*Siberian Mathematical Journal, Vol.* 38, *No.* 4, 1997

# $\varepsilon$-ENTROPY OF COMPACT SETS IN $C$ AND TABULATION OF CONTINUOUS FUNCTIONS[†)]

**V. N. Potapov**

UDC 517.52

## § 1. Introduction

In the present article we consider two interrelated problems. The first of them consists in estimating the $\varepsilon$-entropy of a set of uniformly continuous functions given on an arbitrary compact metric space. The second consists in constructing convenient tables for saving continuous and differentiable functions.

The concept of $\varepsilon$-entropy was introduced by A. N. Kolmogorov: the $\varepsilon$-entropy of a totally bounded space $X$ is defined to be the quantity $H_X(\varepsilon)$ equal to the logarithm of the minimal number of sets of diameter at most $2\varepsilon$ needed to cover $X$. Here and in what follows, the logarithm is to the base 2. $\varepsilon$-Entropy is an important characteristic of a compact space which reflects massiveness of the space. Therefore, there are many articles devoted to studying $\varepsilon$-entropy.

Let $F_\omega$ stand for the set of $\omega$-continuous functions from a compact space $X$ into a compact space $Y$, with $\omega$ some continuity modulus. The following estimates for $H_{F_\omega}(\varepsilon)$ were established under the assumption that $X$ is a connected compact space and $Y$ is an interval in $\mathbb{R}$: In [1] A. N. Kolmogorov and V. M. Tikhomirov proved that the $\varepsilon$-entropy of the space of functions satisfying the Hölder condition with exponent $\alpha$ is asymptotically $O(2^{H_X(\varepsilon^{1/\alpha})})$ (as $\varepsilon \to 0$). In [2] A. G. Vitushkin generalized this estimate for the space $F_\omega$ with an arbitrary continuity modulus $\omega$. It turned out that $2^{H_X(4\delta)} \le H_{F_\omega}(\omega(\delta)/4)$ and $2^{H_X(\delta/2)} \ge H_{F_\omega}(2\omega(\delta))$ as $\delta \to 0$. In [3, 4] A. F. Timan made Vitushkin's estimate more precise under extra conditions, having proved existence of constants $C_0$ and $C_1$ such that

$$C_0 2^{H_X(\delta)} \le H_{F_\omega}(\omega(\delta)/2) \le C_1 2^{H_X(\delta)}.$$

Moreover, A. F. Timan observed that the connectedness of $X$ is not a necessary condition for the $\varepsilon$-entropy of a function class to be asymptotically independent of the target space. He proved that if $X$ is a subset of $\mathbb{R}^k$ of nonzero Lebesgue measure then the preceding inequalities remain valid.

In the present article we give a necessary and sufficient condition on $X$ for the $\varepsilon$-entropy of the function space $F_\omega$ to be asymptotically $O(2^{H_X(\omega^{-1}(\varepsilon))})$. In [1–4] the above-mentioned estimates for the $\varepsilon$-entropy were obtained only for the spaces of $\omega$-continuous functions given on connected compact spaces or on compact spaces of integral metric dimension. Here we derive the same estimates for the spaces over disconnected spaces of fractional or infinite dimension. In particular, the above estimates for the $\varepsilon$-entropy of the space of $\omega$-continuous functions on $X$ are valid in the case when $X$ is the Cantor set.

The second question we discuss in the article is the tabulation problem formulated as follows: for each function in some function class $F$, construct a table (a code for the function) with the possibly least size of the table and the possibly least complexity of decoding, i.e., the number of operations needed for reconstructing an approximate value of the function at an arbitrary point. The set of such tables, approximating the functions in the function space $F$ to within $\varepsilon > 0$, gives rise to a covering of $F$ by sets of diameter at most $2\varepsilon$. For this reason, the number of distinct tables $\varepsilon$-approximating the functions of $F$ cannot be less than $2^{H_F(\varepsilon)}$ and the maximum of the sizes of these tables (in bits) cannot be less than $H_F(\varepsilon)$.

[†)] The research was supported by the Russian Foundation for Basic Research (Grant 95-01-01343).

Novosibirsk. Translated from *Sibirskiĭ Matematicheskiĭ Zhurnal*, Vol. 38, No. 4, pp. 876–891, July–August, 1997. Original article submitted January 29, 1996.

0037-4466/97/3804-0759 $18.00 

In [1], A. N. Kolmogorov and V. M. Tikhomirov obtained an asymptotically exact estimate for the $\varepsilon$-entropy of the function space $F^k_{n,\alpha}$ comprising the real-valued functions given on $\mathbb{R}^k$ and having the $n$th order derivative satisfying the Hölder condition with exponent $\alpha$: $H_{F^k_{n,\alpha}}(\varepsilon) = O(\varepsilon^{-k/(n+\alpha)})$. Here $n \geq 0$ and $k \geq 1$ are integers and $0 < \alpha \leq 1$. The construction of an $\varepsilon$-net for the space $F^k_{n,\alpha}$ proposed by them makes it possible to construct decoding tables of minimal size for differentiable functions; however, the complexity of decoding such a table is of the same order (as $\varepsilon \to 0$) as the length of the table. The tabulation method proposed by E. A. Asarin in [5] for smooth functions makes it possible to construct tables of asymptotically minimal size such that decoding of the values of a function at an arbitrary point requires $O(\log 1/\varepsilon)$ arithmetic operations. The method proposed by the author in [6] enables us to restore the values of a function from a table by finitely many arithmetic operations on preserving the minimal size of the table. However, this method applies suits only to the functions given on parallelepipeds of $\mathbb{R}^k$.

In the present article, we describe some method for tabulating the differentiable functions over an open bounded set in $\mathbb{R}^k$ which enables us to construct $\varepsilon$-approximating tables of minimal size and with the complexity of decoding $O(\log^* 1/\varepsilon)$ times as large as the minimal complexity. Here $\log^*$ is the iterative logarithm, an extremely slowly growing function whose precise definition is given below. We also propose an analogous tabulation method for continuous functions defined on wider classes of sets. Thus, the tabulation methods we propose in the article are not worse than the known methods as regards the size of a table and, at the same time, they have either less complexity of decoding or a larger scope of applicability.

The simplest way for tabulating $\omega$-continuous functions results from the construction of an $\varepsilon$-net for the set of uniformly continuous functions which is used in the proof of the Arzelà Theorem (see, for instance, [7]). Namely, let $F_\omega$ be the space of $\omega$-continuous functions from a compact space $X$ into a compact space $Y$. For every $\delta > 0$, there exist a finite covering $R$ of $X$ by sets of diameter at most $\delta$ and a finite $\omega(\delta)$-net $S$ for $Y$. Denote the sizes of $R$ and $S$ by $|R|$ and $|S|$. Given a function $f$ in $F_\omega$, there is a function $f_\delta$ constant over every element of the covering $R$, taking values in $S$, and approximating the initial function to within $2\omega(\delta)$.. We have thus constructed a covering of $F_\omega$ by sets of diameter at most $4\omega(\delta)$. The number of piecewise constant functions generating this covering equals $|R|^{|S|}$. The *complexity* of a function is defined as the minimum of the sizes (in bits) of the tables specifying the function (see, for instance, [8]). In each function class $F$ comprising finitely many functions, there is a function whose complexity is not less than $\log|F|$. Therefore, for some function $f \in F_\omega$ the complexity of the function $f_\delta$ is not less than $|R|\log|S|$. Writing down the values of $f_\delta$ on the elements of the covering $R$, we obtain a $2\omega(\delta)$-approximating table, for the $\omega$-continuous function $f$, whose size equals $|R|\log|S|$; i.e., the complexity of an arbitrary piecewise constant function $f_\delta$ does not exceed $|R|\log|S|$. The complexity of reconstructing the value of the function at a point from such a table is minimal; however, the size of the table cannot be less than $H_Y(\omega(\delta))2^{H_X(\delta/2)}$, which considerably exceeds the estimate for the $\varepsilon$-entropy of the set $F_\omega$: $H_{F_\omega}(2\omega(\delta)) = O(2^{H_X(\delta/2)})$.

The main idea behind the tabulation method we propose here consists in successively approximating an $\omega$-continuous function by piecewise constant functions with increasing accuracy. Having approximated an arbitrary $\omega$-continuous function with a prescribed accuracy $\varepsilon$ by the sum of finitely many piecewise constant functions whose total complexity does not exceed some number $H$, we thereby obtain a table of the $\omega$-continuous function having size $H$ and representing the union of the tables of these piecewise constant functions. The complexity of decoding such a table is directly proportional to the number of piecewise constant functions in the approximating sum. Moreover, as was observed above, the maximum of the sizes of the tables of the functions, i.e., the number $H$, represents an upper bound for the $\varepsilon$-entropy of the function space $F_\omega$. We show in Theorem 1 that the so-obtained estimate for $\varepsilon$-entropy generalizes the results of [2–5]. In Theorem 2 we reveal the extent to which the number of summands in the approximating sum may be reduced on preserving the asymptotic minimality of the size of the table. Likewise, in Theorem 3 we construct tables of differentiable functions by using piecewise polynomial functions in place of piecewise constant functions.

We illustrate the main idea of the proposed tabulation method on the set of real-valued Lipschitz continuous functions. First, we define the sequence of the functions $\log^{(i)}(t)$: $\log^{(1)}(t) = \log t$,

$\log^{(i+1)}(t) = \log(\log^{(i)}(t))$. The *iterative logarithm* ($\log^*$) of a number $t$ is the smallest integer $h$ such that $\log^{(h)} t \le 0$. Divide the interval $[0,1]$ into $1/(\varepsilon \log(1/\varepsilon) \log^{(2)}(1/\varepsilon))$ subintervals of equal lengths. Given an arbitrary function $f : [0,1] \to [0,1]$ satisfying the Lipschitz condition with constant 1, we can find a function $f_1$ that is constant over each of the subintervals and approximates $f$ to within $\frac{1}{2}\varepsilon \log(1/\varepsilon) \log^{(2)}(1/\varepsilon)$. The complexity of $f_1$ is at most $(\log(1/\varepsilon) + 1)/(\varepsilon \log(1/\varepsilon) \log^{(2)}(1/\varepsilon))$. The difference between the initial function $f$ and $f_1$ does not exceed $\varepsilon \log(1/\varepsilon) \log^{(2)}(1/\varepsilon)$. Now, divide the interval $[0,1]$ into $1/(\varepsilon \log^{(2)}(1/\varepsilon) \log^{(3)}(1/\varepsilon))$ parts and analogously construct a piecewise constant function $f_2$ that approximates $f - f_1$ to within $\frac{1}{2}\varepsilon \log^{(2)}(1/\varepsilon) \log^{(3)}(1/\varepsilon)$. The complexity of $f_2$ does not exceed $(\log^{(2)}(1/\varepsilon) + 1)/(\varepsilon \log^{(2)}(1/\varepsilon) \log^{(3)}(1/\varepsilon))$. Likewise we can define piecewise constant functions $f_3, f_4, \dots, f_{h-2}$, with $h = \log^* 1/\varepsilon$, so that the complexity of each function $f_i$ does not exceed $\log^{(i)}(1/\varepsilon)/(\varepsilon \log^{(i)}(1/\varepsilon) \log^{(i+1)}(1/\varepsilon))$ and the difference between $f$ and $\sum_{j=1}^{i} f_j$ does not exceed $\varepsilon \log^{(i)}(1/\varepsilon) \log^{(i+1)}(1/\varepsilon)$. We have thus approximated $f$ by the sum of $\log^* 1/\varepsilon - 2$ piecewise constant functions with accuracy $2\varepsilon$; moreover, the total complexity of $f_i$'s is a finite number of times as large as $1/\varepsilon$. As the table of $f$ we can take the union of the tables of the values of the piecewise constant functions $f_1, f_2, \dots, f_{h-2}$.

## § 2. $\varepsilon$-Entropy and Coverings of Totally Bounded Spaces

In this section we state general assertions pertinent to the concept of the $\varepsilon$-entropy of a space.

Let $X$ be a totally bounded metric space; i.e., for every positive $\varepsilon$ there exists a finite covering of $X$ by sets of diameter at most $\varepsilon$. Given $\varepsilon > 0$, let $\mathcal{R}$ stand for the collection of coverings of $X$ by sets of diameter at most $2\varepsilon$. Denote by $N_X(\varepsilon)$ the minimal number of elements in such covering of $X$; i.e., $N_X(\varepsilon) = \min_{R \in \mathcal{R}} |R|$, where $|R|$ is the number of elements (the size) of a covering $R \in \mathcal{R}$. The (*absolute*) *$\varepsilon$-entropy* of a space $X$ is the quantity $H_X(\varepsilon) = \log N_X(\varepsilon)$. The *$\varepsilon$-entropy* of a space $X$ *relative to* an ambient metric space $W$ is the quantity $H_X^W(\varepsilon)$ equal to the logarithm of the size of a minimal $\varepsilon$-net for $X$ in $W$.

A subset $A$ of $X$ is called *$\varepsilon$-distinguishable* if the inequality $\rho_X(x_1, x_2) > \varepsilon$ holds for arbitrary $x_1, x_2 \in A, x_1 \neq x_2$, with $\rho_X$ the distance in $X$. Let $\mathcal{A}$ stand for the collection of all $\varepsilon$-distinguishable subsets of $X$. Denote by $M_X(\varepsilon)$ the size of a maximal $\varepsilon$-distinguishable subset of $X$; i.e., $M_X(\varepsilon) = \max_{A \in \mathcal{A}} |A|$.

Proofs of the following two well-known assertions can be found in [1].

**Assertion 1.** *If $X$ is a totally bounded metric space and if $\varepsilon > 0$ then*

$$N_X(2\varepsilon) \le M_X(2\varepsilon) \le N_X(\varepsilon).$$

Observe that $N_X$ and $M_X$, viewed as functions of $\varepsilon$, are monotone decreasing by definition.

**Assertion 2.** *If $X$ is an open bounded subset of a normed $k$-dimensional vector space and if $\varepsilon > 0$ then $H_X(\varepsilon) = (1 + o(1))k \log(\frac{1}{\varepsilon})$ as $\varepsilon \to 0$.*

Let $V$ be a finite-dimensional normed vector space. Then every bounded subset of $V$, in particular a unit ball, is totally bounded. Define the function $h_V(\varepsilon) : \mathbb{R}^+ \to \mathbb{R}^+$, $h_V(\varepsilon) = H^V_{B(0,1)}(\varepsilon)$, where $B(0,1)$ is the ball in $V$ with radius 1 and center 0. Since $V$ is a normed vector space, the $(r\varepsilon)$-entropy (relative to $V$) of the ball with radius $r$ and center an arbitrary point $x \in V$ equals $h_V(\varepsilon)$; i.e., $H^V_{B(x,r)}(\varepsilon) = h_V(\frac{\varepsilon}{r})$.

The next lemma provides a lower bound for the growth of the $\varepsilon$-entropy of a connected set as $\varepsilon$ decreases. Here and in what follows, by a connected set we mean a set that comprises at least two points and is not representable as the union of two closed (relative to the set) nonempty subsets.

**Lemma 1.** *If $X$ is a connected totally bounded metric space then there exists an $\varepsilon_0 > 0$ such that the inequality*

$$H_X(\varepsilon) \ge 1 + H_X(4\varepsilon)$$

*holds for every $\varepsilon$, $\varepsilon_0/4 > \varepsilon > 0$.*

PROOF. By hypothesis, $X$ comprises at least two points. Let $\varepsilon_0$ be a positive number such that $\varepsilon_0 < \rho_X(x_1, x_2)$ for some $x_1, x_2 \in X, x_1 \neq x_2$. Suppose that $\varepsilon_0 > \varepsilon > 0$. Consider a $4\varepsilon$-distinguishable subset $P$ of $X$ such that $|P| \geq N_X(4\varepsilon)$. The existence of such $P$ is guaranteed by Assertion 1. Let $R$ be a covering of $X$ by sets of diameter $2\varepsilon$ such that $|R| = N_X(\varepsilon)$. We may assume that the elements of $R$ are closed, since the diameter of a set equals that of its closure.

Let $R'$ be the collection of the elements of $R$ containing points of $P$. Since the diameter of each element $r \in R$ does not exceed $2\varepsilon$ and the distance between arbitrary two points of $P$ is certainly greater than $4\varepsilon$, each element $r \in R'$ contains exactly one point of $P$ and the elements of $R'$ are disjoint pairwise. On the other hand, $R$ is a covering of $X$ and so of $P$. Therefore, $|R'| = |P|$.

Given $r \in R \setminus R'$, denote by $s$ the union of $r$ and all elements of $R'$ meeting $r$. Let $S$ denote the collection of all $s$ corresponding to the various $r \in R \setminus R'$. The diameter of each $s \in S$ is at most $6\varepsilon$.

Suppose that there exists $x \in X$ such that $x$ belongs to none of the elements $s \in S$. Then there exists $r' \in R'$ containing $x$ and disjoint from any element of $R \setminus R'$. Since the elements of $R'$ are pairwise disjoint, $X$ can be represented as the union of two closed sets: $r'$ and $\cup r$ over all $r \in R \setminus \{r'\}$. The latter is nonempty by the choice of $\varepsilon$. This fact contradicts connectedness of $X$. Hence, the supposition of existence of $x \in X$ lying in no $s \in S$ is false and so $S$ is a $6\varepsilon$-covering of $X$. It follows that $|S| \geq N_X(3\varepsilon)$. However, $|S| = |R \setminus R'|$ and $|R'| \geq N_X(4\varepsilon)$. Therefore,

$$N_X(\varepsilon) = |R| = |R \setminus R'| + |R'| \geq N_X(3\varepsilon) + N_X(4\varepsilon) \geq 2N_X(4\varepsilon),$$

because $N_X$ is monotone decreasing. Thus, $H_X(\varepsilon) \geq 1 + H_X(4\varepsilon)$ and the proof of Lemma 1 is complete.

Lemma 1 shows that the order of growth of the $\varepsilon$-entropy of a connected space is bounded from below as $\varepsilon$ decreases. We distinguish a special class of totally bounded spaces whose $\varepsilon$-entropy increases steadily and rather quickly as $\varepsilon$ decreases.

DEFINITION. A totally bounded metric space $X$ is a *convex-type space* if there exists a (downward) convex function $f : \mathbb{R}^+ \to \mathbb{R}^+$, $f(0) = 0$, such that $f(t) \asymp N_X(1/t)$ as $t \to \infty$.

Here and below, $f(t) \asymp g(t)$ as $t \to t_0$ means that $f(t) = O(g(t))$ and $g(t) = O(f(t))$ as $t \to t_0$.

Examples of convex-type spaces are an arbitrary convex set in $\mathbb{R}^k$, a set of positive Lebesgue measure in $\mathbb{R}^k$ (see [3]), and the spaces of Lipschitz continuous, differentiable, and analytic functions (see [1]).

**Lemma 2.** *If $X$ is a totally bounded convex-type space then there exist $C > 1$ and $\varepsilon_0 > 0$ such that the inequality*

$$\frac{N_X(\varepsilon_2)}{N_X(\varepsilon_1)} \leq C^2 \frac{\varepsilon_1}{\varepsilon_2}$$

*holds for arbitrary $\varepsilon_1$ and $\varepsilon_2, \varepsilon_1 < \varepsilon_2 < \varepsilon_0$.*

PROOF. For an arbitrary convex function $f : \mathbb{R}^+ \to \mathbb{R}^+$ with $f(0) = 0$ and arbitrary numbers $\varepsilon_2 > \varepsilon_1 > 0$, the following inequality holds:

$$\frac{f(1/\varepsilon_2)}{f(1/\varepsilon_1)} \leq \frac{\varepsilon_1}{\varepsilon_2}. \tag{1}$$

The definition of a convex-type space implies that there are a positive convex function $f, f(0) = 0$, and a constant $C > 1$ such that

$$\frac{f(t)}{C} \leq N_X(1/t) \leq Cf(t) \tag{2}$$

for $t > 1/\varepsilon_0$, with $\varepsilon_0$ some positive number.

From (1) and (2) we infer that the inequalities

$$\frac{N_X(\varepsilon_2)}{N_X(\varepsilon_1)} \leq C^2 \frac{f(1/\varepsilon_2)}{f(1/\varepsilon_1)} \leq C^2 \frac{\varepsilon_1}{\varepsilon_2}$$

hold for $0 < \varepsilon_1 < \varepsilon_2 < \varepsilon_0$. This completes the proof of Lemma 2.

It follows from Lemma 2 that

$$\frac{N_X(2C^2\varepsilon)}{N_X(\varepsilon)} \leq \frac{1}{2}, \quad H_X(2C^2\varepsilon) + 1 \leq H_X(\varepsilon)$$

for $\varepsilon < \varepsilon_0/2C^2$. Thus, the $\varepsilon$-entropy of totally bounded convex-type spaces satisfies a condition similar to that proven in Lemma 1 for connected spaces.

We now introduce some class of totally bounded spaces $X$ which contains convex-type spaces as well as connected spaces.

DEFINITION. A totally bounded metric space $X$ is a *uniform-type space* if the $\varepsilon$-entropy of $X$ enjoys the following property: there are real numbers $\beta > 1$, $C > 0$, and $\varepsilon_0 > 0$ such that the inequality

$$H_X(\beta\varepsilon) + C \leq H_X(\varepsilon)$$

holds for every $\varepsilon < \varepsilon_0$.

The following lemma provides an example of a disconnected uniform-type space. By the Cantor set we mean the subset of the interval $[0, 1]$ which results from dividing the interval into three parts and deleting the middle parts.

**Lemma 3.** *The Cantor set is a uniform-type space.*

PROOF. Denote the Cantor set by $X$. The definition of the Cantor set easily implies that $N_X(\varepsilon) = 1$ for $\varepsilon \geq 1/2$ and $N_X(1/(2 \cdot 3^n)) = 2^n$ for a natural $n$. The last equality and monotonicity of $N_X$ yield $2^n \leq N_X(\varepsilon) \leq 2^{n+1}$ for $1/(2 \cdot 3^{n+1}) \leq \varepsilon \leq 1/(2 \cdot 3^n)$. Therefore,

$$\lfloor \log_3(1/2\varepsilon) \rfloor \leq H_X(\varepsilon) \leq \lfloor \log_3(1/2\varepsilon) \rfloor + 1,$$

where $\lfloor \ \rfloor$ stands for the integral part of a number. Assume that $\varepsilon < 1/27$. Then the preceding inequalities imply that

$$H_X(27\varepsilon) \leq \log_3 \frac{1}{2\varepsilon} - 2 \leq H_X(\varepsilon) - 1;$$

i.e., $X$ is a uniform-type space, which completes the proof of Lemma 3.

A covering $R$ of a set $X$ is said to be an *$\varepsilon$-covering* if the diameter of each elements of $R$ is at most $2\varepsilon$. Suppose that $\varepsilon_1 > \varepsilon_0 > 0$. An $\varepsilon_1$-covering $R_1$ of a set $X$ is said to be a *covering generated by an $\varepsilon_0$-covering* $R_0$ of $X$ if each element of $R_1$ is the union of some elements of $R_0$.

Our next lemma shows that if $\varepsilon_1$ is much larger than $\varepsilon_0$ then for every $\varepsilon_0$-covering of a set $X$ there is a generated $\varepsilon_1$-covering of $X$ whose size is close to the minimal size.

**Lemma 4.** *Let $X$ be a totally bounded metric space, let $R_0$ be an $\varepsilon_0$-covering of $X$, and let $0 < \varepsilon_0 < \varepsilon_1/2$. Then there exists an $\varepsilon_1$-covering $R_1$ of $X$ generated by $R_0$ and such that $|R_1| \leq N_X(\varepsilon_1 - 2\varepsilon_0)$.*

PROOF. Consider an $(\varepsilon_1 - 2\varepsilon_0)$-covering $R$ of $X$ such that $|R| = N_X(\varepsilon_1 - 2\varepsilon_0)$. Enumerate the elements of $R$, $R = \{r_i\}$.

Denote by $s_1$ the union of the elements of $R_0$ meeting $r_1 \in R$. The diameter of $s_1$ does not exceed the sum of the maximal diameter of the elements of $R$ and the doubled maximal diameter of the elements of $R_0$, i.e., $2\varepsilon_1$.

Among the remaining elements of $R_0$, choose the elements that meet $r_2 \in R$ and denote their union by $s_2$. Continuing likewise, construct the finite sequence $s_i$, with $s_i$ representing the union of the elements of $R_0$ meeting $r_i$ but disjoint from any $r_j$ with $j < i$.

Since $R$ is a covering of $X$, each element of $R_0$ meets at least one $r_i$. Therefore, the collection of $s_i$ is a covering of $X$. Denote it by $R_1$. By construction, $R_1$ is an $\varepsilon_1$-covering of $X$ generated by $R_0$; furthermore, $|R_1| \leq |R| \leq N_X(\varepsilon_1 - 2\varepsilon_0)$, which completes the proof of Lemma 4.

It is seen from the proof of Lemma 4 that if the elements of $R_0$ are pairwise disjoint then we can construct a covering $R_1$ generated by $R_0$ so that all elements of $R_1$ be pairwise disjoint. Henceforth we consider only coverings with disjoint elements.

## § 3. ε-Entropy of Compact Sets in $C$

Let $X$ and $Y$ be totally bounded metric spaces with respective metrics $\rho_X$ and $\rho_Y$. Denote by $C(X,Y)$ the space composed of continuous functions from $X$ into $Y$ and endowed with the metric

$$\rho_C(f,g) = \sup_{x\in X} \rho_Y(f(x), g(x)).$$

If $f \in C(X,Y)$ then the *continuity modulus* of $f$ is the function

$$\omega_f(t) = \sup_{\rho_X(x_1,x_2)\le t} \rho_Y(f(x_1), f(x_2)).$$

The continuity modulus $\omega_f$ of an arbitrary continuous function $f$ is a nonnegative semi-additive function (i.e., $\omega_f(t_1+t_2) \le \omega_f(t_1)+\omega_f(t_2)$) vanishing and continuous at zero. If an arbitrary function $\omega$ possesses these properties then we call $\omega$ a *continuity modulus.*

**Assertion 3.** *Let $\omega$ be a continuity modulus. Then $\omega$ is nondecreasing, $\omega(nt) \le n\omega(t)$ (for every integer $n > 0$), and $\omega(At) \le (A+1)\omega(t)$ (for every real $A > 0$).*

Call $\omega$ a *polynomial-type continuity modulus* if $\omega$ is a continuity modulus and $\omega(t) \asymp t^\alpha$ as $t \to 0$ for some $\alpha$, $0 < \alpha \le 1$.

**Assertion 4.** *Let $\omega$ be a polynomial-type continuity modulus. Then there are numbers $a_\omega$ ($a_\omega \ge 1$), $\alpha$ ($0 < \alpha \le 1$), and $t_0$ ($t_0 > 0$) such that the inequality*

$$\omega(t) \ge \frac{A^\alpha}{a_\omega}\omega\left(\frac{t}{A}\right) \text{ for all } t, \quad t_0 > t > 0$$

*is valid for all real numbers $A > 0$.*

PROOF. Since $\omega$ is a polynomial-type continuity modulus, there are constants $C > 1$ and $t_0 > 0$ such that the inequalities $Ct^\alpha \ge \omega(t) \ge t^\alpha/C$ hold for $t < t_0$. These inequalities imply that

$$\omega(t/A) \le \frac{Ct^\alpha}{A^\alpha} \le \frac{C^2}{A^\alpha}\omega(t),$$

which completes the proof of Assertion 4.

Given a continuity modulus $\omega$, denote by $F_\omega(X,Y)$, or by $F_\omega$ in case this notation does not lead to misunderstanding, the subset of $C(X,Y)$ that consists of the functions whose continuity moduli do not exceed $\omega$:

$$F_\omega(X,Y) = \{f \in C(X,Y) : (\forall x_1, x_2 \in X)\ \rho_Y(f(x_1), f(x_2)) \le \omega(\rho_X(x_1,x_2))\}.$$

Call the members of $F_\omega(X,Y)$ *$\omega$-continuous* functions.

The Arzelà Theorem asserts that $F_\omega$ is a totally bounded set for every continuity modulus $\omega$ and conversely: each totally bounded set in $C(X,Y)$ lies in $F_\omega$ for some continuity modulus $\omega$.

However, some totally bounded sets in $C(X,Y)$ are essentially smaller than the ambient sets $F_\omega$. For example, let $X$ be an open bounded subset of $\mathbb{R}^k$ with the norm $\|x-x'\| = \max_{i=1,\dots,k}|x_i - x'_i|$, and let $Y$ be an interval in $\mathbb{R}$. Consider the set $F^k_{n,\alpha}$ comprising the functions of $C(X,Y)$ whose $n$th order derivatives satisfy the Hölder condition with constant 1 and exponent $\alpha$ ($0 < \alpha \le 1$):

$$F^k_{n,\alpha}(X,Y) = \{f \in C(X,Y) : |D^\mu f(x_1) - D^\mu f(x_2)| \le \|x_1 - x_2\|^\alpha\},$$

where $\mu$ are multi-indices of length $n$ ($|\mu| = n$).

In [2], A. G. Vitushkin derived the following lower bound on the size of a $\delta$-distinguishable set in $F_\omega(X,Y)$ for an arbitrary $X$ and $Y = [0,1]$:

$$M_{F_\omega}(\omega(\delta/2)) \ge 2^{M_X(4\delta)}.$$

A. G. Vitushkin's proof of this result can be extended to wider classes of continuous functions. This is implemented in the following lemma.

**Lemma 5.** *Let $X$ be a totally bounded metric space, let $Y$ be a totally bounded set in a normed vector space which includes a segment, and let $\omega$ be a polynomial-type continuity modulus. Then there are numbers $a_\omega$ ($a_\omega \geq 1$), $\delta_0$ ($\delta_0 > 0$), and $\alpha$ ($0 < \alpha \leq 1$) such that the inequality*

$$M_{F_\omega}(\omega(\delta)) \geq \left(\frac{A^\alpha}{a_\omega}\right)^{M_X(2A\delta)}$$

*holds for all $A$ ($A^\alpha > a_\omega$) and $\delta$ ($\delta_0 > \delta > 0$).*

PROOF. Let $y_0, y_1 \in Y$ be such that the segment $ty_0 + (1-t)y_1$, $t \in [0,1]$, lies in $Y$. Let $P$ be a $2A\delta$-distinguishable subset of $X$ such that $|P| = M_X(2A\delta)$. Since $\omega$ is a polynomial-type continuity modulus, there are $a_\omega \geq 1$, $\delta_0 > 0$, and $0 < \alpha \leq 1$ such that $\omega(A\delta) \geq \frac{A^\alpha}{a_\omega}\omega(\delta)$ for every $\delta$, $\delta_0 > \delta > 0$.

Enumerate the elements of $P$ somehow to obtain $P = \{x_i, i = 1, \dots, |P|\}$. Given a multi-index $k$, $k \in \mathbb{N}^{|P|}$, such that $k_i = 0, 1, \dots, K$ with $K = \lfloor \frac{A^\alpha}{a_\omega} \rfloor$, define the function $f^k : X \to Y$ as follows:

$$f^k(x) = \begin{cases} \frac{(y_1 - y_0)}{\|y_1 - y_0\|}\frac{k_i}{K}\omega(A\delta - \rho_X(x, x_i)) + y_0 & \text{if } \rho_X(x, x_i) \leq A\delta; \\ y_0 & \text{if } \rho_X(x, x_i) > A\delta \text{ for all } y \in P. \end{cases}$$

The definition of $f^k$ is sound since $P$ is a $2A\delta$-distinguishable subset.

Prove that $f^k \in F_\omega$. We have to examine four possible cases.

1. Consider $x', x'' \in X$ satisfying $\rho_X(x', x_i) \leq A\delta$ and $\rho_X(x'', x_i) \leq A\delta$ for some $x_i \in P$. Then the semi-additivity of a continuity modulus implies that

$$\begin{aligned} \|f^k(x') - f^k(x'')\| &\leq \frac{k_i}{K}|\omega(A\delta - \rho_X(x', x_i)) - \omega(A\delta - \rho_X(x'', x_i))| \\ &\leq \omega(|\rho_X(x', x_i) - \rho_X(x'', x_i)|) \leq \omega(\rho_X(x', x'')). \end{aligned}$$

2. Consider $x', x'' \in X$ satisfying $\rho_X(x', x_i) \leq A\delta$ for some $x_i \in P$ and $\rho_X(x'', y) > A\delta$ for every $y \in P$. Then

$$\|f^k(x') - f^k(x'')\| \leq \omega(A\delta - \rho_X(x', x_i)).$$

From the monotonicity of $\omega$ and the inequality $\rho_X(x'', x_i) > A\delta$ we infer that

$$\omega(A\delta - \rho_X(x', x_i)) \leq \omega(\rho_X(x'', x_i) - \rho_X(x', x_i)) \leq \omega(\rho_X(x', x'')).$$

3. Consider $x', x'' \in X$ satisfying $\rho_X(x', x_i) \leq A\delta$ and $\rho_X(x'', x_j) \leq A\delta$ for some $x_i, x_j \in P, i \neq j$. Then $\rho_X(x', x_j) \geq A\delta$ and $\rho_X(x'', x_i) \geq A\delta$. The monotonicity of $\omega$ implies that

$$\begin{aligned} \|f^k(x') - f^k(x'')\| &\leq \left|\frac{k_i}{K}\omega(A\delta - \rho_X(x', x_i)) - \frac{k_j}{K}\omega(A\delta - \rho_X(x'', x_j))\right| \\ &\leq \max(\omega(A\delta - \rho_X(x', x_i)), \omega(A\delta - \rho_X(x'', x_i))) \leq \omega(\rho_X(x', x'')). \end{aligned}$$

4. Consider the case of $\rho_X(x', y) \geq A\delta$ and $\rho_X(x'', y) \geq A\delta$ for some $y \in P$. Then

$$\|f^k(x') - f^k(x'')\| = 0.$$

We have thus demonstrated that $f^k \in F_\omega$. Prove that $\rho_C(f^k, f^l) > \omega(\delta)$ for $k \neq l$. Indeed, if $k \neq l$ then there exists $i$, $1 \leq i \leq |P|$, such that $k_i \neq l_i$. From the definition of the number $K$ and the functions $f^k$ and $f^l$ we infer that

$$\rho_C(f^k, f^l) \geq \|f_i^k(x_i) - f_i^l(x_i)\| \geq |k_i - l_i|\omega(A\delta)/K \geq \omega(\delta).$$

Thus, we constructed an $\omega(\delta)$-distinguishable subset in $F_\omega$ whose size equals $(1 + K)^{|P|} \geq \left(\frac{A^\alpha}{a_\omega}\right)^{M_X(2A\delta)}$. This completes the proof of Lemma 5.

In the next lemma we give an estimate for the complexity of a piecewise constant function approximating an $\omega$-continuous function with accuracy $\varepsilon$. Recall that the complexity of a function that is constant over each element of a covering $R$ and takes $|S|$ distinct values does not exceed $|R| \log |S|$. The complexity of a piecewise polynomial function on $R$ does not exceed $(n+1)|R| \log |S|$, where $n$ is the maximal degree of polynomials each of whose coefficients takes at most $|S|$ values.

**Lemma 6.** *Suppose that $\delta > 0$, $X$ is a totally bounded metric space, $R$ is a $\delta$-covering of $X$, $V$ is a finite-dimensional vector space, $\omega$ is a continuity modulus, and $f : X \to V$ is a bounded function (i.e., $\|f(x)\| \le d$ for all $x \in X$ with some constant $d$) $\omega$-continuous over each element of $R$. Then there exists a piecewise constant function $f_\delta$, whose complexity does not exceed $h_V(2\omega(\delta)/d)|R|$ ($h_V(t) = H^V_{B(0,1)}(t)$), such that $\rho_C(f, f_\delta) \le 4\omega(\delta)$.*

PROOF. Let $Q$ be a $2\omega(\delta)$-net in the ball $B_V(0,d)$ such that $|Q| = 2^{h_V(2\omega(\delta)/d)}$. For each element $r_i$ of $R$, choose an arbitrary point $x_i \in r_i$. There is $y_i \in Q$ such that $\|f(x_i) - y_i\| \le 2\omega(\delta)$. Define the piecewise constant $f_\delta : X \to V$ as identically equal to $y_i$ on each element $r_i$ of the covering $R$. The definition is correct, because the elements of $R$ are pairwise disjoint. The complexity of $f_\delta$ does not exceed $h_V(2\omega(\delta)/d)|R|$. We demonstrate that $\rho_C(f, f_\delta) \le 4\omega(\delta)$. Suppose that $x \in r_i$. Then

$$\begin{aligned}\|f_\delta(x) - f(x)\| &\le \|f_\delta(x) - f(x_i)\| + \|f(x_i) - f(x)\| \le \|y_i - f(x_i)\| + \|f(x_i) - f(x)\| \\ &\le 2\omega(\delta) + \omega(2\delta) \le 4\omega(\delta),\end{aligned}$$

which completes the proof of Lemma 6.

We can prove an analogous result on approximating differentiable functions by piecewise polynomial functions. Let $R$ be a covering of a set $X$. We write $f \in F^k_{n,\alpha}[R](X,Y)$ if $f|_r \in F^k_{n,\alpha}(r,Y)$ for all $r \in R$.

**Lemma 7.** *Suppose that $X$ is an open bounded subset of $\mathbb{R}^k$, $\delta > 0$; $R$ is a $\delta/2$-covering of $X$; $Y$ is an interval in $\mathbb{R}$; and $f : X \to Y$ is a function of the class $F^k_{n,\alpha}[R](X,Y)$; moreover, the inequality*

$$\left(\frac{|D^\mu f(x)|}{1+e^k}\right)^{\frac{n+\alpha}{n+\alpha-|\mu|}} \le d^{(n+\alpha)} \quad \text{over each } r \in R$$

*holds for all $x \in X$ and all multi-indices $\mu$ ($0 \le |\mu| \le n$), with $e$ the base of the natural logarithm and $d$ some constant. Then there exists a piecewise polynomial function $f_\delta$ whose complexity does not exceed $Ck|R|\log\frac{d}{\delta}$, with $C$ some constant, and for which the inequality*

$$\left(\frac{|D^\mu (f-f_\delta)(x)|}{1+e^k}\right)^{\frac{n+\alpha}{n+\alpha-|\mu|}} \le \delta^{(n+\alpha)} \quad \text{over each } r \in R.$$

*holds for all $\mu$ ($0 \le |\mu| \le n$) and $x \in X$.*

We omit the proof of Lemma 7, since the only difference between it and the proof of Lemma 6 is that the initial function is approximated over each element of the covering by a partial sum of the Taylor series rather than a constant. In [1, 6], there is a proof of an assertion analogous to Lemma 7.

In [2], the following estimates were obtained for the $\varepsilon$-entropy of the set $F_\omega(X,Y)$:

$$H_{F_\omega}(\omega(\delta)/4) \ge N_X(4\delta) \tag{3}$$

if $X$ is an arbitrary totally bounded space and

$$H_{F_\omega}(4\omega(\delta)) \le CN_X(\delta), \tag{4}$$

with $C$ some constant, if $X$ is a connected compact space.

The following theorem gives a necessary and sufficient condition on $X$ for the $\varepsilon$-entropy of $F_\omega(X,Y)$ to satisfy (4) under some constraints on $\omega$ and $Y$.

The proof of the theorem relies on the successive approximation method. In this aspect it differs principally from A. G. Vitushkin's approach to proving (4), which is based on using the connectedness of $X$.

**Theorem 1.** *Suppose that $X$ a totally bounded metric space, $V$ is a finite-dimensional vector space, $Y$ is a bounded subset of $V$ including some segment, $\omega$ is a polynomial-type continuity modulus, and $F_\omega$ is the set of $\omega$-continuous function from $X$ into $Y$. Then*

$$H_{F_\omega}(4\omega(\delta)) = O(2^{H_X(\delta)}) \quad \text{as } \delta \to 0$$

*if and only if $X$ is a uniform-type space.*

PROOF. *Sufficiency*: By definition, a totally bounded uniform-type space $X$ possesses the following property: there are $\beta > 1$, $C > 0$, and $\varepsilon_0 > 0$ such that the inequality

$$H_X(\beta\varepsilon) + C \leq H_X(\varepsilon) \tag{5}$$

holds for all $\varepsilon < \varepsilon_0$. Without loss of generality, we may assume that $\beta \geq 3$, since if (5) holds for some $\beta$ then it also holds (with other $\varepsilon_0$ and $C$) for $\beta^i$, with $i$ natural. Since $Y$ is a bounded set, there exists a number $d$ such that $\|f(x)\| \leq d < \infty$ for all $x \in X$.

Suppose that $\delta < \varepsilon_0$ and let $R_1$ be a $\delta$-covering of $X$ such that $|R_1| = N_X(\delta)$. Define the following sequence: $\delta_1 = \delta$ and $\delta_{i+1} = \beta\delta_i$. Let $n$ be such that $\varepsilon_0 \leq \delta_n < \beta\varepsilon_0$. Then $n = O(\log \frac{1}{\delta})$.

By Lemma 4, to each $i \leq n$ there is a $\delta_i$-covering $R_i$ of $X$ generated by $R_{i-1}$ and such that

$$|R_i| \leq N_X(\delta_i - 2\delta_{i-1}). \tag{6}$$

Suppose that $f \in F_\omega(X, Y)$. According to Lemma 6, there exists a function $g_n$ constant over each element of the covering $R_n$ and such that $\rho_C(f, g_n) \leq 4\omega(\delta_n)$ and the complexity of $g_n$ does not exceed

$$h_V\left(\frac{2\omega(\delta)}{d}\right)|R_n|. \tag{7}$$

The function $f_1 = f - g_n$ is $\omega$-continuous over each element of the covering $R_{n-1}$, since the covering $R_n$ is generated by $R_{n-1}$ and the function $g_n$ is constant over each element of $R_n$. According to Lemma 6, there exists a function $g_{n-1}$ which is constant over each element of the covering $R_{n-1}$, whose complexity does not exceed $h_V(\frac{\omega(\delta_{n-1})}{2\omega(\delta_n)})|R_{n-1}|$, and for which the inequality $\rho_C(f_1, g_{n-1}) \leq 4\omega(\delta_{n-1})$ holds.

By analogy, we construct the sequence $g_i$, $i = n-2, \ldots, 1$. By Lemma 6 we have

$$\rho_C\left(f, \sum_{i=1}^{n} g_i\right) \leq 4\omega(\delta),$$

and the complexity of each function $g_i$ does not exceed

$$h_Y\left(\frac{\omega(\delta_i)}{2\omega(\delta_{i+1})}\right)|R_i|. \tag{8}$$

Assertion 3 implies that $\omega(\delta_{i+1}) = \omega(\beta\delta_i) \leq (\beta + 1)\omega(\delta_i)$, while the monotonicity of $h_V$ implies the inequality

$$h_V\left(\frac{\omega(\delta_i)}{2\omega(\delta_{i+1})}\right) \leq h_V\left(\frac{1}{2(\beta+1)}\right). \tag{9}$$

From (5) and (6) we infer that

$$\frac{|R_{i+1}|}{|R_i|} = \frac{N_X(\delta_i - 2\delta_{i-1})}{N_X(\delta_{i+1} - 2\delta_i)} \leq 2^{-C}. \tag{10}$$

Since $\beta \geq 3$ and $N_X$ does not increase, we have

$$|R_1| = N_X(\delta) \geq N_X(\beta\delta - 2\delta) = |R_2|. \tag{11}$$

From (8)–(11) we infer that the sum of the complexities of the functions $g_2, \ldots, g_{n-1}$ does not exceed

$$\sum_{i=1}^{n-1} h_V\left(\frac{\omega(\delta_i)}{\omega(2\delta_{i+1})}\right)|R_i| \le |R_1|\sum_{i=1}^{n-1} h_V\left(\frac{\omega(\delta_i)}{\omega(2\delta_{i+1})}\right)\frac{|R_i|}{|R_1|}$$

$$\le h_V\left(\frac{1}{2(\beta+1)}\right)|R_1|\left(1+\sum_{i=0}^{n-3} 1/2^{iC}\right) = C'N_X(\delta), \tag{12}$$

where

$$C' = h_V(\frac{1}{2(\beta+1)})\left(1+\sum_{i=0}^{n-3}\frac{1}{2^{iC}}\right).$$

The complexity of $g_n$ does not exceed the constant $h_V(\frac{2\omega(\varepsilon_0)}{d})N_X(\varepsilon_0 - 2\varepsilon_0/\beta)$; denote this constant by $C''$.

As was observed in the introduction, the $\varepsilon$-entropy of the set $F_\omega$ is dominated by the maximum of the sums of the complexities of the functions $g_1, \ldots, g_n$. Thus, if $X$ is a uniform-type space then (12) implies that

$$H_{F_\omega}(4\omega(\delta)) \le C'N_X(\delta) + C'' \quad \text{for } \delta < \varepsilon_0.$$

*Necessity*: Suppose that $X$ is a totally bounded but not uniform-type space. Then, for arbitrary $\beta > 1$, $C > 0$, and $\varepsilon_0 > 0$, there exists an $\varepsilon < \varepsilon_0$ such that

$$H_X(\beta\varepsilon) + C \ge H_X(\varepsilon). \tag{13}$$

Since $\omega$ is a polynomial-type continuity modulus, Assertion 4 implies existences of constants $0 < \alpha \le 1$, $\delta_0 > 0$, and $a_\omega \ge 1$ such that

$$\omega(\delta) \ge \frac{A^\alpha}{a_\omega}\omega\left(\frac{\delta}{A}\right)$$

for all $\delta < \delta_0$ and $A > A_0 = a_\omega^{1/\alpha}$. Introduce the notation $b_\omega = 1/(8a_\omega)^{1/\alpha}$. Then

$$\omega(\delta)/2 \ge 4\omega(b_\omega\delta). \tag{14}$$

From Lemma 5 and Assertion 1 we infer that

$$H_{F_\omega}\left(\frac{\omega(\delta)}{2}\right) \ge \log\left(\frac{A^\alpha}{a_\omega}\right)N_X(2A\delta) \tag{15}$$

for all $A > A_0$ and $\delta_0 > \delta > 0$. If follows from (14), (15), and the monotonicity of $H_{F_\omega}$ that

$$H_{F_\omega}(4\omega(b_\omega\delta)) \ge \log\left(\frac{A^\alpha}{a_\omega}\right)N_X\left(\frac{2A}{b_\omega}b_\omega\delta\right). \tag{16}$$

Insert (13) in (16) with $\beta = 2A/b_\omega$, $\delta_1 = \min(\varepsilon_0, \delta_0/A)$, and $\varepsilon = b_\omega\delta$. Then for arbitrary $A > A_0$ and $C > 1$ we can find $0 < \delta < \delta_1$ such that

$$H_{F_\omega}(4\omega(b_\omega\delta)) \ge 2^{-C}\log\left(\frac{A^\alpha}{2a_\omega}\right)N_X(b_\omega\delta).$$

Since $b_\omega < 1$, whereas $A > A_0$ and $C > 1$ are arbitrary; therefore, for any $C' > 0$ and $\delta_1 > 0$ there exists a $\delta' < \delta_1$ such that

$$H_{F_\omega}(4\omega(\delta')) \ge C'N_X(\delta').$$

This completes the proof of Theorem 1.

In the proof of sufficiency, we did not use the conditions that $\omega$ is a polynomial-type continuity modulus and that $Y$ includes a segment. Since by Lemma 1 every connected totally bounded space is a uniform-type space, Theorem 1 thus generalizes estimate (4) of A. G. Vitushkin to a wider set of function spaces. In particular, Lemma 3 and Theorem 1 imply that the $\varepsilon$-entropy of the space of Lipschitz continuous functions on the Cantor set is $O((1/\varepsilon)^{\log_3 2})$ as $\varepsilon \to 0$.

## § 4. Tabulation of Continuous and Differentiable Functions

In the proof of Theorem 1, an $\omega$-continuous function was approximated to within $4\omega(\delta)$ by the sum of $O(\log 1/\delta)$ piecewise constant functions. In the following theorem we propose a method for approximating a continuous function with the same accuracy by the sum of $O(\log^* 1/\delta)$ piecewise constant functions whose total complexity is $O(N_X(\delta))$.

**Lemma 8.** *Let the sequence of numbers $\alpha_1, \dots \alpha_i, \dots$ be given recursively as follows: $\alpha_1 = 1$ and $\alpha_{i+1} = 2^{\alpha_i - i+1}$. Then $\alpha_{2(\log^* N)+4} \geq N$ for all $N > 0$.*

PROOF. Consider the sequence $\beta_1, \dots, \beta_i, \dots$ given recursively as follows: $\beta_1 = 1$ and $\beta_{i+1} = 2^{\beta_i}$. Define $\gamma_1, \dots, \gamma_i, \dots$ by the rule $\gamma_{2i} = \beta_{i-1}$, $\gamma_{2i+1} = \gamma_{2i}$. The definitions of $\beta_i$ and $\gamma_i$ imply that $\gamma_{2(\log^* N)+4} = \beta_{(\log^* N)+1} \geq N$.

We prove by induction that $\gamma_{2i} < \alpha_{2i}$. Straightforward verification demonstrates that $\alpha_i > \gamma_i$ for $i \leq 6$ while $\gamma_{2i-2} > 6i$ and $2(\gamma_{2i-2} - 2i + 1) \leq 2^{\gamma_{2i-2} - 2i+1}$ for $i > 6$.

Assume that $i \geq 7$. Suppose that $\gamma_{2i-2} < \alpha_{2i-2}$. Then

$$\log \gamma_{2i} = \gamma_{2i-2} \leq 2\gamma_{2i-2} - 6i = 2(\gamma_{2i-2} - 2i - 1) + 2 - 2i \leq 2^{\gamma_{2i-2} - 2i-1} + 2 - 2i$$
$$< 2^{\gamma_{2i-2} - (2i-2)+1} - (2i - 1) + 1 < 2^{\alpha_{2i-2} - (2i-2)+1} - (2i - 1) + 1.$$

These inequalities and the definition of the sequence $\alpha_1, \dots, \alpha_i, \dots$ imply that $\gamma_{2i} < \alpha_{2i}$ for all $i$. Since $\gamma_{2i+1} = \gamma_{2i} < \alpha_{2i} < \alpha_{2i+1}$; therefore, $\gamma_i < \alpha_i$ for all $i$. Hence, $\alpha_{2(\log^* N)+4} \geq N$, as desired.

**Theorem 2.** *Suppose that $X$ is a totally bounded convex-type space, $V$ is a finite-dimensional vector space, $Y$ is a bounded subset of $V$, $\omega$ is a continuity modulus, and $f : X \to Y$ is an $\omega$-continuous function. Then there is $\delta_0 > 0$ such that, for every $0 < \delta < \delta_0$, there exists a function $f_\delta$ with the following properties:*

(1) *$f_\delta$ is the sum of $O(\log^* \frac{1}{\delta})$ piecewise constant functions;*

(2) *the sum of the complexities of the functions in item* (1) *is $O(2^{H_X(\delta)})$ as $\delta \to 0$;*

(3) *$\rho_C(f_\delta, f) \leq 4\omega(\delta)$.*

PROOF. Since $Y$ is bounded, there exists $d$ such that $\|f(x)\| \leq d < \infty$ for all $x \in X$. Since $X$ is a convex-type space, Lemma 2 implies existence of $\varepsilon_0$ such that the inequality

$$\frac{N_X(\varepsilon_2)}{N_X(\varepsilon_1)} \leq C' \frac{\varepsilon_1}{\varepsilon_2} \tag{17}$$

holds for all $0 < \varepsilon_1 < \varepsilon_2 < \varepsilon_0$ with some constant $C'$. Assume that $0 < \delta < \varepsilon_0$ and let $\alpha_1, \dots, \alpha_i, \dots$ be the sequence constructed in Lemma 8. Define the sequence $\delta_1, \dots, \delta_i, \dots$ as follows: $\delta_0 = \delta$ and $\delta_i = 2^{\alpha_i}\delta$. Let $n$ be the smallest number for which $\delta_{n+1} \geq \varepsilon_0$. Then $n = O(\log^* \frac{1}{\delta})$ as $\delta \to 0$, since by Lemma 8 $\alpha_{2\log^* \varepsilon_0/\delta+4} \geq \varepsilon_0/\delta$ and $\delta_{2\log^* \varepsilon_0/\delta+4} = \delta 2^{\alpha_{2\log^* \varepsilon_0/\delta+4}} \geq \varepsilon_0$.

Construct a finite sequence $R_0, \dots, R_n$, where $R_i$ is a $\delta_i$-covering of $X$, $i = 0, \dots, n$, such that the covering $R_i$ is generated by $R_{i-1}$; moreover,

$$|R_0| = N_X(\delta), \quad |R_i| = N_X(\delta_i - 2\delta_{i-1}) \quad \text{for} \quad i = 1, \dots, n.$$

According to Lemma 6, there is a sequence of piecewise constant functions $g_0, \dots, g_n$ such that

$$\rho_C\left(f, \sum_{i=0}^{n} g_i\right) \leq 4\omega(\delta)$$

and the complexity of each function $g_i$ does not exceed

$$h_V\left(\frac{\omega(\delta_i)}{2\omega(\delta_{i+1})}\right)|R_i| \quad \text{for} \quad i = 0, \dots, n-1. \tag{18}$$

The complexity of $g_n$ does not exceed

$$h_V\left(\frac{2\omega(\delta_n)}{d}\right)|R_n| \le h_V\left(\frac{\omega(\delta_n)}{2\omega(\delta_{n+1})}\frac{4\varepsilon_0}{d}\right)|R_n|,$$

because $\omega(\varepsilon_0) \le \omega(\delta_{n+1})$. Since $\omega$ is a continuity modulus, the following inequalities hold:

$$\omega(\delta_{i+1}) = \omega(2^{\alpha_{i+1}}\delta) = \omega(2^{\alpha_{i+1}-\alpha_i}\delta_i) \le 2^{\alpha_{i+1}-\alpha_i}\omega(\delta_i), \quad \frac{\omega(\delta_i)}{\omega(\delta_{i+1})} \ge \frac{2^{\alpha_i}}{2^{\alpha_{i+1}}} \ge \frac{1}{2^{\alpha_{i+1}}}.$$

Since $V$ is finite-dimensional; the monotonicity of $h_V$, Assertion 2, and the last inequality imply that

$$h_V\left(\frac{\omega(\delta_i)}{2\omega(\delta_{i+1})}\right) \le h_V\left(\frac{1}{2^{1+\alpha_{i+1}}}\right) \le C\alpha_{i+1}, \tag{19}$$

with some constant $C$. In accordance with (18) and (19), the sum of the complexities of the functions $g_0, \dots, g_n$ does not exceed

$$C\log\frac{d}{4\varepsilon_0} + C\sum_{i=0}^{n}\alpha_{i+1}|R_i| = C\log\frac{d}{4\varepsilon_0} + C|R_0|\sum_{i=0}^{n}\alpha_{i+1}\frac{|R_i|}{|R_0|}. \tag{20}$$

From (17) we infer that

$$\frac{|R_i|}{|R_0|} = \frac{N_X(\delta_i - 2\delta_{i-1})}{N_X(\delta)} \le \frac{C'\delta}{\delta(2^{\alpha_i} - 2^{\alpha_{i-1}+1})} \le \frac{C'}{2^{\alpha_i}}. \tag{21}$$

Inserting (21) in (20) and using the definition of the sequence $\alpha_1, \dots, \alpha_i, \dots$, we conclude that the sum of the complexities of the functions $g_0, \dots, g_n$ is $O(N_X(\delta))$ as $\delta \to 0$, which completes the proof of Theorem 2.

A similar assertion is valid for differentiable functions.

**Theorem 3.** *Suppose that $X$ is an open bounded subset of $\mathbb{R}^k$, $Y$ is a segment in $\mathbb{R}$, and $f \in F^k_{n,\alpha}(X, Y)$. For some $\delta_0 > 0$ and each $0 < \delta < \delta_0$ there exists a function $f_\delta$ with the following properties:*

(1) *$f_\delta$ is the sum of $O(\log^* \frac{1}{\delta})$ piecewise polynomial functions;*

(2) *the sum of the complexities of the functions of item* (1) *is* $O(2^{H_X(\delta/2)})$ as $\delta \to 0$;

(3) $\rho_C(f_\delta, f) \le (1 + e^k)\delta^{n+\alpha}$, *with $e$ the base of the natural logarithm.*

The proof of Theorem 3 is analogous to that of Theorem 2, the difference being that Lemma 7 is used in place of Lemma 6 and the role of $h_V$ is played by log.

The complexity of the approximating function in Theorem 3 is minimal, since, as shown in [1], $H_{F^k_{n,\alpha}}(\delta^{n+\alpha}) \asymp (1/\delta)^k \asymp 2^{H_X(\delta/2)}$, provided that the departure space $X$ of the function class $F^k_{n,\alpha}$ is an open bounded subset of $\mathbb{R}^k$.

In conclusion, I express my sincere gratitude to Yu. G. Reshetnyak for the helpful remarks he made while discussing the results of the article and to R. E. Krichevskiĭ and V. A. Aleksandrov for numerous pieces of advice in preparation of the results of the article to publication.

## References

1. A. N. Kolmogorov and V. M. Tikhomirov, "ε-Entropy and ε-capacity of sets in function spaces," Uspekhi Mat. Nauk, **14**, No. 2, 3–86 (1959).
2. A. G. Vitushkin, Estimating the Complexity of the Tabulation Problem [in Russian], Fizmatgiz, Moscow (1959).

3. A. F. Timan, "$\varepsilon$-Entropy of Arzelà compacta of functions given on closed sets of positive Lebesgue measure," Dokl. Akad. Nauk SSSR, **177**, No. 2, 285–287 (1967).
4. A. F. Timan, "The exact order of $\varepsilon$-entropy and $\varepsilon$-capacity of Arzelà compacta of any degree of massiveness," in: Metric Questions of the Theory of Functions and Mappings [in Russian], Naukova Dumka, Kiev, 1977, pp. 131–142.
5. E. A. Asarin, "Complexity of uniform approximations to continuous functions," Uspekhi Mat. Nauk, **39**, No. 3, 157–169 (1984).
6. V. N. Potapov, "Tabulation of smooth functions," in: Analysis and Discrete Mathematics [in Russian], Novosibirsk Univ., Novosibirsk, 1995, pp. 115–125.
7. J. A. Dieudonné, Foundations of Modern Analysis [Russian translation], Mir, Moscow (1964).
8. A. K. Zvonkin and L. A. Levin, "Complexity of finite objects and verification of the concepts of information and randomness by means of the theory of algorithms," Uspekhi Mat. Nauk, **25**, No. 6, 85–127 (1970).