The Ribe Programme

Keith Ball

December 6, 2011

Introduction

In 1976 Ribe [R] proved that uniformly homeomorphic Banach spaces have uniformly *linearly* isomorphic finite-dimensional subspaces: that if there is a uniformly continuous homeomorphism between two Banach spaces, with uniformly continuous inverse, then every finite-dimensional subspace of one space, is linearly isomorphic to a subspace of the other, with an implied constant of isomorphism independent of the dimension. This remarkable rigidity theorem guarantees that finite-dimensional properties are determined up to isomorphism by the metric structure of the space. Thus in principle, any property of a space that depends upon finite collections of points has an equivalent formulation which makes no reference to the linear structure of the space and involves only the distances between points.

In view of this, Bourgain [Bo1] proposed an ambitious programme which came to be known as "The Ribe Programme": to find explicit metric descriptions of the most important invariants of normed linear spaces. He himself kick-started the programme by characterising superreflexivity. A Banach space X is called *superreflexive* if every space whose finite-dimensional spaces embed uniformly well into X, must automatically be reflexive. Bourgain showed this holds precisely if the space does not contain copies of arbitrarily large binary trees. Needless to say the general aim of the Ribe programme is not merely to find metric equivalents of linear properties but to transfer the subtle and well-developed theory of normed spaces to the nonlinear setting, and in this broader sense the programme had been anticipated in a prescient paper of Johnson and Lindenstrauss, [JL].

Within a decade or two the Ribe programme acquired an importance that would have been hard to predict at the outset, as the insights provided by the programme became powerful tools in the theory of algorithms and more recently in geometric group theory. Data sets often come equipped with a metric structure but rarely a linear one. For a number of central algorithmic problems the most effective procedure is to embed the data-set into a familiar linear geometry, without distorting it too much, and then use the linear structure to analyse the embedded copy of the data. Plainly, data that fail to possess a metric property that holds in the linear geometry cannot be embedded: so it is essential to understand what these properties are and where they appear. The study of groups as metric spaces and in particular their embeddability into simpler geometric structures has become a subject of intense interest in the last 5-10 years because of its connection with several famous problems such as the Novikov conjecture.

In this article I will describe what is currently by far the most successful family of achievements within the Ribe programme: the development of metric equivalents of type and cotype. In order to make the article reasonably self-contained I shall discuss the linear invariants as well as the non-linear ones and describe three classical principles (due in various combinations to Kwapien, Maurey and Pisier) that provided the inspiration for the non-linear development and serve as test cases to check that the non-linear invariants are useful. The non-linear theory proceeded in two (or even three) distinct phases that were quite widely separated in time. In the first phase Bourgain and others introduced metric type and initiated the theory of non-linear embedding. Formally the theory of embedding does not form part of the Ribe programme but it goes hand in hand with it. It is hard to imagine that the recent subtle development of the non-linear Dvoretzky Theorem by Mendel and Naor [MN2] would have emerged without something like the Ribe programme. Indeed, it still seems extraordinary that there is a subtle structure theory for objects as diverse as general metric spaces, mirroring the structure theory for normed spaces. In the second phase, or second half of the first phase, the present author introduced the Markov type and cotype properties for the specific purpose of studying Lipschitz extensions. The article had an effect opposite to the one that mathematicians always hope for: it seemed to halt the programme instead of encouraging further development. This was probably because we were unable to prove that the Markov type property held in any space other than Hilbert space. Ten years later this embarrassing open problem was adopted by Naor who solved it in collaboration with Peres, Schramm and Sheffield and who then produced a series of deep articles with Mendel that form the current phase of the development. They introduced a metric form of cotype, showed that for linear spaces it agreed with linear cotype and proved a non-linear analogue of the Maurey-Pisier Theorem. This is the most technically difficult part of the Ribe programme to date.

The story begins with the linear invariants.

1 Linear (or Rademacher) type and cotype

The type and cotype invariants were introduced simultaneously by Maurey in the study of factorisation of linear maps and by Hoffmann-Jørgensen for the study of vector-valued central limit theorems. A detailed history of how these ideas grew out of earlier work of James and others can be found in the article of Maurey, [M2]. A normed space X has type p if there is a constant T so that for every sequence of vectors $x_1, \ldots, x_n$ in X

$$\mathrm{Ave}\| \pm x_1 \pm x_2 \pm \cdots \pm x_n\|^p \leq T^p \sum_1^n \|x_i\|^p$$

where the average is taken over all choices of sign in the vector sum. It has cotype q with constant C if

$$\sum_1^n \|x_i\|^q \leq C^q \mathrm{Ave}\| \pm x_1 \pm x_2 \pm \cdots \pm x_n\|^q$$

for every such sequence. In Hilbert space, both statements hold with equality if $p = 2$ or $q = 2$ and with constants $T = 1$ or $C = 1$. (The identity that results is the parallelogram identity.) Consideration merely of the real line shows that if the respective properties are to hold in any space then necessarily $p \leq 2$ and $q \geq 2$. Using the Khintchine inequality for the L_p norms of sign averages it is straightforward to show that if $1 \leq p \leq 2$ the space L_p has type p and cotype 2, while if $2 \leq q < \infty$, L_q has type 2 and cotype q. Every space has type 1 and (with the obvious convention) cotype ∞. Thus there are spaces other than Hilbert space that possess either the optimal type (type 2) or optimal cotype (cotype 2). For this reason these properties have a special place in the theory. However, a gorgeous result of Kwapien [K] from the early days of the theory shows that only Hilbert space can possess both type 2 *and* cotype 2.

2 Metric type and a non-linear ℓ_1 theorem

The first metric analogue of type appeared in the work of Enflo [E1] (actually before the linear version was introduced). He (in effect) asked: for which spaces X is it true that there is a constant T so that for every n and every embedding of the corners of the cube $\{-1,1\}^n$ into X, the average squared length of the cube's 2^n diagonals is at most nT^2 times the average squared length of the cube's $n2^{n-1}$ edges? Such a space clearly has type 2 (with the same constant) since Enflo's definition allows folded (non-linear) embeddings of the cube, but asks for the same inequality. Enflo observed that Hilbert space *does* have this property: that folding the cube shortens the diagonals (on average) more than the edges. But the question already illustrates some of the difficulties that one encounters in the Ribe programme. It is still unknown whether Enflo's property holds for linear spaces with type 2 and it is clear that we cannot hope to find a metric version of cotype just by allowing folded cubes, since we can fold the diagonals to nothing while keeping the edges of fixed length.

Following Bourgain's enunciation of the Ribe programme, Bourgain, Milman and Wolfson [BMW] wrote an article in which they chose a modification of Enflo's property as their definition of metric type (for $1 < p \leq 2$) and showed that a space with linear type p has metric type r for all $r < p$. They also proved a metric analogue of Pisier's ℓ_1 theorem (see [P1]) which states that the finite-dimensional L_1-spaces are the only possible obstruction to type:

Theorem 1 (Pisier). *If a normed space X fails to have type p for every $p > 1$ (the space has* no *non-trivial type) then there is a constant C so that for every n, X has a subspace Y which is C-isomorphic to the n-dimensional L_1-space, ℓ_1^n: in other words there is a linear isomorphism $T : Y \to \ell_1^n$ with $\|T\|\,\|T^{-1}\| \leq C$.*

The theorem of [BMW] guarantees that a *metric* space which has no non-trivial *metric* type contains uniformly Lipschitz-equivalent copies of the discrete cube $\{-1,1\}^n$ equipped with the metric it inherits from ℓ_1^n; the *Hamming cube* as it is usually known.

Theorem 2 (Bourgain, Milman, Wolfson). *If a metric space X fails to have type p for every $p > 1$ (the space has* no *non-trivial metric type) then there is a constant C so that for every n, X has a subset which is C-lipschitz equivalent to the metric space $\{-1,1\}^n$ with the Hamming metric it inherits from ℓ_1^n.*

As in Pisier's paper, the crucial point is a submultiplicativity property for type constants in terms of the number of vectors. This means that if the space contains a cube of large dimension which looks a bit like a Hamming cube it must contain cubes of smaller dimension which look very like Hamming cubes.

At about the same time, Bourgain [Bo2] proved that every n-point metric space can be embedded into Hilbert space with the distances between points of the space being distorted by at most a constant multiple of $\log n$. It had been known since the work of Fritz John [J] that two n-dimensional normed spaces are linearly isomorphic with an isomorphism constant at most $\sqrt{n}$. Bourgain's Theorem was clearly inspired by this fact about linear spaces together with a view which emerged at the time that the number of points of a metric space would play a role something like the exponential of the dimension of a normed space. This view was prompted by a fact related to sphere-packing that had by then become a standard tool in the linear theory: the unit ball of an n-dimensional normed space contains an ε-net with no more than $(5/\varepsilon)^n$ elements.

3 Smoothness, convexity and martingale type

One of the oldest principles in functional analysis states that a space is reflexive if it is uniformly convex: for any two unit vectors u and v with $\|u-v\| \geq \varepsilon$, the average $(u+v)/2$ is significantly inside the ball:

$$\left\|\frac{u+v}{2}\right\| \leq 1-\delta(\varepsilon) \tag{1}$$

where $\delta(\varepsilon) > 0$ depends upon ε but not on the particular u and v. Classical spaces like L_p spaces possess strong quantitative forms of this property and the corresponding dual property, uniform smoothness.

A normed space is said to be p-smooth with constant K if for any two points x and y in the space,

$$\frac{\|x+y\|^p + \|x-y\|^p}{2} \leq \|x\|^p + K^p\|y\|^p$$

or q-convex with constant K if

$$\|x\|^q + \frac{1}{K^q}\|y\|^q \leq \frac{\|x+y\|^q + \|x-y\|^q}{2}. \tag{2}$$

If p or q is 2, these properties can be thought of as intrinsic (in the sense that they use the norm of the space) ways to measure the curvature of the unit ball: the first says that the curvature is not too tight, the second that the ball is not too flat. Euclidean spaces possess both properties with $K = 1$. For $2 < q < \infty$ the space L_q is 2-smooth with constant $K = \sqrt{q-1}$ while for $1 < p < 2$ the space L_p is

2-convex with constant $K = \frac{1}{\sqrt{p-1}}$. An overview of these properties in the L_p and non-commutative L_p spaces can be found in the article [BCL].

To see how these properties are related to type and cotype observe that if $x_1, x_2, \ldots, x_n$ are points in a p-smooth space then

$$\text{Ave}\| \pm x_1 \pm x_2 \pm \cdots \pm x_{n-1} \pm x_n\|^p \leq \text{Ave}\| \pm x_1 \pm x_2 \pm \cdots \pm x_{n-1}\|^p + K^p\|x_n\|^p$$

and we can peel off the individual vectors one by one at the cost of a factor K^p in front of each term $\|x_i\|^p$. So the space has type p with the same constant K. In a similar way, q-convexity implies cotype q. These quantitative forms of smoothness and convexity can be thought of as precise or "local" forms of type and cotype but they are quite a bit stronger. Moreover, for example, L_1 has cotype 2 but is not uniformly convex.

Since uniformly convex spaces are reflexive and uniform convexity depends only upon pairs of points, such spaces are automatically superreflexive. In [E2] Enflo had shown that a space is superreflexive if and only if it can be equipped with an equivalent norm that is uniformly convex. A remarkable (in my view absolutely shocking) result of Pisier [P2] showed that in this case there will necessarily be some finite q for which the space has an equivalent q-convex norm: it isn't possible that the optimal function δ in (1) is of the form $e^{1/\varepsilon}$ for *all* equivalent norms.

Pisier's argument goes via inequalities for martingales in the space in question. A sequence $(M_k)_{k=0}^n$ of integrable random vectors in a normed space is a martingale if for each k, the increment $M_k - M_{k-1}$ has expectation 0, conditional on the value of M_{k-1}. The simplest examples can be built from independent choices of sign as follows. Suppose $U_1, U_2, \ldots, U_n$ are IID random variables each taking the values 1 and -1 each with probability 1/2. Let $x_1, x_2, \ldots, x_n$ be random vectors (on the same probability space) where for each k the value of x_k depends only upon the earlier U_i namely: $U_1, U_2, \ldots, U_{k-1}$. Then the sums

$$M_k = \sum_1^k x_i U_i$$

form a martingale. This sequence builds a binary tree in the space as shown in Figure 1. The leaves of the tree form a non-linear embedding of the cube which has some features in common with linear embeddings.

The easy part of Pisier's argument depends upon the fact that in a p-smooth space the type inequality can be strengthened to hold for all martingales (and similarly for the cotype inequality in a q-convex space).

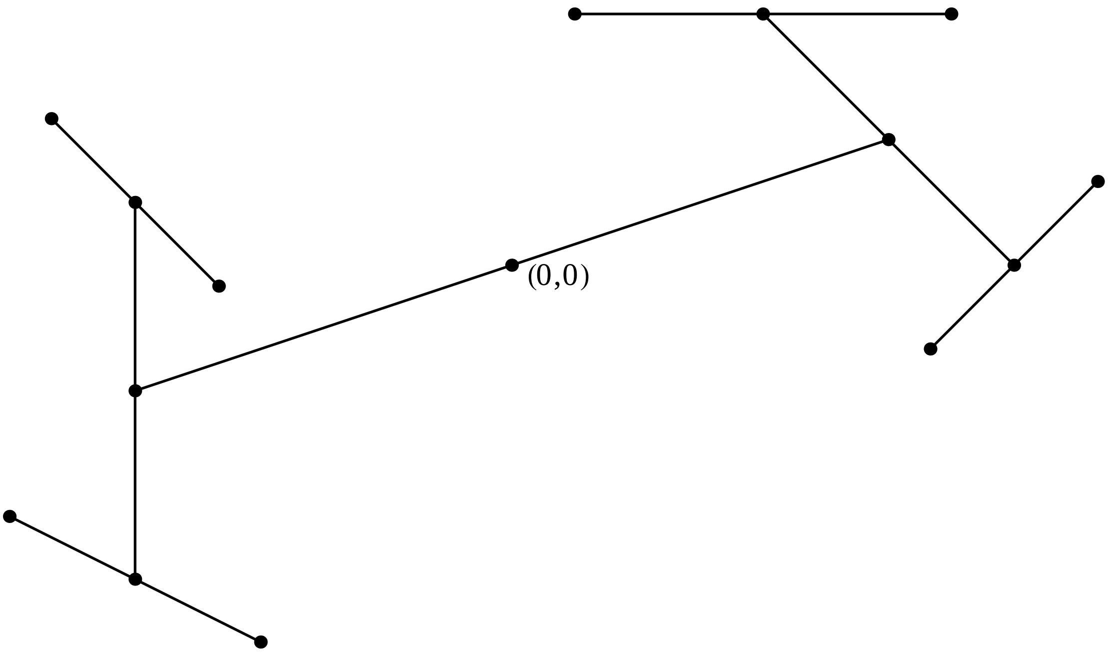

Figure 1: A martingale in the plane

Lemma 3. *If X is a p-smooth space with constant K and $(M_k)_{k=1}^m$ is a martingale in X then*

$$E\|M_m - M_0\|^p \leq f(K)^p \sum_{k=1}^m E\|M_k - M_{k-1}\|^p$$

where the function $f(K)$ depends only on K and not on the space (or the martingale).

For a proof of this lemma see [B1].

4 Markov type and cotype and extensions of Lipschitz maps

Maurey's extension theorem (which generalises Kwapien's Theorem referred to earlier) [M1], states the following.

Theorem 4 (Maurey)**.** *If A is a subspace of a normed space X with type 2, Y is a normed space with cotype 2 and $S : A \to Y$ is a bounded linear map, then there is an bounded linear map $\tilde{S} : X \to Y$ which extends S: so $\tilde{S}(a) = S(a)$ for each $a \in A$. (Moreover the least possible norm of such an $\tilde{S}$ can be estimated in terms of the norm of S and the type and cotype constants of the respective spaces.)*

In their article, [JL] Johnson and Lindenstrauss studied extensions of Lipschitz maps from metric spaces into Hilbert space and proved the celebrated dimension-reduction lemma stating that any n points in Euclidean space can be embedded in an Euclidean space of dimension only about $\log n$, with very little distortion of the distances between points. This lemma has been used repeatedly in the theory of algorithms since data embedded in a space of low dimension are much easier to search than those in a space of higher dimension. In the same article, the authors asked whether an analogue of Maurey's extension theorem holds for Lipschitz maps initially defined on arbitrary subsets of X (as opposed to subspaces), at least if X is L_q with $2 \leq q < \infty$ and Y is L_p with $1 \leq p \leq 2$? The purist might object that if one is going to study *Lipschitz* maps one should be studying them on general metric spaces rather than on linear spaces, but one may interpret the question of Johnson and Lindenstrauss as asking: can you find appropriate metric versions of type and cotype 2, that imply an analogue of Maurey's extension theorem, and that are possessed at least by the correct L_p spaces?

At the time this was a rather bold question. The only real evidence for their suggestion was an old theorem of Kirszbraun which guarantees that a Lipschitz map from a subset of one Hilbert space H_1 into another H_2, can be extended to the whole of H_1. The linear analogue of this statement is an immediate consequence of the existence of orthogonal projections in Hilbert space. The proof of Kirszbraun's theorem is completely unrelated to the linear case and the extension is carried out one point at a time, so it is essential that there is no increase in the Lipschitz constant at each step. In contrast it is easy to see that no theorem can hold for other L_p spaces without such an increase in Lipschitz constant, so a one point extension is useless.

In the context of the extension problem it was easy to see why metric cotype was more difficult to invent than metric type. If you want to extend a map into Y then you must have some hypothesis that prevents Y from being a collection of isolated points: you need that any collection of points in Y are at least joined by some sort of tree-like structure onto which you can map the new points. So the cotype property we are after must involve some existential assertion on points, rather than merely an inequality on distances. One solution to the problem was introduced in the author's paper [B1] where the properties of Markov type and cotype were defined. A metric space X has Markov type p with constant K if every time-reversible stationary Markov chain $(M_k)_{k=1}^m$ in X, running in steady state, satisfies what amounts to the same inequality as in Lemma 3:

$$\mathrm{E}\, d(M_m, M_0)^p \leq K^p m \mathrm{E}\, d(M_1, M_0)^p.$$

Thus, a space has Markov type 2 if time-reversible Markov chains only wander about $\sqrt{m}$ times as far in m steps as they do in one step: if they experience the same sort of cancellation as appears in the central limit theorem.

The Markov type p condition can be rephrased as follows. A space X has Markov type p if there is a constant K so that for every symmetric stochastic matrix (a_{ij}), every sequence (x_i) of points in X, and each natural number m,

$$\sum (A^m)_{ij} d(x_i, x_j)^p \leq K^p m \sum a_{ij} d(x_i, x_j)^p.$$

(The matrix A is the transition matrix for the Markov chain.) To check this condition for Hilbert space (with $p = 2$ and $K = 1$) we need only check that for a sequence of real numbers (x_i)

$$\sum (A^m)_{ij} (x_i - x_j)^2 \leq m \sum a_{ij} (x_i - x_j)^2.$$

This is equivalent to the statement that the matrix $(m-1)I - mA + A^m$ is positive semi-definite, where I is the identity matrix of the correct dimension. But the matrix A has all its eigenvalues in the interval $[-1, 1]$ and the function $\lambda \mapsto (m-1) - m\lambda + \lambda^m$ is non-negative on this interval.

The Markov *cotype* condition is more complicated because of the existential assertion referred to above but in the linear setting it can be replaced by a simpler condition involving the Green's matrices for the solution of difference equations based on stochastic matrices. As was mentioned earlier, at the time these properties were introduced it was not known whether the Markov type 2 property held in any normed space other than Hilbert space. This problem was solved some 10 years later in [NPSS]: 2-smooth spaces have Markov type 2. Their result combined with those from [B1] gives the following non-linear analogue of Maurey's Extension Theorem.

Theorem 5 (Ball, Naor, Peres, Schramm, Sheffield)**.** *If $1 < p \leq 2 \leq q < \infty$, A is a subset of L_q and $S : A \to L_p$ is a Lipschitz map, then there is a Lipschitz map $\tilde{S} : L_q \to L_p$ which extends S: so $\tilde{S}(a) = S(a)$ for each $a \in A$.*

5 Metric cotype and the non-linear Maurey-Pisier Theorem

Although Markov cotype seems to be the correct non-linear analogue of cotype for the purpose of Lipschitz extensions it evidently does not fulfil the precise demand of the Ribe programme for a metric version of the linear condition. In their paper [MN1] Mendel and Naor finally found an appropriate metric analogue, some 40 years after Enflo came up with the analogue of type. Just as Pisier's ℓ_1 Theorem and Maurey's Extension Theorem provided a way to test the metric type and Markov type and cotype properties, the test problem for cotype is the Maurey-Pisier Theorem for ℓ_∞:

Theorem 6 (Maurey, Pisier)**.** *If a normed space X fails to have cotype q for every $q < \infty$ (the space has* no *non-trivial cotype) then there is a constant C so that for every n, X has a subspace Y which is C-isomorphic to the n-dimensional L_∞-space, ℓ_∞^n: in other words there is a linear isomorphism $T : Y \to \ell_\infty^n$ with $\|T\|\,\|T^{-1}\| \leq C$.*

In view of the fact that the Hamming cube is the only obstruction to metric type, one might at first guess that the discrete cube lying in ℓ_∞^n would be the analogous obstruction to non-trivial metric cotype. But a moment's thought shows that this is absurd since any two distinct points of this set are distance 2 apart: so the set can be embedded in Hilbert space. Mendel and Naor not only found the metric version of cotype and showed that for linear spaces it is equivalent to linear cotype, they also found the right obstructions. The two parts of the problem should be thought of as being closely linked: the cotype condition depends on the distances between points in sets (more complicated than cubes) which mirror the structure of the obstructions.

The sets in question are grids rather than cubes and the simplest way to describe the cotype property is by using inequalities for embeddings of the discrete torus $\mathbf{Z}_m^n$. A metric space X has cotype q with constant K if for every n there is an even integer m so that for every embedding $f : \mathbf{Z}_m^n \to X$

$$\sum_{j=1}^{n} \mathrm{E}\, d\left(f(x + m/2\, e_j), f(x)\right)^q \leq K^q m^q \mathrm{E}\, d\left(f(x+\varepsilon), f(x)\right)^q$$

where the averages are taken over all x in the torus and (in the second case) over all choices $\varepsilon \in \{-1, 0, 1\}^n$. The vectors e_j are the standard basis vectors so the expression on the left is similar to edges of the cube (as in linear cotype) but with jumps increased by a factor of $m/2$. The expression on the right is built from the diagonals of the cubes that belong to the grid. The main theorem in [MN1] is the following:

Theorem 7 (Mendel, Naor). *A normed space X has metric cotype q if an only if it has linear cotype q.*

Needless to say, the really difficult part of the theorem is to show that linear cotype implies metric cotype. The metric cotype property makes clearer than the linear version that the inequality relates two different discretisations of the gradient of a function. The key difference in using grids rather than simply cubes is that the expression that comes from diagonals is now impossible to collapse to zero without doing the same on the left of the inequality: large steps in the coordinate directions can be built from small diagonals.

Mendel and Naor prove an analogue of the Maurey-Pisier Theorem for a variant of metric cotype, analogous to the variant used by Bourgain et al. in [BMW]. The obstruction that they build in the absence of metric cotype is not just a discrete cube in ℓ_∞ but a large grid.

6 The non-linear Dvoretzky Theorem

The famous theorem of Dvoretzky [D] states that for each k, every normed space of large enough dimension contains a subspace that is nearly isometric to the Euclidean space of dimension k. In the article [BFM] Bourgain, Figiel and Milman proved an analogue of this for finite metric spaces: every n-point metric space contains a subset of at least about $\log n$ elements that can be almost perfectly embedded in Euclidean space. During the 1980's a great deal was learned about what dimension of Euclidean subspace could be found in what spaces. One of the simplest facts is that the space ℓ_∞^n does *not* have subspaces isomorphic to Euclidean spaces of dimension larger than $\log n$: (see [B2] for an elementary discussion of the problem). Based on this fact it was natural to believe that the result of [BFM] gave the correct dependence on n whatever distortion of the metric we allow on the subset. So it was something of a shock when Bartal, Linial, Mendel and Naor [BLMN] discovered a remarkable threshold phenomenon which does not exist in the linear theory.

Theorem 8 (Bartal, Linial, Mendel, Naor). *For every $\varepsilon > 0$ there is a constant $C(\varepsilon)$ so that every n point metric space contains a subset of size at least $n^{1-\varepsilon}$ which can* *be embedded in Euclidean space with its metric distorted by a factor of no more than $C(\varepsilon)$. Moreover, for every $C > 2$ there is a constant $\alpha(C) > 0$ so that every n-point metric space contains subsets of size at least n^α which are C-Lipschitz equivalent to subsets of Euclidean space.*

In this paper the authors show that the threshold $C > 2$ is sharp: if we insist that the distortion of the metric on the subset is less than 2, we can find Euclidean subsets only of logarithmic size, as in the earlier theorem of [BFM]. However, once we allow the metric to be distorted by a factor of more than 2 the size jumps to a power of n and by increasing the distortion we can take this power as close to 1 as we wish. As one might expect, these very much larger subsets are of significance in applications to the theory of algorithms.

Following the renewed interest in the non-linear Dvoretzky Theorem, Tao asked the wonderfully provocative question: "Why was the *cardinality* of a finite metric space chosen to be the analogue of the dimension of a normed space, instead of the *Hausdorff dimension* of the metric space?" In their remarkable recent article [MN2] Mendel and Naor solved the problem of Tao in the following very strong way:

Theorem 9 (Mendel, Naor)**.** *There is a constant C so that for each $\varepsilon > 0$, every compact metric space X contains a closed subset S whose Hausdorff dimension is at least $1 - \varepsilon$ times that of X and which embeds in Hilbert space with a distortion at most C/ε. The dependence of the distortion on the dimension is optimal.*

When Tao first posed the problem it was tempting to think that it could be tackled using methods similar to those used for the cardinality version. This now looks implausible. In an article in preparation [MN3] Mendel and Naor have found a very considerable generalisation of Theorem 9 which not only implies the cardinality version as well as the Hausdorff dimension version but also implies the famous majorising measure theorem of Talagrand, [T]. The statement is roughly the following:

Theorem 10 (Mendel, Naor)**.** *There is a constant C and for each $\varepsilon > 0$ a constant $K = K(\varepsilon)$ so that for any metric space X and probability measure μ on X there is a subset S of X which embeds in Hilbert space with a distortion at most C/ε and a probability measure ν on S for which*

$$\nu\left(B(x,r)\right) \leq \mu\left(B(x,Kr)\right)^{1-\varepsilon}$$

for each metric ball $B(x,r)$ in X.

References

[B1] K. M. Ball, Markov chains, Riesz transforms and Lipschitz maps, *GAFA* **2**, (1992), 137–172.

[B2] K. M. Ball, An Elementary Introduction to Modern Convex Geometry, in Flavors of Geometry (Silvio Levy ed.), *MSRI lecture notes*, CUP, 1997.

[BCL] K. Ball, E. A. Carlen and E. H. Lieb, Sharp uniform convexity and smoothness inequalities for trace norms, *Invent. Math.* **115**, (1994), 463–482.

[BLMN] Y. Bartal, N. Linial, M. Mendel and A. Naor, On metric Ramsey-type phenomena, *Ann. of Math.* **162**, (2005), 643–709.

[Bo1] J. Bourgain, The metrical interpretation of superreflexivity in Banach spaces, *Israel J. Math.* **56**, (1986), 222–230.

[Bo2] J. Bourgain, On Lipschitz embedding of finite metric spaces in Hilbert space, *Israel J. Math.* **52** (1985), 46–52.

[BFM] J. Bourgain, T. Figiel and V. Milman, On Hilbertian subsets of finite metric spaces, *Israel J. Math.* **55**, (1986), 147–152.

[BMW] J. Bourgain, V. Milman, and H. Wolfson, On type of metric spaces, *Trans. Amer. Math. Soc.* **294** (1986), 295–317.

[D] A. Dvoretzky, Some results on convex bodies and Banach spaces, *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, (1961) Jerusalem Academic Press, 123–160.

[E1] P. Enflo, On the nonexistence of uniform homeomorphisms between L_p-spaces, *Ark. Mat.* **8**, (1969), 103–105.

[E2] P. Enflo, Banach spaces which can be given an equivalent uniformly convex norm, *Israel J. Math.* **13**, (1973), 281–288.

[J] F. John, Extremum problems with inequalities as subsidiary conditions, Studies and Essays Presented to R. Courant on his 60th Birthday, (1948), Interscience N.Y., 187–204

[JL] W. B. Johnson and J. Lindenstrauss, Extensions of lipschitz mappings into Hilbert space, *Contemp. Math.* **26**, (1984), 189–206.

[K] S. Kwapień, Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients, *Studia Math.* **44** (1972), 583–595.

[M1] B. Maurey, Théorèmes de factorisation pour les opérateurs linéaires à valeurs dans les espaces L_p, *Astérisque* **11** (1974).

[M2] B. Maurey, Type, cotype and K-convexity, Handbook of the Geometry of Banach Spaces, Vol. 2, (W. B. Johnson and J. Lindenstrauss, eds.), (2003) North-Holland Amsterdam, 1299–1332.

[MN1] M. Mendel and A. Naor, Metric cotype, *Annals of Mathematics* **168** (2008), 247–298.

[MN2] M. Mendel and A. Naor, Ultrametric subsets with large Hausdorff dimension, submitted.

[MN3] M. Mendel and A. Naor, Ultrametric skeletons, in preparation.

[NPSS] A. Naor, Y. Peres, O. Schramm and S. Sheffield, Markov chains in smooth banach spaces and Gromov-hyperbolic metric spaces, *Duke. Math J.* **134**, (2006), 165–197.

[P1] G. Pisier, Sur les espaces de Banach qui ne contiennent pas uniformément de ℓ_1^n, *C. R. Acad. Sci. Paris, série A* **277**, (1973), 991–994.

[P2] G. Pisier, Martingales with values in uniformly convex spaces, *Israel J. Math.* **20**, (1975), 326–350.

[R] M. Ribe, On uniformly homeomorphic normed spaces, *Ark. Mat.* **14** (1976), 237–244.

[T] M. Talagrand, Regularity of Gaussian processes, *Acta Math.* **159** (1987) 99–149.