The Legacy of Jean Bourgain in Geometric Functional Analysis

Keith Ball

December 1, 2020

Introduction

Jean Bourgain began his career working in the geometry of Banach spaces. At that time, the late '70s and early '80s, several of the most prominent figures in the field had begun to focus on the study of quantitative problems in finite-dimensional spaces rather than on the more traditional infinite-dimensional theory which was, at least in spirit, qualitative. These quantitative problems suited Jean's mathematical temperament perfectly[1] and in the early '80s he produced some of the most crucial ideas that underpin the subject and some of his most striking and beautiful results. Since this part of Jean's work is all over 30 years old it is natural that the areas in question have developed considerably since the time he was involved with them. So this article will discuss the relevant parts of Jean Bourgain's early work and describe at least briefly how they were followed up.

The study of finite-dimensional normed spaces has fed into at least 3 areas of mathematics that have been very successful in recent decades. Metric geometry has become a staple of mathematical computer science and the theory of algorithms: the geometry of convex domains has been revitalised and now provides a subtle probabilistic picture of high-dimensional space: and the study of random operators is central to the field of data-compression. (The latter field also has deep connections to number theory and quantum mechanics but research in those directions is in a rather different spirit.) Jean played a seminal role in each of these developments.

This article discusses aspects of each of these three areas. Section 1 starts with a brief background on normed spaces and convex domains. It then discusses the important reverse Santaló inequality of Bourgain and Milman, Dvoretzky's Theorem, the "slicing problem" made popular by Jean, the conjecture of Kannan, Lovász and

---

[1]This viewpoint is echoed in the article by Tao [T] in this volume.

Simonovits concerning the Poincaré inequality for convex domains and the "central limit problem" raised by the present author and solved by Klartag. Section 2 describes the origins of metric geometry, the "Ribe programme" proposed by Jean, and his striking results about how close a metric space must be to a subset of Hilbert space. It then discusses the subsequent development of the Ribe programme and the non-linear Dvoretzky Theorem by a number of mathematicians and mathematical computer scientists, especially Mendel and Naor. Section 3 discusses a rather specific topic within the huge field of random matrices or operators: restricted invertibility. Jean proved some extremely delicate results in this area which it would be impossible to discuss in detail and many of the methods he developed are closely related to his work in harmonic analysis (which forms the subject of other articles in this collection). However one part of this work, that relates specifically to the famous Kadison-Singer problem, seems to me to be so characteristic of Jean's style and so enjoyable to read that I will explain the proof. At the end of the section I will briefly mention the recent solution of the Kadison-Singer problem by Marcus, Spielman and Srivastava.

Jean Bourgain is most famous for long and extremely complex papers in which he brings his enormous mastery of analytic techniques to bear on deep problems. He did write a number of such articles in the areas I will be discussing but in the end his most enduring contributions rely more on striking insights and his vision of the subject's future. So I have taken the opportunity to sketch several proofs that are not especially long, in sufficient detail that the general mathematical reader can really appreciate them. Bourgain was an enormous inspiration to analysts of my generation. Rather than try to give a comprehensive account of these fields I decided that it was more important to convey some of the delight we felt watching him produce these beautiful arguments.

1 High-dimensional geometry

Functional analysis was originally concerned mainly with infinite-dimensional spaces and operators between them: most especially the spaces of functions that appear in PDE, quantum mechanics and harmonic analysis. All normed spaces of a given *finite* dimension are isomorphic to one another so in a crude sense they are all Euclidean spaces and it might seem that there is nothing more to say. But at a quantitative level this is far from true.

In [Jo] Fritz John proved that if X is a d-dimensional normed space then there is a linear bijection $T : X \to \ell_2^d$ with $\|T\|.\|T^{-1}\| \leq \sqrt{d}$. We express this by saying that X is $\sqrt{d}$-isomorphic to ℓ_2^d. John proved the theorem by regarding X as $\mathbf{R}^d$ equipped with a norm and considering the ellipsoid of largest volume $\mathcal{E}$ inside the unit ball

B_X of X. (He showed among other things that this ellipsoid is indeed unique.) John proved that that $B_X \subset \sqrt{d}\mathcal{E}$. Since the ellipsoid is the unit ball of an Euclidean norm, the inclusions

$$\mathcal{E} \subset B_X \subset \sqrt{d}\mathcal{E}$$

imply the theorem. Thus John proved a theorem in functional analysis by asking geometric questions involving volume and tangency of convex domains: the unit balls of normed spaces. His theorem is sharp in that there are normed spaces such as ℓ_∞^d and ℓ_1^d whose "distance" from Euclidean space in this sense, is $\sqrt{d}$: spaces which look very unlike Euclidean spaces. Since every symmetric convex set with non-empty interior in $\mathbf{R}^d$ is the unit ball of a norm on the space, and vice versa, the study of finite-dimensional normed spaces is in principle equivalent to the study of these convex sets but in practice the viewpoints are quite different and in spite of John's Theorem the two areas of mathematics remained largely separate for many years.

The classical theory of convex geometry developed by Minkowski and Blaschke grew out of the isoperimetric principle: that among domains of a given volume the Euclidean balls have the smallest surface area. The Brunn-Minkowski inequality, which generalises the isoperimetric inequality, states that if C and D are non-empty compact subsets of $\mathbf{R}^d$ then the set $C + D$ of sums

$$\{c + d : c \in C, d \in D\}$$

has large volume: $\operatorname{vol}(C + D)^{1/d} \geq \operatorname{vol}(C)^{1/d} + \operatorname{vol}(D)^{1/d}$. While the statement makes no mention of convexity, much of the power of the inequality stems from the following fact. Suppose C is a compact convex set in $\mathbf{R}^d$ and e is a unit vector. Scan across C with hyperplanes perpendicular to e, measuring the $(d-1)$-dimensional volume of the slices of C. The volume of the slice regarded as a function of the hyperplane's position, has a concave logarithm. (See [B2] for a detailed discussion.)

1.1 Bodies and their polars

After the Brunn-Minkowski Theorem, probably the most famous classical fact about convex domains is the Blaschke-Santaló inequality, [Bla] and [S]. If C is a convex set in $\mathbf{R}^d$ then its polar is

$$C^\circ = \{y \in \mathbf{R}^d : \langle x, y \rangle \leq 1, \text{ for all } x \in C\}.$$

The inequality states that if C has the origin as its centre of mass then the product of the volumes of C and its polar is at most the corresponding product for the standard Euclidean ball B_d in $\mathbf{R}^d$:

$$\operatorname{vol}(C) \operatorname{vol}(C^\circ) \leq \operatorname{vol}(B_d)^2.$$

3

One of Minkowski's motivations for studying convex domains was what is now known as the geometry of numbers, which involves the interplay between lattices in $\mathbf{R}^d$ and domains, usually convex ones. In this context the number theorist K. Mahler asked what lower bound can be put on these volume products. By John's Theorem it is clear that if C is a symmetric convex set then

$$\operatorname{vol}(C)\operatorname{vol}(C^{\circ}) \geq \left(\frac{1}{\sqrt{d}}\right)^{d} \operatorname{vol}(B_d)^2.$$

However if we consider the pair in which C is the unit ball of ℓ_{∞}^d and C° the ball of ℓ_1^d the volume product is $\frac{4^d}{d!}$ and this is more than

$$\left(\frac{2}{\pi}\right)^{d} \operatorname{vol}(B_d)^2.$$

Thus, as far as Mahler's question is concerned, the spaces that seem the furthest from being Euclidean, appear to behave as if they are rather close to Euclidean. Mahler conjectured that among symmetric convex domains the cube and its polar, the unit balls of ℓ_{∞}^d and ℓ_1^d, have the smallest volume product. The fact that for this pair, the product is much larger than you might guess, indicates just how powerful such a result would be. Mahler's conjecture is still unproved but for most purposes what matters is an approximate statement: the reverse Santaló inequality of Bourgain and Milman [BM].

Theorem 1 (Bourgain-Milman). *There is a constant K, independent of dimension, so that for all symmetric convex domains C*

$$\operatorname{vol}(C)\operatorname{vol}(C^{\circ}) \geq \left(\frac{1}{K}\right)^{d} \operatorname{vol}(B_d)^2.$$

The assumption of symmetry was quickly removed in subsequent works: the more general statement, for convex domains that may not be symmetric, doesn't really add anything to the original. This reverse Santaló inequality is a theorem that is used so frequently that people forget they are doing it. Duality is central to the study of normed spaces and this theorem allows you to flip back and forth between spaces and their duals without losing significant information about the volumes of the unit balls.

The proof of this remarkable fact depends upon a longish programme that drew together the two viewpoints: the linear theory of finite-dimensional normed spaces and the geometry of convex domains. The programme is described in loving detail in the book by Pisier [P1]. The real starting point for this unification was a seminal theorem of Dvoretzky [D] from the late '50s that answered a question of Grothendieck.

Theorem 2 (Dvoretzky). *For each natural number k and each $\epsilon > 0$ there is a natural number d with the property that every normed space of dimension d contains a subspace of dimension k which is $(1 + \epsilon)$-isomorphic to ℓ_2^k.*

The original proof of the Bourgain-Milman Theorem used a subtle but rather technical estimate[2] of Milman's [Mi2] depending upon the distribution of the standard rotation-invariant measure on the Euclidean sphere together with the theory of type and cotype developed principally by Kwapień, Maurey and Pisier: see in particular [K], [Ma2] and [P2]. Both Dvoretzky's Theorem and the theory of type and cotype were also important motivations for the development of metric geometry which will be discussed in the second section of this article.

1.2 Slicing

In the introduction it was mentioned that we now have a view of domains in high-dimensional space, especially convex domains, that is inspired by probability theory. Suppose C is a convex domain in $\mathbf{R}^d$ of volume 1. Its indicator $\mathbf{1}_C$ is the density of a random vector: X say. For each unit vector θ we can consider the random variable $\langle X, \theta \rangle$. After an appropriate affine transformation we can arrange that each of these random variables has mean 0 (the centre of mass of C is the origin) and that they all have the same variance (the inertia tensor of C is a multiple of the identity). If so then we call the domain *isotropic*. Now if our domain is the cube $[-1/2, 1/2]^d$ in $\mathbf{R}^d$ then its indicator function is the joint density of IID random variables $(X_1, X_2, \ldots, X_d)$ uniformly distributed on $[-1/2, 1/2]$. Its marginal in the direction $\theta = (1/\sqrt{d}, 1/\sqrt{d}, \ldots, 1/\sqrt{d})$ is

$$\frac{1}{\sqrt{d}} \sum_{i=1}^{d} X_i$$

and so by the Central Limit Theorem it is almost Gaussian (with variance independent of d). During the late '80s and early '90s it became more and more apparent that the indicator of any convex domain should exhibit features that we expect of the joint density of independent random variables.

The Brunn-Minkowski inequality stated earlier implies that each of the marginals $\langle X, \theta \rangle$ of C has a density on the line, f_θ say, whose logarithm is concave. That guarantees that the density decreases exponentially fast as we move away from the function's maximum value. This in turn ensures that if L^2 is the variance of f_θ then the maximum value of f_θ must be around $1/L$. If C is a *symmetric* isotropic convex domain then for each θ the density f_θ attains its maximum at 0 where its value is

---

[2]Milman calls it the "lower M^* estimate"

the $(d-1)$-dimensional volume of the slice of C by the subspace orthogonal to θ. Hensley [H] pointed out that as a result, any symmetric isotropic convex domain in $\mathbf{R}^d$ has the property that the volumes of its central slices are all about the same as one another. Formally

Theorem 3 (Hensley). *There are constants c_1 and c_2 so that the following is true. Let d be a natural number and C an isotropic convex domain in $\mathbf{R}^d$ with*

$$\int_C \langle x, \theta \rangle^2 \, dx = L^2$$

for all unit vectors θ. Then for every 1-codimensional subspace H of $\mathbf{R}^d$

$$\frac{c_1}{L} \leq \text{vol}(C \cap H) \leq \frac{c_2}{L}.$$

If the indicator of the domain C does indeed behave like the joint density of IID one-dimensional random variables then it should look like the product of those densities: in particular its value at 0 (which is 1) should be roughly the product of the d values $f_{e_1}(0), f_{e_2}(0), \ldots, f_{e_d}(0)$ of its marginals, in a set of orthogonal directions. Each of these values is roughly $1/L$ so we would expect L to be roughly 1, or roughly constant, independent of dimension.

Clearly the expectation of the square of the length of the random vector X is

$$\int_C |x|^2 \, dx = \sum_{i=1}^{d} \int_C \langle x, e_i \rangle^2 \, dx = dL^2.$$

So for any C the value of L is at least as large as that of the Euclidean ball: the Euclidean ball packs a given volume as close as possible to the origin. For this set the value of L is approximately $1/\sqrt{2\pi e}$ and so it is indeed roughly a constant independent of dimension. The conjecture that there is a constant *upper* bound as well as a constant lower bound for the values of L is known as the "slicing problem" and is often attributed to Bourgain although I'm not convinced that he ever really believed it.

Conjecture 4 (Bourgain). *There is a constant K independent of d so that for any isotropic convex domain of volume 1, $L \leq K$.*

Soon after the slicing problem started attracting attention it became clear that more or less equivalent problems had been asked about convex domains over many years dating back as far as the work of Sylvester. The conjecture is known to be true for quite a few families of unit balls, in the sense that there is a uniform bound for all members of the family independent of dimension: for example it is quite easy to prove for the unit balls of finite-dimensional ℓ_p spaces.

6

It is easy to see that for any symmetric isotropic domain in d dimensions the value of L is at most about $\sqrt{d}$ just because the domain cannot have larger width than the Euclidean ball in every direction. The question is whether this upper bound can be improved to a constant. In one of his most celebrated early works Bourgain [Bo3] improved the bound to $d^{1/4}(\log d)^\alpha$ for some power α.

Theorem 5 (Bourgain). *For some $\alpha > 0$, if C is an isotropic convex domain in $\mathbf{R}^d$ then*

$$\left(\frac{1}{d}\int_C |x|^2\right)^{1/2} \le d^{1/4}(\log d)^\alpha.$$

The logarithmic factor in Bourgain's Theorem was removed by Klartag by combining his "isomorphic slicing theorem" [Kl1] with a deep result of Paouris [Pao].

1.3   Subsequent developments

Since Bourgain and Milman proved the reverse Santaló inequality there have been a number of other, very different, proofs for example by Kuperberg [Ku] and Nazarov [Na]. Fairly recently Giannopoulos, Paouris and Vritsiou [GPV] gave a proof that is much more in the spirit of convex geometry and is surprisingly elementary. The development of high-dimensional geometry during the last 30 years has been guided by the Bourgain-Milman Theorem and its proof, by the slicing problem and by two other conjectures. The first of these, proposed in [KLS], concerns the spectral gap for the Neumann Laplacian on a convex domain. In concrete terms it states the following:

Conjecture 6 (Kannan, Lovász and Simonovits). *There is a constant K independent of d so that for any isotropic convex domain $C \subset \mathbf{R}^d$ of volume 1, and every differentiable $f : C \to \mathbf{R}$ with mean $\bar{f} = \int_C f$ we have*

$$\int_C (f - \bar{f})^2 \le K^2 L^2 \int_C \|\nabla f\|^2.$$

A Poincaré inequality of this kind is known to hold for every open domain but the value of the constant K can be huge for nasty domains. Even if the domain is convex the constant can be large if the domain is long and thin. The conjecture is that once we apply a linear map to make our domain isotropic the constant is bounded independently of C and the dimension d. If f is a *linear* function then the inequality is an identity with $K = 1$ because it is just the definition of L. So the conjecture states that, up to a constant, the linear functions are the worst for the Poincaré inequality.

The second conjecture, publicised by the author and also Brehm and Voigt (see [ABP] and [BV]), was known as the central limit problem for convex domains and

describes a sense in which their indicator functions do indeed behave like the joint densities of independent random variables. Its rough statement is the following.

Conjecture 7 (Ball, Brehm-Voigt). *The 1-dimensional marginals of an isotropic domain are approximately Gaussian in all but a tiny proportion of directions.*

It was explained in [ABP] that this conjecture is very much weaker than the KLS conjecture (applied to the function $x \mapsto \|x\|_2$). As was mentioned in the introduction, a proof of the conjecture was found by Klartag [Kl2]. A completely different proof was given by Fleury, Guédon and Paouris [FGP]. The KLS conjecture on the other hand is very far from being proved.

The slicing problem and the KLS conjecture are also related to one another. Formally they are independent: the Poincaré inequality takes for granted that the natural length scale in C is L (however large that is) while the slicing conjecture claims that L is always at most a fixed number. However the natural feeling is that the two are effectively opposites of one another. The Poincaré inequality says that linear functions deviate most rapidly: the slicing conjecture says that linear functions do not deviate rapidly at all. However it was demonstrated by the present author that the KLS conjecture actually implies the slicing conjecture and an extended version of the argument appears in [BN]. With hindsight it may not be so surprising that KLS implies the slicing conjecture. KLS implies that most of the volume of the body lies in a fairly thin shell. The isotropic property guarantees that this volume is distributed rather uniformly around the shell. If the slicing conjecture fails then the shell has large radius and it is plausible that the convexity of the domain fills in so much that the volume of the domain must be larger than 1.

Currently the best bound known in the KLS conjecture is $d^{1/4}$ which follows from the work of Eldan [El] and Lee and Vempala [LV]. The latter article provides a very clear and readable overview of the state of the art. My personal feeling is that the slicing problem and the KLS conjecture are too optimistic but I would be very reluctant to guess the correct order of the constants.

2 Metric geometry

In 1976 Ribe [R] proved a rigidity theorem which effectively states that finite-dimensional properties of a Banach space are determined up to isomorphism by the metric structure of the space without any reference to its linear structure. Prompted by this, Bourgain [Bo1] proposed a research programme which is now known as "The Ribe Programme": to find explicit metric descriptions of the most important invariants of normed linear spaces. He himself started the programme by providing a

metric characterisation of superreflexivity. A Banach space X is called *superreflexive* if every space whose finite-dimensional spaces embed into X, must automatically be reflexive. Elementary functional analysis courses often contain the theorem that uniformly convex spaces are reflexive. Since uniform convexity is a finite-dimensional property they are therefore superreflexive. Clearly any space which is linearly isomorphic to a uniformly convex space will also be superreflexive. A series of works by James and Enflo [J] and [E2] demonstrated that actually these are the only ones. Another characterisation of superreflexivity, in terms of the growth of martingales in the space, was found by Pisier in [P3]. These beautiful results provide *geometric* characterisations of superreflexivity but still rely on the linear structure. Bourgain [Bo1] showed that superreflexivity holds precisely if the space does not contain copies of arbitrarily large binary trees (viewed as metric spaces in the obvious way). Needless to say the real point of the Ribe programme is not merely to find metric equivalents of linear properties. The broader aim is to mimic the well-developed theory of normed spaces in the non-linear setting.

In studying normed spaces it is frequently useful to know that the unit ball of a d-dimensional normed space has an ε-net with no more than about $\left(\frac{5}{\varepsilon}\right)^d$ elements (and that you can't make do with significantly fewer). So a metric space with around 10^d elements should be capable of mimicking all the bad behaviour of a d-dimensional normed space: that an n-point metric space is as bad as a $\log n$-dimensional normed space. One's first instinct is that general metric spaces must by potentially much wilder than normed spaces. There is *some* limit to this wildness since every metric space with n points can be isometrically embedded into ℓ_∞^n. But that is a far cry from a $\log n$-dimensional normed space. Nevertheless there was a feeling in the early '80s that in studying finite-metric spaces the logarithm of the number of elements should play a role something like the dimension of a normed space.

This feeling got a boost from a famous paper of Johnson and Lindenstrauss [JL]. They studied extensions of Lipschitz maps from metric spaces into Hilbert space and in order to do so they proved the 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 a dataset embedded in a space of low dimension is much easier to search than one in a space of higher dimension. Following this article Joram Lindenstrauss especially, promoted the idea that there should be a subtle analogue of the theory of normed spaces but for general metric spaces.

In view of John's Theorem it was natural to ask how far could an n-point metric space be from a subset of Hilbert space. If an n-point metric space does indeed look like a subset of a $\log n$-dimensional normed space then the answer should be at most

$\sqrt{\log n}$. This is too much to ask but in [Bo2] Bourgain found essentially the correct order: $\log n$. His argument is not very long but it is one of his most striking and has inspired a huge range of subsequent arguments in the field of metric geometry and the theory of algorithms in mathematical computer science. This will be the subject of Section 2.1 below.

In the following year Bourgain wrote two other articles that played a key role in starting the field of metric geometry. Just as John's Theorem concerning the distance between normed spaces prompts one to ask about the distance of a finite metric space from a subset of Euclidean space, so Dvoretzky's Theorem prompts one to ask how large a subset must a metric space contain, that is actually embeddable in Hilbert space with at most a constant distortion. Again, the first result in this direction was proved by Bourgain, this time in collaboration with Figiel and Milman, [BFM]. They proved that every n-point metric space contains a subset of at least about $\log n$ elements that can be almost perfectly embedded in Euclidean space. This non-linear Dvoretzky Theorem was the start of a long development which will be discussed in Section 2.3 below. The second of the two articles [BMW] discussed a metric version of a linear invariant: the type of a normed space. This and subsequent work in this area by the present author and Mendel and Naor will be discussed in Section 2.2. The Ribe programme as a whole is explained in more detail in the present author's article [B3] and the much more detailed survey by Naor [N].

The field of metric geometry is now extremely active not least because of its connections to theoretical computer science. It was founded in the early '80s and I think it is fair to say that J. Lindenstrauss and J. Bourgain were the principal architects.

2.1 How Euclidean is a metric space?

As explained above, Bourgain [Bo2] proved the following remarkable theorem which has hugely influenced subsequent developments in metric geometry.

Theorem 8 (Bourgain)**.** *There is a constant K so that if X is an n-point metric space there is a map $f : X \to \ell_2$ so that for every pair of points $x, y \in X$*

$$\frac{1}{K \log n} d(x, y) \le \|f(x) - f(y)\| \le d(x, y).$$

Bourgain constructs the map f as follows. He indexes the coordinates of an Euclidean space (of the correct dimension) by the non-empty subsets A of X. Each point $x \in X$ is mapped to the vector whose A-coordinate is the distance $d(x, A)$ scaled by a weight factor $\sqrt{p_A}$. The weights are chosen so that $\sum_{A \subset X} p_A = 1$ which automatically ensures that if x and y are points in X the distance between their

images satisfies

$$\|f(x) - f(y)\|^2 = \sum_{A \subset X} p_A \left(d(x, A) - d(y, A)\right)^2 \leq \sum_{A \subset X} p_A \, d(x, y)^2 \leq d(x, y)^2.$$

Thus a subset A is chosen randomly according to the distribution determined by the probabilities p_A and the problem is to show that for every pair x and y the expectation of $(d(x, A) - d(y, A))^2$ is at least about $\frac{1}{(\log n)^2} d(x, y)^2$.

The really remarkable thing about the proof is that the weight p_A depends only upon the size of A and not in any way upon the metric space in question. Somehow, a typical set in X can detect more or less how far apart are the points x and y, regardless of how they sit in the metric space. The random set is chosen as follows. Pick an integer j between 1 and $1 + \log_2 n$ (each integer equally likely) and now choose a random set by including each point of the space with probability $p = 1/2^j$, independently of the others (and excluding it with probability $1 - p$). The following proof that this works is a slight modification of Bourgain's original argument, due to Matousek and Naor.

If $B_1 \subset X$ is a set in X with 2^j elements then the random set chosen with $p = 1/2^j$ will miss B_1 with probability $(1 - p)^{2^j}$ which is about $1/e$. If B_2 has only 2^{j-1} elements then the random set is more likely to miss it but still with probability only $1/\sqrt{e}$. As long as the sets are disjoint (so their points are chosen independently) there is a fair chance that the random set will miss the larger, B_1, but meet the smaller, B_2.

Now fix x and y in X and for each j let r_j be the smallest radius for which the balls $B(x, r_j)$ and $B(y, r_j)$ around x and y both contain at least 2^j points of the metric space. At least one of the open balls $B^\circ(x, r_j)$ and $B^\circ(y, r_j)$ contains fewer than 2^j points: let's assume that the first one does so. Up to the point where r_j is about say $d(x, y)/3$ the sets $B^\circ(x, r_j)$ and $B(y, r_{j-1})$ will be disjoint and so there is a fixed probability that a random set whose points are chosen with probability $1/2^j$ misses the first and hits the second. For such a set A

$$\left(d(x, A) - d(y, A)\right)^2 \geq (r_j - r_{j-1})^2.$$

If we replace each r_j by $\tilde{r}_j$, the minimum of r_j and $d(x, y)/3$, we get that for *all* j

$$\left(d(x, A) - d(y, A)\right)^2 \geq (\tilde{r}_j - \tilde{r}_{j-1})^2.$$

Averaging over the (at most) $1 + \log_2 n$ different values of j we get

$$\|f(x) - f(y)\|^2 \geq \frac{c}{1 + \log_2 n} \sum_j (\tilde{r}_j - \tilde{r}_{j-1})^2$$

11

for some constant c. By the Cauchy-Schwarz inequality the latter is at least

$$\frac{c}{(1 + \log_2 n)^2} \left(\sum_j (\tilde{r}_j - \tilde{r}_{j-1}) \right)^2 .$$

Since the greatest value of $\tilde{r}_j$ is $d(x,y)/3$ and the least is 0, this gives the lower bound we want. $\qquad\square$

In harmonic analysis it is common to consider what is happening on many different length scales: indeed one might almost regard this as the definition of harmonic analysis. Usually one chooses the lengths, 1, 1/2 1/4 and so on, from the start. Here Bourgain gives the idea a twist. He chooses the numbers of points in the balls $B(x, r_j)$ to be powers of 2 and lets the radii take care of themselves. This idea has reappeared quite regularly in metric geometry.

The idea and the theorem itself have also been used in many articles in theoretical computer science: the following are just a smattering. Linial, London and Rabinovich [LLR] find a polynomial-time random embedding of a metric space into Hilbert space that achieves Bourgain's bound and, as the authors point out, the overall structure of the proof follows Bourgain's. The max-flow/min-cut problem for a single commodity is famously easy to solve but if there are several commodities it becomes NP-hard. In [AR], which uses the Linial, London and Rabinovich result, Aumann and Rabani obtain a cut which approximates the optimal one (in size). The way they do this illustrates rather well one of the reasons that embeddings of metric spaces play a role in the theory of algorithms. The problem naturally associates vectors in $\mathbf{R}^k$ to the vertices of the network and their aim is to find a cut by using a hyperplane to partition these vectors into two subsets. In order for the method to work they need to embed the metric space consisting of the vectors equipped with the ℓ_∞ norm, into a better space. The use of hyperplanes to find cuts in combinatorial settings itself appears quite frequently in the theory of algorithms: most famously in the paper of Goemans and Williamson [GW]. Feige [F] uses an extension of the embedding method to tackle what looks like a completely different problem: the so-called bandwidth problem for graphs: how badly the edges of the graph must be stretched to peg the vertices at points 1, 2, 3 and so on.

2.2 Metric type and cotype

The most successful family of achievements within the Ribe programme has been the development of metric equivalents of the linear invariants known as type and cotype. As metioned earlier the original *linear* theory was due predominantly to Kwapień, Maurey and Pisier. 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

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

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

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

for every such sequence. The parallelogram identity shows that Hilbert space has type and cotype 2: the best possible. 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. A result of Kwapień [K] shows that only Hilbert space can possess both type 2 *and* cotype 2.

The non-linear theory began formally with an article by Bourgain, Milman and Wolfson [BMW] who introduced a metric form of the type invariant, but it had been anticipated by an article of Enflo [E1]. He 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? Bourgain, Milman and Wolfson [BMW] 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 [P2]) which states that the finite-dimensional L_1-spaces are the only obstructions to type:

Theorem 9 (Pisier)**.** *If a normed space X fails to have type p for every $p > 1$ 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.*

The main theorem in [BMW] states that a *metric* space which has no non-trivial *metric* type contains a non-linear analogue of the finite-dimensional ℓ_1 spaces.

Theorem 10 (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.*

Thus although there need not be a copy of anything like a linear space, a metric space with no type must include a copy of the Hamming cube. This was the first really convincing sign that metric spaces might have a subtle structure that parallels that of normed spaces.

The obvious question arising from this article was "What is metric cotype?" This question is made difficult by the fact that one cannot simply compare edges and diagonals of a cube in the opposite order to Enflo: the resulting inequality never holds in any space. One answer was given in the author's paper [B1] which introduces the properties of *Markov* type and cotype. These describe the behaviour of time-reversible stationary Markov chains in the space: for example 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. The properties were introduced in order to study extensions of Lipschitz maps. At the time 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]. Their result combined with those from [B1] gives the following non-linear analogue of a linear extension theorem of Maurey [Ma1] that generalises the result of Kwapień.

Theorem 11 (Ball, Naor, Peres, Schramm, Sheffield). *If $1 < p \le 2 \le q < \infty$, 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 $a \in A$.*

Although Markov cotype is well adapted for the purpose of Lipschitz extensions it does not fulfil the demand of the Ribe programme for a metric version of the linear cotype. In their paper [MN1] Mendel and Naor finally found an appropriate metric analogue. The test for cotype is the Maurey-Pisier Theorem which states that if a normed space X fails to have cotype q for every $q < \infty$ (the space has *no* non-trivial cotype) then X contains uniformly isomorphic copies the finite-dimensional L_∞-spaces, ℓ_∞^n. Mendel and Naor not only found the metric version of cotype they also found the right obstructions by proving an analogue of the Maurey-Pisier Theorem for a variant of metric cotype. The two parts of the problem are closely linked: the cotype condition depends on grids rather than cubes and the obstruction they build in the absence of metric cotype is not just a discrete cube in ℓ_∞ but a large grid.

2.3   The non-linear Dvoretzky Theorem

The optimal form of Dvoretzky's Theorem proved by Milman [Mi1] guarantees that each normed space of dimension d contains an almost Euclidean subspace of dimension $k \approx \log d$. If a metric space of size $n \approx e^d$ points roughly corresponds to a normed space of dimension d then it should have a subset of size about $e^k = \log n$ points. Bourgain, Figiel and Milman [BFM] proved that this is indeed so.

Theorem 12 (Bourgain, Figiel, Milman). *For every $\epsilon > 0$ there is a constant K so that for every n, every n-point metric space contains a subset of size at least $(\log n)/K$ which is $1 + \epsilon$ Lipschitz equivalent to a subset of Euclidean space.*

Dvoretzky's Theorem is proved by finding a subspace of fairly large dimension in which a norm is almost constant on the Euclidean unit sphere and then using linearity to deduce that the norm and the Euclidean norm are almost the same on the whole subspace. No such method is available in a general metric space so it is extremely hard to imagine how one can detect an almost Euclidean subset of a metric space (let alone find it). The proof of Theorem 12 constructs a subset of each metric space which is much more special than being almost Euclidean: it is almost an *ultrametric* space, meaning that every triangle in the space has its two longer sides equal. Such a metric space has a highly restricted structure: it can be represented as the leaves of a tree in which the distance between any two leaves depends only upon the identity of their nearest common ancestor. It is a nice exercise to check that every ultrametric space can indeed be embedded in Hilbert space. As a result, at first sight Theorem 12 looks like a cheat: that it doesn't really have anything much to do with Euclidean space. It looks as though what is happening is that $\log n$ points is such a small part of an n-point space that you can find them satisfying an incredibly restrictive condition that happens to imply that they sit in Euclidean space but implies much more. In fact the opposite is true. Theorem 12 was the beginning of a remarkable story showing that finite metric spaces actually contain quite large ultrametric subsets that act as a kind of skeleton of the space: the space is a fleshed-out version of its ultrametric skeletons.

It is easy to check that the space ℓ_∞^d does *not* have approximately Euclidean subspaces of dimension larger than $\log d$: (see [B2] for an elementary discussion of the problem). So it was natural to believe that the result of [BFM] gave the correct dependence on d whatever distortion of the metric we allow on the subset. However Bartal, Linial, Mendel and Naor [BLMN] discovered that if the distortion is allowed to be more than 2, there are very much larger subsets.

Theorem 13 (Bartal, Linial, Mendel, Naor)**.** *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 an ultrametric space.*

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. The ultrametric skeleton idea appears in an article of Mendel and Naor [MN2] who prove that

Theorem 14 (Mendel, Naor)**.** *For each $\varepsilon > 0$ there is a constant $C = C(\varepsilon)$ so that*

for any metric space X and probability measure μ on X there is a subset S of X which is $1/\varepsilon$ equivalent to an ultrametric space and a probability measure ν on S for which

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

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

So, however you measure the sizes of different parts of X, the skeleton S is well represented in each part: it really is a skeleton not just a subset.

3 Restricted invertibilty and the Kadison-Singer problem

In 1959 Kadison and Singer [KS] asked a question about the uniqueness of extensions of certain linear functionals called pure states, on C^*-algebras. In 1979 Anderson [A] found a reformulation of the problem solely in terms of matrices acting on finite-dimensional Euclidean spaces. The question can be reformulated further in terms of finite families of vectors as follows. (This reformulation was mentioned at the end of the article [BT] without proof. It is proved in the article of Casazza and Tremain [CT].)

Problem 15 (Kadison-Singer). *Is there a number r so that for any m and any unit vectors $u_1, u_2, \ldots, u_m$ in Hilbert space satisfying*

$$\left\|\sum_i \lambda_i u_i\right\|^2 \leq 2 \sum_i \lambda_i^2$$

we can partition the sequence into r subsequences $\sigma_1, \sigma_2, \ldots, \sigma_r$ so that each subsequence satisfies

$$\left\|\sum_{i \in \sigma_j} \lambda_i u_i\right\|^2 \geq \frac{1}{2} \sum_{i \in \sigma_j} \lambda_i^2?$$

So the question is asking whether the original sequence can be broken into a bounded number of subsequences each of which looks rather like an orthonormal basis.

At the time this problem looked far out of reach but a natural question that is obviously weaker is whether it is possible to select a fixed *proportion m/r* of the vectors satisfying the lower bound. It is trivial to check that one can select about $\sqrt{m}$ with the desired property but getting a fixed proportion of m is much harder. Bourgain and Tzafriri published a series of papers on this question: in the first [BT] they proved the following selection theorem which is an absolute gem.

Theorem 16 (Bourgain-Tzafriri)**.** *There is a constant $c > 0$ so that if $u_1, u_2, \ldots, u_m$ are unit vectors in Euclidean space satisfying*

$$\left\| \sum_1^m \lambda_i u_i \right\| \leq M \left(\sum_1^m \lambda_i^2 \right)^{1/2}$$

for some M and every sequence of scalars $(\lambda_1, \ldots, \lambda_m)$, then there is a subset $\sigma \subset \{1, 2, \ldots, m\}$ containing at least cm/M^2 of the indices so that for this subset

$$\left\| \sum_{i \in \sigma} \lambda_i u_i \right\| \geq c \left(\sum_{i \in \sigma} \lambda_i^2 \right)^{1/2}. \tag{1}$$

Thus the map $T : \ell_2^m \to \ell_2$ which takes the standard unit vectors to the vectors $u_1, u_2, \ldots, u_m$ respectively, when restricted to a proportion c/M^2 of the standard basis, is nicely invertible (as a map onto its image). The proof of this theorem is quite remarkable. It employs a combination of random and non-random selections and invokes a combinatorial principle which appears to have nothing whatsoever to do with the problem. Moreover the proof proceeds in 3 steps, the first two of which look far too weak to be of any real use.

In the first step the authors randomly select a proportion k of the vectors σ_1 satisfying the much weaker ℓ_∞ estimate:

$$\left\| \sum_{i \in \sigma_1} \lambda_i u_i \right\| \geq \frac{1}{2} \max_{i \in \sigma_1} |\lambda_i|. \tag{2}$$

In the second step they select about half of the vectors in σ_1 so that for this new subset σ_2 they get a lower ℓ_1 estimate:

$$\left\| \sum_{i \in \sigma_2} \lambda_i u_i \right\| \geq \frac{1}{2\sqrt{k}} \sum_{i \in \sigma_2} |\lambda_i|. \tag{3}$$

This selection is non-random and depends upon the Sauer-Shelah lemma from combinatorics [Sau] and [Sh]. Finally in the third step Bourgain and Tzafriri use another selection (which might be called semi-random) of about half the vectors in σ_2 for which they get the bound in equation (1).

The ℓ_∞ lower bound in equation (2) is equivalent to the statement that each u_i, for $i \in \sigma_1$, is at least distance $1/2$ from the span of the other vectors indexed by σ_1. Once you have decided to prove this seemingly extremely weak statement it is easy to see that a random selection will do it. Choose indices independently at random with probability about $1/(10M^2)$. The expected square of the length of the projection of any of the unit vectors onto the span of the randomly chosen ones will

be small. So at least half of the random collection will have the property required. Let this set of indices be σ_1 and k be the number of indices chosen.

The second step is rather magical and it depends upon the Sauer-Shelah Lemma:

Theorem 17 (Sauer-Shelah). *Suppose $\mathcal{F}$ is a collection of choices of sign $(\epsilon_i) \in \{-1, 1\}^n$ with more than*

$$\sum_{j=0}^{r-1} \binom{n}{j}$$

elements. Then there is a set $\tau \subset \{1, 2, \ldots, n\}$ of at least r indices with the property that for any choice of sign $(\delta_j)_{j \in \tau}$ indexed by τ there is an element of $\mathcal{F}$, whose restriction to τ is δ.

Now, yet another way to state the ℓ_∞ lower bound (2) is that the biorthogonal vectors $(\phi)_{i \in \sigma_1}$ to the $(u_i)_{i \in \sigma_1}$ each have norm at most 2. By the parallelogram law at least 2^{k-1} of the sums

$$\sum_{i \in \sigma_1} \varepsilon_i \phi_i$$

(as (ε_i) runs over all 2^k choices of sign) have norm at most $2\sqrt{k}$. So there is a set of indices σ_2 comprising at least half of the indices in σ_1 so that every choice of sign $(\delta_i)_{i \in \sigma_2}$ can be extended to a choice of sign (ε_i) on σ_1. This ensures that the original u_i indexed by σ_2 satisfy the ℓ_1 estimate

$$\left\| \sum_{i \in \sigma_2} \lambda_i u_i \right\| \geq \frac{1}{2\sqrt{k}} \sum_{i \in \sigma_2} |\lambda_i|.$$

At this point in the argument Bourgain and Tzafriri originally used a direct exhaustion argument to select a further subset which satisfies the conclusion of the theorem. They explain that N. Kalton pointed out that an important theorem of Grothendieck [G] could be used instead. We actually need what Pisier calls the "little Grothendieck Theorem" (see [P4] Theorem 5.2). This consequence of Grothendieck's Theorem was first stated in this form by Lindenstrauss and Pełczynski [LP].

Theorem 18 (little GT). *There is a constant K with the following property. Let $T : \ell_\infty^m \to \ell_2$ be linear. Then there is a probability measure $\{p_1, p_2, \ldots, p_m\}$ on the set of indices so that for every $\theta = (\theta_i) \in \ell_\infty^m$*

$$\|T\theta\| \leq K\|T\| \left(\sum p_i \theta_i^2 \right)^{1/2}.$$

If we now consider the biorthogonal vectors $(\psi_i)_{i \in \sigma_2}$ for the smaller set of indices, they satisfy an upper ℓ_∞ bound,

$$\left\| \sum_{i \in \sigma_2} \theta_i \psi_i \right\| \leq 2\sqrt{k} \max_{i \in \sigma_2} |\theta_i|.$$

Hence if we define a linear map $T : \ell_\infty^{|\sigma_2|} \to \ell_2$ by mapping the standard basis vectors to the vectors (ψ_i) then T is bounded by $2\sqrt{k}$. By the little GT we can find positive numbers p_i whose sum is 1 satisfying

$$\left\| \sum_{i \in \sigma_2} \theta_i \psi_i \right\| \le 2K\sqrt{k} \left(\sum_{i \in \sigma_2} p_i \theta_i^2 \right)^{1/2}.$$

At least half of the p_i must be at most $4/k$ since there are at least $k/2$ of them and they add up to 1. If we take σ to be the set of these indices then on this set

$$\left\| \sum_{i \in \sigma} \theta_i \psi_i \right\| \le 4K \left(\sum_{i \in \sigma} \theta_i^2 \right)^{1/2}.$$

$\square$

In their article [BT] Bourgain and Tzafriri prove a related result for more general L_p spaces.

Theorem 19 (Bourgain-Tzafriri). *For every $p \in [1, \infty]$ and every real M and positive ϵ there is a constant c with the following property. If the matrix (a_{ij}) acts on ℓ_p^m with norm at most M and its diagonal entries are equal to 0, then there is a set $\sigma \subset \{1, 2, \dots, m\}$ containing at least a proportion cm of the indices so that the principal submatrix $(a_{ij})_{i,j \in \sigma}$ has norm on $\ell_p^{|\sigma|}$ at most ϵ.*

The case $p = 2$ follows quite easily from the Theorem 16 just discussed. The general case is quite a bit more intricate. It uses similar ideas to the ℓ_2 case together with the sort of delicate computations in L_p spaces that Bourgain employed in other problems such as the Λ_p problem.

In 2008 a completely different proof of Theorem 16 was published by Spielman and Srivastava [SS] based on step by step modification of a quadratic form: a method that they had used with Batson [BSS] for graph sparsification. Finally in 2014 Marcus, Spielman and Srivastava [MSS] combined this with a delicate use of interlacing polynomials to solve the full Kadison-Singer problem.

References

[A]   Anderson, Restrictions and representations of states on c^*-algebras, *Trans. AMS* **249**, 1979(), 303–329.

[ABP]   M. Anttila, K. M. Ball and I. Perissinaki, The central limit problem for convex bodies, *Trans. AMS* **355**, (2003), 4723–4735.

[AR]   Y. Aumann and Y. Rabani, An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm, *J. Comput.* **27**, (1998), 291–301.

[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.

[B3]   K. M. Ball, The Ribe Programme, *Séminaire Bourbaki* **1047**, (2012).

[BN]   K. M. Ball and V. H. Nguyen, Entropy jumps for isotropic log-concave random vectors and spectral gap, *Studia Math.* , (2012), 81–96.

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

[BSS]   J. Batson, D.A. Spielman and N. Srivastava, Twice Ramanujan sparsifiers, *SIAM J. Comput.* **41**, (2012), 1704–1721.

[Bla]   W. Blaschke, Über affine Geometrie VII: Neue Extremeingenschaften von Ellipse und Ellipsoid, *Ber. Verh. Sächs. Akad. Wiss. Math. Phys. Kl.* **69**, (1917), 412–420.

[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.

[Bo3]   J. Bourgain, On the distribution of polynomials on high dimensional convex sets, *Lecture Notes in Math.* **1469**, (1991), 127–137, Springer-Verlag.

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

[BM]   J. Bourgain and V. Milman, New volume ratio properties for convex symmetric bodies in $\mathbf{R}^n$, *Inventiones Math.* **88** (1987), 319-340.

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

[BT]   J. Bourgain and L. Tzafriri, Invertibility of "large" submatrices with applications to the geometry of Banach spaces and harmonic analysis, *Israel J. Math.* **57** (1987), 137–224.

[BV] U. Brehm and J. Voigt, Asymptotics of cross sections for convex bodies, *Beitr. Alg. Geom.* **41**, (2000), 437–454.

[CT] P. Casazza and J. Tremain, Revisiting the Bourgain-Tzafriri restricted invertibility theorem, *Operators and Matrices* **3**, (2009), 97–110.

[D] A. Dvoretzky, A theorem on convex bodies and applications to Banach spaces, *Proc. Natl Acad. Sci.* **45**, (1959), 223–226.

[El] R. Eldan, Thin shell implies spectral gap up to a polylog via a stochastic localization scheme, *GAFA* **23**, (2013), 532–569.

[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.

[F] U. Feige, Approximating the bandwidth via volume respecting embeddings, 30th Annual ACM Symposium on Theory of Computing (Dallas, TX, 1998). *Comput. System Sci.* **60**, (2000), 510–539.

[FGP] B. Fleury, O. Guédon, G. Paouris, A stability result for mean width of ℓ_p-centroid bodies, *Adv. Math.* **214**, (2007), 865–877.

[GPV] A. Giannopoulos, G. Paouris and B. Vritsiou, The isotropic position and the reverse Santaló inequality, *Israel J. Math.* **203**, (2014), 1–22.

[GW] M. X. Goemans and D. P. Williamson, .878-approximation algorithms for MAX CUT and MAX 2SAT, *Proc. of the 26th Annual ACM Symposium on Theory of Computing, ACM, New York*, (1994), 422–431.

[G] A. Grothendieck, Résumé de la théorie métrique des produits tensoriels topologiques, *Boll. Soc. Mat. São-Paulo* **8**, (1953), 401–480.

[H] D. Hensley, Slicing convex bodies- bounds for slice area in terms of the bodies' covariance, *Proc. AMS* **79**, (1980), 619–625.

[J] R. C. James, Superreflexive Banach spaces, *Canad. J. Math.* **24**, (1972), 896–904.

[Jo] 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.

[KS]   R. Kadison and I. Singer, Extensions of pure states, *Amer. J. Math.* **81**, (1959), 383–400.

[KLS]   R. Kannan, L. Lovász and M. Simonovits, Isoperimetric problems for convex bodies and a localization lemma, *Disc. Comp. Geom.* **13**, (1995), 541–559.

[Kl1]   B. Klartag, On convex perturbations with a bounded isotropic constant, *GAFA* **16**, (2006), 1274–1290.

[Kl2]   B. Klartag, A central limit theorem for convex sets, *Invent. Math.* **168**, (2007), 91–131.

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

[Ku]   G. Kuperberg, From the Mahler conjecture to Gauss linking integrals, *GAFA* **18**, (2008), 870–892.

[LLR]   N. Linial, E. London and Y. Rabinovich, The geometry of graphs and some of its algorithmic applications, *Combinatorica* **15**, (1995), 215–245.

[LP]   J. Lindenstrauss and A. Pełczyński, Absolutely summing operators in $\mathcal{L}_o$-spaces and their applications, *Studia Math.* **29**, (1968), 275–326.

[LV]   Y. T. Lee and S. Vempala, The Kannan-Lovász-Simonovits Conjecture, preprint.

[MSS]   A. Marcus, D. A. Spielman and N. Srivastava, Ramanujan graphs and the solution of the Kadison-Singer problem, *Proc. ICM* **2**, (2014), 375–386.

[Ma1]   B. Maurey, Théorèmes de factorisation pour les opérateurs linéaires à valeurs dans les espaces L_p, *Astérisque* **11** (1974).

[Ma2]   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 skeletons, *PNAS* **110** (2013), 19,256–19,262.

[Mi1]   V. Milman, A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies, *Funkcional. Anal. i Priložen* **5**, (1971), 28–37. In Russian.

[Mi2]   V. Milman, Random subspaces of proportional dimension of finite dimensional normed spaces: approach through the isoperimetric inequality, Banach Spaces, LNM, volume 1166, (N. Kalton and E. Saab, eds.), (1985) Springer Berlin Heidelberg, 106–115.

[N]   A. Naor, An introduction to the Ribe program, *Japanese J. Math.* **7**, (2012), 167–233.

[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.

[Na]   F. Nazarov, The Hörmander proof of the Bourgain-Milman Theorem, *Lecture Notes in Math.* **2050**, (2012), 335–343, Springer-Verlag.

[Pao]   G. Paouris, Concentration of mass on convex bodies, *GAFA* **16**, (2006), 1021–1049.

[P1]   G. Pisier, The volume of convex bodies and Banch space geometry, *Cambridge Tracts in Math.* **94**, (1997), CUP.

[P2]   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.

[P3]   G. Pisier, Martingales with values in uniformly convex spaces, *Israel J. Math.* **20**, (1975), 326–350.

[P4]   G. Pisier, Grothendieck's Theorem, past and present, *Bull. AMS* **49** (2012), 237–323.

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

[S]   L. A. Santaló, Un invariante afin para los cuerpos convexsos del espacio des n dimensiones, *Portugaliae Math.* **8**, (1949), 155–161.

[Sau]   N. Sauer, On the density of families of sets, *J. Comb. Th. Ser. A* **13**, (1972), 145–147.

[Sh]   S. Shelah, A combinatorial problem; stability and order for models and theories in infinitary languages, *Pacific. J. Math.* **41**, (1972), 247–261.

[SS]   D.A. Spielman and N. Srivastava, An elementary proof of the restricted invertibility theorem, *Israel J. Math.* **190**, (2012), 83–91.

[T]   T. Tao, Exploring the toolkit of Jean Bourgain, preprint.