METRIC ENTROPY AND APPROXIMATION[1]

BY G. G. LORENTZ

1. Introduction. The notion of metric entropy (called also ϵ-entropy) has been invented by Kolmogorov [16], [19] in order to classify compact metric sets according to their massivity. The basic definitions are as follows.

Let A be a subset of a metric space X, and let $\epsilon>0$ be given. A family $U_1, \cdots, U_n$ of subsets of X is an *ϵ-covering of A* if the diameter of each U_k does not exceed 2ϵ and if the sets U_k cover A. For a given $\epsilon>0$, the number n depends upon the covering family, but $N_\epsilon(A)=\min n$ is an invariant of the set A. The logarithm

$$H_\epsilon(A) = \log N_\epsilon(A) \tag{1}$$

is the *entropy* of A. (Sometimes this definition is modified by assuming that the sets U_k are balls of radius ϵ.)

Points $y_1, \cdots, y_m$ of A are called *ϵ-distinguishable* if the distance between each two of them exceeds ϵ. The number $M_\epsilon(A)=\max m$ is an invariant of the set A, and

$$C_\epsilon(A) = \log M_\epsilon(A) \tag{2}$$

is called the *capacity of A*. The main general fact about $C_\epsilon(A)$ and $H_\epsilon(A)$ is the simple set of inequalities

$$C_{2\epsilon}(A) \leqq H_\epsilon(A) \leqq C_\epsilon(A). \tag{3}$$

In general, $C_\epsilon(A)$ and $H_\epsilon(A)$ increase rapidly to $+\infty$ as $\epsilon\to 0$; their asymptotic behavior serves to describe the compact set A.

For the computation of the entropy of concrete sets of functions, Kolmogorov [16], [19], Vituškin [37] and others, have used different special devices. The results obtained were mainly valid for the uniform metric, and for sets A, whose approximation properties by polynomials, or by arbitrary linear combinations of fixed functions were well known. More precisely, the sets A under consideration were sets $A(\Delta, \Phi)$ described below, or at least approximable by such sets.

Let $\Phi=\{\phi_1, \cdots, \phi_n, \cdots\}$ be a *fundamental sequence* of points in

[1] Based on an address entitled *Applications of entropy to approximation* delivered by Professor Lorentz by invitation of the Committee to Select Hour Speakers for Summer and Annual Meetings, on Friday, September 3, 1965; received by the editors May 30, 1966. This work has been supported, in part, by the Contract no. AF 49 (638)-1401 of the Office of Scientific Research, U. S. Air Force.

a Banach space X, that is, a sequence of points whose linear combinations are dense in X.

Let $\Delta=\{\delta_0, \cdots, \delta_n, \cdots\}$ be a sequence of numbers for which $\delta_n>0$, $\delta_0 \geqq \delta_1 \geqq \cdots$, $\delta_n \to 0$. The nth degree of approximation of an element $f\in X$ by the system Φ is given by

$$E_n^{\Phi}(f) = \min_{\{a_k\}} \left\| f - \sum_{k=1}^{n} a_k\phi_k \right\|, \qquad n = 0, 1, \cdots. \tag{4}$$

The set $A(\Delta, \Phi)$ consists of all elements $f\in X$ which satisfy the inequalities $E_n^{\Phi}(f) \leqq \delta_n$, $n=0, 1, \cdots$. A set $A(\Delta, \Phi)$ is sufficiently rich, because by a theorem of S. Bernšteĭn (Davis [8, p. 332]), for each Δ there is an $f\in X$ for which $E_n^{\Phi}(f)=\delta_n$, $n=0, 1, \cdots$. We shall call the sets $A(\Delta, \Phi)$ the *full approximation sets.*

Is it possible to develop a method for the evaluation of the entropy of each set $A(\Delta, \Phi)$? We shall give a positive answer to this question. It is remarkable that neither the structure of the sequence Φ, nor the properties of the norm of the space X are important for the final results. In this way we obtain a uniform derivation of known estimates (Kolmogorov's for the classes $\Lambda_{r,\alpha}$, Vituškin's for analytic functions) as well as new ones (for example, classes $\Lambda_{r,\alpha}^p$ in the L^p-norm).

As forerunners of our results we mention the computation of entropies of ellipses by Kolmogorov and Tihomirov [19, p. 40] and Mitjagin [25], and papers of Brudnyĭ and Timan [5] and of Helemskiĭ and Henkin [14], which deal with arbitrary compacts in Hilbert and Banach spaces.

For a general theory of entropy, the reader can consult the article [19] and the books of Vituškin [37] and Lorentz [24, Chapter 10].

The plan of this paper is as follows. In §2, we study some geometric properties of compacts in finitely and infinitely dimensional Banach spaces. In §3, our main results (Theorems 2 and 3) are established; they concern the entropy of full approximation sets. Their usefulness for the computation of entropies of concrete sets is illustrated in §5. Another application, in §4.2, is to the "stability" of the approximation of full approximation sets. The estimate $E_n^{\Phi}(f) \leqq \delta_n$ will not be essentially improved for most $f\in A(\Delta, \Phi)$ if Φ is replaced by some other sequence Ψ, and even if one is allowed to select for each f the most favorable from a countable set of sequences $\Psi_1, \Psi_2, \cdots$.

Another application of ideas of §3 is to theorems of Vituškin's type in Banach spaces: in §4.1 for linear approximation, in §7 to (piecewise) polynomial approximation. A very simple proof of Vituškin's theorem for rational approximation in uniform norm is offered in §6. Finally, §5 contains a review of recent results on entropy.

2. Properties of ellipsoids and of full approximation sets. In this section we shall study the geometric and the measure-theoretic properties of certain sets. Let X be a Banach space, $\Phi=\{\phi_1, \phi_2, \cdots\}$ a fundamental sequence in X. Let X_n be the n-dimensional subspace of X spanned by $\phi_1, \cdots, \phi_n$. To certain subsets B of X we assign their "euclidean volume" $|B|$, by identifying it with the volume of the image of B in R^n under the map $a_1\phi_1+\cdots+a_n\phi_n\to(a_1, \cdots, a_n)$. For $k\leqq n$, the volume in the subspace X_k of X_n (and in subspaces of X_n parallel to X_k) will be denoted by $|B|_k$.

In X_n we consider the unit ball U, balls U_r with center origin and radius $r>0$, and "ellipsoids" $E=E(\delta_0, \cdots, \delta_{n-1})$, which consist of all points $a_1\phi_1+\cdots+a_n\phi_n$ for which

$$\frac{a_1}{\delta_0}\phi_1+\cdots+\frac{a_n}{\delta_{n-1}}\phi_n\in U.$$

If $|U|=\lambda_n$, then $|U_r|=\lambda_n r^n$, and

$$|E(\delta_0, \cdots, \delta_{n-1})| = \lambda_n\delta_0, \cdots, \delta_{n-1}. \tag{1}$$

A set $A=A(\delta_0, \cdots, \delta_{n-1})$ is the set of all points $f\in X_n$ for which $E_k^\Phi(f)\leqq\delta_k$, $k=0, 1, \cdots, n-1$.

LEMMA 1. *For the ball U_r of X_n, and $0<\epsilon\leqq r$,*

$$\frac{1}{2^n}\left(\frac{r}{\epsilon}\right)^n \leqq N_\epsilon(U_r) \leqq 3^n\left(\frac{r}{\epsilon}\right)^n. \tag{2}$$

PROOF. Let $y_1, \cdots, y_m$, $m=M_\epsilon(U_r)$ be a maximal set of ϵ-distinguishable points of U_r. The closed balls with centers y_i and radii ϵ cover U_r. Comparing the euclidean volumes, we see that $\lambda_n\epsilon^n M_\epsilon(U_r)\geqq\lambda_n r^n$.

On the other hand, balls with centers y_i and radii $\epsilon/2$ are disjoint, and contained in the ball $U_{r+\epsilon/2}\subset U_{3r/2}$. It follows $\lambda_n(\epsilon/2)^n M_\epsilon(U_r)\leqq\lambda_n(3r/2)^n$. Thus,

$$\left(\frac{r}{\epsilon}\right)^n \leqq M_\epsilon(U_r) \leqq 3^n\left(\frac{r}{\epsilon}\right)^n. \tag{3}$$

Relation (2) follows from this and the inequalities $M_{2\epsilon}(A)\leqq N_\epsilon(A)\leqq M_\epsilon(A)$.

LEMMA 2. *Let E be the ellipsoid $E(\delta_0, \cdots, \delta_{n-1})$ $\delta_0\geqq\cdots\geqq\delta_{n-1}>0$ in X_n, let $E_{a_k^0, \ldots, a_n^0}$ be the $(k-1)$-dimensional section of E, given by the relations $a_i=a_i^0$, $i=k, \cdots, n$. If $|\lambda|\leqq 1$ and $a_i'=\lambda a_0^0$, $i=k, \cdots, n$, then*

$$(4) \qquad |E_{a'_k,\cdots,a'_n}| \geqq |E_{\overset{0}{a}_k,\cdots,\overset{0}{a}_n}|.$$

PROOF. We put

$$F = E_{\overset{0}{a}_k,\cdots,\overset{0}{a}_n}, \qquad F_- = E_{-\overset{0}{a}_k,\cdots,-\overset{0}{a}_n},$$

$$F' = E_{a'_k,\cdots,a'_n}, \qquad F'' = \frac{1+\lambda}{2}F + \frac{1-\lambda}{2}F_-.$$

The set F_- is obtained from F by the mapping $x\to -x$ and X_n onto itself, hence $|F_-| = |F|$. If $x=(a_1, \cdots, a_n)\in F''$, then

$$a_i = \frac{1+\lambda}{2}\overset{0}{a}_i - \frac{1-\lambda}{2}\overset{0}{a}_i = \lambda\overset{0}{a}_i = a'_i, \quad i = k, \cdots, n.$$

Moreover, since E is convex, $F''\subset E$. It follows that $F''\subset E_{a'_k,\cdots,a'_n} = F'$.

By the theorem of Brun-Minkowski about mixed volumes [3, p. 88],

$$|F''| \geqq \frac{1+\lambda}{2}|F| + \frac{1-\lambda}{2}|F_-| = |F|.$$

Hence $|F'|\geqq |F|$.

THEOREM 1. *Let $E=E(\delta_0, \cdots, \delta_{n-1})$ and $A=A(\delta_0, \cdots, \delta_{n-1})$ be subsets of X_n. Then their Euclidean volumes satisfy the relation*

$$(5) \qquad |E(\delta_0, \cdots, \delta_{n-1})| \leqq |A(\delta_0, \cdots, \delta_{n-1})|.$$

PROOF. The set A consists of all points $x=a_1\phi_1+ \cdots +a_n\phi_n = (a_1, \cdots, a_n)\in X_n$ for which

$$(6) \qquad \begin{aligned} &\min_{x_1,\cdots,x_{n-1}} \left\| x_1\phi_1 + \cdots + x_{n-1}\phi_{n-1} + \frac{a_n}{\delta_{n-1}}\phi_n \right\| \leqq 1, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &\min_{x_1} \left\| x_1\phi_1 + \frac{a_2}{\delta_1}\phi_2 + \cdots + \frac{a_n}{\delta_1}\phi_n \right\| \leqq 1, \\ &\left\| \frac{a_1}{\delta_0}\phi_1 + \cdots + \frac{a_n}{\delta_0}\phi_n \right\| \leqq 1; \end{aligned}$$

while E is given by one single inequality

$$(7) \qquad \left\| \frac{a_1}{\delta_0}\phi_1 + \cdots + \frac{a_n}{\delta_{n-1}}\phi_n \right\| \leqq 1.$$

We introduce intermediate sets A_k, $k=1, \cdots, n-1$ in X_n. A point $(a_1, \cdots, a_n)$ belongs to A_k if and only if

$$\min_{x_1,\cdots,x_{n-1}} \left\| x_1\phi_1 + \cdots + x_{n-1}\phi_{n-1} + \frac{a_n}{\delta_{n-1}}\phi_n \right\| \leqq 1$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\min_{x_1,\cdots,x_k} \left\| x_1\phi_1 + \cdots + x_k\phi_k + \frac{a_{k+1}}{\delta_k}\phi_{k+1} + \cdots + \frac{a_n}{\delta_k}\phi_n \right\| \leqq 1 \tag{8}$$

$$\left\| \frac{a_1}{\delta_0}\phi_1 + \cdots + \frac{a_{k-1}}{\delta_{k-2}}\phi_{k-1} + \frac{a_k}{\delta_{k-1}}\phi_k + \cdots + \frac{a_n}{\delta_{k-1}}\phi_n \right\| \leqq 1.$$

Conditions (8) contain $n-k+1$ inequalities. All but the last one disappear if $k=n$. Therefore, $A_1=A$, $A_n=E$. Hence all that we have to show is that $|A_k| \leqq |A_{k-1}|$ for $k=2, \cdots, n$. Let G_k, $k=2, \cdots, n$ be the $n-k$-dimensional region of the change of $a_{k+1}, \cdots, a_n$, given by the first $n-k$ inequalities (8). Then

$$|A_k| = \int_{G_k} |E_{\delta_k a_{k+1}/\delta_{k-1},\cdots,\delta_{n-1}a_n/\delta_{k-1}}|_k da_{k+1} \cdots da_n$$

$$= \int_{G_k} da_{k+1} \cdots da_n \int |E_{a_k,\delta_k a_{k+1}/\delta_{k-1},\cdots,\delta_{n-1}a_n/\delta_{k-1}}|_{k-1} da_k,$$

where the last integral is extended over all a_k for which

$$\min_{x_1,\cdots,x_{k-1}} \left\| x_1\phi_1 + \cdots + x_{k-1}\phi_{k-1} + \frac{a_k}{\delta_{k-1}}\phi_k + \cdots + \frac{a_n}{\delta_{k-1}}\phi_n \right\| \leqq 1.$$

Applying Lemma 2, we obtain

$$|A_k| \leqq \int_{G_{k-1}} |E_{\delta_{k-1}a_k/\delta_k,\cdots,\delta_{n-1}a_n/\delta_k}|_{k-1} da_k \cdots da_n = |A_{k-1}|.$$

We now return to the *full approximation sets* $A=A(\Delta, \Phi)$ in separable Banach spaces X. It is easy to see that each set A is compact. Their *widths* ([22], [33], [24]) can also be determined:

$$d_n(A) = \delta_n, \qquad n = 0, 1, \cdots. \tag{9}$$

Indeed, from the definition of A and the widths, $d_n(A) \leqq \delta_n$. On the other hand, $A_n = A(\delta_0, \cdots, \delta_{n-1}) \subset X_n$ contains the ball $U_{\delta_{n-1}}$ of X_n. By Tihomirov's theorem [33, p. 84], [24, p. 137], in the space X, $d_{n-1}(U_{\delta_{n-1}}) = \delta_{n-1}$. Hence

$$d_{n-1}(A) \geqq d_{n-1}(A_n) \geqq \delta_{n-1}, \qquad n = 1, 2, \cdots.$$

LEMMA 3. *Let $A(\Delta, \Phi)$ be a full approximation set in the Banach space X, let $A' = A(\delta_0 + 2\delta_n, \cdots, \delta_{n-1} + 2\delta_n)$ be a subset of X_n. Let B be the full approximation set of the factor space $Y = X/X_n$, which corresponds to the sequences $\delta_n, \delta_{n+1}, \cdots$ and $\psi_n, \psi_{n+1}, \cdots$, where ψ_i, $i = n, n+1, \cdots$ is the factor class which contains ϕ_i. If $\epsilon = \epsilon_1 + \epsilon_2$, $\epsilon_1 > 0$, $\epsilon_2 > 0$, then*

$$H_\epsilon(A(\Delta, \Phi)) \leqq H_{\epsilon_1}(A') + H_{\epsilon_2}(B). \tag{10}$$

PROOF. Let B_0 be the set of all $\eta \in Y$ for which $\eta \cap A(\Delta, \Phi) \neq 0$. Then

$$B_0 \subset B. \tag{11}$$

Indeed, let $x \in \eta \cap A(\Delta) \neq 0$. Since $x \in A(\Delta)$, there exists for each $m \geqq n$ a linear combination $a_1\phi_1 + \cdots + a_m\phi_m$ for which $\|x - (a_1\phi_1 + \cdots + a_m\phi_m)\| \leqq \delta_m$. This implies

$$\|\eta - (a_n\psi_n + \cdots + a_m\psi_m)\|_Y \leqq \delta_m, \; m \geqq n.$$

Hence $\eta \in B$.

Assume now that A_0 is the subset of X that consists of all points x with the property $E_k(x) \leqq \delta_k'$, $k = 0, \cdots, n-1$. Then for each class $\eta \in Y$, the set $\eta \cap A_0$ has an ϵ-net containing at most $N_\epsilon(A(\delta_0' + \delta_n', \cdots, \delta_{n-1}' + \delta_n'))$ points.

For the proof, let $\eta \cap A_0 \neq 0$. If $x \in \eta \cap A_0$, we have $\rho(x, X_n) \leqq \delta_n'$. This implies that there exists a point $x_0 \in \eta$ with $\|x_0\| \leqq \delta_n'$. Then $\eta = x_0 + X_n$. The sets $\eta \cap A_0$ and $C = \eta \cap A_0 - x_0$ are isometric, and it is sufficient to construct a required ϵ-net for C in X_n. However, if $y \in C$, then $y = x - x_0$, $x \in A_0$, and $E_k(y) \leqq \delta_k' + \|x_0\| \leqq \delta_k' + \delta_n'$, $k < n$. Therefore, $C \subset A(\delta_0' + \delta_n', \cdots, \delta_{n-1}' + \delta_n')$, and the statement follows.

We can now prove (10). Let $\epsilon_2 < \delta_n$, let $\eta_1, \cdots, \eta_m$, $m = N_{\epsilon_2}(B_0)$ be an ϵ_2-net for B_0. For each i, let x_{ij}, $j = 1, \cdots, N_{\epsilon_1}(\eta_i \cap A_0)$ be an ϵ_1-net for the set $\eta_i \cap A_0$, where A_0, with $\delta_k' = \delta_k + \delta_n$, $k = 0, \cdots, n-1$, has been described above. Then the x_{ij} form an ϵ-net for $A(\Delta)$. In fact, let $y \in A(\Delta)$, and let η be the class to which y belongs. Then for some i, $\|\eta - \eta_i\| \leqq \epsilon_2$, and we can find a $y_i \in \eta_i$ for which $\|y - y_i\| \leqq \epsilon_2 (< \delta_n)$. Then $E_k(y_i) \leqq \delta_k + \delta_n = \delta_k'$, $k = 0, \cdots, n-1$; hence $y_i \in \eta_i \cap A_0$. Then for some j, $\|y - x_{ij}\| \leqq \epsilon$.

The number of points x_{ij} does not exceed

$$N_{\epsilon_1}(A(\delta_0' + \delta_n', \cdots, \delta_{n-1}' + \delta_n'))m,$$

and we obtain (10).

The case $\epsilon_2 \geqq \delta_n$ is simpler. Here B_0 is contained in the ball U_{ϵ_2} of the space Y, and the set $\eta_1, \cdots, \eta_m$ can be replaced by the zero element of Y.

REMARK. The spaces X and X_n have been *real* Banach spaces. Similar considerations hold for *complex* Banach spaces. In this case, X_n is the set of all linear combinations $\alpha_1\phi_1 + \cdots + \alpha_n\phi_n$ with complex α_k or of all real linear combinations of $2n$ elements ϕ_k, $i\phi_k$, $k=1, \cdots, n$. In this case, n should be replaced by $2n$ in the formula (2), while instead of (1) we have $|E(\delta_1, \cdots, \delta_n)| = \lambda_n \delta_1^2, \cdots, \delta_n^2$. No changes are necessary in Lemmas 2, 3 and Theorem 1.

3. **The main results.** Let $A = A(\Delta, \Phi)$ be a full approximation set in a Banach space X. Let $C > 1$ be a fixed constant. We define

$$N_0 = 0,\ N_i = \min\{k : \delta_k \leqq C^{-i}\}, \qquad i = 1, 2, \cdots. \tag{1}$$

The sequence N_i increases to $+\infty$, and

$$C^{-(i+1)} < \delta_k \leqq C^{-i} \quad \text{if} \quad N_i \leqq k < N_{i+1}, \qquad i = 1, 2, \cdots. \tag{2}$$

Let $\Delta N_i = N_{i+1} - N_i$, $i = 0, 1, \cdots$.

THEOREM 2. *For a given $\epsilon > 0$, $\epsilon \leqq 1$, let j be defined by*

$$C^{-(j-1)} < \epsilon \leqq C^{-(j-2)}. \tag{3}$$

Then

$$(N_1 + \cdots + N_{j-3}) \log C \leqq H_\epsilon(A); \tag{4}$$

$$H_\epsilon(A) \leqq (N_1 + \cdots + N_j) \log C + N_j \log \frac{9}{C-1} + \sum_{i=0}^{j-1} N_i \log \frac{N_j}{\Delta N_i} + N_1 \log \delta_0. \tag{5}$$

PROOF. We begin with (4). Let n be arbitrary, and let $A_n = A \cap X_n$. Let $y_1, \cdots, y_m$, $m = M_\epsilon(A_n)$, be a maximal set of ϵ-distinguishable points in A_n. The closed balls of X_n with centers y_i and radii ϵ cover A_n. Therefore $m\lambda_n\epsilon^n \geqq |A_n|$. By Theorem 1 and 2(1) we obtain

$$M_\epsilon(A) \geqq M_\epsilon(A_n) \geqq \frac{\delta_0 \cdots \delta_{n-1}}{\epsilon^n}, \quad C_\epsilon(A) \geqq \sum_{k=0}^{n-1} \log \frac{\delta_k}{\epsilon} \cdot$$

We take $n = N_{j-2}$. By (2) and (3),

$$\begin{aligned} C_\epsilon(A) &\geqq \sum_{i=1}^{j-3} \sum_{N_i \leqq k < N_{i+1}} \log \frac{\delta_k}{\epsilon} \\ &\geqq \sum_{i=1}^{j-3} \Delta N_i \log (C^{-i-1}\, C^{j-2}) = \sum_{i=1}^{j-3} (j - i - 3) \Delta N_i \log C \\ &= (N_1 + N_2 + \cdots + N_{j-3}) \log C. \end{aligned}$$

To prove (5), we apply several times Lemma 3. Let $\epsilon_i>0$, and let $\sum_0^{j-1}\epsilon_i=(1-1/C)\epsilon$. We have

$$H_\epsilon(A) \leqq \sum_{i=0}^{j-1} H_{\epsilon_i}(B_i) + H_{\epsilon/C}(B), \tag{6}$$

where B is the set $A(\delta_{N_j}, \delta_{N_{j+1}}, \cdots)$ in some Banach space Y, while B_i is the set $A(3\delta_{N_i}, \cdots, 3\delta_{N_{i+1}-1})$ in some ΔN_i-dimensional space Y_i. [If $\Delta N_i=0$ for some i, then the space Y_i is 0-dimensional, and $H_\delta(B_i)=0$ for each $\delta>0$. For these i, we shall interpret the terms of the sum in (6), and of the last sum in (5) to be zero, and also put $\epsilon_i=0$.] By (2) and (3), $\delta_{N_j}<\epsilon/C$, and so B is contained in the ball $U_{\epsilon/C}$ of the space Y. Hence $H_{\epsilon/C}(B)=0$. On the other hand, B_i is contained in the ball U_{r_i} of Y_i with $r_i=3\delta_{N_i}$. We estimate $H_{\epsilon_i}(U_{r_i})$ by means of 2(2) and obtain

$$H_\epsilon(A) \leqq \sum_{i=0}^{j-1} \Delta N_i \log (9\delta_{N_i}/\epsilon_i). \tag{7}$$

We can select

$$\epsilon_i = (1-1/C)\epsilon\Delta N_i/N_j \quad i=0,1,\cdots,j-1, \tag{8}$$

then

$$\begin{aligned} H_\epsilon(A) &\leqq \sum_{i=0}^{j-1} \Delta N_i \log\left(\frac{9}{C-1}\,\frac{\delta_{N_i}N_j}{\Delta N_i}\,C^j\right) \\ &\leqq N_1\log\delta_0 + \left\{\Delta N_0 \log\frac{9C^j}{C-1} + \sum_{i=1}^{j-1}\Delta N_i\log\frac{9C^{j-i}}{C-1}\right\} \\ &\qquad + \sum_{i=0}^{j-1}\Delta N_i \log\frac{N_j}{\Delta N_i}\cdot \end{aligned} \tag{9}$$

According to (2), the expression in the braces does not exceed

$$\begin{aligned} \sum_{i=0}^{j-1}\Delta N_i\log\frac{9C^{j-i}}{C-1} &\leqq N_j\log\frac{9}{C-1} + \sum_{i=0}^{j-1}(j-i)\Delta N_i\log C \\ &= (N_1+\cdots+N_j)\log C + N_j\log\frac{9}{C-1}, \end{aligned}$$

and (5) follows.

In many cases, we can put $C=e$ in (4) and (5); then we have

$$\begin{aligned} N_1+\cdots+N_{j-3} \leqq H_\epsilon(A) &\leqq N_1+\cdots+N_j+2N_j \\ &\qquad + \sum_{i=0}^{j-1} N_i\log\frac{N_j}{\Delta N_i} + N_1\log\delta_0. \end{aligned} \tag{10}$$

In other cases, it is necessary to take C sufficiently close to 1, so that the N_i form a fairly dense net.

Theorem 2 should be compared with weaker results of Brudnyĭ and Timan [5]. Their Theorem 1 reads in our notations (but with logarithms in the definition of $H_\epsilon(A)$ to the base 2)

$$H_\epsilon(A) \geqq n + 1 \quad \text{if} \quad 0 < \epsilon < \tfrac{1}{4}\delta_{n-1}.$$

This is a weaker version of (4); for if in (1) and (3) $C=2$, one obtains $N_{j-3} \geqq n$. In the opposite direction these authors have [35, p. 392], [5, Theorem 7]:

$$H_{e^2\epsilon}(A) \leqq (N_j + 1) \log \frac{e^2\delta_0}{\epsilon} + (N_j + 1) \log (N_j + 2).$$

An inequality of this type follows from (5). In fact, we have

$$N_1 + \cdots + N_j \leqq jN_j \leqq \left(\log \frac{1}{\epsilon} + 2\right) N_j.$$

We now turn our attention to some more or less precise asymptotic formulas for $H_\epsilon(A)$, which can be derived from Theorem 2 for special classes of sequences δ_n.

LEMMA 4. *One has*

$$(11) \quad S_j = \sum_{i=0}^{j-1} \Delta N_i \log \frac{N_j}{\Delta N_i} \leqq N_1 + \cdots + N_j, \quad j = 1, 2, \cdots.$$

For the proof we note that, if $S_0 = 0$, then for $k = 1, 2, \cdots$,

$$(12) \quad \begin{aligned} S_k - S_{k-1} &= \Delta N_{k-1} \log \frac{N_k}{\Delta N_{k-1}} + \sum_{i=0}^{k-2} \Delta N_i \log \frac{N_k}{N_{k-1}} \\ &= (N_k - N_{k-1}) \log \frac{N_k}{N_k - N_{k-1}} + N_{k-1} \log \frac{N_k}{N_{k-1}} \cdot \end{aligned}$$

Now the function $f(x) = (b-x) \log (b/(b-x)) + x \log (b/x)$ has its maximum on $[0, b]$ equal to $b \log 2 < b$. This proves that $S_k - S_{k-1} \leqq N_k$, hence the inequality (11).

LEMMA 5. *Let* $C = e$ *in* (1) *and* (3). (i) *If*

$$(13) \quad \delta_{2n} \leqq c\delta_n, \quad n = 0, 1, \cdots \text{ for some } 0 < c < 1,$$

then

$$(14) \quad N_{i+1} \leqq \text{Const. } N_i.$$

(ii) *If*

$$\frac{\delta_{[1+\epsilon)n]}}{\delta_n} \to 0 \text{ as } n \to \infty \text{ for each } \epsilon > 0, \tag{15}$$

then

$$\frac{N_{i+1}}{N_i} \to 1, \tag{16}$$

$$S_j = \sum_{i=1}^{j-1} \Delta N_i \log \frac{N_j}{\Delta N_i} = o(N_1 + \cdots + N_j). \tag{17}$$

PROOF. (i) Let p be a natural number such that $c^p \leqq e^{-1}$. We have, by (13),

$$\begin{aligned} N_i &= \min \{k: \delta_k \leqq e^{-i}\} \geqq \min \{k: \delta_{2k} \leqq ce^{-i}\} \\ &= \tfrac{1}{2} \min \{2k: \delta_{2k} \leqq ce^{-i}\} \geqq \tfrac{1}{2} \min \{k: \delta_k \leqq ce^{-i}\}. \end{aligned}$$

Repeating this, we obtain

$$N_i \geqq \frac{1}{2^p} \min \{k: \delta_k \leqq c^p e^{-i}\} \geqq \frac{1}{2^p} N_{i+1}.$$

(ii) Let $\epsilon > 0$ be given. For all sufficiently large i, the inequality $\delta_k \leqq e^{-i}$ implies that $\delta_{[(1+\epsilon)k]}/\delta_k \leqq e^{-1}$. For all such i,

$$\begin{aligned} N_i &= \min \{k: e^{-1}\delta_k \leqq e^{-i-1}\} \geqq \min \{k: \delta_{[(1+\epsilon)k]} \leqq e^{-i-1}\} \\ &\geqq \frac{1}{1+\epsilon} \min \{[(1+\epsilon)k]: \delta_{[(1+\epsilon)k]} \leqq e^{-i-1}\} \\ &\geqq \frac{1}{1+\epsilon} N_{i+1}. \end{aligned}$$

Hence for all sufficiently large i, $1 \leqq N_{i+1}/N_i \leqq 1+\epsilon$, and (16) is established.

From (12), with $x_k = N_{k-1}/N_k$ we have

$$S_k - S_{k-1} = N_k \left\{(1 - x_k) \log \frac{1}{1 - x_k} + x_k \log \frac{1}{x_k}\right\}.$$

For large k, $0 \leqq 1 - x_k \leqq \epsilon$. The functions $(1-x) \log (1/1-x)$ and $x \log (1/x)$ approach zero for $x \to 0+$. Therefore, for all large k, the expression in the braces is $\leqq \epsilon_1$, where $\epsilon_1 > 0$ is arbitrarily small. Thus $S_k - S_{k-1} = o(N_k)$. Therefore we obtain (7).

THEOREM 3. *Let $C=e$. Then*

$$H_\epsilon(A) = N_1 + \cdots + N_j + O(S_j); \tag{18}$$

If δ_n satisfies (13), *then*

$$H_\epsilon(A) \approx N_1 + \cdots + N_j; \tag{19}$$

If δ_n satisfies (15), *then*

$$H_\epsilon(A) \sim N_1 + \cdots + N_j. \tag{20}$$

PROOF. (18) follows from (10) and the inequality $N_j \leqq S_j$; (19) and (20) follow from (10) and Lemmas 4 and 5. More generally, we have (19) whenever the N_i satisfy $N_{i+1}=O(N_1+\cdots+N_i)$.

Theorem 3 applies, roughly, to sequences δ_n that decrease to zero as n^α, $\alpha>0$, or faster. The counterpart, Theorem 5, will deal with sequences δ_n that decrease to zero slower than n^α, $\alpha>0$. For the applications in §4.2, however, we insert its more complicated version, Theorem 4.

LEMMA 6. (i) *If, for some* $0<c<1$,

$$\delta_{2n} \geqq c\delta_n, \tag{21}$$

then, for the N_i defined by (1) *with* $C=c^{-1}$,

$$N_{i+1} \geqq 2N_i - 1. \tag{22}$$

(ii) *If*

$$\lim_{n\to\infty} \frac{\delta_{2n}}{\delta_n} = 1, \tag{23}$$

then for each $C>1$,

$$\lim_{i\to\infty} \frac{N_{i+1}}{N_i} = +\infty. \tag{24}$$

PROOF. We prove (ii); (i) is simpler. Let $C>1$ and a large integer p be given; we define c, $0<c<1$ by $c^p=C^{-1}$. For all large k, $\delta_{2k}\geqq c\delta_k$, hence for all large i

$$\begin{aligned} N_i &\leqq \min\{k: c^{-1}\delta_{2k} \leqq C^{-i}\} \\ &\leqq \tfrac{1}{2}\min\{k: \delta_k \leqq cC^{-i}\} + \tfrac{1}{2}. \end{aligned}$$

Repeating this p times, we obtain

$$N_i \leqq \frac{1}{2^p}\min\{k: \delta_k \leqq c^pC^{-i}\} + 1 \leqq \frac{1}{2^p} N_{i+1} + 1.$$

A statement weaker than (22) has been derived by Brudnyĭ and Timan [**5**, Theorem 2] from the assumption that $\delta_{n[\log n]} \geqq c\delta_n$.

The following theorem compares the entropies of two full approximation sets $A = A(\Delta, \Phi)$ and $A' = A(\Delta', \Psi)$ if δ_n decreases slowly.

THEOREM 4. (i) *Assume that* δ_n *satisfies* (21) *and that* $\delta_n' \leqq \delta_n$ *for all large* n. *If* $0 < q < 1$, *then there is a* $c_1 > 1$ (*which depends only on* q *and* c) *such that for all small* $\epsilon > 0$,

$$H_{c_1\epsilon}(A') \leqq qH_\epsilon(A). \tag{25}$$

(ii) *If the sequence* δ_n *satisfies* (23) *and if* $\delta_n' \leqq \delta_n$ *for large* n, *then for each* $c_1 > 1$,

$$\lim_{\epsilon \to 0} \{H_{c_1\epsilon}(A')/H_\epsilon(A)\} = 0. \tag{26}$$

PROOF. We shall establish (i); the proof of (ii) is simpler. We take $C = c^{-1}$. Let r be a large integer, to be fixed later. We take $c_1 \geqq C^{4+r}$. If j_1 corresponds to $\epsilon_1 = c_1\epsilon$ according to (3), then

$$C^{-(j-1)} < \epsilon \leqq c_1^{-1}C^{-(j_1-2)} \leqq C^{-2-r-j_1},$$

hence

$$j_1 \leqq j - r - 3. \tag{27}$$

Let N_i and N_i' denote the numbers (1) that correspond to the sets A and A'. By (5), (11), (27), (22) and (4), if $C_1 = 1 + \log(9C/(C-1))$,

$$\begin{aligned}
H_{\epsilon_1 c}(A') &\leqq O(1) + (1 + \log C)(N_1' + \cdots + N_{j_1}') + N_{j_1}' \log \frac{9}{C-1} \\
&\leqq O(1) + \left(1 + \log \frac{9C}{C-1}\right)(N_1' + \cdots + N_{j_1}') \\
&\leqq O(1) + C_1(N_1 + \cdots + N_{j-r-3}) \\
&\leqq O(j) + 2^{-r}C_1(N_1 + \cdots + N_{j-3}) \\
&\leqq \frac{o(1) + 2^{-r}C_1}{\log C} H_\epsilon(A) \leqq qH_\epsilon(A),
\end{aligned}$$

if r is selected large enough.

From Theorem 5 one can obtain relations that describe the behavior $H_\epsilon(A)$ for slowly decreasing δ_n. Taking $A' = A$, we have:

THEOREM 5. *If the sequence* δ_n *satisfies* (21), *then for each* $0 < q < 1$ *and properly chosen* $c_1 > 1$,

$$H_{c_1\epsilon}(A) \leqq qH_\epsilon(A). \tag{28}$$

If the sequence δ_n *satisfies* (23), *then for each* $c_1>1$,

$$\lim_{\epsilon\to 0}\{H_{c_1\epsilon}(A)/H_\epsilon(A)\} = 0. \tag{29}$$

REMARK. Similar statements hold for sets $A(\Delta)$ in *complex* Banach spaces. However (compare §2, Remark) one has then to replace the N_i in the formulas (4), (5), (11) by $2N_i$.

4. Applications of main theorems.

4.1. *Results of Vituškin's type for linear approximation.* Vituškin [37] has proved several theorems which provide lower bounds for the degree of approximation for some sets A of functions. They apply to general, not necessarily linear, approximation (for example, to rational approximation), and will be discussed in §§6 and 7. For the linear approximation, his results reduce to statements of the following type. If for some n and some $\epsilon>0$, the nth width (for the definition and properties of widths see [24, Chapter 9], [33]) of A satisfies $d_n(A)<\epsilon$, then n is at least as large as some simple lower bound, that depends on ϵ, and is, roughly, $H_\epsilon(A)$.

In this section we deduce results of this type from Theorem 2. They will be better than the special cases of Vituškin's theorems, mainly because our sets A are arbitrary, or restricted by the asymptotic behavior of $H_\epsilon(A)$, while Vituškin restricts the structure of the class A.

THEOREM 6. *Let A be an arbitrary compact set in a separable Banach space X, and let $d_n(A)<\epsilon$. Then*

$$n \geqq \frac{H_{e^2\epsilon}(A) - C_0}{2+\log\dfrac{1}{\epsilon}+\log\left[H_{e^2\epsilon}(A)-C_0\right]}\,. \tag{1}$$

If, in particular,

$$H_{2\epsilon}(A)/H_\epsilon(A)\to 1 \quad as \quad \epsilon\to 0, \tag{2}$$

then $d_n(A)<\epsilon$ implies

$$n \geqq \frac{(1-o(1))H_\epsilon(A)}{\log\dfrac{1}{\epsilon}}\,. \tag{3}$$

If for each q, $0<q<1$ there is a $c_1>0$ for which $H_{c_1\epsilon}(A)\leqq qH_\epsilon(A)$, then $d_n(A)<\epsilon$ implies

$$n \geqq C_1 H_{e^2\epsilon}(A) - C_2. \tag{4}$$

PROOF. There exists a fundamental sequence Φ in X, for which $d_n(A) \leqq E_n^\Phi(A) < \epsilon$. Let $\delta_k = E_k^\Phi(A)$, $k=0, 1, \cdots$, then $A \subset A' = A(\Delta, \Phi)$. In view of the monotonicity of the right-hand side of (1), it is sufficient to prove that if $\delta_n \leqq \epsilon$, then n is not less than the right-hand side of (1) with A replaced by A'. We take $C=e$ in 3(1) and 3(3), and have by 3(10)

$$(5) \qquad H_{e^2\epsilon}(A') - N_1 \log \delta_0 \leqq N_1 + \cdots + N_j + 2N_j + N_j \log N_j,$$

where j is defined by $e^{-j-1} < \epsilon \leqq e^{-j}$. Since $\delta_n \leqq e^{-j}$, we have $n \geqq N_j$. Hence

$$H_{e^2\epsilon}(A') - C_0 \leqq (j + 2 + \log n)n \leqq (\log (1/\epsilon) + 2 + \log n)n.$$

Now we use the simple fact that if $\alpha \leqq n(\beta + \log n)$, where $\alpha > 0$, $\beta > 0$, $\beta + \log \alpha > 1$, then $n > \alpha/(\beta + \log \alpha)$. This proves (1). Relation (3) is a special case, since (2) implies that

$$\log H_\epsilon(A) = o\left(\log \frac{1}{\epsilon}\right).$$

To prove (4), we derive from 3(10) and 3(11)

$$(6) \qquad H_{e^2\epsilon}(A') - C_0 \leqq 4(N_1 + \cdots + N_j).$$

By 3(10) and the assumption with $q = 1/5$ we have, if r is sufficiently large,

$$N_1 + \cdots + N_{j-r-1} \leqq H_{e^r\epsilon}(A') \leqq \tfrac{1}{5} H_{e^2\epsilon}(A').$$

Therefore (6) yields

$$(7) \qquad \tfrac{1}{5} H_{e^2\epsilon}(A') - C_0 \leqq 4(r+1)n,$$

and (4) follows.

The result (3) should be compared with Vituškin's [**37**, p. 177, Theorem 1], where this inequality is proved for classes of analytic functions, and (4) with his [**37**, p. 182, Theorem 1], which holds for classes of "smooth" functions, such as $\Lambda_{p,\omega}^s$ (see §5). Results of Brudnyĭ and Timan [**5**, Theorems 3 and 5] are very special cases of (1), for compacts in a Hilbert space.

4.2. *Comparison of sets* $A(\Delta, \Phi)$ *for different* Δ. The purpose of this section is to show that the set $A(\Delta, \Phi)$ decreases drastically if the numbers δ_n are replaced by smaller numbers δ_n', even if Φ is replaced by another set Ψ in this process. For the sets $A = A(\Delta, \Phi)$ and $A' = A(\Delta', \Psi)$ we prove that $H_\epsilon(A')$ is asymptotically smaller than $H_\epsilon(A)$. In this context, entropy plays a role that is usually reserved for measure, category or cardinality in theorems about the comparison of sets. Results of this type allow us to establish the existence of elements $f \in X$, for which the Ψ-approximation is not essentially better than the Φ-approximation.

THEOREM 7. *Let $A = A(\Delta, \Phi)$ and $A' = A(\Delta', \Psi)$ be two full approximation sets in a Banach space X.* (i) *If*

$$\delta_{2n} \geqq c\delta_n, \qquad n = 1, 2, \cdots c > 0, \tag{8}$$

and if $0 < q < 1$, then there is a $q_1 > 0$ (which depends only on c and q) with the property that $\limsup(\delta_n'/\delta_n) < q_1$ *implies*

$$\limsup_{\epsilon \to 0} \{H_\epsilon(A')/H_\epsilon(A)\} < q; \tag{9}$$

(ii) *If δ_n satisfies* (8) *and if* $\lim(\delta_n'/\delta_n) = 0$, *then*

$$\lim_{\epsilon \to 0} \{H_\epsilon(A')/H_\epsilon(A)\} > 0; \tag{10}$$

(iii) *Relation* (10) *holds also if* $\limsup(\delta_n'/\delta_n) < q_1$, $q_1 < 1$ *and if*

$$\lim_{n \to \infty} (\delta_{2n}/\delta_n) = 1. \tag{11}$$

PROOF (i). Take c_1 from Theorem 4 (i), and put $q_1 = c_1^{-1}$. Then $\delta_n'' = \delta_n'/q_1 \leqq \delta_n$ for all large n. According to Theorem 4 (i) we have $H_{c_1\epsilon}(A(\Delta'')) \leqq qH_\epsilon(A)$, for all small $\epsilon > 0$. But $H_{c_1\epsilon}(A(\Delta'')) = H_\epsilon(A(\Delta'))$. Statement (ii) follows from (i), while (iii) can be proved in a like manner, with the help of Theorem 5 (ii).

We shall use Theorem 7 to show that even an introduction of countably many sequences Ψ_p instead of Φ cannot essentially reduce the degree of approximation of the elements $f \in A(\Delta, \Phi)$.

THEOREM 8. *Let $\Phi, \Psi_p, p = 1, 2, \cdots$ be fundamental sequences in X. Assume that the sequence δ_n satisfies* (8), *and let*

$$\limsup \delta_n^{(p)} = 0, \quad p = 1, 2, \cdots, \delta_n. \tag{12}$$

Then there exists an element $f \in X$ such that

$$E_n^{\Phi}(f) \leqq \delta_n, \quad n = 0, 1, \cdots \tag{13}$$

but the condition

$$E_n^{\Psi_p}(f) \leqq \delta_n^{(p)}, \quad n = 0, 1, \cdots \tag{14}$$

is violated for each p.

PROOF. Let Y be the set of all $f \in X$ with the property that $E_n^{\Phi}(f) = O(\delta_n)$, and the norm $\rho(f) = \sup_n \{E_n^{\phi}(f)/\delta_n\}$. Standard arguments show that Y is a Banach space with this norm. Let also V_p, $p=1, 2, \cdots$ be the set of $f \in Y$ for which $E_n^{\Psi_p}(f) \leqq \delta_n^{(p)}$, $n=0, 1, 2, \cdots$.

Now $E_n^{\Psi_p}(f)$ is a continuous function of $f \in Y$ in the norm ρ, for $\rho(f_s - f) \to 0$, $s \to +\infty$ implies $\|f_s - f\| \to 0$, and $E_n^{\Psi_p}(f)$ is continuous on X. Hence each V_p is a closed subset of Y. From Theorem 7 (ii) it follows that $H_\epsilon(V_p) < H_\epsilon(U)$, for all small $\epsilon > 0$, for each ball U in Y (with an arbitrary center and radius). Thus, V_p does not exhaust U. In other words, V_p is nowhere dense in Y. It follows, by Baire's theorem, that $Y - \bigcup_p V_p \neq 0$.

THEOREM 9. *Assume that δ_n satisfies* (8). *There is a $\rho > 0$ with the property that for each two fundamental sequences Φ, Ψ, there is an element $f \in X$ which satisfies* (13) *and the inequality*

$$\overset{\Psi}{E}_n(f) \geqq \rho\delta_n \quad \textit{for infinitely many } n. \tag{15}$$

PROOF. We select ρ, $0 < \rho < 1$ so that $q_1 = 4\rho/c$ works in Theorem 7(i) for some $q < 1$. By induction we define increasing sequences of integers n_i, m_i, $i = 1, 2, \cdots$ $(m_0 = 0)$, and elements f_i such that

$$m_i > 2i, m_i \leqq n_i < m_{i+1}, i = 1, 2, \cdots ; \delta_{m_k} \leqq \frac{\rho}{2^{k-i}} \delta_{m_i} \quad \text{for } i < k. \tag{16}$$

The $(k+1)$st step of the induction is as follows. We define two sequences of positive numbers Δ_{k+1}, $\bar{\Delta}_{k+1}$, putting

$$\begin{aligned} \delta_n^{(k+1)} &= \tfrac{1}{2}\delta_{m_k} \text{ for } n \leqq m_k, & \bar{\delta}_n^{(k+1)} &= \delta_0 \text{ for } n < k + m_k, \\ &= \tfrac{1}{2}\delta_n \text{ for } n > m_k, & &= 2\rho\delta_{n-k} \text{ for } n \geqq k + m_k. \end{aligned} \tag{17}$$

We take $\Psi_{k+1} = \{f_1, \cdots f_k, \psi_0, \psi_1, \cdots\}$. Then

$$\limsup_{n\to\infty} \frac{\bar{\delta}_n^{(k+1)}}{\delta_n^{(k+1)}} = \limsup_{n\to\infty} \frac{2\rho\delta_{n-k}}{\frac{1}{2}\delta_n} \leqq \limsup_{n\to\infty} \frac{4\rho\delta_n}{\delta_{2n}} \leqq \frac{4\delta}{c} = q_i.$$

By Theorem 7(i) there exists an element $f = f_{k+1}$ of $A(\Delta_{k+1}, \Phi)$ which does not belong to $A(\bar{\Delta}_{k+1}, \Psi_{k+1})$; we can assume that $f_{k+1} \in X_{m_{k+1}}$ for some large m_{k+1}. We may further assume that

$$m_{k+1} > 2k + 2; \quad \delta_{m_{k+1}} \leqq \frac{\rho}{2^{k+1-i}} \delta_{m_i} \qquad \text{for } i < k + 1. \tag{18}$$

Then

$$E_n^{\Phi}(f_{k+1}) \leqq \frac{\rho}{2^{k+1-i}}\,\delta_n \qquad \text{for } m_i \leqq n < m_{i+1},\ i \leqq k, \tag{19}$$
$$E_n^{\Psi_{k+1}}(f_{k+1}) = 0 \qquad \text{for } n \geqq m_{i+1}.$$

On the other hand,

$$E_n^{\Psi_{k+1}}(f_{k+1}) > \bar{\delta}_n^{(k+1)} \qquad \text{for at least one } n. \tag{20}$$

For $n<k+m_k$, $\bar{\delta}_n^{k+1}=\delta_0$ and since

$$E_n^{\Psi_{k+1}}(f_{k+1}) \leqq \|f_{k+1}\| = E_0^{\Phi}(f_{k+1}) \leqq \delta_0,$$

we must have $n \geq k+m_k$ for our n. Selecting one such n that satisfies (20), we put $n=k+n_k$. Increasing m_{k+1}, if necessary, we have $m_k \leq n_{k+1} < m_{k+1}$. For $n=n_{k+1}$,

$$E_n^{\Psi}(f_1 + \cdots + f_k + f_{k+1}) \geqq E_{n+k}^{\Psi_{k+1}}(f_{k+1}) > 2\rho\delta_n.$$

We put $f=\sum_{k=1}^{\infty} f_k$. If $m_i \leq n < m_{i+1}$, then by (19),

$$E_n^{\Phi}(f) = E_n^{\Phi}\left(\sum_{k=i+1}^{\infty} f_k\right) \leqq \sum_{k=i+1}^{\infty} E_n^{\Phi}(f_k) \leqq \sum_{k=i+1}^{\infty} \frac{\rho}{2^{k-i}}\delta_n = \rho\delta_n < \delta_n.$$

This shows that $f \in A(\Delta, \Phi)$. On the other hand, for $n=n_k$, $k=1, 2, \cdots$,

$$E_n^{\Psi}(f) \geqq E_n^{\Psi}(f_1 + \cdots + f_k) - \sum_{i=k+1}^{\infty} \|f_i\| \geqq 2\rho\delta_n - \sum_{i=k+1}^{\infty} \frac{\rho}{2^{i+1-k}}\delta_n \geqq \rho\delta_n.$$

Theorem 7 (for a Hilbert space X) was given in Lorentz [23]. For special spaces X ($X=C$ and $X=L^p$) and $\delta_n=n^{-\alpha}$, $\alpha>0$, Theorem 9 appears in Lorentz [21]. Olevskiĭ [27] has an interesting counterpart of this when $X=C[a, b]$, $A=\text{Lip}\,\alpha$, $0<\alpha<1$: For each system Φ there exists a function $f \in \text{Lip}\,\alpha$, with the property that $E_n^{\Phi}(f) \geqq C\omega(f, 1/n)$, $C>0$ for all $n=1, 2, \cdots$. Here $\omega(f, 1/n)$ is, possibly, much smaller than $n^{-\alpha}$, which serves as δ_n.

5. Review of recent results on entropy.

5.1. *Computation of entropy.*

5.1.1. *Continuous and differentiable functions.* First we introduce some notations (compare [24, Chapter 3]). Let $r=0, 1, \cdots$ be an integer, ω a modulus of continuity, S an s-dimensional parallelepiped,

$a_i \leqq x_i \leqq b_i$, $i=1, \cdots, s$. Let $\Lambda^s_{r\omega} = \Lambda^s_{r\omega}(S; M_0, \cdots, M_r)$ be the class of all functions $f(x_1, \cdots, x_s)$ on S which have all partial derivatives $|D^{(k)}f| \leqq M_k$, $k=0, \cdots, r$. In addition, the modulus of continuity of each rth derivative should not exceed $\omega(h)$. We write Λ_ω for the class $\Lambda_{0\omega}$, Lip α for Λ_ω, $\omega(h)=Mh^\alpha$. In case $s=0$, functions f may be defined on an arbitrary compact metric space B. The theorems about the degree of approximation of the classes $\Lambda^s_{r\omega}(S)$ by algebraic polynomials are well known (see [35, pp. 279 and 363], [24, Chapters 4, 5, 6]). They imply that Λ^s_{rw} is contained between two full approximation sets of the space $C[S]$ of continuous functions: $A(\Delta') \subset \Lambda^s_{rw} \subset A(\Delta)$, where both δ_n and δ_n' are of the form $Cn^{-r/s}\omega(C'n^{-1/s})$. An application of Theorem 2 leads to the result:

$$\frac{C_1}{\delta(\beta\epsilon)^s} \leqq H_\epsilon(\Lambda^s_{r\omega}) \leqq \frac{C_2}{\delta(\gamma\epsilon)^s}, \tag{1}$$

where C_1, C_2, β, γ are positive constants, and $\delta=\delta(\epsilon)$ is defined by the equation $\delta^r\omega(\delta)=\epsilon$. If $r \geqq 1$, or if $r=0$, $\omega(h)=h^\alpha$, $0<\alpha\leqq 1$, (1) leads to an asymptotic determination of $H_\epsilon(A)$, namely

$$H_\epsilon(A) \approx \frac{1}{\delta(\epsilon)^s}. \tag{2}$$

With a different proof, results (1) and (2) are due to Kolmogorov (in [19], they are proved for $\omega(h)=h^\alpha$). For moduli of continuity $\omega(h)$ which increase more rapidly than h^α at $h=0$, one does not get in this way precise formulas for the entropy. According to Timan [36], however, (2) still holds for each concave modulus of continuity, in the corrected form

$$H_\epsilon(\Lambda^s_\omega) \approx \left\{\frac{1}{\omega^{-1}(2\epsilon)}\right\}^s. \tag{3}$$

This is obtained by finding an $M=M(\epsilon)>0$ for which the class $\Lambda^s_\omega(B)$ is approximated by $\Lambda^s_w(B)$, $w(h) = Mh$ with an error $\leqq \frac{1}{2}[\omega(\epsilon)-M\epsilon]$ (compare [24, p. 122]). It remains to take advantage of the very precise formulas for the entropy, that are valid for the class Λ^s_w ([19, pp. 14, 78]).

Vosburg [40] improves a little the inequalities (1) if $\Lambda=\text{Lip}(\alpha, B)$ and if B is an arbitrary compact set contained in $[-1, +1]$. In this way he shows: $H_\epsilon(\text{Lip}(\alpha, B))=o(\epsilon^{-1/\alpha})$ if and only if the set B has measure zero.

The set of continuous functions is not compact in the uniform norm. However, the set G of all graphs of continuous functions in the unit square is compact in the Hausdorff metric space of closed sets in the plane. Its entropy is $H_\epsilon(G) \approx (1/\epsilon) \log (1/\epsilon)$ (Clements [**6**], Penkov and Sendov [**29**]). Another result of Clements [**6**] concerns the class V_α of functions of bounded variation on $[0, 1]$, which belong to Lip α, in the uniform norm. Then $H_\epsilon(V_\alpha) \approx (1/\epsilon) \log (1/\epsilon)$ for each α, $0 < \alpha < 1$. An interesting unsolved question is the determination of the widths of the sets V_α. Brudnyĭ and Kotljar [**4**], [**20**] study the entropy of classes A of functions $f(x_1, \cdots, x_s)$ that have different properties with respect to each variable x_i. They assume that for each i, the k_ith difference of f with respect to x_i satisfies

$$\left| \Delta_h^{k_i} f(x_1, \cdots, x_s) \right| \leqq M_i h^{\beta_i}, \tag{4}$$

where $0 < \beta_i \leqq k_i$, and derive, for this class A, $H_\epsilon(A) \approx (1/\epsilon)^{\sum_1^n \beta_i^{-1}}$. This also follows from our Theorem 2 and results about polynomial approximation (Timan [**35**, p. 279]).

5.1.2. *The L^p-norm.* Analogues of classes Λ of 5.1.1 can be defined in the L^p-norm. They are denoted $\Lambda_{r\omega}^{sp}(S)$. The known approximation theorems [**35**] and Theorem 2 lead to statements of the type (2). For example we have:

THEOREM 10. *The set $A = \mathrm{Lip}(\alpha, p)$, $0 < \alpha \leqq 1$, $p \geqq 1$ of functions f on $[0, 1]$ for which $\int_0^1 |f|^p dx \leqq 1$ and $\int_0^{+1} |f(x+h) - f(x)|^p dx \leqq Ch^{p\alpha}$ has the entropy $H_\epsilon(A) \approx (1/\epsilon)^{s/\alpha}$. Similarly,*

$$H_\epsilon(\Lambda_{r\alpha}^{sp}) \approx (1/\epsilon)^{s/(r+\alpha)}. \tag{5}$$

These results seem to be new. For similar classes in L^2, for which our Theorem 2 can also be used, Brudnyĭ and Timan [**5**] obtain as an upper estimate only $(1/\epsilon)^{s/(r+\alpha)} \log (1/\epsilon)$.

Golovkin [**13**], for certain classes A, similar to $\Lambda_{r\alpha}^{sp}$, in spaces of functions with monotone and translation-invariant norm ($\|f\|$ is monotone if it increases whenever $|f|$ increases), announces estimates of $H_\epsilon(A)$ from above and below. They are of great generality, but not very precise. In many cases he determines the limit

$$\lim_{\epsilon \to 0} \left\{ \log H_\epsilon(A) / \log \frac{1}{\epsilon} \right\}.$$

Smoljak [**32**] considers sets of functions on the s-dimensional torus,

$$f(x_1, \cdots, x_s) = \sum_{m_1, \ldots, m_s = -\infty}^{+\infty} c_{m_1 \ldots m_s} \exp i(m_1 x_1 + \cdots + m_s x_s),$$

in the L^2-metric, that are restricted by some asymptotic properties of their Fourier coefficients. For the class $E_s^{\alpha k}(M)$ this restriction is

$$|c_{m_1 \ldots m_s}| \leqq M(\bar{m}_1 \cdots \bar{m}_s)^{-\alpha}[\log^k (\bar{m}_1 \cdots \bar{m}_s) + 1], \tag{6}$$

where $\bar{m} = |m|$ if $m \neq 0$, $0 = 1$, $\alpha > \frac{1}{2}$, $k \geqq 0$, and for $W_s^\alpha(M)$,

$$\sum |(\bar{m}_1 \cdots \bar{m}_s)^\alpha c_{m_1 \ldots m_s}|^2 \leqq M^2. \tag{7}$$

He finds:

$$\begin{aligned} H_\epsilon(E_s^{\alpha k}) &\approx \left(\frac{1}{\epsilon}\right)^{1/(\alpha - 1/2)} \log^{(2k+2\alpha(s-1))/(2\alpha-1)} \frac{1}{\epsilon}, \\ H_\epsilon(W_s^\alpha) &\approx \left(\frac{1}{\epsilon}\right)^{1/\alpha} \log^{s-1} \frac{1}{\epsilon} \cdot \end{aligned} \tag{8}$$

Arranging the functions $\phi = \exp i(k_1 x_1 + \cdots + k_s x_s)$ in proper order, we can compute from (6) and (7) upper bounds for the degrees of approximation δ_n, and imbed the above classes into sets $A(\Delta, \Phi)$. In this way, the upper (harder) estimates for the entropies given by (8) follow from our Theorem 2.

An interesting problem is to derive, from the behavior of $H_\epsilon(A)$, properties of the Fourier coefficients of some, or of most functions $f \in A$, $A \subset L^p$ or $A \subset C$. To give only one example: Is it true, for a subset A of C, that $H_\epsilon(A) \geqq \text{Const.}(1/\epsilon)^2$ implies the existence of a function $f \in A$ whose Fourier series is not absolutely convergent?

5.1.3. *Analytic functions.* For analytic functions, excellent approximation theorems are known. They can be used for the computation of entropies. This applies to Vituškin's estimates of classes A of functions defined on polycylinders [**37**, Chapter 2], [**24**, pp. 156–157]. Each of these classes is contained between two full approximation sets. The upper class is obtained from the approximation of functions $f \in A$ by partial sums of their Taylor series; the lower class is derived from inverse theorems of Bernšteĭn's type. The term S_j of 3(11) gives the remainder of Vituškin's formula. Erohin [**12**] proves the formula

$$H_\epsilon(A(K, G; M)) \sim \frac{1}{\log R} \log^2 \left(\frac{1}{\epsilon}\right), \tag{9}$$

where $A = A(K, G; M)$ is the set of functions, analytic in the open set G, with $|f(z)| \leqq M$, $z \in G$, in the uniform norm on the compact set $K \subset G$; R is the conformal radius of the pair K, G. This follows at once from our Theorem 2 and Theorem 3 in Erohin [**11**], which deals with approximation of functions $f \in A$ by linear combinations of elements of a properly chosen basis Φ. [The last quoted theorem should probably contain two different constants, for a lower and an upper estimate, instead of a single one $C(\delta)$.]

For some classes of functions analytic in a strip and for classes of harmonic functions, Tihomirov [**34**] has a better remainder, namely $O(\log(1/\epsilon))$, instead of Vituškin's $O(\log(1/\epsilon)\log\log(1/\epsilon))$. The new method of proof is based on some geometric properties of sets of Fourier coefficients.

Al′per [**2**] studies classes T_q of functions f analytic in a region G, bounded by a curve Γ and continuous on $\overline{G}$; the functions satisfy the conditions $|f^{(k)}(z)| \leqq C_k$, $k=0, \cdots, r$, $f^{(r)} \in \mathrm{Lip}\,\alpha$ on $\overline{G}$; $q=r+\alpha$, $0<\alpha\leqq 1$. Al′per refers to his own approximation theorems of functions $f \in T_q$; at present, more powerful results of Dzjadyk [**9**, Theorem 4.1]; [**10**, Theorem 1.7] are known, which allow even angular points of Γ. One obtains in this way $H_\epsilon(T_q) \approx (1/\epsilon)^{1/q}$, if Γ is smooth except for angles; Al′per also has $\log H_\epsilon(T_q) \sim (1/q) \log (1/\epsilon)$, if Γ has a continuously turning tangent, and $\log H_\epsilon(T_q) \approx \log(1/\epsilon)$, if Γ is a rectifiable Jordan curve.

5.1.4. *Compact sets in Banach spaces.* Compact sets in Hilbert and Banach spaces were treated by Brudnyĭ and Timan [**5**], see also [**35**]. Most of their results are special cases of ours. (Compare §3, §4.) Helemskiĭ and Henkin [**14**] use entropy in order to characterize the size of ellipsoids D which contain a given compact subset A of a Hilbert space. One of their results is that if $\limsup\{H_\epsilon(A)/\log(1/\epsilon)\} < q < 2$, then there exists an ellipsoid $D \supset A$ with

$$\limsup_{\epsilon\to 0} \left\{ H_\epsilon(D)/\log \frac{1}{\epsilon} \right\} < \frac{2q}{2-q} \cdot$$

It should be noted, that for ellipsoids D in spaces l^p, $1 \leqq p \leqq +\infty$, Mitjagin [**25**, Theorem 3] obtained inequalities similar to or even better than the inequalities of our Theorem 2. An ellipsoid in l^p is the set of all points $x=(x_1, x_2, \cdots)$ for which $\sum_{n=1}^{\infty} |x_n/\delta_n|^p \leqq 1$, where $\delta_n > 0$, $\delta_n \to 0$ is a given decreasing sequence. For arbitrary compacts he gives inequalities that connect $H_\epsilon(A)$ and the widths $d_n(A)$ [**25**, Theorem 4], [**24**, p. 164]. They do not imply 3(5).

5.2. *Applications of entropy.*

5.2.1. *Invariants of linear topological spaces.* Entropy-theoretic notions can be used in order to construct invariants of linear topological spaces. The *approximate dimension* of a linear topological space, defined by Kolmogorov [**18**] is less fine but easier to handle than Banach's dimension. All infinite-dimensional Banach spaces have the same (maximal) approximate dimension. This means that the new notion is fruitful only for topological spaces that are in some sense close to finite dimensional spaces, for example, for spaces of analytic functions. Pelczyński [**28**] also defined an invariant of this kind. Computing the approximate dimension, Kolmogorov showed that the spaces of analytic functions of different number of variables are not linearily homeomorphic. Similarly, Pelczyński [**28**] proved that the spaces of entire functions of one variable and of functions analytic in $|z|<1$ are topologically different. Compare the article of Rolewicz [**30**].

Another interesting application of entropic notions is a characterization of nuclear spaces among all spaces of type F, by Gel'fand and Mitjagin [**25**].

5.2.2. *Superpositions of functions.* According to Kolmogorov ([**17**], slightly improved in [**24**, Chapter 11]) there exist fixed continuous functions ϕ_i, ψ_i, $i=1, \cdots, 5$, that map $[0, 1]$ into itself and have the property that each function $f(x, y)$, continuous on the square $0 \leqq x$, $y \leqq 1$, is representable in the form

$$f(x, y) = \sum_{i=1}^{5} g(\phi_i(x) + \psi_i(y)). \tag{10}$$

Here $g(u)$, $0 \leqq u \leqq 2$ is some continuous function, depending on f. One can assume that the functions ϕ_i, ψ_i belong to the class Lip α *for each* $0<\alpha<1$. There exists a conjecture of Kolmogorov,[2] which, in its general form, states the following: Not all analytic functions of two variables are representable by means of superpositions of continuously differentiable functions of one variable and of addition; not all analytic functions of three variables are representable by superpositions of continuously differentiable functions of two variables. For "linear superpositions," such as (10), this has been proved by Vituškin [**38, 39**]. One of Henkin's [**15**] improvements of his results is the following interesting theorem. Assume that the functions p_i, ϕ_i, $i=1, \cdots, N$ are continuous in the whole plane, and that the ϕ_i are continuously differentiable. Then for each region D in the plane there are natural numbers n, m for which the polynomial $(x+ny)^m$ is not equal to any superposition

[2] Communication of D. Sprecher.

$$\sum_{i=1}^{N} p_i(x, y) g_i(\phi_i(x, y)), \tag{11}$$

where g_i are arbitrary bounded measurable functions. The proof depends upon the comparison of entropies $H_\epsilon^\delta(F)$ of sets F of functions $f(x, y)$ of the form (11) and of analytic functions, in the metric

$$\|f\| = \sup_{U_\delta \subset D} \left| \frac{1}{\pi\delta^2} \int_{-U_\delta}\int f(u, v) du\, dv \right|,$$

where $\delta > 0$ is a properly chosen function of ϵ, and U_δ is an arbitrary circle of radius δ.

As a corollary, representation (10) is not always possible if ϕ_i, ψ_i are continuously differentiable.

6. **Vituškin's theorems in the uniform metric.** The last two sections of this paper are devoted to the exposition of results from Vituškin's theory of nonlinear approximation, developed by him in his book [**37**]. In this section, we deal with sets of continuous and differentiable functions in the uniform metric. The main advantage of our approach in this section is its simplicity: we avoid the use of multidimensional variations [**37**, Chapter 4]. The simple combinatorial device which we use appears also in H. S. Shapiro [**31**]; see Remark at the end of §6.2. As Vituškin does, we use in an essential way Oleĭnik's estimates [**26**]; [**37**, p. 132]. In this way we obtain Theorem 13 below, which is essentially identical with Theorem 3 [**37**, p. 186] of Vituškin about piecewise rational approximation with moving barrier. This method is such that each improvement of Oleinĭk's estimate would automatically lead to some improvement of the final inequalities. For the case of a constant barrier and for "rational continuous" approximation, Vituškin has slightly better results (Theorem 12 below), which do not seem to follow by our method.

In §7 we will show that theorems of Vituškin's type are not restricted to uniform metric. Results obtained there are valid for polynomial approximation in Banach and L^p spaces. Here again we use the method of §3 for the computation of entropies.

6.1. *Definitions and lemmas.* Let $P(t)$ be a polynomial in $t = (t_1, \cdots, t_n)$, of degree p in the variables t_i jointly, which is not a constant. The equation $P(t) = 0$ splits the space R^n into a collection of connected sets. Important for us is an upper bound for their number, contained in the following lemma.

LEMMA 7 (OLEĬNIK). *The equation* $P(t)=0$ *defines at most* $2p^n$ *bounded components of* R^n, *and at most*

$$(1) \qquad 2(p+2)^n$$

connected bounded or unbounded components ($p\geqq 2$, $n\geqq 1$).

According to Oleĭnik [26], the number of bounded components of R^n defined by $P(t)=0$ is at most $(p-1)^n+2n$ if p is odd, and at most

$$\frac{1}{2}\,\frac{(p-1)^{n+1}-(p-1)}{p-2}+\frac{n(n+1)}{4}$$

if p is even. Each of these numbers does not exceed $2p^n$ if $p\geqq 2$, $n\geqq 1$. The second statement of the lemma follows from the fact that the number of components of R^n, defined by $P(t)=0$, which intersect the ball $t_1^2+\cdots+t_n^2\leqq a^2$ does not exceed the number of bounded components of R^n, defined by the equation $P(t)(t_1^2+\cdots+t_n^2-a^2)=0$.

In the classical theory of linear approximation, functions $f(x)$, $x\in B$, which belong to a given class, are approximated by *linear expressions*

$$t_1\phi_1(x)+\cdots+t_n\phi_n(x),$$

where the functions $\phi_i(x)$, $x\in B$ are given, and $t=(t_1,\cdots,t_n)\in R^n$ can be selected for a given f. With Vituškin [37], we consider more general expressions. A *rational expression* of degree p is a quotient

$$(2) \qquad R(x,t)=\frac{P(x,t)}{Q(x,t)},\quad x\in B,\ t\in R^n,$$

where P and Q are polynomials of degree p in t_i, $i=1,\cdots,n$, with coefficients which are given functions of x.[3] There may exist some x, t for which $R(x,t)$ is not defined, in other words, $Q(x,t)=0$. More general are piecewise rational expressions $R(x,t)$. Let

$$(3) \qquad \tilde{P}(x,t)$$

be a polynomial of degree q in $t=(t_1,\cdots,t_n)$, with coefficients which are given functions of x. According to Lemma 7, for each x, $\tilde{P}(x,t)=0$ decomposes R^n into at most $2(q+2)^n$ sets $\Gamma_j=\Gamma_j(x)$. On the closure of each Γ_j we put

$$(4) \qquad R(x,t)=\frac{P_j(x,t)}{Q_j(x,t)},$$

[3] Vituškin [37] formulates his theorems for the case when P, Q are polynomials of degree p in each t_i, $i=1,\cdots,n$ separately. Since his proofs are based on Lemma 7, this seems to be unjustified.

where P_j, Q_j are polynomials of degree p in t. This $R(x, t)$ is called a *piecewise rational expression of degree p with barrier* (3) *of degree q*. One of the justifications for the study of piecewise rational approximations lies in the fact that operations of taking maxima or minima, performed on rational expressions lead to piecewise rational expressions.

Sometimes simpler results are valid for piecewise rational approximations $R(x, t)$ with barrier $\tilde{P}(t)=0$ *independent of x*.

If $Q(x, t)\equiv 1$ in (2), we obtain a *polynomial expression* of degree p, or a *piecewise polynomial expression* of degree p with barrier of degree q. This case lends itself to generalizations for arbitrary Banach spaces.

We shall say that $R(x, t)$ is an ϵ-approximation of a class A of functions $f\in C(B)$, if for each $f\in A$ there is a $t\in R^n$ for which $R(x, t)$ is defined for all $x\in B$ and satisfies

$$|f(x) - R(x, t)| < \epsilon, \quad x \in B. \tag{5}$$

If A is a subset of a Banach space X, and $P(t)$ is a polynomial (or piecewise polynomial) expression whose coefficients are elements of X, we can say that P provides an ϵ-approximation for A if for each $f\in A$ there is a $t\in R^n$ with the property that

$$\|f - P(t)\| < \epsilon. \tag{6}$$

6.2. *Approximation of sets Λ in the uniform norm.* For an ϵ-approximation $R(x, t)$ of a set $A\subset C(B)$ we would like to obtain a lower bound for the numbers n, p, q.

Let A be a set of continuous functions on a set B, let $\epsilon>0$, and assume that B contains $M=M(\epsilon)$ points $x_1, \cdots, x_M$ such that for each distribution of signs $\lambda=(\lambda_\mu)$, $\mu=1, \cdots, M$, $\lambda_\mu=\pm 1$, there is a function $f\in A$ for which

$$|f(x_\mu)| \geqq \epsilon, \quad \operatorname{sign} f(x_\mu) = \lambda_\mu, \quad \mu = 1, \cdots, M. \tag{7}$$

THEOREM 11. *If a set A of the above kind has an ϵ-approximation $R(x, t)$, $t=(t_1, \cdots, t_n)$, which is piecewise rational of degree p with barrier of degree q, then*

$$n \log (p + q + M) \geqq CM, \tag{8}$$

where C is an absolute constant.

PROOF. For each of the 2^M distributions of signs λ, let $f_\lambda\in A$ be the function f which satisfies (7). Let $t^\lambda\in R^n$ be the value of t which satisfies (5). Then $|f_\lambda(x_\mu)-R(x_\mu, t^\lambda)|<\epsilon$ shows that all values $R(x_\mu, t^\lambda)$ are different from zero and of the same sign as the $f_\lambda(x_\mu)$. It follows that for each pair $\lambda\neq\lambda'$, there is a μ for which $R(x_\mu, t^\lambda)$ and $R(x_\mu, t^{\lambda'})$

are of different sign.

We consider the equation of degree Mq,

$$\prod_{\mu=1}^{M} \tilde{P}(x_\mu, t) = 0,$$

which, by Lemma 7, decomposes R^n into $N \leqq 2(Mq+2)^n$ connected sets Γ_j, $j \leqq N$. On each Γ_j, each $R(x_\mu, t)$ is the quotient of two polynomials,

$$R(x_\mu, t) = P_{j\mu}(t)/Q_{j\mu}(t), \quad \mu = 1, \cdots, M, \quad t \in \Gamma_j.$$

For a fixed j, the equation

$$\prod_{\mu=1}^{M} P_{j\mu}(t) Q_{j\mu}(t) = 0 \tag{9}$$

decomposes R^n into at most $2(2Mp+2)^n$ components, on the interior of each of which the $R(x_\mu, t)$ do not vanish; on their closures the $R(x_\mu, t)$ keep a constant sign. Intersecting these closures with the set Γ_j, we obtain sets $\Delta_{\nu j}$, at most $4(Mq+2)^n(2Mp+2)^n$ in number, with the following property: if $t, t' \in \Delta_{\nu j}$, then $R(x_\mu, t)$ and $R(x_\mu, t')$ are of the same sign for all μ.

It follows that the number of points t^λ cannot exceed the number of the sets $\Delta_{\nu j}$, or that

$$2^M \leqq 4(Mq + 2)^n (2Mp + 2)^n. \tag{10}$$

This implies (8). Our Theorem 11 is essentially equivalent to Vituškin's Theorem 3, p. 186 in [**37**]. For certain representations, however, Vituškin has a better result. He calls a piecewise rational function $R(x, t)$ *continuous* if the denominator $Q_j(x, t)$ of the representation $R(x, t) = P_j(x, t)/Q_j(x, t)$, $t \in \Gamma_j$, does not vanish for $x \in B$, $t \in \Gamma_j$. On the other hand, we have assumed in Theorem 11 only that $Q_j(x, t) \neq 0$ for $x \in B$ and for t actually used in (5).

THEOREM 12 (VITUŠKIN [**37**, THEOREM 1, p. 182]). *If $R(x, t)$ is a continuous piecewise rational ϵ-approximation of the set A with barrier independent of x, then*

$$n \log (p + q + 1) \geqq C_1 M, \tag{11}$$

where C_1 is an absolute constant.

We return to some applications of Theorem 11. Let $A = \Lambda_\omega(B, c)$ be the class of continuous functions defined in §5. Then, if Λ has an ϵ-approximation $R(x, t)$,

$$n \log (p + q + M) \geqq \text{Const. } H_{4\epsilon}(\Lambda). \tag{12}$$

To prove this, we note that we can take M equal to the maximal number M_δ, $\delta = \omega^{-1}(2\epsilon)$ of points of B at distances $\geqq \delta$ from each other. By [**19**, p. 77, (234)],

$$M_{\frac{1}{2}\omega^{-1}(\frac{1}{2}\epsilon)} \geqq H_\epsilon(\Lambda) - \log \frac{C}{\epsilon} \cdot$$

Therefore

$$M = M_\delta \geqq H_{4\epsilon}(\Lambda) - \log \frac{C}{\epsilon} \geqq \frac{1}{3} H_{4\epsilon}(\Lambda),$$

as $H_\epsilon(\Lambda) \geqq 2 \log (1/\epsilon)$, since B has at least two points. Thus (12) follows from Theorem 11.

If we assume that moreover

$$H_\epsilon(\Lambda) \leqq C\epsilon^{-C_1}, \tag{13}$$

then from the inequality [**19**, p. 77, (235)] we have

$$M_{2\omega^{-1}(2\epsilon)} \leqq \frac{1}{\log 2} H_\epsilon(\Lambda),$$

and using the properties of a modulus of continuity and (13),

$$M = M_{\omega^{-1}(2\epsilon)} \leqq M_{2\omega^{-1}(\epsilon)} \leqq \text{Const. } \epsilon^{-C_1},$$

so that (12) implies

$$n \log \left(p + q + \frac{1}{\epsilon}\right) \geqq \text{Const. } H_{4\epsilon}(\Lambda), \tag{14}$$

which is equivalent to Theorem 2 [**37**, p. 195] of Vituškin.

Similarly, let $\Lambda = \Lambda^s_{r,\omega}(S;\, c)$, where $r \geqq 1$ and S is an s-dimensional parallelepiped. Then (Vituškin [**37**, Theorem 2, p. 191]),

$$n \log \left(p + q + \frac{1}{\epsilon}\right) \geqq \text{Const. } H_\epsilon(\Lambda). \tag{15}$$

It follows from [**24**, p. 136] that one can take $M \approx \delta^{-s}$, if δ is defined by $\delta^r\omega(\delta) = \epsilon$. Also, it is known that for some constants K, L, β, γ [**24**, p. 153],

$$K\delta(\beta\epsilon)^{-s} \leqq H_\epsilon(\Lambda) \leqq L\delta(\gamma\epsilon)^{-s}.$$

It follows from this and the properties of the modulus of continuity that $H_\epsilon(\Lambda) \approx \delta(\epsilon)^{-s} \approx M$, so that (15) is a consequence of Theorem 11.

REMARK. Let $P(t)$, $t=(t_1, \cdots, t_n)$ be a product of p linear forms in t. In this case, Shapiro [**31**] found a better upper bound than that given in Lemma 7 for the number N of components of R^n. He showed that $N \leqq p^n/(n-1)!$ if $n \leqq p$. This has the following consequence. Let $R(x, t)=P(x, t)/Q(x, t)$ be a rational ϵ-approximation of Λ of degree 1 without barriers. Then

$$(16) \qquad n \geqq CM, \quad C > 0.$$

For the proof consider the product of $2M$ linear factors

$$P(t) = \prod_{\mu=1}^{M} P(x_\mu, t)Q(x_\mu, t).$$

We can assume that $n<2M$. Then we can use the estimate given above, and obtain, as in the proof of Theorem 11,

$$(2M)^n/(n-1)! \geqq 2^M.$$

This, after a simple computation, gives (16).

It should be noted, that Shapiro's inequality (16) does not follow from Vituškin's Theorem 12, for $p=1$, $q=0$, since Vituškin assumes that the denominators $Q(x, t)$ do not vanish.

7. **Vituškin's theorems for Banach spaces.** Results similar to Theorem 11 hold also for sets A in Banach spaces X. Of course, we must now replace rational expressions by polynomial expressions, since division, in general, has no meaning in X. The method of §7.1 is close to that of §6 and is valid in L^p-spaces. The method of §7.2 is somewhat different. It applies to sets $A(\Delta)$ in arbitrary spaces X, if the δ_n tend to zero sufficiently rapidly.

7.1. *Approximation in L^p-norm.* Let A be a subset of the space L^p of functions $f(x)$ defined on an s-dimensional parallelepiped S. Let $\tilde{P}(t)$ be a polynomial of degree q in $t=(t_1, \cdots, t_n)$, then $\tilde{P}(t)=0$ decomposes the space R^n into $N \leqq 2(q+2)^n$ connected sets Γ_j. For $x \in B$, $t \in \Gamma_j$ let $R(x, t)=P_j(x, t)$ be a polynomial in t of degree p. The piecewise polynomial function $R(x, t)$ is an ϵ-approximation of A if for each $f \in A$ there is a $t \in R^n$ with the property

$$(1) \qquad \int_S |f(x) - R(x, t)|^p dx < \epsilon^p.$$

We shall use a method which appears in [**21**].

LEMMA 8. *Let Q be the set of all 2^M possible distributions of signs $\lambda=\{\lambda_\mu\}$, $\mu=1, \cdots, M$, $\lambda_\mu=\pm 1$, and Q' a subset of Q which consists of not more than ρ^μ* elements where $\rho<\frac{1}{3}2^{5/3}$. *Then for all large M, there* *is at least one distribution of signs* λ *which is different from each* $\lambda' \in Q'$ *in at least* $M/3$ *places.*

Proof. Let $N=[M/3]+1$. We count the number of all λ which can be obtained from a given λ' by changes in at most N places. This number T does not exceed

$$T \leqq 1 + M + \frac{M(M-1)}{2} + \cdots + \binom{M}{N} \leqq (N+1)\frac{M!}{N!(M-N)!} \cdot$$

By Stirling's formula, the right-hand side is asymptotically $\sim M \log (3/2^{2/3})$, so that $T<(2/\rho)^M$ for large M. The number of λ obtainable in this way from *some* $\lambda' \in Q'$, does not exceed $T\rho^M < 2^M$, which completes the proof.

Theorem 13. *Let* $\Lambda = \Lambda^s_{r\omega}(S, c)$, $r \geqq 1$, *be the subset of* $C[S]$, *and let* $R(x, t)$, *be a piecewise polynomial function of degree* p *with barrier* $\tilde{P}(t)=0$ *of degree* q, *which is an* ϵ*-approximation of* Λ *in the* L^1*-norm on* S. *Then*

$$n \log \left(p + q + \frac{1}{\epsilon}\right) \geqq \text{Const. } H_\epsilon(\Lambda). \tag{2}$$

We assume here that $r \geqq 1$ or that $r=0$ and $\omega(t)=t^\alpha$, $0<\alpha\leqq 1$; the entropy in (2) is in the uniform norm.

Proof. For each small $\delta>0$, S contains $M \approx \delta^{-s}$ disjoint balls U_μ of radius δ. Let δ be defined by $\delta^r\omega(\delta)=a\epsilon$; $a>0$. Then we can find functions f_λ, $\lambda=\{\lambda_\mu\}$, $\lambda_\mu = \pm 1$ with

$$\left| \int_{U_\mu} f\, dx \right| \geqq C\delta^s a\epsilon \geqq C_1 \frac{a\epsilon}{M}, \quad \operatorname*{sign}_{x \in U_\mu} f(x) = \lambda_\mu. \tag{3}$$

(See [24, p. 136].) We have

$$\int_S |f(x) - R(x, t)|\, dx \geqq \sum_{\mu=1}^{M} \left| \int_{U_\mu} f\, dx - \int_{U_\mu} R(x, t) dx \right|. \tag{4}$$

For $j=1, \cdots, N$, $\mu=1, \cdots, M$, we put

$$\pi_{\mu j}(t) = \int_U P_j(x, t) dx = \int_{U_\mu} R(x, t) dx. \tag{5}$$

The equation

$$\pi_j(t) = \prod_{\mu=1}^{M} \pi_{\mu j}(t) = 0$$

decomposes Γ_j into at most $2(Mp+2)^n$ sets Δ_{lj}, on each of which the

$\pi_{\mu j}(t)$ are of constant sign. There are $T \leqq 4(Mp+2)^n(q+2)^n$ sets Δ_{lj}, and on each of them, each integral $\int_{U_\mu} R(x, t)dx$ is of constant sign. We show that $T > \rho^M$. Assume that $T \leqq \rho^M$, then by Lemma 8, there is a distribution of signs $\lambda = \{\lambda_\mu\}$ which is different from any distribution of signs (5) in at least $\frac{1}{3}M$ places. Then for the corresponding f_λ, by (3) and (4),

$$\int_S |f_\lambda(x) - R(x, t)| dx \geqq \frac{1}{3} MC_1 \frac{a\epsilon}{M} > \epsilon, \quad \text{for each } t,$$

if $a > 0$ is selected properly. Hence $T > \rho^M$, and we obtain

$$n \log (M + p + q) \leqq \text{Const. } M.$$

Because of the expressions for δ and M, and of the formula [**24**, p. 153] for the entropy of Λ, this is equivalent to (2).

The formulation of Theorem 12 was restricted to $p = 1$ for the approximation and to $p = +\infty$ for the entropy because a similar statement for other values of the p's is a simple corollary.

COROLLARY. *If R is a piecewise polynomial ϵ-approximation of $\Lambda^{ps}_{r\omega}(S, c)$ in the L^{p_1}-norm, then under the same conditions on r and ω as in Theorem* 13, *we have relation* (2), *with $H_\epsilon(\Lambda)$ in L^{p_1}-norm.*

For small $c_1 > 0$, $\Lambda^{s}_{r\omega_1}(S, c_1c) \subset \Lambda^{ps}_{r\omega}(S, c)$, where $\omega_1 = c_1\omega$; also, $\|f - P\|_{L^{p_1}} < \epsilon$ implies $\|f - P\|_{L^1} < c_1^{-1}\epsilon$. Therefore our statement follows from Theorem 13 and the formula for $H_\epsilon(\Lambda^{ps}_{r\omega})$ in §5.

7.2. *Classes $A(\Delta)$ with rapidly decreasing δ_n.* If δ_n decreases to zero rapidly, results better than those of §7.1 can be obtained. For linear approximation, we could observe this phenomenon in §4, compare 4(3) and 4(4). However, while 4(4) was true for the δ_n which satisfy 3(15) (this follows from Theorem 3 and 3(16)), we have now to assume more (see condition (7) below). The results of this section should be compared with Vituškin's theorem [**37**, Theorem 1, p. 177] about analytic functions. The advantages of our results as compared to Vituškin's are the following: we treat arbitrary classes $A(\Delta, \Phi)$ (in arbitrary Banach spaces), which are described only in terms of the behavior of δ_n, while Vituškin uses structural properties of the class A, in a space of continuous functions. He is able later to check these properties for classes of analytic functions. On the other hand, Vituškin can treat rational, instead of polynomial approximation.

THEOREM 14. *Let $P(t)$ be a polynomial ϵ-approximation for $A(\Delta, \Phi)$ in X, of degree p in $t = (t_1, \cdots, t_n)$. Assume that the numbers δ_k satisfy*

$$\frac{\delta_{[n+\sigma n/\log n]}}{\delta_n} \to 0 \quad \textit{for each } \sigma > 0. \tag{6}$$

Then

$$n \log\left(p + \frac{1}{\epsilon}\right) \geqq (1 - o(1))H_\epsilon(A), \quad \epsilon \to 0. \tag{7}$$

PROOF. Let the N_i be defined by 3(1) with $C=e$. We note some consequences of the assumption (6). It follows from 3(2) that $\delta_{N_{i+1}} \geqq e^{-2}\delta_{N_i}$, hence (6) implies $N_{i+1} \leqq N_i + \sigma N_i/\log N_i$, for each $\sigma > 0$, therefore

$$\Delta N_i \log \Delta N_i = o(N_i), \quad i \to \infty. \tag{8}$$

Next, condition (6) implies 3(15), and we derive from 3(20) that

$$H_{a\epsilon}(A)/H_\epsilon(A) \to 1, \quad \epsilon \to 0, \quad \text{for each } a > 0. \tag{9}$$

Easy calculations show that (6) implies

$$\delta_n n^\alpha \to 0 \quad \text{for each } \alpha > 0. \tag{10}$$

Let $\epsilon > 0$ be given, we define j by 3(3) and put $N = N_j - 1$. Without changing the set $A(\Delta)$, we can replace the functions $\phi_1, \cdots, \phi_N$ by any linearly independent set of their linear combinations. Using also a standard result about finite dimensional Banach spaces (the lemma of Auerbach, see for example [**35**, p. 393]) we can assume the existence of linear functionals $L_1, \cdots, L_N$ on X_N, that together with the ϕ_k satisfy

$$\begin{aligned} &\|L_k\| = \|\phi_k\| = 1, \\ &L_k(\phi_l) = 0,\ k \neq l,\ L_k(\phi_k) = 1, \quad k, l = 1, \cdots, N. \end{aligned} \tag{11}$$

The set $B \subset X$ consists of all $g \in X$ of the form

$$g = 3\epsilon \sum_{i=0}^{j-1} \sum_{N_i \leqq k < N_{i+1}} \lambda_k \phi_k, \tag{12}$$

where the λ_k are integers which satisfy

$$|\lambda_k| \leqq \frac{1}{3\epsilon \Delta N_i} e^{-i-2}, \quad N_i \leqq k < N_{i+1}, \quad i = 0, \cdots, j-1. \tag{13}$$

For g of the form (12), $L_k(g) = 3\epsilon t_k$. We note some properties of the set B:

(a) If $g, g' \in B, g \neq g'$, then $|L_k(g) - L_k(g')| \geqq 3\epsilon$ for at least one k.
(b) We have $B \subset A(\Delta)$. This follows from the inequality

$$\rho(g, X_k) \leqq \left\| 3\epsilon \sum_{k<l\leqq N} \lambda_l \phi_l \right\| \leqq 3\epsilon \sum_{i=i_0}^{j-1} \sum_{N_i \leqq l < N_i+1} |\lambda_l|$$

$$\leqq \sum_{i=i_0}^{j-1} e^{-j-2} \leqq e^{-i_0-1} \leqq \delta_k, \quad k = 0, \cdots, N,$$

where i_0 is defined by $N_{i_0} \leqq k < N_{i_0+1}$.

(c) The number M of elements of B satisfies

(14) $$\log M \geqq (1 - o(1))H_\epsilon(A(\Delta)), \quad \epsilon \to 0.$$

Since $\epsilon \geqq e^{-j+1}$, the number of integers λ_k which satisfy (13) is $\geqq e^{j-i-3}/(3\Delta N_i)$, and the number of all sets $(\lambda_1, \cdots, \lambda_N)$ is $\geqq \prod_{i=0}^{j-3} (e^{j-i-3}/3\Delta N_i)^{\Delta N_i}$, therefore

$$\log M \geqq \sum_{i=1}^{j-3} \Delta N_i(j - i - 3) - \sum_{i=1}^{j-3} \Delta N_i(\log 3\Delta N_i).$$

The first sum is $= N_1 + \cdots + N_{j-3} \sim H_\epsilon(A(\Delta))$ by 3(20); the second sum is $o(1)H_\epsilon(A(\Delta))$ by (8), 3(16) and 3(20).

LEMMA 9. *Let $\pi_1, \cdots, \pi_N$ be polynomials of degree p in $t = (t_1, \cdots, t_n)$. Let $\epsilon > 0$ and let t^μ, $\mu = 1, \cdots, M$ be points of R^n with the properties that*

(15) $$|\pi_k(t^\mu) - \pi_k(t^{\mu'})| > \epsilon \text{ for at least one } k \text{ if } \mu \neq \mu':$$

(16) $$|\pi_k(t^\mu)| \leqq M_k, \quad k = 0, \cdots, M.$$

Then

(17) $$2\left[p \sum_{k=1}^{N} \left(\frac{2M_k}{\epsilon} + 1\right) + 2\right]^n \geqq M.$$

PROOF. We can assume that none of the polynomials π_k is constant. Let $\pi(t)$ be defined by

(18) $$\pi(t) = \prod_{k=1}^{N} \prod_{|s| \leqq M_k/\epsilon} \{\pi_k(t) - s\epsilon\}.$$

The degree of π does not exceed $\sum_{k=1}^{N} (2M_k/\epsilon + 1)p$, hence, by Lemma 7, the space R^n can be decomposed into at most

$$M_0 = 2\left[p \sum_{k=1}^{n} (2M_k/\epsilon + 1) + 2\right]^n$$

connected sets Γ_j, on each of which π does not change sign. In the interior of each Γ_j, π does not vanish, hence each π_k satisfies there an inequality

$$s'\epsilon \leqq \pi_k(t) \leqq (s'+1)\epsilon.$$

The same inequality holds on the whole of Γ_j by continuity. It follows that no two points t^μ, $t^{\mu'}$, $\mu \neq \mu'$, can belong to the same Γ_j. Hence $M_0 \geqq M$.

PROOF OF THEOREM 14. We consider the set $B \subset A(\Delta)$ defined by (12) and (13). The number M of points $g_\mu \in B$ satisfies (14).

Putting $\pi_k(t) = L_k(P(t))$, $k=1, \cdots, N$, we obtain polynomials of degree p in $t=(t_1, \cdots, t_n)$. Since $P(t)$ is an ϵ-approximation of $A(\Delta)$, for each $g_\mu \in B$, $\mu=1, \cdots, M$ there is a $t^\mu \in R^n$ with $\|g_\mu - P(t^\mu)\| \leqq \epsilon$. Since $L_k(g_\mu) = 3\epsilon\lambda_k^\mu$, where λ_k^μ are the coordinates of g_μ in (12), we have

$$(19) \qquad |3\epsilon\lambda_k^\mu - \pi_k(t^\mu)| < \epsilon, \quad k = 1, \cdots, N, \quad \mu = 1, \cdots, M.$$

Let $\mu \neq \mu'$. Since the λ_k^μ are integers, it follows from (19) that

$$|\pi_k(t^\mu) - \pi_k(t^{\mu'})| > \epsilon$$

must hold for at least one k. To estimate the $\pi_k(t^\mu)$, we note that (19) and (13) imply

$$(20) \quad |\pi_k(t^\mu)| \leqq \epsilon + 3\epsilon|\lambda_k^\mu| \leqq \epsilon + e^{-i-2}/\Delta N_i, \quad \text{if} \quad N_i \leqq k < N_{i+1}.$$

If we denote the term on the right by M_k, then by 3(3), $M_k/\epsilon \leqq 1+e^{j-i-3}/\Delta N_i$. Applying Lemma 3 and estimating the sum in (17) we obtain

$$(21) \qquad 2(3pN_j + pe^{j-2} + 2)^n \geqq M.$$

By (10), $N_j\delta_{N_j} \to 0$, hence $N_je^{-j} \to 0$, and from (21) and (14) it is easy to derive (7).

REMARK. A similar theorem is valid for a piecewise polynomial ϵ-approximation of $A(\Delta)$ with a barrier $\tilde{P}(t)=0$ of degree q. Then one has, instead of (7),

$$(22) \qquad n \log (p + q + 1/\epsilon) \geqq (1 - o(1))H_\epsilon(A).$$

We omit the proof, which is based on ideas used for the derivation of the last two theorems; the polynomials (5) are now to be replaced by $\pi_{kj}(t) = L_k(P_j(t))$.

BIBLIOGRAPHY

1. N. I. Achiezer, *Theory of approximation*, F. Ungar, New York, 1956.

2. S. Ja. Al'per, *On ε-entropy of certain classes of functions*, Dokl. Akad. Nauk SSSR **132** (1960), 977–979 = Soviet Math Dokl. 1 (1960), 667–669.

3. T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Ergebnisse der Math., Bd. 3, no. 1, Springer, Berlin 1934.

4. Ju. A. Brudnyĭ and B. D. Kotljar, *The order of growth of ϵ-entropy on certain compact classes of functions*, Dokl. Akad. Nauk SSSR **148** (1963), 1001–1004 = Soviet Math. Dokl. **4** (1963), 196–199.

5. Ju. A. Brudnyĭ and A. F. Timan, *Constructive characteristics of compact sets in Banach spaces and ϵ-entropy*, Dokl. Akad. Nauk SSSR **126** (1959), 927–930.

6. G. F. Clements, *Entropies of sets of functions of bounded variation*, Canad. J. Math. **15** (1963), 422–432.

7. ———, *Entropies of several sets of real valued functions*, Pacific J. Math. **13** (1963), 1085–1095.

8. Ph. J. Davis, *Interpolation and approximation*, Blaisdell, New York, 1963.

9. V. K. Dzjadyk, *On the problem of S. M. Nikol'skiĭ*, Izv. Akad. Nauk SSSR, Ser. Mat. **23** (1959), 697–736.

10. ———, *To the theory of approximation of analytic functions*, Izv. Akad. Nauk SSSR, Ser. Mat. **27** (1963), 1135–1164; English transl., Amer. Math. Soc. Transl. (2) **53** (1966), 253–284.

11. V. D. Erohin, *Conform mappings of rings and the fundamental basis of the space of functions that are analytic in the elementary neighborhood of an arbitrary continuum*, Dokl. Akad. Nauk SSSR **120** (1958), 689–692.

12. ———, *The asymptotic behavior of ϵ-entropy of analytic functions*, Dokl. Akad. Nauk SSSR **120** (1958), 949–952.

13. K. K. Golovkin, *The ϵ-entropy of some compact sets of differentiable functions in spaces with a monotone norm*, Dokl. Akad. Nauk SSSR **158** (1964), 261–263 = Soviet Math. Dokl. **5** (1964), 1209–1211.

14. A. Ja. Helemskiĭ and G. M. Henkin, *Imbeddings of compacts into ellipsoids*, Vestnik Moskov. Univ., Ser. 1, Mat. Meh, no. 2 (1963), 3–12.

15. G. M. Henkin, *Linear superpositions of continuously differentiable functions*, Dokl. Akad. Nauk SSSR **157** (1964), 288–290 = Soviet Math. Dokl. **5** (1964), 948–950.

16. A. N. Kolmogorov, *Asymptotic characteristics of some completely bounded metric spaces*, Dokl. Akad. Nauk SSSR **108** (1956), 585–589.

17. ———, *On the representation of continuous functions of several variables*, Dokl. Akad. Nauk SSSR **114** (1957), 953–956.

18. ———, *Linear dimension of topological vector spaces*, Dokl. Akad. Nauk SSSR **120** (1958), 239–241.

19. A. N. Kolmogorov and V. M. Tihomirov, *ϵ-entropy and ϵ-capacity of sets in function spaces*, Uspehi Mat. Nauk no. 2 (86), **14** (1959), 3–86; English transl., Amer. Math. Soc. Transl. (2) **17** (1961), 277–364.

20. B. D. Kotljar, *The order of growth of ϵ-entropy on classes of quasi-smooth functions*, Uspehi Mat. Nauk no. 2 (110), **18** (1963), 135–138.

21. G. G. Lorentz, *Lower bounds for the degree of approximation*, Trans. Amer. Math. Soc. **97** (1960), 25–34.

22. ———, *Metric entropy, widths, and superpositions of functions*, Amer. Math. Monthly **69** (1962), 469–485.

23. ———, *Entropy and its applications*, J. Soc. Indust. Appl. Math. Ser. B, Numer. Anal. **1** (1964), 97–103.

24. ———, *Approximation of functions*, Holt, Rinehart, Winston, New York 1966.

25. B. S. Mitjagin, *The approximative dimension and bases in nuclear spaces*, Uspehi Mat. Nauk no. 4 (100), **16** (1961), 63–132.

26. O. A. Oleĭnik, *Estimations of Betti numbers of real algebraic hypersurfaces*, Mat. Sb. **28** (**70**) (1951), 635–640.

27. A. M. Olevskiĭ, *On a problem of P. L. Ul'janov*, Uspehi Mat. Nauk no. 2 (122), **20** (1965), 197–202.

28. A. Pelczynski, *On the approximation of S-spaces by finite dimensional spaces*, Bull. Acad. Polon. Sci. Cl. III **5** (1957), 879–881.

29. B. Penkov and Bl. Sendov, *Entropy of the set of continuous functions of several variables*, C. R. Acad. Bulgare Sci. **17** (1964), 335–337.

30. S. Rolewicz, *On spaces of holomorphic functions*, Studia Math. **21** (1962), 135–160.

31. H. S. Shapiro, *Some negative theorems of approximation theory*, Michigan Math. J. **11** (1964), 211–217.

32. S. A. Smoljak, *The ϵ-entropy of classes $E_s^{\alpha,k}(B)$ and $W_s^{\alpha}(B)$ in the L^2-metric*, Dokl. Akad. Nauk SSSR **131** (1960), 30–33 = Soviet Math. Dokl. **1** (1960), 192–195,

33. V. M. Tihomirov, *Widths of sets in functional spaces and the theory of best approximations*, Uspehi Mat. Nauk no. 3 (93), **15** (1960), 81–120.

34. V. M. Tihomirov, *The ϵ-entropy of some classes of periodic functions*, Uspehi Mat. Nauk no. 6 (108), **17** (1962), 163–169.

35. A. F. Timan, *Theory of approximation of functions of a real variable*, Macmillan, New York, 1963.

36. ———, *The order of growth of ϵ-entropy of spaces of real continuous functionals, defined on a connected compactum*, Uspehi Mat. Nauk no. 1 (115), **19** (1964), 173–177.

37. A. G. Vituškin, *Theory of transmission and processing of information*, Pergamon Press, New York, 1961.

38. ———, *Some properties of linear superpositions of smooth functions*, Dokl. Akad. Nauk SSSR **156** (1964), 1003–1006 = Soviet Math Dokl. **5** (1964), 741–744.

39. ———, *Proof of existence of analytic functions of several variables, not representable by linear superpositions of continuously differentiable functions of fewer variables*, Dokl. Akad. Nauk SSSR **156** (1964), 1258–1261 = Soviet Math. Dokl. **5** (1964), 793–796.

40. A. Vosburg, *Metric entropy of certain classes of Lipschitz functions*, Proc. Amer. Math Soc. **17** (1966), 665–669.

ACKNOWLEDGMENT

The preparation of the following two papers, which appeared in the May, 1966 issue of the BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY, was supported by the National Science Foundation under grant NSF G-16428 to the American Mathematical Society.

Whitehead torsion, by J. Milnor
Singular integrals, by A. P. Calderón

This information was inadvertently omitted from the BULLETIN.