



Soviet-era science, translated into English

B. S. MITYAGIN

1960

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196001.90139>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

B. S. MITYAGIN

**THE RELATION BETWEEN ϵ -ENTROPY,
THE RATE OF APPROXIMATION, AND
THE NUCLEARITY OF A COMPACT SET IN
A LINEAR SPACE**

(Presented by Academician A. N. Kolmogorov, 23 IV 1960)

A. N. Kolmogorov ⁽¹⁾ and A. Pelczynski ⁽²⁾ introduced the concept of the *approximative dimension* of linear topological spaces. In the present note there is obtained, in particular, in terms of this concept, a complete characterization of nuclear spaces ⁽³⁾, which are very important in analysis. Our results give a complete solution of a problem of I. M. Gelfand (see ⁽⁴⁾, p. 6), posed at the Third All-Union Conference on Functional Analysis (January 1956).

In what follows we shall be concerned with *locally convex* linear topological spaces, real or complex, although, for simplicity, some computations are given only for the real case.

1. Let E be a linear space, S a centrally symmetric convex (c.s.c.) set in E , and A an arbitrary set in E .

Definition 1 (see ^(1,5,6)). Put

$$N(A, S; \varepsilon) = \inf_{x_k \in E} \left\{ N : \bigcup_1^N (x_k + \varepsilon S) \supset A \right\},$$

$$M(A, S; \varepsilon) = \sup_{x_k \in A} \{ M : x_k - x_j \notin \varepsilon S; j \neq k; k, j = 1, \dots, M \}.$$

The number $H(A, S; \varepsilon) = \log N(A, S; \varepsilon)$ will be called the ε -*entropy* of A relative to S , and the number $\log M(A, S; \varepsilon)$ its ε -*capacity*.

Definition 2 (cf. ⁽⁶⁾, p. 22). We shall call the *order* $\rho(A, S)$ of the set A relative to the c.s.c. set S the number

$$\rho(A, S) = \lim_{\varepsilon \rightarrow 0} \left\{ \log H(A, S; \varepsilon) / \log \frac{1}{\varepsilon} \right\}.$$

Lemma 1 ((⁶), p. 8).

$$M(A, S; \varepsilon) \geq N(A, S; \varepsilon) \geq M(A, S; 2\varepsilon).$$

Lemma 2 (⁷). Let U_i , $i = 1, \dots, n$, be c.s.c. sets in E and

$$\rho_{ki} = \rho(U_k, U_i).$$

Then

$$\frac{1}{\rho_{n1}} \geq \sum_1^{n-1} \frac{1}{\rho_{k+1,k}}.$$

Let $\{u_1, \dots, u_n\}$ be a basis in the n -dimensional space E_n , and let p_n be the octahedron generated by the vectors $a_k u_k$, $k = 1, \dots, n$ ($a_k > 0$ and fixed), i.e.

$$p_n = \left\{ \sum_1^n \xi_k a_k u_k : \sum_1^n |\xi_k| \leq 1 \right\}.$$

Lemma 3. The number of points of the form

$$y = \sum_1^n \gamma_i u_i,$$

where γ_i are integers, lying in p_n , is not less than

$$\prod_1^n \frac{a_k}{k}.$$

Indeed, if each point $y = \sum_1^n \gamma_i u_i$, γ_i integers, lying in p_n , is surrounded by a cube K_y with side length "two," then $\bigcup_{y \in p_n} K_y \supset p_n$. Counting the volumes of these sets leads to the assertion of the lemma.

Definition 3 (see (⁸)). The n -dimensional **width** of a set A in a Banach space B with unit ball S is

$$d_n = d_n(A, S) = \inf r(A, E_n),$$

where

$$r(A, E_n) = \sup_{x \in A} \inf_{y \in E_n} |x - y|_S$$

and the infimum is taken over all n -dimensional subspaces E_n in B .

2. Computation of the functions $H(\varepsilon)$ and d_n for concrete compacta in many function spaces shows ⁽⁶⁾ that the asymptotics of these functions are closely connected. Below this connection is investigated for any centrally symmetric convex compactum K in an arbitrary Banach space B .

Let us consider the functions

$$m(t) = \max \left\{ n : d_n \geq \frac{1}{t} \right\}$$

and

$$l(t) = \max \left\{ n : \frac{d_n}{n} \geq \frac{1}{t} \right\},$$

which characterize the compactum K .

Lemma 4 (⁽⁹⁾, p. 20).

$$\alpha = \inf \left\{ \beta : \sum d_n^\beta < \infty \right\} = \overline{\lim}_{t \rightarrow 0} \{ \log m(t) / \log t \}.$$

Theorem 1. In an arbitrary Banach space B , for every centrally symmetric convex compactum K the inequalities

$$m \left(\frac{2}{\varepsilon} \right) \log \frac{8(d_0 + \varepsilon)}{\varepsilon} \geq H(K, S; \varepsilon) \geq \int_0^{1/\varepsilon} \frac{l(t)}{t} dt. \quad (1)$$

The **left-hand inequality**^{*}, valid for any compactum, is easily obtained by using the discussion in § 4 of ⁽⁶⁾.

The **right-hand inequality**. Let $a_1 = d_0 = \sup_{x \in K} |x|$; since K is compact, for some $x_1 \in K$ we have $|x_1| = a_1$. By induction one constructs a system of vectors x_i , $i = 1, 2, \dots$, such that $x_i \in K$ and

$$a_{n+1} = r(x_{n+1}, L_n) \geq d_n$$

(1), where $L_n = L\{x_1, \dots, x_n\}$; this construction is possible, since (see Definition 3) $r(K, L_n) \geq d_n$ and K is compact.

By the convexity and central symmetry of K , for all n it contains the octahedra

$$\sigma_n = \left\{ \sum_1^n \xi_k x_k : \sum_1^n |\xi_k| \leq 1 \right\}.$$

For any** two distinct points

$$y_1 = \sum_1^n \alpha_k \varepsilon a_k^{-1} x_k, \quad y_2 = \sum_1^n \beta_k \varepsilon a_k^{-1} x_k,$$

where α_k, β_k are integers, we have

$$z = y_1 - y_2 = \sum_1^l \gamma_k \varepsilon a_k^{-1} x_k, \quad 1 \leq l \leq n, \quad |\gamma_l| \geq 1,$$

and

$$|z| \geq r(z, L_{l-1}) = r(\gamma_l \varepsilon a_l^{-1} x_l, L_{l-1}) \geq \varepsilon,$$

so that, by Lemma 3,

$$M(\sigma_n, S; \varepsilon) \geq \prod_1^n \frac{a_k}{\varepsilon k}.$$

In view of Lemma 1, inequalities (1), and the inclusions $K \supset \sigma$ for all n , we have

$$N(K, S; \varepsilon) \geq \prod_{\varepsilon k \leq d_k} \frac{d_k}{\varepsilon k}.$$

* In a less precise form this inequality was noted in the note ⁽¹⁶⁾.

** In some details our computations are close to estimates of the ε -entropy of concrete compacta in the works of V. D. Erokhin ⁽¹⁵⁾.

This is the fundamental inequality. From it it follows that:

$$H(K, S; \varepsilon) \geq \sum_{\varepsilon k < d_k} \log \frac{d_k}{\varepsilon k} = \int_0^{1/\varepsilon} \log \frac{1}{\varepsilon t} dl(t) = \int_0^{1/\varepsilon} \frac{l(t)}{t} dt.$$

The theorem is proved.

Let us note that the inequalities of Theorem 1 are sharp in the sense that one can indicate compact sets for which the right- or left-hand inequality turns into an asymptotic equality. We also point out that, in fact, Theorem 2 of the note (10) has been proved above.

From the last inequality it follows, in particular, that

$$H(K, S; \varepsilon) \geq \int_{1/\varepsilon}^{1/\varepsilon} \frac{l(t)}{t} dt \geq l\left(\frac{1}{e\varepsilon}\right).$$

Using the fact that if $\sum \xi_n < \infty$ and $\xi_n \downarrow 0$, then $n\xi_n \rightarrow 0$, from Theorem 1 and Lemma 4 we obtain:

Lemma 5. *Let α be the convergence exponent of the sequence d_n ; then $\rho(K, S) \leq \alpha$. If $\rho(K, S) = \rho < 1$, then $\alpha \leq \rho/(1 - \rho)$; in particular, if $\rho(K, S) < 1/4$, then $\sum n^2 d_n < \infty$.*

3. Lemma 6* (Auerbach, (11), p. 205). *In every n -dimensional Banach space L_n there exists an "orthonormalized basis," i.e. systems of elements $\{g_i, i = 1, \dots, n\}$ and functionals $\{g'_i, i = 1, \dots, n\}$ such that $|g_i| = |g'_i| = 1$, $g'_i(g_j) = \delta_{ij}$ for $1 \leq i, j \leq n$.*

If L_n is an n -dimensional subspace in a Banach space B , and $\{g_i, g'_i\}$ is an Auerbach basis in L_n , then, extending g'_i to all of B with preservation of the norm, we obtain a linear operator

$$P_n, \quad P_{nx} = \sum_1^n g'_i(x)g_i,$$

projecting onto L_n , with norm $|P_n| \leq n$. The usual arguments (cf. (13), p. 193) show that

$$|x - P_{nx}| \leq (1 + n)r(x, L_n)$$

for every $x \in B$.

Definition 4 (see (3)). A compact set K is called **nuclear** with respect to S if there exists a system $\{f_i, i = 1, 2, \dots\}$ of linear functionals on $L(K)$ and a system $\{h_i, i = 1, 2, \dots\}$ of vectors in B such that

$$\sum_1^\infty |f_i|_K |h_i|_S < \infty$$

and for every $x \in K$ in B the representation

$$x = \sum_1^\infty f_i(x)h_i$$

holds. A linear topological space E is called **nuclear** if for every neighborhood U of zero in E there is a neighborhood V nuclear with respect to it.

The operators P_n constructed above make it possible to prove:

Lemma 7. *If for a compact set K the series $\sum n^2 d_n(K, S)$ converges, in particular, if $\rho(K, S) < 1/4$, then K is nuclear with respect to S .*

Definition 5 (cf. (1), § 2). To every linear topological space E we assign a class $\Psi(E)$ of functions defined for $\varepsilon > 0$ by the following condition: $\psi(\varepsilon) \in \Psi(E)$ if

$$\forall(\delta > 0) \forall U \exists V \left\{ \lim_{\varepsilon \rightarrow 0} \frac{N(V, U; \delta\varepsilon)}{\psi(\varepsilon)} = 0 \right\}^{**}.$$

The class $\Psi(E)$ has the properties of Kolmogorov linear dimension (see (1), § 1) and, as a theorem of Bessaga and Pełczyński shows (see (14), Theorem 1), in the case of a metrizable space $\Psi(E)$ coincides with the class $\Phi(E)$ of functions defining approximative

* For its proof see (12).

** $\forall(\delta > 0)$ means “for every $\delta > 0$,” etc.; $\exists V$ means “there exists such a neighborhood V of zero that ...” .

dimension in the sense of Kolmogorov (see ⁽¹⁾, § 2). From Lemmas 7 and 2 (cf. ⁽⁷⁾), where Theorem 2 is proved for some metrizable spaces) it follows:

Theorem 2 (criterion of nuclearity). For the nuclearity of a linear topological space E it is necessary that $\exp \varepsilon^{-\lambda} \in \Psi'(E)$ for all $\lambda > 0$, and sufficient that $\exp \varepsilon^{-\lambda_0} \in \Psi'(E)$ for some $\lambda_0 > 0$.

Another, as follows from all that has been said above, equivalent formulation of the criterion is possible:

Theorem 2a. For the nuclearity of the space E it is necessary that

$$\forall(\lambda > 0) \forall U \exists V \{n^\lambda d_n(V, U) \rightarrow 0\},$$

and sufficient that

$$\exists(\lambda_0 > 0) \forall U \exists V \{n^{\lambda_0} d_n(V, U) \rightarrow 0\}.$$

In conclusion I express my gratitude to M. A. Krasnosel' skii, who pointed out Auerbach' s lemma to me, and to V. Erokhin and V. Tikhomirov for useful discussions and remarks.

Moscow State University
named after M. V. Lomonosov

Received
20 IV 1960

CITED LITERATURE

- ¹ A. N. Kolmogorov, DAN, 120, No. 2, 239 (1958).
- ² A. Pełczyński, Bull. Acad. Pol. Sci., 5, No. 9, 879 (1957).
- ³ A. Grothendieck, Mem. Am. Math. Soc., No. 16 (1959); D. A. Raikov, UMN, 12, No. 5, 231 (1957).
- ⁴ I. M. Gel' fand, UMN, 11, No. 6, 3 (1956).
- ⁵ A. N. Kolmogorov, DAN, 108, No. 3, 585 (1956).
- ⁶ A. N. Kolmogorov, V. M. Tikhomirov, UMN, 14, No. 2, 3 (1959).
- ⁷ A. Dynin, V. Mitiagin, Bull. Acad. Pol. Sci., 8 (1960).
- ⁸ A. Kolmogoroff, Ann. Math., 37, 107 (1936); V. M. Tikhomirov, DAN, 130, No. 4, 734 (1960).
- ⁹ B. Ya. Levin, *Distribution of Zeros of Entire Functions*, Moscow, 1956.
- ¹⁰ B. S. Mityagin, UMN, 15 (1960).
- ¹¹ S. Banach, *Course of Functional Analysis*, Kiev, 1948.
- ¹² M. Day, Trans. Am. Math. Soc., 62, 315 (1947); A. Taylor, Bull. Am. Math. Soc., 53, 614 (1947).

¹³ I. P. Natanson, *Constructive Theory of Functions*, Moscow-Leningrad, 1949.

¹⁴ S. Rolewicz, *Bull. Acad. Pol. Sci.*, 7, 195 (1959).

¹⁵ V. D. Erokhin, *DAN*, 120, No. 4, 689; 120, No. 5, 949 (1958).

¹⁶ Yu. A. Brudnyi, A. F. Timan, *DAN*, 126, No. 5 (1959).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.