APPROXIMATION THEORY:
A volume dedicated to Blagovest Sendov
(B. Bojanov, Ed.), DARBA, Sofia, 2002, pp. 241-257.

On Lower Estimates for n-term Approximation in Hilbert Spaces

B. S. KASHIN

1. Introduction

Let B be a real normed space, let $\Phi \subset B$ be a subset of B — we will call it a *dictionary*, and let f be any element of B.

Definition 1. *n-term approximation* of an element $f \in B$ with respect to the dictionary Φ is

$$e_n(f, \Phi, B) \equiv \inf_{P \in \Sigma_n} \|f - P\|_B, \tag{1}$$

where for $n = 1, 2, \dots$

$$\Sigma_n \equiv \left\{ \sum_{j=1}^{n} a_j x_j,\ a_j \in \mathbb{R},\ x_j \in \Phi \right\}.$$

Further, if K is a subset of B, then

$$e_n(K, \Phi, B) \equiv \sup_{f \in K} e_n(f, \Phi, B). \tag{2}$$

At present, research in n-term approximations has become a separate branch of approximation theory. The best-investigated situation is when K is some class of smooth functions defined on the domain Ω in $\mathbb{R}^d$ and $B = L^p(\Omega)$, $1 \le p \le \infty$. Various dictionaries Φ have been considered. Here are the most important examples:

a) Φ is a complete orthonormal set of functions (O.N.S.) in $L^2(\Omega)$; in particular, Φ can be a trigonometric system;

b) for the approximation of functions of d variables, defined on the cube $I^d = (0,1)^d$, we can consider the dictionary Φ_k that consists of all functions of the type

$$u(x_1, \dots, x_k) \cdot v(x_{k+1}, \dots, x_d), \qquad 1 \le k < d;$$

c) for the approximation of functions of d variables by "free knots splines" we can consider dictionaries that consist of functions of the type

$$P \cdot \chi_\Delta,$$

where P is a polynomial of the degree $\leq r$ of d variables and χ_Δ is the characteristic function of the segment $\Delta \subset \mathbb{R}^d$;

d) Φ is the set of ridge functions, i.e., functions in $L^p(\Omega)$, $\Omega \subset \mathbb{R}^d$, of the type

$$u(x) = f(\langle x, \theta \rangle),$$

where f is a function of one variable, $x \in \Omega$, $\theta \in \mathbb{R}^d$, $|\theta| = 1$, and $\langle \cdot , \cdot \rangle$ is an inner product in $\mathbb{R}^d$.

For each of those families of dictionaries there are research results that are valuable from the theoretical as well as practical point of view (see [1], [2], [3]).

Research related to the example a), i.e., the situation when Φ is a complete O.N.S., has become more active due to the development of wavelet theory. As to Definition 1, for orthonormal dictionaries in Hilbert space H it was introduced back in 1955 by S. B. Stechkin [4] when he was investigating the absolute convergence of series with respect to general complete O.N.S. Note that in this case

$$e_n(f, \Phi, H) = \left(\sum_{k \geq n+1} [c_k^*(f)]^2 \right)^{1/2}, \tag{3}$$

where $\{c_k^*(f)\}$ is a non-increasing rearrangement of the sequence of absolute values of the Fourier coefficients of the function f with respect to the complete O.N.S. Φ.

The author in [5] suggested a geometric approach to the proof of lower estimates for values (2); this approach can be applied to any orthonormal dictionary Φ in a Hilbert space H. More precisely, it was shown in [5] that if for some $n \in \mathbb{N}$ K contains the set Q of all vertices of a $2n$-dimensional cube:

$$Q = \left\{ \sum_{i=1}^{2n} \varepsilon_i \psi_i, \ \varepsilon_i = \pm 1, \ \{\psi_i\}_{i=1}^{2n} \text{ is an O.N.S.} \right\}, \tag{4}$$

then

$$e_n(K, \Phi, H) \geq c \cdot n^{1/2}, \qquad c > 0. \tag{5}$$

To apply this result we need to inscribe a big enough cube in a given set K (i.e., for a given n we need to find a big enough number λ and a set Q of type (4) such that $\lambda \cdot Q \subset K$). This problem is not difficult to solve for classical function classes K and it allows us in certain cases to obtain order-sharp lower estimates for n-term approximations. Estimates analogous to (5) and its generalizations were obtained in [6], [7].

In 1993 S. V. Konyagin suggested a problem of estimating the values (2) when $B = L^2(I^d)$, Φ was an O.N.S., and K was the family of characteristic functions of convex subsets of the unit cube $I^d \subset \mathbb{R}^d$. This problem seems natural from both theoretical and practical point of view. Even for $d = 1$ it remained unsolved. In that case (i.e., $d = 1$), the problem essentially can be reduced to the estimating quantities (2) for "one-parametric" family of functions

$$K = \mathbb{X} \equiv \{\chi_t\}_{t\in[0,1]}, \qquad \chi_t(x) = \begin{cases} 0, & \text{if } 0 \le x < t, \\ 1, & \text{if } t \le x \le 1. \end{cases} \tag{6}$$

S. V. Konyagin brought to the author's attention the problem of lower estimates for n-term approximation of the family (6); he also noted that there were exponential upper estimates for those values. More precisely, if Φ_0 is the Haar system, then

$$e_n(\mathbb{X}, \Phi_0, L^2(0,1)) \le C \cdot 2^{-n/2}. \tag{7}$$

This estimate can be checked by applying a standard error estimate for L^2-approximation of functions χ_t $(0 \le t \le 1)$ by partial sums of Fourier–Haar series (see, for example, [8], p. 75) noticing that each dyadic block of the Fourier–Haar series of χ_t has at most one non-zero coefficient.

The family $\mathbb{X}$ is very "thin". That is why the above mentioned geometric approach to lower estimates for n-term approximation is not applicable to this set. It turned out that instead we can use a technique from the theory of general orthogonal series.

In this article, using the approach from author's paper [9], we establish the following two theorems.

Theorem 1. *There exists an absolute constant $C > 0$ such that for each $n = 1, 2, \dots$ and for any orthonormal system $\Phi \subset L^2(0,1)$*

$$e_n(\mathbb{X}, \Phi, L^2(0,1)) \ge C^{-n}.$$

Theorem 2. *If Φ is a complete uniformly bounded orthonormal system:*

$$\Phi = \{\varphi_j\}_{j=1}^{\infty} \subset L^2(0,1),$$
$$\|\varphi_j\|_{L^\infty(0,1)} \le M, \quad j = 1, 2, \dots,$$

then for each $n = 1, 2, \dots$

$$e_n(\mathbb{X}, \Phi, L^2(0,1)) \ge \frac{C_M}{n^{1/2}} > 0. \tag{8}$$

Remark 1. The exact value of the constant C in Theorem 1 is unknown.

Remark 2. Theorem 2 shows that the uniform boundedness condition for the orthonormal dictionary Φ substantially changes (compared with the general case) the behavior of n-term approximations of functions from the family $\mathbb{X}$. The sharpness of estimate (8) is demonstrated by the example of the trigonometric system T : if $\Phi = T$, then

$$e_n(\mathbb{X}, T, L^2(0,1)) \le C \cdot n^{-1/2}.$$

The results of this paper are announced in the note [10].

2. Proof of Theorem 1

Below we denote by $\#\Lambda$ the cardinality (the number of elements) of a set Λ.

Lemma 1. *Suppose $N \in \mathbb{N}$, $\Lambda \subset \mathbb{N}$, $\#\Lambda < \infty$, $\{\varphi_j\}_{j\in\Lambda}$ is a normalized system of functions in $L^2(0,1)$ (i.e. $\|\varphi_j\|_{L^2(0,1)} = 1 \ \forall j \in \Lambda$), and for each $k = 1, 2, \dots, N$*

$$\chi_{k/N} = \sum_{j\in\Lambda} a_{k,j}\varphi_j + \Delta_k, \qquad \|\Delta_k\|_{L^1} \le \frac{1}{N},$$

where χ_t is defined in (6). *Then*

$$\sum_{j\in\Lambda}\left(\frac{1}{N}\sum_{k=1}^{N} a_{k,j}^2\right)^{1/2} \ge \frac{1}{2\pi}(\ln N - 7) \ge \frac{\log_2 N}{3\pi} - \frac{7}{2\pi}.$$

Lemma 1, up to the constants, is a special case of Theorem 1 of the paper [9]. Still, let us prove it for the sake of completeness. For each $N > 2$ we define a system of functions $\{f_k^N\}_{k=1}^N$ on $(0,1)$ by the following formula:

$$f_k^N(x) = \begin{cases} \dfrac{1}{i-k}, & \text{if } x \in \left(\dfrac{i-1}{N}, \dfrac{i}{N}\right), \ i \ne k, \ 1 \le i \le N, \\ 0, & \text{if } x \in \left(\dfrac{k-1}{N}, \dfrac{k}{N}\right). \end{cases}$$

Using the classical inequality for a Hilbert bilinear form

$$\left|\sum_{k=1}^{N}\sum_{i=1, i\ne k}^{N} \frac{a_k b_i}{i-k}\right| \le \pi\left(\sum_{k=1}^{N} a_k^2\right)^{1/2} \cdot \left(\sum_{i=1}^{N} b_i^2\right)^{1/2}$$

(see [11], p. 256, th. 294) for any coefficients $\{a_k\}_{k=1}^N$, we get

$$\begin{aligned} \left\|\sum_{k=1}^N a_k f_k^N\right\|_{L^2(0,1)} &= \frac{1}{N^{1/2}}\left[\sum_{i=1}^{N}\left(\sum_{k=1}^{N} a_k f_k^N\left(\frac{i-1/2}{N}\right)\right)^2\right]^{1/2} \\ &= \frac{1}{N^{1/2}} \sup_{\sum_{i=1}^N b_i^2 = 1} \sum_{i=1}^{N} b_i\left(\sum_{k\ne i}\frac{a_k}{i-k}\right) \qquad (9) \\ &\le \pi\left(\frac{\sum_{j=1}^N a_j^2}{N}\right)^{1/2}. \end{aligned}$$

Now consider the integral

$$\int_0^1 \sum_{k=1}^N \chi_{k/N}(x) f_k^N(x)\,dx = \sum_{k=1}^N \int_0^1 \chi_{k/N}\cdot f_k^N\,dx \geq \sum_{k=1}^{N-1} \frac{1}{N} \sum_{i=k+1}^N \frac{1}{i-k}$$
$$= \frac{1}{N}\sum_{k=1}^{N-1}\sum_{r=1}^{N-k}\frac{1}{r} \geq \frac{1}{N}\sum_{k=1}^{[N/2]}\sum_{r=1}^{[N/2]}\frac{1}{r} \geq \frac{1}{2}(\ln N - 5).$$

But
$$\sum_{k=1}^N \chi_{k/N} f_k = \sum_{k=1}^N \Big[\sum_{j\in\Lambda} a_{k,j}\varphi_k + \Delta_k\Big] f_k$$
and for each $k = 1, 2, \ldots, N$ $\|f_k\|_{L^\infty} \leq 1$; therefore
$$\sum_{k=1}^N \left|\int_0^1 f_k^N \Delta_k\,dx\right| \leq \sum_{k=1}^N \|\Delta_k\|_{L^1} \leq 1.$$

Thus
$$\begin{aligned}\frac{1}{2}(\ln N - 7) &\leq \sum_{k=1}^N \int_0^1 f_k^N(x)\Big(\sum_{j\in\Lambda} a_{k,j}\varphi_j(x)\Big)dx\\ &= \sum_{j\in\Lambda}\int_0^1 \varphi_j(x)\Big(\sum_{k=1}^N a_{k,j} f_k^N(x)\Big)dx\\ &\leq \sum_{j\in\Lambda}\Big(\int_0^1 \varphi_j^2\,dx\Big)^{1/2}\Big(\int_0^1\Big(\sum_{k=1}^N a_{k,j} f_k^N\Big)^2 dx\Big)^{1/2}.\end{aligned}$$

By (9) and since the system $\{\varphi_j\}$ is normalized, the latter sum is bounded from above by
$$\pi \sum_{j\in\Lambda}\Big(\frac{1}{N}\sum_{k=1}^N a_{k,j}^2\Big)^{1/2}.$$
This completes the proof of Lemma 1.

Lemma 2. *Let f be an absolutely continuous function on $(0,1)$ such that $\|f'\|_{L^2(0,1)} \leq 1$. If for some $N \in \mathbb{N}$*
$$\Big(\frac{1}{N}\sum_{k=1}^N f^2\Big(\frac{k}{N}\Big)\Big)^{1/2} \equiv \varepsilon \geq \frac{3}{N},$$
then
$$\frac{2}{3}\varepsilon \leq \|f\|_{L^2(0,1)} \leq \frac{4}{3}\varepsilon.$$

Proof. First we shall show that for each $\tau \in (0, 1/N)$

$$S = S_\tau \equiv \left\{ \frac{1}{N} \sum_{k=1}^{N} \left[f\left(\frac{k}{N}\right) - f\left(\frac{k}{N} - \tau\right) \right]^2 \right\}^{1/2} \leq \frac{1}{N}. \tag{10}$$

Indeed,

$$\left| f\left(\frac{k}{N}\right) - f\left(\frac{k}{N} - \tau\right) \right| = \left| \int_{k/N-\tau}^{k/N} f'(x)\, dx \right| \leq \tau^{1/2} \left(\int_{k/N-\tau}^{k/N} (f')^2\, dx \right)^{1/2};$$

therefore,

$$S^2 \leq \frac{\tau}{N} \sum_{k=1}^{N} \int_{k/N-\tau}^{k/N} (f')^2\, dx \leq \frac{1}{N^2},$$

which proves the estimate (10). Then

$$\begin{aligned} \|f\|^2_{L^2(0,1)} &= \sum_{k=1}^{N} \int_{(k-1)/N}^{k/N} f^2(x)\, dx = \sum_{k=1}^{N} \int_0^{1/N} f^2\left(\frac{k}{N} - \tau\right) d\tau \\ &= N \int_0^{1/N} \left[\frac{1}{N} \sum_{k=1}^{N} f^2\left(\frac{k}{N} - \tau\right) \right] d\tau = N \int_0^{1/N} \|v_\tau\|^2_{L^2_N}\, d\tau, \end{aligned}$$

where $v_\tau \in \mathbb{R}^N$, $(v_\tau)_k = f(k/N - \tau)$, $k = 1, 2, \ldots, N$, and for any vector $v = \{(v)_k\}_{k=1}^N \in \mathbb{R}^N$

$$\|v\|_{L^2_N} \equiv \left(\frac{1}{N} \sum_{k=1}^{N} (v)_k^2 \right)^{1/2}.$$

Therefore, we get

$$\|f\|^2_{L^2} = N \int_0^{1/N} \|v_\tau\|^2_{L^2_N}\, d\tau; \tag{11}$$

besides, by (10),

$$\|v_0 - v_\tau\|_{L^2_N} \leq \frac{1}{N}, \qquad 0 < \tau < \frac{1}{N},$$

and by assumption, $\|v_0\|_{L^2_N} \geq 3/N$. Using the triangle inequality, for each $\tau \in (0, 1/N)$ we get

$$\frac{2}{3}\|v_0\|_{L^2_N} \leq \|v_0\|_{L^2_N} - \frac{1}{N} \leq \|v_\tau\|_{L^2_N} \leq \|v_0\|_{L^2_N} + \frac{1}{N} \leq \frac{4}{3}\|v_0\|_{L^2_N};$$

that is, for each $\tau \in (0, 1/N)$ we have

$$\frac{4}{9}\|v_0\|^2_{L^2_N} \leq \|v_\tau\|_{L^2_N} \leq \frac{16}{9}\|v_0\|^2_{L^2_N},$$

and Lemma 2 now follows from (11) and the definition of the vector v_0.

Lemma 3. *Suppose f is an absolutely continuous function on $(0,1)$, $f(1)=0$, $\|f'\|_{L^2(0,1)} \le 1$ and $\|f\|_{L^2(0,1)} = \delta$. Then there exists an interval $\omega \subset (0,1)$ such that its measure $|\omega| \ge \frac{1}{2}\delta$ and*

$$|f(x)| \ge \frac{1}{3}\delta \qquad \forall x \in \omega.$$

Proof. We assume that $\delta > 0$. Consider

$$E = \left\{x \in (0,1) : |f(x)| > \frac{1}{3}\delta\right\}.$$

Clearly, E is an open set, and it can be written in the form

$$E = \bigcup \omega_i,$$

where ω_i are non-overlapping intervals. The set E is not empty and, moreover,

$$\int_E f^2\,dx = \int_{(0,1)} f^2\,dx - \int_{(0,1)\setminus E} f^2\,dx \ge \delta^2 - \frac{1}{9}\delta^2 = \frac{8}{9}\delta^2. \tag{12}$$

Let z_i be the right-hand endpoint of the interval ω_i. Then $|f(z_i)| = \frac{1}{3}\delta$. Further, for each $x \in E$ there exists i such that $x \in \omega_i$, and then

$$|f(x) - f(z_i)| = \left|\int_x^{z_i} f'(u)\,du\right| \le |\omega_i|^{1/2}\left(\int_{\omega_i} |f'(u)|^2\,du\right)^{1/2}.$$

Hence

$$|f(x)| \le \frac{1}{3}\delta + |\omega_i|^{1/2}\left(\int_{\omega_i} |f'(u)|^2\,du\right)^{1/2}, \qquad x \in \omega_i,$$

therefore

$$|f(x)|^2 \le \frac{2}{9}\delta^2 + 2|\omega_i|\int_{\omega_i} |f'(u)|^2\,du, \qquad x \in \omega_i,$$

and (see (12))

$$\begin{aligned}
\frac{8}{9}\delta^2 &\le \int_E f^2\,dx = \sum_i \int_{\omega_i} |f|^2\,dx \\
&\le \sum_i |\omega_i|\left[\frac{2}{9}\delta^2 + 2|\omega_i|\int_{\omega_i} |f'(u)|^2\,du\right] \\
&\le \frac{2}{9}\delta^2 + \sum_i 2|\omega_i|^2 \int_{\omega_i} |f'(u)|^2\,du.
\end{aligned}$$

By definition, put $\lambda_i = \displaystyle\int_{\omega_i} |f'(u)|^2\,du$.

Then the previous inequality yields

$$\frac{1}{3}\delta^2 \le \sum_i |\omega_i|^2 \lambda_i. \tag{13}$$

Since

$$\sum_i \lambda_i \le \int_0^1 |f'(u)|^2\, du \le 1,$$

by (13) we get that there exists an interval ω_i such that $|\omega_i|^2 \ge \frac{1}{3}\delta^2$, but in that case we have $|\omega_i| \ge \frac{1}{\sqrt{3}}\delta > \frac{1}{2}\delta$. This completes the proof of Lemma 3.

Now we can start the proof of Theorem 1. Obviously, we can complete the system Φ if necessary, thus we can assume that Φ is a complete O.N.S. It is also clear that it suffices to consider only the case when $n \ge n_0$. Let $n \ge 4$ be given and

$$N = 2^{60\,n}. \tag{14}$$

It is easy to check that in this case

$$\frac{\log_2 N}{3\pi} - \frac{7}{2\pi} \ge \frac{\log_2 N}{10}. \tag{15}$$

For the given O.N.S. Φ we obviously have one of the following two cases:

$$1) \qquad \max_{1\le k\le N} e_n(\chi_{k/N}, \Phi, L^2) \ge \frac{1}{2N} > \frac{1}{2^{61\,n}};$$

$$2) \qquad \max_{1\le k\le N} e_n(\chi_{k/N}, \Phi, L^2) < \frac{1}{2N}.$$

Let us prove that in case 2)

$$\max_{1\le k\le N} e_n(\chi_{k/N}, \Phi, L^2) \ge \frac{1}{N^{1+1/100}} \ge \frac{1}{2^{61\,n}}. \tag{16}$$

Thus, for each $n \ge 4$ the estimate in the Theorem 1 will be established if we take $C = 2^{61}$.

Now, if condition 2) holds, then for each $k = 1, 2, \ldots, N$,

$$\chi_{k/N}(x) = P_\Phi^{k,N}(x) + \Delta_{k,N}(x),$$

where $P_\Phi^{k,N}(x)$ is a polynomial with respect to the system Φ such that the number of its non-zero coefficients is $\le n$ and

$$\|\Delta_{k,N}\|_{L^2(0,1)} \le \frac{1}{2N}, \qquad k = 1, 2, \ldots, N. \tag{17}$$

Suppose

$$P_\Phi^{k,N}(x) = \sum_{j\in E_{k,N}} a_{k,j}\varphi_j(x), \qquad \#E_{k,N} \le n. \tag{18}$$

Clearly, we can assume below that

$$a_{k,j} = \int_0^1 \chi_{k/N}\varphi_j\,dx = \int_{k/N}^1 \varphi_j\,dx$$

is a Fourier coefficient of the function $\chi_{k/N}$. This follows directly from the extremal properties of the Fourier coefficients.

By definition, put

$$\mathbb{N} \supset \Lambda \equiv \bigcup_k \mathrm{spectrum} P_\Phi^{k,N},$$

where $\mathrm{spectrum} P_\Phi = \{j : a_j \neq 0\}$ if $P_\Phi = \sum_{j\in E} a_j\varphi_j$ is a polynomial with respect to the system Φ. Clearly, $\#\Lambda \leq n\cdot N$. Now let us consider

$$\Lambda' = \left\{ j \in \Lambda : \max_{1\leq k\leq N} |a_{k,j}| = \max_{1\leq k\leq N}\left|\int_{k/N}^1 \varphi_j\,dx\right| \geq \frac{1}{2N\cdot n^{1/2}} \right\}$$

and

$$L' = \mathrm{span}\{\varphi_j,\ j\in\Lambda'\}$$

— a subspace generated by the functions φ_j, $j\in\Lambda'$. Then for each $k = 1,2,\ldots,N$ for the polynomial

$$\widetilde{P}^{k,N} = \sum_{j\in\Lambda'} a_{k,j}\varphi_j$$

we have

$$\chi_{k,N} = \widetilde{P}^{k,N} + \widetilde{\Delta}_{k,N}, \qquad \|\widetilde{\Delta}_{k,N}\|_{L^2} \leq \frac{1}{N}. \tag{19}$$

Indeed,

$$\chi_{k,N} = \widetilde{P}^{k,N} + \sum_{j\in E_{k,N}\setminus\Lambda'} a_{k,j}\varphi_j + \Delta_{k,N},$$

but by the very definition of Λ',

$$\left\| \sum_{j\in E_{k,N}\setminus\Lambda'} a_{k,j}\varphi_j \right\|_{L^2} \leq \left[n\frac{1}{(2N)^2 n} \right]^{1/2} = \frac{1}{2N}$$

from which, using also (17), we get (19).

By Lemma 1 (see also (15)), it follows from (19) that

$$\sum_{j\in\Lambda'} \varepsilon_j \geq \frac{1}{10}\log_2 N, \tag{20}$$

where

$$\varepsilon_j = \|\{a_{k,j}\}\|_{L_N^2} \equiv \left(\frac{1}{N}\sum_{k=1}^N a_{k,j}^2\right)^{1/2}.$$

Let us show that

$$\left.\begin{array}{l}\text{there exists a number } k_0 \in \{1,\dots,N\} \text{ such that the set} \\ \text{of coefficients } a_{k_0,j},\ j \in \Lambda', \text{ contains more than } n \\ \text{coefficients whose absolute value is greater than } \dfrac{1}{N^{101/100}}.\end{array}\right\} \tag{21}$$

The required relation (16) follows directly from (21).

By definition, for $K \subset \{1,\dots,N\}$ put

$$\mu(K) = \frac{\#K}{N}.$$

The relation (21) will be proved if we check that

$$\sum_{j\in\Lambda'} \mu\left\{k : |a_{k,j}| > \frac{1}{N^{1.01}}\right\} > n. \tag{22}$$

We consider the decomposition

$$\Lambda' = \Lambda'' \cup \Lambda''',$$

where

$$\Lambda'' = \left\{j \in \Lambda' : \varepsilon_j > \frac{6}{N}\right\}, \qquad \Lambda''' = \Lambda' \setminus \Lambda''.$$

By the definition of Λ' and the inequality $N^{1.01} \geq 2N \cdot n^{1/2}$ (see (14)), we obtain that for each $j \in \Lambda'''$

$$\left\{k : |a_{k,j}| \geq \frac{1}{N^{1.01}}\right\} \neq \emptyset;$$

therefore, for each $j \in \Lambda'''$,

$$\mu\left\{k : |a_{k,j}| \geq \frac{1}{N^{1.01}}\right\} \geq \frac{1}{N} \geq \frac{1}{6}\varepsilon_j$$

and, thus,

$$\sum_{j\in\Lambda'''} \mu\left\{k : |a_{k,j}| \geq \frac{1}{N^{1.01}}\right\} \geq \frac{1}{6} \sum_{j\in\Lambda'''} \varepsilon_j. \tag{23}$$

On the other hand, if $j \in \Lambda''$, then $\|\{a_{k,j}\}\|_{L^2_N} = \varepsilon_j > 6/N$ and the application of Lemma 2 to the function

$$f(x) = \int_x^1 \varphi_j(t)\,dt$$

yields

$$\|f\|_{L^2(0,1)} \equiv \left\|\int_x^1 \varphi_j(t)\,dt\right\|_{L^2(0,1)} \geq \frac{2}{3}\varepsilon_j > \frac{4}{N}.$$

Therefore, by Lemma 3, there exists an interval $\omega \subset (0,1)$ such that its measure is

$$|\omega| \geq \frac{1}{2}\|f\|_{L^2(0,1)} > \frac{2}{N}$$

and for each $x \in \omega$

$$\left|\int_x^1 \varphi_j(t)\,dt\right| \geq \frac{1}{3}\|f\|_{L^2(0,1)} \geq \frac{2}{9}\varepsilon_j. \tag{24}$$

Note that the lower estimate for the number of points of the form k/N, $1 \leq k \leq N$, within the interval ω is $|\omega| \cdot N - 1$, so we can see that for each $j \in \Lambda''$ this number is not less than $\dfrac{N}{3}\varepsilon_j - 1 > \dfrac{N}{6}\varepsilon_j$.

Hence

$$\sum_{j\in\Lambda''} \mu\{k : |a_{k,j}| \geq N^{-1.01}\} \geq \sum_{j\in\Lambda''} \mu\left\{k : |a_{k,j}| \geq \frac{2}{9}\varepsilon_j\right\} > \sum_{j\in\Lambda''} \frac{\varepsilon_j}{6} \tag{25}$$

(we used the estimate $N^{-1.01} \leq N^{-1} \leq \frac{2}{9}\varepsilon_j$ for $j \in \Lambda''$).

Finally, we have (see (23), (25))

$$\sum_{j\in\Lambda'} \mu\{k : |a_{k,j}| \geq N^{-1.01}\} = \sum_{j\in\Lambda''} + \sum_{j\in\Lambda'''} > \sum_{j\in\Lambda''} \frac{1}{6}\varepsilon_j + \sum_{j\in\Lambda'''} \frac{1}{6}\varepsilon_j = \frac{1}{6}\sum_{j\in\Lambda'} \varepsilon_j.$$

The right-hand side of this relation, by (20) and (14), is ~~not less~~ bigger than n. Thus the inequality (22) is established, and this completes the proof of Theorem 1.

3. Proof of Theorem 2

The considerations below are in some sense similar to the proof of the estimate (9) in [9]. Let us check that for each O.N.S. Φ that satisfies the conditions of Theorem 2, for each $n \in \mathbb{N}$ and $N = [2000\,M^2 n]$ the following inequality holds

$$\max_{1\leq k\leq N} e_n\big(\chi_{k/N}, \Phi, L^2(0,1)\big) \geq \frac{1}{1000\,M} \cdot \frac{1}{n^{1/2}}. \tag{26}$$

It is clear that Theorem 2 follows from (26). In order to prove the estimate (26) we define (for $N > 10$) the system of functions $\{f_k^N\}_{k=1}^N$ on the segment $[0,1]$ such that these functions are constant on each interval $\left(\dfrac{r-1}{N}, \dfrac{r}{N}\right)$, $1 \leq r \leq N$, and

$$f_k^N\left(\frac{r-1/2}{N}\right) = \begin{cases} \dfrac{1}{2}, & \text{if } 2 \leq r-k \leq 3,\ 4 < r \leq N-4, \\ -\dfrac{1}{2}, & \text{if } 2 \leq k-r \leq 3,\ 4 < r \leq N-4, \\ 0 & \text{for other } k,\ r. \end{cases} \tag{27}$$

It follows from definition (27) that

$$\begin{aligned}
&\text{a)}\ \ f_k^N(x) = 0 \text{ when } 4 \le k \le N-4 \text{ and } x \notin \left[\frac{k-4}{N}, \frac{k+3}{N}\right];\\
&\text{b)}\ \ f_k^N(x) \ge 0 \text{ when } x \ge \frac{k}{N};\\
&\text{c)}\ \ \sum_{k=1}^{N} f_k^N \equiv 0;\\
&\text{d)}\ \ \|f_k^N\|_{L^1} \le \frac{2}{N}, \ \|f_k^N\| \le \frac{1}{N^{1/2}}.
\end{aligned} \tag{28}$$

For $k = 1, 2, \ldots, N$ put

$$a_{k,j} = \int_0^1 \chi_{k/N} \varphi_j \, dx, \qquad j = 1, 2, \ldots.$$

Note that

$$|a_{k,j} - a_{k+1,j}| \le \frac{M}{N}, \qquad 1 \le k \le N-1, \tag{29}$$

which follows from the estimate

$$|a_{k,j} - a_{k+1,j}| = \left| \int_0^1 \varphi_j (\chi_{k/N} - \chi_{(k+1)/N}) \, dx \right| = \left| \int_{k/N}^{(k+1)/N} \varphi_j \, dx \right| \le \frac{M}{N}.$$

From the definitions and properties of the functions f_k^N, (see (27), (28)) it follows directly that

$$\int_0^1 \sum_{k=1}^{N} \chi_{k/N} f_k^N \, dx \ge \sum_{k=4}^{N-8} \int_{k/N}^1 f_k^N(x) \, dx \ge \frac{N-12}{N} \ge \frac{1}{2} \tag{30}$$

(we used here that $N \ge 2000$).

Thus

$$\begin{aligned}
\frac{1}{2} &\le \int_0^1 \sum_{k=1}^{N} \chi_{k/N} f_k^N \, dx \\
&= \int_0^1 \sum_{k=1}^{N} f_k^N \Big(\sum_{j=1}^{\infty} a_{k,j} \varphi_j \Big) \, dx \\
&= \sum_{j=1}^{\infty} \sum_{k=1}^{N} a_{k,j} \int_0^1 f_k^N \varphi_j \, dx \equiv R_1 + R_2,
\end{aligned} \tag{31}$$

where

$$R_1 = \sum_{j=1}^{\infty} \int_0^1 \varphi_j \Big(\sum_{k:\, |a_{k,j}| > 2M/N} a_{k,j} f_k^N \Big) \, dx, \tag{32}$$

$$R_2 = \sum_{k=1}^{N} \sum_{j:\,|a_{k,j}|\le 2M/N} a_{k,j} \int_0^1 \varphi_j f_k^N \, dx. \tag{33}$$

Now let us estimate R_1 and R_2. For $j = 1, 2, \ldots$, consider

$$F_j(x) \equiv \sum_{k:\,|a_{k,j}|>2M/N} a_{k,j} f_k^N. \tag{34}$$

Then (using the estimate $|\varphi_j(x)| \le M$, $j = 1, 2, \ldots$) we obtain

$$R_1 \le M \sum_{j=1}^{\infty} \int_0^1 |F_j(x)| \, dx. \tag{35}$$

Now let us estimate separately each term in the sum in the right-hand side of (35). By definition, for each $r \in \{1, \ldots, N\}$ put $\underline{r} = \max\{1, r-4\}$, $\overline{r} = \min\{N, r+4\}$. Then

$$\int_0^1 |F_j(x)| \, dx = \int_{E_1^j} |F_j| \, dx + \int_{E_2^j} |F_j| \, dx,$$

where

$$E_1^j = \left\{ x : x \in \left(\frac{r-1}{N}, \frac{r}{N} \right), \ \max_{\underline{r} \le k \le \overline{r}} |a_{k,j}| > \frac{16\,M}{N} \right\},$$

$$E_2^j = (0,1) \setminus E_1^j.$$

Since (see (27))

$$f_k^N \left(\frac{r-1/2}{N} \right) = 0 \quad \text{when} \quad k < \underline{r} \quad \text{or} \quad k > \overline{r} \tag{36}$$

we see that by definition of E_2^j and (28d),

$$\begin{aligned} \int_{E_2^j} |F_j| \, dx &\le \int_0^1 \sum_{k:\,(2M)/N \le |a_{k,j}| \le (16M)/N} |a_{k,j}| \, |f_k^N(x)| \, dx \\ &\le \frac{2}{N} \sum_{k:\,(2M)/N \le |a_{k,j}| \le (16M)/N} |a_{k,j}|. \end{aligned} \tag{37}$$

Further, the definition of E_1^j yields

$$\min_{\underline{r} \le k \le \overline{r}} |a_{k,j}| > \frac{8M}{N} \quad \text{if} \quad \frac{r-1/2}{N} \in E_1^j. \tag{38}$$

Indeed, for $x = \dfrac{r-1/2}{N} \in E_1^j$,

$$\max_{\underline{r} \le k \le \overline{r}} |a_{k,j}| > \frac{16\,M}{N}.$$

Therefore (see (29))

$$\min_{\underline{r}\le k\le \overline{r}} |a_{k,j}| > \frac{16\,M}{N} - \frac{8M}{N} = \frac{8M}{N}\,,$$

and (38) is established. Using (36) and (38) we get that for $x = \dfrac{r-1/2}{N} \in E_1^j$

$$F_j(x) = \sum_{k=1}^{N} a_{k,j} f_k^N(x) = \sum_{k=\underline{r}}^{\overline{r}} a_{k,j} f_k^N(x) = \sum_{k=\underline{r}}^{\overline{r}} (a_{k,j} - a_{\underline{r},j}) f_k^N(x) + a_{\underline{r},j} \sum_{k=\underline{r}}^{\overline{r}} f_k^N(x).$$

The latter sum is equal to zero by (28c) and (28a). Besides (see (29)),

$$|a_{k,j} - a_{\underline{r},j}| \le \frac{8M}{N} \quad \text{if} \quad \underline{r} \le k \le \overline{r}, \quad r = 1,2,\ldots,N,$$

and $\|f_k^N\|_{L^\infty} \le \frac{1}{2}$. Hence, for each point $x = \dfrac{r-1/2}{N} \in E_1^j$ there is an estimate

$$|F_j(x)| \le \frac{8M}{N} \cdot \frac{1}{2} \cdot \sum_{\underline{r}\le k\le \overline{r}} 1$$

and, therefore (see (38)),

$$|F_j(x)| \le \frac{4M}{N} \sum_{\{k\in\{\underline{r},\ldots,\overline{r}\}:\, |a_{k,j}|\ge 8M/N\}} 1. \tag{39}$$

It follows from (39) that

$$\begin{aligned} \int_{E_1^j} |F_j(x)|\,dx &\le \sum_{r:\,(r-1/2)/N\in E_1^j} \frac{1}{N} \cdot \frac{4M}{N} \cdot \sum_{\{k\in\{\underline{r},\ldots,\overline{r}\}:\, |a_{k,j}|\ge 8M/N\}} 1 \\ &\le \frac{4M}{N^2} \cdot 9 \cdot \sum_{k:\, |a_{k,j}|\ge 8M/N} 1 = \frac{36\,M}{N^2} \sum_{k:\, |a_{k,j}|\ge 8M/N} 1. \end{aligned} \tag{40}$$

Combining (40) and (37), we obtain that for each $j = 1,2,\ldots$

$$\begin{aligned} \int_0^1 |F_j|\,dx &\le \frac{36\,M}{N^2} \sum_{k:\, |a_{k,j}|\ge 8M/N} 1 + \frac{2}{N} \cdot \frac{16\,M}{N} \cdot \sum_{k:\, |a_{k,j}|\ge 2M/N} 1 \\ &\le \frac{70\,M}{N^2} \sum_{k:\, |a_{k,j}|\ge 2M/N} 1. \end{aligned}$$

Hence (see (35))

$$R_1 \le \frac{70\,M^2}{N^2} \sum_{(k,j):\, |a_{k,j}|\ge 2M/N} 1. \tag{41}$$

Note that in (41) the summation is done over all pairs (k, j) such that $j \in \mathbb{N}$, $k \in \{1, \dots, N\}$ and $|a_{k,j}| \geq 2M/N$.

Now let us estimate R_2 (see (33)). By the Cauchy inequality,

$$R_2 \leq \sum_{k=1}^{N} \Big(\sum_{j:\, |a_{k,j}| \leq 2M/N} a_{k,j}^2 \Big)^{1/2} \cdot \Big(\sum_{j=1}^{\infty} c_{k,j}^2 \Big)^{1/2},$$

where for each $j = 1, 2, \dots$ and $k \in \{1, \dots, N\}$

$$c_{k,j} = \int_0^1 f_k^N(x) \varphi_j(x)\, dx.$$

At the same time, by Bessel inequality and (28d), for each $k = 1, 2, \dots, N$,

$$\Big(\sum_{j=1}^{\infty} c_{k,j}^2 \Big)^{1/2} \leq N^{-1/2}.$$

Thus,

$$R_2 \leq \sum_{k=1}^{N} N^{-1/2} \Big(\sum_{j:\, |a_{k,j}| \leq 2M/N} a_{k,j}^2 \Big)^{1/2}. \tag{42}$$

To conclude the proof of (26) and, therefore, Theorem 2, we use the inequality

$$\frac{70\, M^2}{N^2} \sum_{(k,j):\, |a_{k,j}| \geq 2M/N} 1 + \sum_{k=1}^{N} N^{-1/2} \Big(\sum_{(k,j):\, |a_{k,j}| \leq 2M/N} a_{k,j}^2 \Big)^{1/2} \geq \frac{1}{2} \tag{43}$$

which we proved above (see (31), (41), (42)). Assuming that n is fixed and $N = [2000\, M^2 n]$, suppose

$$\max_{1 \leq k \leq N} e_n(\chi_{k/N}, \Phi, L^2(0,1)) < \frac{1}{400\, M n^{1/2}}; \tag{44}$$

(otherwise, for a given n (26) evidently holds) and let $\Lambda_k \subset \mathbb{N}$, $\#\Lambda_k = n$, be a set of indices such that

$$\|P_k - \chi_{k/N}\|_{L^2} \leq \frac{1}{400\, M n^{1/2}}, \qquad P_k = \sum_{j \in \Lambda_k} a_{k,j} \varphi_j.$$

Then

$$\sum_{j:\, |a_{k,j}| \leq 2M/N} a_{k,j}^2 \leq \sum_{\{j:\, |a_{k,j}| \leq 2M/N\} \cap \Lambda_k} a_{k,j}^2 + \sum_{\{j:\, |a_{k,j}| \leq 2M/N\} \cap C\Lambda_k} a_{k,j}^2, \tag{45}$$

where $C\Lambda_k = \mathbb{N} \setminus \Lambda_k$.

Combining (45) with the inequality

$$\sum_{j\in C\Lambda_k} a_{k,j}^2 \le \frac{1}{(400\,M)^2 n},$$

we obtain

$$\sum_{j:\,|a_{k,j}|\le 2M/N} a_{k,j}^2 \le n\cdot\frac{(2M)^2}{N^2}+\frac{1}{(400\,M)^2 n}.$$

Hence, the second sum in (43) is not greater than

$$\sum_{k=1}^{N} N^{-1/2}\left[\sqrt{2}\,\frac{2M}{N}n^{1/2}+\frac{\sqrt{2}\cdot 1}{400\,Mn^{1/2}}\right] \le 2\sqrt{2}\,M\left(\frac{n}{N}\right)^{1/2}+\frac{\sqrt{2}}{400\,M}\left(\frac{N}{n}\right)^{1/2}\le\frac{1}{4}.$$

Thus, assuming that (44) holds, we may claim (see (43)) that

$$\frac{70\,M^2}{N^2}\sum_{(k,j):\,|a_{k,j}|\ge 2M/N} 1 \ge \frac{1}{4}. \tag{46}$$

From (46) and the definition of N it follows that there exists a number $k_0\in\{1,\dots,N\}$ such that

$$\#\left\{j : |a_{k,j}| \ge \frac{2M}{N}\right\} \ge \frac{N}{280\,M^2} \ge 2n. \tag{47}$$

The inequality (47) yields the following estimate for the n-term approximation of the function $\chi_{k_0/N}$:

$$\begin{aligned} e_n(\chi_{k_0/N},\Phi,L^2(0,1)) &\ge \left[\left(\frac{2M}{N}\right)^2\cdot n\right]^{1/2} = \frac{2M}{N}\cdot n^{1/2}\\ &= \frac{2M\cdot n^{1/2}}{[2000\,M^2]\cdot n} \ge \frac{1}{1000\,M}\cdot n^{-1/2}, \end{aligned}$$

which means that under the assumption (44) the estimate (26) also holds. This completes the proof of Theorem 2.

References

[1] R. A. DeVore, Nonlinear approximation, *Acta Numerica* **7** (1998), 51–150.

[2] V. N. Temlyakov, Nonlinear Methods of Approximation, *IMI Preprint 2001:09,* University of South Carolina, 2001.

[3] K. I. Oskolkov, Linear and nonlinear methods in ridge approximation, *in* "Metric Function Theory and Related Topics" (dedicated to P. L. Ulyanov's 70th birthday), pp. 165–195, AFC, Moscow, 1999. [In Russian]

[4] S. B. STECHKIN, On absolute convergence of orthonormal series, *Dokl. Akad. Nauk SSSR* **102**, 1 (1955), 37–40. [In Russian]

[5] B. S. KASHIN, On approximation properties of complete orthonormal systems, *Trudy Mat. Inst. Steklov* **172** (1985), 187–191; *Proc. Steklov Inst. Math.* **3** (1987), 207–211. [English translation]

[6] B. S. KASHIN AND V. N. TEMLYAKOV, On best m-term approximations and the entropy of sets in the space L^1, *Math. Notes* **56**, 5 (1994), 57–86.

[7] D. L. DONOHO, CART and best-ortho-basis: a connection, *Ann. of Statistics* **25**, 5 (1997), 1870–1911.

[8] B. S. KASHIN AND A. A. SAAKYAN, "Orthogonal Series," 2nd edition, AFC, Moscow, 1999. [In Russian]

[9] B. S. KASHIN, On the coefficients of expansion of functions from a certain class with respect to complete systems, *Siberian J. Math.* **18**, 1 (1977), 122–131.

[10] B. S. KASHIN, On lower estimates for n-term approximations, *Math. Notes* **70**, 4 (2001), 636–638.

[11] G. HARDY, D. LITTLEWOOD, AND G. POLIA, "Inequalities," Inostrannaya Literatura, Moscow, 1948. [In Russian]

BORIS KASHIN
Steklov Mathematical Institute
Vavilov str. 42
Moscow 117 966
GSP-1, RUSSIA
E-mail: `kashin@ipsun.ac.msk.su`