  $k$ -cube as  $C_2^k = \{-1, 1\}^k$ .

Let  $(X, \|\cdot\|)$  be an infinite-dimensional Banach space of metric type  $p$  with type constants  $\alpha_p(k; X)$ . Then for any  $\psi: C_2^k \rightarrow X$ ,  $\text{diag}_\psi(\varepsilon) = \|\psi(\varepsilon) - \psi(-\varepsilon)\|$  and an edge differing on the  $j$ th coordinate is

$$\text{edge}(\varepsilon, \varepsilon') = \|\psi(\varepsilon_1, \dots, \varepsilon_j, \dots, \varepsilon_k) - \psi(\varepsilon_1, \dots, -\varepsilon_j, \dots, \varepsilon_k)\| \stackrel{\text{def}}{=} \|\Delta_j \psi(\varepsilon)\|.$$

Hence, the metric type condition is

(5.2.1)

$$\left\{ \int_{C_2^k} \|\psi(\varepsilon) - \psi(-\varepsilon)\|^2 d\varepsilon \right\}^{1/2} \leq \alpha_p(k; X) k^{1/p-1/2} \left\{ \sum_{j=1}^k \int_{C_2^k} \|\Delta_j \psi(\varepsilon)\|^2 d\varepsilon \right\}^{1/2}.$$

**THEOREM.** *Let  $p_X$  be the supremum of types of  $X$  and suppose  $p_X > 1$ . Then  $\lim_{k \rightarrow \infty} \alpha_p(k; X) = 0$  whenever  $p < p_X$ .*

5.3. **PROOF.** Let  $\Omega_j$  ( $j = 1, \dots, k$ ) denote arbitrary probability spaces (not necessarily  $\{-1, 1\}$ ). For any  $k$  and  $\psi: \Omega = \Omega_1 \times \dots \times \Omega_k \rightarrow X$  define  $\beta_p(k)$  as the smallest  $\beta > 0$  fulfilling

$$\|\psi - E[\psi]\|_2 \leq \beta k^{1/p-1/2} \left\{ \sum_{j=1}^k \|\psi - E_j[\psi]\|_2^2 \right\}^{1/2}.$$

( $E_j$  is the conditional expectation with respect to the  $\sigma$ -algebra generated by  $\{\omega_i\}_{i \neq j}$ ; namely, the averaging is on the  $j$ th coordinate.)

Since  $\psi(\varepsilon) - E_j[\psi(\varepsilon)] = \frac{1}{2} \Delta_j[\psi(\varepsilon)]$  we get

$$\begin{aligned} \left( \int \|\psi(\varepsilon) - \psi(-\varepsilon)\|^2 d\varepsilon \right)^{1/2} &\leq 2 \|\psi - E[\psi]\|_2 \\ &\leq 2\beta k^{1/p-1/2} \left\{ \sum_{j=1}^k \int \|\Delta_j \psi(\varepsilon)\|^2 d\varepsilon \right\}^{1/2}. \end{aligned}$$

Hence,  $\beta_p(k) \geq \frac{1}{2} \alpha_p(k; X)$ . By a method similar to the one used in 2.3 and using the fact that  $\Omega_j$  were arbitrary we get the submultiplicativity property  $\beta_p(k \cdot l) \leq \beta_p(k) \beta_p(l)$ .

Now, fix  $p < p_X$  and assume  $\overline{\lim} \alpha_p(k; X) > 0$ . Then  $\overline{\lim} \beta_p(k) > 0$  and by submultiplicativity  $\beta_p(k) \geq 1$  for each  $k$ .

5.4. Let  $\tau > 0$  be a (small) positive number. There are arbitrary large integers  $k$  satisfying

$$(5.4.1) \quad \beta_p(k_1) \leq (1 + \tau) \beta_p(k) \quad \text{if } k/2 < k_1 \leq k.$$

Suppose this statement is false. Fix some  $\bar{k}$  and a large  $k$ . Negation of (5.4.1) gives a decreasing sequence  $k > k_1 > k_2 > \dots > k_j > k_{j+1} > \dots > \bar{k}$  such that  $\beta_p(k_{j+1}) > (1 + \tau) \beta_p(k_j)$  and  $k_{j+1} > k_j/2$ . This implies  $\beta_p(k) \leq (1 + \tau)^{-\log k \bar{k}^{-1}} \beta_p(\bar{k})$  which is a contradiction.



Now fix  $k$  satisfying (5.4.1) and choose  $\psi: \Omega_1 \times \dots \times \Omega_k \rightarrow X$  with

$$\|\psi - E[\psi]\|_2 \sim \beta_p(k) k^{1/p-1/2} \left\{ \sum_{j=1}^k \|\psi - E_j[\psi]\|_2^2 \right\}^{1/2}.$$

5.5. Define  $L_2^0(\Omega_i, R)$  to be all the  $L_2$ -functions  $\alpha: \Omega_i \rightarrow R$ , which are orthogonal to the constant functions (i.e.  $\int \alpha = 0$ ). Any function  $f: \Omega = \Omega_1 \times \dots \times \Omega_k \rightarrow X$  has a generalized Walsh expansion  $f = f_0 + f_1 + \dots + f_k$ , where  $f_j = \sum_{|A|=j} w_A x_A$ ,  $A \subset \{1, \dots, k\}$ ,  $x_A \in X$ ,  $w_A \in \text{span}\{\prod_{i \in A} \alpha_i | \alpha_i \in L_2^0(\Omega_i, R)\}$ .

Since  $p_X > 1$ ,  $X$  is  $K$  convex. By [P2]  $S_t = \prod_{j=1}^k [E_j + e^{-t}(I - E_j)]$  defines a holomorphic semigroup on  $L_2(X)$  and, hence, there is a constant  $C_1 = C_1(X)$  such that

$$(5.5.1) \quad \|f_j\|_2 \leq C_1^j \|f\|_2 \quad \text{for each } j.$$

5.6. Fix  $\delta = \delta(C_1) > 0$ . Consider independent  $\{0, 1\}$ -valued random variables  $\xi_1(\nu), \dots, \xi_k(\nu)$  of mean  $\delta$  and define the operator

$$\pi(\nu) = \prod_{j=1}^k [E_j + \xi_j(\nu)(I - E_j)] = E_{\{j | \xi_j(\nu)=1\}}$$

(which means the averaging is performed over  $j$ 's satisfying  $\xi_j(\nu) = 0$ ).

Let  $\psi = \psi_0 + \psi_1 + \dots + \psi_k$  be the Walsh expansion of  $\psi$ . Then

$$\int \pi(\nu)[\psi] d\nu = \psi_0 + \delta\psi_1 + \dots + \delta^k\psi_k.$$

By (5.5.1) and the fact that  $E[\psi_i] = 0$  for  $i > 0$  we get

$$(5.6.1) \quad \begin{aligned} \delta\|\psi_1\|_2 &\geq \left\| \int \pi(\nu)[\psi] d\nu - E[\psi] \right\|_2 - \sum_{j=2}^k (C_1\delta)^j \|\psi - E[\psi]\|_2, \\ \delta\|\psi_1\|_2 &\geq \|\psi - E[\psi]\|_2 - \int \|\psi - E_{\{j | \xi_j(\nu)=1\}}[\psi]\|_2 d\nu \\ &\quad - o(\delta)\|\psi - E[\psi]\|_2, \end{aligned}$$

where  $\delta$  is chosen as  $\delta \leq C_1^{-2}$ .

5.7. We now estimate  $\int \|\psi - E_{\{j | \xi_j(\nu)=1\}}[\psi]\|_2 d\nu$ . Take some  $\nu$  and define  $A(\nu) = \{j | \xi_j(\nu) = 1\}$  and  $\bar{A}(\nu) = \{j | \xi_j(\nu) = 0\}$ . Consider  $\psi$  as a function of these variables  $\omega_j$  with  $j \in \bar{A}(\nu)$ . By definition of  $\beta_p(n)$  we may write

$$\begin{aligned} \|\psi - E_{A(\nu)}[\psi]\|_{L_2(\omega_j | j \in \bar{A}(\nu))}^2 &\leq \left[ \beta_p \left( k - \sum_{j=1}^k \xi_j(\nu) \right) \right]^2 \left[ k - \sum \xi_j(\nu) \right]^{2(1/p-1/2)} \\ &\quad \times \sum_{j \in \bar{A}(\nu)} \|\psi - E_j[\psi]\|_{L_2(\omega_j | j \in \bar{A}(\nu))}^2. \end{aligned}$$

Integrate, now, in  $\omega_j$  for  $j \in A(\nu)$ .

$$\begin{aligned} \|\psi - E_{A(\nu)}[\psi]\|_2 &\leq \beta_p \left( k - \sum \xi_j(\nu) \right) \left[ k - \sum \xi_j(\nu) \right]^{1/p-1/2} \\ &\quad \times \left\{ \sum_{j=1}^k \|\psi - E_{\hat{j}}[\psi]\|_2^2 \right\}^{1/2}. \end{aligned}$$

Proceed by integrating both sides of the last inequality in  $\nu$ . First estimate the right side:

$$\begin{aligned} &\int \beta_p \left( k - \sum \xi_j(\nu) \right) \left[ k - \sum \xi_j(\nu) \right]^{1/p-1/2} d\nu \\ &\leq \left[ \int \beta_p \left( k - \sum \xi_j(\nu) \right)^2 d\nu \right]^{1/2} \left\{ \int \left[ k - \sum \xi_j(\nu) \right]^{2(1/p-1/2)} d\nu \right\}^{1/2} \end{aligned}$$

by the Cauchy-Schwartz inequality. Moreover, since  $0 \leq 2(\frac{1}{p} - \frac{1}{2}) \leq 1$ , we get

$$\begin{aligned} \left\{ \int \left[ k - \sum \xi_j(\nu) \right]^{2(1/p-1/2)} d\nu \right\} &\leq \left\{ \int \left( k - \sum \xi_j(\nu) \right) d\nu \right\}^{2(1/p-1/2)} \\ &= [(1-\delta)k]^{2(1/p-1/2)} \end{aligned}$$

and

$$\begin{aligned} \int \left[ \beta_p \left( k - \sum \xi_j(\nu) \right) \right]^2 d\nu &= \sum_{i=0}^k [\beta_p(k-i)]^2 P\left(\sum \xi_j = i\right) \\ &\leq \sum_{i=0}^{k/2} [\beta_p(k-i)]^2 P\left(\sum \xi_j = i\right) + \left(\frac{k}{2}\right)^{2/q} P\left(\sum \xi_j > \frac{k}{2}\right) \end{aligned}$$

since  $\beta_p(k) \leq k^{1/q}$  ( $\frac{1}{q} + \frac{1}{p} = 1$ ) by definition of  $\beta$ . Using (5.4.1) and the fact that  $P(\sum \xi_j > \frac{k}{2})$  is a tail of a normal distribution we get

$$\int \left[ \beta_p \left( k - \sum \xi_j(\nu) \right) \right]^2 d\nu \leq (1+\tau)^2 [\beta_p(k)]^2 + k^{2/q} e^{-ck}$$

for some constant  $c = c(\delta)$  and provided  $\delta < \frac{1}{4}$ . Collecting estimations

$$\begin{aligned} &\int \|\psi - E_{A(\nu)}[\psi]\|_2 d\nu \\ &\leq (1+2\tau)(1-\delta)^{1/p-1/2} \beta_p(k) k^{1/p-1/2} \left\{ \sum_{j=1}^k \|\psi - E_{\hat{j}}[\psi]\|_2^2 \right\}^{1/2} \\ &\leq (1-c_p\delta) \beta_p(k) k^{1/p-1/2} \left\{ \sum_{j=1}^k \|\psi - E_{\hat{j}}[\psi]\|_2^2 \right\}^{1/2} \\ &\approx (1-c_p\delta) \|\psi - E[\psi]\|_2 \end{aligned}$$

for some constant  $c_p$  and for  $\tau$  and  $\psi$  chosen as in 5.4.

5.8. Substitute (5.7.1) and (5.6.1) to obtain

$$\delta \|\psi_1\|_2 \geq c_p \delta \|\psi - E[\psi]\|_2 - o(\delta) \|\psi - E[\psi]\|_2.$$

Hence

$$\begin{aligned} \left\| \sum_{j=1}^k (E_{\omega_j}[\psi] - E[\psi]) \right\|_2 &= \|\psi_1\|_2 \geq c_p \|\psi - E[\psi]\|_2 \\ (5.8.1) \quad &\approx c_p \beta_p(k) k^{1/p-1/2} \left\{ \sum_{j=1}^k \|\psi - E_j[\psi]\|_2^2 \right\}^{1/2} \\ &\geq c_p \beta_p(k) k^{1/p-1/2} \left\{ \sum_{j=1}^k \|E_{\omega_j}[\psi] - E[\psi]\|_2^2 \right\}^{1/2} \end{aligned}$$

where  $E_{\omega_j}[\psi] - E[\psi]$  are independent  $L_2(\Omega, X)$  functions. Choose  $p < p_1 < p_X$ . Then, using the (standard) type condition on  $L_2(X)$  (recall that  $L_2(X)$  has the same type constant as  $X$ ) we get

$$\begin{aligned} \left\| \sum_{j=1}^k (E_{\omega_j}[\psi] - E[\psi]) \right\|_2 &\leq 2 \int \left\| \sum_{j=1}^k \varepsilon_j (E_{\omega_j}[\psi] - E[\psi]) \right\|_2 d\varepsilon \\ &\leq 2T_{p_1}(X) \left( \sum_{j=1}^k \|E_{\omega_j}[\psi] - E[\psi]\|_2^{p_1} \right)^{1/p_1} \\ &\leq 2T_{p_1}(X) k^{1/p_1-1/2} \left( \sum_{j=1}^k \|E_{\omega_j}[\psi] - E[\psi]\|_2^2 \right)^{1/2} \end{aligned}$$

which contradicts (5.8.1) for large  $k$ .

5.9. COROLLARY. (i) A Banach space  $X$  contains Hamming cubes  $C_2^n$  uniformly iff it contains  $l_1^n$ 's uniformly.

(ii) If a Banach space  $X$  contains uniformly  $l_p$   $n$ -cubes for some  $1 < p < 2$  then it contains  $l_p^n$ 's uniformly.

5.10. COROLLARY. If a Banach space  $X$  is of (standard) type  $p$  then for any  $p' < p$ , there is a constant  $\alpha_{p'}(X)$  such that for any  $n$  and any map  $\psi: C_2^n \rightarrow X$ :

$$\left\{ \sum_{\varepsilon \in C_2^n} \|\psi(\varepsilon) - \psi(-\varepsilon)\|^2 \right\}^{1/2} \leq \alpha_{p'}(X) n^{1/p'-1/2} \left\{ \sum_{j=1}^n \sum_{\varepsilon \in C_2^n} \|\Delta_j \psi(\varepsilon)\|^2 \right\}^{1/2}.$$

REMARK. After this text was typed, we were informed that G. Pisier gave another proof of Theorem 5.2 His argument is substantially simpler. The proof given above has, however, the advantage of giving power-type estimations on the size, when passing from nonlinear to linear. (Some questions are still not fully clarified at this moment and deserve further investigation.)

## 6. Embeddings of $l_p$ -cubes.

6.1. In this section the problem of almost-isometric embedding of  $l_p$  cubes into  $l_s$  cubes for  $1 \leq s < p \leq \infty$  is considered.

We show that for any  $\varepsilon > 0$  and any  $s, p$  as above there is a constant  $\lambda(s, p, \varepsilon)$  such that the  $l_s^N$  cube contains an  $(1 + \varepsilon)$ -isomorphic copy of the  $l_p^k$  cube, where  $N \leq \lambda(s, p, \varepsilon)k^3$ . (We recall that in the linear theory  $l_p^k$  can be embedded into  $l_1^N$  for  $1 \leq p \leq 2$  only.) The proof uses results [J-S, J-L] obtained by probabilistic methods which prove only the existence of a "good" embedding.

In the second part of this section we show a constructive embedding of any  $l_p^k$  cube in a  $l_s^N$  cube; however, as often happens in such cases, the estimate on the dimensions is  $k \sim \log N$ .

6.2. We start with the following

REMARK. If there is an  $(1 + \varepsilon)$  embedding of the  $l_p^k$  cube into the  $l_1^N$  cube for some  $p > 1$ , then for any  $s, t \geq 1$  satisfying  $s = pt$  there is an  $(1 + \varepsilon)^{1/t}$  embedding of the  $l_s^k$  cube into the  $l_t^N$  cube.

PROOF. Let  $\{y_i\}_{i=0}^{2^k-1}$  be the points of an  $(1 + \varepsilon)$  embedding of the  $l_p$   $k$ -cube into the  $l_1$   $N$ -cube. Hence there is some constant  $M > 0$  such that

$$(6.2.1) \quad \sum_{j=1}^N |y_{i,j} - y_{k,j}|^{1+\varepsilon} \sim M h(i, k)^{1/p} \quad \text{for every } i, k = 0, 1, \dots, 2^k - 1.$$

Here  $y_i = (y_{i,j})_{j=1}^N$  and  $h(i, k)$  is the Hamming distance between  $i$  and  $k$  considered as binary vectors.

Since in the left side of (6.2.1) we are summing only 0's and 1's which remain unchanged by taking power  $t$ , we get the conclusion by taking power  $t^{-1}$  on both sides.  $\square$

6.3. PROPOSITION. Let  $\varepsilon > 0$  and suppose  $1 < p \leq 2$ . Then there exists a constant  $\lambda = \lambda(p, \varepsilon)$  such that for any positive integer  $k$  there is an integer  $N \leq \lambda(p, \varepsilon)k^{2+1/p}$  and the  $l_1$   $N$ -cube contains an  $(1 + \varepsilon)$ -isometric copy of the  $l_p$   $k$ -cube.

PROOF. Let  $(C_2^k, \rho_p)$  be the  $l_p^k$ -cube. By the [J-S] result there is a constant  $\lambda'(p, \varepsilon)$  such that  $(C_2^k, \rho_p)$  can be  $(1 + \frac{\varepsilon}{2})$ -isometrically embedded into  $l_1^n$ , where  $n = \lambda'(p, \varepsilon)k$ . Hence, it can be  $(1 + \varepsilon)$ -isometrically embedded into  $C_t^n = \{(z_i)_{i=1}^n | z_i \in \{0, 1, \dots, t-1\}\}$ ,  $t = [\frac{2}{\varepsilon}k^{1/p}n] + 1$ , equipped with the  $l_1$ -metric. This set is isometric to a subset of  $C_2^N$   $l_1$ -cube, where  $N = (t-1)n = \lambda(p, \varepsilon)k^{2+1/p}$  for a constant  $\lambda(p, \varepsilon)$  independent of  $k$ . (For the infinite-dimensional variant of the last consideration see [Ox, p. 44].)

6.4. A result similar to Proposition 6.3 can be obtained for  $2 < p < \infty$  also. Indeed, let  $2^j < p \leq 2^{j+1}$  for some integer  $j \geq 1$ . By applying Remark 6.2 and Proposition 6.3  $j$  times we can  $(1 + \varepsilon)$ -embed the  $l_p^k$ -cube into the  $l_2^N$  cube, with  $N \leq \lambda(p, \varepsilon)k^{3^j}$ .

By [J-L]  $2^k$  points in  $l_2$  can be  $(1 + \varepsilon)$ -embedded into  $l_2^n$  with  $n$  proportional to  $k$ . As in Proposition 6.3 the points can now be  $(1 + \varepsilon)$ -embedded into  $l_1$   $N$ -cube with  $N$  proportional to  $k^{2.5}$ .

6.5. If  $p = \infty$ , then the  $2^k$  points of the  $l_\infty$   $k$ -cube are isometrically embedded into the  $l_2$   $n$ -cube with  $n = 2^k$ , by mapping them onto the unit binary vectors

of the  $l_2$   $n$ -cube. Now, as above, we can  $(1 + \varepsilon)$ -embed the former set into the  $l_1$   $N$ -cube with  $N$  proportional to  $k^2$ .

6.6. Summarizing 6.3–6.5 and applying Remark 6.2 we obtain

**THEOREM.** *Let  $\varepsilon > 0$  and  $1 \leq s < p \leq \infty$ . Then there is a constant  $\lambda(p, s, \varepsilon)$  such that for any integer  $k$  there is an integer  $N \leq \lambda(p, s, \varepsilon)k^3$  such that the  $l_p$   $k$ -cube can be  $(1 + \varepsilon)$ -embedded into the  $l_s$   $N$ -cube.*

6.7. Since the embedding in Theorem 6.6 is nonconstructive we add an example of a constructive embedding, which gives, however, a bad estimate on  $N$ . By Remark 6.2 it is enough to show an embedding of the  $l_p$   $k$ -cube for some  $p > 1$  into the  $l_1$   $N$ -cube.

6.8. Consider the  $2^k \times 2^k$  zero-one Walsh matrix  $Z = (z_{ij}) = (\langle i, j \rangle)$ , where  $i, j = 0, 1, \dots, 2^k - 1$  and  $\langle i, j \rangle$  is the modulo 2 scalar product of the binary representations (as  $k$ -dimensional vectors) of the indices  $i, j$ .

The  $j$ th column of  $Z$  is repeated  $n_j$  times (the  $n_j$ 's are calculated later) and a new  $2^k \times N$  ( $N = \sum_{j=0}^{2^k-1} n_j$ ) matrix is obtained. Let  $y_i = (y_{i,j})_{j=1}^N$ ,  $i = 0, \dots, 2^k - 1$ , be the rows of this matrix.

Consider  $Y = \{y_i\}_{i=0}^{2^k-1}$  with the  $l_1$ -metric. If  $Y$  (in the natural order) was isometric to the  $l_p$   $k$ -cube, then there was a constant  $M > 0$  such that the following system of equations would hold:

$$(6.8.1) \quad \sum_{j=1}^N |y_{i_1,j} - y_{i_2,j}| = Mh(i_1, i_2)^{1/p} \quad \text{for every } i_1, i_2 = 0, \dots, 2^k - 1.$$

The system (6.8.1) can be obviously reduced to

$$(6.8.2) \quad \sum_{j=1}^N y_{i,j} = Mh(i, 0)^{1/p}, \quad i = 1, \dots, 2^k - 1,$$

or in our notations

$$(6.8.3) \quad \sum_{j=0}^{2^k-1} n_j z_{ij} = Mh(i, 0)^{1/p}, \quad i = 1, \dots, 2^k - 1.$$

Existence of a (positive) integer solution  $\{n_i\}_{i=0}^{2^k-1}$  to (6.8.3) would imply that  $Y$  is isometric to the  $l_p$   $k$ -cube. We will find a positive real-valued solution of (6.8.3), and by a small perturbation we will pass to a (positive) integer “ $\varepsilon$ -solution”. Since the perturbation of an equation will be kept less than  $\varepsilon M$ , we will get a  $(1 + \varepsilon)$  isomorphic set.

6.9. Let  $W = \{w_{ij}\}_{i,j=0}^{2^k-1} = \{(-1)^{\langle i,j \rangle}\}$  be the standard  $2^k \times 2^k$  Walsh matrix,  $J$  the matrix with all its entries 1 and define two vectors  $\bar{x} = (x_j)_{j=0}^{2^k-1} = (M^{-1}n_j)$ ,  $\bar{b} = (b_i)_{i=0}^{2^k-1} = (h(0, i)^{1/p})$ . Then (6.8.3) becomes

$$(6.9.1) \quad \frac{1}{2}(J - W)\bar{x} = \bar{b}.$$

Multiply both sides by  $W$ :

$$(6.9.2) \quad (WJ - 2^k I)\bar{x} = 2W\bar{b}.$$

Since the entries of the first row of  $WJ$  equal  $2^k$  and all other entries are 0 we obtain

$$(6.9.3) \quad \begin{cases} 2^k \sum_{j=1}^{2^k-1} x_j = 2 \sum_{i=0}^{2^k-1} b_i, \\ -2^k x_i = 2 \sum_{j=0}^{2^k-1} w_{ij} b_j, \quad i = 1, \dots, 2^k - 1. \end{cases}$$

The system is consistent and its solution is

$$(6.9.4) \quad x_i = -2^{-k+1} \sum_{j=0}^{2^k-1} w_{ij} b_j = -2^{-k+1} \sum_{l=1}^k a_{i,l} l^{1/p}, \quad i = 1, \dots, 2^k - 1,$$

where  $a_{il} = \sum_j w_{ij}$  and summation is extended over those indices  $j$  having Hamming weight  $l$  in their binary representation. Denote by  $|j|$  the Hamming weight of  $j$  and define

$$P_{i,l} = \{0 \leq j \leq 2^k - 1 \mid |j| = l, \langle j, i \rangle = 0\}$$

and

$$N_{i,l} = \{0 \leq j \leq 2^k - 1 \mid |j| = l, \langle j, i \rangle = 1\}.$$

Then  $a_{i,l} = |P_{i,l}| - |N_{i,l}|$  and it is easily verified that  $a_{i,l}$  are the coefficients of the following generating function

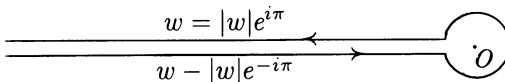
$$(6.9.5) \quad (1 - \xi)^{|i|} (1 + \xi)^{k-|i|} = \sum_l a_{i,l} \xi^l.$$

6.10. We have to show that  $\{x_i\}$  (see (6.9.4)) are positive. The following calculation is due to M. Dubiner.

Consider the generating function (6.9.5)  $f(\xi) = (1 - \xi)^{|i|} (1 + \xi)^{k-|i|} = \sum_l a_{i,l} \xi^l$  and substitute  $z = \ln \xi$ . Then

$$f(e^z) = (1 - e^z)^{|i|} (1 + e^z)^{k-|i|} = \sum_l a_{i,l} e^{lz}.$$

It is well known (see for example [F-S, p. 350]) that

$$(6.10.1) \quad \int_C e^w w^{-(\alpha+1)} dw = 2\pi i / \Gamma(\alpha + 1),$$


where  $C$  is the contour around the branch point 0 with the branch cut along the negative real axis.

Hence, substituting  $w = zl$  we obtain

$$(6.10.2) \quad \int_C e^{lz} z^{-(\alpha+1)} dz = \int_C e^w \left(\frac{w}{l}\right)^{-(\alpha+1)} l^{-1} dw = l^\alpha 2\pi i / \Gamma(\alpha + 1).$$

Hence in order to evaluate  $\sum_l l^\alpha a_{i,l}$  it is enough to evaluate

$$\begin{aligned} I &= (2\pi i)^{-1} \Gamma(\alpha + 1) \int_C \frac{(1 - e^z)^{|i|} (1 + e^z)^{k-|i|}}{z^{\alpha+1}} dz \\ &= \frac{\Gamma(\alpha + 1)}{2\pi i} [e^{-i\pi\alpha} - e^{i\pi\alpha}] \int_0^\infty \frac{(1 - e^{-u})^{|i|} (1 + e^{-u})^{k-|i|}}{u^{\alpha+1}} du \\ &= -\frac{\Gamma(\alpha + 1)}{\pi} \sin \pi\alpha \int_0^\infty \frac{(1 - e^{-u})^{|i|} (1 + e^{-u})^{k-|i|}}{u^{\alpha+1}} du. \end{aligned}$$

Since  $0 < \alpha = \frac{1}{p} < 1$ ,  $I$  is always negative and it is easy to show that

$$(6.10.3) \quad I < -\Gamma(\alpha + 1) \frac{\sin \pi \alpha}{\pi} \cdot \frac{1}{2e^2} (\ln k/2)^{-(\alpha+1)}.$$

Hence for every  $i = 1, \dots, 2^k - 1$ ,  $x_i > C(p)2^{-k+1}(\ln k)^{-(1/p+1)}$  for a constant  $C(p) > 0$ .

6.11. Take  $M > \max\{\varepsilon^{-1}2^{k-1}, C(p)^{-1}2^{k-1}(\ln k)^{1+1/p}\}$  and define  $a_i = Mx_i$ . By (6.9.4) and (6.10.3)  $1 < a_i < 2Mk$ . Define

$$n_i = \begin{cases} [a_i], & |a_i - [a_i]| \leq 1/2, \\ [a_i] + 1, & \text{otherwise.} \end{cases}$$

Clearly,  $1 \leq n_i \leq 2Mk + 1$ . Substituting  $n_i$  in (6.8.3) every equation is perturbed at most by  $2^{k-1} < \varepsilon M$ . Hence  $Y \subset (C_2^N, h)$  with  $N = \sum n_i$  is  $(1 + \varepsilon)$ -isomorphic to the  $l_p$   $k$ -cube and  $N \leq k2^{2k} \max(\frac{1}{\varepsilon}, C(p)^{-1}(\ln k)^{1/p+1})$  or  $k \sim \log N$ .

## REFERENCES

- [A-M] D. Amir and V. D. Milman, *Unconditional and symmetric sets in  $n$ -dimensional normed spaces*, Israel J. Math. **37** (1980), 3–20.
- [E] P. Enflo, *On the nonexistence of uniform homeomorphisms between  $L_p$ -spaces*, Ark. Mat. **8** (1969), 103–105.
- [F-S] B. A. Fuchs and B. V. Shabat, *Functions of a complex variable and some of their applications*, Pergamon Press, Oxford, 1961.
- [G] M. Gromov, *Filling Riemannian manifolds*, J. Differential Geom. **18** (1983), 115–117.
- [H] L. Harper, *Optimal numberings and isoperimetric problems on graphs*, J. Combinatorial Theory **1** (1966), 385–393.
- [J-L] W. B. Johnson and J. Lindenstrauss, *Extensions of Lipschitz mappings into a Hilbert space*, Contemp. Math. **26** (1984), 189–206.
- [J-S] W. B. Johnson and G. Schechtman, *Embedding  $l_p^m$  into  $l_1^n$* , Acta Math. **149** (1982), 71–85.
- [Mar-P] M. B. Marcus and G. Pisier, *Characterizations of almost surely continuous  $p$ -stable random Fourier series and strongly stationary processes* (to appear).
- [M] B. Maurey, *Espaces des Banach: Construction de suites symétriques*, C. R. Acad. Sci. Paris Sér. A **288** (1979), 679–681.
- [M-P] B. Maurey et G. Pisier, *Séries de v.a. vectorielles indépendantes et propriétés géométriques des espaces de Banach*, Studia Math. **58** (1976), 45–90.
- [M-S] V. D. Milman and G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Lecture Notes in Math., Springer, Berlin and New York (to appear).
- [Ox] J. C. Oxtoby, *Measure and category*, Springer-Verlag, New York, 1971.
- [P1] G. Pisier, *Sur les espaces de Banach qui ne contiennent pas uniformément de  $l_1^n$* , C. R. Acad. Sci. Paris Sér. A **277** (1973), 991–994.
- [P2] G. Pisier, *Holomorphic semi-groups and the geometry of Banach spaces*, Ann. of Math. (2) **115** (1982), 375–392.

DEPARTMENT OF MATHEMATICS, VRIJE UNIVERSITEIT BRUSSEL, PLEINLAAN 2-F7, 1050 BRUSSELS, BELGIUM

SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, RAMAT AVIV, TEL AVIV (Current address of V. Milman)

*Current address* (H. Wolfson): Robotics Laboratory, Computer Science Division, Courant Institute of Mathematical Sciences, New York University, New York, New York 10012