IMI

Preprint Series

Industrial Mathematics Institute

1997:13

Nonlinear m-term approximation with regard to the multivariate Haar system

V. Temlyakov

Department of Mathematics
University of South Carolina

NONLINEAR m-TERM APPROXIMATION WITH REGARD TO THE MULTIVARIATE HAAR SYSTEM[1]

V.N. Temlyakov

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

Abstract. We study efficiency of L_p-greedy algorithm with regard to a multivariate system which is equivalent to the multivariate Haar system. In a place of multivariate Haar system we take the corresponding tensor product of univariate Haar systems. We prove that for $1 < p < \infty$ the L_p-greedy algorithm G^p provides an error of m-term approximation of any function in L_p such that it is at most $C(p,d)(\log m)^d$ times bigger than the best m-term approximation of that function. We also prove that in the case $p = 1$ or $p = \infty$ efficiency of G^p is not as good as in the case $1 < p < \infty$. Namely, the extra factor jumps from $(\log m)^d$ in the case $1 < p < \infty$ to m in the case $p = 1, \infty$.

1. Introduction

This paper is a follow up to the paper [T1]. We recall the most important notations from [T1] and formulate one result from [T1] which is a starting point for this paper. Denote the univariate Haar system by $\mathcal{H} := \{H_I\}_I$, where I are dyadic intervals of the form $I = [(j-1)2^{-n}, j2^{-n}), \quad j = 1, \dots, 2^n; \quad n = 0, 1, \dots$ and $I = [0,1]$ with

$$H_{[0,1]}(x) = 1 \quad \text{for } x \in [0,1) \quad ,$$

$$H_{[(j-1)2^{-n}, j2^{-n})} = \begin{cases} 2^{n/2}, & x \in [(j-1)2^{-n}, (j-1/2)2^{-n}) \\ -2^{n/2}, & x \in [(j-1/2)2^{-n}, j2^{-n}) \\ 0, & \text{otherwise.} \end{cases}$$

Consider the multivariate Haar basis $\mathcal{H}^d := \mathcal{H} \times \dots \times \mathcal{H}$ which consists of functions

$$H_I(x) = \prod_{j=1}^d H_{I_j}(x_j), \qquad I = I_1 \times \dots \times I_d, \quad x = (x_1, \dots, x_d).$$

For the Haar basis $\mathcal{H}^d$ we define for each $1 \le p \le \infty$ the Greedy Algorithm G^p which acts as follows. Denote

$$f_I := \langle f, H_I \rangle = \int_{[0,1]^d} f(x) H_I(x) dx \quad .$$

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

Let Λ be a set of m dyadic intervals I for which $\|f_I H_I\|_p$ take the biggest values. We set

$$G^p_m(f, \mathcal{H}^d) := \sum_{I \in \Lambda} f_I H_I \quad .$$

In order to get an idea about efficiency of Greedy Algorithm G^p we compare the error of its approximation $\|f - G^p_m(f, \mathcal{H}^d)\|_p$ with the best possible error of approximation by a linear combination of m Haar functions. Denote

$$\sigma_m(f, \mathcal{H}^d)_p := \inf \|f - \sum_{I \in \Lambda_m} c_I H_I\|_p, \quad m = 1, 2, \dots, \tag{1.1}$$

where inf is taken over coefficients c_I and sets Λ_m of dyadic intervals of cardinality $\#\Lambda_m = m$.

It was proved in [T1] that for any $1 < p < \infty$ we have for $f \in L_p$

$$\|f - G^p_m(f, \mathcal{H})\|_p \leq C(p)\sigma_m(f, \mathcal{H})_p \quad . \tag{1.2}$$

What means that the Greedy Algorithm G^p realizes near best m-term approximation. In this paper we study efficiency of Greedy Algorithm G^p with regard to multivariate Haar system $\mathcal{H}^d$. A simple example suggested by R. Hochmuth (see Section 4) shows that we do not have the relation (1.2) for $\mathcal{H}^d, \quad d \geq 2$, anymore and

$$e_m(G^p, \mathcal{H}^d) := \sup_{f \in L_p} (\|f - G^p_m(f, \mathcal{H}^d)\|_p / \sigma_m(f, \mathcal{H}^d)_p) \geq \tag{1.3}$$

$$C(p, d)(\log m)^{|1/2 - 1/p|(d-1)},$$

for $1 < p < \infty$.

In Section 2 we use one general inequality to prove that

$$e_m(G^p, \mathcal{H}^d) \leq C(p, d)(\log m)^d.$$

This inequality shows that efficiency of G^p with regard to $\mathcal{H}^d$ is not far from the optimal.

An interesting open question here is the following. What is the correct order of the quantity $e_m(G^p, \mathcal{H}^d)$?

Conjecture. *For $1 < p < \infty$ we have*

$$e_m(G^p, \mathcal{H}^d) \asymp (\log m)^{|1/2 - 1/p|(d-1)}. \tag{1.4}$$

The lower estimate in (1.4) follows from (1.3) what means we need to prove the complimentary upper bound in (1.4) in order to prove the conjecture. In Section 3 we prove the conjecture in the case $d = 2$ and $4/3 \leq p \leq 4$.

Similarly to [T1] we prove our results in a more general setting: for a system Ψ L_p-equivalent to $\mathcal{H}^d$ and for Greedy Algorithm G^p replaced by t-thresholding Greedy Algorithm $G^{p,t}$. For details see Sections 2 and 3.

In Section 5 we illustrate how the inequality (1.2) and some other results from [T1] can be used for proving direct and inverse theorems in m-term approximation. We prove two lemmas (see Lemmas 5.1 and 5.2) which establish the inequalities between the errors $\sigma_n(f)_p$ of best m-term approximation and the nonincreasing rearragement of $\{\|f_I H_I\|_p\}$. These lemmas allow us to reduce the problem of studying approximation errors to the problem of comparing two decreasing sequences. We use standard technique to derive from these lemmas some new and also some known results on direct and inverse theorems.

In Section 6 we consider the case $p = 1$ or $p = \infty$ and prove that efficiency of greedy algorithms G^p is not good in this case. Namely, we prove that

$$e_m(G^p, \mathcal{H}^d)_p \asymp m, \qquad p = 1, \infty.$$

2. Efficiency of multivariate Greedy Algorithm

Let $\Psi := \{\psi_I\}_I$ be a basis in $L_p([0,1]^d)$ indexed by dyadic intervals I. We say that Ψ is L_p-equivalent to $\mathcal{H}^d$ if there exist two positive constants $C_1(p,d)$ and $C_2(p,d)$ such that for any finite set of coefficients c_I we have

$$C_1(p,d)\|\sum_I c_I H_I\|_p \le \|\sum_I c_I \psi_I\|_p \le C_2(p,d)\|\sum_I c_I H_I\|_p. \tag{2.1}$$

For a given basis Ψ and $0 < t \le 1$ we define the t-thresholding Greedy Algorithm $G^{p,t}(\cdot, \Psi)$ as follows. Let

$$f = \sum_I c_I(f, \Psi)\psi_I$$

and

$$c_I(f, p, \Psi) := \|c_I(f, \Psi)\psi_I\|_p.$$

Then $c_I(f,p,\Psi) \to 0$ as $|I| \to 0$. Denote $\Lambda_m(t)$ a set of m dyadic intervals I such that

$$\min_{I \in \Lambda_m(t)} c_I(f,p,\Psi) \ge t \max_{J \notin \Lambda_m(t)} c_J(f,p,\Psi). \tag{2.2}$$

We define $G^{p,t}(\cdot, \Psi)$ by formula

$$G_m^{p,t}(f, \Psi) := \sum_{I \in \Lambda_m(t)} c_I(f, \Psi)\psi_I.$$

Denote

$$e_m(G^{p,t}, \Psi) := \sup_{f \in L_p} (\|f - G_m^{p,t}(f, \Psi)\|_p / \sigma_m(f, \Psi)_p), \tag{2.3}$$

where $\sigma_m(f, \Psi)_p$ is the best m-term approximation of f with regard to Ψ which is defined similarly to (1.1).

Theorem 2.1. *Let $1 < p < \infty$ and a basis Ψ be L_p-equivalent to $\mathcal{H}^d$. Then there exist two positive constants $C(p,d)$ and $c(p,d)$ such that*

$$e_m(G^{p,t}, \Psi) \le C(p,d) \sup_{G^{p,\tau}, \tau \ge tc(p,d)} e_m(G^{p,\tau}, \mathcal{H}^d).$$

Proof. This proof repeats the arguments in the proof of Theorem 2.1 from [T1]. For a given function $f \in L_p$ we define

$$g(f) := \sum_I c_I(f, \Psi) H_I.$$

It is clear that $g(f) \in L_p$ and

$$\sigma_m(g(f), \mathcal{H}^d)_p \le C_1(p,d)^{-1} \sigma_m(f, \Psi)_p. \tag{2.4}$$

Next, for any two intervals $I \in \Lambda_m(t)$, $J \notin \Lambda_m(t)$ by the definition of $\Lambda_m(t)$ we have

$$c_I(f,p,\Psi) \ge t c_J(f,p,\Psi).$$

Using (2.1) we get from here

$$\|g(f)_I H_I\|_p = \|c_I(f,\Psi) H_I\|_p \ge C_2(p,d)^{-1} \|c_I(f,\Psi)\psi_I\|_p = \tag{2.5}$$

$$= C_2(p,d)^{-1} c_I(f,p,\Psi) \ge t C_2(p,d)^{-1} c_J(f,p,\Psi) =$$

$$= t C_2(p,d)^{-1} \|c_J(f,\Psi)\psi_J\|_p \ge t C_1(p,d) C_2(p,d)^{-1} \|g(f)_J H_J\|_p.$$

This inequality implies that choosing $c(p,d) := C_1(p,d)C_2(p,d)^{-1}$ we have for any m that for $I \in \Lambda_m(t)$ and $J \notin \Lambda_m(t)$

$$\|g(f)_I H_I\|_p \ge tc(p,d) \|g(f)_J H_J\|_p,$$

and, therefore,

$$\|f - G_m^{p,t}(f,\Psi)\|_p \le C_2(p,d) \sup_{G^{p,\tau}, \tau \ge tc(p,d)} \|g(f) - G_m^{p,\tau}(g(f), \mathcal{H}^d)\|_p. \tag{2.6}$$

The relations (2.4) and (2.6) imply Theorem 2.1.

Theorem 2.2. *Let $1 < p < \infty$ and $0 < t \le 1$. Then for any $f \in L_p$ we have*

$$\|f - G_m^{p,t}(f, \mathcal{H}^d)\|_p \le C(p,t,d)(\log m)^d \sigma_m(f, \mathcal{H}^d)_p.$$

Proof. Denote

$$g_{\Lambda,p} := \sum_{I \in \Lambda} |I|^{1/2-1/p} H_I$$

and

$$\mu(m,p,d) := \sup_{k \le m} (\sup_\Lambda \|g_{\Lambda,p}\|_p / \inf_\Lambda \|g_{\Lambda,p}\|_p)$$

where sup and inf are taken over all sets Λ of dyadic intervals I with the same cardinality $\#\Lambda = k$. We prove first the following inequality.

Lemma 2.1. *Let $1 < p < \infty$ and $0 < t \leq 1$. Then for any $f \in L_p$ we have*

$$\|f - G_m^{p,t}(f, \mathcal{H}^d)\|_p \leq C(p,t,d)\mu(m,p,d)\sigma_m(f, \mathcal{H}^d)_p.$$

Proof. Let T_m be the m-term Haar polynomial of best m-term approximation to f in L_p (for existence see [D]):

$$T_m = \sum_{I \in \Lambda} a_I H_I, \quad \#\Lambda = m.$$

For any finite set Q of dyadic intervals we denote by S_Q the projector

$$S_Q(f) := \sum_{I \in Q} f_I H_I.$$

The Littlewood-Paley Theorem for Haar system (see for instance [KS]) gives for $1 < p < \infty$

$$C_3(p,d)\|(\sum_I |f_I H_I|^2)^{1/2}\|_p \leq \|f\|_p \leq C_4(p,d)\|(\sum_I |f_I H_I|^2)^{1/2}\|_p. \tag{2.7}$$

From (2.7) we get

$$\|f - S_\Lambda(f)\|_p = \|f - T_m - S_\Lambda(f - T_m)\|_p \leq \|Id - S_\Lambda\|_{p\to p}\sigma_m(f, \mathcal{H}^d)_p \leq \tag{2.8}$$

$$C_4(p,d)C_3(p,d)^{-1}\sigma_m(f, \mathcal{H}^d)_p,$$

where Id denotes the identical operator. Further, we have

$$G_m^{p,t}(f) = S_{\Lambda_m(t)}(f),$$

and

$$\|f - G_m^{p,t}(f)\|_p \leq \|f - S_\Lambda(f)\|_p + \|S_\Lambda(f) - S_{\Lambda_m(t)}(f)\|_p. \tag{2.9}$$

The first term in the right side of (2.9) has been estimated in (2.8). We estimate now the second term. We represent it in the form

$$S_\Lambda(f) - S_{\Lambda_m(t)}(f) = S_{\Lambda\setminus\Lambda_m(t)}(f) - S_{\Lambda_m(t)\setminus\Lambda}(f)$$

and remark that similarly to (2.8) we get

$$\|S_{\Lambda_m(t)\setminus\Lambda}(f)\|_p \leq C_5(p,d)\sigma_m(f, \mathcal{H}^d)_p. \tag{2.10}$$

We estimate now $\|S_{\Lambda\setminus\Lambda_m(t)}(f)\|_p$ and prove that

$$\|S_{\Lambda\setminus\Lambda_m(t)}(f)\|_p \leq C(p,t,d)\mu(m,p,d)\|S_{\Lambda_m(t)\setminus\Lambda}(f)\|_p. \tag{2.11}$$

Denote

$$A := \max_{I \in \Lambda \setminus \Lambda_m(t)} \|f_I H_I\|_p,$$

and

$$B := \min_{I \in \Lambda_m(t) \setminus \Lambda} \|f_I H_I\|_p.$$

Then by the definition of $\Lambda_m(t)$ we have

$$B \geq tA. \tag{2.12}$$

Using the Littlewood-Paley inequality (2.7) we get on one hand

$$\|S_{\Lambda \setminus \Lambda_m(t)}(f)\|_p \ll A \|g_{\Lambda \setminus \Lambda_m(t),p}\|_p \tag{2.13}$$

and on the other hand

$$\|S_{\Lambda_m(t) \setminus \Lambda}(f)\|_p \gg B \|g_{\Lambda_m(t) \setminus \Lambda,p}\|_p, \tag{2.14}$$

with the constants depending on p and d. Using (2.12) we obtain (2.11) from (2.13) and (2.14).

Theorem 2.2 follows from Lemma 2.1 and the following Lemma 2.2.

Lemma 2.2. *For any Λ, $\#\Lambda = m$, we have for*

$$g_{\Lambda,p} := \sum_{I \in \Lambda} |I|^{1/2-1/p} H_I$$

the inequalities

$$m^{1/p} \ll \|g_{\Lambda,p}\|_p \ll m^{1/p} (\log m)^d, \quad 2 \leq p < \infty; \tag{2.15}$$

$$m^{1/p} (\log m)^{-d} \ll \|g_{\Lambda,p}\|_p \ll m^{1/p}, \quad 1 < p \leq 2; \tag{2.16}$$

with constants depending on p and d.

Proof. The following two simple corollaries of the Littlewood-Paley Theorem are well known

$$\|f\|_p \leq C_6(p,d) (\sum_I \|f_I H_I\|_p^p)^{1/p}, \qquad 1 < p \leq 2, \tag{2.17}$$

$$\|f\|_p \geq C_7(p,d) (\sum_I \|f_I H_I\|_p^p)^{1/p}, \qquad 2 \leq p < \infty. \tag{2.18}$$

These inequalities imply the lower estimate in (2.15) and the upper estimate in (2.16). We prove now the upper estimate in (2.15) and then by duality arguments derive from it the lower estimate in (2.16). This proof is based on one embedding type inequality. In order to formulate it we introduce some notations. For a given l denote D_l the set of dyadic intervals of length 2^{-l}, i.e.,

$$D_l := \{I = [(j-1)2^{-l}, j2^{-l}), j = 1, \dots, 2^l\}, \quad l \geq 1, \qquad D_0 := \{[0,1], [0,1)\}.$$

For a vector $s = (s_1, \dots, s_d)$, where $s_1, \dots, s_d$ are nonnegative integers, denote

$$P_s := \{I = I_1 \times \dots \times I_d, \quad I_j \in D_{s_j}, j = 1, \dots, d\}.$$

Let $f \in L_p([0,1]^d)$, denote

$$\delta_s(f) := \sum_{I \in P_s} f_I H_I.$$

The following inequality is known (see [T2], Lemma 2.3). Let $1 \leq q < p < \infty$; then for $f \in L_q([0,1]^d)$ we have

$$\|f\|_p \leq C(q,p,d)(\sum_s (\|\delta_s(f)\|_q 2^{(1/q-1/p)\|s\|_1})^p)^{1/p}. \tag{2.19}$$

We will also need an upper bound on the constant $C(q,p,d)$ in this inequality. Examining the proof of inequality (2.19) (see [T3], pp.24–26) we get

$$C(q,p,d) \leq C(d)(1/q - 1/p)^{-d}. \tag{2.20}$$

Now we are in the position to prove the upper estimate in (2.15). Let $2 \leq p < \infty$ and

$$g_{\Lambda,p} = \sum_{I \in \Lambda} |I|^{1/2-1/p} H_I, \qquad \#\Lambda = m.$$

Denote

$$\Lambda_s := \Lambda \cap P_s, \qquad n_s := \#\Lambda_s.$$

Then

$$\sum_s n_s = m.$$

Next, for any s and q we have

$$\| \sum_{I \in \Lambda_s} |I|^{1/2-1/p} H_I \|_q = n_s^{1/q} 2^{\|s\|_1(1/p-1/q)}. \tag{2.21}$$

Using (2.19) and (2.21) with $q < p$ we obtain

$$\|g_{\Lambda,p}\|_p \leq C(q,p,d)(\sum_s n_s^{p/q})^{1/p} \leq C(q,p,d) m^{1/q-1/p} m^{1/p}.$$

Taking into account (2.20) and specifying q as $1/q - 1/p = (\ln m)^{-1}$ we get

$$\|g_{\Lambda,p}\|_p \leq C(d)(\ln m)^d m^{1/p},$$

what completes the proof of upper estimate in (2.15).

The lower estimate in (2.16) follows from the upper estimate in (2.15) which has been proved and the inequality

$$m = \int_{[0,1]^d} g_{\Lambda,p} g_{\Lambda,p'} \leq \|g_{\Lambda,p}\|_p \|g_{\Lambda,p'}\|_{p'}.$$

Theorem 2.2 is proved now.

Proof of Conjecture in the case $d = 2$ and $4/3 \le p \le 4$

The upper estimate

$$e_m(G^p, \mathcal{H}^2) \ll (\log m)^{|1/2-1/p|}$$

for $4/3 \le p \le 4$ follows from Lemma 2.1 and the following Lemma 3.1.

Lemma 3.1. *Let $4/3 \le p \le 4$ and $d = 2$. For any Λ, $\#\Lambda = m$ we have*

$$m^{1/p} \ll \|g_{\Lambda,p}\|_p \ll m^{1/p}(\log m)^{1/2-1/p}, \quad 2 \le p \le 4; \tag{3.1}$$

$$m^{1/p}(\log m)^{1/2-1/p} \ll \|g_{\Lambda,p}\|_p \ll m^{1/p}, \quad 4/3 \le p \le 2. \tag{3.2}$$

Proof. We prove only the upper estimate in (3.1) and remark that it implies the lower estimate in (3.2) in the same way as in the proof of Lemma 2.2.

We consider first the case $p = 4$. Let

$$g_{\Lambda,4} = \sum_{I\in\Lambda} |I|^{1/4} H_I, \qquad \#\Lambda = m.$$

Denote as above $\Lambda_s := \Lambda \cap P_s$, $n_s := \#\Lambda_s$ and $A_s := \cup_{I\in\Lambda_s} I$. Then

$$\sum_{I\in\Lambda_s} (|I|^{1/4}H_I)^2 = 2^{\|s\|_1/2}\chi_{A_s},$$

where χ_A is a characteristic function of a set A. By the Littlewood-Paley Theorem we have

$$\|g_{\Lambda,4}\|_4^4 \ll \int_{[0,1]^2} (\sum_{I\in\Lambda}(|I|^{1/4}H_I)^2)^2 dx = \int_{[0,1]^2} (\sum_s 2^{\|s\|_1/2}\chi_{A_s})^2 dx = \tag{3.3}$$

$$\int_{[0,1]^2} \sum_s \sum_{s'} 2^{\|s\|_1/2}\chi_{A_s} 2^{\|s'\|_1/2}\chi_{A_{s'}} dx.$$

Denote

$$J_{s,s'} := \int_{[0,1]^2} \chi_{A_s}\chi_{A_{s'}} dx.$$

Then we get from (3.3)

$$\|g_{\Lambda,4}\|_4^4 \ll \sum_s \sum_{s'} 2^{(\|s\|_1+\|s'\|_1)/2} J_{s,s'}. \tag{3.4}$$

Let us estimate $J_{s,s'}$. It is obvious that

$$J_{s,s'} \le \min\{\operatorname{mes} A_s, \operatorname{mes} A_{s'}\} = \min\{n_s 2^{-\|s\|_1}, n_{s'} 2^{-\|s'\|_1}\}. \tag{3.5}$$

Next for any $I \in \Lambda_s$ and $I' \in \Lambda_{s'}$ we have

$$\operatorname{mes} I \cap I' \le 2^{-\max(s_1,s_1')-\max(s_2,s_2')}$$

and therefore

$$J_{s,s'} \le n_s n_{s'} 2^{-\max(s_1,s_1')-\max(s_2,s_2')}. \tag{3.6}$$

We estimate now

$$a(s) := \sum_{s'} J_{s,s'} 2^{\|s'\|_1/2}.$$

Denote $\|s\|_1 = n$. Then we have

$$a(s) = \sum_k 2^{k/2} \sum_{\|s'\|_1=k} J_{s,s'} = \sum_{k<n} 2^{k/2} \sum_{\|s'\|_1=k} J_{s,s'} + \sum_{k\ge n} 2^{k/2} \sum_{\|s'\|_1=k} J_{s,s'} =: a_1(s) + a_2(s).$$

For s and s' such that $\|s\|_1 = n$, $\|s'\|_1 = k$ we have

$$\max(s_1,s_1') + \max(s_2,s_2') \ge \min(n,k) + |s_1 - s_1'|.$$

For fixed s with $\|s\|_1 = n$ introduce the notations

$$V(s) := \{s' : \quad \|s'\|_1 = k; \quad |s_1' - s_1| \le |k-n| + \log m\},$$

$$U(s) := \{s' : \quad \|s'\|_1 = k\} \setminus V(s).$$

We estimate $a_1(s)$ now. Let $k < n$. Then we have

$$\sum_{\|s'\|_1=k} J_{s,s'} \le \sum_{\|s'\|_1=k} \min\{n_s n_{s'} 2^{-k-|s_1-s_1'|}, n_s 2^{-n}\} \le$$

$$\sum_{s'\in V(s)} n_s 2^{-n} + n_s \sum_{s'\in U(s)} m 2^{-k-|s_1-s_1'|} \ll (n-k+\log m) n_s 2^{-n} + n_s 2^{-n}$$

and therefore

$$a_1(s) \ll \sum_{k<n} 2^{k/2} \sum_{\|s'\|_1=k} J_{s,s'} \ll n_s 2^{-n/2} \log m = n_s 2^{-\|s\|_1/2} \log m.$$

Hence

$$\sum_s a_1(s) 2^{\|s\|_1/2} \ll m \log m. \tag{3.7}$$

We proceed to estimating $a_2(s)$ now. We have $k \ge n$ and

$$\sum_{\|s'\|_1=k} J_{s,s'} \le \sum_{\|s'\|_1=k} \min\{n_s n_{s'} 2^{-n-|s_1-s_1'|}, n_{s'} 2^{-k}\} \le$$

$$\sum_{s'\in V(s)} n_{s'}2^{-k} + n_s \sum_{s'\in U(s)} m2^{-n-|s_1-s_1'|} \ll$$

$$2^{-k}\sum_{s'\in V(s)} n_{s'} + n_s 2^{-k}.$$

Therefore

$$a_2(s) \ll \sum_{k\ge n} 2^{-k/2} n_s + \sum_{k\ge n} 2^{-k/2} \sum_{s'\in V(s)} n_{s'}$$

and

$$\sum_s 2^{\|s\|_1/2} a_2(s) \ll \sum_s n_s + \sum_n 2^{n/2} \sum_{k\ge n} 2^{-k/2} \sum_{\|s\|_1=n} \sum_{s'\in V(s)} n_{s'} \ll \tag{3.8}$$

$$m + \sum_n 2^{n/2} \sum_{k\ge n} 2^{-k/2} \sum_{\|s'\|_1=k} n_{s'}(k-n+\log m).$$

Introducing the notation $N_k := \sum_{\|s\|_1=k} n_s$ we get from (3.8)

$$\sum_s 2^{\|s\|_1/2} a_2(s) \ll m + \sum_k 2^{-k/2} \sum_{n\le k} 2^{n/2}(k-n+\log m)N_k \ll m\log m.$$

This completes the proof in the case $p=4$.

Let now $2<p<4$. By Littlewood-Paley Theorem we have

$$A := \|g_{\Lambda,p}\|_p \ll \|(\sum_s \sum_{I\in\Lambda_s} 2^{2\|s\|_1/p}\chi_I)^{1/2}\|_p = (\int_{[0,1]^2}(\sum_s 2^{2\|s\|_1/p}\chi_{A_s})^{p/2}dx)^{1/p}.$$

Using the Hölder inequality with $1/q = 4/p-1$ we get

$$\sum_s 2^{2\|s\|_1/p}\chi_{A_s} \le (\sum_s 2^{\|s\|_1}\chi_{A_s})^{4/p-1}(\sum_s 2^{\|s\|_1/2}\chi_{A_s})^{2-4/p}. \tag{3.9}$$

Denote

$$h_p := \sum_s 2^{2\|s\|_1/p}\chi_{A_s}.$$

Using (3.9) and applying the Hölder inequality with $1/q = 2-p/2$ we get

$$A^p \ll \int_{[0,1]^2} h_p^{p/2}dx \le \int_{[0,1]^2} h_2^{2-p/2}h_4^{p-2}dx \tag{3.10}$$

$$\le (\int_{[0,1]^2} h_2 dx)^{2-p/2}(\int_{[0,1]^2} h_4^2 dx)^{(p-2)/2}.$$

Next,

$$\int_{[0,1]^2} h_2 = m,$$

and by the above considered case $p=4$

$$\int_{[0,1]^2} h_4^2 dx \ll m\log m.$$

Therefore we have from (3.10)

$$A^p \ll m(\log m)^{p/2-1}.$$

This completes the proof in the case $2\le p\le 4$.

4. The lower estimate in Conjecture

We will give here an example which provides the estimate (1.3). As we mentioned in Introduction this example was constructed by R. Hochmuth. For each $n \in \mathbb{N}$ we define two sets A and B of dyadic intervals I as follows

$$A := \{I : \quad \operatorname{mes} I = 2^{-n}\};$$

$$B := \{I : \quad I \notin A, \forall I \neq I' \text{we have} I \cap I' = \emptyset; \quad \#B = \#A\}.$$

Let $2 \le p < \infty$ be given. Denote $m = \#A$ and consider

$$f = g_{A,p} + 2g_{B,p}$$

where $g_{\Lambda,p}$ is defined in Section 2 (see for instance Lemma 2.2). Then we have on one hand

$$G_m^p(f, \mathcal{H}^d) = 2g_{B,p}$$

and

$$\|f - G_m^p(f, \mathcal{H}^d\|_p = \|g_{A,p}\|_p \gg m^{1/p}(\log m)^{(1/2-1/p)(d-1)}. \tag{4.1}$$

On the other hand we have

$$\sigma_m(f, \mathcal{H}^d)_p \le \|2g_{B,p}\|_p \ll m^{1/p}. \tag{4.2}$$

The relations (4.1) and (4.2) imply the required lower estimate in the case $2 \le p < \infty$. The remaining case $1 < p \le 2$ can be handled in the same way considering the function $f = 2g_{A,p} + g_{B,p}$.

5. Some direct and inverse theorems in m-term approximation

In this Section we prove necessary and sufficient conditions for f to have a prescribed decay of $\{\sigma_n(f, \Psi)_p\}$. These conditions are formulated in terms of $\|c_I(f, \Psi)\psi_I\|_p$, what is convenient in numerical applications. We formulate a general statement and then consider several important particular examples of rate of decrease of $\{\sigma_n(f, \Psi)_p\}$. We use a method which is based on results from [T1]. The same method can be used for approximation in the multivariate case of tensor product of univariate bases, for instance, in the case of $\mathcal{H}^d$. The necessary and sufficient conditions are also can be given in terms of $\|c_I(f, \Psi)\psi_I\|_p$ but to the contrary to the one-dimensional case or to the case of multivariate wavelet bases with isotropic supports these conditions do not coincide. In this case instead of results from [T1] one is supposed to use results from Sections 2 and 3 of this paper. We give a presentation of results in this section in terms of Haar basis keeping in mind that all results hold for any system Ψ L_p-equivalent to $\mathcal{H}$.

We begin by introducing some notations. For a monotonically decreasing to zero sequence $\mathcal{E} = \{\epsilon_k\}_{k=0}^{\infty}$ of positive numbers (we write $\mathcal{E} \in MDP$) we define inductively a sequence $\{N_s\}_{s=0}^{\infty}$ of nonnegative integers:

(5.1)

$$N_0 = 0; \quad N_{s+1} \quad \text{is the smallest satisfying} \quad \epsilon_{N_{s+1}} < \frac{1}{2}\epsilon_{N_s}; \qquad n_s := N_{s+1} - N_s.$$

We are going to consider the following examples of sequences.

Example A. *Take* $\epsilon_0 = 1$ *and* $\epsilon_k = k^{-r}, \quad r > 0, \quad k = 1, 2, \dots$. *Then*

$$N_{s+1} = [2^{1/r} N_s] + 1 \quad and \quad n_s = [2^{1/r} N_s] + 1 - N_s.$$

What implies

$$N_s \asymp 2^{s/r} \quad and \quad n_s \asymp 2^{s/r}.$$

Example B. *Fix* $0 < b < 1$ *and take* $\epsilon_k = 2^{-k^b}, \quad k = 0, 1, 2, \dots$. *Then*

$$N_s = s^{1/b} + O(1) \quad and \quad n_s \asymp s^{1/b - 1}.$$

Let $f \in L_p$. Rearrange the sequence $\|f_I H_I\|_p$ in decreasing order

$$\|f_{I_1} H_{I_1}\|_p \geq \|f_{I_2} H_{I_2}\|_p \geq \dots$$

and denote

$$a_k(f, p) := \|f_{I_k} H_{I_k}\|_p.$$

We prove now some inequalities for $a_n(f, p)$ and $\sigma_m(f, \mathcal{H})_p$. In this section we use brief notation $\sigma_m(f)_p := \sigma_m(f, \mathcal{H})_p$ and $\sigma_0(f)_p := \|f\|_p$.

Lemma 5.1. *For any two positive integers* $N < M$ *we have*

$$a_M(f, p) \leq C(p) \sigma_N(f)_p (M - N)^{-1/p}.$$

Proof. By Theorem 2.1 from [T1] we have for all m

$$\|f - G^p_m(f, \mathcal{H})\|_p \leq C(p) \sigma_m(f)_p.$$

From here and definition of G^p_m we get

$$J := \| \sum_{k=N+1}^{M} f_{I_k} H_{I_k} \|_p \leq C(p)(\sigma_N(f)_p + \sigma_M(f)_p). \tag{5.2}$$

Next, we have for $k \in (N, M]$

$$\|f_{I_k} H_{I_k}\|_p \geq \|f_{I_M} H_{I_M}\|_p = a_M(f, p)$$

and by Lemma 2.2 from [T1] we get from here

$$a_M(f, p)(M - N)^{1/p} \leq C(p) J. \tag{5.3}$$

Relations (5.2) and (5.3) imply the conclusion of Lemma 5.1.

Lemma 5.2. *For any sequence $m_0 < m_1 < m_2 < \dots$ of nonnegative integers we have*

$$\sigma_{m_s}(f)_p \le C(p) \sum_{l=s}^{\infty} a_{m_l}(f,p)(m_{l+1} - m_l)^{1/p}.$$

Proof. We have

$$\sigma_{m_s}(f)_p \le \| \sum_{k>m_s} f_{I_k} H_{I_k} \|_p \le \sum_{l=s}^{\infty} \| \sum_{k\in(m_l,m_{l+1}]} f_{I_k} H_{I_k} \|_p.$$

Using Lemma 2.1 from [T1] we get from here

$$\sigma_{m_s}(f)_p \le \sum_{l=s}^{\infty} a_{m_l}(f,p)(m_{l+1} - m_l)^{1/p}$$

what proves the lemma.

Theorem 5.1. *Assume a given sequence $\mathcal{E} \in MDP$ satisfies the conditions*

$$\epsilon_{N_s} \ge C_1 2^{-s}, \qquad n_{s+1} \le C_2 n_s, \qquad s = 0, 1, 2, \dots.$$

Then we have the equivalence

$$\sigma_n(f)_p \ll \epsilon_n \quad \Longleftrightarrow \quad a_{N_s}(f,p) \ll 2^{-s} n_s^{-1/p}.$$

Proof. We prove first $\Rightarrow$. We use Lemma 5.1 with $M = N_{s+1}$ and $N = N_s$

$$a_{N_{s+1}}(f,p) \le C(p)\sigma_{N_s}(f)_p n_s^{-1/p} \le C(p) 2^{-s-1}(n_{s+1}/C_2)^{-1/p}$$

what implies the statement of Theorem 5.1 in this direction.

We prove the inverse statement now $\Leftarrow$. Using Lemma 5.2 we get

$$\sigma_{N_s}(f)_p \ll \sum_{l=s}^{\infty} a_{N_l}(f,p)(N_{l+1} - N_l)^{1/p} \ll \sum_{l=s}^{\infty} 2^{-l} \ll 2^{-s} \ll \epsilon_{N_s}$$

and for $n \in [N_s, N_{s+1})$

$$\sigma_n(f)_p \le \sigma_{N_s}(f)_p \ll \epsilon_{N_s}(f)_p \ll 2^{-s} \ll \epsilon_{N_{s+1}}(f)_p \le \epsilon_n(f)_p.$$

Corollary 5.1. *Theorem 5.1 applied to Examples A, B gives the following relations:*

(5.A) $$\sigma_m(f)_p \ll (m+1)^{-r} \quad \Longleftrightarrow \quad a_n(f,p) \ll n^{-r-1/p},$$

(5.B) $$\sigma_m(f)_p \ll 2^{-m^b} \quad \Longleftrightarrow \quad a_n(f,p) \ll 2^{-n^b} n^{(1-1/b)/p}.$$

Remark 5.1. *Making use of Lemmas 5.1 and 5.2 we can prove a version of Corollary 5.1 with sing $\ll$ replaced by $\asymp$.*

Theorem 5.1 and Corollary 5.1 are in spirit of classical Jackson-Bernstein direct and inverse theorems in linear approximation theory, where conditions on the corresponding sequences of approximating characteristics are imposed in the form

$$E_n(f)_p \ll \epsilon_n, \quad \text{or} \quad \|E_n(f)_p/\epsilon_n\|_{l_\infty} < \infty. \tag{5.4}$$

It is well known that in studying many questions of approximation theory it is convenient to consider along with restriction (5.4) the following its generalization

$$\|E_n(f)_p/\epsilon_n\|_{l_q} < \infty. \tag{5.5}$$

Lemmas 5.1 and 5.2 are also useful in considering this more general case. For instance, in the particular case of Example A one gets the following statement.

Theorem 5.2. *Let $1 < p < \infty$ and $0 < q < \infty$. Then for any positive r we have the equivalence relation*

$$\sum_m \sigma_m(f)_p^q m^{rq-1} < \infty \quad \Longleftrightarrow \quad \sum_n a_n(f,p)_p^q n^{rq-1+q/p} < \infty.$$

Proof. Using Lemma 5.1 with $M = 2^{s+1}$ and $N = 2^s$ we get

$$\sum_s a_{2^s}(f,p)_p^q 2^{s(rq+1/p)} \le C(p) \sum_s \sigma_{2^s}(f)_p^q 2^{srq}$$

what proves the implication $\Rightarrow$ in the theorem.

Using Lemma 5.2 with $m_0 = 0$ and $m_s = 2^s$ for $s = 1, 2, \dots$ we get

$$\sum_s \sigma_{2^s}(f)_p^q 2^{srq} \le C(p) \sum_s (\sum_{l \ge s} a_{2^l}(f,p)_p 2^{l/p})^q 2^{rqs} \le$$

$$C(p) \sum_s 2^{rqs} (\sum_{l \ge s} a_{2^l}(f,p)_p 2^{l(r+1/p)} 2^{-lr})^q \le$$

by Hölder inequality

$$C(p) \sum_s (\sum_{l \ge s} a_{2^l}(f,p)_p^q 2^{l(r+1/p)q})^q \le C(p) \sum_l a_{2^l}(f,p)_p^q 2^{l(r+1/p)q}$$

what completes the proof of Theorem 5.2.

Remark 5.2. *The condition*

$$\sum_n a_n(f,p)_p^q n^{rq-1+q/p} < \infty$$

with $q = \tau := (r + 1/p)^{-1}$ takes a very simple form

$$\sum_n a_n(f,p)_p^\tau = \sum_I \|f_I H_I\|_p^\tau < \infty. \tag{5.6}$$

Rewriting

$$\|f_I H_I\|_p = \|f_I H_I\|_\tau |I|^{1/p-1/\tau} = \|f_I H_I\|_\tau |I|^{-r}$$

we get that the condition (5.6) is equivalent to f is in Besov space $B_\tau^r(L_\tau)$.

Corollary 5.2. *Theorem 5.2 implies the following relation*

$$\sum_m \sigma_m(f)_p^\tau n^{r\tau-1} < \infty \qquad \Longleftrightarrow \qquad f \in B^r_\tau(L_\tau),$$

where $\tau := (r+1/p)^{-1}$.

The statement similar to Corollary 5.2 for free knots spline approximation was proved by P. Petrushev [P]. Corollary 5.2 and further results in this direction can be found in [DP] and [DJP]. We want to remark here that conditions in terms of $a_n(f,p)$ are convenient in applications. For instance, the relation (5.A) can be rewritten using the idea of thresholding. For a given $f \in L_p$ denote

$$T(\epsilon) := \#\{a_k(f,p) \quad : \quad a_k(f,p) \geq \epsilon\}.$$

Then (5.A) is equivalent to

$$\sigma_m(f)_p \ll (n+1)^{-r} \qquad \Longleftrightarrow \qquad T(\epsilon) \ll \epsilon^{-(r+1/p)^{-1}}.$$

6. Efficiency of Greedy Algorithm in the cases $p=1$ or $p=\infty$

In this section we consider approximation with regard to the Haar multivariate system $\mathcal{H}^d$. It turns out that efficiency of greedy algorithms G^p, $p=1,\infty$, drops down dramatically comparing to the case $1<p<\infty$.

Theorem 6.1. *Let* $p=1$ *or* $p=\infty$. *Then we have*

$$e_m(G^p,\mathcal{H}^d) \asymp m \quad .$$

Proof. We first prove the upper estimates. Let

$$t_\Lambda = \sum_{I\in\Lambda} c_I H_I$$

be a best m-term approximant to a given $f \in L_p$, $p=1$ or $p=\infty$ (for existence see [D]). Denote Λ_m a set of m dyadic intervals I for which $\|f_I H_I\|_p$ take the biggest values. We need to estimate

$$\delta := \|f - G^p_m(f,\mathcal{H}^d)\|_p = \|\sum_{I\notin\Lambda_m} f_I H_I\|_p \quad .$$

We have

$$\delta = \|\sum_{I\notin\Lambda} f_I H_I - \sum_{I\in\Lambda_m\setminus\Lambda} f_I H_I + \sum_{I\in\Lambda\setminus\Lambda_m} f_I H_I\|_p \leq \tag{6.1}$$

$$\|\sum_{I\notin\Lambda} f_I H_I\|_p + \|\sum_{I\in\Lambda_m\setminus\Lambda} f_I H_I\|_p + \|\sum_{I\in\Lambda\setminus\Lambda_m} f_I H_I\|_p =: \delta_1+\delta_2+\delta_3.$$

Let p' denote the dual to p ($p' = \infty$ if $p = 1$ and $p' = 1$ if $p = \infty$). Then we have

$$(6.2) \qquad |(f - t_\Lambda)_I| \le \sigma_m(f, \mathcal{H}^d)_p \|H_I\|_{p'},$$

and for $I \notin \Lambda$ we get

$$(6.3) \qquad |f_I| \le \sigma_m(f, \mathcal{H}^d)_p \|H_I\|_{p'}.$$

Next, by the deffinition of Λ_m and by (6.3) we have

$$\max_{I \in \Lambda \setminus \Lambda_m} \|f_I H_I\|_p \le \min_{J \in \Lambda_m \setminus \Lambda} \|f_J H_J\|_p \le \sigma_m(f, \mathcal{H}^d)_p \quad .$$

Therefore, for $\delta_i, i = 2, 3$ we get

$$(6.4) \qquad \delta_i \le \#(\Lambda \setminus \Lambda_m) \sigma_m(f, \mathcal{H}^d)_p, \quad i = 2, 3.$$

It remains to estimate δ_1. We have by (6.2)

$$(6.5) \qquad \delta_1 \le \|f - t_\Lambda\|_p + \|\sum_{I \in \Lambda} (f - t_\Lambda)_I H_I\|_p \le \sigma_m(f, \mathcal{H}^d)_p + \#\Lambda \sigma_m(f, \mathcal{H}^d)_p.$$

Combining (6.1),(6.4) and (6.5) we obtain

$$\delta \le (3m + 1) \sigma_m(f, \mathcal{H}^d)_p \quad .$$

We prove now the lower bounds. We consider the two cases $p = 1$ and $p = \infty$ separately. In both cases we construct an example for $d = 1$.

Case 1: $p = 1$. Let m be given. Consider two functions f_1 and f_2. Denote $I_k := [0, 2^{-k})$ and define

$$f_1 := \sum_{k=1}^{m} |I_k|^{-1/2} H_{I_k} \quad .$$

It is easy to check that

$$f_1 = \begin{cases} 2^{m+1} - 2, & x \in [0, 2^{-m-1}) \\ -2, & x \in [2^{-m-1}, 1/2). \end{cases}$$

Let A be any set of m disjoint dyadic intervals J such that $J \cap [0, 1/2) = \emptyset$. Denote

$$f_2 := \sum_{J \in A} |J|^{-1/2} H_J \quad .$$

Consider m-term approximation in L_1 of the function $f = 2f_1 + f_2$. We have

$$(6.6) \qquad \sigma_m(f, \mathcal{H})_1 \le 2\|f_1\|_1 \le 4,$$

and

$$(6.7) \qquad \|f - G_m^1(f, \mathcal{H})\|_1 = \|f_2\|_1 = m.$$

Case 2: $p = \infty$**.** We use functions similar to those from the previous case. Define

$$g_1 := \sum_{k=1}^{m} |I_k|^{1/2} H_{I_k}$$

and

$$g_2 := \sum_{J \in A} |J|^{1/2} H_J \quad .$$

Consider the function $g = g_1 + 2g_2$. Then

$$\sigma_m(g, \mathcal{H})_\infty \leq 2\|g_2\|_\infty = 2 \tag{6.8}$$

and

$$\|g - G_m^\infty(g, \mathcal{H})\|_\infty = \|g_1\| = m. \tag{6.9}$$

The relations (6.6), (6.7) and (6.8), (6.9) imply the lower estimates in Theorem 6.1.

References

[D] V.V. Dubinin, *Greedy Algorithms and Applications*, Dissertation USC, 1997.

[DP] R.A. DeVore and V.A. Popov, *Interpolation spaces and non-linear approximation*, Lecture Notes in Mathematics **1302** (1988), 191–205.

[DJP] R. DeVore, B. Jawerth, V. Popov, *Compression of wavelet decompositions*, American Journal of Math. **114** (1992), 737-785.

[KS] B.S. Kashin and A.A. Saakyan, *Orthogonal Series*, American Math. Soc., Providence, R.I., 1989.

[P] P. Petrushev, *Direct and converse theorems for spline and rational approximation and Besov spaces*, Lecture Notes in Mathematics **1302** (1988), 363–377.

[T1] V.N. Temlyakov, *The best m-term approximation and Greedy Algorithms* (to appear).

[T2] V.N. Temlyakov, *Some inequalities for multivariate Haar polynomials*, East Journal on Approximations **1** (1995), 61–72.

[T3] V. Temlyakov, *Approximation of functions with bounded mixed derivative*, Proc. Steklov Institute (1989, Issue 1).