Industrial Mathematics Institute

1999:12

On the size of approximately convex sets in normed spaces

S.J. Dilworth, R. Howard and J.W. Roberts

IMI

Preprint Series

Department of Mathematics
University of South Carolina

ON THE SIZE OF APPROXIMATELY CONVEX SETS IN NORMED SPACES

S. J. DILWORTH, RALPH HOWARD AND JAMES W. ROBERTS

Abstract. Let X be a normed space. A set $A \subseteq X$ is *approximately convex* if $d(ta + (1-t)b, A) \leq 1$ for all $a, b \in A$ and $t \in [0,1]$. We prove that every n-dimensional normed space contains approximately convex sets A with $\mathcal{H}(A, \mathrm{Co}(A)) \geq \log_2 n - 1$ and $\mathrm{diam}(A) \leq C\sqrt{n}(\ln n)^2$, where $\mathcal{H}$ denotes the Hausdorff distance. These estimates are reasonably sharp. For every $D > 0$, we construct worst possible approximately convex sets in $C(0,1)$ such that $\mathcal{H}(A, \mathrm{Co}(A)) = \mathrm{diam}(A) = D$. Several results pertaining to the Hyers-Ulam stability theorem are also proved.

Contents

1. Introduction

Let $(X, \|\cdot\|)$ be a normed space. In the following definition $d(x, A) = \inf\{\|x - a\| : a \in A\}$ denotes the distance from x to the set A.

Definition 1.1. A set $A \subseteq X$ is *approximately convex* if

$$d(tx + (1-t)y, A) \leq 1$$

for all $x, y \in A$ and $t \in [0,1]$.

Date: 16 August 1999.

The work of the second author was supported in part by DoD Grant No. N00014-97-1-0806.

Recall that the Hausdorff distance between subsets A and B of X is defined by

$$\mathcal{H}(A,B) = \sup\{d(x,B), d(y,A) : x \in A, y \in B\}.$$

Thus, A is approximately convex if and only if

$$\sup_{t\in[0,1]} \mathcal{H}(tA + (1-t)A, A) \leq 1.$$

The aim of this article is to study the relationship betwen the size of an approximately convex set, as measured by its *diameter*

$$\operatorname{diam}(A) = \sup\{\|x-y\| : x, y \in A\},$$

and the extent to which A fails to be convex, as measured by the Hausdorff distance $\mathcal{H}(A, \operatorname{Co}(A))$ from A to its *convex hull* $\operatorname{Co}(A)$.

In Section 3 we extend some of the results of [6] to the case of approximately convex sets. In particular, it is shown that if X is an n-dimensional normed space then the quantity

$$C(X) = \sup\{\mathcal{H}(A, \operatorname{Co}(A)) : A \subseteq X \text{ is approximately convex}\}$$

satisfies

$$\log_2 n \leq C(X) \leq \lceil \log_2(n+1) \rceil, \tag{1}$$

where $\lceil x \rceil$ denotes the smallest integer $n \geq x$. For the Euclidean spaces $\mathbb{R}^n$, we prove that $C(\mathbb{R}^n) = \log_2 n$ for infinitely many values of n. Thus, the lower bound in (1) is sharp.

We also prove in Section 3 that *every* infinite-dimensional normed space contains an approximately convex set A with $\mathcal{H}(A, \operatorname{Co}(A)) = \infty$. This is used to show that the Hyers-Ulam stability theorem fails rather spectacularly in every infinite-dimensional normed space.

In our previous paper [6] we studied the quantity $\mathcal{H}(A, \operatorname{Co}(A))$ for the class of *approximately Jensen-convex sets* defined as follows.

Definition 1.2. A set $A \subseteq X$ is *approximately Jensen-convex* if

$$d\left(\frac{x+y}{2}, A\right) \leq 1$$

for all $x, y \in A$.

Suppose again that X is an n-dimensional normed space. In the construction of approximately convex sets $A \subseteq X$ presented in Section 3, we find that $\operatorname{diam}(A) \to \infty$ as $\mathcal{H}(A, \operatorname{Co}(A))$ approaches $C(X)$. Section 4 refines this construction to produce such sets whose diameters are not too large in an asymptotic sense as $n \to \infty$. To make this precise, let us say that an approximately convex set A is *bad* if $\mathcal{H}(A, \operatorname{Co}(A)) \geq \log_2 n - 1$. Then our main result says that every

n-dimensional normed space contains bad approximately convex sets of diameter $O(\sqrt{n}(\log n)^2)$. The proof uses a result of Bourgain and Szarek [1] from the local theory of Banach spaces.

In Section 5 we show that the factor $\sqrt{n}$ in the latter result is sharp by demonstrating the *lower* bound $\operatorname{diam}(A) \geq 0.76\sqrt{n}$ for *all* bad approximately convex sets in the Euclidean space $\mathbb{R}^n$ when n is sufficiently large. We also construct nearly extremal approximately convex sets in $\mathbb{R}^n$ of diameter $O(\sqrt{n \log n})$, which is better than our estimate in the general normed space case.

Our constructions uses the clasical *entropy function*

$$E_n(t_1, \ldots, t_{n+1}) = \sum_{i=1}^{n+1} t_i \log_2(1/t_i)$$

defined on the standard n-simplex. In particular, we make heavy use of the fact that E_n is an *approximately convex function.* This observation seems to be new, and we include its short proof in Section 2. As a corollary we obtain the *best constants* in the classical Hyers-Ulam stability theorem [10] when $n+1$ is a power of 2.

The last two sections concern approximately convex sets in infinite-dimensional spaces. The results in Section 6 are in principle not new: they are essentially reformulations of known results of Larsson [11] and of Casini and Papini [3] (also of Bruck [2]). It is shown that X is B-convex if and only if there exists $c > 0$ such that

$$\operatorname{diam}(A) \geq c \exp(c\mathcal{H}(A, \operatorname{Co}(A)))$$

for every approximately convex set $A \subseteq X$. A similar bound with a sharp exponent is given for spaces of type p.

Our deepest and perhaps most interesting result is Theorem 7.1 of Section 7, which says that the trivial inequality

$$\operatorname{diam}(A) \geq \mathcal{H}(A, \operatorname{Co}(A))$$

is actually *best possible* in general Banach spaces. More precisely, we show that for every $M > 0$ there exists a Banach space X (which is isomorphic to ℓ_1) and an approximately convex set $A \subseteq X$ such that

$$\operatorname{diam}(A) = \mathcal{H}(A, \operatorname{Co}(A)) = M.$$

The space X is obtained from a rather complicated combinatorial construction which may conceivably have other applications in Banach space theory. Theorem 7.1 and its proof may be read independently of the rest of the paper.

Finally, a few words about notation. All normed spaces are assumed to be *real*. The closed unit ball $\{x \in X : \|x\| \leq 1\}$ of a normed space

X is denoted B_X. The closed ball of radius R is denoted $B_R(X)$. The dual space of X is denoted X^*. A closed subspace Y of X has a *finite-dimensional decomposition* if there exist finite-dimensional subspaces $F_n \subseteq Y$ $(n \geq 1)$ such that every $y \in Y$ admits a unique representation as a convergent series $y = \sum_{n=1}^{\infty} y_n$ with $y_n \in F_n$. This implies that the finite-dimensional projections $P_n(y) = \sum_{i=1}^{n} y_i$ are uniformly bounded in the operator norm. We write $Y = \sum_{n=1}^{\infty} \oplus F_n$. The sequence spaces ℓ_p, the finite-dimensional spaces ℓ_p^n, the Lebesgue spaces $L_p(0,1)$ $(1 \leq p \leq \infty)$, and the space $C(0,1)$ of all continuous functions on $[0,1]$, are all equipped with their classical norms. More specialized terminology from Banach space theory will be introduced as needed.

2. Approximately convex functions

Hyers and Ulam [10] introduced the notion of an ε-convex function.

Definition 2.1. Let C be a convex subset of X and let $\varepsilon \geq 0$. A function $f : C \to \mathbb{R}$ is ε-convex if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) + \varepsilon \tag{2}$$

for all $x, y \in C$ and $t \in [0,1]$.

Note that if f is ε-convex then the function λf is $\lambda\varepsilon$-convex for each $\lambda > 0$. Thus, ε merely plays the role of a scaling factor. For our results it is convenient to normalize by taking $\varepsilon = 1$ as follows.

Definition 2.2. Let C be a convex subset of X. A function $f : C \to \mathbb{R}$ is *approximately convex* if

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) + 1 \tag{3}$$

for all $x, y \in C$ and $t \in [0,1]$.

For $n \geq 1$, let $\Delta_n = \{t = (t_i)_{i=1}^{n+1} : t_i \geq 0, \sum_{i=1}^{n+1} t_i = 1\}$ be the standard n-simplex. Let e_i $(1 \leq i \leq n+1)$ be the vertices of Δ_n and let $\mathcal{F}_n$ be the collection of all approximately convex functions $f : \Delta_n \to \mathbb{R}$ satisfying $f(e_i) \leq 0$ for $1 \leq i \leq n+1$. Now define

$$\kappa(n) = \sup_{f \in \mathcal{F}_n} \sup_{x \in \Delta_n} f(x). \tag{4}$$

Cholewa [5] (cf. [9]) proved the following sharp version of the famous Hyers-Ulam stability theorem [10].

Theorem A. [5] *Let $U \subseteq \mathbb{R}^n$ be a convex set and let $\varepsilon > 0$. For every ε-convex function $f : U \to \mathbb{R}$ there exist convex functions g and g_0 such that*

$$f(x) \leq g(x) \leq f(x) + \kappa(n)\varepsilon \qquad \text{and} \qquad |f(x) - g_0(x)| \leq \frac{\kappa(n)}{2}\varepsilon. \tag{5}$$

Moreover, $\kappa(n)$ is the sharp constant in (5) *and satisfies the upper bound $\kappa(n) \le k$ for $2^{k-1} \le n < 2^k$, i.e. $\kappa(n) \le \lceil \log_2(n+1) \rceil$.*

Remark 2.3. Lazckovich [12] observed that $\kappa(n)$ is the sharp constant for *every* convex U with nonempty interior.

The following lemma will be used repeatedly.

Lemma 2.4. *Let $f : C \to \mathbb{R}$ be approximately convex, where $C \subset X$ is convex. Suppose that $n \ge 1$ and that $x_1, \dots, x_{n+1} \in C$. Then*

$$f\left(\sum_{i=1}^{n+1} t_i x_i\right) \le \sum_{i=1}^{n+1} t_i f(x_i) + \kappa(n) \tag{6}$$

for all $(t_i)_{i=1}^{n+1} \in \Delta_n$.

Proof. Define $F : \Delta_n \to \mathbb{R}$ by

$$F(t) = f\left(\sum_{i=1}^{n+1} t_i x_i\right) - \sum_{i=1}^{n+1} t_i f(x_i).$$

Then F is approximately convex and $F(e_i) = 0$ for $1 \le i \le n+1$. So $F(t) \le \kappa(n)$ for all $t \in \Delta_n$, which gives (6). □

For our results on approximately convex sets we require a good *lower bound* for $\kappa(n)$: we shall show that $\kappa(n) \ge \log_2(n+1)$, which improves the bound $\kappa(n) \ge (1/2)\log_2(n+1)$ given in [12].

We require the following lemma from [6] concerning the function $\phi(t)$ defined by $\phi(0) = 0$ and $\phi(t) = -t\log_2 t$ ($t \in (0,1]$). For completeness we include the proof.

Lemma 2.5. *For all $t, x, y \in [0,1]$, we have*

$$0 \le \phi(tx + (1-t)y) - t\phi(x) - (1-t)\phi(y) \le \phi(t)x + \phi(1-t)y.$$

Proof. The left-hand inequality just says that ϕ is concave (to see this note that $\phi''(t) = -1/(t\ln 2) < 0$). To prove the right-hand inequality, first consider the case $0 < x \le y \le 1$. For fixed t and y, let

$$\psi(x) = \phi(tx + (1-t)y) - t\phi(x) - (1-t)\phi(y).$$

Then

$$\psi'(x) = \frac{t}{\ln 2}(\ln x - \ln(tx + (1-t)y)) \le 0.$$

Thus $\psi(x)$ is decreasing on $[0, y]$ and attains its maximum at $x = 0$. But

$$\begin{aligned}\psi(0) &= \phi((1-t)y) - (1-t)\phi(y)\\ &= -(1-t)y\log_2((1-t)y) + (1-t)y\log_2 y\\ &= -(1-t)y\log_2(1-t) = \phi(1-t)y.\end{aligned}$$

Thus, if $x \le y$, then

$$\phi(tx+(1-t)y) - t\phi(x) - (1-t)\phi(y) \le \phi(1-t)y \le \phi(t)x + \phi(1-t)y.$$

Similarly, if $y \le x$, then

$$\phi(tx+(1-t)y) - t\phi(x) - (1-t)\phi(y) \le \phi(t)x \le \phi(t)x + \phi(1-t)y.$$

□

The approximately convex sets which we construct in the next section are essentially graphs of the *entropy* functions

$$E_n(t_1, \dots, t_{n+1}) = \sum_{i=1}^{n+1} t_i \log_2(1/t_i) \qquad ((t_i)_{i=1}^{n+1} \in \Delta_n).$$

The following crucial observation seems to be new.

Proposition 2.6. *E_n is a continuous concave approximately convex function on Δ_n. In particular, E_n is approximately affine, i.e.*

$$|E_n(tx+(1-t)y) - tE_n(x) - (1-t)E_n(y)| \le 1 \tag{7}$$

for all $x, y \in \Delta_n$ and $t \in [0,1]$.

Proof. $E_n(t) = \sum_{i=1}^{n+1} \phi(t_i)$ is a sum of concave functions (by Lemma 2.5) and so E_n is concave. For $x = (x_i)_{i=1}^{n+1}$ and $y = (y_i)_{i=1}^{n+1}$ in Δ_n and $t \in [0,1]$, we can use Lemma 2.5 for the first inequality to get

$$\begin{aligned} E_n(tx+(1-t)y) &- tE_n(x) - (1-t)E_n(y) \\ &= \sum_{i=1}^{n+1} (\phi(tx_i + (1-t)y_i) - t\phi(x_i) - (1-t)\phi(y_i)) \\ &\le \sum_{i=1}^{n+1} (\phi(t)x_i + \phi(1-t)y_i) \\ &= \phi(t)\sum_{i=1}^{n+1} x_i + \phi(1-t)\sum_{i=1}^{n+1} y_i \\ &= \phi(t) + \phi(1-t). \end{aligned}$$

The function $\phi(t) + \phi(1-t)$ is concave and symmetric about $t = 1/2$. Thus,

$$\phi(t) + \phi(1-t) \le 2\phi(1/2) = 1,$$

with equality in the last inequality only if $t = 1/2$. □

Remark 2.7. The fact that E_n has the weaker property of being approximately *Jensen-convex* (which corresponds to setting $t = 1/2$ in Definition 2.2) is well-known and has been observed by various authors, e.g. [12].

Note that the following theorem gives the sharp constant in the Hyers-Ulam stability theorem when $n+1$ is a power of 2.

Theorem 2.8. *The constants $\kappa(n)$ satisfy the bounds*

$$\log_2(n+1) \le \kappa(n) \le \lceil \log_2(n+1) \rceil. \tag{8}$$

In particular, $\kappa(n) = \log_2(n+1)$ when $n+1$ is a power of 2.

Proof. The upper bound is due to Cholewa [5]. For the lower bound, since $E_n \in \mathcal{F}_n$, we have

$$\kappa(n) \ge \max_{t \in \Delta_n} E_n(t) = E_n(1/(n+1), \ldots, 1/(n+1)) = \log_2(n+1).$$

□

Remark 2.9. Obviously, $\kappa(1) = 1$. Green [8] showed that $\kappa(2) = 5/3$. In a later paper we shall show that, for $n \ge 1$,

$$\kappa(n) = [\log_2(n+1)] + 2 - \frac{2^{1+[\log_2(n+1)]}}{n+1},$$

where $[x]$ is the greatest integer function. The proof is too long to be included here. The corresponding constants for bounded Jensen-convex functions were computed in [6].

3. Approximately convex sets

Theorem 3.1. *let X be an n-dimensional normed space. There is a least positive constant $C(X)$ such that*

$$\mathcal{H}(A, \mathrm{Co}(A)) \le C(X) \sup_{t \in [0,1]} \mathcal{H}(tA + (1-t)A, A) \tag{9}$$

for every nonempty $A \subseteq X$. Moreover, $C(X)$ satisfies

$$\log_2 n \le C(X) \le \kappa(n). \tag{10}$$

In particular, $\log_2 n \le C(X) \le \lceil \log_2(n+1) \rceil \le \log_2 n + 1$.

Proof. We may assume that the right-hand side of (9) is finite, otherwise there is nothing to prove. Observe that the effect of replacing A by λA is to multiply both sides of (9) by $|\lambda|$. So, by choosing λ appropriately, we may assume that

$$\sup_{t \in [0,1]} \mathcal{H}(tA + (1-t)A, A) = 1.$$

The right-hand estimate for $C(X)$ is due to Casini and Papini [3]. For completeness we recall the proof. Let $f(x) = d(x, A)$ $(x \in X)$. First note that f is 1-Lipschitz and non-negative. To see that f is approximately convex, note that for $x, y \in X$, $a, b \in A$, and $t \in [0,1]$, we have

$$\begin{aligned} f(tx+(1-t)y) &= d(tx+(1-t)y, A) \\ &\le \|(tx+(1-t)y)-(ta+(1-t)b)\| \\ &+ d(ta+(1-t)b, A) \\ &\le t\|x-a\|+(1-t)\|(y-b)\|+1. \end{aligned}$$

Taking the infimum of this expression over all choices of a and b yields

$$f(tx+(1-t)y) \le tf(x)+(1-t)f(y)+1.$$

Now suppose that $x \in \mathrm{Co}(A)$. By Carathéodory's Theorem (see e.g. [18, Thm. 17.1]), $x = \sum_{i=1}^{n+1} t_i a_i$, a convex combination of $n+1$ elements $a_i \in A$. Then Lemma 2.4 yields

$$f(x) \le \sum t_i f(a_i) + \kappa(n) = \kappa(n),$$

since $f(a) = 0$ for all $a \in A$. The left-hand inequality uses the entropy functions E_n. Let $(e_i)_{i=0}^{n-1}$ be an Auerbach basis for X (see e.g. [14, p. 16]). Recall that this means that

$$\max |a_i| \le \left\| \sum_{i=0}^{n-1} a_i e_i \right\| \le \sum_{i=0}^{n-1} |a_i|. \tag{11}$$

for all scalars $a_0, \ldots, a_{n-1}$. Set $e_n = 0$ so that $\mathrm{Co}\{e_i : 1 \le i \le n\}$ is an $(n-1)$-simplex. For each $M > 0$, we define a set A_M thus:

$$A_M = \left\{ M \sum_{i=1}^{n-1} t_i e_i + E_{n-1}(t_1, \ldots, t_n) e_0 : (t_i)_{i=1}^n \in \Delta_{n-1} \right\}$$

First let us verify that A_M is approximately convex. Suppose that $0 \le t \le 1$ and that $a = M \sum_{i=1}^{n-1} x_i e_i + E_{n-1}(x) e_0$ and $b = M \sum_{i=1}^{n-1} y_i e_i + E_{n-1}(y) e_0$ belong to A_M, where $x = (x_i)_{i=1}^n$ and $y = (y_i)_{i=1}^n$ belong to Δ_{n-1}. Then $c = M \sum_{i=1}^{n-1} z_i e_i + E_{n-1}(z) e_0$ also belongs to A_M, where $z = tx + (1-t)y$. Since e_0 is a unit vector and E_{n-1} is *approximately affine* (7), we have

$$\begin{aligned} &\|ta+(1-t)b-c\| \\ &\qquad = |tE_{n-1}(x)+(1-t)E_{n-1}(y)-E_{n-1}(tx+(1-t)y)| \\ &\qquad \le 1, \end{aligned}$$

and so A_M is approximately convex. Note that $x_0 = (M/n) \sum_{i=1}^{n-1} e_i \in \mathrm{Co}(A_M)$. We shall show that $d(x_0, A_M) \to \log_2 n$ as $M \to \infty$. To see this, fix $\varepsilon > 0$. By continuity of E_{n-1} there exists $\alpha > 0$ such that if

$\max_{1\le i\le n-1}|t_i-1/n|\le\alpha$ then $E_{n-1}(t_1,\dots,t_n)\ge\log_2 n-\varepsilon$, whence by (11)

$$\begin{aligned}\left\|x_0-\left(M\sum_{i=1}^{n-1}t_ie_i+E_{n-1}(t_1,\dots,t_n)e_0\right)\right\| &\ge E_{n-1}(t_1,\dots,t_n)\\ &\ge \log_2 n-\varepsilon.\end{aligned}$$

Now suppose, on the other hand, that $\max_{1\le i\le n-1}|t_i-1/n|\ge\alpha$. By (11)

$$\begin{aligned}\left\|x_0-\left(M\sum_{i=1}^{n-1}t_ie_i+E_{n-1}(t_1,\dots,t_n)e_0\right)\right\| &\ge M\max_{1\le i\le n-1}|t_i-1/n|\\ &\ge M\alpha\to\infty\end{aligned}$$

as $M\to\infty$. Thus, for all sufficiently large M, we have $d(x_0,A_M)\ge\log_2 n-\varepsilon$. Since $\varepsilon>0$ is arbitrary, this gives the lower bound $C(X)\ge\log_2 n$. □

For large n the lower bound $C(X)\ge\log_2 n$ is actually attained for certain Euclidean spaces (e.g. for $X=\mathbb{R}^{16}$).

Theorem 3.2. *Suppose that $n=2^k$, where $k\ge 4$. Then $C(\mathbb{R}^n)=\log_2 n$.*

Proof. For $n=2^k$, we have $\kappa(n-1)=\log_2 n$. The argument used to prove Theorem 3.7 of [6] (too lengthy to recall here) shows that the result will follow provided $n=2^k$ is large enough to ensure that

$$\kappa(n-1)\ge\frac{\sqrt{2n}(\sqrt{2n}+\sqrt{n-1})}{n+1}.$$

This holds for $k\ge 4$. □

Remark 3.3. The calculation of $C(\mathbb{R}^n)$ for small n seems problematic. Clearly $C(\mathbb{R})=1$, and examples show that $C(\mathbb{R}^2)>1.37$. In [6] the corresponding constants for approximately Jensen-convex sets in $\mathbb{R}^n$ were computed in all dimensions.

Before turning to infinite-dimensional spaces, let us make the following definition (the analogue of Definition 2.1).

Definition 3.4. Let $\varepsilon>0$. A set $A\subseteq X$ is ε-convex if

$$d(ta+(1-t)b,A)\le\varepsilon$$

for all $a,b\in A$.

Theorem 3.5. *Let X be an infinite-dimensional normed space. There exists an approximately convex set $A\subseteq X$ such that $\mathcal{H}(A,\mathrm{Co}(A))=\infty$.*

Proof. We shall use the following consequence of Theorem 3.1. Let $\varepsilon > 0$ and $M > 0$. Then every normed space of sufficiently large dimension contains a compact ε-convex set A such that $\mathcal{H}(A, \mathrm{Co}(A)) > M$. Using this fact repeatedly, a routine argument (cf. [14, p. 4]) shows that X contains a subspace Y with a *finite-dimensional decomposition* $\sum_{n=1}^{\infty} \oplus F_n$ and sets $A_n \subseteq F_n$ ($n \geq 1$) such that A_n is a 2^{-n}-convex set containing zero and $\mathcal{H}(A_n, \mathrm{Co}(A_n)) > n$. Let A be the collection of all vectors of the form $\sum_n x_n$, where $x_n \in A_n$ and only finitely many of the x_n's are nonzero.

First let us verify that A is approximately convex. Suppose that $x = \sum_n x_n$ and $y = \sum_n y_n$ are in A and that $0 \leq t \leq 1$. Since A_n is 2^{-n}-convex and compact, there exists $z_n \in A_n$ with $\|z_n - (tx_n + (1-t)y_n)\| \leq 2^{-n}$. Moreover, we may choose the z_n's so that only finitely many are nonzero, ensuring that $z = \sum_n z_n$ belongs to A. By the triangle inequality

$$\|z - (tx + (1-t)y)\| \leq \sum_n \|z_n - (tx_n + (1-t)y_n)\| \leq \sum_n 2^{-n} = 1.$$

Let us verify that $\mathcal{H}(A, \mathrm{Co}(A)) = \infty$. Since $\sum_{n=1}^{\infty} \oplus F_n$ is a finite-dimensional decomposition, the natural projection maps from $\sum_{n=1}^{\infty} \oplus F_n$ onto F_n are uniformly bounded in operator norm by K, say. Since $\mathcal{H}(A_n, \mathrm{Co}(A_n)) > n$, there exists $w_n \in \mathrm{Co}(A_n)$ such that $d(w_n, A_n) \geq n$, and since $A_n \subseteq F_n$, we have

$$d(w_n, A) \geq (1/K) d(w_n, A_n) \geq n/K.$$

Thus, $\mathcal{H}(A, \mathrm{Co}(A)) = \infty$. □

As an application of the last result we show that the Hyers-Ulam stability theorem (Theorem A above) fails rather dramatically in *every* infinite-dimensional normed space (cf. [3]).

Corollary 3.6. *Let X be an infinite-dimensional normed space. There exists a 1-Lipschitz approximately convex function $f : X \to \mathbb{R}$ with the following property. For all $M > 0$ there exists $R > 0$ such that for every convex function $g : B_R(X) \to \mathbb{R}$, we have*

$$\sup_{x \in B_R(X)} |f(x) - g(x)| > M.$$

In particular, $\sup\{|f(x) - g(x)| : x \in X\} = \infty$ *for every convex function* $g : X \to \mathbb{R}$.

Proof. Using the notation of Theorem 3.5, we prove that $f(x) = d(x, A)$ has the required property. It was shown in Theorem 3.1 that f is approximately convex and 1-Lipschitz. Choose R so that $\mathrm{Co}(A_n) \subseteq$

$B_R(X)$. Suppose that $g : B_R(X) \to \mathbb{R}$ is a convex function satisfying $|g(x) - f(x)| \leq M$. Since $f(x) = 0$ for all $x \in A_n$, it follows that $g(x) \leq M$ for all $x \in A_n$, and hence $g(x) \leq M$ for all $x \in \mathrm{Co}(A_n)$. But $f(w_n) > n$, and so $M > n/2$. ☐

Recall that a normed space X is *B-convex* if X does not 'contain ℓ_1^n's uniformly', i.e., if there exist $n \geq 2$ and $\alpha > 0$ such that

$$\min_{\pm} \left\| \sum_{i=1}^{n} \pm x_i \right\| \leq n - \alpha$$

for all $x_i \in B(X)$ $(1 \leq i \leq n)$.

For general normed spaces, Corollary 3.6 is close to optimal in view of the following positive result on the approximation of Lipschitz ε-convex functions on *bounded* sets from [4]. (Here (a)⇒(c) is [4, Thm. 1] and (b)⇒(a) is implicit in [4, Props. 1,2]. The other implication (c)⇒(b) is trivial.)

Theorem B. [4] *Let X be a normed space. The following are equivalent:*

(a) *X is B-convex;*
(b) *there exist $k < 1/2$ and $\alpha > 0$ such that for every $\varepsilon < \alpha$ and for every ε-convex 1-Lipschitz function $f : B(X) \to \mathbb{R}$ there exists a convex function $g : B(X) \to \mathbb{R}$ such that*

$$|g(x) - f(x)| \leq k \qquad (x \in B(X));$$

(c) *there exist $c > 0$ and $\alpha > 0$ such that for every $\varepsilon < \alpha$ and for every ε-convex 1-Lipschitz function $f : B(X) \to \mathbb{R}$ there exists a convex function $g : B(X) \to \mathbb{R}$ such that*

$$|g(x) - f(x)| \leq c\varepsilon \log_2(1/\varepsilon) \qquad (x \in B(X)).$$

Remark 3.7. Condition (b) of this result is very pertinent to Section 7 below, where we prove (Corollary 7.13) that for $X = C(0,1)$ there is no constant $k < 1$ such that (b) holds. This is clearly an optimal result since every 1-Lipschitz function f on $B(X)$ satisfies $|f(x) - c| \leq 1$, where $c = (\inf f + \sup f)/2$, i.e. (b) holds for $k = 1$.

4. Diameter of approximately convex sets

Our next goal is to prove that *every* n-dimensional normed space contains a "bad" approximately convex set (that is, $\mathcal{H}(A, \mathrm{Co}(A)) \geq \log_2(n+1) - \varepsilon$) of diameter $O(\sqrt{n}(\log n)^2)$. In the next section we shall prove that for Euclidean spaces this estimate for the diameter is fairly sharp.

For two isomorphic Banach spaces X and Y recall that their Banach-Mazur distance $d(X,Y)$ is defined thus:

$$d(X,Y) = \inf\{\|T\|\|T^{-1}\| : T : X \to Y \text{ is an isomorphism}\}.$$

Theorem 4.1. *Let $\varepsilon \in (0,3)$. For all sufficiently large n and all normed spaces X of dimension n there exists an approximately convex set $A \subseteq X$ such that*

$$\mathcal{H}(A, \mathrm{Co}(A)) \geq \log_2 n - \varepsilon \tag{12}$$

and

$$\mathrm{diam}(A) \leq \frac{25}{\varepsilon}(\log_2 n)^2 d(X, \ell_1^n). \tag{13}$$

Proof. In order to simplify notation we shall prove the result for all normed spaces X of dimension $n+1$ (with $n+1$ replacing n in (12) and (13)). Note that X contains a subspace Z of codimension one such that $d(Z,\ell_1^n) \leq d(X,\ell_1^{n+1})$ (since ℓ_1^{n+1} contains subspaces isometric to ℓ_1^n). Let F be a linear functional in X^* of unit norm such that $Z = \ker(F)$. Let e_0 be an unit vector in X which is normed by F, i.e., such that $F(e_0) = \|e_0\| = 1$. Note that by the triangle inequality

$$\|z + \lambda e_0\| \geq \max(\|z\| - |\lambda|, |\lambda|) \geq \max\left(\frac{\|z\|}{2}, |\lambda|\right) \tag{14}$$

for all $z \in Z$ $(= \ker(F))$ and $\lambda \in \mathbb{R}$. Since $d(Z,\ell_1^n) \leq d(X,\ell_1^{n+1})$, Z has a basis $(e_k)_{k=1}^n$ satisfying

$$\sum_{k=1}^n |a_k| \leq \left\| \sum_{k=1}^n a_k e_k \right\| \leq d(X,\ell_1^{n+1}) \sum_{k=1}^n |a_k|, \tag{15}$$

for all choices of scalars $(a_k)_{k=1}^n$. For each $M > 0$, define $A_M \subseteq X$ by

$$A_M = \left\{ M\left(\sum_{k=1}^n t_k e_k\right) + E_{n-1}(t_1,\ldots,t_n)e_0 : (t_1,\ldots,t_n) \in \Delta_{n-1} \right\}.$$

It was proved in Theorem 3.1 that A_M is approximately convex for all choices of M. Observe also that

$$x_0 = \frac{M}{n}\sum_{k=1}^n e_k \in \mathrm{Co}(A_M).$$

In order to verify (12), it suffices to show that

$$d(x_0, A_M) \geq \log_2(n+1) - \varepsilon$$

for a suitable choice of M. To that end, fix $\alpha \in (0,1)$ and fix $y = M(\sum_{k=1}^n t_k e_k) + (\sum_{k=1}^n t_k \log_2(1/t_k))\, e_0 \in A_M$. Let

$$B_1 = \{k : t_k \geq (1+\alpha)/n\}, \qquad B_2 = \{k : t_k < (1+\alpha)/n\},$$

and set $\mu(B_i) = \sum_{k\in B_i} t_k$ $(i = 1, 2)$. Then, by (14) for the first inequality and the left-hand side of (15) for the second, we have

$$\begin{aligned}
\|y - x_0\| &= \left\| \sum_1^n M(t_k - (1/n))e_k + \left(\sum_{k=1}^n t_k \log_2(1/t_k)\right) e_0 \right\| \\
&\geq \max\left(\frac{1}{2}\left\| \sum_1^n M(t_k - (1/n))e_k \right\|, \sum_{k=1}^n t_k \log_2(1/t_k) \right) \\
&\geq \max\left(\frac{M}{2} \sum_1^n |t_k - (1/n)|, \sum_{k=1}^n t_k \log_2(1/t_k) \right) \\
&\geq \max\left(\sum_{k\in B_1} \frac{M}{2} |t_k - (1/n)|, \sum_{k\in B_2} t_k |\log_2(t_k)| \right) \\
&\geq \max\left(\frac{M}{2}\left(\frac{\alpha}{2}\sum_{k\in B_1} t_k\right), \left(\sum_{k\in B_2} t_k\right)(\log_2 n - \log_2(1+\alpha)) \right)
\end{aligned}$$

(since $t_k - (1/n) \geq (\alpha/(1+\alpha))t_k \geq (\alpha/2)t_k$ for $k \in B_1$)

$$\geq \max\left(\frac{M\alpha}{4}\mu(B_1), (\log_2 n - (3\alpha/2))\mu(B_2) \right),$$

where at the last step we use the fact that $\log_2(1+\alpha) \leq 3\alpha/2$ for $\alpha \in [0, 1]$. Now set $M = 4(\log_2 n)^2/\alpha$. There are two cases to consider. First, if $\mu(B_2) \geq 1 - \alpha/\log_2 n$, then

$$\begin{aligned}
\|y - x_0\| &\geq \left(\log_2 n - \frac{3\alpha}{2}\right)\mu(B_2) \\
&\geq \left(\log_2 n - \frac{3\alpha}{2}\right)\left(1 - \frac{\alpha}{\log_2 n}\right) \geq \log_2 n - \frac{5\alpha}{2}.
\end{aligned}$$

Secondly, if $\mu(B_1) \geq \alpha/\log_2 n$, then

$$\|y - x_0\| \geq \frac{M}{4}\frac{\alpha}{\log_2 n} = \log_2 n.$$

Hence $\|y - x_0\| \geq \log_2 n - (5\alpha/2)$. Setting $\alpha = \varepsilon/3$ we see that (12) is satisfied (with n replaced by $n + 1$) by $A = A_M$ whenever n is large enough to ensure that $\log_2(n + 1) - \log_2(n) \leq \alpha/2$. Finally, the right-hand side of (15) yields

$$\begin{aligned}
\operatorname{diam}(A) &\leq 2d(X, \ell_1^{n+1})M + \log_2 n \\
&\leq d(X, \ell_1^{n+1})(4(\log_2 n)^2/\alpha) + \log_2 n \\
&\leq 25(\log_2 n)^2 d(X, \ell_1^{n+1})/\varepsilon,
\end{aligned}$$

for all sufficiently large n, and so A satisfies condition (13). $\square$

Since $d(\ell_p^n, \ell_1^n) = n^{(p-1)/p}$ for $1 \le p \le 2$, we get the following corollary.

Corollary 4.2. *Let $1 < p \le 2$ and let $\varepsilon \in (0, 3)$. For all sufficiently large n there exists an approximately convex set $A \subseteq \ell_p^n$ such that*

$$\mathcal{H}(A, \mathrm{Co}(A)) \ge \log_2 n - \varepsilon$$

and

$$\mathrm{diam}(A) \le \frac{25}{\varepsilon} n^{(p-1)/p} (\log_2 n)^2.$$

Remark 4.3. For $p = 2$, a much stronger result will be proved in the next section.

For $p = 1$, we may reduce the exponent of $\log_2 n$.

Proposition 4.4. *Let $\varepsilon \in (0, 2)$. For all sufficiently large n there exists an approximately convex set $A \subset \ell_1^n$ such that*

$$\mathcal{H}(A, \mathrm{Co}(A)) \ge \log_2 n - \varepsilon \tag{16}$$

and

$$\mathrm{diam}(A) \le \left(\frac{8}{\varepsilon} + 1\right) \log_2 n. \tag{17}$$

Proof. Setting $X = \ell_1^{n+1}$, we follow Theorem 4.1 taking advantage of some simplifications in the proof which we now indicate. First, we may choose $(e_0, \dots, e_n)$ to be the standard unit vector basis of ℓ_1^{n+1}, so that (14) becomes simply $\|z + \lambda e_0\| = \|z\| + |\lambda|$, for all $z = \sum_{i=1}^n a_i e_i \in Z$. The estimate for $\|y - x_0\|$ then becomes

$$\|y - x_0\| \ge \left(\frac{M\alpha}{2}\mu(B_1) + (\log_2 n - (3\alpha/2))\mu(B_2)\right).$$

Setting $M = 2(\log_2 n)/\alpha$, we obtain

$$\|y - x_0\| \ge \left(\log_2 n - \frac{3\alpha}{2}\right)(\mu(B_1) + \mu(B_2)) = \log_2 n - \frac{3\alpha}{2}.$$

Setting $\alpha = \varepsilon/2$ we see that (16) is satisfied (with n replaced by $n+1$) by $A = A_M$ whenever n is large enough to ensure that $\log_2(n+1) - \log_2(n) \le \alpha/2$. Finally,

$$\mathrm{diam}(A) \le 2M + \log_2 n \le \left(\frac{4}{\alpha} + 1\right) \log_2 n,$$

which yields (17). $\square$

Remark 4.5. In particular, ℓ_1^n contains "bad" approximately convex sets of "small" diameter $O(\log n)$. Indeed, the trivial lower bound

$$\operatorname{diam}(A) \geq \mathcal{H}(A, \operatorname{Co}(A)) \geq \log_2 n - \varepsilon$$

shows that the diameter must grow at least logarithmically with n.

Finally, we come to the main result of this section.

Theorem 4.6. *Let $\varepsilon \in (0,6)$. For all sufficiently large n and all normed spaces X of dimension n there exists an approximately convex set $A \subseteq X$ such that*

$$\mathcal{H}(A, \operatorname{Co}(A)) \geq \log_2 n - \varepsilon \tag{18}$$

and

$$\operatorname{diam}(A) \leq \frac{K(\log_2 n)^2\sqrt{n}}{\varepsilon^3}, \tag{19}$$

where K is an absolute constant.

Proof. Fix $\theta \in (0,1)$. Bourgain and Szarek [1] (cf. also [19]) proved that every n-dimensional normed space X contains a subspace Y, with $\dim Y = k > [\theta n]$, satisfying

$$d(Y, \ell_1^k) \leq C(1-\theta)^{-2}\sqrt{n}, \tag{20}$$

where C is a constant. Set $\theta = 1 - \varepsilon/6$. Then, for $\varepsilon < 1$,

$$\log_2 k \geq \log_2 n - \log_2(1/\theta) \geq \log_2 n - \varepsilon/2. \tag{21}$$

Applying Theorem 4.1 to Y and to $\varepsilon/2$ yields an approximately convex set $A \subseteq Y$ satisfying (18) (from (21) and (12)) and (19) (from (20) and (13)). □

5. Bounds in Euclidean spaces

In this section we prove that the "bad" approximately convex sets constructed in Theorem 4.6 necessarily have diameter larger than $0.76\sqrt{n}$ in n-dimensional Euclidean spaces when n is large. The proof uses only elementary geometry. Along the way we prove a result about Hilbert space (Theorem 5.2) which may be of independent interest because of its sharp constants. We also improve the upper bound of Corollary 4.2 by constructing a nearly extremal approximately convex set in $\mathbb{R}^n$ of diameter $O(\sqrt{n \log n})$.

Recall that a simplex $\Sigma \subseteq \mathbb{R}^n$ is *regular* if its edges all have the same Euclidean length.

Lemma 5.1. *Let Σ be an n-simplex which contains the origin in its interior and whose vertices lie on the Euclidean unit sphere S^{n-1}. For each $0 \le k \le n-1$ there exists a k-face F_k of Σ such that*

$$d(0, F_k) \le \alpha_{n,k} = \sqrt{\frac{n-k}{n(k+1)}},$$

with equality if Σ is a regular simplex.

Proof. First we prove the result for $k = n-1$. Let V be one of the vertices for which the corresponding barycentric coordinate of the origin is at most $1/(n+1)$. Let the line segment through the origin joining V to the opposite $(n-1)$-face F intersect F in a point P, say. Then the origin divides the line joining V to P into two segments bearing a ratio of not less than n to 1. Since V lies on the unit sphere, it follows that $d(0, P) \le 1/n$. Thus, $d(0, F) \le 1/n = \alpha_{n,n-1}$, which completes the proof for the case $k = n-1$.

The proof for $0 < k < n-1$ is by induction on n. Suppose that the result holds for $n-1$ and for $0 < k < n-1$. Let F_{n-1} be an $(n-1)$-face of Σ nearest to the origin and let Q be the point in F_{n-1} nearest to the origin. Then $0 \le d = d(0, Q) = d(0, F_{n-1}) \le 1/n$. The largest Euclidean ball inscribed in Σ with center the origin touches F_{n-1} at Q. Hence Q is in the interior of the $(n-1)$-simplex F_{n-1} whose vertices lie on the $(n-2)$-sphere with center Q and radius $\sqrt{1-d^2}$. Fix $0 < k < n-1$. By the inductive hypothesis applied to Q and F_{n-1} there exists a k-face F_k of F_{n-1} such that

$$d(Q, F_k) \le \alpha_{n-1,k}\sqrt{1-d^2}.$$

So

$$d(0, F_k)^2 = d(0, Q)^2 + d(Q, F_k)^2 \le d^2 + (1-d^2)\alpha_{n-1,k}^2,$$

where $0 \le d \le 1/n$. The right-hand side is greatest when $d = 1/n$, which gives

$$d(0, F_k)^2 \le \frac{1}{n^2} + \left(1 - \frac{1}{n^2}\right)\alpha_{n-1,k}^2 = \alpha_{n,k}^2.$$

□

The following theorem is perhaps of independent interest because of the sharp constants.

Theorem 5.2. *Let $(x_i)_{i=1}^{n+1}$ be elements from the unit ball of a Hilbert space H and suppose that $0 \in \mathrm{Co}(\{x_i : 1 \le i \le n+1\})$. For each*

$1 \leq j \leq n$, *there exists* $J \subseteq \{i : 1 \leq i \leq n+1\}$ *such that* $|J| = j$ *and*

$$d(0, \mathrm{Co}(\{x_i : i \in J\})) \leq \sqrt{\frac{n+1-j}{nj}}.$$

Proof. By slightly perturbing the elements, if necessary, we may assume that the set $\{x_i : 1 \leq i \leq n+1\}$ is affinely independent and that the origin lies in the interior of the simplex $\mathrm{Co}(\{x_i : 1 \leq i \leq n+1\})$. Let $y_i = x_i / \|x_i\|$. Clearly,

$$d(0, \mathrm{Co}(\{x_i : i \in A\})) \leq d(0, \mathrm{Co}(\{y_i : i \in A\}))$$

for all $A \subseteq \{i : 1 \leq i \leq n+1\}$. Now Lemma 5.1 applied to the simplex Σ with vertices $\{y_i : 1 \leq i \leq n+1\}$ yields the desired result. $\square$

Theorem 5.3. *Suppose that* $A \subseteq \mathbb{R}^n$ *is approximately convex and satisfies* $\mathcal{H}(A, \mathrm{Co}(A)) \geq \log_2 n - 1$. *Then, for any integer* j *with* $1 \leq j \leq n$ *we have*

$$\mathrm{diam}(A) \geq \left(\frac{(\log_2 n - 1 - \lceil \log_2 j \rceil)\sqrt{j}}{\sqrt{n-j+1}} \right) \sqrt{n} \tag{22}$$

In particular, A *satisfies the (nontrivial) lower bounds* $\mathrm{diam}(A) \geq 0.7525\sqrt{n}$ *for all* $n \geq 20$, *and* $\mathrm{diam}(A) \geq 0.768\sqrt{n}$ *for all sufficiently large* n.

Proof. Assuming (as we may) that A is compact, there exists $x_0 \in \mathrm{Co}(A)$ with $d(x_0, A) \geq \log_2 n - 1$. By translating A, we may assume that $x_0 = 0$. Thus, $0 \in \mathrm{Co}(A)$ and $d(0, A) \geq \log_2 n - 1$. The fact that $\mathrm{diam}(A) = D$ now implies that $\|x\| \leq D$ for all $x \in A$. By Carathéodory's Theorem, there exist $(x_i)_{i=1}^{n+1}$ in A such that $0 \in \mathrm{Co}(\{x_i : 1 \leq i \leq n+1\})$. Let $1 \leq j \leq n$, then by Theorem 5.2 there exists $J \subseteq \{i : 1 \leq i \leq n+1\}$ such that $|J| = j$ and

$$d(0, \mathrm{Co}(\{x_i : i \in J\})) \leq \left(\sqrt{\frac{n-j+1}{nj}} \right) D = \left(\sqrt{\frac{n-j+1}{j}} \right) \frac{D}{\sqrt{n}}.$$

Let y_0 be the point in $\mathrm{Co}(\{x_i : i \in J\})$ nearest the origin. Because A is approximately convex, the function $d(x, A)$ is an approximately convex function which vanishes at each x_i. So, by Lemma 2.4, $d(y_0, A) \leq \kappa(j-1) \leq \lceil \log_2 j \rceil$ for $2 \leq j \leq n$ and if $j = 1$ then $y_0 \in A$ and $d(y_0, A) = 0 = \lceil \log_2 1 \rceil$. Thus $d(y_0, A) \leq \lceil \log_2 j \rceil$ for $1 \leq j \leq n$. Therefore

$$\log_2 n - 1 \leq d(0, A) \leq \|y_0\| + d(y_0, A) \leq \left(\sqrt{\frac{n-j+1}{j}} \right) \frac{D}{\sqrt{n}} + \lceil \log_2 j \rceil$$

which yields

$$D \geq \frac{\big(\log_2 n - 1 - \lceil \log_2 j \rceil\big)\sqrt{j}}{\sqrt{n-j+1}}\,\sqrt{n} = f(j,n)\sqrt{n}$$

where this defines $f(j,n)$. If k is a non-negative integer with $2^k \leq n$ then $\lceil \log_2 2^k \rceil = k = \log_2(2^k)$. Therefore

$$\begin{aligned} f(2^k,n) &= \frac{\big(\log_2 n - 1 - \log_2 2^k\big)\sqrt{2^k}}{\sqrt{n-2^k+1}} = \frac{\log_2(n/2^k) - 1}{\sqrt{(n/2^k) - 1 + 2^{-k}}} \\ &= F(n/2^k) + r(k,n) \end{aligned}$$

where $F(\alpha) = (\log_2(\alpha) - 1)/(\sqrt{\alpha - 1})$ and $r(k,n) \to 0$ as $n, k \to \infty$. For each n and $\alpha > 0$ there is an integer k so that $\alpha \leq n/2^k \leq 2\alpha$, and for $\alpha_0 = 9.109883742$

$$\alpha_0 \leq \alpha \leq 2\alpha_0 \quad \text{implies} \quad F(\alpha) \geq 0.76811996\ .$$

Therefore if n is sufficiently large and k is chosen so that $\alpha_0 \leq n/2^k \leq 2\alpha_0$ then

$$\max_{1 \leq j \leq n} f(j,n) \geq f(2^k,n) \geq 0.768$$

and thus $D \geq 0.768\sqrt{n}$.

For any n and $k \geq 1$

$$f(2^k,n) \geq \frac{\log_2(n/2^k) - 1}{\sqrt{(n/2^k) - 1 + 1/2^1}} = G(n/2^k)$$

where $G(\beta) = (\log_2(\beta) - 1)/\sqrt{\beta - 1/2}$. If $\beta_0 = 9.919205826$, then $\beta_0 \leq \beta \leq 2\beta_0$ implies $G(\beta) \geq 0.7525$. Now assume that $n \geq 20$, and that $\beta_0 \leq n/2^k \leq 2\beta_0$. Then $1.008 < 20/(2\beta_0) \leq n/(2\beta_0) \leq 2^k$, so $k \geq 1$. Therefore the argument above implies that for $n \geq 20$ the bound $D \geq 0.7525\sqrt{n}$ holds. For this lower bound to be nontrivial we also require $0.7525\sqrt{n} \geq \log_2 n - 1$. However this holds for all $n \geq 1$ and so the lower bound on D holds and is nontrivial for all $n \geq 20$. ☐

Remark 5.4. A similar argument shows that there exists $\varepsilon > 0$ such that if $A \subseteq \mathbb{R}^n$ is approximately convex and satisfies $\mathcal{H}(A, \mathrm{Co}(A)) \geq \log_2 n - \varepsilon$, then $\mathrm{diam}(A) \geq 1.16\sqrt{n}$ for infinitely many n.

Finally, we improve the upper estimate for the diameter provided by Corollary 4.2.

Theorem 5.5. *Let $(e_i)_{i=0}^n$ be the unit vector basis of ℓ_2^{n+1}. Then, for $n \geq 4$ and $M = \sqrt{(2/\ln 2) n \log_2 n}$, the set*

$$A = \left\{ M \sum_{t=1}^{n} t_i e_i + E_{n-1}(t_1, \ldots, t_n) e_0 : (t_1, \ldots, t_n) \in \Delta_{n-1} \right\}$$

is approximately convex and satisfies the following:

$$\mathcal{H}(A, \mathrm{Co}(A)) = \log_2 n \quad \textit{and} \quad \mathrm{diam}(A) \le \frac{2}{\sqrt{\ln 2}}\sqrt{n \log_2 n} + \log_2 n.$$

Remark 5.6. Theorem 5.5 is a significant improvement on Corollary 4.2 as it eliminates the dependence on ε and reduces the exponent of $\log n$ in the estimate for $\mathrm{diam}(A)$. When $n+1 = 2^k$, the set A is very nearly extremal, since in this case $\mathcal{H}(A, \mathrm{Co}(A)) \le \log_2(n+1) = C(\mathbb{R}^{n+1})$ by Theorem 3.2.

The proof of this result is a consequence of the solution to a constrained optimization problem. Consider the following functional:

$$I(y) = M^2 \int_0^n y(x)^2\,dx + \left(\int_0^n \phi(y(x))\,dx\right)^2,$$

where $y(x)$ is a non-negative function defined on the open interval $(0, n)$. (Recall that $\phi(t) = t\log_2(1/t)$.) The problem is to minimize $I(y)$ subject to the following constraints on y:

$$0 \le y \le 1 \quad \text{and} \quad \int_0^n y(x)\,dx = 1.$$

We prove in Lemma 5.10 below that, for $M^2 = (2/\ln 2)n\log_2 n$, $I(y)$ is minimized by $y_0 = (1/n)\chi_{(0,n)}$.

Assuming this result, let us complete the proof of Theorem 5.5.

Proof of Theorem 5.5. Clearly,

$$\mathcal{H}(A, \mathrm{Co}(A)) \le \max_{t\in\Delta_{n-1}} E_{n-1}(t) = \log_2 n.$$

To establish the reverse inequality, we show that $d(x_0, A) = \log_2 n$ for $x_0 = (M/n)\sum_{i=1}^n e_i$. Observe that

$$d(x_0, A)^2 = \min\left\{M^2 \sum_{i=1}^n \left(t_i - \frac{1}{n}\right)^2 + E_{n-1}(t_1, \ldots, t_n)^2 : (t_1, \ldots, t_n) \in \Delta_{n-1}\right\},$$

and also that

$$M^2 \sum_{i=1}^n \left(t_i - \frac{1}{n}\right)^2 + E_{n-1}(t_1, \ldots, t_n)^2 = g(t_1, \ldots, t_n) - \frac{M^2}{n},$$

where

$$g(t_1, \ldots, t_n) = M^2 \sum_{i=1}^n t_i^2 + \left(\sum_{i=1}^n \phi(t_i)\right)^2.$$

Hence

$$d(x_0, A)^2 = \min\{g(t_1, \ldots, t_n) : (t_1, \ldots, t_n) \in \Delta_{n-1}\} - \frac{M^2}{n}.$$

But $g(t_1,\ldots,t_n)=I(\tilde{g})$, where $\tilde{g}(x)=\sum_{k=1}^{n} t_k\chi_{[k-1,k)}$. (Note that $\tilde{g}(x)$ satisfies the constraints for the optimization problem.) Since $I(y)$ is minimized by $y_0=(1/n)\chi_{(0,n)}$ (see Lemma 5.10), we get

$$g(t_1,\ldots,t_n)=I(\tilde{g})\geq I(y_0)=g(1/n,\ldots,1/n).$$

Hence

$$d(x_0,A)^2=g(1/n,\ldots,1/n)-\frac{M^2}{n}=(\log_2 n)^2.$$

Thus, $\mathcal{H}(A,\mathrm{Co}(A))\geq \log_2 n$. The estimate for $\mathrm{diam}(A)$ is straightforward. □

The next four lemmas solve the constrained optimization problem.

Lemma 5.7. *Let $M>0$ and $n\geq 1$. There exists a right-continuous non-increasing function y_0 on $(0,n)$ which solves the constrained optimization problem.*

Proof. Let m be the infimum of $I(y)$ taken over all y which satisfy the constraints. There exist y_n $(n\geq 1)$ satisfying the constraints such that $I(y_n)\to m$ as $n\to\infty$. By replacing each y_n by its *non-increasing rearrangement*, we may assume that each y_n is right-continuous and non-increasing. By Helly's selection theorem (see e.g. [15, p. 221]), we may also assume (by passing to a subsequence) that $y_n(x)\to\tilde{y}_0(x)$ pointwise. Since $0\leq y_n\leq 1$, it follows from the Bounded Convergence Theorem that $\tilde{y}_0$ satisfies the constraints and that $I(\tilde{y}_0)=\lim_n I(y_n)=m$. Finally, let y_0 be the right-continuous modification of $\tilde{y}_0$. □

Lemma 5.8. *There exists $\alpha\in(0,1)$ such that the set of values taken by y_0 is a subset of $\{0,1,\alpha\}$.*

Proof. In the notation of Lemma 5.7, we may assume that y_k is a step function minimizing $I(y)$ over all step functions of the form $\sum_{j=1}^{k} a_j\chi_{[(j-1)n/k,jn/k)}$ satisfying the constraints. A value $\lambda\in(0,1)$ taken by y_k must satisfy the following Lagrange multiplier equation for a local minimum:

$$2M^2\lambda+\left(2\int_0^n \phi(y_k)\,dx\right)\phi'(\lambda)=2A, \tag{23}$$

where A is a constant. It is easily seen that this equation has at most two roots in $(0,1)$. By the pointwise convergence of y_k to $\tilde{y}_0$, it follows that y_0 takes at most two values in $(0,1)$. Therefore we may apply the method of Lagrange multipliers again to deduce that these values must also satisfy (23) (with y_k replaced by y_0). Equivalently, setting $B=\int_0^n \phi(y_0)\,dx>0$,

$$M^2\lambda+B(\log_2(1/\lambda)-(1/\ln 2))=A. \tag{24}$$

Suppose that there are two distinct roots, α and β, with $0<\alpha<\beta<1$, and suppose that y_0 takes one of these values, α say, on an interval J. (The argument is similar if y_0 takes the value β.) Let g take the value 0 on the complement of J, and the values 1 and -1 on the left-hand and right-hand halves of J, respectively. Since $\alpha\in(0,1)$, it follows that $y_0+\varepsilon g$ satisfies the constraints, provided $\varepsilon>0$ is sufficiently small. Moreover,

$$I(y_0+\varepsilon g)=I(y_0)+|J|\left(M^2-\frac{B}{(\ln 2)\alpha}\right)\varepsilon^2+o(\varepsilon^2).$$

Since y_0 minimizes I,

$$M^2-\frac{B}{(\ln 2)\alpha}\geq 0. \tag{25}$$

To derive a contradiction, suppose that y_0 also takes the value β on an interval. Then, by the same argument,

$$M^2-\frac{B}{(\ln 2)\beta}\geq 0.$$

Since (24) is satisfied by $\lambda=\alpha$ and $\lambda=\beta$, the Mean Value Theorem implies the existence of $\gamma\in(\alpha,\beta)$ such that

$$M^2-\frac{B}{(\ln 2)\gamma}=0.$$

Thus,

$$M^2-\frac{B}{(\ln 2)\alpha}<M^2-\frac{B}{(\ln 2)\gamma}=0.$$

But this contradicts (25). Thus, y_0 cannot take the value β, which completes the proof. □

Lemma 5.9. *Suppose that* $n\geq 4$ *and that*

$$5n<M^2\leq\frac{2}{\ln 2}n\log_2 n.$$

Then y_0 *does not take the value* 1.

Proof. For $n\geq 4$, we have

$$I(y_0)\leq I\left(\frac{1}{n}\chi_{[0,n]}\right)=\frac{M^2}{n}+(\log_2 n)^2\leq\frac{2}{\ln 2}\log_2 n+(\log_2 n)^2<\frac{M^2}{2}. \tag{26}$$

Suppose that y_0 takes the value 1 on $[0,x]$ and the nonzero value $k\in(0,1)$ on an interval of length $(1-x)/k\leq n-x$. If $k\geq 1/2$ then

$I(y_0) \geq (1/2)M^2$, which contradicts (26). So we may assume that $k \in (0, 1/2)$. Now

$$I(y_0) = M^2(x + k(1-x)) + ((1-x)\log_2(1/k))^2.$$

So

$$\begin{aligned}\frac{\partial I(y_0)}{\partial x} &= M^2(1-k) - 2\log_2(1/k)^2(1-x) \\ &\geq \frac{M^2}{2} - 2\log_2(1/k)^2 k(n-x)\end{aligned}$$

(since $(1-x) \leq k(n-x)$)

$$\begin{aligned}&\geq \frac{M^2}{2} - 2\left(\max_{0 \leq k \leq 1/2} k\log_2(1/k)^2\right) n \\ &\geq \left(\frac{5}{2} - \frac{8}{e^2(\ln 2)^2}\right) n > 0.\end{aligned}$$

Since $I(y_0)$ minimizes $I(y)$, it follows that $x = 0$, as desired. □

Lemma 5.10. *Suppose that $n \geq 4$ and that $M^2 = (2/\ln 2)n\log_2 n$. Then $y_0 = (1/n)\chi_{(0,n)}$ and $I(y_0) = 2\log_2 n + (\log_2 n)^2$.*

Proof. By Lemma 5.9, y_0 takes only one nonzero value $k \in [1/n, 1)$ on an interval of length $1/k$. So $I(y_0) = M^2 k + (\log_2(1/k))^2$. Thus,

$$\begin{aligned}\frac{\partial I(y_0)}{\partial k} &= M^2 - \frac{2(\log_2(1/k))}{(\ln 2)k} \\ &= \frac{2}{\ln 2}\left(n\log_2 n - \frac{1}{k}\log_2(1/k)\right) \geq 0,\end{aligned}$$

with equality if and only if $k = 1/n$. Since y_0 minimizes $I(y)$, it follows that $k = 1/n$, which gives the result. □

Remark 5.11. Setting $M^2 = 6n$ in Lemma 5.10 yields an approximately convex set set $A \subset \mathbb{R}^{n+1}$ with $\operatorname{diam}(A) = O(\sqrt{n})$ and $\mathcal{H}(A, \operatorname{Co}(A)) \geq \log_2 n - c\log_2\log_2 n$ for some constant c.

6. Lower bounds in spaces of type p

First we recall the notion of *type*. In the following definition $(\varepsilon_i)_{i=1}^{\infty}$ is a sequence of independent Bernoulli random variables, with $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = 1/2$, defined on a probability space (Ω, Σ, P). The *expected value* of a random variable Y is denoted $\mathbb{E}Y$.

Definition 6.1. Let $1 \le p \le 2$. A normed space X is of type p if there exists a constant $T_p(X)$ (the 'type p constant') such that

$$\left(\mathbb{E}\left\|\sum_{i=1}^{n} \varepsilon_i x_i\right\|^p\right)^{1/p} \le T_p(X)\left(\sum_{i=1}^{n} \|x_i\|^p\right)^{1/p}$$

for all $n \ge 1$ and for all choices of $x_i \in X$ $(1 \le i \le n)$.

The following theorem can be deduced from (and in fact is essentially equivalent to) [3, Thm. 3.6]. For completeness we give a short direct proof. We show in Corollary 6.6 below that the exponent of $(p-1)/p$ in this theorem is sharp.

Theorem 6.2. *Let $1 < p \le 2$ and let X be a normed space of type p. Suppose that $A \subseteq X$ is approximately convex. Let $D = \operatorname{diam}(A)$ and let $d = \mathcal{H}(A, \operatorname{Co}(A))$. Then, provided $d \ge 2$, we have*

$$D \ge \frac{8^{1/p}}{16 T_p(X)} (2^d)^{(p-1)/p} \tag{27}$$

Proof. We may assume (cf. Theorem 5.3) that $0 \in \operatorname{Co}(A)$, that $d = d(0, A)$, and that $\|a\| \le D$ for all $a \in A$. Since $0 \in \operatorname{Co}(A)$ there exist $m \ge 1$, $a_i \in A$ and $p_i > 0$ $(1 \le i \le m)$, with $\sum_{i=1}^m p_i = 1$ and $\sum_{i=1}^m p_i a_i = 0$.

Let $(Y_j)_{j=1}^\infty$ be a sequence of independent identically distributed X-valued random variables defined by

$$P(Y_j = a_i) = p_i \qquad (1 \le i \le m).$$

Then $\|Y_j(\omega)\| \le D$ $(\omega \in \Omega)$ and $\mathbb{E}\ Y_j = \sum_{i=1}^m p_i a_i = 0$. Thus, [13, Prop. 9.11] yields (for each n)

$$\begin{aligned}\left(\mathbb{E}\left\|\sum_{i=1}^{n} Y_i\right\|^p\right)^{1/p} &\le 2T_p(X)\left(\sum_{i=1}^{n} \mathbb{E}\|Y_i\|^p\right)^{1/p} \\ &\le 2T_p(X) n^{1/p} D.\end{aligned}$$

So there exist $b_i^n \in A$ $(1 \le i \le n)$ with

$$\left\|\frac{1}{n}\sum_{i=1}^{n} b_i^n\right\| \le 2T_p(X) n^{(1-p)/p} D \tag{28}$$

Since A is approximately convex,

$$d\left(\frac{1}{n}\sum_{i=1}^{n} b_i^n, A\right) \le \kappa(n-1) \le \log_2 n + 1.$$

So

$$\begin{aligned} d(0,A) &\leq \left\| \frac{1}{n} \sum_{i=1}^{n} b_i^n \right\| + d\left(\frac{1}{n} \sum_{i=1}^{n} b_i^n, A \right) \\ &\leq 2T_p(X) n^{(1-p)/p} D + \log_2 n + 1 \end{aligned}$$

Put $n = 2^{[d]-2}$ (noting that $n \geq 1$ since $d \geq 2$ by assumption) so that $\log_2 n + 1 \leq d - 1$. Then

$$d = d(0,A) \leq 2T_p(X) D (2^{d-3})^{(1-p)/p} + d - 1,$$

which yields (27). □

Remark 6.3. (28) and its probabilistic proof are from [2]. It is proved in [2] that X has the *convex approximation property* if and only if X has type p for some $p > 1$. When X is a Hilbert space, Theorem 5.2 above gave a deterministic proof of (28) with the sharp constants.

Corollary 6.4. *Let X be a Banach space. The following are equivalent:*

(a) *X is B-convex;*
(b) *there exists $c > 0$ such that for every approximately convex set $A \subseteq X$, we have*

$$\operatorname{diam}(A) \geq c \exp(c\mathcal{H}(A, \operatorname{Co}(A))). \tag{29}$$

Proof. It is known that X is B-convex if and only if X has type p for some $p > 1$ [17]. Thus, (a)⇒(b) follows from Theorem 6.2. Now suppose that X is not B-convex. By definition (see Section 3), X contains 'almost isometric' copies of ℓ_1^n for all n. So by Remark 4.5 X contains approximately convex sets A_n such that $\mathcal{H}(A_n, \operatorname{Co}(A_n)) \geq \log_2 n - 1$ and $\operatorname{diam}(A_n) \leq C \log_2 n$, where C is an absolute constant. Clearly, (29) cannot hold in X, and so (b)⇒(a). □

Remark 6.5. The above result is essentially equivalent to [3, Thm. 3.7], which was first obtained in [11].

The following corollary is a partial converse to Corollary 4.2. When combined with the latter it shows that the factor $n^{(p-1)/p}$ in Corollary 4.2 and the exponent of $(p-1)/p$ in Theorem 6.2 are both sharp.

Corollary 6.6. *Let $1 < p < \infty$. There exists a constant $c_p > 0$ such that if $A \subseteq L_p(0,1)$ is approximately convex and satisfies $\mathcal{H}(A, \operatorname{Co}(A)) \geq \log_2 n - 1$, then*

$$\operatorname{diam}(A) \geq \begin{cases} c_p n^{(p-1)/p} & (1 < p \leq 2), \\ c_p n^{1/2} & (2 \leq p < \infty). \end{cases}$$

Proof. It is known that $L_p(0,1)$ has type $\min(p,2)$. Setting $d = \log_2 n - 1$ in Theorem 6.2 gives the result. □

7. Sets with $\operatorname{diam}(A) = \mathcal{H}(A, \operatorname{Co}(A))$

In this section we show that there exists an infinite-dimensional Banach space Y such that for every prescribed diameter D there exists an approximately convex set $A \subseteq Y$ such that $\operatorname{diam}(A) = \mathcal{H}(A, \operatorname{Co}(A)) = D$. This is clearly "worst possible". More precisely, we shall prove the following theorem.

Theorem 7.1. *Let $M > 0$. There exist a Banach space $(X, \|\cdot\|)$ that is linearly isomorphic to ℓ_1 and an approximately convex set $A \subseteq B_M(X)$ such that $\mathcal{H}(A, \operatorname{Co}(A)) = \operatorname{diam}(A) = 2M$.*

First observe that Theorem 7.1 admits the following reformulation in terms of ε-convex sets.

Theorem 7.2. *Let $\varepsilon > 0$. There exist a Banach space $(X, \|\cdot\|)$ that is linearly isomorphic to ℓ_1 and an ε-convex set $A' \subseteq B(X)$ such that $\mathcal{H}(A', \operatorname{Co}(A')) = \operatorname{diam}(A') = 2$.*

Proof. Let $M = 1/\varepsilon$ and let X and A satisfy the conclusion of Theorem 7.1. Then $A' = \varepsilon A$ has the required properties. □

The following lemma is known [3], but for completeness we outline the proof.

Lemma 7.3. *Suppose that $A \subseteq X$ is approximately Jensen-convex. Then A is 2-convex. In particular, $(1/2)A$ is approximately convex.*

Proof. Let $f(x) = d(x, A)$ $(x \in X)$. Then f is a continuous approximately Jensen-convex function, i.e.

$$f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f(x) + f(y)) + 1.$$

By [16] f is a 2-convex function, which implies that A is a 2-convex set. □

Lemma 7.3 shows that Theorem 7.1 is equivalent to the following result.

Theorem 7.4. *Let $M \in \mathbb{N}$. There exist a Banach space $(X, \|\cdot\|)$ that is linearly isomorphic to ℓ_1 and an approximately Jensen-convex set $A \subseteq B_M(X)$ such that $\mathcal{H}(A, \operatorname{Co}(A)) = \operatorname{diam}(A) = 2M$.*

Remark 7.5. The restriction $M \in \mathbb{N}$ is made only to simplify notation in the proof. Clearly the result will hold for all $M > 0$ by scaling.

The rest of the paper is devoted to the lengthy proof of Theorem 7.4. To construct the space X appearing in the conclusion of the theorem, let us begin with the 'tree-like' combinatorial structure which will form a Schauder basis for X. Fix $M \in \mathbb{N}$. Let $L_1 = \mathbb{N}$, and for $n > 1$ define L_n recursively as follows:

$$L_n = \{(a,b) : a \in L_i, b \in L_j, i + j = n, 1 \leq i, j < n\}.$$

Let $L = \cup_{n=1}^{\infty} L_n$, and, for $a \in L$, let e_a denote the indicator function of $\{a\}$. For $R \subseteq L$, let $c_{00}(R)$ denote the vector subspace of $\ell_\infty(R)$ spanned by the set $\{e_a : a \in R\}$. For $x = \sum_{a \in L} \lambda_a e_a \in c_{00}$, let $\operatorname{supp}(a) = \{a \in L : \lambda_a \neq 0\}$

We introduce two norms, $\|\cdot\|_1$ and $\|\cdot\|_1'$, on $c_{00}(L)$:

$$\left\| \sum_{a \in L} \lambda_a e_a \right\|_1 = \sum_{a \in L} |\lambda_a|$$

and

$$\left\| \sum_{a \in L} \lambda_a e_a \right\|_1' = M \sum_{a \in L_1} |\lambda_a| + \sum_{a \in L \setminus L_1} |\lambda_a|.$$

Note that $\|\cdot\|_1$ is the usual ℓ_1 norm and that $\|\cdot\|_1'$ is a weighted ℓ_1 norm with respect to the basis $\{e_a : a \in L\}$. A linear mapping $T : c_{00}(L) \to c_{00}(L)$ is defined (extending linearly) thus:

$$T(e_a) = \begin{cases} 0 & \text{if } a \in L_1, \\ \dfrac{e_b + e_c}{2} & \text{if } a \in \cup_{n=2}^{\infty} L_n \text{ and } a = (b,c). \end{cases}$$

Note that $T(c_{00}(L_n)) \subseteq c_{00}(\cup_{k=1}^{n-1} L_k)$ and that $T^n(x) = 0$ for all $x \in c_{00}(L_n)$. Hence $S = I - T$ is an invertible operator on $c_{00}(L)$ with inverse $S^{-1} = \sum_{k=0}^{\infty} T^k$. Note also that

$$\sum_{a \in L} T(x)(a) \leq \sum_{a \in L} x(a)$$

if $x(a) \geq 0$ for all $a \in A$, with equality if $\operatorname{supp}(x) \subset \cup_{n=2}^{\infty} L_n$.

Define a norm $\|\cdot\|$ on $c_{00}(L)$ thus:

$$\|x\| = \inf\{M\|y\|_1 + \|S^{-1}(z)\|_1' : x = y + z\} \qquad (x \in c_{00}(L)).$$

Let $(X, \|\cdot\|)$ be the completion of $(c_{00}(L), \|\cdot\|)$ and let $A = \{e_a : a \in L\} \subseteq X$.

The verification that X and A satisfy the conclusion of Theorem 7.4 will be broken down into four lemmas.

Lemma 7.6. *Suppose that $F \in B(X^*)$. Then the mapping $\phi : L \to \mathbb{R}$ defined by $\phi(a) = F(e_a)$ satisfies the following:*

(a) $|\phi(a)| \leq M$ *for all* $a \in L$;
(b)

$$\left|\phi(a) - \frac{\phi(b)+\phi(c)}{2}\right| \leq 1$$

for all $a = (b,c) \in \cup_{n=2}^{\infty} L_n$.

Conversely, every ϕ *which satisfies (a) and (b) corresponds to a unique* $F \in B(X^*)$.

Proof. ¿From the definition of $\|\cdot\|$ we see that $F \in B(X^*)$ if and only if

$$|F(x)| \leq \min(M\|x\|_1, \|S^{-1}(x)\|_1') \qquad (x \in c_{00}(L)). \tag{30}$$

Indeed, if F satisfies (30), then for every $x \in c_{00}(L)$, we have

$$\begin{aligned}\|x\| &= \inf\{M\|y\|_1 + \|S^{-1}(z)\|_1' : x = y + z\}\\ &\geq \inf\{F(y) + F(z) : x = y + z\} = F(x),\end{aligned}$$

and so $\|F\| \leq 1$. Conversely, if $\|F\| \leq 1$, then

$$F(x) \leq \|x\| \leq \min(M\|x\|_1, \|S^{-1}(x)\|_1'),$$

and (30) is satisfied.

The condition $|F(x)| \leq M\|x\|_1$ is clearly equivalent to (a). Since $\|\cdot\|_1$ is a weighted ℓ_1 norm, the condition $|F(x)| \leq \|S^{-1}(x)\|_1'$ is equivalent to the condition

$$|F(S(e_a))| \leq \|e_a\|_1' \qquad (a \in L). \tag{31}$$

Suppose that $a \in L_1$. Then $S(e_a) = e_a$ and $\|e_a\|_1' = M$, and so (31) becomes $|\phi(a)| \leq M$. Now suppose that $a = (b,c) \in \cup_{n=2}^{\infty} L_n$. Then $S(e_a) = e_a - (1/2)(e_b + e_c)$ and $\|e_a\|_1' = 1$, and so (31) becomes

$$\left|\phi(a) - \frac{\phi(b)+\phi(c)}{2}\right| \leq 1,$$

and so (b) is satisfied. Conversely, if ϕ satisfies (a) and (b), then the mapping $F(e_a) = \phi(a)$ will extend linearly to an element of $B(X^*)$. □

Remark 7.7. From the description of X^* it follows that

$$\frac{1}{2}\|x\|_1 \leq \|x\| \leq M\|x\|_1.$$

So $(X, \|\cdot\|)$ is isomorphic to ℓ_1 (and the Banach-Mazur distance from X to ℓ_1 is at most $2M$).

Lemma 7.8. *Suppose that $E \subseteq L$ has the property that whenever $a = (b,c) \in E$, then $b, c \in E$. If $\phi_0 : E \to [-M, M]$ satisfies*

$$\left|\phi_0(a) - \frac{\phi_0(b) + \phi_0(c)}{2}\right| \leq 1 \tag{32}$$

for all $a = (b,c) \in E$, then ϕ_0 admits an extension $\phi : L \to [-M, M]$ satisfying

$$\left|\phi(a) - \frac{\phi(b) + \phi(c)}{2}\right| \leq 1 \tag{33}$$

for all $a = (b,c) \in \cup_{n=2}^{\infty} L_n$.

Proof. We define ϕ recursively. First define ϕ from L_1 into $[-M, M]$ to be an arbitrary extension of the restriction of ϕ_0 to L_1. Suppose that $n > 1$ and that ϕ has been defined on $\cup_{k=1}^{n-1} L_k$ to extend the restriction of ϕ_0 to $\cup_{k=1}^{n-1} L_k$. Let $a = (b,c) \in L_n$. Then $b, c \in \cup_{k=1}^{n-1} L_k$, and so $\phi(b)$ and $\phi(c)$ have already been defined. If $a \in E$, then $b, c \in E$, and so $\phi(b) = \phi_0(b)$ and $\phi(c) = \phi_0(c)$. It follows from (32) that (33) will be satisfied with $\phi(a) = \phi_0(a)$. If $a \notin E$, define $\phi(a) = (1/2)(\phi(b) + \phi(c))$, so that (33) is trivially satisfied. This completes the definition of ϕ on L_n. □

Now fix $a \in L$ and let $E_a = \cup_{n=0}^{\infty} \operatorname{supp}(T^n(e_a))$ $(= \cup_{n=0}^{N-1} \operatorname{supp}(T^n(e_a))$ for $a \in L_N)$. For $d \in E_a$, we define the *a-order* of d, denoted $\mathbf{o}_a(d)$, thus:

$$\mathbf{o}_a(d) = \min\{n \geq 0 : d \in \operatorname{supp}(T^n(e_a))\}.$$

Lemma 7.9. *Given $a \in L$, there exists $\phi : L \to [-M, M]$ satisfying* (33) *such that*

$$\phi(d) = -M \qquad (d \in L_1 \setminus E_a), \tag{34}$$

$$\phi(d) = \max(M - \mathbf{o}_a(d), -M) \qquad (d \in L_1 \cap E_a), \tag{35}$$

and

$$\phi(d) \geq \max(M - \mathbf{o}_a(d), -M) \qquad (d \in E_a). \tag{36}$$

Proof. First we define a mapping $\phi_0 : E_a \cup L_1 \to [-M, M]$. If $d \in L_1 \setminus E_a$, let $\phi_0(d) = -M$, and if $d \in E_a \cap L_1$, let

$$\phi_0(d) = \max(M - \mathbf{o}_a(d), -M),$$

so that (34) and (35) are satisfied. Now extend to the rest of E_a recursively as follows. Suppose that $n > 1$ and that ϕ_0 has been defined on $E_a \cap (\cup_{k=0}^{n-1} L_k)$ to satisfy (33) and (36). Let $d \in E_a \cap L_n$. Then $d = (b, c)$ for some $b, c \in E_a \cap (\cup_{k=0}^{n-1} L_k)$. Define

$$\phi_0(d) = \min\left(M, \frac{\phi_0(b) + \phi_0(c)}{2} + 1\right).$$

If $\phi_0(d) = M$, then, as $\phi_0(b) \leq M$ and $\phi_0(c) \leq M$, we have

$$\phi_0(d) = M \leq \frac{\phi_0(b) + \phi_0(c)}{2} + 1 \leq \frac{M + M}{2} + 1 = M + 1,$$

so that

$$\left|\phi_0(d) - \frac{\phi_0(b) + \phi_0(c)}{2}\right| \leq 1,$$

i.e., (33) is satisfied by $d = (b, c)$. Also, if $\phi_0(d) = M$, then (36) is trivially satisfied.

On the other hand, if $\phi_0(d) = (\phi_0(b) + \phi_0(c))/2 + 1$, then (33) is trivially satisfied by $d = (b, c)$. In order to verify (36), suppose that $\mathbf{o}_a(d) = k$. Then both $\mathbf{o}_a(b) \leq k+1$ and $\mathbf{o}_a(c) \leq k+1$. Moreover, both b and c satisfy (36) by the recursive hypothesis. Thus,

$$\begin{aligned}
\phi_0(d) &= \frac{\phi_0(b) + \phi_0(c)}{2} + 1 \\
&\geq \frac{\max(M - \mathbf{o}_a(b), -M) + \max(M - \mathbf{o}_a(c), -M)}{2} + 1 \\
&\geq \frac{\max(M - (k+1), -M) + \max(M - (k+1), -M)}{2} + 1 \\
&\geq \max(M - k, -M) \\
&= \max(M - \mathbf{o}_a(d), -M).
\end{aligned}$$

Thus, (36) is satisfied by d, which completes the recursive definition of ϕ_0. Now ϕ_0 and $E_a \cup L_1$ (replacing E) satisfy the hypotheses of Lemma 7.8. Let ϕ be the extension of ϕ_0 given by Lemma 7.8. □

The following lemma completes the proof of Theorem 7.4.

Lemma 7.10. *Let $A = \{e_a : a \in L\}$. Then A satisfies the following:*

(i) *$A \subseteq B_M(X)$;*
(ii) *A is approximately Jensen-convex;*
(iii) *$\mathcal{H}(A, \mathrm{Co}(A)) = 2M$.*

Proof. Suppose that $F \in B(X^*)$. By Lemma 7.6, $|F(e_a)| \leq M$ for all $a \in A$, and so (i) follows from the Hahn-Banach Theorem. Suppose that $b, c \in A$. Then $a = (b, c) \in A$, and by Lemma 7.6

$$|F(e_a) - F((1/2)(e_b + e_c))| \leq 1,$$

which gives (ii). To prove (iii), note that (i) implies that $\mathrm{Co}(A) \subseteq B_M(X)$ (since $B_M(X)$ is convex), and hence

$$\mathcal{H}(A, \mathrm{Co}(A)) \leq \mathrm{diam}(B_M(X)) = 2M.$$

So it suffices to prove that $\mathcal{H}(A, \mathrm{Co}(A)) \geq 2M$. Fix $N \geq 1$ and choose distinct elements $a_1, \ldots, a_N \in L_1$. We shall prove that

$$d\left(\frac{1}{N}\sum_{k=1}^{N} e_{a_k}, A\right) \geq 2M - \varepsilon_N,$$

where $\varepsilon_N \to 0$ as $N \to \infty$. Let $a \in L$. If $d \in E_a$ and $\mathbf{o}_a(d) = k$, then $T^k(e_a)(d) \geq 2^{-k}$. Since $\sum_{b \in L} T^k(e_a)(b) \leq 1$, it follows that $E_k = \{d \in L : \mathbf{o}_a(d) = k\}$ has cardinality at most 2^k. Thus

$$\left|\bigcup_{k=0}^{2M-1} E_k\right| \leq \sum_{k=0}^{2M-1} 2^k = 2^{2M} - 1.$$

Let $\phi : L \to [-M, M]$ be the function associated to a defined in Lemma 7.9, and let $F \in B(X^*)$ be the linear functional corresponding to ϕ. If $a_i \in L_1 \setminus E_a$, then $\phi(a_i) = -M$ by (34). If $a_i \in E_a$ and $\mathbf{o}_a(a_i) \geq 2M$, then $\phi(a_i) = -M$ by (35). Hence if $a_i \notin G = \cup_{k=0}^{2M-1} E_k$, then $\phi(a_i) = -M$. Moreover, $\phi(a) = M$ by (36), since $\mathbf{o}_a(a) = 0$. So

$$\begin{aligned} F\left(e_a - \frac{1}{N}\left(\sum_{k=1}^{N} e_{a_k}\right)\right) &= \phi(a) - \frac{1}{N}\sum_{k=1}^{N}\phi(a_i) \\ &\geq M - \frac{1}{N}((N - |G|)(-M) + |G|M) \\ &= 2M - \frac{2}{N}|G|M \\ &\geq 2M - \frac{2^{2M+1}M}{N}, \end{aligned}$$

and so

$$\left\|\frac{1}{N}\sum_{k=1}^{N} e_{a_k} - e_a\right\| \geq 2M - \varepsilon_N,$$

where $\varepsilon_N = 2^{2M+1}M/N \to 0$ as $N \to \infty$ as desired. □

Theorem 7.11. *There exists a Banach space Y such that for every $\varepsilon > 0$ there exists an ε-convex set $A \subseteq B(Y)$ with $\mathcal{H}(A, \mathrm{Co}(A)) = 2$.*

Proof. Let X_n denote the space constructed above for $M = n$. Then the ℓ_2-sum $Y = (\sum_{n=1}^{\infty} \oplus X_n)_2$ has the required property. □

Remark 7.12. Since X_n is isomorphic to ℓ_1 (Remark 7.7), it has both the *Radon-Nikodým* property (see e.g. [7]) and the *approximation property* (see e.g. [14, p. 29]). Hence $Y = (\sum_{n=1}^{\infty} \oplus X_n)_2$ has the Radon-Nikodým property [7, p. 219] and (as is easily verified) the approximation property.

Since $C(0,1)$ is a universal space for separable Banach spaces (Mazur's theorem), it satisfies the conclusion of Theorem 7.11. So, finally, let us reformulate Theorem 7.2 to make good the claim made in Remark 3.7.

Corollary 7.13. *Let $\varepsilon > 0$. There exists a (non-negative) ε-convex 1-Lipschitz function on $B(C(0,1))$ such that*

$$\sup\{|f(x) - g(x)| : x \in B(C(0,1))\} \geq 1$$

for every convex function g.

Proof. By Theorem 7.2 there exists $A \subseteq B(C(0,1))$ such that A is ε-convex and $\mathcal{H}(A, \mathrm{Co}(A)) = 2$. Then $f(x) = d(x, A)$ has the required properties. □

References

[1] J. Bourgain and S. J. Szarek, The Banach-Mazur distance to the cube and the Dvoretzky-Rogers Factorization, *Israel J. Math.* **62** (1988), 169–180.

[2] Ronald E. Bruck, On the convex approximation property and the asymptotic behavior of nonlinear contractions in Banach spaces, *Israel J. Math.* **34** (1981), 304–314.

[3] Emanuele Casini and Pier Luigi Papini, Almost convex sets and best approximation, *Ricerche Mat.* **40** (1991), 299–310.

[4] Emanuele Casini and Pier Luigi Papini, A counterexample to the infinity version of the Hyers and Ulam stability theorem, *Proc. Amer. Math. Soc.* **118** (1993), 885–890.

[5] Piotr W. Cholewa, Remarks on the stability of functional equations, *Aequationes Math.* **27** (1984), 76–86.

[6] S. J. Dilworth, Ralph Howard and James W. Roberts, Extremal approximately convex functions and estimating the size of convex hulls, *Adv. in Appl. Math.* (to appear).

[7] J. Diestel and J. J. Uhl, Jr., Vector Measures, Amer. Math. Soc. Providence, R.I., 1977.

[8] John W. Green, Approximately subharmonic functions, *Duke Math. J.* **19** (1952), 499-504.

[9] Donald H. Hyers, George Isac and Themistocles M. Rassias, Stability of Functional Equations in Several Variables, Birkhauser, Boston, 1998.

[10] D. H. Hyers and S. M. Ulam, Approximately convex functions, *Proc. Amer. Math. Soc.* **3** (1952), 821–828.

[11] J.-O Larsson, Studies in the geometrical theory of Banach spaces; part 5: Almost convex sets in Banach spaces of type p, $p > 1$, Ph.D. Thesis, Uppsala Univ., 1987.

[12] Miklós Laczkovich, The local stability of convexity, affinity and of the Jensen equation, *Aequationes Math.* (to appear).

[13] Michel Ledoux and Michel Talagrand, Probability in Banach Spaces, Springer-Verlag, Berlin-Heidelberg, 1991.

[14] J. Lindenstrauss and L. Tzafriri, Classical Banach Spaces I Sequence Spaces, Springer-Verlag, Berlin-Heidelberg-New York, 1977.

[15] I. P. Natanson, Theory of functions of a real variable, Vol. 1, Ungar, New York, 1961.

[16] C. T. Ng and K. Nikodem, On approximately convex functions, *Proc. Amer. Math. Soc.* **118** (1993), 103–108.

[17] G. Pisier, Sur les espaces qui ne contiennent pas de ℓ_n^1 uniformément, Séminaire Maurey-Schwartz 1973-74, Ecole Polytechnique, Paris 1974.

[18] R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.

[19] S. J. Szarek, On the geometry of the Banach-Mazur compactum, *Functional Analysis (Austin, TX, 1987/1989), 48-59, Lecture Notes in Math., 1470*, Springer-Verlag, Berlin, 1991.

Department of Mathematics, University of South Carolina, Columbia, SC 29208, U.S.A.

E-mail address: `dilworth@math.sc.edu, howard@math.sc.edu, roberts@math.sc.edu`