GAFA, Geom. funct. anal.
Vol. 12 (2002) 183 – 198
1016-443X/02/010183-16 $ 1.50+0.20/0

GAFA Geometric And Functional Analysis

EMBEDDING LEVY FAMILIES INTO BANACH SPACES

M. Rudelson and R. Vershynin

Abstract

We prove that if a metric probability space with a usual concentration property embeds into a finite dimensional Banach space X, then X has a Euclidean subspace of a proportional dimension. In particular this yields a new characterization of weak cotype 2. We also find optimal lower estimates on embeddings of metric spaces with concentration properties into l_∞^k, generalizing estimates of Bourgain–Lindenstrauss–Milman, Carl–Pajor and Gluskin.

1 Introduction

The concentration of measure phenomenon in various classes of probability metric spaces is a remarkable and well-known theme in Geometric Functional Analysis. Discovered by V. Milman, it has been crucial in proofs of many results in the asymptotic theory of finite dimensional normed spaces.

Consider a probability metric space (T, μ, d) with the following concentration property for some positive constant c. For every $\varepsilon > 0$ and every subset A of T of measure at least $1/2$, the ε-neighborhood $A_\varepsilon = \{t \in T : d(t, A) < \varepsilon\}$ has measure at least $1 - 3\exp(-c\varepsilon^2 n)$. Here n is a parameter, usually an integer. If n varies, $n = 1, 2, \dots$, then the family of such metric probability spaces (T_n, μ_n, d_n) is called a *normal Levy family.* The constant 3 is not important in this definition.

The notion of normal Levy family, introduced by M. Gromov and V. Milman [GrM], is by now standard, as it covers many natural families of spaces. Important examples of normal Levy families include the Euclidean spheres $(S^{n-1}, \sigma_n, \rho_n)$ with normalized geodesic distance and normalized Lebesgue measure, the orthogonal groups $(O(n), \mu_n, \rho_n)$ with Hilbert–Schmidt metric and normalized Haar measure, and all homogeneous spaces of $O(n)$, like Stiefel manifolds and Grassmanian manifolds. A

The research of the first author was supported in part by NSF grant DMS-0070458.

remarkable class of discrete normal Levy families is given by the powers $(T^n, \mu^{\otimes n}, d_n)$, where (T, μ) is arbitrary probability space, $\mu^{\otimes n}$ is the product measure and d_n is the normalized Hamming distance. More examples and references can be found in [MiS1] and [T].

Let (T_n, μ_n, d_n) be a normal Levy family. To eliminate trivial cases, where the whole measure is concentrated in one atom or in its small neighborhood, let us assume that the ε-neighborhood of any point in T_n has measure smaller than $1 - \delta$, for some positive ε and δ independent of n. We prove that if (T_n, d_n) can be C-Lipschitz embedded into an n-dimensional Banach space X, then X has a Euclidean subspace of dimension proportional to n. In other words, there exists $k \geq cn$ so that the Euclidean sphere S^{k-1}, which itself is a member of a normal Levy family, linearly embeds into X. This result highlights the importance of the concentration of measure phenomenon in the Euclidean sphere: if *some* metric space with a standard concentration property embeds into X, then so does the Euclidean sphere.

This result gives a new characterization for the Banach spaces of weak cotype 2. Recall that X has weak cotype 2 iff there exist constants c_1, c_2 such that every finite dimensional subspace Y of X contains in turn a subspace c_1-isomorphic to l_2^k with $k > c_2 \dim Y$. In particular, S^{k-1} linearly embeds into Y with k proportional to $\dim Y$. The result stated above yields that if this definition of weak cotype 2 holds for *some* normal Levy family (T_n, μ_n, d_n) instead of the Euclidean spheres S^{k-1}, then it must also hold for the Euclidean spheres, i.e. X must have weak cotype 2.

Our second result states that normal Levy families poorly embed into l_∞^k. If (T_n, μ_n, d_n) is a non-trivial normal Levy family, then for any map $F : T_n \to l_\infty^k$

$$\|F\|_{\mathrm{Lip}} \|F^{-1}|_{F(T_n)}\|_{\mathrm{Lip}} \geq C \sqrt{\frac{n}{\log(2 + k/n)}} \,. \tag{1}$$

In the case when $T = S^{n-1}$ is a Euclidean sphere and F is a linear operator, estimate (1) was proved independently by J. Bourgain, J. Lindenstrauss and V. Milman [BLM], B. Carl, A. Pajor [CP] and E. Gluskin [G2]: any n-dimensional subspace of l_∞^k has distance to l_2^n at least $\psi(k, n)$, where $\psi(k, n)$ denotes the right side of (1). Estimate (1) matches the "isomorphic Dvoretzky theorem" [MiS2] that states that every k-dimensional Banach space contains an n-dimensional subspace $\psi(k, n)$-isomorphic to the Euclidean space, This again distinguishes the family of Euclidean spheres among all normal Levy families: if the n-th term of *some* non-trivial normal

Levy family embeds into l_∞^k with Lipschitz constant K, then so does the Euclidean sphere S^{n-1} with Lipschitz constant proportional to K.

Our approach easily carries over to probability metric spaces with different concentration behavior, including uniformly convex Banach spaces. We make no restrictions on the nature of the metric space besides the measure concentration property. This distinguishes our perspective from the earlier results on embedding finite metric spaces into normed spaces, where different *particular* classes of metric spaces were considered: a generic n point metric space [B], [M1,2], [JLS], certain classes of graphs (expanders, trees) [LiLR], [M2,3].

We thank V. Milman for his encouragement and interest in this work and his numerous and important remarks.

2 Proportional Euclidean Subspaces

Let $T = (T, \mu, d)$ be a probability metric space and A be a subset of T. For $\varepsilon > 0$, the ε-neighborhood of A is defined as $A_\varepsilon = \{x \in T : d(x, A) \leq \varepsilon\}$.

Let c and n be positive numbers. We consider the following concentration property of T:

$$A \subset T,\ \mu(A) \geq \tfrac{1}{2} \implies \mu(A_\varepsilon) \geq 1 - 3\exp(-c\varepsilon^2 n) \quad \text{for all } \varepsilon > 0\,. \tag{2}$$

A family $(T_n, \mu_n, d_n)_{n=1}^\infty$ of probability metric spaces is a *normal Levy family* with constant $c > 0$ if, for each $n = 1, 2, \dots$, property (2) holds for (T_n, μ_n, d_n).

A trivial example of a space satisfying (2) is given by any metric space (T, d) equipped with the measure μ concentrated in one (arbitrary) point of T. To eliminate such degenerate cases we will say that, for $\varepsilon, \delta > 0$, a probability metric space (T, μ, d) is (ε, δ)-*regular* if the ε-neighborhood of any point in T has measure smaller than $1-\delta$. Next, a family $(T_n, \mu_n, d_n)_{n=1}^\infty$ is called regular if there exist $\varepsilon, \delta > 0$ such that each term of the family is (ε, δ)-regular. Note that all the examples of normal Levy families mentioned in the introduction are regular.

We say that a metric space (T, d) is K-Lipschitz embedded into a Banach space X if there exists a mapping $F : T \to X$ and constants A, B such that $A \cdot B \leq K$ and

$$A^{-1} d(x, y) \leq \big\|F(x) - F(y)\big\| \leq B d(x, y)$$

for all $x, y \in T$.

Our first (simple) observation is that if the n-th term of a regular Levy family can be Lipschitz embedded into a Banach space X, then $\dim X \geq Cn$.

PROPOSITION 1. *Let X be a Banach space and (T,μ,d) be an (ε,δ)-regular metric probability space, which satisfies concentration property* (2). *Assume that (T,d) can be K-Lipschitz embedded into X. Then*

$$\dim X \geq C \cdot \left(\tfrac{\varepsilon\delta}{K}\right)^2 \cdot cn\,,$$

Henceforth C, C_1, C_2 denote universal constants. Before the proof, let us recall a key fact that under (2) every 1-Lipschitz function $F : T \to \mathbf{R}$ concentrates:

$$\mu_n\big\{|F - \mathbb{E}F| > \varepsilon\big\} \leq 8\exp(-c\varepsilon^2 n) \quad \text{for all } \varepsilon > 0 \tag{3}$$

(see [MiS1, 6.1.1 and V.4]).

This fact is conveniently described in terms of the norm in the space L_{ψ_2}. The norm of a function f in $L_{\psi_2}(T,\mu)$ is usually defined as the minimal number λ for which $\mathbb{E}\exp(f^2/\lambda^2) \leq 8$. We shall work with the quantity that is equivalent to the L_{ψ_2}-norm and is defined as the minimal number λ such that

$$\mu\{|f| > s\} \leq 8\exp(-s^2/\lambda^2) \quad \text{for all } s > 0\,.$$

We denote this number λ by $\|f\|_{L_{\psi_2}(T_n,\mu_n)}$. We see from (3) that if (T,μ,d) satisfies concentration property (2) then for any mean zero 1-Lipschitz function $f : T_n \to \mathbf{R}$

$$\|f\|_{L_{\psi_2}(T,\mu)} \leq (cn)^{-1/2}. \tag{4}$$

This inequality easily generalizes to Banach space valued functions. Let X be a Banach space, and $F : T \to X$ be a 1-Lipschitz mean zero map. Then for every $x^* \in B_{X^*}$ the function $f : T \to \mathbf{R}$ defined as $f(\omega) = \langle F(\omega), x^*\rangle$, $w \in T$, is a 1-Lipschitz scalar valued mean zero function. Therefore under (2) one has

$$\big\|\langle F, x^*\rangle\big\|_{L_{\psi_2}(T,\mu)} \leq (cn)^{-1/2} \quad \text{for every } x^* \in B_{X^*}\,. \tag{5}$$

Proof of Proposition 1. Let $F : T \to X$ be a K-Lipschitz embedding. Consider the map $\Phi = (F - \mathbb{E}F)/\|F\|_{\mathrm{Lip}}$. Then Φ is mean zero, 1-Lipschitz, and $\|\Phi^{-1}|_{\Phi(T)}\|_{\mathrm{Lip}} \leq K$.

Let $k = \dim X$. Then there exists a 1/2-net of B_{X^*} of cardinality 5^k, i.e. there is a set $\mathcal{N} \subset B_{X^*}$ such that

$$B_{X^*} \subset \mathcal{N} + \tfrac{1}{2}B_{X^*} \quad \text{and} \quad |\mathcal{N}| \leq 5^k$$

(see [MiS1, 2.6]). By (5),

$$\big\|\big\langle \sqrt{cn}\Phi, x\big\rangle\big\|_{L_{\psi_2}(T,\mu)} \leq 1 \;\text{ for all } x^* \in \mathcal{N}\,.$$

This implies by the standard estimate (see [LT] (3.13)) that

$$\mathbb{E}\max_{x^*\in\mathcal{N}} \big\langle \sqrt{cn}\Phi, x\big\rangle \leq C\big(\log|\mathcal{N}|\big)^{1/2} \leq C_1\sqrt{k}\,.$$

Therefore

$$\begin{aligned}\mathbb{E}\|\Phi\|_X &= \mathbb{E}\sup_{x^*\in B_{X^*}}\langle\Phi,x\rangle \le \mathbb{E}\sup_{x^*\in\mathcal{N}}\langle\Phi,x\rangle + \mathbb{E}\sup_{x^*\in\frac{1}{2}B_{X^*}}\langle\Phi,x\rangle\\ &\le C_1\sqrt{\tfrac{k}{cn}} + \tfrac{1}{2}\mathbb{E}\|\Phi\|_X\,.\end{aligned}$$

This shows that

$$\mathbb{E}\|\Phi\|_X \le 2C_1\sqrt{\tfrac{k}{cn}}\,.$$

On the other hand, let Φ' be an independent copy of Φ, then using the (ε,δ)-regularity assumption we have

$$\begin{aligned}\mathbb{E}\|\Phi\|_X &\ge \frac{1}{2}\mathbb{E}\|\Phi-\Phi'\|_X \ge \frac{1}{2K}\int_{T\times T} d(\omega,\omega')d\mu(\omega)d\mu(\omega')\\ &\ge \frac{\varepsilon\delta}{2K}\,.\end{aligned}$$

Hence $2C_1\sqrt{k/cn} \ge \varepsilon\delta/2K$, which proves the proposition. □

If, under the hypotheses of Proposition 1, the dimension of X is n, then X necessarily contains an Euclidean subspace of proportional dimension. This is the main result in this section. We denote by $k(X)$ the maximal dimension of a subspace of X which is 2-isomorphic to l_2^k.

Theorem 2. *Let X be an n-dimensional Banach space and let (T,μ,d) be an (ε,δ)-regular metric probability space, which satisfies the concentration property* (2)*. Assume that (T,d) can be K-Lipschitz embedded in X. Then*

$$k(X) \ge C\cdot\left(\tfrac{\varepsilon\delta}{K}\right)^4\cdot cn\,.$$

The proof is again based on estimate (5), which motivates the following definition. Let (T,μ,d) be a probability metric space and X be a Banach space. We will call a map $F:T\to X$ a *subgaussian random vector* if

$$\big\|\langle F,x^*\rangle\big\|_{L_{\psi_2}(T,\mu)} \le 1 \quad \text{for every } x^*\in B_{X^*}\,. \tag{6}$$

A "canonical" example of a subgaussian random vector is the standard Gaussian vector g in $(\mathbf{R}^n,\|\cdot\|_X)$, whose coordinates are independent $N(0,1)$ random variables, under the assumption that the unit Euclidean ball B_2^n is contained in B_X. In this case, up to an absolute constant, $\frac{1}{\sqrt{n}}\mathbb{E}\|g\|_X \sim M_X$, where $M_X = \int_{S^{n-1}}\|x\|_X\,d\sigma_n(x)$. The example considered is in some sense extremal. Namely, we prove the following result, which may be of independent interest.

Theorem 3. *Let $X=(\mathbf{R}^n,\|\cdot\|_X)$ be a Banach space such that the maximal volume ellipsoid in B_X is the standard Euclidean ball B_2^n. Let F be a subgaussian random vector in X, i.e.* (6) *holds. Then*

$$\tfrac{1}{\sqrt{n}}\mathbb{E}\|F\|_X \le CM_X^{1/2}.$$

Proof. We write the average of $\|F\|_X$ as the expectation of the supremum of a subgaussian process:

$$\mathbb{E}\|F\|_X = \mathbb{E} \sup_{x^* \in B_{X^*}} \langle F, x^* \rangle .$$

We will cover B_{X^*} by translates of small Euclidean balls. Fix an $\varepsilon > 0$. By Sudakov's inequality ([LT, 3.3]),

$$\big(\log N(B_{X^*}, B_2^n, \varepsilon)\big)^{1/2} \le C\varepsilon^{-1}\mathbb{E}\|g\|_X ,$$

where $N(B_{X^*}, B_2^n, \varepsilon)$ denotes the minimal number of translates of εB_2^n needed to cover εB_{X^*}. Let $\mathcal{N}$ be the set of points guaranteed by this entropy bound, i.e.

$$B_{X^*} \subset \mathcal{N} + \varepsilon B_2^n \quad \text{and} \quad (\log|\mathcal{N}|)^{1/2} \le C\varepsilon^{-1}\mathbb{E}\|g\|_X .$$

Then

$$\mathbb{E} \sup_{x^* \in B_{X^*}} \langle F, x^* \rangle \le \mathbb{E} \sup_{x^* \in \mathcal{N}} \langle F, x^* \rangle + \varepsilon \cdot \mathbb{E} \sup_{x^* \in B_2^n} \langle F, x^* \rangle . \tag{7}$$

We can certainly assume that $\mathcal{N} \subset B_{X^*}$, so by (6) we have

$$\big\| \langle F, x^* \rangle \big\|_{L_{\psi_2}(T,\mu)} \le 1 \quad \text{for all } x^* \in \mathcal{N} . \tag{8}$$

Then (8) alone implies that $\mathbb{E}\sup_{x^* \in \mathcal{N}} \langle F, x^* \rangle \le C(\log|\mathcal{N}|)^{1/2}$ (see [LT, (3.13)]). Thus the first summand in (7) is bounded by

$$C\varepsilon^{-1}\mathbb{E}\|g\|_X .$$

Next, the second summand in (7) is

$$\varepsilon \cdot \mathbb{E} \sup_{x^* \in B_2^n} \langle F, x^* \rangle = \varepsilon \cdot \mathbb{E}\|F\|_2$$

(by $\|\cdot\|_2$ we denote the Euclidean norm in $\mathbf{R}^n$). We will use John's decomposition of the identity on X. Namely, since B_2^n is the maximal volume ellipsoid inscribed in B_X, then the identity operator on X can be decomposed as

$$id_X = \sum_{j=1}^m x_j \otimes x_j ,$$

where $x_j/\|x_j\|_X$ are some contact points between the surfaces of B_2^n and B_X, i.e. $\|x_j\|_X = \|x_j\|_{X^*} = \|x_j\|_2$, $j = 1, \dots, m$, and $\sum_{j=1}^m \|x_j\|_2^2 = n$. We have

$$\begin{aligned}
\mathbb{E}\|F\|_2 \le (\mathbb{E}\|F\|_2^2)^{1/2} &= \left(\mathbb{E}\sum_{j=1}^m \langle F, x_j\rangle^2\right)^{1/2} \\
&= \left(\sum_{j=1}^m \|\langle F, x_j\rangle\|_{L_2(T,\mu)}^2\right)^{1/2} \\
&\le C\left(\sum_{j=1}^m \|\langle F, x_j\rangle\|_{L_{\psi_2}(T,\mu)}^2\right)^{1/2} \\
&\le C\left(\sum_{j=1}^m \|x_j\|_{X^*}^2\right)^{1/2} \qquad \text{by (6)} \\
&= C\left(\sum_{j=1}^m \|x_j\|_2^2\right)^{1/2} = Cn^{1/2}.
\end{aligned}$$

As a consequence, (7) is bounded by

$$C\big(\varepsilon^{-1}\mathbb{E}\|g\|_X + \varepsilon n^{1/2}\big)\,.$$

Therefore

$$\frac{1}{\sqrt{n}}\mathbb{E}\|F\|_X \le C\left(\varepsilon^{-1}\frac{\mathbb{E}\|g\|_X}{\sqrt{n}} + \varepsilon\right) \le C(\varepsilon^{-1}M_X + \varepsilon)\,.$$

Now taking $\varepsilon = M_X^{1/2}$ we obtain the required estimate. The proof is complete. $\square$

It is a difficult question whether the estimates obtained are optimal. In particular, it is natural to ask whether the bound $CM_X^{1/2}$ can be improved to CM_X. One of the ways to improve it is to replace the maximal volume ellipsoid by an ellipsoid $\mathcal{E} \subset B_X$ for which M_X is maximal. An example in the paper by A. Giannopoulos, V. Milman and M. Rudelson [GiMR] shows that this ellipsoid can be very far from that of the maximal volume.

Now we derive Theorem 2 from Theorem 3.

Proof of Theorem 2. We can assume B_2^n is the ellipsoid of maximal volume contained in the unit ball of X. By [MiS1, 4.2],

$$k(X) \ge CM_X^2 n\,, \tag{9}$$

Now Theorem 3 will provide a lower bound for M_X.

Indeed, let $F : T \to X$ be a K-Lipschitz embedding. Defining $\Phi = (F - \mathbb{E}F)/\|F\|_{\mathrm{Lip}}$ we see that Φ is mean zero and 1-Lipschitz, so by (5) $c\sqrt{n}\Phi$ is a subgaussian random vector. By Theorem 3

$$\mathbb{E}\|\Phi\|_X \le CM_X^{1/2}$$

and, as in the proof of Proposition 1,

$$\mathbb{E}\|\Phi\|_X \geq \tfrac{\varepsilon\delta}{2K}\,.$$

Hence $CM_X^{1/2} \geq \varepsilon\delta/2K$, which, when combined with (9), completes the proof. □

3 Weak Cotype 2

One of several equivalent definitions of weak cotype 2 of a Banach space X is through a saturation of X by finite dimensional Euclidean subspaces in the following sense. There are constants $\alpha, M > 0$ such that for every n and every n-dimensional subspace Y of X there exist a further subspace $Z \subset Y$ with $k = \dim Z \geq \alpha n$, which is M-isomorphic to l_2^k (see [P3]).

Our aim is to show that the space l_2^k in this definition can be replaced by the k-th term of any regular normal Levy family. The linear embedding is replaced naturally by a Lipschitz embedding or, more generally, by a semi-Lipschitz embedding.

DEFINITION 4. *Let (T_n, d_n) be a sequence of metric spaces, and X be a Banach space. Suppose we have for each n a one-to-one map $F_n : T_n \to X$. We call the family (F_n)* a semi-Lipschitz embedding *if*

(i) $\sup_n \|F_n\|_{\mathrm{Lip}} < \infty$;
(ii) *The family of maps (F_n^{-1}) defined on the images of F_n is equicontinuous.*

We say that a family of metric spaces (T_n, d_n) *semi-Lipschitz saturates* a Banach space X if there is a constant $\alpha > 0$ such that for every sequence of subspaces (X_n) of X with $\dim X_n \geq \alpha n$ there is a semi-Lipschitz embedding (F_n) of (T_n, d_n) into X so that $F_n(T_n) \subset X_n$ for all n.

Theorem 5. *Suppose X is a Banach space, and there exists a regular normal Levy family which semi-Lipschitz saturates X. Then X has weak cotype 2.*

Proof. Consider a regular normal Levy family (T_n, d_n, μ_n) with regularity constants $\varepsilon, \delta > 0$, which semi-Lipschitz saturates X. Let (X_n) be any subspaces of X with $\dim X_n \geq \alpha n$, and consider the corresponding semi-Lipschitz embedding $(F_n : T_n \to X_n)$. Since the family (F_n^{-1}) is equicontinuous, there exists a $\gamma > 0$ such that

$$\big\|F_n(\omega) - F_n(\omega')\big\|_{X_n} \geq \gamma \quad \text{whenever } d_n(\omega, \omega') \geq \varepsilon\,.$$

Let F_n' be an independent copy of F_n; then

$$\mathbb{E}\|F_n - F_n'\|_{X_n} \geq \gamma \cdot \mu_n \times \mu_n\big\{\|F_n(\omega) - F_n(\omega')\|_{X_n} \geq \gamma\big\}$$

$$\geq \gamma \cdot \mu_n \times \mu_n \big\{ d_n(\omega, \omega') \geq \varepsilon \big\} \geq \gamma\delta \,.$$

Then, as in the proofs of Proposition 1 and Theorem 2, $k(X_n) \geq C(c,\varepsilon,\delta,\gamma)n$. Thus X has weak cotype 2. □

In general, it is impossible to interchange the Lipschitz estimate and the equicontinuity property of F_n and F_n^{-1} in the definition of the semi-Lipschitz embedding. This is illustrated by the next two propositions.

Consider the discrete cube $C_2^n = \{-1, 1\}^n$, endowed with the normalized Hamming metric $d_n(x, y) = \frac{1}{n} |\{i : \ x(i) \neq y(i)\}|$.

Proposition 6. *C_2^n cannot be semi-Lipschitz embedded into any normed space with type $p > 1$.*

Proof. This follows from a result of J. Bourgain, V. Milman and H. Wolfson [BMW]. Assume there exists a semi-Lipschitz embedding $F_n : C_2^n \to X$, where X is a normed space. We will show that X fails to have metric type $p > 1$. We can certainly assume that $\|F_n\|_{\mathrm{Lip}} = 1$. Put $X_\theta = F_n(\theta)$ for all $\theta \in C_2^n$. Since $d_n(\theta, -\theta) = 1$, it follows from the equicontinuity of the family (F_n^{-1}) that $\|X_\theta - X_{-\theta}\|_X \geq \delta$, where δ is some positive constant independent of n. Given a vertex $\theta \in C_2^n$, let $\theta[i]$ be the vertex in C_2^n differing from θ in the i-th coordinate only. The unordered pair $(\theta, \theta[i])$ is called an *edge*. There are $n2^{n-1}$ edges in C_2^n. We have $d_n(\theta, \theta[i]) = 1/n$, hence $\|X_\theta - X_{\theta[i]}\|_X \leq 1/n$. Then

$$D := \Big(\sum_{\theta \in C_2^n} \|X_\theta - X_{-\theta}\|_X^2 \Big)^{1/2} \geq 2^{n/2}\delta$$

and

$$E := \Big(\sum_{\text{edges}} \|X_\theta - X_{\theta[i]}\|_X^2 \Big)^{1/2} \leq (n2^{n-1})^{1/2}(1/n) = \frac{2^{n/2}}{\sqrt{2n}} \,.$$

Assume that X has a metric type $p > 1$. Then by the definition [BMW], $D \leq \alpha n^{1/p - 1/2} E$ for some constant α independent of n. But for n large enough this clearly contradicts to the estimates above. Therefore X has no metric type $p > 1$, and, consequently, it has no type $p > 1$ [BMW]. This completes the proof. □

In particular, C_2^n cannot be semi-Lipschitz embedded into l_p $(1<p<\infty)$. However, we have

Proposition 7. *For $1 < p < \infty$ there is a sequence of mappings $F_n : C_2^n \to l_p^n$ such that*

(i) *The family (F_n) is equicontinuous;*
(ii) $\sup_n \|F_n^{-1}\|_{\mathrm{Lip}} < \infty$.

Proof. Define F_n by $F_n(\theta) = n^{-1/p}\theta$ for $\theta \in C_2^n$. Pick any $\theta, \xi \in C_2^n$. Since the coordinates of θ, ξ are ± 1, we have

$$\big\| F_n(\theta) - F_n(\xi) \big\|_p = n^{-1/p} \|\theta - \xi\|_p = 2(d_n(\theta,\xi))^{1/p}.$$

Then both (i) and (ii) are easily verified. □

4 Embedding into ℓ_∞^k

We begin with a result stating that a Lipschitz map from a normal Levy family into l_∞^k concentrates. Given an k-dimensional normed space X, we denote by $d_\infty(X)$ its Banach-Mazur distance to l_∞^k.

Define the function

$$\varphi(k,n) = \sqrt{\frac{\ln(2 + k/n)}{n}}\,.$$

Theorem 8. *Let (T_n, d_n, μ_n) be a normal Levy family with constant c. Let (X_n) be a sequence of finite dimensional normed spaces with*

$$d_\infty(X_n)\varphi(\dim X_n, n) \to 0 \quad \text{as } n \to \infty\,. \tag{10}$$

Then for any $\varepsilon > 0$ there exists a number n_0 (depending on c, ε, and the rate of convergence in (10)) such that for $n > n_0$ every 1-Lipschitz map $F : T_n \to X_n$ concentrates:

$$\mu_n\big\{ \|F - \mathbb{E}F\|_{X_n} > \varepsilon \big\} \le 6 \exp\left(-\tfrac{c}{18}\varepsilon^2 n\right).$$

In general, n_0 must depend on ε. This can be seen by a straightforward calculation for the Levy family $T_n = S^{n-1}$ and a sequence of normed spaces $X_n = l_3^n$ by using for every n the map $F_n : T_n \to X_n$ defined as the formal identity on $\mathbf{R}^n$ restricted to S^{n-1}.

We postpone the proof of Theorem 8 to the end of the section, and discuss some of its conclusions. First, Theorem 8 applies to Lipschitz embeddings.

Theorem 9. *Let (T_n, d_n, μ_n) be a regular normal Levy family. Then any map F from T_n into l_∞^k satisfies*

$$\|F\|_{\mathrm{Lip}} \cdot \|F^{-1}|_{F(T_n)}\|_{\mathrm{Lip}} \ge C(c,\varepsilon,\delta)\varphi(k,n)^{-1}\,, \tag{11}$$

where c, ε, δ are the constants from the definition of the regular normal Levy family.

In particular, letting $T_n = S^{n-1}$ we get

Corollary 10. *Let X be an n-dimensional subspace of l_∞^k. Then $d(X, l_2^n) \ge C\varphi(k,n)^{-1}$, where c is an absolute constant.*

Equivalently, let X be an n-dimensional normed space whose unit ball has at most k extreme points. Then $d(X, l_2^n) \ge C\varphi(k,n)^{-1}$.

The estimates are exact up to absolute constants, see [G1,2]. Moreover, *any* k-dimensional normed space Y has an n-dimensional subspace X, $n \geq \log k$, satisfying $d(X, l_2^n) \leq C\varphi(k,n)^{-1}$ [MiS2].

Corollary 10 was obtained by different methods by Bourgain, Lindenstrauss and Milman ([BLM, Corollary 9.5]), Carl and Pajor ([CP, Theorem 3.3]) and Gluskin ([G2, Corollary 3]). More generally, using Maurey's observation (see [P2]) Carl and Pajor obtained a precise bound for the Banach-Mazur distance from a subspace of l_∞^k to a space whose dual satisfies a type condition. Namely ([CP, Theorem 4.2]):

If $X \subset l_\infty^k$ is an n-dimensional subspace and Y is an n-dimensional Banach space whose dual has type $p > 1$ constant $T_p(Y^*) = K$, then $d(X,Y) \geq (c/K)(n/\log(2+k/n))^{1/p}$, where $1/p + 1/p' = 1$ and $c > 0$ is a universal constant.

It might be of interest that our approach gives the same bound in terms of the modulus of uniformly convexity, which is however a stronger assumption.

COROLLARY 11. *Let X be an n-dimensional subspace of l_∞^k, and let Y be a n-dimensional space whose modulus of uniform convexity satisfies $\delta_Y(t) \leq Kt^p$ for $t > 0$. Then $d(X,Y) \geq c(n/\log(2+k/n))^{1/p}$, where c depends on K only.*

The proof follows from a modification of our technique to "non-normal Levy families" through the following result of Gromov and Milman. Under the assumption on space Y in Corollary 11 there exists a (natural) probability measure on the unit sphere S_Y such that any 1-Lipschitz function $F : S_Y \to \mathbf{R}$ concentrates,

$$\mathbb{P}\big\{|F - \mathbb{E}F| > \varepsilon\big\} \leq 4\exp(-c\varepsilon^p n) \quad \text{for } \varepsilon > 0\,.$$

Now we pass to the proofs. The key tool will be the following lemma.

LEMMA 12. *Let (T_n, d_n, μ_n) be a normal Levy family with constant c, and let $\varepsilon > 0$. Suppose $F : T_n \to l_\infty^k$ is a 1-Lipschitz function. Then there is a set $A \subset T_n$ such that*

$$\mu_n(A) \geq \tfrac{1}{2}\exp(-c\varepsilon^2 n) \ \ \textit{and} \ \ \operatorname{diam}(F(A)) \leq C(c,\varepsilon)\varphi(k,n)\,.$$

Proof. One can assume that $k/n \geq e$ by embedding l_∞^k into some $l_\infty^{k'}$ with some larger k'. So we can substitute $\varphi(k,n)$ by $\sqrt{\ln(k/n)/n}$ in the statement of the lemma. Let $t = t(c,\varepsilon) > 0$ be a parameter, which will be specified later. Write $F = (f_1, \ldots, f_k)$, where all f_i are real valued 1-Lipschitz functions; we can also assume that they are mean zero.

Define a map $T : T_n \to \mathbb{Z}^k$ as follows: $T(\omega) = (b_1(\omega), \dots, b_k(\omega))$, where

$$b_i(\omega) \ \text{ is the nearest integer to } \ \frac{f_i(\omega)}{t\varphi(k,n)}\,.$$

That is

$$b_i(\omega) = s \qquad \Longleftrightarrow \qquad s - 1/2 < \frac{f_i(\omega)}{t\varphi(k,n)} \le s + 1/2\,.$$

Then for every $i \in \{1, \dots, k\}$ and $s \in \mathbb{Z}_+$

$$\begin{aligned}\mu_n\big\{|b_i(\omega)| \ge s\big\} &= \mu_n\left\{s - 1/2 < \frac{f_i(\omega)}{t\varphi(k,n)} \ \text{ or } \ \frac{f_i(\omega)}{t\varphi(k,n)} \le -s + 1/2\right\}\\ &\le \mu_n\big\{|f_i(\omega)| \ge (s-1/2)t\varphi(k,n)\big\}\\ &\le 8\exp\big(-c(s-1/2)^2t^2\varphi(k,n)^2n\big) =: P\,,\end{aligned}$$

by (3). Since expectation is linear,

$$\mathbb{E}\big|\{i : |b_i(\omega)| \ge s\}\big| \le kP\,.$$

Set

$$\alpha_s = P \cdot 2^{s+1}\,.$$

By Chebyshev's inequality

$$\mu_n\big\{|\{i : |b_i(\omega)| \ge s\}| \ge k\alpha_s\big\} \le 1/2^{s+1}.$$

Now we define a set $B \subset \mathbb{Z}^n$ by

$$(b_1, \dots, b_n) \in B \qquad \Longleftrightarrow \qquad \big|\{i : |b_i| \ge s\}\big| \le k\alpha_s \ \text{ for all } s \in \mathbb{Z}_+\,.$$

Then

$$\mu_n\big\{T(\omega) \in B\big\} \ge 1 - \sum_{s=1}^{\infty} 1/2^{s+1} = 1/2\,. \tag{12}$$

Claim. $\quad |B| \le \exp(\varepsilon^2 n)$.

Once the claim is proved, we can apply the pigeonhole principle to (12). There exists a set $A \subset T_n$ with $\mu_n(A) \ge \frac{1}{2}\exp(-\varepsilon^2 n)$ such that $T(A)$ is a singleton. This will clearly complete the proof of the lemma.

By the definition,

$$\alpha_s = 8 \cdot 2^{s+1}(k/n)^{-c(s-1/2)^2t^2}.$$

Now we proceed by a counting argument from [Sp]. By taking t large enough we can assume that $1/2 > \alpha_1 > \alpha_2 > \dots$ Then

$$|B| \le \prod_{s=1}^{\infty}\left[\left(\sum_{i=0}^{k\alpha_s}\binom{k}{i}\right)2^{k\alpha_s}\right].$$

Indeed, $\{i : |b_i| = s\}$ can be chosen in at most $\sum_{i=0}^{k\alpha_s}\binom{k}{i}$ ways, and, having been selected, can be split into $\{i : b_i = s\}$ and $\{i : b_i = -s\}$ in at most

$2^{k\alpha_s}$ ways. We bound

$$\sum_{i=0}^{k\alpha_s} \binom{k}{i} \le 2^{kH(\alpha_s)},$$

where $H(\alpha) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha)$, see [Ch]. Therefore

$$|B| \le 2^{bk}, \quad \text{where } b = \sum_{s=1}^{\infty} \big(H(\alpha_s) + \alpha_s\big). \tag{13}$$

Note that $\alpha_{s+1} \le \gamma\alpha_s$ and $\alpha_s \le \alpha_1 \le \gamma$ for all $s \ge 1$, where $\gamma = 32\exp(-ct^2/4) \le 1/100$ (by an appropriate choice of t). Then

$$H(\alpha_{s+1}) + a_{s+1} \le 3\gamma\big(H(\alpha_s) + a_s\big) \quad \text{for all } s \ge 1.$$

This yields that the series for b in (13) is dominated by the first term:

$$\begin{aligned} b &\le \frac{H(\alpha_1) + \alpha_1}{1 - 3\gamma} \\ &\le -3\alpha_1 \log_2(\alpha_1) \qquad \text{since } \alpha_1 \le \gamma \le 1/100, \\ &= 24ct^2(k/n)^{-ct^2/4} \log_2(32k/n). \end{aligned} \tag{14}$$

Then $|B| \le \exp(\beta n)$, where $\beta = (\ln 2)(k/n)b$. If $t = t(c,\varepsilon)$ is chosen large enough, then $\beta \le c\varepsilon^2$, since $k/n \ge e$. This proves the Claim and completes the proof of the lemma. □

Next, we recall a standard but quite useful observation, which is successfully applied for normal Levy families, see e.g. [AM]. For a probability metric space (T,μ,d) we define the concentration function as

$$\alpha(T,\varepsilon) = 1 - \inf\big\{\mu(A_\varepsilon) : A \subset T \text{ with } \mu(A) \ge 1/2\big\}, \quad \varepsilon > 0.$$

So, if (T_n,μ_n,δ_n) is a normal Levy family with constant c, then $\alpha(T_n,\varepsilon) \le 3\exp(-c\varepsilon^2 n)$ for all $n = 1,2,\dots$ and $\varepsilon > 0$.

Lemma 13. *Let $\varepsilon > 0$. Let (T,d,μ) be a metric probability space. If $A \subset T$ with $\mu(A) > \alpha(T,\varepsilon/2)$, then $\mu(A_\varepsilon) \ge 1 - \alpha(T,\varepsilon/2)$.*

In particular, for any subset $A \subset T$ and every $\varepsilon > 0$

$$\mu(A_\varepsilon) \ge 1 - \frac{\alpha(T,\varepsilon/2)}{\mu(A)}.$$

Proof. Let $\mu(A) > \alpha(T,\varepsilon/2)$. We claim that $\mu(A_{\varepsilon/2}) \ge 1/2$. Assume the converse. That is, assume $\mu(A_{\varepsilon/2})^c > 1/2$. Then $\mu((A_{\varepsilon/2})^c)_{\varepsilon/2} \ge 1 - \alpha(T,\varepsilon/2)$. Clearly, $((A_{\varepsilon/2})^c)_{\varepsilon/2} \cap A = \emptyset$, thus $\mu(A) \le \alpha(T,\varepsilon/2)$. This contradicts the assumption and proves the claim. Now $\mu(A_\varepsilon) = \mu(A_{\varepsilon/2})_{\varepsilon/2} \ge 1 - \alpha(T,\varepsilon/2)$. This proves the first statement of the lemma. The second statement follows from the first one. □

Theorem 14. *Let (T_n, d_n, μ_n) be a normal Levy family with constant c, and let $\varepsilon > 0$. Let X be a k-dimensional Banach space, and $F : T_n \to X$ be a 1-Lipschitz map. Then*

$$\mu_n\big\{\|F - \mathbb{E}F\|_X > \varepsilon + \bar{C}(c,\varepsilon) d_\infty(X)\varphi(k,n)\big\} \le 6\exp\left(-\tfrac{c}{17}\varepsilon^2 n\right). \tag{15}$$

Proof. Denote $\bar{C} = C(\varepsilon/10, c)$ where $C(\varepsilon, c)$ is a constant from Lemma 12. We get a set A from Lemma 12 so that

$$\mu_n(A) \ge \tfrac{1}{2}\exp\big(-c(\varepsilon/10)^2 n\big) \quad\text{and}\quad \operatorname{diam}(F(A)) \le \bar{C} d_\infty(X)\varphi(k,n).$$

Then Lemma 13 gives for $\delta \ge \varepsilon$

$$\mu_n(A_{\delta/2}) \ge 1 - 6\exp\left(-\tfrac{c}{17}\delta^2 n\right). \tag{16}$$

As F is 1-Lipschitz, it stabilizes not only on A but also on A_δ, so that we have for all $\delta \ge \varepsilon$ and for all $\omega \in A_{\varepsilon/2} \subset A_{\delta/2}$

$$\begin{aligned}\mu_n\big\{\omega' : \|F(\omega) - F(\omega')\|_X > \delta + \bar{C} d_\infty(X)\varphi(k,n)\big\} &\le \mu_n((A_{\delta/2})^c)\\ &\le 6\exp\left(-\tfrac{c}{17}\delta^2 n\right).\end{aligned}$$

Then for every $\omega \in A_{\varepsilon/2}$ we bound

$$\begin{aligned}\big\|F(\omega) - \mathbb{E}F\big\|_X &\le \int_{T_n} \big\|F(\omega) - F(\omega')\big\|_X d\mu_n(\omega')\\ &\le \varepsilon + \bar{C} d_\infty(X)\varphi(k,n) + \int_\varepsilon^\infty 6\exp\left(-\tfrac{c}{17}\delta^2 n\right) d\delta\\ &\le \varepsilon + \bar{C} d_\infty(X)\varphi(k,n) + 6\sqrt{\tfrac{17}{cn}}\tfrac{\sqrt{\pi}}{2}\\ &\le \varepsilon + \bar{C}(\varepsilon, c) d_\infty(X)\varphi(k,n)\end{aligned}$$

since $\varphi(k,n) \ge 1/\sqrt{n}$. Then the measure in (15) does not exceed $\mu_n((A_{\varepsilon/2})^c)$ which, in turn, is majorized by (16). This concludes the proof. □

Now Theorem 8 follows immediately from Theorem 14.

Proof of Theorem 9. We can assume that $\|F\|_{\mathrm{Lip}} = 1$.

Choose ε and δ from the definition of regularity. Let $\varepsilon_0 = \varepsilon/4$. Clearly, we may assume that $n > n_0$, where $n_0 = n_0(c, \varepsilon, \delta)$ is large enough. Then we find a set A from Lemma 12 so that

$$\mu_n(A) \ge \alpha(T_n, \varepsilon_0/2)$$

and

$$\operatorname{diam}(F(A)) \le C(c,\varepsilon)\varphi(k,n). \tag{17}$$

Then Lemma 13 yields

$$\begin{aligned}\mu_n(A_{\varepsilon_0}) &\ge 1 - \alpha(T_n, \varepsilon_0/2)\\ &\ge 1 - 3\exp\big(-(c/4)\varepsilon_0^2 n\big) \ge 1 - \delta\end{aligned}$$

provided n_0 was chosen sufficiently large, and $n > n_0$. Then we get from the regularity of (T_n, d_n, μ_n) that $\operatorname{diam}(A_{\varepsilon_0}) \geq \varepsilon$. Thus $\operatorname{diam}(A) \geq \varepsilon - 2\varepsilon_0 = \varepsilon/2$. Together with (17) this gives $\|F^{-1}|_{F(T_n)}\|_{\mathrm{Lip}} \geq (\varepsilon/2)(C(c,\varepsilon)\varphi(k,n))^{-1}$, completing the proof. □

References

[AM] N. Alon, V. Milman, Embedding of l_k^∞ in finite-dimensional Banach spaces, Israel J. Math. 45 (1983), 265–280.

[B] J. Bourgain, On Lipschitz embeddings of finite metric spaces in Hilbert space, Israel J. Math. 52 (1992), 103–111.

[BLM] J. Bourgain, J. Lindenstrauss, V. Milman, Approximation of zonoids by zonotopes, Acta Math. 162 (1989), 73–141.

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

[CP] B. Carl, A. Pajor, Gelfand numbers of operators with values in Hilbert spaces, Invent. Math. 94 (1988), 79–103.

[Ch] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, Ann. Math. Statist. 23 (1952), 493–509.

[G1] E. Gluskin, An octahedron is poorly approximated by random subspaces, Funktsional. Anal. i Prilozhen. 20:1 (1986), 14–20; English transl. in Functional Anal. Appl. 20 (1986), 11-16.

[G2] E. Gluskin, Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces, Math. USSR Sbornik 64 (1989), 85–96.

[GiMR] A.A. Giannopoulos, V.D. Milman, M. Rudelson, Convex bodies with minimal mean width, Geometric Aspects of Functional Analysis 1996-2000, Springer Lecture Notes in Math. 1745 (2000), 81–93.

[GrM] M. Gromov, V. Milman, Generalization of the spherical isoperimetric inequality to uniformly convex Banach spaces, Compositio Math. 62 (1987), 263–282.

[JLS] W. Johnson, J. Lindenstrauss, G. Schechtman, On Lipschitz embedding of finite metric spaces in low dimensional normed spaces, in Geometrical Aspects of Functional Analysis (J. Lindenstrauss and V. Milman, eds.), Springer Lecture Notes in Math. 1267 (1987), 1774-184.

[LT] M. Ledoux, M. Talagrand, Probability in Banach Spaces, A Series of Modern Surveys in Math., 23, Springer Verlag, 1991.

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

[M1] J. Matoušek, On the distortion required for embedding finite metric spaces into normed spaces, Israel J. Math. 93 (1996), 333–344.

[M2] J. Matoušek, On embedding expanders into ℓ_p spaces, Israel J. Math. 102 (1997), 189–197.

[M3] J. Matoušek, On embedding trees into uniformly convex Banach spaces, Israel J. Math. 114 (1999), 221–237.

[MiS1] V. Milman, G. Schechtman, Asymptotic Theory of Finite Dimensional Normed Spaces, Springer Lecture Notes in Math. 1200 (1986).

[MiS2] V. Milman and G. Schechtman, An "isomorphic" version of Dvoretzky's theorem, II, to appear in Convex Geometry, MSRI Publications 34 (1998), 159–164.

[P1] G. Pisier, Martingales with values in uniformly convex spaces, Israel J. Math. 20 (1975), 326–350.

[P2] G. Pisier, Remarques sur un résultat non publie de B.Maurey, Sem. d'Analyse Fonct. 1980/81, Exp. V.

[P3] G. Pisier, The Volume of Convex Bodies and Banach Space Geometry, Cambridge Tracts in Math. 94. Cambridge University Press, Cambridge, 1989

[S] M. Schmuckenschlaeger, A concentration of measure phenomenon on uniformly convex bodies, Geometric Aspects of Functional Analysis (Israel, 1992-1994), Springer Oper. Theory, Adv. Appl. 77 (1995), 275–287.

[Sp] J. Spencer, Six standard deviations suffice, Trans. AMS 289 (1985), 679–706.

[T] M. Talagrand, A new look at independence, Annals of Probability 24 (1996), 1–34.

Mark Rudelson, Department of Mathematics, University of Missouri, Columbia, MO 65211, USA rudelson@math.missouri.edu

Roman Vershynin, Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel vershyn@wisdom.weizmann.ac.il

Submitted: February 2001
Revision: August 2001