ZEROS OF ANALYTIC FUNCTIONS AND NORMS OF INVERSE MATRICES

ΒY

E. GLUSKIN*

The Raymond and Beverly Sackler Faculty of Exact Sciences Tel Aviv University, Tel Aviv 69978, Israel

AND

M. Meyer

Equipe d'Analyse, Université Paris VI 4 Place Jussieu, 75252 Paris Cedex 05, France

AND

A. Pajor

Equipe de Mathématiques, Université de Marne-la-Vallée 2 rue de la Butte Verte, 93166 Noisy-le-Grand Cedex, France

ABSTRACT

Let k_n be the smallest constant such that for any *n*-dimensional normed space X and any invertible linear operator $T \in \mathcal{L}(X)$ we have

 $|\det(T)| \cdot ||T^{-1}|| \le k_n ||T||^{n-1}.$

Let A_+ be the Banach space of all analytic functions $f(z) = \sum_{k\geq 0} a_k z^k$ on the unit disk D with absolutely convergent Taylor series, and let $||f||_{A_+} = \sum_{k\geq 0} |a_k|$; define φ_n on \overline{D}^n by

$$\varphi_n(\lambda_1, \dots, \lambda_n) = \inf \left\{ \|f\|_{A_+} - |f(0)|; \ f(z) = g(z) \prod_{i=1}^n (\lambda_i - z), \ g \in A_+, \ g(0) = 1 \right\}$$

* Research done while this author was visiting the University of Paris VII and I.H.E.S. in September 1991, partially supported by GIF grant. Received January 16, 1993 and in revised form June 21, 1993 We show that $k_n = \sup \left\{ \varphi_n(\lambda_1, \ldots, \lambda_n) ; (\lambda_1, \ldots, \lambda_n) \in \overline{D}^n \right\}$. Moreover, if S is the left shift operator on the space ℓ_{∞} : $S(x_0, x_1, \ldots, x_p, \ldots) = (x_1, \ldots, x_p, \ldots)$ and if $\mathbf{J}_n(S)$ denotes the set of all S-invariant n-dimensional subspaces of ℓ_{∞} on which S is invertible, we have

$$k_n = \sup\{ |\det(S|_E)| | ||(S|_E)^{-1}|| ; E \in \mathbf{J}_n(S) \}.$$

J. J. Schäffer (1970) proved that $k_n \leq \sqrt{en}$ and conjectured that $k_n = 2$, for $n \geq 2$. In fact $k_3 > 2$ and using the preceding results, we show that, up to a logarithmic factor, k_n is of the order of \sqrt{n} when $n \to +\infty$.

Introduction

Let X be an n-dimensional normed space. The number k(X) is defined as the smallest constant k such that the following inequality

$$\left\|\det(T)\right\| \|T^{-1}\| \le k \|T\|^{n-1}$$

holds for any invertible linear operator $T \in \mathcal{L}(X)$. This characteristic k(X)of X was introduced and studied by Schäffer [S] for real normed spaces. It is interesting to observe that the operator $(\det(T))T^{-1}$, also called the adjugate operator of T, coincides with the differential at T of the function $S \to \det(S)$. The starting point of Schäffer's investigation was the remark that for a finite dimensional Hilbert space H we have k(H) = 1. He proved that, if ℓ_1^n is \mathbb{R}^n equipped with the norm $||x|| = \sum_{i=1}^n |x_i|$, one has $k(\ell_1^n) = 2$, for $n \ge 2$. He conjectured that ℓ_1^n was an extremal case, that is, if we define k_n by

 $k_n = \sup\{k(X); X \text{ is an } n \text{-dimensional normed space}\},\$

then $k_n = k(\ell_1^n) = 2$. He proved this conjecture for n = 2 and then obtained the following estimate

$$k(X) \leq c\sqrt{n}$$

for any real *n*-dimensional normed space X (from now on c will be a constant independent of n).

The question whether the sequence k_n is bounded remained open. As we shall see further, the answer is negative, and we have already $k_3 \ge 9/4 > 2$ (see Example 11 below).

The problem of the computation of k_n has an analytic reformulation as follows. Denote by A_+ the Banach space of all analytic functions on the unit disk $D = \{z \in \mathbf{C}; |z| < 1\}$, with absolutely convergent Taylor series. The norm of the function $f(z) = \sum_{k\geq 0} a_k z^k$ of A_+ is given by

$$\|f\|_{A_+} = \sum_{k \ge 0} |a_k|.$$

Let us define a function $\varphi_n : \overline{D}^n \to \mathbb{R}$ as follows: for $(\lambda_1, \ldots, \lambda_n) \in \overline{D}^n$

$$\varphi_n(\lambda_1,\ldots,\lambda_n) = \inf \{ \|f\|_{A_+} - |f(0)|; \ f(z) = g(z) \prod_{i=1}^n (\lambda_i - z) \},$$

where the infimum is taken over all polynomials g satisfying g(0) = 1.

We shall prove (Theorem 1) that in the complex case

(1)
$$k_n = \sup \left\{ \varphi_n(\lambda_1, \ldots, \lambda_n); \ (\lambda_1, \ldots, \lambda_n) \in \bar{D}^n \right\}.$$

A very similar formula holds also in the real case (see Remark 2 after Theorem 1). Thus the problem of estimating k_n is reduced to finding the good choice of a vector $(\lambda_1, ..., \lambda_n)$ in \overline{D}^n and to computing $\varphi_n(\lambda_1, ..., \lambda_n)$.

This new analytic representation of k_n allows one also to give immediately a negative answer to Schäffer's conjecture.

CLAIM: The sequence k_n is unbounded.

Proof: Suppose that on the contrary there exists a universal constant c' such that $k_n \leq c'$ for any n. Let $(\lambda_k)_{k\geq 1}$ be a sequence in the open unit disk D for which the set of limit points is the unit circle S^1 and $\prod_{k\geq 1} |\lambda_k| = \beta > 0$. Then from our assumption we have

$$\varphi_n(\lambda_1, ..., \lambda_n) \leq c', \quad \text{for every } n.$$

So that there exists polynomials $q_n(z)$ such that

$$q_n(0) = \prod_{k=1}^n |\lambda_k|, \ q_n(\lambda_k) = 0 \quad \text{for } k = 1, ..., n \quad \text{and} \quad ||q_n||_{A_+} \le c' + 1.$$

Since the sequence q_n is uniformly bounded in the space A_+ , it has a subsequence (q_{n_i}) which converges to some $f \in A_+$ in the weak *-topology on $A_+ = \ell_1(\mathbb{N}) =$

Isr. J. Math.

 $c_0(\mathbb{N})^*$. Consequently q_{n_i} converges to f uniformly on compact subsets of D. We have $f(0) = \lim q_n(0) = \beta \neq 0$ and for any $k \ge 1$, $f(\lambda_k) = \lim_i q_{n_i}(\lambda_k) = 0$. By the continuity of f on all \overline{D} , we have thus f(z) = 0 for all $z \in S^1$; since f is analytic over D, it must be 0 identically; but this contradicts $f(0) = \beta \neq 0$.

The question about the precise behaviour of k_n is more delicate. Using probabilistic arguments, we proved that if $(\lambda_1, ..., \lambda_n)$ are taken independently on the circle of radius 1 - 1/n then, for some c > 0, the event

$$\left\{ (\lambda_1,...,\lambda_n); \; \varphi_n(\lambda_1,...,\lambda_n) > rac{c}{\log(\log n)} \sqrt{rac{n}{\log n}} \;
ight\}$$

has a strictly positive probability. This shows that

$$k_n > \frac{c}{\log(\log n)} \sqrt{\frac{n}{\log n}}.$$

When this paper was almost complete, the above result was communicated to J. Bourgain and he gave a very nice and short proof of it, in the stronger form:

$$k_n \ge c' \sqrt{\frac{n}{\log n}}.$$

We are indebted to Bourgain for his permission to present the argument of his proof (Theorem 5)*.

The original argument of the authors is somewhat more complicated, but it gives more information concerning the class A_+ . We present it in the last section.

In this problem, there is some difference between the real and the complex case. To simplify the exposition, we shall deal mainly with the complex case. If needed, we shall explain how to deal with the real case.

1. Preliminaries

As usual D, \overline{D} , S^1 will be the open unit disk, closed unit disk and unit circle, respectively. Together with the notations given in introduction, we shall use also the following ones. Let M_n be the space of $n \times n$ matrices. For $T \in M_n$, we shall denote by P_T its characteristic polynomial, $P_T(z) = \det(T - zI)$, and by

^{*} After this paper was submitted, Theorem 6 was improved by Hervé Queffelec, who proved that $k_n \ge \sqrt{n/2e}$ for every $n \ge 1$, and that $\underline{\lim}_n (k_n/\sqrt{n}) \ge 1$ (to appear in Notes C.R. Acad. Sci. Paris).

 Q_T its minimal polynomial; spec (T) is the set of all eigenvalues of T. If all eigenvalues of T have multiplicity 1, we shall say that T is a matrix with simple spectrum. If the matrix T is invertible, the matrix $\operatorname{adj}(T) = \operatorname{det}(T) T^{-1}$ is called the **adjugate matrix** of T. The matrix $\operatorname{adj}(T)$ is the comatrix of T, with (j, i) cofactor of T as (i, j) entry. It follows that the adjugate of T can be defined for any matrix T. If the rank of T is less than n - 1, then $\operatorname{adj}(T) = 0$ and if $\operatorname{rank}(T) = n - 1$, then $\operatorname{adj}(T)$ is the rank one matrix such that

$$\operatorname{adj}(T) \circ T = 0$$
 and $\operatorname{adj}(T)x = \alpha . x$, for every $x \in \ker(T)$

where α is the product of all nonzero eigenvalues of T. In the sequel, we shall identify the spaces M_n and $\mathcal{L}(\mathbb{C}^n)$ so that, for example, a simple spectrum operator will be an operator with simple spectrum matrix.

Let X be a linear normed space. We denote by $\|.\|_X$ or simply by $\|.\|$ the norm on X. For given normed linear spaces X and Y, $\mathcal{L}(X, Y)$ is the space of all linear operators from X to Y. We denote the usual norm on $\mathcal{L}(X, Y)$ by $\|.\|_{X\to Y}$ or simply by $\|.\|$. We shall write $\mathcal{L}(X)$ instead of $\mathcal{L}(X, X)$. An operator $T \in \mathcal{L}(X)$ is called a **contraction** if $\|T\| \leq 1$. By homogeneity, we rewrite the definition of k(X) given in the introduction as follows:

$$k(X) = \sup\{\|\operatorname{adj}(T)\|; \ T \in \mathcal{L}(X), \ \|T\| \le 1\}.$$

By continuity, this supremum can be restricted to any dense set of contractions. We shall work then with operators with simple spectrum.

Let E^n be a *n*-dimensional linear space (e.g. $E^n = \mathbb{R}^n$ or \mathbb{C}^n). We denote by \mathcal{N} the collection of all norms on E^n . For a given $T \in \mathcal{L}(E^n)$, the subset $\mathcal{N}(T)$ of \mathcal{N} is defined as

$$\mathcal{N}(T) = \{ \|.\| \in \mathcal{N}; \|T\| \le 1 \}.$$

Observe that $\mathcal{N}(T)$ is not empty if and only if the set $\{T^k; k \geq 1\}$ is bounded in $\mathcal{L}(E^n)$ (if $|| \cdot || \in \mathcal{N}$, define $|||x||| = \sup_k ||T^kx||$, then $||| \cdot ||| \in \mathcal{N}(T)$). In this case, we shall say that T is **power bounded** (abr. p.b.). It is easy to see that T is power bounded if and only if spec $(T) \subset \overline{D}$ and Q_T has no multiple zeros on S^1 (use the Jordan canonical form of T).

Definition: For a power bounded operator $T \in \mathcal{L}(E^n)$, define

$$\kappa(T) = \sup\{\|\operatorname{adj}(T)\|; \|.\| \in \mathcal{N}(T)\}$$

Of course, $\kappa(T)$ is a function of the mapping T and does not depend on matrix representation. In particular, $\kappa(T)$ only depends on the Jordan canonical form of T.

From the definition of κ , it follows that

(2)
$$k_n = \sup\{\kappa(T); T \in \mathcal{L}(\mathbb{C}^n), T \text{ is p.b. and has simple spectrum}\}.$$

Equality (1) is proved in the following theorem.

THEOREM 1: Let $T \in \mathcal{L}(\mathbb{C}^n)$ be an invertible power bounded operator with simple spectrum; then if $P_T(z) = \prod_{i=1}^n (\lambda_i - z)$, we have

$$\kappa(T) = \varphi_n(\lambda_1, ..., \lambda_n).$$

In particular,

$$k_n = \sup \left\{ \varphi_n(\lambda_1, \dots, \lambda_n); \ \lambda_i \in \overline{D}, \ \lambda_i \neq \lambda_j \text{ for } i \neq j \right\}.$$

Proof: Let $P_T(z) = \sum_{k=0}^n a_k z^k$ be the characteristic polynomial of T, then $a_0 = \det(T)$ and, by the Cayley-Hamilton theorem, we get $0 = \sum_{k=0}^n a_k T^k$ and so

$$\operatorname{adj}\left(T\right) = -\sum_{k=1}^{n} a_{k} T^{k-1}.$$

Let g be a polynomial such that g(0) = 1 and let the polynomial $f(z) = \sum_{k\geq 0} b_k z^k$ be defined as $f(z) = g(z)P_T(z)$. Then we have $b_0 = a_0 = \prod_{i=1}^n \lambda_i$ and $\sum_{k\geq 0} b_k T^k = 0$. By multiplying the last equality by T^{-1} , we obtain

$$\operatorname{adj}(T) = b_0 T^{-1} = -\sum_{k \ge 1} b_k T^k.$$

Then for all $\|.\| \in \mathcal{N}(T)$, the following inequality holds:

$$\|\operatorname{adj}(T)\| \le \sum_{k\ge 1} |b_k| \|T\|^k \le \sum_{k\ge 1} |b_k| = \|f\|_{A_+} - |f(0)|.$$

Taking infimum over all polynomials g such that g(0) = 1, we obtain

$$\|\operatorname{adj}(T)\| \leq \varphi_n(\lambda_1, ..., \lambda_n)$$

for any $\|.\| \in \mathcal{N}(T)$. It follows that

$$\kappa(T) \leq \varphi_n(\lambda_1, ..., \lambda_n)$$

Let now V be the subspace of $\mathcal{L}(\mathbf{C}^n)$ generated by $\{T^k; k \geq 0\}$; since T has a simple spectrum, V is n-dimensional with basis $\{I, T, \ldots, T^{n-1}\}$, where $I = T^0$ denotes the identical operator. Let \widetilde{T} be the linear operator on V which transforms T^k into T^{k+1} for all $0 \leq k \leq n-1$. Then $\widetilde{T}(T^k) = T^{k+1}$, for all $k \geq 0$. Observe that for any polynomial R and any $k \geq 0$, we have $R(\widetilde{T})(T^k) = T^k \circ R(T)$; since T has a simple spectrum it follows that T and \widetilde{T} have the same characteristic polynomial. As a consequence, the operators T and \widetilde{T} are similar, in the sense that there exists an invertible operator S from \mathbf{C}^n onto V such that $\widetilde{T} = STS^{-1}$. In particular we have $\kappa(T) = \kappa(\widetilde{T})$. Indeed if $\| \cdot \| \in \mathcal{N}(T)$, define for $y \in V$, $\| \|y \| \| = \|S^{-1}y\|$; then $\| \operatorname{adj}(\widetilde{T}) \| = \| \| \operatorname{adj}(T) \| \|$.

Let B be the closed convex disked hull of $\{T^k; k \ge 0\}$ in $\mathcal{L}(\mathbb{C}^n)$,

$$B = \overline{\operatorname{conv}}\{b_k T^k; k \ge 0, b_k \in \overline{D}\}.$$

Since T is power bounded, B is bounded. We equip V with the norm associated to B; in other words, for $x \in V$, we have

$$||x||_V = \inf\{\lambda > 0 \ ; \ x \in \lambda B\}.$$

By the definition of \widetilde{T} , we have $\left\|\widetilde{T}\right\|_{V\to V} \leq 1$ and since $I \in B$, $\|I\|_{V} \leq 1$. Recall that $\operatorname{adj}(T) = -\sum_{k=1}^{n} a_{k} T^{k-1}$ and so $\operatorname{adj}(\widetilde{T}) = -\sum_{k=1}^{n} a_{k} \widetilde{T}^{k-1}$; we get then

$$\kappa(T) = \kappa(\widetilde{T}) \ge \left\| \operatorname{adj} (\widetilde{T}) \right\|_{V \to V} \ge \left\| \operatorname{adj} (\widetilde{T}) (I) \right\|_{V} = \left\| \sum_{k=1}^{n} a_{k} T^{k-1} \right\|_{V}$$

And by the definition of $\|.\|_{V}$, we have

$$\left\|\sum_{k=1}^{n} a_k T^{k-1}\right\|_{V} = \inf\left\{\sum_{k\geq 0} |b_k|; \sum_{k\geq 0} b_k T^k = \sum_{k=1}^{n} a_k T^{k-1}\right\}$$

where the infimum is taken over all sequences (b_k) with a finite number of nonzero entries. For such a sequence (b_k) , let $b(z) = \sum_{k\geq 0} b_k z^k - \sum_{k=1}^n a_k z^{k-1}$. The equality $\sum_{k\geq 0} b_k T^k = \sum_{k=1}^n a_k T^{k-1}$ means that b(T) = 0, or in other words, that the minimal polynomial of T divides b. In our case, the minimal and characteristic polynomials of T coincide, and thus there exists a polynomial h such that $b(z) = h(z)P_T(z)$, or in terms of (b_k) that $a_0 + z(\sum_{k\geq 0} b_k z^k) = (zh(z) + 1)P_T(z)$. It follows that

$$\left\|\operatorname{adj}(\widetilde{T})(I)\right\|_{V} = \left\|\sum_{k=1}^{n} a_{k} T^{k-1}\right\|_{V} = \varphi_{n}(\lambda_{1}, ..., \lambda_{n}).$$

We get thus $\kappa(T) \geq \varphi_n(\lambda_1, ..., \lambda_n)$.

The conclusion follows using equality (2).

Remarks: (1) The last theorem is also true, with the same proof, for any p.b. operator T such that $P_T = Q_T$. When $P_T \neq Q_T$, the following correction is needed: let $Q_T(z) = \prod_{i=1}^m (z - \mu_i)$ and let the polynomial R be defined by $P_T(z) = R(z)Q_T(z)$, then

$$\kappa(T) = |R(0)| \varphi_m(\mu_1, \dots, \mu_m).$$

(2) With the same proof, the first part of Theorem 1 holds in the real case. To get an estimate of k_n , we have to describe the class of characteristic polynomials of real matrices. Namely, let $\bar{D}_{\mathbb{R}}^n$ be the set of all $(\lambda_1, ..., \lambda_n) \in \bar{D}^n$ such that the polynomial $\prod_{i=1}^n (\lambda_i - z)$ has real coefficients (in other words the sets $\{\lambda_1, ..., \lambda_n\}$ and $\{\bar{\lambda}_1, ..., \bar{\lambda}_n\}$ coincide). For any $(\lambda_1, ..., \lambda_n) \in \bar{D}_{\mathbb{R}}^n$, there exists a p.b. operator $T \in \mathcal{L}(\mathbb{R}^n)$ for which $\lambda_1, ..., \lambda_n$ is the spectrum of T. Of course the converse is also true. Consequently, we get the following real analogue of (1):

$$\begin{split} \sup\{k(X); \ X \ n\text{-dimensional real normed space}\} \\ = \sup\{\varphi_n(\lambda_1, ..., \lambda_n); \ \lambda_i \neq \lambda_j \ \text{for} \ i \neq j, (\lambda_1, ..., \lambda_n) \in \bar{D}^n_{\mathbb{R}}\} \end{split}$$

(3) Let $T \in \mathcal{L}(E^n)$ be a p.b. operator such that $P_T = Q_T$ and let $x \in E^n$ be such that $\{T^p x; p \ge 0\}$ spans E^n . Using the proof of Theorem 1, one can show that $\kappa(T) = \inf\{\lambda > 0; |\det(T)| x \in \lambda \overline{\operatorname{conv}}\{a_p T^p x; p \ge 1, a_p \in \overline{D}\}\}$. In particular, if $A \in M_n$ is a Jordan matrix of T, if $\mathbf{1} = (1, ..., 1) \in \mathbb{C}^n$ and $\Delta_p = A^p \mathbf{1}, p \ge 1$, then

$$\kappa(T) = \inf\{\lambda > 0; |\det(A)| 1 \in \lambda, \overline{\operatorname{conv}}\{a_p \Delta_p; p \ge 1, a_p \in \overline{D}\}\}.$$

Let now S be the left shift operator on the space ℓ_{∞} consisting of bounded sequences:

$$S(x_0, x_1, ..., x_p, ...) = (x_1, ..., x_p, ...).$$

Let us denote by $\mathbf{I}_n(S)$ the set of all S-invariant *n*-dimensional subspaces of ℓ_{∞} and, for $E \in \mathbf{I}_n(S)$, let $S|_E \in \mathcal{L}(E)$ be the restriction of S to E. It is easy to see that any p.b. matrix $T \in M_n$, with $P_T = Q_T$, can be realized as a restriction of the left shift S, in the sense that there exists an *n*-dimensional S-invariant subspace E of ℓ_{∞} such that $S|_E$ is similar to T. For estimating $\kappa(T)$, it is enough to consider only $S|_E$. Namely we have: LEMMA 2: For any power bounded matrix $T \in M_n$ which satisfies $P_T = Q_T$, there exists an invariant subspace $E \in \mathbf{I}_n(S)$ such that

$$\kappa(T) = \|\operatorname{adj}\left(S|_E\right)\|.$$

Proof: For simplicity, we shall only consider the case when T has a simple spectrum and is invertible. Let $P_T(z) = \prod_{i=1}^n (\lambda_i - z)$ and, if $1 \le i \le n$, define $f_i \in \ell_{\infty}$ by $f_i(p) = \lambda_i^p$, for every integer $p \ge 0$. Then $\{f_1, ..., f_n\}$ is a basis of the unique space $E \in I_n(S)$ such that T is similar to $S|_E$. Let $x \in E$ be written $x = \sum_{i=1}^n x_i f_i$; then

$$\|x\| = \sup_{p \ge 0} \left| \sum_{i=1}^n \lambda_i^p x_i \right| \quad \text{and} \quad \|\operatorname{adj} (S|_E) x\| = \left| \prod_{i=1}^n \lambda_i \right| \sup_{p \ge 0} \left| \sum_{i=1}^n \lambda_i^{p-1} x_i \right|.$$

Hence

$$\|\operatorname{adj} (S|_E)\| = \inf \left\{ \rho > 0 \left| \left(\prod_{i=1}^n \lambda_i\right) \sum_{i=1}^n x_i \right| \le \rho \cdot \sup_{p \ge 1} \left| \sum_{i=1}^n \lambda_i^p x_i \right| \text{ for all } (x_i)_{i=1}^n \in \mathbf{C}^n \right\}.$$

The lemma follows by duality, in view of Remark 3 after Theorem 1.

The next proposition follows from (2) and the preceding lemma:

PROPOSITION 3: With the previous notation, $k_n = \sup\{\|\operatorname{adj}(S|_E)\|; E \in \mathbf{I}_n(S)\}$.

2. Estimates for k_n

LEMMA 4: For any integer $n \ge 1$ there exist n points $\zeta_1, \ldots, \zeta_n \in S^1$ such that

$$|\sum_{i=1}^{n} \zeta_i^k| \le 4\sqrt{n\log(1+k)}$$
 for every integer $k \ge 1$.

Proof: Let Z_1, \ldots, Z_n be n independent complex valued random variables uniformly distributed on the unit circle S^1 . Then by the Bernstein inequality, one has

$$\mathbf{P}(\{|\sum_{i=1}^{n} Z_i| > t\}) < 4e^{-t^2/4n}.$$

But for any integer $k \ge 1$, the sequence of random variables Z_1^k, \ldots, Z_n^k has the same distribution as the sequence Z_1, \ldots, Z_n and consequently

$$\mathbf{P}\Big(\Big\{\Big|\sum_{i=1}^{n} Z_{i}^{k}\Big| \le 4\sqrt{n\log(1+k)} \text{ for every } k \ge 1\Big\}\Big) \ge 1 - 4\sum_{k\ge 1} e^{-4\log(1+k)} \ge \frac{1}{2}.$$

The lemma follows.

THEOREM 5 (J. Bourgain): For any $n \ge 1$, there exist n points $\lambda_1, \ldots, \lambda_n \in \overline{D}$ such that

$$arphi_n(\lambda_1,\ldots,\lambda_n) \geq c \sqrt{rac{n}{\log(1+n)}} \; ,$$

where c > 0 is a numerical constant (independent of n).

Proof: Let $\lambda_i = 2^{-1/n} \zeta_i$ for every i = 1, ..., n where $(\zeta_i)_{1 \le i \le n}$ is given from Lemma 4. Then for any function $f(z) = \sum_{k \ge 0} a_k z^k \in A_+$ which satisfies

$$a_0 = f(0) = \prod_{i=1}^n \lambda_i$$
 and $f(\lambda_i) = 0$, for $i = 1, ..., n$,

one has

$$0 = \left|\sum_{i=1}^{n} f(\lambda_{i})\right| = \left|\sum_{k\geq 0} a_{k} \left(\sum_{i=1}^{n} \lambda_{i}^{k}\right)\right| \ge n|a_{0}| - \sum_{k\geq 1} |a_{k}| 2^{-k/n} \left|\sum_{i=1}^{n} \zeta_{i}^{k}\right|$$
$$\ge n 2^{-1} - 4 \sum_{k\geq 1} 2^{-k/n} \sqrt{n \log(1+k)} |a_{k}|.$$

It follows that

$$\|f\|_{A_{+}} - |f(0)| = \sum_{k \ge 1} |a_{k}| \ge \frac{\sqrt{n}}{8} \left(\sup_{k \ge 1} 2^{-\frac{k}{n}} \sqrt{\log(1+k)} \right)^{-1} \ge c \sqrt{\frac{n}{\log(n+1)}} ,$$

for some c > 0.

۲

In view of Theorem 1, Theorem 5 gives the following estimate for k_n : THEOREM 6: There exists a constant c > 0 such that for every $n \ge 1$,

$$k_n \ge c \sqrt{\frac{n}{\log(1+n)}}$$

Remark: It is clear that

$$\varphi_{2n}(z_1,\ldots,z_n,\overline{z}_1,\ldots,\overline{z}_n) \geq \left|\prod_{i=1}^n z_i\right| \varphi_n(z_1,\ldots,z_n).$$

It follows from this observation and from Remark 2 after Theorem 1 that Theorem 6 is also true in the real case. Indeed, with $\lambda_1, \ldots, \lambda_n$ of Proposition 5, we have

$$\varphi_{2n}(\lambda_1,\ldots,\lambda_n,\bar{\lambda}_1,\ldots,\bar{\lambda}_n)\geq 2^{-1}\varphi_n(\lambda_1,\ldots,\lambda_n),$$

and therefore, in the real case,

$$k_n \geq \frac{c}{2} \sqrt{\frac{n}{\log(1+n)}} \; .$$

We investigate now an upper bound for k_n . Let ℓ_2^n be \mathbb{C}^n or \mathbb{R}^n equipped with the Euclidean norm. Given $A \in \mathcal{L}(\ell_2^n)$, let $s_1 \geq \cdots \geq s_n \geq 0$ be its singular numbers (that is s_1^2, \ldots, s_n^2 are the eigenvalues of the operator A^*A). It is well known that $||A|| = s_1$ and that $||A^{-1}|| = s_n^{-1}$ if A is invertible; it follows that $|| \operatorname{adj} A || = \prod_{k=1}^{n-1} s_k$, and this gives $k(\ell_2^n) = 1$.

Let us now recall that the Hilbert-Schmidt norm of $A \in \mathcal{L}(\ell_2^n)$ is defined by

$$||A||_{HS} = \sqrt{\operatorname{trace}(A^*A)} = \left(\sum_{k=1}^n s_k^2\right)^{1/2}.$$

It follows that for any $A \in \mathcal{L}(\ell_2^n)$, we have

$$\| \operatorname{adj} A \| \le \left((n-1)^{-1/2} \|A\|_{HS} \right)^{n-1}$$

If $T \in \mathcal{L}(X)$ and $u \in \mathcal{L}(X, Y)$ is an isomorphism, then for the operator $S = uTu^{-1} \in \mathcal{L}(Y)$ we have adj $S = u \circ (\operatorname{adj} T) \circ u^{-1}$. So, for any *n*-dimensional normed space X and for any operator $T \in \mathcal{L}(X)$, the following inequality holds:

$$\| \operatorname{adj} T \| \le (n-1)^{-\frac{n-1}{2}} \inf\{ \|u^{-1}\| \| u T u^{-1} \|_{HS}^{n-1} \| u \|; u \text{ isomorphism: } X \to \ell_2^n \}.$$

The following proposition is a consequence of this inequality and of the next lemma:

PROPOSITION 7 ([S] in the real case): For any $n \ge 1$ we have

$$k_n \leq \sqrt{en}$$
.

LEMMA: Let X be an n-dimensional normed space. Then there exists an operator $u \in \mathcal{L}(X, \ell_2^n)$ such that

- (1) $||u^{-1}|| \leq 1$,
- (2) $||u|| \leq \sqrt{n}$ and, moreover, for any operator $R \in \mathcal{L}(\ell_2^n, X)$, we have $||uR||_{HS}$ $\leq \sqrt{n}||R||$.

We omit the proof of this lemma (see e.g. [P] or [T]). In fact the image by u^{-1} of the unit ball of ℓ_2^n is the ellipsoid of maximal volume contained in the unit ball

of the normed space X, and the lemma is equivalent to the result of F. John [J] about contact points of the unit ball with this ellipsoid.

In other terms, Proposition 7 says that for any power bounded matrix $T \in M_n$, one has $\kappa(T) \leq \sqrt{en}$ and Theorem 6 shows that this bound is essentially sharp. Of course, for some special classes of matrices, this estimate may be improved.

LEMMA 8: Let T be a power bounded operator with $|\det(T)| = 1$. Then $\kappa(T) = 1$.

Proof: Consider first the real case. Let $||.|| \in \mathcal{N}(T)$ and denote

$$B = \{x; ||x|| \le 1\}.$$

Then $T(B) \subset B$ and, since $|\det(T)| = 1$, we have $\operatorname{vol}(T(B)) = \operatorname{vol}(B)$. So T(B) = B and $||\operatorname{adj} T|| = ||T^{-1}|| = 1$, which gives the desired equality. In the complex case, it is enough to consider \mathbb{C}^n as \mathbb{R}^{2n} , and to check that the real operator T' corresponding to T in such a representation still satisfies $|\det(T')| = 1$.

As consequences of Theorem 1 and Proposition 7, we obtain the following corollaries.

COROLLARY 9: Let z_1, \ldots, z_n be *n* points in the unit disk \overline{D} , then $\varphi_n(z_1, \ldots, z_n) \leq \sqrt{en}$.

In the same way, using Lemma 8, we get the following result which can also be proved directly.

COROLLARY 10: For any $\lambda_1, \ldots, \lambda_n \in S^1$ and for every $\varepsilon > 0$ there exists a polynomial f such that f(0) = 1, $f(\lambda_i) = 0$ for $i = 1, \ldots, n$ and $||f||_{A^+} \leq 2 + \varepsilon$.

Remarks: (1) Some information about $\kappa(T)$ remains even if, instead of requiring $|\det(T)| = 1$, we know that $|\det(T)|$ is very close to 1. For example, in the real case we have

$$\kappa(T) \le \frac{|\det(T)|}{1 - (1 - |\det(T)|)^{1/n}}$$

Indeed let B_1 and B_2 be two centrally symmetric convex bodies of \mathbb{R}^n such that $B_2 \subset B_1$, and define $\rho_0 = \inf\{\rho > 0 ; B_1 \subset \lambda B_2\}$. We get from the Brunn-Minkowski theorem that

$$\frac{\text{vol}(B_2)}{\text{vol}(B_1)} \le \frac{\int_0^{1/\rho_0} f(t)dt}{\int_0^1 f(t)dt}$$

for some continuous positive function f such that $f^{1/(n-1)}$ is concave on [0, 1]. One can deduce from this property of f that

$$\operatorname{vol}(B_2) \leq \left(1 - \left(1 - \frac{1}{\rho_0}\right)^n\right) \operatorname{vol}(B_1).$$

Suppose now that X is an *n*-dimensional real normed space, and let $T \in \mathcal{L}(X)$ be invertible and norm 1; we have then $||T^{-1}|| \leq (1 - (1 - |\det(T)|)^{1/n})^{-1}$. The desired inequality for $\kappa(T)$ follows.

(2) If T is an invertible p.b. matrix in M_n , define a matrix $T' \in M_{n+1}$ by

 $T'(x_1,...,x_n,x_{n+1}) = T(x_1,...,x_n)$ for $(x_1,...,x_{n+1}) \in \mathbb{C}^{n+1}$.

Then it is easy to see that $\kappa(T') = \kappa(T) + |\det(T)|$. It follows that the sequence $(k_n)_{n\geq 0}$ is increasing ([S]).

Problem: Give a direct proof of Corollary 9, that is, given $z_1, ..., z_n \in \overline{D}$, find a polynomial f such that $|f(0)| = \prod_{k=1}^n |z_k|$, $f(z_k) = 0$ for every k = 1, ..., n, and $||f||_{A^+} \leq 1 + \sqrt{en}$. The Blaschke products are natural candidates. But recently, answering a question of the authors, I. Vidensky proved that there exist Blaschke products of n terms, with distinct zeros in D and A_+ -norm of the order of n.

Example 11: We have $k_3 \ge 9/4$. Let $(a, b, 1) \in \mathbb{R}^3$ be such that 0 < a < b < 1, and let $\Delta_n = (a^n, b^n, 1)$, for $n \ge 0$; let also C be the closed disked convex hull in \mathbb{R}^3 of $\{\Delta_n, n \ge 0\}$; we consider the operator $T (= T_{a,b})$: $\mathbb{R}^3 \to \mathbb{R}^3$ with diagonal matrix:

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

for the norm on \mathbb{R}^3 given by the gauge of C. It is clear that ||T|| = 1, and we have

$$\|\operatorname{adj}(T)\| = \inf\{\rho > 0 ; (b, a, ab) \in \rho C\}.$$

Now, if a > 0, b > 0, $a \neq b$, a+b < 1, it is not difficult to prove, using Remark 3 after Theorem 1, that ||adj(T)|| = 2+ab, which tends to 9/4 when a and $b \to 1/2$. It follows that $k_3 \ge 9/4$ and, by the last remark, since det $(T) \to 1/4$, that $k_4 \ge 5/2$.

This example shows also that the mapping $T \longrightarrow \kappa(T)$ is not continuous on the set of p.b. operators. In fact when $a, b \rightarrow 1/2$, $T_{a,b}$ tends to an operator U with diagonal matrix

$$\begin{pmatrix} \frac{1}{2} & 0 & 0\\ 0 & \frac{1}{2} & 0\\ 0 & 0 & 1 \end{pmatrix}$$

By Remark 1 after Theorem 1, we have $\kappa(U) = \frac{1}{2}\varphi(1/2, 1) \leq \frac{1}{2}k_2 \leq 1$ (since $k_2 = 2$). But $\kappa(T_{a,b}) \to 9/4$.

3. Entropy estimates for A_+ and lower bound for k_n

We give here our original proof of the estimate

$$k_n \ge rac{c}{\log(\log n)} \sqrt{rac{n}{\log n}}.$$

This proof uses some tools which may be useful for other purposes: Proposition 14 gives an estimate for the entropy of the unit ball of A_+ . We need first some notation.

For every number $0 < r \leq 1$, let μ_r be the uniform probability measure on $rS^1 = \{z \in \mathbf{C}; |z| = r\}$ and extend μ_r on the unit disk by $\mu_r(E) = \mu_r(E \cap rS^1)$ for any measurable subset E of \overline{D} . Let $W = \{f; f \in A_+, ||f||_{A_+} \leq 1\}$. For every 0 < r < 1, we define a new metric on A_+ by $||f||_r = \sup_{\xi} |f(re^{i\xi})|$, for $f \in A_+$. For every $t \geq 1$ let $V(t) = \{f \in tW; f(0) = 1\}$. If (E, d) is a metric space and $\varepsilon > 0$, an ε -net of (E, d) is a subset Λ of E such that

$$E = \bigcup_{y \in \Lambda} \{x; \ d(x,y) < \varepsilon\}.$$

Denote by $N(E, d, \varepsilon)$ the smallest cardinality of an ε -net of (E, d). Observe that if (E, d) is a metric subspace of another metric space (E', d), and if we have a covering of E by ε -balls of E', that is, $E \subset \bigcup_{y \in \Lambda} \{x; d(x, y) < \varepsilon\}$ with $\Lambda \subset E'$, then

$$(3) N(E,d,2\varepsilon) \le |\Lambda|$$

where $|\Lambda|$ denotes the cardinality of Λ . Finally, for a subset V of A_+ , let $N(V, r, \epsilon) = N(V, ||.||_r, \epsilon)$.

LEMMA 12: Let $\varepsilon, t > 0, 0 < r < 1$ and $n \ge 1$ be an integer, such that

(4)
$$\left(\sup_{f\in V(t)}\mu_r(\{|f|\leq 2\varepsilon\})\right)^n N(V(t),r,\varepsilon)<1.$$

238

Then there exist n distinct complex numbers z_1, z_2, \ldots, z_n of moduli r such that

$$\max_{1 \le i \le n} |f(z_i)| > \varepsilon, \quad \text{for every } f \in V(t).$$

Proof: For a function $g \in A_+$ and a number $\delta > 0$, let us define

$$\Omega_{g,\delta} = \left\{ (\zeta_1, ..., \zeta_n) \in (rS^1)^n; \max_{1 \le i \le n} |g(\zeta_i)| \le \delta \right\}.$$

Choose a minimal ε -net Λ_{ε} for the metric space $(V(t), \|.\|_r)$ and set

$$\Omega = \bigcup_{g \in \Lambda_{\varepsilon}} \Omega_{g, 2\varepsilon}.$$

Then

$$\mu_r^n(\Omega) \le |\Lambda_{\varepsilon}| \max_{g \in \Lambda_{\varepsilon}} \mu_r^n(\Omega_{g,2\varepsilon}) \le |\Lambda_{\varepsilon}| \Big(\sup_{f \in V(t)} \mu_r\{|f| \le 2\varepsilon\} \Big)^n < 1$$

Thus there exist distinct complex numbers $z_1, ..., z_n \in rS^1$ such that for every $g \in \Lambda_{\varepsilon}$, we have $\max_{1 \leq i \leq n} |g(z_i)| > 2\varepsilon$. Let now $f \in V(t)$. Since Λ_{ε} is an ε -net of V(t), there exists $g \in \Lambda_{\varepsilon}$ such that $||f - g||_r < \varepsilon$. Therefore

$$\max_{1 \le i \le n} |f(z_i)| \ge \max_{1 \le i \le n} |g(z_i)| - ||f - g||_r > 2\varepsilon - \varepsilon = \varepsilon.$$

For satisfying condition (4), we study each of its factors.

LEMMA 13: Let $0 < u < 1 \le t$ and 0 < r < 1. For every function $f \in V(t)$, we have

$$\mu_r(\{|f| \le u\}) \le 1 - \frac{\log \frac{1}{u}}{\log \frac{t}{u}}.$$

Proof: Let $f \in V(t)$; since f is holomorphic in the disk of radius r, Jensen's inequality gives $0 = \log |f(0)| \leq \int \log |f| d\mu_r$. Now since $|f| \leq t$ on D, we get

$$0 \le \mu_r(\{|f| \le u\}) \log u + (1 - \mu_r(\{|f| \le u\})) \log t;$$

hence

$$\mu_r(\{|f| \le u\}) \le \frac{\log t}{\log t - \log u} \cdot \quad \blacksquare$$

In the next proposition we estimate the entropy of A_+ for the uniform metric on a circle of radius r. We ignore its exact order of magnitude. In what follows, we shall only use the upper bound, but the lower bound that we get gives a good idea of the order of this entropy. **PROPOSITION 14:** Let W be the unit ball of A_+ and $\frac{1}{2} < r < 1$. Then, with the preceding notation, for every $0 < \theta < 1$, we have

$$\log N(W, r, \theta) \leq \frac{\alpha}{\theta^2} \log^2 \left(\beta \frac{\theta^2}{\log_2 \frac{1}{r}} + 2\right)$$

where $\alpha, \beta > 0$ are universal constants. Moreover for $0 < \theta < 1/4\sqrt{2}$, we have

$$\log N(W, r, \theta) \ge \frac{1}{32\theta^2} \log \left(\frac{8\theta^2}{\log_2 \frac{1}{r}}\right).$$

Proof: (1) The upper estimate. Let r and θ be given as above. Define $m = [3/\log_2 \frac{1}{r}]$, where [x] denotes the integral part of $x \in \mathbb{R}$, so that $\frac{1}{8} \leq r^m \leq \frac{1}{4}$. Then set

We observe also that

$$\frac{\theta}{4} \cdot 2^p \le \theta_p \le \frac{\theta}{4} \cdot 4^p.$$

Now let W_m be the subset of W consisting of all the complex polynomials of degree less than m, and for every integer $p \ge 0$, let $\Lambda_{p,m}$ be a θ_p -net of W_m for the uniform metric on the circle of radius r.

We shall show that if Λ is the set of all polynomials of the form $\sum_{p\geq 0} z^{mp}g_p$, where $g_p \in \Lambda_{p,m}$, then for any $f \in W$ there exists $g \in \Lambda$ such that $||f-g||_r < \theta/2$. Indeed if $f(z) = \sum_{k\geq 0} \alpha_k z^k \in W$ and $f_p(z) = \sum_{k=0}^{k=m-1} \alpha_{mp+k} z^k$, write

$$f(z) = \sum_{p \ge 0} z^{mp} \left(\sum_{k=mp}^{mp+m-1} \alpha_k z^{k-mp} \right) = \sum_{p \ge 0} z^{mp} f_p(z).$$

Since $f_p \in W_m$, there exists $g_p \in \Lambda_{p,m}$ such that $||f_p - g_p||_r < \theta_p$. Set $g(z) = \sum_{p \ge 0} z^{mp} g_p(z)$. Using (5) we get $||f - g||_r < \theta/2$. By (3) we have then

(6)
$$\log N(W, r, \theta) \leq \sum_{p \geq 0} \log N(W_m, r, \theta_p).$$

Denote $\rho_q = e^{2i\pi q/2m}$, $q = 1, \ldots, 2m$. From Bernstein's theorem, we have

(7)
$$||P||_r \le \sqrt{2} \max_{1 \le q \le 2m} |P(r\rho_q)|, \text{ for every } P \in W_m.$$

Now, for k = 0, ..., m - 1, let $A_k \in \mathbb{C}^{2m}$, $A_k = (\rho_1^k, ..., \rho_{2m}^k)$, and define $K \subset \mathbb{C}^{2m}$ to be the disked convex hull of $\{A_k, k = 0, ..., m - 1\}$. It follows from (7) that

(8)
$$\log N(W_m, r, \theta_p) \le \log N\left(K, \| \|_{\infty}, \frac{\theta_p}{\sqrt{2}}\right)$$

where $\| \|_{\infty}$ is defined on \mathbb{C}^{2m} by $\|(z_1, \ldots, z_{2m})\|_{\infty} = \max_{1 \leq q \leq 2m} |z_q|$.

Notice that K is the disked convex hull of m points of \mathbb{C}^{2m} with $\| \|_{\infty}$ -norm one. By a result of B. Carl ([C]) we have then

(9)
$$\log N(K, \| \|_{\infty}, \eta) \leq \frac{\alpha}{\eta^2} \log^2 \left(\beta m \eta^2 + 2\right), \quad 0 < \eta < 1$$

where α and β are universal constants.

It follows from (8) and (9) that

$$\log N(W_m, r, \theta_p) \leq \frac{2\alpha}{\theta_p^2} \log^2 \left(\frac{\beta m}{2} \theta_p^2 + 2\right) \leq \frac{64\alpha}{4^p \theta^2} \log^2 \left(\frac{3\beta}{32} \cdot \frac{8^p \theta^2}{\log_2 \frac{1}{r}} + 2\right).$$

Using (6), it is now easy to get the desired upper estimate.

(2) The lower estimate. Set $m = 1 + [1/\log_2 \frac{1}{r}]$, so that $r^m < \frac{1}{2} \le r^{m-1}$. Define

$$S = \left\{ a = (a_1, \dots, a_m) \in \{0, \pm 2\theta\}^m; |\{i; a_i \neq 0\}| = \left[\frac{r^{2(m-1)}}{4\theta^2}\right] \right\}$$

and denote $\sigma_q = e^{2i\pi q/m}$, $1 \le q \le m$. For every $a = (a_1, \ldots, a_m) \in \mathbb{C}^m$, let L(a) be the polynomial of degree m-1 such that $L(a)(r\sigma_q) = a_q$, for $1 \le q \le m$. Set $\Lambda = \{L(a); a \in S\}$, and observe that

(10)
$$||f - g||_r \ge 2\theta$$
 for every $f \ne g$ in Λ .

Fix $a = (a_1, \ldots, a_m) \in S$ and let $L(a)(z) = \sum_{k=0}^{m-1} b_k z^k$. It is well known that

$$\frac{1}{\sqrt{m}} \left(\sum_{q=1}^{m} |a_q|^2 \right)^{\frac{1}{2}} = \left(\frac{1}{2\pi} \int_0^{2\pi} |L(a)(re^{i\xi})|^2 d\xi \right)^{\frac{1}{2}} = \left(\sum_{k=0}^{m-1} |b_k r^k|^2 \right)^{\frac{1}{2}}.$$

Since $a \in S$, we have

$$\sum_{q=1}^{m} |a_q|^2 = 4\theta^2 \left[\frac{r^{2(m-1)}}{4\theta^2}\right] \le r^{2(m-1)}.$$

We conclude that

$$||L(a)||_{A_{+}} = \sum_{k=0}^{m-1} |b_{k}| \le \frac{1}{\sqrt{m}} \left(\sum_{q=1}^{m} |a_{q}|^{2}\right)^{\frac{1}{2}} \left(\sum_{k=0}^{m-1} r^{-2k}\right)^{\frac{1}{2}} \le 1$$

which means that $\Lambda \subset W$. From (10), we get $N(W, r, \theta) \ge |\Lambda| = |S|$. If

$$p = \Big[\frac{r^{2(m-1)}}{4\theta^2}\Big],$$

then

$$|S| = 2^p \binom{m}{p} \ge \left(\frac{2m}{p}\right)^p;$$

since $0 < \theta < 1/4\sqrt{2}$, we get

$$\log N(W, r, \theta) \ge \left(\frac{1}{16\theta^2} - 1\right) \log \left(\frac{8\theta^2}{\log_2 \frac{1}{r}}\right) \ge \frac{1}{32\theta^2} \log \left(\frac{8\theta^2}{\log_2 \frac{1}{r}}\right).$$

Proof of the estimate $k_n \geq \frac{c}{\log(\log n)} \sqrt{\frac{n}{\log n}}$: We observe first that if ε, t, r, n satisfy the conditions of Lemma 12, then $k_n \geq (t-1)r^n$. Indeed, let $z_1, ..., z_n$ be obtained by this lemma, and let h be any function in A_+ such that h(0) = 1and $h(z_1) = \cdots = h(z_n) = 0$. Then $||h||_{A_+} > t$, and taking into account the normalization $\prod_{i=1}^{n} |z_i| = r^n$, we get $\varphi_n(z_1, ..., z_n) \geq (t-1)r^n$, and the inequality $k_n \geq (t-1)r^n$ follows from Theorem 1. To conclude the proof, we verify that (4) is satisfied for $\varepsilon = \frac{1}{4}, r = 2^{-\frac{1}{n}}, t = \gamma \sqrt{n}(\sqrt{\log n} \log(\log n))^{-1}$, where γ is some universal constant. Indeed, since $V(t) \subset tW$, we have

$$N(V(t), r, \varepsilon) \leq N\left(tW, r, \frac{\varepsilon}{2}\right) = N\left(W, r, \frac{\varepsilon}{2t}\right),$$

and then the result follows from Lemma 13 and Proposition 14.

References

- [C] B. Carl, Inequalities of Bernstein-Jackson type and the degree of compactness of operators in Banach spaces, Ann. Inst. Fourier 35 (1985), 79-118.
- [J] F. John, Extremum problems with inequalities as subsidiary conditions, in Studies and Essays Presented to R. Courant on his 60th Birthday, 1948, Interscience, New York, 1948, pp. 187-204.
- [P] G. Pisier, The Volume of Convex Bodies and Banach Space Geometry, Cambridge University Press, 1989.
- [S] J. J. Schäffer, Norms and determinants of linear mappings, Math. Z. 118 (1970), 331-339.
- [T] N. Tomczak-Jaegermann, Banach-Mazur Distances and Finite Dimensional Operator Ideals, Pitman, London, 1988.