

Estimates of Covering Numbers

B. CARL

*Fachbereich Mathematik, Universität Oldenburg,
D-2900 Oldenburg, Germany*

AND

A. HESS

*Sektion Mathematik, Friedrich Schiller Universität Jena,
Universitätshochhaus 17. OG, D-6900 Jena, Germany*

Communicated by Oved Shisha

Received May 8, 1989; revised July 20, 1990

For operators between Banach spaces, we study certain s -covering numbers, which are a kind of combination between s -numbers and entropy numbers. We prove inequalities between s -covering numbers and various s -numbers. As an application, minoration theorems involving the ℓ -norm and γ -norm are given.

© 1991 Academic Press, Inc.

1. INTRODUCTION

In this paper we establish inequalities involving s -numbers and so-called s -covering numbers of operators between Banach spaces. One of the main results, which is proved in Section 3 (Theorem 3.4), is the following weak type inequality, in which the a -covering numbers of operators are dominated by the approximation numbers

$$\sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{n,l}(T) \leq C(p) \sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{l+n-1}(T), \quad (1.1)$$

for $0 < p < \infty$, $1 \leq l, m < \infty$. Various inequalities of weak type, used to majorize other s -covering numbers by s -numbers, can be verified with the help of this inequality. In order to prove (1.1) we refer to a striking result of Pisier [P1, P2] concerning the existence of isomorphisms from arbitrary n -dimensional Banach spaces into the n -dimensional Hilbert space such

that certain s -numbers of these isomorphisms can be appropriately majorized.

A number of useful applications of (1.1) are given in Section 4. We derive so-called minoration theorems for Gaussian averages (l -norms) and Rademacher averages (r -norms) of operators from ℓ_2 into Banach spaces. The results generalize the classical minoration theorem of Sudakov for the ℓ -norm, which plays an important role in the theory of stochastic processes. The estimates can be considered as a kind of interpolation or combination between Sudakov's minoration theorem and a minoration theorem of the ℓ -norm, which originates in the paper of V. Milman [M1] and which was improved by A. Pajor and N. Tomczak-Jaegermann [PaT].

Throughout this paper we use standard definitions and notations of Banach space theory. For the sake of convenience we recall some of them. Throughout this paper E , F , and G denote (real or complex) Banach spaces. Given a Banach space E we denote its closed unit ball by U_E and its dual space by E' . Moreover $L(E, F)$ denotes the Banach space of all (bounded linear) operators from E into F equipped with the usual operator norm.

In order to describe the covering concept let us recall the definition of the dyadic entropy numbers $e_n(T)$. The n th dyadic entropy number $e_n(T)$, $n \geq 1$, of an operator $T \in L(E, F)$ is defined to be the infimum of all $\varepsilon > 0$ such that there exist $y_1, \dots, y_{2^{n-1}} \in F$ for which

$$T(U_E) \subseteq \bigcup_{i=1}^{2^{n-1}} \{y_i + \varepsilon U_F\}$$

holds. For the definition of s -numbers we refer to [Pi] or [CS].

2. s -COVERING NUMBERS

According to the concepts of s -numbers and dyadic entropy numbers we define an s -covering number function. Let L denote the class of all (bounded linear) operators, \mathbb{N} the set of natural numbers, and $s: L \rightarrow \ell_\infty(\mathbb{N})$ any s -number function. An s -covering number function is a map

$$s = (s_{n,k}): L \rightarrow \ell_\infty(\mathbb{N} \times \mathbb{N})$$

which associates with every operator T its s -covering numbers $s_{n,k}(T)$ and satisfies the following properties:

(SC i) *Combination Property.*

$$s_{n,1}(T) = e_n(T), \quad s_{1,n}(T) = s_n(T) \quad \text{for } T \in L(E, F) \text{ and } 1 \leq n < \infty.$$

(SC ii) *Monotonicity.*

$$\|T\| = s_{1,1}(T) \geq s_{m,j}(T) \geq s_{n,k}(T) \geq 0$$

for $T \in L(E, F)$ and $n \geq m \geq 1$ and $k \geq j \geq 1$.

(SC iii) *Additivity.*

$$s_{m+n-1, j+k-1}(S+T) \leq s_{m,j}(S) + s_{n,k}(T)$$

for $S, T \in L(E, F)$ and $1 \leq j, k, m, n < \infty$.

(SC iv) *Ideal Property.*

$$s_{n,k}(RST) \leq \|R\| s_{n,k}(S) \|T\|$$

for $R \in L(G, H), S \in L(F, G), T \in L(E, F)$, and $1 \leq k, n < \infty$.

Moreover, some s -covering number functions satisfy an additional property called

(SC v) *C-Multiplicativity.* There exists a constant $C \geq 1$ such that

$$s_{m+n-1, j+k-1}(ST) \leq C \cdot s_{m,j}(S) s_{n,k}(T)$$

for $S \in L(F, G), T \in L(E, F)$, and $1 \leq j, k, m, n < \infty$.

If $C = 1$, then the s -covering number function is said to be multiplicative. This concept was introduced in [CM] in a slightly modified version.

The basic examples are the

a-Covering numbers.

$$a_{n,k}(T) = \inf\{e_n(T-A) : \text{rank}(A) < k\}. \quad (2.1)$$

This means that for $T \in L(E, F)$, $a_{n,k}(T)$ is the infimum of all $\rho > 0$ such that there exists an operator $A \in L(E, F)$ with $\text{rank}(A) < k$ and elements $y_i \in F$, $1 \leq i \leq 2^{n-1}$ for which

$$(T-A)(U_E) \subseteq \bigcup_{i=1}^{2^{n-1}} \{y_i + \rho U_F\}.$$

Analogously:

c-Covering numbers.

$$c_{n,k}(T) = \inf\{e_n(TI_M^E) : M \subseteq E, \text{codim}(M) < k\}. \quad (2.2)$$

d-Covering numbers.

$$d_{n,k}(T) = \inf\{e_n(Q_N^F T) : N \subseteq F, \dim(N) < k\}. \tag{2.3}$$

t-Covering numbers.

$$t_{n,k}(T) = a_{n,k}(I_x TQ_1), \tag{2.4}$$

where $Q_1: \ell_1(U_E) \rightarrow E$ and $I_x: F \rightarrow \ell_x(U_{F'})$ are the canonical surjection and injection, respectively. (For the *t*-covering numbers we have a weaker form of the combination property, namely

$$\frac{1}{2} e_n(T) \leq t_{n,1}(T) \leq e_n(T).$$

These *s*-covering number functions were already considered in [CM] together with the so-called

e-Covering numbers.

$$e_{n,k}(T) = \inf\{\varepsilon > 0: \text{there exist subspaces } M_i \subseteq F, \dim(M_i) < k, \text{ and } y_i \in F, 1 \leq i \leq 2^{n-1}, \text{ such that } T(U_E) \subseteq \bigcup_{i=1}^{2^{n-1}} \{y_i + M_i + \varepsilon U_F\}\}. \tag{2.5}$$

Note that the combination property (SCi) coincides for the *e*- and *d*-covering numbers since $e_{n,1}(T) = e_n(T)$ and $e_{1,n}(T) = d_n(T)$. The difference between (2.3) and (2.5) becomes clear if we rewrite (2.3) into

$$d_{n,k}(T) = \inf\{\varepsilon > 0: \text{there exists a subspace } M \subseteq F, \dim(M) < k, \text{ and } y_i \in F, 1 \leq i \leq 2^{n-1}, \text{ such that } T(U_E) \subseteq \bigcup_{i=1}^{2^{n-1}} \{y_i + M_i + \varepsilon U_F\}\},$$

which means that in the definition of $d_{n,k}$ the coverings consist of “cylinders” $\{y_i + M_i + \varepsilon U_F\}$ with one common “direction” M while the directions may be different in (2.5). From this we obviously obtain

$$e_{n,k}(T) \leq d_{n,k}(T).$$

3. INEQUALITIES BETWEEN *s*-COVERING NUMBERS AND *s*-NUMBERS

This section deals with basic estimates between *s*-covering numbers and *s*-numbers.

LEMMA 3.1. *Let $A \in L(E, F)$ be an operator of finite rank, $\text{rank}(A) = r$, acting between real Banach spaces E and F . Furthermore, let $N = N(A)$ and*

$R = R(A)$ be the null space and range of A , respectively. For arbitrary isomorphisms

$$X: E/N \rightarrow \ell'_2 \quad \text{and} \quad Y: R \rightarrow \ell'_2$$

we have

$$a_{3n-2,l}(A) \leq 7e_n(X) e_n(Y^{-1}) \sup_{1 \leq k \leq M} \left[2^{-n(3k+3)} \times \left\{ \prod_{i=1}^k c_i(Y) d_i(X^{-1}) \right\}^{1:k} \left\{ \sum_{i=1}^k t_{l+i-1}(A) \right\}^{1:k} \right], \quad (3.1)$$

for $1 \leq l, n < \infty$ with

$$M = \max \left\{ 1, \left[1 + \frac{r-l}{3} \right] \right\}, \quad (3.2)$$

where $[x]$ denotes the integer part of x .

Proof. Without loss of generality we may assume that

$$r = \text{rank}(A) \geq 1, \quad (3.3)$$

since $a_{3n-2,l}(A) \leq a_{1,l}(A) = a_1(A) = 0$ in the contrary case. We factorize A canonically through the quotient map $Q: E \rightarrow E/N$ and the imbedding $I_R^F: R \rightarrow F$,

$$A = I_R^F A_0 Q, \quad (3.4)$$

and use the isomorphisms $X: E/N \rightarrow \ell'_2$ to introduce $S: \ell'_2 \rightarrow \ell'_2$ by

$$S = Y A_0 X^{-1} \quad \text{or, equivalently,} \quad A_0 = Y^{-1} S X. \quad (3.5)$$

Next we estimate the a -covering numbers of S . According to the Schmidt representation formula there exist isometries $U, V: \ell'_2 \rightarrow \ell'_2$ and a diagonal operator D with positive entries such that

$$S = U D V^{-1} \quad \text{and} \quad D = U^{-1} S V. \quad (3.6)$$

If the diagonal elements σ_i of D are ordered in a non-increasing sequence we then can express σ_i by

$$\sigma_i = a_i(S) = c_i(S) = d_i(S) = t_i(S), \quad 1 \leq i \leq r. \quad (3.7)$$

The ideal property (SC iv) and (3.6) imply that

$$a_{n,l}(S) = a_{n,l}(D), \quad 1 \leq l, n < \infty. \quad (3.8)$$

An estimate for the $a_{n,l}(D)$ can be obtained as follows. Consider $D_l: \ell'_2 \rightarrow \ell'_2$, $D_l(\xi_1, \xi_2, \dots, \xi_r) = (\sigma_1 \xi_1, \dots, \sigma_{l-1} \xi_{l-1}, 0, \dots, 0)$. According to the definition of $a_{n,l}$ and a result of Gordon, Föning, and Schütt (cf. [CS, GKS]) we obtain

$$a_{n,l}(D) \leq e_n(D - D_l) \leq 6 \sup_{1 \leq k \leq r-l+1} 2^{-(n-1)/k} \left\{ \prod_{i=1}^k \sigma_{l+i-1} \right\}^{1/k}. \quad (3.9)$$

It follows from (3.7) and (3.8) that

$$a_{n,l}(S) \leq 6 \sup_{1 \leq k \leq r-l+1} 2^{-(n-1)/k} \left\{ \prod_{i=1}^k t_{l+i-1}(S) \right\}^{1/k}. \quad (3.10)$$

Note that the range of k is not empty because of (3.3).

In order to replace the operator S by A we use (3.5) and the multiplicativity, injectivity, and surjectivity of the symmetrized approximation numbers t_i . The multiplicativity of the t_i allows us to write

$$t_{i+j+k-2}(RST) \leq c_i(R) t_j(S) d_k(T)$$

(cf. [CS]). Hence we obtain from (3.5)

$$t_{l+3i-3}(S) \leq c_i(Y) t_{l+i-1}(A_0) d_i(X^{-1}) = c_i(Y) t_{l+i-1}(A) d_i(X^{-1}). \quad (3.11)$$

Now we insert (3.11) into (3.10). For this purpose let

$$\gamma_j = \left\{ \prod_{i=1}^j t_{l+i-1}(S) \right\}^{1/j}$$

be the non-increasing sequence of geometric means of $t_l(S), t_{l+1}(S), \dots$. Then

$$2^{-(n-1)j/\gamma_j} \leq 2^{-(n-1)/6} \gamma_{3k} \leq 2^{-(n-1)/6} t_l(S) \quad \text{for } j = 4, 5, 6;$$

$$2^{-(n-1)j/\gamma_j} \leq 2^{-(n-1)(3k+3)/\gamma_{3k}} \leq 2^{-(n-1)(3k+3)} \left\{ \prod_{i=1}^k t_{l+3i-3}(S) \right\}^{1/k}$$

$$\text{for } j = 3k + 1, 3k + 2, 3k + 3.$$

Furthermore

$$2^{-(n-1)j/\gamma_j} \leq 2^{-(n-1)/6} t_l(S) \quad \text{for } j = 1, 2, 3.$$

Hence (3.10) becomes

$$\begin{aligned}
 a_{n,l}(S) &\leq 6 \sup_{1 \leq k < \infty} 2^{-(n-1)(3k+3)} \left\{ \prod_{i=1}^k t_{l+3i-3}(S) \right\}^{1/k} \\
 &\leq 7 \sup_{1 \leq k < \infty} 2^{-n(3k+3)} \left\{ \prod_{i=1}^k t_{l+3i-3}(S) \right\}^{1/k}, \tag{3.12}
 \end{aligned}$$

where non-zero terms occur only for $l + 3k - 3 \leq r$, i.e., $1 \leq k \leq M$, cf. (3.2). From (3.11) and (3.12) we obtain

$$a_{n,l}(S) \leq 7 \sup_{1 \leq k \leq M} 2^{-n(3k+3)} \left\{ \prod_{i=1}^k c_i(Y) d_i(X^{-1}) \right\}^{1/k} \left\{ \prod_{i=1}^k t_{l+i-1}(A) \right\}^{1/k}.$$

The multiplicativity (SC v) and (3.4), (3.5) imply that

$$a_{3n-2,l}(A) \leq a_{3n-2,l}(A_0) \leq e_n(Y^{-1}) a_{n,l}(S) e_n(X);$$

therefore

$$\begin{aligned}
 a_{3n-2,l}(A) &\leq 7 e_n(X) e_n(Y^{-1}) \sup_{1 \leq k \leq M} \left[2^{-n(3k+3)} \right. \\
 &\quad \left. \times \left\{ \prod_{i=1}^k c_i(Y) d_i(X^{-1}) \right\}^{1/k} \left\{ \prod_{i=1}^k t_{l+i-1}(A) \right\}^{1/k} \right]. \blacksquare
 \end{aligned}$$

Remark 3.1. If E and F are complex Banach spaces we must replace

$$2^{-n(3k+3)} \text{ by } 2^{-n(6k+6)} \text{ in (3.1).}$$

From this lemma we derive some conclusions. Let us first recall the definition of local injective, resp. surjective distances, based on the Banach-Mazur distance $d(E, F)$ of Banach spaces E, F :

$$d(E, F) = \begin{cases} \inf \{ \|T\| \|T^{-1}\| \text{ for all isomorphisms } T \in L(E, F) \} \\ \infty \text{ if } E \text{ and } F \text{ are not isomorphic.} \end{cases}$$

The n th local injective distance $\delta_n(E)$ is given as

$$\delta_n(E) = \sup \{ d(M, \ell_2^m) : M \subseteq E, m = \dim(M) \leq n \},$$

and the n th local surjective distance $\delta^{(n)}(E)$ as

$$\delta^{(n)}(E) = \sup \{ d(E/M, \ell_2^m) : M \subseteq E, m = \text{codim}(M) \leq n \}.$$

COROLLARY 3.1. *Under the assumptions of Lemma 3.1 the following inequality holds*

$$a_{n,l}(A) \leq 7 \delta^{(r)}(E) \delta_r(F) \sup_{1 \leq k \leq M} 2^{-n(9k+9)} \left\{ \prod_{i=1}^k t_{l+i-1}(A) \right\}^{1/k}. \tag{3.13}$$

Proof. Given $\varepsilon > 0$ we choose X and Y in Lemma 3.1 with

$$\|X\| \|X^{-1}\| \leq (1 + \varepsilon) d(E/N, \ell_2^r) \leq (1 + \varepsilon) \delta^{(r)}(E)$$

$$\|Y\| \|Y^{-1}\| \leq (1 + \varepsilon) d(R, \ell_2^r) \leq (1 + \varepsilon) \delta_r(F).$$

All terms in (3.1) containing X or Y are estimated by the operator norms. Hence for $3m - 2 \leq n \leq 3m$ we have

$$\begin{aligned} a_{n,l}(A) &\leq a_{3m-2,l}(A) \leq 7(1 + \varepsilon)^2 \delta^{(r)}(E) \delta_r(F) \\ &\quad \times \sup_{1 \leq k \leq M} 2^{-m/(3k+3)} \left\{ \prod_{i=1}^k t_{l+i-1}(A) \right\}^{1/k}. \end{aligned}$$

Since $m \geq n/3$ the proof is complete.

Now an important result of Pisier [P1, P2] is used to obtain further consequences of Lemma 3.1.

THEOREM 3.1. (Pisier). *For each $\alpha > \frac{1}{2}$ there is a constant $C(\alpha)$ such that for any n -dimensional (real or complex) Banach space E there is an isomorphism X from E onto ℓ_2^n (real or complex, respectively), such that*

$$d_k(X) \leq C(\alpha) \left(\frac{n}{k}\right)^\alpha \quad \text{and} \quad d_k(X^{-1}) \leq C(\alpha) \left(\frac{n}{k}\right)^\alpha \quad (3.14)$$

for all k , $1 \leq k \leq n$. (For $k > n$ we have $d_k(X) = c_k(X^{-1}) = 0$ in any case.) Moreover, the constant $C(\alpha)$, only depending on α , can be chosen of order $O(\alpha - \frac{1}{2})^{-1}$ for $\alpha \downarrow \frac{1}{2}$. ■

The corresponding result concerning the dyadic entropy numbers of X follows immediately from Theorem 3.1 and the following inequality from [C] (cf. [CS, P2]).

THEOREM 3.2. *For each p , $p > 0$, there is a constant c_p such that for all operators T we have*

$$\sup_{1 \leq k \leq n} k^{1/p} e_k(T) \leq c_p \sup_{1 \leq k \leq n} k^{1/p} t_k(T), \quad 1 \leq n < \infty, \quad (3.15)$$

and c_p remains bounded if p varies in a compact subset of $(0, \infty)$.

If we note that $t_k(X) = d_k(I_\infty X) \leq d_k(X)$ and analogously $t_k(X^{-1}) = c_k(X^{-1}Q_1) \leq c_k(X^{-1})$ we obtain from (3.14) and (3.15), by setting $p = 1/\alpha$,

$$\begin{aligned} k^\alpha e_k(T) &\leq c_{1/\alpha} \sup_{1 \leq k \leq n} k^\alpha t_k(T) \leq c_{1/\alpha} \sup_{1 \leq k \leq n} k^\alpha C(\alpha) \left(\frac{n}{k}\right)^\alpha \\ &= c_{1/\alpha} C(\alpha) n^\alpha, \quad 1 \leq k \leq n. \end{aligned}$$

Hence we may state the following corollary.

COROLLARY 3.2. *Let $X: E \rightarrow \ell_2^n$ be the isomorphism from Theorem 3.1. Then*

$$\max\{e_k(X), e_k(X^{-1})\} \leq c_{1,\alpha} C(\alpha) \left(\frac{n}{k}\right)^\alpha \quad \text{for all } k, 1 \leq k < \infty, \quad (3.16)$$

where $C(\alpha)$ and $c_{1,\alpha}$ are the constants of (3.14) and (3.15). Moreover, for a modified $C(\alpha)$, which is also of order $O(\alpha - \frac{1}{2})^{-1/2}$ for $\alpha \downarrow \frac{1}{2}$, we have

$$\max\{e_k(X), e_k(X^{-1}), d_k(X), c_k(X^{-1})\} \leq C(\alpha) \left(\frac{n}{k}\right)^\alpha \quad \text{for } 1 \leq k < \infty. \quad (3.17)$$

These considerations lead to a more sophisticated corollary of Lemma 3.1.

COROLLARY 3.3. *For any $\beta, \beta > 1$, there exists a constant $C(\beta)$ of order $C(\beta) = O(\beta - 1)^{-2}$ for $\beta \downarrow 1$ such that for any operator $A \in L(E, F)$ with $\text{rank}(A) = r$ between real Banach spaces E and F*

$$a_{n,l}(A) \leq C(\beta) \left(\frac{r}{n}\right)^{2\beta} \sup_{1 \leq k \leq M} 2^{-n \cdot (9k+9)} \left(\frac{n}{9k+9}\right)^\beta \left\{ \prod_{i=1}^k t_{l+i-1}(A) \right\}^{1/k}, \quad (3.18)$$

for all $l, n, 1 \leq l, n < \infty, M$ being given by (3.2).

Proof. Put $\alpha = \beta/2 > \frac{1}{2}$. Then by 3.17 there are X and Y such that

$$\max\{e_k(X), e_k(Y^{-1}), c_k(Y), d_k(X^{-1})\} \leq C(\alpha) \left(\frac{r}{k}\right)^\alpha, \quad 1 \leq k < \infty. \quad (3.19)$$

We know that $C(\alpha) = O(\alpha - \frac{1}{2})^{-1/2} = O(\beta - 1)^{-1/2}$ for $\beta \downarrow 1$. Hence it follows from (3.1) that

$$a_{3n-2,l}(A) \leq 7e^2 C^4(\alpha) \left(\frac{r}{n}\right)^\beta \sup_{1 \leq k \leq M} 2^{-n \cdot (3k+3)} \left(\frac{r}{k}\right)^\beta \left\{ \prod_{i=1}^k t_{l+i-1}(A) \right\}^{1/k},$$

where $k! \geq (k/e)^k$ is used. Estimating

$$\left(\frac{1}{k}\right)^\beta \leq 18^\beta \left(\frac{n}{9k+9}\right)^\beta$$

and passing from $a_{3n-2,l}(A)$ to $a_{n,l}(A)$ as in the proof of Corollary 3.1, we

finally obtain the inequality (3.18) with the constant $C(\beta) = 7e^2 18^\beta C^4(\alpha)$ which is indeed of order $O(\beta - 1)^{-2}$ for $\beta \downarrow 1$. ■

Remark 3.2. If the Banach spaces X and Y are complex Banach spaces then $2^{-n/(9k+9)}$ in (3.13) and (3.18) must be replaced by $2^{-n/(18k+18)}$.

Now let us remove the restriction on A being a finite rank operator.

PROPOSITION 3.1. *Let $\beta > 1$ and $T \in L(E, F)$ be an operator between real Banach spaces. Then*

$$a_{n,l}(T) \leq a_{3m+l}(T) + D(\beta) \left(\frac{m+l}{n}\right)^{2\beta} \sup_{1 \leq k \leq m} \left[2^{-n/(9k+9)} \times \left(\frac{n}{9k+9}\right)^\beta \left\{ \prod_{i=1}^k a_{l+i-1}(T) \right\}^{1:k} \right] \quad (3.20)$$

for $1 \leq l, m, n < \infty$ with a constant $D(\beta)$, depending only on β and being of order $O(\beta - 1)^{-2}$ for $\beta \downarrow 1$.

Proof. Given $\varepsilon > 0$ we determine an operator $A \in L(E, F)$ with $r = \text{rank}(A) < 3m + 1$ such that

$$\|T - A\| \leq (1 + \varepsilon) a_{3m+l}(T). \quad (3.21)$$

From

$$a_{n,l}(T) \leq \|T - A\| + a_{n,l}(A)$$

(cf. (SC iii) and by applying (3.18) to A and using $r < 3(m+l)$ and $M = \max\{1, [1 + (r-l)/3]\} \leq m$ we obtain

$$a_{n,l}(T) \leq (1 + \varepsilon) a_{3m+l}(T) + 9^\beta C(\beta) \left(\frac{m+l}{n}\right)^{2\beta} \sup_{1 \leq k \leq m} \left[2^{-n/(9k+9)} \times \left(\frac{n}{9k+9}\right)^\beta \left\{ \prod_{i=1}^k t_{l+i-1}(A) \right\}^{1:k} \right]. \quad (3.22)$$

In order to eliminate $t_{l+i-1}(A)$ we derive from (3.21)

$$\begin{aligned} t_{l+i-1}(A) &\leq \|A - T\| + t_{l+i-1}(T) \leq (1 + \varepsilon) a_{3m+l}(T) + a_{l+i-1}(T) \\ &\leq (2 + \varepsilon) a_{l+i-1}(T) \quad \text{for } 1 \leq i \leq m. \end{aligned}$$

Hence, since for $\beta \downarrow 1$

$$D(\beta) = 2 \cdot 9^\beta C(\beta) = O(\beta - 1)^{-2}$$

(3.22) implies the desired estimate (3.20). ■

The next inequality relates $a_{n,l}(T)$ to the $\ell_{p,\infty}$ -norm of a finite subsequence of the approximation number sequence.

PROPOSITION 3.2. *Let $0 < p < \infty$ and $T \in L(E, F)$ (E, F real Banach spaces). Then*

$$a_{n,l}(T) \leq a_{l+m-1}(T) + C(p) \left(\frac{m+l}{n}\right)^2 \log^2\left(\frac{m+l}{n} + 2\right) \times n^{-1/p} \sup_{1 \leq k \leq m} k^{1/p} a_{l+k-1}(T), \tag{3.23}$$

for all $l, m, n, 1 \leq l, m, n < \infty$, where $C(p)$ is a constant depending only on p .

Proof. Because of $k! \geq (k/e)^k$ we have

$$\begin{aligned} & \left\{ \prod_{i=1}^k a_{l+i-1}(T) \right\}^{1/k} \\ & \leq (k!)^{-1/pk} \sup_{1 \leq j \leq k} j^{1/p} a_{l+j-1}(T) \\ & \leq \left(\frac{e}{k}\right)^{1/p} \sup_{1 \leq k \leq m} k^{1/p} a_{l+k-1}(T) \quad \text{for } 1 \leq k \leq m. \end{aligned} \tag{3.24}$$

We now insert (3.24) into (3.20) and choose an appropriate $\beta > 1$. Note that for the first summand on the right-hand side of (3.20)

$$a_{l+m-1}(T) \geq a_{3m+l}(T).$$

For the second summand we apply (3.24) and get from (3.20)

$$\begin{aligned} a_{n,l}(T) & \leq a_{l+m-1}(T) + e^{1/p} D(\beta) \left(\frac{m+l}{n}\right)^{2\beta} \left\{ \sup_{1 \leq k \leq m} k^{1/p} a_{l+k-1}(T) \right\} \\ & \times \left\{ \sup_{1 \leq k \leq m} 2^{-n \cdot (9k+9)} \left(\frac{n}{9k+9}\right)^\beta k^{-1/p} \right\}. \end{aligned} \tag{3.25}$$

Estimating $k^{-1/p}$ by

$$k^{-1/p} \leq 18^{1/p} n^{-1/p} \left(\frac{n}{9k+9}\right)^{1/p}$$

and setting $x = n/(9k+9)$, we may rewrite (3.25) as

$$\begin{aligned} a_{n,l}(T) & \leq a_{l+m-1}(T) + (18e)^{1/p} D(\beta) \left(\frac{m+l}{n}\right)^{2\beta} n^{-1/p} \\ & \times \left\{ \sup_{1 \leq k \leq m} k^{1/p} a_{l+k-1}(T) \right\} \left\{ \sup_{0 < x < \infty} 2^x x^{\beta+1/p} \right\}. \end{aligned} \tag{3.26}$$

With the function $\sigma(t) := \sup_{0 < x < \infty} 2^{-x} x^t = (t/e \cdot \ln 2)^t$ and the constant $C(\beta, p) := (18e)^{1/p} D(\beta) \sigma(\beta + 1/p)$ we get from (3.26)

$$a_{n,l}(T) \leq a_{l-m-1}(T) + C(\beta, p) \left(\frac{m+l}{n}\right)^{2(\beta-1)} \left(\frac{m+l}{n}\right)^2 n^{-1/p} \times \left\{ \sup_{1 \leq k \leq m} k^{1/p} a_{l+k-1}(T) \right\}. \tag{3.27}$$

Fixing p , we see from the definition of $C(\beta, p)$ that $C(\beta, p) = O(\beta - 1)^{-2}$ for $\beta \downarrow 1$. If we set

$$\beta = 1 + \left[\log \left(\frac{m+l}{n} + 2 \right) \right]^{-1} \tag{3.28}$$

then $1 < \beta < 3$ for all $l, m, n \geq 1$ and thus there is a constant $C(p)$ such that

$$C(\beta, p) \leq C(p) (\beta - 1)^{-2} \quad \text{for } 1 < \beta < 3.$$

With β given in (3.28) it is easy to verify that

$$C(\beta, p) \left(\frac{m+n}{n}\right)^{2(\beta-1)} \leq e^2 C(p) \left[\log \left(\frac{m+l}{n} + 2 \right) \right]^2.$$

Together with (3.27) this completes the proof. ■

Remark 3.3. In case of complex Banach spaces the value $2^{-n/(9k+9)}$ must be replaced by $2^{-n/(18k+8)}$ in (3.20), whereas (3.23) holds in the real and complex cases (clearly $C(p)$ must be modified).

The next theorems deal with inequalities related to Theorem 3.2.

THEOREM 3.3. *Let $0 < p < \infty$ and $T \in L(E, F)$, where E and F can be either real or complex Banach spaces. Then there exists a constant $C(p)$, only depending on p , such that*

$$\sup_{1 \leq n \leq m} n^{1/p} a_{n,l}(T) \leq C(p) \sup_{1 \leq n \leq m} n^{1/p} a_{l+n-1}(T) \tag{3.29}$$

for all $l, m, 1 \leq l \leq m < \infty$.

THEOREM 3.4. *Let $0 < p < \infty$ and $T \in L(E, F)$. Then there exists a constant $C(p)$ such that*

$$\sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{n,l}(T) \leq C(p) \sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{l+n-1}(T) \tag{3.30}$$

for all $l, m, 1 \leq l, m < \infty$.

Proofs. (Theorem 3.3) Suppose that n is given such that $1 \leq n \leq m$. We apply (3.23) (if E and F are complex Banach spaces cf. Remark 3.3) for l, m, n with $m = n$

$$\begin{aligned} n^{1/p} a_{n,l}(T) &\leq n^{1/p} a_{l+n-1}(T) \\ &\quad + 4 \log^2 4 \cdot C(p) \sup_{1 \leq k \leq n} k^{1/p} a_{l+k-1}(T). \end{aligned}$$

For $\tilde{C}(p) = 1 + 4 \log^2 4 \cdot C(p)$ we obtain

$$\begin{aligned} n^{1/p} a_{n,l}(T) &\leq \tilde{C}(p) \sup_{1 \leq k \leq n} k^{1/p} a_{l+k-1}(T) \\ &\leq \tilde{C}(p) \sup_{1 \leq n \leq m} n^{1/p} a_{l+n-1}(T), \end{aligned}$$

that is, (3.29).

(Theorem 3.4) According to (3.29) we have for $1 \leq m$

$$\begin{aligned} &\sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{n,l}(T) \\ &\leq \sup_{1 \leq n \leq m} (2n)^{1/p} a_{n,l}(T) \leq 2^{1/p} C(p) \sup_{1 \leq n \leq m} n^{1/p} a_{l+n-1}(T) \\ &\leq 2^{1/p} C(p) \sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{l+n-1}(T). \end{aligned} \quad (3.31)$$

Without any restrictions on l and m , we have

$$\begin{aligned} &\sup_{1 \leq n < l} (l+n-1)^{1/p} a_{n,l}(T) \\ &\leq (2l)^{1/p} a_{1,l}(T) = (2l)^{1/p} a_l(T) \leq 2^{1/p} \sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{l+n-1}(T). \end{aligned} \quad (3.32)$$

Combining (3.31) and (3.32) we obtain (3.30). More precisely, for $l \leq m$ we use (3.31) and (3.32), and for $l > m$ we start from

$$\sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{n,l}(T) \leq \sup_{1 \leq n < l} (l+n-1)^{1/p} a_{n,l}(T)$$

and apply (3.32). ■

The preceding two theorems are also valid for other s -covering numbers. We only mention the following result which corresponds to Theorem 3.4.

THEOREM 3.5. *Let $0 < p < \infty$ and $T \in L(E, F)$. Then there exists a constant $C(p)$ such that*

$$\sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{n,l}(T) \leq C(p) \sup_{1 \leq n \leq m} (l+n-1)^{1/p} a_{l+n-1}(T) \quad (3.33)$$

for $s = c, d, t$ (cf. (2.2)–(2.4)) and all $l, m, 1 \leq l, m < \infty$.

Proof. For $s = t$ the desired estimate is an obvious consequence of Theorem 3.4 and the definition of t . For the other cases $s = c, d$ recall the definition of Gelfand and Kolmogorov numbers based on the approximation numbers

$$c_n(T) = a_n(I_\infty T),$$

$$d_n(T) = a_n(TQ_1),$$

where Q_1 and I_∞ are the canonical surjection and injection, respectively (cf. [CS]). In order to prove (3.33) for $s = c, d$ we show that

$$c_{n,l}(T) \leq 2a_{n,l}(I_\infty T) \quad (3.34)$$

$$d_{n,l}(T) \leq a_{n,l}(TQ_1). \quad (3.35)$$

These inequalities are consequences of

$$c_{n,l}(S) \leq a_{n,l}(S) \quad (3.36)$$

$$d_{n,l}(S) \leq a_{n,l}(S) \quad (3.37)$$

for any operator S . Since (3.37) is already proved in [CM] we only give the argument for the proof of (3.36). Let A be an operator with $\text{rank}(A) < l$ such that

$$e_n(S - A) \leq a_{n,l}(S) + \varepsilon = \rho.$$

Hence

$$(S - A)(U_E) \subseteq \bigcup_{i=1}^{2^{n-1}} \{y_i + \rho U_F\}$$

for $y_i \in F, 1 \leq i \leq 2^{n-1}$. For $M = N(A)$, the null space of A , we have $\text{codim}(M) < l$ and

$$(SI_M^E)(U_E) \subseteq \bigcup_{i=1}^{2^{n-1}} \{y_i + \rho U_F\},$$

hence

$$c_{n,l}(S) \leq e_n(SI_M^E) \leq a_{n,l}(S) + \varepsilon.$$

Because of the surjectivity (resp. injectivity) up to a factor 2 of the dyadic entropy numbers we have

$$\begin{aligned}c_{n,t}(T) &\leq 2c_{n,t}(I_x T) \leq 2a_{n,t}(I_x T) \\d_{n,t}(T) &\leq d_{n,t}(TQ_1) \leq a_{n,t}(TQ_1).\end{aligned}$$

These inequalities complete the proof of (3.33) for $s = c, d$. ■

4. INEQUALITIES BETWEEN s -COVERING NUMBERS, GAUSSIAN AVERAGES, AND RADEMACHER AVERAGES

This section is devoted to the investigation of inequalities between d -covering numbers of operators from a Hilbert space into a Banach space on one side and Gaussian or Rademacher averages on the other side. These inequalities complement and generalize V. Milman's discovering that the Gaussian average or the ℓ -norm is an appropriate parameter for estimating Gelfand and Kolmogorov numbers [M1, M2].

For this purpose recall the definition of the so-called Gaussian average or ℓ -norm of an operator T from ℓ_2^n into an arbitrary Banach space E . The ℓ -norm $\ell(T)$ of T is defined as

$$\ell(T) = \left(\int_{\mathbb{R}^n} \|Tx\|^2 d\gamma_n(x) \right)^{1/2}, \quad (4.1)$$

where γ_n denotes the canonical (normalized) Gaussian measure on the euclidean space \mathbb{R}^n . Moreover, for any operator T from ℓ_2 into E we define $\ell(T)$ as

$$\ell(T) = \sup \{ \ell(TX) : X \in L(\ell_2^n, \ell_2) \text{ for some } n, \|X\| \leq 1 \}. \quad (4.2)$$

We use a minoration of $\ell(T)$ which originated in the paper of Milman [M1] and was improved by A. Pajor and N. Tomczak-Jaegermann [PaT] (cf. [G, P2]) in the following theorem.

THEOREM 4.1. *Let E be a Banach space and let $T \in L(\ell_2, E)$ be a compact operator. Then*

$$\sup_{1 \leq k < \infty} k^{1/2} d_k(T) \leq C \cdot \ell(T), \quad (4.3)$$

where $C \geq 1$ is a universal constant.

Remark 4.1. Note that there is a dual version (4.3) for compact operators $T \in L(E, \ell_2)$, namely

$$\sup_{1 \leq k < \infty} k^{1/2} c_k(T) \leq C \cdot \ell(T'), \quad (4.4)$$

which is based upon $c_k(T) = d_k(T')$ for all operators T . But (4.3) is also a consequence of (4.4) because of $d_k(T) = c_k(T')$ for compact operators and $\ell(T) = \ell(T')$, for $T \in L(\ell_2, E)$.

Combining (4.3) with Theorem 3.5 for $s = d$ we may state the following minoration theorem of the Gaussian average (ℓ -norm) by d -covering numbers.

THEOREM 4.2. *Let E be a Banach space and let $T \in L(\ell_2, E)$ be a compact operator. Then*

$$\sup_{1 \leq l, n < \infty} (l+n-1)^{1/p} d_{n,l}(T) \leq C \cdot \ell(T), \quad (4.5)$$

where $C \geq 1$ is a universal constant.

The inequality (4.5) includes the classical Sudakov minoration theorem [Su] for $l = 1$ and the inequality (4.3) for $n = 1$. The version corresponding to (4.4) is the minoration of $\ell(T')$, $T \in L(E, \ell_2)$ by c -covering numbers:

$$\sup_{1 \leq l, n < \infty} (l+n-1)^{1/p} c_{n,l}(T) \leq C \cdot \ell(T). \quad (4.6)$$

Next we want to derive similar inequalities for Rademacher averages instead of Gaussian averages. For this purpose let $\Phi = \{f_1, \dots, f_m\}$ be an orthonormal basis of ℓ_2^m and $T \in L(\ell_2^m, E)$ be an operator. The Rademacher average or r -norm of T with respect to Φ is given by

$$\iota_\Phi(T) = \left(\text{Average}_{\varepsilon_i = \pm 1} \left\| \sum_{i=1}^m \varepsilon_i T(f_i) \right\|^2 \right)^{1/2}.$$

It is well-known that $\iota_\Phi(T) \leq c \cdot \ell(T)$ for some universal constant c , independent of the special choice of Φ . The following minoration theorem for the Rademacher averages was proved in [CP].

THEOREM 4.3. *Let E be a Banach space and let $T \in L(\ell_2^m, E)$ be an operator of rank n . Then*

$$\sup_{1 \leq k \leq n} \left(\log \left(1 + \frac{n}{k} \right) \right)^{1/2} k^{1/2} d_k(T) \leq C \cdot \iota_\Phi(T) \quad (4.7)$$

for any orthonormal basis Φ in ℓ_2^m . The constant C is a universal constant.

Combining Theorem 3.5 with Theorem 4.3, we obtain the Rademacher version of Theorem 4.2.

THEOREM 4.4. *Let E be a Banach space and let $T \in L(\ell_2^m, E)$ be an operator of rank n . Then*

$$\sup_{1 \leq k, l < j} \left(\log \left(1 + \frac{n}{k+l-1} \right) \right)^{1.2} (k+l-1)^{1.2} d_{k,l}(T) \leq C \cdot \gamma_\Phi(T) \quad (4.8)$$

for a universal constant C and any orthonormal basis Φ in ℓ_2^m .

Proof. Let $k \geq 1$. From (3.33) we get

$$\begin{aligned} & (l+k-1) d_{k,l}(T) \\ & \leq C(1) \sup_{1 \leq j \leq k} (l+j-1) d_{l+j-1}(T) \\ & \leq C(1) \sup_{1 \leq j \leq k} (l+j-1)^{1.2} \left(\log \left(1 + \frac{n}{k+j-1} \right) \right)^{1.2} \\ & \quad \times \sup_{1 \leq j \leq k} (l+j-1)^{1.2} \left(\log \left(1 + \frac{n}{l+j-1} \right) \right)^{1.2} d_{l+j-1}(T). \end{aligned}$$

The first supremum on the right-hand side equals

$$(l+k-1)^{1.2} \left(\log \left(1 + \frac{n}{l+k-1} \right) \right)^{1.2}$$

since $x \rightarrow x \cdot \log(1 + n/x)$ is an increasing function. The second supremum can be estimated by $C \cdot \gamma_\Phi(T)$ according to (4.7). Hence

$$(l+k-1)^{1.2} \left(\log \left(1 + \frac{n}{l+k-1} \right) \right)^{1.2} d_{k,l}(T) \leq C \cdot C(1) \cdot \gamma_\Phi(T). \quad \blacksquare$$

There are dual versions of (4.7) and (4.8) for operators $T \in L(E, \ell_2^m)$, of rank n , namely

$$\sup_{1 \leq k \leq n} \left(\log \left(1 + \frac{n}{k} \right) \right)^{1.2} k^{1.2} c_k(T) \leq C \cdot \gamma_\Phi(T') \quad (4.9)$$

and

$$\sup_{1 \leq l, k < j} \left(\log \left(1 + \frac{n}{k+l-1} \right) \right)^{1.2} (k+l-1)^{1.2} c_{k,l}(T) \leq C \cdot \gamma_\Phi(T'). \quad (4.10)$$

The last inequality can be interpreted as follows:

Let $B \subset \mathbb{R}^n$ be a compact, convex, and symmetric set with non-empty interior. Equipped with the Minkowski functional $\|\cdot\|_B$ of B the vector space \mathbb{R}^n becomes a normed vector space which is isomorphic to ℓ_2^n and admitting B as its unit ball. Let us apply (4.10) to the canonical isomorphism

$$T: (\mathbb{R}^n, \|\cdot\|_B) \rightarrow (\mathbb{R}^n, |\cdot|_2) = \ell_2^n.$$

COROLLARY 4.1. For $0 < p < \infty$ there exists a constant C_p , only depending on p , and for each pair (k, l) , $1 \leq k < \infty$, $1 \leq l < n$, there exists an l -dimensional subspace $M \subset \mathbb{R}^n$ such that

$$e_k(M \cap B) \leq C_p \cdot \left\{ \text{Average}_{\varepsilon_i = \pm 1} \sup_{t = (t_1, \dots, t_n) \in B} \left| \sum_{i=1}^n \varepsilon_i t_i \right|^p \right\}^{1/p} \times \left(\log \left(1 + \frac{n}{n+k-1} \right) \right)^{1/2} (n+k-1)^{-1/2}. \quad (4.11)$$

Therefore $M \cap B$ is covered by 2^{k-1} euclidean balls of radius ρ , where ρ is bounded by the right-hand side of (4.11). In particular, for each l , $1 \leq l \leq n$, there is an l -dimensional subspace $M \subset \mathbb{R}^n$ such that

$$e_l(M \cap B) \leq C_p \cdot \left\{ \text{Average}_{\varepsilon_i = \pm 1} \sup_{t = (t_1, \dots, t_n) \in B} \left| \sum_{i=1}^n \varepsilon_i t_i \right|^p \right\}^{1/p} n^{-1/2}. \quad (4.12)$$

For operators $T \in L(\ell_1^n, \ell_2)$ we have the following modification of (4.9), which was proved in [CP].

THEOREM 4.5. There exists a universal constant C such that

$$\sup_{1 \leq k \leq n} \left(\log \left(1 + \frac{n}{k} \right) \right)^{-1/2} k^{1/2} c_k(T) \leq C \|T\| \quad (4.13)$$

for all $T \in L(\ell_2^n, \ell_2)$ and all $n \in \mathbb{N}$.

As in Theorem 4.4, we derive from (4.13) and Theorem 3.5

$$\sup_{1 \leq k, l < \infty} \left(\log \left(1 + \frac{n}{k+l-1} \right) \right)^{-1/2} (k+l-1)^{1/2} c_{k,l}(T) \leq C \|T\| \quad (4.14)$$

for $T \in L(\ell_1^n, \ell_2)$.

If we take the volume ratio inequality

$$\left[\frac{\text{vol}_m(A)}{\text{vol}_m(U_{\ell_2^n})} \right]^{1/m} \leq \inf_{1 \leq k < \infty} k^{1/m} \varepsilon_k(A)$$

into consideration and denote the absolutely convex hull of the points $x_1, \dots, x_n \in \ell_2^n$, i.e., the set

$$\left\{ \sum_{i=1}^n \lambda_i x_i : \sum_{i=1}^n |\lambda_i| \leq 1 \right\},$$

by $\text{aconv}(x_1, \dots, x_n)$, then we can derive the following corollary from (4.14).

COROLLARY 4.2. Let $x_1, \dots, x_n \in \ell_2^n$ be given, such that $B = \text{aconv}(x_1, \dots, x_n)$ has a non-empty interior. For each m , $1 \leq m < n$, there exists an m -dimensional subspace $M \subset \ell_2^n$ with

$$\left[\frac{\text{vol}_m(M \cap B)}{\text{vol}_m(U_{\ell_2^m})} \right]^{1:m} \leq C \cdot \left(\log \left(1 + \frac{n}{n-m+1} \right) \right)^{+1.2} \\ \times (n-m+1)^{-1.2} \max_{1 \leq i \leq n} \|x_i\|_2,$$

where C is a universal constant.

REFERENCES

- [C] B. CARL, Entropy numbers, s -numbers, and eigenvalue problems, *J. Funct. Anal.* **41** (1981), 290–306.
- [CP] B. CARL AND A. PAJOR, Gelfand numbers of operators with values in a Hilbert space, *Invent. Math.* **94** (1988), 479–504.
- [CM] B. CARL AND E. MAKAI, JR., New covering numbers and spectral properties of operators in Banach spaces, Preprint No. 39, Math. Inst. Hungarian Acad. Sciences, 1988.
- [CS] B. CARL AND I. STAPHANI, "Entropy, Compactness and the Approximation of Operators," Cambridge Univ. Press, London/New York, 1990.
- [G] Y. GORDON, On Milman's inequality and random subspaces which escape through a mesh in \mathbb{R}^n , in "GAFA Israel Funct. Analysis Seminar (1986/1987)," Springer Lecture Notes 1317, pp. 84–106, Springer-Verlag, Berlin/New York, 1988.
- [GKS] Y. GORDON, H. KÖNIG, AND C. SCHÜTT, Geometric and probabilistic estimates for entropy and approximation numbers of operators, *J. Approx. Theory* **49** (1987), 219–239.
- [M1] V. MILMAN, Random subspaces of proportional dimension of finite dimensional normed spaces, approach through the isoperimetric inequality, in "Séminaire d'Analyse Fonction (1984/1985)," Université Paris VI, Paris.
- [M2] V. MILMAN, Volume approach and iteration procedures in local theory of normed spaces, in "Proceedings, Missouri Conf. 1984," Springer Lecture Notes 1166, pp. 99–105, Springer-Verlag, Berlin/New York, 1985.
- [PaT] A. PAJOR AND N. TOMCZAK-JAEGERMANN, Subspaces of small co-dimension of finite dimensional Banach spaces, *Proc. Amer. Math. Soc.* **97** (1986), 637–642.
- [P1] G. PISIER, A new approach to several results of V. Milman, *J. Reine Angew. Math.* **393** (1989), 115–131.
- [P2] G. PISIER, "Volume Inequalities in the Geometry of Banach Spaces," Cambridge Univ. Press, London/New York, 1989.
- [Pi] A. PIETSCH, "Operator Ideals," AdW, Berlin, 1978.
- [Su] V. N. SUDAKOV, Gaussian processes and measures of solid angles in Hilbert space, *Soviet Math. Dokl.* **12** (1971), 412–415.