Industrial Mathematics Institute

1997:12

Nonlinear Kolmogorov's widths

V. Temlyakov

IMI
Preprint Series

Department of Mathematics
University of South Carolina

NONLINEAR KOLMOGOROV'S WIDTHS[1]

V.N. Temlyakov

Department of Mathematics, University of South Carolina, Columbia, SC 29208

Abstract. We investigate a generalization of Kolmogorov's width which is suitable for estimating best m-term approximation. We generalize the Carl's inequality which gives lower estimate of Kolmogorov widths in terms of the entropy numbers. Application of these new inequalities gives some progress in the problem of estimating best m-term trigonometric approximation of multivariate functions.

1. Introduction

A number of different widths are being studied in approximation theory: Kolmogorov widths, linear widths, Fourier widths, Gelfand widths, Alexandrov widths and others. All these widths were introduced in approximation theory as characteristics of function classes (more generally compact sets) which give the best possible eccuracy of algorithms with certain restrictions. For instance, Kolmogorov's n-width for centrally symmetric compact set F in Banach space X is defined as follows

$$d_n(F,X) := \inf_L \sup_{f\in F} \inf_{g\in L} \|f-g\|_X$$

where $\inf_L$ is taken over all n-dimensional subspaces of X. In other words the Kolmogorov n-width gives the best possible error in approximating a compact set F by n-dimensional linear subspaces.

There has been an increasing interest last decade in nonlinear m-term approximation with regard to different systems. The present paper contains an attempt to generalize the concept of classical Kolmogorov's width in order to be used in estimating best m-term approximation. For this purpose we introduce a nonlinear Kolmogorov's (N,m)-width:

$$d_m(F,X,N) := \inf_{\mathcal{L}_N, \#\mathcal{L}_N \le N} \sup_{f\in F} \inf_{L\in\mathcal{L}_N} \inf_{g\in L} \|f-g\|_X,$$

where $\mathcal{L}_N$ is a set of at most N m-dimensional subspaces L. It is clear that

$$d_m(F,X,1) = d_m(F,X).$$

[1]This research was supported by the National Science Foundation Grant DMS 9622925 and by ONR Grant N0014-96-1-1003

Key words and phrases. entropy, nonlinear approximation, m-term approximation, width.

The new feature of $d_m(F,X,N)$ is that we allow to choose a subspace $L \in \mathcal{L}_N$ depending on $f \in F$. It is clear that the bigger N the more flexibility we have to approximate f. It turns out that from the point of view of our applications the following two cases

$$\text{(I)} \qquad N \asymp K^m,$$

where $K > 1$ is a constant, and

$$\text{(II)} \qquad N \asymp m^{am},$$

where $a > 0$ is a fixed number, play an important role.

We intend to use the (N,m)-widths for estimating from below the best m-term approximations. There are several general results (see[L], [C]) which give lower estimates of the Kolmogorov widths $d_n(F,X)$ in terms of the entropy numbers $\epsilon_k(F,X)$. In Section 2 we will generalize the following Carl's (see [C]) inequality: for any $r > 0$ we have

$$(1.1) \qquad \max_{1\le k\le n} k^r \epsilon_k(F,X) \le C(r) \max_{1\le m\le n} m^r d_{m-1}(F,X).$$

We denote here for integer k

$$\epsilon_k(F,X) := \inf\{\epsilon : \exists f_1,\dots,f_{2^k} \in X : F \subset \cup_{j=1}^{2^k}(f_j + \epsilon B(X))\},$$

where $B(X)$ is the unit ball of Banach space X. For noninteger k we set $\epsilon_k(F,X) := \epsilon_{[k]}(F,X)$ where $[k]$ is the integral part of number k. It is clear that

$$d_1(F,X,2^n) \le \epsilon_n(F,X).$$

In Section 2 we prove the inequality

$$(1.2) \qquad \max_{1\le k\le n} k^r \epsilon_k(F,X) \le C(r,K) \max_{1\le m\le n} m^r d_{m-1}(F,X,K^m),$$

where we denote

$$d_0(F,X,N) := \sup_{f\in F} \|f\|_X.$$

This inequality is a generalization of inequality (1.1). In Section 2 we also prove the following inequality

$$(1.3) \qquad \max_{1\le k\le n} k^r \epsilon_{(a+r)k\log k}(F,X) \le C \max_{1\le m\le n} m^r d_{m-1}(F,X,m^{am})$$

and give an example showing that $k\log k$ in this inequality can not be replaced by slower growing function on k.

In Section 3 we apply inequalities (1.2) and (1.3) for estimating the best m-term trigonometric approximation from below. Let $\mathcal{T} := \{e^{i(k,x)}\}$ be the trigonometric system. Denote

$$\sigma_m(f)_p := \sigma_m(f,\mathcal{T})_p := \inf_{k^1,\dots,k^m;c_1,\dots,c_m} \|f - \sum_{j=1}^m c_j e^{i(k^j,x)}\|_p$$

the best m-term trigonometric approximation in L_p. For a function class F we denote

$$\sigma_m(F)_p := \sigma_m(F,\mathcal{T})_p := \sup_{f\in F} \sigma_m(f,\mathcal{T})_p.$$

As a corollary to a version of (1.2) (see Remark 2.1) we give a new proof (see [DT]) for the estimate

$$\sigma_m(W^r_\infty,\mathcal{T})_1 \gg m^{-r},$$

where W^r_∞ is a standard Sobolev class with the restriction imposed in the L_∞-norm. We use a version of (1.3) to get some new lower estimates in m-term trigonometric approximation of multivariate classes MW^r_∞ of functions with bounded mixed derivative in the L_1-norm. We prove that

$$\sigma_m(MW^r_\infty,\mathcal{T})_1 \gg m^{-r}(\log m)^{r(d-2)}. \tag{1.4}$$

The inequality (1.4) gives a new estimate for small r. In Section 3 we discuss some results on $\sigma_m(MW^r_q,\mathcal{T})_p$. We remark here that the correct order of the quantity $\sigma_m(MW^r_\infty,\mathcal{T})_1$ is unknown.

In Section 4 we apply the method developed in Section 3 to a general system Ψ instead of trigonometric system $\mathcal{T}$. We consider best m-term approximations with regard to a general system $\Psi := \{\psi_j\}_{j=1}^\infty$

$$\sigma_m(f,\Psi)_X := \inf_{c_k,j_k,k=1,\dots,m} \|f-\sum_{k=1}^m c_k\psi_{j_k}\|_X \quad ,$$

$$\sigma_m(F,\Psi)_X := \sup_{f\in F}\sigma_m(f,\Psi)_X.$$

We prove that for good systems Ψ the estimate

$$\epsilon_n(F,X) \gg n^{-a}(\log n)^b, \quad a>0, b\in\mathbb{R},$$

for the entropy numbers implies the same estimate for best m-term approximation:

$$\sigma_m(F,\Psi)_X \gg m^{-a}(\log m)^b \quad .$$

2. Some general inequalities for nonlinear Kolmogorov's widths

We begin this section by the proof of the inequality (1.2).

Theorem 2.1. *For any compact $F\subset X$ and any $r>0$ we have for all $n\in\mathbb{N}$*

$$\max_{1\le k\le n} k^r\epsilon_k(F,X) \le C(r,K)\max_{1\le m\le n} m^r d_{m-1}(F,X,K^m).$$

Proof. Let $X(N,m)$ denote the union of some N subspaces L with $\dim L = m$. Consider a collection $\mathcal{K}(K,l) := \{X(K^{2^{s+1}},2^{s+1})\}_{s=1}^l$ and denote

$$H^r(\mathcal{K}(K,l)) := \{f\in X : \exists L_1(f),\dots,L_l(f) : L_s(f)\in X(K^{2^{s+1}},2^{s+1}),$$

and $\exists t_s(f)\in L_s(f)$ such that

$$\|t_s(f)\|_X \le 2^{-r(s-1)}, \quad s=1,\dots,l; \quad \|f-\sum_{s=1}^l t_s(f)\|_X \le 2^{-rl}\}.$$

Lemma 2.1. *We have for $r > 0$*

$$\epsilon_{2^l}(H^r(\mathcal{K}(K,l)), X) \le C(r,K)2^{-rl}.$$

Proof. We use the following well-known (see[P]) estimate for $\epsilon_n(B,X)$ of the unit ball B in the d-dimensional space X :

$$\epsilon_n(B,X) \le 3(2^{-n/d}). \tag{2.1}$$

Take any sequence $\{n_s\}_{s=1}^{l(r)}$ of $l(r) \le l$ nonnegative integers. Construct ϵ_{n_s}-nets for all unit balls of the spaces in $X(K^{2^{s+1}}, 2^{s+1})$. Then the total number of the elements y_j^s in these ϵ_{n_s}-nets does not exceed

$$M_s := K^{2^{s+1}} 2^{n_s}.$$

We consider now the set A of elements of the form

$$y_{j_1}^1 + 2^{-r} y_{j_2}^2 + \dots + 2^{-r(l(r)-1)} y_{j_{l(r)}}^{l(r)}, \quad j_s = 1, \dots, M_s, \quad s = 1, \dots, l(r).$$

The total number of these elements does not exceed

$$M = \prod_{s=1}^{l(r)} M_s \le K^{2^{l(r)+2}} 2^{\sum_{s=1}^{l(r)} n_s}.$$

It is clear it suffices to consider the case of big $l \ge l(r,K)$. We take now

$$n_s = [(r+1)(l-s)2^{s+1}], \quad s = 1, \dots, l(r),$$

where $[x]$ denotes the integer part of a number x. We choose $l(r) \le l$ as a maximal natural number satisfying

$$\sum_{s=1}^{l(r)} n_s \le 2^{l-1}$$

and

$$2^{l(r)+2} \log K \le 2^{l-1}.$$

It is clear that

$$l(r) \ge l - C(r,K).$$

Then we have

$$M \le 2^{2^l}.$$

For the error $\epsilon(f)$ of approximation of $f \in H^r(\mathcal{K}(K,l))$ by elements of A we have

$$\begin{aligned}
\epsilon(f) &\le 2^{-rl} + \sum_{s=1}^{l(r)} \|t_s(f) - 2^{-r(s-1)} y_{j_s}^s\|_X + \sum_{s=l(r)+1}^{l} \|t_s(f)\|_X \\
&\le C(r,K)2^{-rl} + \sum_{s=1}^{l(r)} 2^{-r(s-1)} \epsilon_{n_s}(B(L_s(f), X) \\
&\le C(r,K)2^{-rl} + 3\sum_{s=1}^{l(r)} 2^{-r(s-1)} 2^{-n_s/2^{s+1}} \\
&\le C(r,K)2^{-rl}.
\end{aligned}$$

Lemma 2.1 is proved now.

We continue the proof of Theorem 2.1. Assume

$$\max_{1\le m\le n}(m)^r d_{m-1}(F,X,K^m) < 1/2.$$

Then for $s=1,2,\dots,l$; $l\le[\log(n-1)]$ we have

$$d_{2^s}(F,X,K^{2^s}) < 2^{-rs-1}.$$

This means that for each $s=1,2,\dots,l$, there is a collection $\mathcal{L}_{K^{2^s}}$ of K^{2^s} 2^s-dimensional spaces $L^s_j, j=1,\dots,K^{2^s}$, such that for each $f\in F$ there exists a subspace $L^s_{j_s}(f)$ and an approximant $a_s(f)\in L^s_{j_s}(f)$ such that

$$\|f-a_s(f)\|\le 2^{-rs-1}.$$

Consider

$$t_s(f) := a_s(f)-a_{s-1}(f),\quad s=1,2,\dots,l. \tag{2.2}$$

Then we have

$$t_s(f)\in L^s_{j_s}(f)\oplus L^s_{j_{s-1}}(f),\quad \dim(L^s_{j_s}(f))\oplus L^s_{j_{s-1}}(f)\le 2^s+2^{s-1}<2^{s+1}.$$

Let $X(K^{2^{s+1}},2^{s+1})$ denote the collection of all $L^s_{j_s}\oplus L^s_{j_{s-1}}$ over various $1\le j_s\le K^{2^s}$; $\quad 1\le j_{s-1}\le K^{2^{s-1}}$. For $t_s(f)$ defined by (2.2) we have

$$\|t_s(f)\|\le 2^{-rs-1}+2^{-r(s-1)-1}\le 2^{-r(s-1)}.$$

Next, for $a_0\in L^0$ we have

$$\|f-a_0\|\le 1/2$$

and from $d_0(F,X)\le 1/2$ we get

$$\|a_0\|\le 1.$$

Take $t_0(f)=a_0(f)$. Then we have $F\subset H^r(\mathcal{K}(K,l))$ and Lemma 2.1 gives the required bound

$$\epsilon_{2^l}(F)\le C(r,K)2^{-rl},\quad 1\le l\le[\log(n-1)].$$

It is clear that these inequalities imply the conclusion of Theorem 2.1.

Remark 2.1. *Examining the proof of Theorem 2.1 one can check that the inequality holds for K^m replaced by bigger function. For example we have*

$$\max_{1\le k\le n} k^r\epsilon_k(F,X)\le C(r,K)\max_{1\le m\le n} m^r d_{m-1}(F,X,(Kn/m)^m).$$

We proceed now to the case (II) when $N\asymp m^{am}$. We prove a lemma which will imply the inequality (1.3).

Lemma 2.2. *For any compact set $F \subset X$ and any real numbers $0 < a < b$ we have*

$$\epsilon_{bm\log m}(F,X) \le C(d_0(F,X)m^{a-b} + d_m(F,X,m^{am})).$$

Proof. Let $N := [m^{am}]$. For a given $\delta > 0$ denote $\mathcal{L}_N$ a collection of m-dimensional subspaces $L_j, j = 1,\dots,N$, such that for each $f \in F$ there exists $j(f) \in [1,N]$ and an element $g(f) \in L_{j(f)}$ with the approximating property

$$\|f - g(f)\| \le d_m(F,X,m^{am}) + \delta. \tag{2.3}$$

Then

$$\|g(f)\|_X \le \|f\|_X + \|f - g(f)\|_X \le d_0(F,X) + d_m(F,X,m^{am}) + \delta =: \alpha.$$

Thus we got to estimate ϵ-entropy of the union U of m-dimensional balls of radius α in L_j over $j \in [1,N]$. By (2.1) we have

$$\epsilon_{n+[\log N]+1}(U,X) \le \alpha 3(2^{-n/m}).$$

If $(b-a)m\log m < 1$ the statement of Lemma 2.2 is trivial. Assume $(b-a)m\log m \ge 1$ and choose n such that $n = [(b-a)m\log m] - 1$, then we have

$$n + [\log N] + 1 \le bm\log m$$

and

$$\alpha 3(2^{-n/m}) \le C\alpha m^{a-b}. \tag{2.4}$$

Combining (2.3) and (2.4) we get the statement of Lemma 2.2.

It is easy to see that this lemma implies the inequality (1.3). We give now an example showing that we can not get rid of $\log k$ in (1.3).

Example. Let $r > 0$ and $a > 0$ be given. Consider a partition of $[0,1]$ into $N = [n^{a+1}/3]$ segments $I_j := [\frac{j-1}{N}, \frac{j}{N})$, $j = 1,\dots,N$, and form the set of all n-dimensional subspaces of the form

$$X_Q := \operatorname{span}\{\chi_{I_j}\}_{j\in Q}, \quad Q \subset \{1,2,\dots,N\}, \quad \#Q = n$$

where χ_I denotes the characteristic function of a segment I. The number of these subspaces is

$$N(n) = \binom{N}{n}; \quad ((n^a - 1)/3)^n \le \binom{N}{n} \le n^{an}.$$

Consider

$$F_n := \cup_Q B_\infty(X_Q)n^{-r}.$$

Then we have $F_n \subset n^{-r}B(L_\infty([0,1])$, what implies for all m

$$d_m(F_n, L_\infty([0,1])) \le n^{-r}.$$

We also have

$$d_n(F_n, L_\infty([0,1]), n^{an}) \le d_n(F_n, L_\infty([0,1]), N(n)) = 0.$$

This implies that for any $s \in \mathbb{N}$ we have for the right hand side of (1.3)

$$\max_{1\le m\le s} m^r d_m(F_n, L_\infty([0,1]), m^{am}) \le 1.$$

Next, consider the set of functions χ_{G_Q}, $G_Q = \cup_{j\in Q} I_j$, where χ_G is the characteristic function of G. Then for any $Q \ne Q'$ we have

$$\|\chi_{G_Q} - \chi_{G_{Q'}}\|_\infty \ge 1.$$

The number of functions $\{\chi_{G_Q}\}$ is equal to $N(n)$. This implies that

$$\epsilon_{[\log N(n)]}(F_n, L_\infty([0,1])) \ge n^{-r}/2.$$

Assume we can replace $\log k$ in (1.3) by a slower growing function $\psi(k)$. Take any $n \in \mathbb{N}$ and let $\mu_n \in \mathbb{N}$ be the biggest number satisfying the inequality

$$(a+r)\mu_n\psi(\mu_n) \le [\log N(n)].$$

Then our assumption implies

$$\lim_{n\to\infty} \mu_n/n = \infty.$$

Thus for the left hand side of (1.3) we get

$$\max_{1\le k\le \mu_n} k^r \epsilon_{(a+r)k\psi(k)}(F_n.X) \ge \mu_n^r \epsilon_{(a+r)\mu_n\psi(\mu_n)}(F_n, X) \ge (\mu_n/n)^r/2 \to \infty$$

as $n \to \infty$. We got a contradiction to (1.3).

3. Best m-term trigonometric approximation in the L_1-norm

In this section we apply Remark 2.1 and Lemma 2.2 to get some lower estimates for best m-term trigonometric approximation. In order to orient the reader we first give a new proof of the following estimate obtained in [DT]

$$\sigma_m(W^r_\infty, L_1) \gg m^{-r}, \quad r > 0. \tag{3.1}$$

Let us remind the definition of W^r_∞ in the case of univariate functions

$$W^r_\infty := \{f = B_r * \varphi, \quad \|\varphi\|_\infty \le 1, \quad B_r(x) := 1 + 2\sum_{k=1}^\infty k^{-r}\cos(kx - r\pi/2)\}.$$

It is known that

$$\epsilon_n(W^r_\infty, L_1) \gg n^{-r}, \quad r > 0.$$

Moreover, the same estimate is known for the intersection of W^r_∞ and the space $\mathcal{T}(n)$ of trigonometric polynomials of degree n. We formulate this estimate in the following way using the notation $\mathcal{T}(n)_\infty$ for the unit L_∞-ball in $\mathcal{T}(n)$

$$\epsilon_n(\mathcal{T}(n)_\infty, L_1) \gg 1. \tag{3.2}$$

This relation can be derived from the estimate of volume of the set of Fourier coefficients of polynomials from $\mathcal{T}(n)_\infty$. The corresponding volume estimate was obtained in [K]. For multivariate case and the way of deriving estimates for the entropy numbers from the corresponding volume estimates see [T3]. We will not go in these details here.

Consider k-term approximation of functions from $\mathcal{T}(n)_\infty$. Using the de la Vallée-Possin operator we confine ourselves to the case of approximation by harmonics with frequences in $(-2n, 2n)$. Then the number of possible combinations of k frequences from this set is $\binom{4n-1}{k} =: N_k \le (Kn/k)^k$. And, consequently

$$\sigma_k(\mathcal{T}(n)_\infty, L_1) \ge C d_k(\mathcal{T}(n)_\infty, L_1, N_k). \tag{3.3}$$

The relation (3.2) and Remark 2.1 to Theorem 2.1 imply

$$\max_{1\le k\le n} k^r d_k(\mathcal{T}(n)_\infty, L_1, N_k) \ge C(r)n^r. \tag{3.4}$$

Using the obvious inequality

$$d_0(\mathcal{T}(n)_\infty, L_1) \le 1$$

we get

$$\max_{C(r)n\le k\le n} k^r d_k(\mathcal{T}(n)_\infty, L_1, N_k) \ge C(r)n^r,$$

what together with (3.3) implies

$$\sigma_{C(r)n}(\mathcal{T}(n)_\infty, L_1) \gg 1,$$

and in turn

$$\sigma_n(W^r_\infty, L_1) \gg n^{-r}.$$

We proceed now to the main part of this section to the best m-term approximation of classes of multivariate functions with bounded mixed derivative. The first result in this direction was obtained in [T]. We remind the definition of these classes

$$MW^r_q := \{f = B_r * \varphi, \quad \|\varphi\|_q \le 1, \quad B_r(x_1,\dots,x_d) := \prod_{i=1}^d B_r(x_i)\}.$$

The first results concerned the case of approximating the classes MW^r_q in the L_p-norm for $1 < q \le p \le 2$. Those results were surprising. In the case of univariate functions it is well known that we have

$$\sigma_m(W^r_q, L_p) \asymp d_m(W^r_q, L_p), \quad 1 < q \le p \le 2,$$

and the standard subspace $\mathcal{T}(m)$ of trigonometric polynomials of degree m provides both an optimal (in the sense of order) subspace in the Kolmogorov $(2m+1)$-width and an optimal (in the sense of order) set of harmonics in nonlinear best $(2m+1)$-term approximation. It turned out that in the multivariate case we have in the same setting the following relations

$$\sigma_m(MW^r_q, L_p)(\log m)^{(d-1)(1/q-1/p)} \asymp d_m(MW^r_q, L_p), \quad 1 < q \le p \le 2.$$

This relation shows that we gain by using nonlinear m-term trigonometric approximation instead of approximating by elements from any m-dimentional subspace.

Some upper and lower bounds for approximating classes MW^r_q in the uniform norm were obtained in [B1]. In [KT1] we proved the estimate

$$\sigma_m(MW^r_q)_p \gg m^{-r}(\log m)^{(d-1)r}, \quad 1 < p \le q < \infty.$$

Combining these estimates with known upper bounds for approximation by trigonometric polynomials with frequences in hyperbolic crosses (see [T2], Ch.II, s.2)

$$E_{Q_n}(MW^r_q, L_q) \ll 2^{-rn}, \quad 1 < q < \infty,$$

we get the asymptotic estimates

$$\sigma_m(MW^r_q)_p \asymp m^{-r}(\log m)^{(d-1)r}, \quad 1 < p \le q < \infty.$$

As we mentioned in Introduction the main goal of this section is to get some new estimates for $\sigma_m(MW^r_q)_p$ in the case $q = \infty$ and $p = 1$. For completeness we will obtain first the asymptotic estimates in the case $1 < q \le p < \infty$, $p \ge 2$.

Theorem 3.1. *For $r > \max(1/2, 2/q - 1/2)$ and $1 < q \le p < \infty$, $p \ge 2$ we have*

$$\sigma_m(MW^r_q)_p \asymp \begin{cases} m^{-(r-1/q+1/2)}(\log m)^{(d-1)(r-2/q+1)}, & 1 < q \le 2; \\ m^{-r}(\log m)^{(d-1)r}, & 2 \le q < \infty. \end{cases}$$

Proof. The lower estimates follow from the known result (see [T2], Ch.IV, Th. 2.1)

$$\sigma_m(MW^r_q)_2 \gg m^{-(r-1/q+1/2)}(\log m)^{(d-1)(r-2/q+1)}, \quad 1 < q \le 2. \tag{3.5}$$

We prove now the upper estimates. It is clear that it is sufficient to consider the case $1 < q \le 2$. We will prove a little stronger result than the upper estimates in Theorem 3.1. Namely, instead of L_p-norm we use a stronger one $B_{\infty,2}$-norm which is defined as follows. Let $f \in L_1$ and

$$\delta_s(f) := \sum_{k\in\rho(s)} \hat{f}(k)e^{i(k,x)}, \quad \rho(s) := \{k \in \mathbb{Z}^d : [2^{s_j-1}] < |k_j| \le 2^{s_j}, \quad j = 1, 2, \dots, d\}.$$

Define the $B_{\infty,2}$-norm by the formula

$$\|f\|_{B_{\infty,2}} := (\sum_s \|\delta_s(f)\|^2_\infty)^{1/2}.$$

Lemma 3.1. *For $r > 2/q - 1/2$ we have*

$$\sigma_m(MW^r_q)_{B_{\infty,2}} \ll m^{-(r-1/q+1/2)}(\log m)^{(d-1)(r-2/q+1)}, \quad 1 < q \le 2.$$

Proof. Denote

$$f_l := \sum_{\|s\|_1 = l} \delta_s(f).$$

Then it is known (see for instance [T2],Ch.II, s.2) that for $f \in MW^r_q$, $1 < q < \infty$, we have

$$\|f_l\|_q \ll 2^{-rl}. \tag{3.6}$$

Next, by Lemma 3.1′ from [T2] we get for $1 < q \le 2$

$$\|f_l\|_q^q \gg \sum_{\|s\|_1 = l} \|\delta_s(f)\|_2^q 2^{l(1/2-1/q)q}$$

what implies

$$(\sum_{\|s\|_1 = l} \|\delta_s(f)\|_2^q)^{1/q} \ll 2^{-l(r+1/2-1/q)}. \tag{3.7}$$

We take small $\kappa > 0$ and specify

$$m_l := [2^{-(1+\kappa)(l-n)} l^{d-1}]; \quad N_l := [2^{n-\kappa(l-n)}].$$

By [DT], Corollary 5.1 we have for $\|s\|_1 = l$

$$\sigma_{N_l}(\delta_s(f))_\infty \ll (2^l/N_l)^{1/2} \log(2^l/N_l)\|\delta_s(f)\|_2. \tag{3.8}$$

Denote $D(f,l)$ the set of m_l indices s, $\|s\|_1 = l$, for which $\|\delta_s(f)\|_2$ take the biggest values. Then by Lemma 2.1 from [T2] and (3.7) we get

$$\left(\sum_{s \notin D(f,l)} \|\delta_s(f)\|_2^2\right)^{1/2} \ll m_l^{1/2-1/q} 2^{-l(r+1/2-1/q)}. \tag{3.9}$$

Denote

$$\nu(l) := \#\{s : \|s\|_1 = l\}; \qquad m := \sum_{l \ge n} (m_l 2^l + N_l \nu(l)) + \#(\cup_{\|s\|_1 < n} \rho(s)).$$

From (3.8) and (3.9) we get

$$\sigma_{m_l 2^l + N_l \nu(l)}(f_l)_{B_{\infty,2}} \ll \left(\sum_{s \notin D(f,l)} \|\delta_s(f)\|_2^2\right)^{1/2} (2^l/N_l)^{1/2} \log(2^l/N_l) \ll$$
$$m_l^{1/2-1/q}(2^l/N_l)^{1/2} \log(2^l/N_l) 2^{-l(r+1/2-1/q)}.$$

Therefore,

$$\sigma_m(f)_{B_{\infty,2}} \le \sum_{l \ge n} \sigma_{m_l 2^l + N_l \nu(l)}(f_l)_{B_{\infty,2}} \ll n^{(d-1)(1/2-1/q)} 2^{-n(r+1/2-1/q)}.$$

It remains to remark that $m \ll n^{d-1} 2^n$.

We proceed now to the case $q = \infty$, $p = 1$.

Theorem 3.2. *For $r > 0$ we have*

$$\sigma_m(MW^r_\infty)_1 \gg m^{-r}(\log m)^{r(d-2)}.$$

Proof. In [KT2] (see also [KT1] and [B2]) we proved

$$\epsilon_m(MW^r_\infty)_1 \gg m^{-r}(\log m)^{r(d-1)}. \tag{3.10}$$

We use the construction from [KT1] and [KT2] in our proof. In [KT1] for each $k \in \mathbb{N}$ we found a number $n(k) \in \mathbb{N}$ such that $k \asymp 2^{n(k)}n(k)^{d-1}$ and constructed a set U^r_K with the following properties

$$U^r_m \subset MW^r_\infty; \tag{3.11}$$

$$\epsilon_k(U^r_k, L_1) \gg k^{-r}(\log k)^{r(d-1)}. \tag{3.12}$$

It follows from the proofs in [KT1] and [KT2] that we can modify (convolve with appropriate de la Vallée-Poussin kernal) the original set U^r_m in a way such that in addition to properties (3.11) and (3.12) we can have one more property

$$U^r_k \subset \mathcal{T}(2^{n(k)+1}), \tag{3.13}$$

which says that all elements of U^r_k are trigonometric polynomials of degree $2^{n(k)+1}$ in each variable. Let $m \in \mathbb{N}$ be given. Define $k := [bm\log m]$ with $b > 0$ to be chosen later on and consider the set U^r_k with properties (3.11), (3.12) and (3.13). Similarly to the proof of (3.2) using the de la Vallée-Poussin operator we confine ouselves to approximation only by frequences in $(-2^{n(k)+2}, 2^{n(k)+2})^d$. Then the number of possible combinations of m frequences is $N(m) := \binom{(2^{n(k)+3}-1)^d}{m} \le (C(d)m)^{dm}$ for $d \ge 2$. We apply now Lemma 2.2 and the inequality

$$\sigma_m(U^r_k)_1 \ge d_m(U^r_k, L_1, N(m)). \tag{3.14}$$

Using the property (3.11) we get

$$d_0(U^r_k, L_1) \ll 1$$

and by Lemma 2.2 and (3.12) we find

$$d_m(U^r_k, L_1, N(m)) \ge C^{-1}\epsilon_{bm\log m}(U^r_k, L_1) - d_0(U^r_k, L_1)m^{a-b} \gg m^{-r}(\log m)^{r(d-2)}$$

for properly chosen a and b.

4. One estimate for general m-term approximation

In this section we use Remark 2.1 to prove a lower estimate for best m-term approximation with regard to a system satisfying some restrictions. Assume a system $\Psi := \{\psi_j\}_{j=1}^\infty$ of elements in X satisfies the condition:

(VP) There exist three positive constants A_i, $i = 1,2,3$, and a sequence $\{n_k\}_{k=1}^\infty$, $n_{k+1} \le A_1 n_k$, $k = 1,2,\dots$ such that there is a sequence of the de la Vallée-Poussin type operators V_k with the properties

$$V_k(\psi_j) = \lambda_{k,j}\psi_j, \qquad \lambda_{k,j} = 1 \quad \text{for} \quad j = 1,\dots,n_k; \qquad \lambda_{k,j} = 0 \quad \text{for} \quad j > A_2 n_k, \tag{4.1}$$

$$\|V_k\|_{X\to X} \le A_3, \quad k = 1,2,\dots \quad . \tag{4.2}$$

Theorem 4.1. *Assume that for some $a > 0$ and $b \in \mathbb{R}$ we have*

$$\epsilon_m(F,X) \ge C_1 m^{-a}(\log m)^b, \quad m = 1,2,\dots \quad .$$

Then if a system Ψ satisfies the condition (VP) and also satisfies the following two conditions

$$E_n(F,\Psi) := \sup_{f\in F}\inf_{c_1,\dots,c_n}\|f - \sum_{j=1}^n c_j\psi_j\|_X \le C_2 n^{-a}(\log n)^b, \quad n = 1,2,\dots; \tag{4.3}$$

$$V_k(F) \subset C_3 F \tag{4.4}$$

we have

$$\sigma_m(F,\Psi)_X \gg m^{-a}(\log m)^b.$$

Proof. Let m be fixed. We find a constant C_4 such that for all n

$$E_{C_4 n}(F,\Psi) \le (2(A_3+1))^{-1}\epsilon_n(F,X). \tag{4.5}$$

Let $k(m)$ be such that

$$n_{k(m)} \ge C_4 m \quad \text{and} \quad n_{k(m)-1} < C_4 m.$$

Then

$$n_{k(m)} \le A_1 C_4 m. \tag{4.6}$$

Consider a new set

$$F_m := V_{k(m)}(F).$$

By our assumption we have $F_m \subset C_3 F$ and, therefore, it is sufficient to prove the required estimate for the set F_m instead of F. We get from (4.2) and (4.5)

$$\epsilon_m(F_m,X) \ge \epsilon_m(F,X) - \|I - V_{k(m)}\|_{X\to X} E_{C_4 m}(F,\Psi) \ge \epsilon_m(F,X)/2. \tag{4.7}$$

Denote

$$\sigma_n(f,\Psi_N)_X := \inf_{\Lambda,\{c_j\}} \|f - \sum_{j\in\Lambda} c_j\psi_j\|,$$

where Λ is a subset of $[1,N]$ of cardinality n. Then using the (VP) property of the system Ψ we get for F_m

$$\sigma_l(F_m,\Psi_{A_2 n_{k(m)}})_X \le A_3\sigma_l(F_m,\Psi)_X. \tag{4.8}$$

Denote $N := [A_2 n_{k(m)}]$ and $K := [A_2A_1C_4] + 1$. Then the total number of l-dimensional subspaces of the space $\mathrm{span}\{\psi_1,\dots,\psi_N\}$ is less than or equal to $\binom{Km}{l} \le (Km/l)^l$. Therefore for each l we have

$$\sigma_l(F_m,\Psi_N)_X \ge d_l(F_m,X,(Km/l)^l). \tag{4.9}$$

Next, from the assumption (4.4) we get

$$d_l(F_m,X) \le C_3 d_l(F,X) \le C_3 E_n(F,\Psi).$$

Using this inequality we obtain from Remark 2.1 with $r = a+1$ that for some C_5 we have

$$\max_{C_5 m\le l\le m} l^r d_l(F_m,X,(Km/l)^l) \gg m(\log m)^b$$

what with (4.9) implies

$$\sigma_{C_5 m}(F_m,\Psi_N)_X \gg m^{-a}(\log m)^b.$$

It remains to use (4.8) in order to complete the proof of Theorem 4.1.

References

[B1] E.S. Belinskii, *Approximation of functions of several variables by trigonometric polynomials with given number of harmonics, and estimates of ϵ-entropy*, Analysis Math. **15** (1989), 67–74.

[B2] E.S. Belinskii, *The asymptotic characteristics of classes of functions with restrictions on the mixed derivative (mixed difference)*, Invest. on theory of multivariate functions of real variables, Yaroslavskii Gos. Universitet, 1990, pp. 22–37.

[DT] R.A. DeVore and V.N. Temlyakov, *Nonlinear approximation by trigonometric sums*, J. Fourier Analysis and Applications **2** (1995), 29–48.

[C] B. Carl, *Entropy numbers, s-numbers, and eigenvalue problems*, J. Funct. Anal. **41** (1981), 290–306.

[K] B.S. Kashin, *On certain properties of the space of trigonometric polynomials with the uniform norm*, Trudy Mat. Inst. Steklov **145** (1980), 111–116; English transl. in Proc. Steklov Inst. Math. **145** (1981), no. 1.

[KT1] B.S. Kashin and V.N. Temlyakov, *On best m-term approximation and the entropy of sets in the space L^1*, Mat. Zametki **56** (1994), 57–86; English transl. in Math. Notes **56** (1994).

[KT2] B.S. Kashin and V.N. Temlyakov, *On estimating the asymptotic characteristics of classes of functions with bounded mixed derivative*, Mat. Zametki **58** (1995), 922–925; English transl. in Math Notes **58** (1995).

[L] G.G. Lorentz, *Metric entropy and approximation*, Bull. Amer. Math. Soc. **72** (1966), 903–937.

[P] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge University Press, 1989.

[T1] V.N. Temlyakov, *On the approximation of periodic functions of several variables*, Dokl. AN SSSR **279** (1984), 301–305; English transl. in Soviet Math. Dokl. **30** (1984).

[T2] V. N. Temlyakov, *Approximation of funcions with bounded mixed derivative*, Trudy Mat. Inst. Steklov **178** (1986); English transl. in Proc. Steklov Inst. Math., (1989,), no. 1.

[T3] V. Temlyakov, *Estimates of the asymptotic characterisitics of classes of functions with bounded mixed derivative or difference*, Trudy Mat. Inst. Steklov **189 yr 1989**, 138–168; English transl. in Proc. Steklov Inst. Math. (1990), no. 4, 161–197.