ON THE DUALITY PROBLEM FOR ENTROPY NUMBERS OF OPERATORS

J. Bourgain — IHES, Bures s/Yvette

A. Pajor — Université Paris VII, Paris

S.J. Szarek* — Case Western, Cleveland

N. Tomczak-Jaegermann** — University of Alberta, Edmonton

It is well known that an operator u acting between two Banach spaces X, Y is compact if and only if dual operator u^* is compact. For any such $u : X \to Y$ and for every $\varepsilon > 0$, denote by $N(u, \varepsilon)$ the minimal cardinality of an ε-net, in the metric of Y, of the image $u(B_X)$ of the unit ball B_X of X. Since now the compactness of an operator may be quantified via its *metric entropy* $\log N(u, \varepsilon)$, one may ask for a quantitative version of the result recalled above, i.e., for a comparison of the metric entropies of u and its dual u^*.

It is a conjecture, promoted by B. Carl and A. Pietsch, that the two metric entropies are equivalent in the sense that there exist universal constants $a, b > 0$ so that

$$a^{-1} \log N(u^*, b^{-1}\varepsilon) \le \log N(u, \varepsilon) \le a \log N(u^*, b\varepsilon)$$

holds for any compact operator u and for any $\varepsilon > 0$. We will refer to it as "the duality conjecture" or the "the duality problem".

Let us observe that for operators acting between Hilbert spaces the metric entropies of u and u^* are *exactly* the same; this can be seen by considering polar decompositions. Other special cases are settled in [Car], [GKS], [KMT] and [P-T]. Also, a form of the duality problem – for operators with fixed rank – was considered in [K-M] (see also [Pi4], Chap. 7). However, in the general setting, the problem of equivalence of the metric entropies is still wide open; even in the form requiring one of the constants a or b (but not both) to be equal to 1.

Let us rephrase the problem in terms of the so-called *entropy numbers*, defined for an operator u by

$$e_k(u) \equiv \inf \left\{ \varepsilon > 0 : N(u, \varepsilon) \le 2^k \right\} .$$

* Research done while this author was visiting IHES. Supported in part by NSF Grants DMS-8401906 and 8702058 and the Sloan Research Fellowship.

** Research partially done while this author was visiting University Paris VI.

The result recalled at the beginning just asserts that the two sequences $\big(e_k(u)\big)$ and $\big(e_k(u^*)\big)$ tend to 0 simultaneously. The duality problem asks (roughly) whether

$$b e_{ak}(u^*) \le e_k(u) \le b^{-1} C e_{k/a}(u^*) \ .$$

It seems reasonable to conjecture that $\big(e_k(u)\big)$ and $\big(e_k(u^*)\big)$ should have at least similar asymptotic behavior, for example in the sense that, for any symmetric (i.e., invariant under permutations) norm $\|\cdot\|$ on the space of sequences,

$$C^{-1}\big\|\big(e_k(u^*)\big)\big\| \le \big\|\big(e_k(u)\big)\big\| \le C\big\|\big(e_k(u^*)\big)\big\|$$

holds for any compact operator u with C independent of u. This is the question we study in this paper. Our main results (Theorems 1 & 3) show that the answer is affirmative if one of the spaces is *uniformly convex* or even *B-convex* (see below for definitions); a variant of the statement is also valid for some quasinorms such as the Lorentz $\ell_{p,r}$-"norms" for all $p, r > 0$. Moreover, if one additionally assumes that the sequences $\big(e_k(u)\big)$ and $\big(e_k(u^*)\big)$ decrease in a "regular" manner (see Theorem 1 for a precise formulation), the stronger form of equivalence from the duality conjecture follows (with $a = 1$). If one of the spaces was a Hilbert space, all this was done in [T-J].

We would like to point out that the problem stated above is essentially equivalent to the following one.

Do there exist constants $\alpha, \beta > 0$ such that whenever $x_1, \ldots, x_N$ are points in the unit ball of ℓ_∞^n which are, say, 1/4-separated (i.e., $\|x_i - x_j\| \ge 1/4$ if $i \ne j$), then there exist $f_1, \ldots, f_M$, $M \ge N^\alpha$, in the unit ball of ℓ_1^n satisfying, for $i \ne j$, $\max_{k \le N} \big|\langle f_i - f_j, x_k\rangle\big| \ge \beta$?

Recall that, for $p \in [1,2]$, a Banach space X is said to be of *type p* iff there is a constant C such that

$$\text{Average}_{\varepsilon_i = \pm 1}\Big\|\sum_{1\le i\le m} \varepsilon_i x_i\Big\| \le C\Big(\sum_{1\le i \le m}\|x_i\|^p\Big)^{1/p}$$

holds for any finite sequence $(x_i) \subset X$. The best constant C that works above – the *type p constant of X* – is denoted $T_p(X)$. If X is determined by its unit ball $K \subset I\!R^n$, we will write $T_p(K) = T_p(X)$. X is said to be *B-convex* if it is of type p for some $p > 1$ (this is equivalent to the so-called *K-convexity*, see [Pi3]).

The duality problem (in any of the versions above) may be reduced to a finite dimensional one, with u a one-to-one operator. In this setting we may consider u to be the formal identity and the problem can be restated in the following geometric language. For K – a compact symmetric convex body in $\mathbb{R}^n$, denote by $\|\cdot\|_K$ its *Minkowski functional* (i.e., the norm on $\mathbb{R}^n$ for which K is the unit ball). We will identify K with the normed space $(\mathbb{R}^n, \|\cdot\|_K)$. Let K° be the polar body of K with respect to the canonical Euclidean structure so that we may identify $\|\cdot\|_{K^\circ}$ with the dual norm of $\|\cdot\|_K$.

For two symmetric convex bodies $U, V \subset \mathbb{R}^n$, denote by $N(U,V)$ the maximal cardinality of a set $\{x_1, \dots, x_N\} \subset U$, which is 1-separated in $\|\cdot\|_V$ (we may say: V-separated), i.e., $x_j - x_k \notin V$ for $j \neq k$. This *packing number* $N(U,V)$ turns out to be more convenient for our purposes than the more commonly used in this context *covering number*

$$N'(U,V) \equiv \min\left\{N : \exists \{x_1, \dots, x_N\} \subset U \ , \ U \subset \bigcup_{1 \leq j \leq N} (x_j + V)\right\} ,$$

used implicitly above to define $N(u, \varepsilon) = N'\big(u(B_X), B_Y\big)$. It should be noted here that $N(\cdot,\cdot)$ and $N'(\cdot,\cdot)$ are very closely related; one always has

$$N'(U,V) \leq N(U,V) \leq N'(U, V/2)$$

and so, for our purposes, one does not need to distinguish between them. In particular, all facts about packing numbers can be restated for covering numbers after inserting proper numerical factors (and vice versa). We will frequently use the relations

$$N(U,W) \leq N(U,V)N(V,W) \ , \qquad N(U,V) = N(U, 2U \cap V) \ ,$$

$$N(U, W + 2V) \leq N(U, V + K)N(K, W) \ ,$$

which can be easily checked.

We begin by studying the duality of entropies in the more restrictive context of uniformly convex spaces. In this case we are able to obtain nicer "pointwise" relations between the entropy numbers $e_k(u)$ and $e_k(u^*)$, even though the results on comparison of their asymptotic behavior are identical in the uniformly convex and B-convex setting.

Recall that the *modulus of (uniform) convexity* of a Banach space X is defined by

$$\delta_X(\varepsilon) = \inf\left\{1 - \|x+y\|/2 : \|x\| = \|y\| = 1 \ , \ \|x - y\| = \varepsilon\right\}$$

for $\varepsilon \in [0,2]$. When the space X is defined by its unit ball K in $\mathbb{R}^n$, we will write $\delta_X(\cdot) = \delta_K(\cdot)$. The space X is said to be uniformly convex iff $\delta_X(\varepsilon) > 0$ for all $\varepsilon > 0$ (X is then necessarily B-convex; see [L-T] for more information on this concept).

We are now ready to state our first two results.

Theorem 1. *Let X be a uniformly convex (or just superreflexive) Banach space. Let $u : X \to Y$ (resp. $u : Y \to X$) be a compact operator. Then*

(i) For every $m \in \mathbb{N}$ and $p \in [1,\infty)$

$$C_0^{-1}\Big(\sum_{0\le k\le m} e_k(u^*)^p\Big)^{1/p} \le \Big(\sum_{0\le k\le m} e_k(u)^p\Big)^{1/p} \le C_0\Big(\sum_{0\le k\le m} e_k(u^*)^p\Big)^{1/p} ,$$

where C_0 depends only on X. Moreover, the same holds with the ℓ_p-norm replaced by any symmetric norm.

(ii) If, for some k,

$$e_k(u) \le Ae_{2k}(u) \quad , \quad e_k(u^*) \le Ae_{2k}(u^*) ,$$

then

$$C_1^{-1}e_k(u^*) \le e_k(u) \le C_1 e_k(u^*) ,$$

where C_1 depends only on X and A.

Theorem 1 is derived from the following technical fact.

Proposition 2. *Let $B \subset \mathbb{R}^n$, $K \subset \mathbb{R}^n$ be compact, convex centrally symmetric bodies, B°, K°-the polar bodies. Let $\varepsilon > 0$.*

(i) If K is uniformly convex with modulus of convexity δ_K, then, for every $\theta \ge \varepsilon$,

$$N(B,\varepsilon K) \le N(B,\theta K)N\big(K^\circ, 2^{-4}\theta\delta_K(\varepsilon/\theta)B^\circ\big) .$$

(ii) If B is uniformly convex with modulus of convexity δ_B, then, for every $\theta \ge \varepsilon$,

$$N(B,5\varepsilon K) \le N(B,\theta K)N\big(K^\circ, 2^{-3}\varepsilon\delta_B(\varepsilon/\theta)B^\circ\big) .$$

In particular, for the formal identity operator $u : B \to K$ and $k \in \mathbb{N}$,

(i') $e_{2k}(u) \le Ce_k(u)^{(q-1)/q}e_k(u^*)^{1/q}$ if $\delta_K(\tau) \ge c\tau^q$.

(ii') $e_{2k}(u) \le Ce_k(u)^{q/q+1}e_k(u^*)^{1/(q+1)}$ if $\delta_B(\tau) \ge c\tau^q$,

where C depends on $c > 0$ and $q \in [2,\infty)$.

Remarks. (1) if, for some $c > 0$ and $q \in [2, \infty)$, one has $\delta_X(\tau) \geq c\tau^q$ and $\delta_{X^*}(\tau) \geq c\tau^q$ for all $\tau > 0$, then, in the statement of Theorem 1, the phrase "depends only on X" may be read "depends only on c and q". On the other hand, by [Pil], every uniformly convex (or superreflexive) Banach space can be renormed to have this property.

(2) The assertion (i) of Theorem 1 remains true also of the ℓ_p-"norms", $p \in (0,1)$ and for all Lorentz $\ell_{p,r}$-"norms", the constants depending additionally on p, r as $p, r \to 0$. This follows, e.g., from the fact that the assertions (i′),(ii′) of Proposition 2 imply similar statements with $e_k(\cdot)$ replaced by $e_k(\cdot)^p$.

Proof of Theorem 1 (Assuming Proposition 2). By Remark (1) above, we may assume that X is uniformly convex and uniformly smooth with $\delta_X(\tau) \geq c\tau^q$, $\delta_{X^*}(\tau) \geq c\tau^q$ for some $c > 0$ and $q \in [2, \infty)$. Clearly it is enough to settle the case of $u : Y \to X$ with u one-to-one and X, Y-finite dimensional. Then the parts (i′),(ii′) of Proposition 2 apply and so we have, for some $C = C(q, c)$,

$$e_{2k}(u) \leq Ce_k(u)^{(q-1)/q} e_k(u^*)^{1/q} \quad , \quad e_{2k}(u^*) \leq Ce_k(u^*)^{q/q+1} e_k(u)^{1/(q+1)} .$$

Then (ii) follows immediately. To prove, e.g., the second inequality in (i) for $p = 1$ observe that

$$\sum_{0 \leq k \leq m} e_k(u) \leq 2 \sum_{0 \leq k \leq m/2} e_{2k}(u)$$

while, for every k,

$$\begin{aligned} e_{2k}(u) &\leq Ce_k(u)^{(q-1)/q} e_k(u^*)^{1/q} \\ &= \left[e_k(u)/4\right]^{(q-1)/q} \left[4^{q-1} C^q e_k(u^*)\right]^{1/q} \\ &\leq e_k(u)/4 + 4^{q-1} C^q e_k(u^*) . \end{aligned}$$

Hence

$$\sum_{0 \leq k \leq m} e_k(u) \leq 1/2 \sum_{0 \leq k \leq m/2} e_k(u) + 4^{q-1} C^q \sum_{0 \leq k \leq m/2} e_k(u^*)$$

whence the required inequality immediately follows with, e.g., $C_0 = 2^{2q-1} C^q$. The variant involving a symmetric norm now follows formally from the result of Hardy-Littlewood-Polya; see, e.g., [L-T], Prop. 2.a.5. □

For the proof of Proposition 2 we need two simple lemmas.

Lemma A. *Let $U \subset \mathbb{R}^n$ be uniformly convex with modulus of convexity $\delta(\cdot)$. For $x \in \mathbb{R}^n$, let x^* denote the supporting functional to U at x (i.e., $\|x^*\|_{U^\circ} = 1$, $x^*(x) = \|x\|_U$) and*

$\phi(x) = \|x\|_U \cdot x^*$. *Let* $0 < \varepsilon \le 1$. *If* $x, y \in U$ *with* $\|x - y\|_U \ge \varepsilon$, *then*

$$\max\left\{ |\langle \phi(x) - \phi(y), x\rangle|, |\langle \phi(x) - \phi(y), y\rangle| \right\} \ge \delta(\varepsilon)/8 .$$

Proof: First consider the case when $\|y\|_U \le \|x\|_U = 1$. If $z \in U$ with $\|z\|_U = 1$ and $\|y - z\|_U \ge \varepsilon$, we get from the definition of $\delta(\cdot)$ that

$$\langle y^*, (y+z)/2\rangle \le \|(y+z)/2\|_U \le 1 - \delta(\varepsilon)$$

and so $\langle y^*, z\rangle \le 1 - 2\delta(\varepsilon)$. Since, by a simple variational argument, $\max\left\{ \langle y^*, z\rangle : \|z\|_U \le 1 \ , \ \|y - z\|_U \ge \varepsilon \right\}$ must be achieved when $\|z\|_U = 1$, it follows that $\langle y^*, x\rangle \le 1 - 2\delta(\varepsilon)$. Consequently (still under the assumption $\|x\|_U = 1$)

$$\begin{aligned} \langle \phi(x) - \phi(y), x\rangle &= \langle x^* - \|y\|_U y^*, x\rangle = 1 - \|y\|_U \langle y^*, x\rangle \\ &\ge 1 - \|y\|_U \left(1 - 2\delta(\varepsilon)\right) \ge 2\delta(\varepsilon) , \end{aligned}$$

(note that $\varepsilon \le 1$ implies $\delta(\varepsilon) \le 1/2$). Since (by [Fig], Corollary 11) $\delta(\varepsilon)/\varepsilon^2 < 16\delta(\tau)/\tau^2$ whenever $\varepsilon > \tau > 0$, the assertion follows by homogeneity (and symmetry). □

Lemma B. *Let* $B \subset \mathbb{R}^n$, $K \subset \mathbb{R}^n$ *be as in Proposition 2. Then every* $x \in B$ *can be written in the form* $x = x' + x''$ *with* $x' \in K$, $x'' \in B$ *and* $f \in K^\circ$, *where* f *satisfies* $\|f\|_{B^\circ} \le 1$ *and* $f(x'') = \|x''\|_B$, *i.e.,* f *is a supporting functional to* B *at* x'', *whenever* $x'' \ne 0$.

Proof: Fix $x \in B$. If $x \in K$, the assertion is trivial. If $x \notin K$, choose $x' \in K$ such that $\|x - x'\|_B = \min_{y \in K} \|x - y\|_B \equiv \delta$ (< 1). Then the interiors of K and $x + \delta B$ are disjoint and so we may separate them with a functional f, which we choose to normalize so that $\|f\|_{B^\circ} = 1$ and $f(x) > 0$ (then necessarily $f(x) \le 1$). It then follows that $f(\delta B) = [-\delta, \delta]$, hence $f(x + \delta B) = \left[f(x) - \delta, f(x) + \delta\right]$ and so $f(K) = \left[-(f(x) - \delta), f(x) - \delta\right]$, i.e., $\|f\|_{K^\circ} = f(x) - \delta < 1$. On the other hand, since $x' \in K \cap (x + \delta B)$, we have $f(x') = f(x) - \delta$. Consequently, setting $x'' \equiv x - x'$ we get $f(x'') = \delta = \|x''\|_B$ as required. Note that the argument yields in fact $\|f\|_{K^\circ} + \|x''\|_B \le 1$. □

Proof of Proposition 2. (i) We clearly have

$$N(B, \varepsilon K) \le N(B, 2B \cap \theta K) N(2B \cap \theta K, \varepsilon K) = N(B, \theta K) N(2B \cap \theta K, \varepsilon K) .$$

Let $\{x_1, \ldots, x_N\}$ be a maximal set in $2B \cap \theta K$ such that $x_j - x_k \notin \varepsilon K$ for $j \ne k$. We now use Lemma A to define, for $j = 1, \ldots, N$, $f_j = \phi(x_j/\theta)$ ($\in K^\circ$) such that, for $j \ne k$,

$$\delta_K(\varepsilon/\theta)/8 \le \max\left\{ |\langle f_j - f_k, x_k\rangle|, |\langle f_j - f_k, x_j\rangle| \right\} \le 2/\theta \|f_j - f_k\|_{B^\circ} .$$

So

$$N(2B \cap \theta K, \varepsilon K) \le N\big(K^\circ, 2^{-4}\theta\delta_K(\varepsilon/\theta)B^\circ\big) \ ,$$

which proves (i).

(ii) We have, for $\eta > 0$,

$$(*) \qquad N(B, 5\varepsilon K) \le N(B, \eta B + 2\varepsilon K)N(\eta B, \varepsilon K) = N(B, \eta B + 2\varepsilon K)N(B, \varepsilon/\eta K) \ .$$

Let $\{x_1, \ldots, x_N\}$, $N = N(B, \eta B + 2\varepsilon K)$, be a maximal set in B such that $x_j - x_k \notin \eta B + 2\varepsilon K$ for $j \neq k$. Apply Lemma B with K replaced by εK to obtain decompositions $x_j = x'_j + x''_j$ with $x'_j \in \varepsilon K$, $x''_j \in B$ and functionals $f_j \in B^\circ \cap \varepsilon^{-1}K^\circ$, $f_j(x''_j) = \|x''_j\|_B$ for $j = 1, \ldots, N$. Now, for $j \neq k$, $x_j - x_k = (x'_j - x'_k) + (x''_j - x''_k)$ and, since $x'_j - x'_k \in 2\varepsilon K$, we must have $\|x''_j - x''_k\|_B > \eta$. Consequently, we may apply Lemma A with $U = B$ and $\varepsilon = \eta$ to get a map $\phi : B \to B^\circ$ such that $\big\|\phi(x''_j) - \phi(x''_k)\big\|_{B^\circ} \ge \delta_B(\eta)/8$ for $j \neq k$ while $\|\phi(x''_j)\|_{K^\circ} \le \varepsilon^{-1}$ for all j. Finally, considering the set $\big\{\varepsilon\phi(x''_1), \ldots, \varepsilon\phi(x''_N)\big\} \subset K^\circ$ we deduce that

$$N\big(K^\circ, \varepsilon\delta_B(\eta)/8\, B^\circ\big) \ge N = N(B, \eta B + 2\varepsilon K) \ .$$

To conclude the argument, set $\eta = \varepsilon/\theta$ and combine the above inequality with $(*)$.

(i′) We will show how (i) implies (i′). Fix $k \in \mathbb{N}$ and set $\varepsilon = e_{2k}(u)$, $\theta = e_k(u)$. This means that, roughly $N(B, \varepsilon K) = 2^{2k}$ and $N(B, \theta K) = 2^k$ (the exact statements would involve arbitrary $\varepsilon' < \varepsilon$, $\theta' > \theta$, inequalities rather than equalities and the covering numbers rather than packing numbers). Combined with (i), this shows that $N\big(K^\circ, 2^{-4}\theta\delta_K(\varepsilon/\theta)B^\circ\big) \ge 2^k$ or $e_k(u^*) \ge 2^{-4}\theta\delta_K(\varepsilon/\theta) \ge 2^{-4}c\theta(\varepsilon/\theta)^q$. Solving for ε and substituting $\varepsilon = e_{2k}(u)$ and $\theta = e_k(u)$ we get (i′).

The proof of (ii′) follows exactly the same pattern. □

We now state the results we obtain with B-convexity hypothesis.

Theorem 3. *Let X be a B-convex Banach space (i.e., X is of type p for some $p > 1$). Let $u : X \to Y$ (rep. $u : Y \to X$) be a compact operator. Then the assertions (i), (ii) of Theorem 1 hold; one only needs to replace "for some k" in (ii) by "for all k". Similarly, the Remarks following the statement of Theorem 1 carry over, the role of c and q from the first of them played now by p and $T_p(X)$.*

Theorem 3 will follow, very much in the same way as Theorem 1 follows from Proposition 2, from the following.

Proposition 4. *Let $B, K \subset \mathbb{R}^n$ and $\varepsilon > 0$ be as in Proposition 2. Let $p \in (1,2]$ and $q = p/(p-1)$. Then, for every $\theta \geq \varepsilon$,*

$$N(B, \varepsilon K) \leq N(B, \theta K)\left[N(K^\circ, \varepsilon/8B^\circ)\right]^s$$

with $s \leq \left(2^6 T_p(B)\theta/\varepsilon\right)^q$ (or $s \leq \left(2^6 T_p(K)\theta/\varepsilon\right)^q$).

Remark. The special case $\theta = 2\varepsilon$ of Proposition 4 carries the same strength as the full statement; one can recover the latter one by iteration obtaining, in fact, better estimates on s like, e.g., $s \leq \left(2^7 T_p(\cdot)\right)^q \left(1 + \log(\theta/\varepsilon)\right)$. See also the Remark following the proof of Theorem 6.

Proof of Theorem 3. (ii) We argue first as in the proof of part (i′) of Theorem 1. Set $\varepsilon = e_{2sk}(u)$, $\theta = e_{sk}(u)$, where $s = \left[(2^6 A T_p(X))^q\right]$. Proposition 4 them implies that $e_k(u^*) \geq \varepsilon/8$ and so

$$e_k(u) \leq (4s)^{\log_2 A} e_{2sk}(u) \leq 8(4s)^{\log_2 A} e_k(u^*) \ ,$$

as needed; the other inequality is proved in the same way.

From the proof of part (i) we need a simple

Lemma C. *Let $\mathbb{L}, \mathbb{L}^* : (0,\infty) \to [0,\infty)$ be nonincreasing functions with bounded support such that, for some $s \geq 1$ and all $x \geq 0$,*

$$(*) \qquad \mathbb{L}(x) \leq \mathbb{L}(2x) + s\mathbb{L}^*(x) \ .$$

Let e (resp. e^*) be the "inverse" defined by

$$e(y) = \inf\left\{\tau : \mathbb{L}(\tau) \leq y\right\} \ .$$

Then, for any $m \in \mathbb{N}$,

$$\sum_{1 \leq k \leq m} e(k) \leq 2s \sum_{0 \leq k \leq m-1} e^*(k) \ .$$

Proof: We have, by $(*)$,

$$\int_0^\infty \mathbb{L}(x)dx \leq \int_0^\infty \mathbb{L}(2x)dx + s\int_0^\infty \mathbb{L}^*(x)dx = 1/2\int_0^\infty \mathbb{L}(x)dx + s\int_0^\infty \mathbb{L}^*(x)dx$$

and so

$$\int_0^\infty \mathbb{L}(x)dx \leq 2s\int_0^\infty \mathbb{L}^*(x)dx$$

provided the first of the integrals is finite. Now fix $m \in \mathbb{N}$ and observe that

$$\sum_{1 \leq k \leq m} e(k) \leq \int_0^m e(y)dy = \int_0^\infty (\mathbb{L}(x) \wedge m)dx \leq \sum_{0 \leq k \leq m-1} e(k)$$

(resp. $\mathbb{L}^*, e^*$). Since it is clear that the condition $(*)$ still holds if we replace $\mathbb{L}, \mathbb{L}^*$ by $\mathbb{L} \wedge m, \mathbb{L}^* \wedge m$, we may as well assume that $\mathbb{L}, \mathbb{L}^*$ are bounded by m. Combining this observation with the last two estimates we get the assertion. □

We are now ready to prove the part (i) of the Theorem. As in the proof of the corresponding part of Theorem 1, it is enough to show that, for any $m \in \mathbb{N}$,

$$\sum_{0\le k\le m} e_k(u) \le C_0 \sum_{0\le k\le m} e_k(u^*) \ .$$

But this follows immediately from Lemma C applied with $\mathbb{L}(x) \equiv \log N(u,x)$, $\mathbb{L}^*(x) \equiv \log N(8u^*,x)$ and $s = \left(2^7 T_p(X)\right)^q$; the fact that the hypothesis $(*)$ of Lemma C is satisfied is then just a restatement of Proposition 4 applied with $\theta = 2\varepsilon$. □

For the proof of Proposition 4 we need another lemma, which is due to B. Maurey (see [Pi2]). We include the proof for completeness.

Lemma D. *Let Y be a Banach space which is of type p for some $p > 1$, $D \subset Y$ and $x \in \operatorname{conv} D$. Set $q = p/(p-1)$. Then, for any $m \in \mathbb{N}$, there exist $x_1, \dots, x_m \in D$ such that*

$$\Big\| x - 1/m \sum_{1\le j\le m} x_j \Big\| \le T_p(Y) m^{-1/q} \cdot \operatorname{diam} D \ .$$

Consequently, if $U = \operatorname{conv}\{z_1, \dots, z_N\} \subset B_Y$, then

$$N\big(U, 4T_p(Y)m^{-1/q}B_Y\big) \le N'\big(U, 2T_p(Y)m^{-1/q}B_Y\big) \le N^m \ .$$

Proof: Since $x \in \operatorname{conv} D$, there exist $y_1, \dots, y_N \in D$ and positive scalars $t_1, \dots, t_N$, $\sum_{1\le j\le N} t_j = 1$, such that $x = \sum_{1\le j\le N} t_j y_j$. Let Z be a D-valued random variable which takes the value y_j with probability t_j. Then $\mathbb{E}Z = x$, where $\mathbb{E}$ stands for the expected value. Let $Z_1, Z_2, \dots, Z_m, Z_1', Z_2', \dots, Z_m'$ be a sequence of independent copies of Z. Then,

$$\begin{aligned} \mathbb{E}\Big\| \sum_{1\le j\le m} (Z_j - x)\Big\| &\le \mathbb{E}\Big\| \sum_{1\le j\le m} (Z_j - Z_j')\Big\| \\ &= \mathbb{E}\,\mathrm{Average}_{\varepsilon_j=\pm1} \Big\| \sum_{1\le j\le m} \varepsilon_j (Z_j - Z_j'\Big\| \le T_p(Y) m^{1/p} \cdot \operatorname{diam} D \ , \end{aligned}$$

where we used consecutively the facts that, for all j, $\mathbb{E}(Z_j' - x) = 0$; that the $(Z_j - Z_J')$'s are symmetric and independent, and the definition of type p. Comparing the first and the last terms shows that, for $x_1, \dots, x_m$ chosen "at random" (possibly with repetitions) from among

$y_1,\ldots,y_M$, the first assertion of the Lemma holds. The estimate on the covering numbers (hence packing numbers) follows now from the fact that, in the situation as in the second part of the Lemma, there are $< N^m$ such choices as we have then of course $\{y_1,\ldots,y_M\} \subset \{z_1,\ldots,z_N\}$. □

Proof of Proposition 4. The arguments are similar to those used to prove Proposition 2. First consider the case when we control the type p constant of K. We have

$$(*) \qquad N(B,\varepsilon K) \le N(B,2B\cap\theta K)N(2B\cap\theta K,\varepsilon K) = N(B,\theta K)N\big((2\theta^{-1}B)\cap K,\varepsilon/\theta K\big)\ .$$

We now claim that, for $\delta \in (0,1]$ and $s = \big[(2^6T_p(K)/\delta)^q\big]$, we have

$$(**) \qquad N(B\cap K,\delta K) \le \big[N(K^\circ,\delta/4B^\circ)\big]^s\ .$$

From this the assertion readily follows: just combine $(**)$ – applied with $\delta = \varepsilon/\theta$ and B replaced by $2\theta^{-1}B$ – with $(*)$.

To prove $(**)$, let $\{x_1,\ldots,x_N\}$ be a maximal set in $B\cap K$ such that $x_j - x_k \notin \delta K$ for $j \ne k$. Choose m so that $4T_p(Y)m^{-1/q} < \delta/2$. It now follows from Lemma D that if $D \subset \{x_1,\ldots,x_N\}$ satisfies $|D| < N^{1/m}$, then, for some $j \le N$, $(\operatorname{conv} D)\cap(x_j+\delta/2K) = \emptyset$. Thus one can construct, by induction, a subset $\{y_1,\ldots,y_M\}$ of $\{x_1,\ldots,x_N\}$ with $M \ge N^{1/m}$ such that, for all $k \le M$,

$$(***) \qquad \big(\operatorname{conv}\{y_j : j<k\}\big)\cap(y_k+\delta/2K) = \emptyset\ .$$

Now produce, by a separation argument, $f_1,\ldots,f_M \in K^\circ$ such that, for each $j<k\le M$,

$$\langle f_k,y_k\rangle > \langle f_k,y_j\rangle + \delta/2\ .$$

This is possible by $(***)$. Divide the interval $[-1+\delta/2,1]$ into $[8/\delta]$ subintervals of length $< \delta/4$. Then for at least one of them, say I, the set $\{k : \langle f_k,y_k\rangle \in I\} \equiv \sigma$ is of cardinality $\ge \delta/8\cdot N$. For all $j,k\in\sigma$ with $j<k$ we have then

$$\langle f_j - f_k, y_j\rangle > \delta/4\ ,$$

i.e., the set $\{f_j\}_{j\in\sigma} \subset K^\circ$ is $\delta/4B^\circ$-separated. In other words, we did show that

$$N(B\cap K,\delta K) \le \big[8/\delta N(K^\circ,\delta/4B^\circ)\big]^m\ .$$

This is nearly (**); one just needs to get rid of the factor $8/\delta$. But it is evident that $N(K^\circ, \delta/4B^\circ) \geq 4/\delta$ if $K^\circ \not\subset \delta/4B^\circ$ ((**) is trivially satisfied otherwise) and so we can drop that factor if we appropriately increase the exponent.

We now sketch the argument in the case when we control the type p constant of B. As in the proof of part (ii) of Proposition 2, the problem reduces to showing that, for $\varepsilon > 0$ and $\eta \in (0,1]$,

$$N(B, \eta B + \varepsilon K) \leq \left[N(K^\circ, \varepsilon/4B^\circ)\right]^s .$$

Again let $\{x_1, \dots, x_N\}$, $N = N(B, \eta B + \varepsilon K)$, be a maximal set in B such that $x_j - x_k \notin \eta B + \varepsilon K$ for $j \notin K$. Let $m \in \mathbb{N}$ be such that $4T_p(X)m^{-1/q} < \eta/4$. As before, if $D \subset \{x_1, \dots, x_N\}$ satisfies $|D| < N^{1/m}$, then, for some $j \leq N$,

$$(\operatorname{conv} D) \cap (x_k + \varepsilon/2K + \eta/4B) = \emptyset .$$

Indeed, otherwise we would get, for all $k \leq N$, $z_k \in (\operatorname{conv} D) \cap (x_k + \varepsilon/2K + \eta/4B)$ (and so $z_j - z_k \notin \eta/2B$ for $j \notin k$) while, by Lemma D, $N(\operatorname{conv} D, \eta/4B) \leq |D|^m < N$ – a contradiction. This leads to a subset $\{y_1, \dots, y_M\}$ of $\{x_1, \dots, x_N\}$ with $M \geq N^{1/m}$ such that, for all $k \leq M$,

$$\left(\operatorname{conv}\{y_j : j < k\}\right) \cap (y_k + \varepsilon/2K + \eta/4B) = \emptyset ,$$

to functionals $f_1, \dots, f_m \in K^\circ \cap 2\varepsilon/\eta B^\circ$ verifying, for $j < k \leq M$,

$$\langle f_k, y_k \rangle > \langle f_k, y_j \rangle + \varepsilon/2$$

and finally to the estimate

$$N(B, \eta B + \varepsilon K) \leq \left[8/\eta N(K^\circ, \varepsilon/4B^\circ)\right]^m ,$$

from which we derive the required inequality in the same manner as in the first part of the proof. □

We conclude by presenting two results which can be obtained, by methods similar to the proofs of the Theorems above, without any assumptions on the spaces involved.

Theorem 5. *Let $u : X \to Y$ be a compact operator. Then, for any $\theta \geq \varepsilon > 0$,*

$$\log N(u, \varepsilon) \leq \log N(u, \theta) + \left[\psi(\theta/\varepsilon) \log N(u^*, \psi(\theta/\varepsilon)^{-1}\varepsilon)\right]^2 ,$$

where $\psi(\cdot)$ is a polynomial function. Consequently, for any $p, r \geq 1$ and $m \in \mathbb{N}$ we have the inequality between the Lorentz "norms":

$$\left\|(e_k(u))_{k \leq m}\right\|_{2p,r} \leq C\left\|(e_k(u^*))_{k \leq m^{1/2}}\right\|_{p,r} ,$$

where C is a universal constant (resp. depending on p, r if $p, r \in (0,1)$).

Proof: (Sketch). The argument from the proof of Proposition 4 shows that it is enough to prove that if B, K are as in Proposition 2 and $\theta \geq \varepsilon > 0$, then

$$\log N(B \cap K, \varepsilon K) \leq \left\{\psi(1/\varepsilon) N(K^\circ, \Psi(1/\varepsilon)^{-1} B^\circ)\right]^2 .$$

To this end, we need to examine the proofs of Propostion 4 and Lemma D. Indeed, all needed in order to show that, for some $m \in \mathbb{N}$,

$$\log N(K^\circ, \delta/4B^\circ) \geq (3m)^{-1} \log N(B \cap K, \delta K)$$

was that, for any $x_1, \ldots, x_m \in B \cap K$,

$$\text{Average}_{\varepsilon_j = \pm 1} \left\| \sum_{1 \leq j \leq m} \varepsilon_j x_j \right\|_K \leq \delta/2m .$$

On the other hand, if this is not the case, then, by [Elt] (or [Paj] in the complex case), there exists $\sigma \in \{1, \ldots, m\}$, $|\sigma| \geq cm$, such that, for any scalars $(t_j)_{j \in \sigma}$,

$$\left\| \sum_{j \in \sigma} t_j x_j \right\|_K \geq \beta \sum_{j \in \sigma} |t_j| ,$$

where c and β depend (polynomially) on δ. This yields $\geq 2^{cm}$ functionals in K°, which are $2\beta B^\circ$-separated, i.e.,

$$\log N(K^\circ, 2\beta B^\circ) \geq cm .$$

Now an analysis of $\max_m \min \left\{(3m)^{-1} \log N(B \cap K, \delta K), cm\right\}$ leads to the assertion.

To derive the part involving the Lorentz norms from the first part, we argue as in the proof of Theorem 3 to obtain, for $m \in \mathbb{N}$,

$$\int_0^\infty \left(\log N(u, \varepsilon) \wedge m\right) d\varepsilon \leq C_0 \int_0^\infty \left(\log N(Cu^*, \varepsilon) \wedge m^{1/2}\right)^2 d\varepsilon ,$$

which is equivalent to

$$\sum_{0 \leq k \leq m} e_k(u) \leq C_1 \sum_{0 \leq k \leq m} e_{[k^{1/2}]}(u^*)$$

(we note here that the latter quantity is of the same order as $\sum_{0 \leq k \leq m^{1/2}} k e_k(u^*)$). Consequently, for any symmetric norm $\|\cdot\|$ on sequences,

$$\left\| \left(e_k(u)\right)_{k \leq m} \right\| \leq C_1 \left\| \left(e_{[k^{1/2}]}(u^*)\right)_{k \leq m} \right\| ,$$

which in the context of the Lorentz norms reduces to the required inequality. □

Theorem 6. *Let $u : X \to Y$ be a norm one operator with* $\operatorname{rank} u \le n$. *Then, for any* $\varepsilon \in (0,1)$,

$$\log N(u,\varepsilon) \le C \log(n/\varepsilon) \log N\big(u^*, C^{-1}\varepsilon/\log(n/\varepsilon)\big) \ .$$

where C is a universal constant.

Proof: (Sketch). Let $A = C_1 \log(n/\varepsilon)$, The choice of C_1 will be indicated by what follows. Our objective is to exhibit, for sufficiently large M, a set $\{x_1, \ldots, x_M\} \subset u(B_X)$ such that, for $k \le M$,

$$(*) \qquad \big(\operatorname{conv}\{x_j : j < k\}\big) \cap (x_k + \varepsilon/A B_Y) = \emptyset \ ;$$

then the argument from the proof of Theorem 3 will show that

$$\log N(u^*, \varepsilon/2A) \ge M/3 \ .$$

Suppose that, for some M, this is not possible. Introduce an auxiliary Euclidean norm $|\cdot|_E$ on the range of u such that $\|\cdot\|_Y \le |\cdot|_E \le n^{1/2}\|\cdot\|_Y$. Let $D, |D| = N(u,\varepsilon)$ be an εB_Y-separated subset of $u(B_X)$. Then $\operatorname{diam}_E D \le n^{1/2}$. By our assumption, there exists $D' \subset D$, $|D'| < M$ such that, for any $x \in D$, $\operatorname{dist}_Y(x, \operatorname{conv} D') < \varepsilon/A$. Applying Lemma D with, say $m = 100$, we see that $\operatorname{conv} D'$ can be covered with M^{100} Euclidean balls of radius $\operatorname{diam}_E D/5 \le n^{1/2}/5$. Consequently, there is a set $D_1 \subset \operatorname{conv} D'$, $|D_1| \ge N(u,\varepsilon)/M^{100}$, $\operatorname{diam}_E D_1 \le n^{1/2}/5$ such that, for any $x \in D_1$, there exists $x' \in D$ with $\|x - x'\|_Y \le C^{-1}\varepsilon/\log n$. Since, *a fortiori*, $\{x_1, \ldots, x_M\}$ verifying $(*)$ cannot be found inside D_1, we can repeat this procedure with D_1 in place of D and so on. If it was possible to perform, say, $s > A/100$ iterations, we would be left with a set D_s, $|D_s| \ge N(u,\varepsilon)/M^{100s} \ge N(u,\varepsilon)/M^A$, still $\varepsilon/2B_Y$-separated, such that $\operatorname{diam}_Y D_s \le \operatorname{diam}_E D_s \le \operatorname{diam}_E D_1/5^s \le \varepsilon/4$. This is only possible if $|D_s| = 1$, i.e., $M^A \ge N(u,\varepsilon)$, which gives the needed lower estimate on M. □

Remark. In the statement of Theorem 6 the "norm one + rank n" hypothesis can be dropped if we replace $\log(n/\varepsilon)$ by $\log\big(\gamma_2(u)/\varepsilon\big)$, where $\gamma_2(\cdot)$ denotes the norm of factorization through the Hilbert space. In fact, it is enough to consider just the "norm of factorization through a B-convex space". This can be used to unify somewhat the proofs of the two parts of Proposition 4; notice that for u – the formal identity $2B \cap \theta K \longrightarrow \varepsilon K$ – such "norm" is θ/ε (resp. η^{-1} for $B \longrightarrow \eta B + \varepsilon K$).

References

[Car] B. Carl, Inequalities of Bernstein-Jackson type and the degree of compactness of operators in Banach spaces, Ann. Inst. Fourier 35 (1985), 79-118.

[Elt] J. Elton, Sign-embeddings of ℓ_1^n, Trans. A.M.S. 279 (1983), 113-124.

[Fig] T. Figiel, On the moduli of convexity and smoothness, Studia Math. 56 (1976), 121-155.

[GKS] Y. Gordon, H. König, C. Schütt, Geometric and probabilistic estimates for entropy and approximation numbers of operators, J. Approx. Theory 49 (1987), 219-239.

[K-M] H. König, V.D. Milman, On the covering numbers of convex bodies, Israel Functional Analysis Seminar GAFA, Lecture Notes in Math. 1267, Springer Verlag 1987, 82-95.

[KMT] H. König, V.D. Milman, N. Tomczak-Jaegermann, Entropy numbers and duality for operators with values in Hilbert spaces, to appear in Proc. Internat. Conf. on Probab. in Banach spaces, June 1986, Lecture Notes in Math., Springer Verlag.

[L-T] J. Lindenstrauss, L. Tzafriri, Classical Banach Spaces II, Ergebnisse der Mathematik and ihrer Grenzgebiete 97, Springer Verlag 1979.

[Paj] A. Pajor, Sous-espaces ℓ_1^n des espaces de Banach, Travaux en Cours, Hermann, Paris 1986.

[P-T] A. Pajor, N. Tomczak-Jaegermann, Volume ratio and other s-numbers related to local properties of Banach spaces, J. Funct. Anal. to appear.

[Pi1] G. Pisier, Martingales with values in uniformly convex spaces, Israel J. Math. 20 (1975), 326-350.

[Pi2] G. Pisier, Remarques sur un résultat non publié de B. Maurey, Séminaire d'Analyse Fonctionnelle 1980/81, Exposé No. V, Ecole Polytechnique, Palaiseau.

[Pi3] G. Pisier, Holomorphic semi-groups and the geometry of Banach spaces, Annals of Math. 115 (1982), 375-392.

[Pi4] G. Pisier, Volume Inequalities in the Geometry of Banach Spaces, Cambridge University Press, to appear.

[T-J] N. Tomczak-Jaegermann, Dualité des nombres d'entropie pour des opérateurs à valeurs dans un espace de Hilbert, C.R. Acad. Sci. Paris 305 (1987), 299-301.