IMI

Preprint Series

Industrial Mathematics Institute

2003:06

Greedy expansions in Banach spaces

V.N. Temlyakov

Department of Mathematics
University of South Carolina

Greedy expansions in Banach Spaces[1]

V.N.Temlyakov

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

Abstract. We study covergence and rate of convergence of expansions of elements in a Banach space X into series with regard to a given dictionary $\mathcal{D}$. For convenience we assume that $\mathcal{D}$ is symmetric: $g \in \mathcal{D}$ implies $-g \in \mathcal{D}$. The primary goal of this paper is to study representations of an element $f \in X$ by a series

$$f \sim \sum_{j=1}^{\infty} c_j(f) g_j(f), \quad g_j(f) \in \mathcal{D}, \quad c_j(f) > 0, \quad j = 1, 2, \dots.$$

In building such a representation we should construct two sequences: $\{g_j(f)\}_{j=1}^{\infty}$ and $\{c_j(f)\}_{j=1}^{\infty}$. In this paper the construction of $\{g_j(f)\}_{j=1}^{\infty}$ will be based on ideas used in greedy-type nonlinear approximation. This explains the use of the term *greedy expansion*. We use a norming functional $F_{f_{m-1}}$ of a residual f_{m-1} obtained after $m-1$ steps of an expansion procedure to select the mth element $g_m(f) \in \mathcal{D}$ from the dictionary. This approach has been used in previous papers on greedy approximation. The new feature of this paper is a way of selecting the mth coefficient $c_m(f)$ of the expansion. An approach developed in the paper works in any uniformly smooth Banach space. For instance, in a Banach space X with the modulus of smoothness $\rho(u)$ we can choose $c_m(f)$ from the equation

$$\|f_{m-1}\| \rho(c_m(f)/\|f_{m-1}\|) = \frac{tb}{2} c_m(f) \sup_{g \in \mathcal{D}} F_{f_{m-1}}(g),$$

where $t \in (0,1]$ is the weakness parameter of an algorithm and $b \in (0,1)$ is its tuning parameter. We prove convergence of such expansions for all $f \in X$ and obtain rate of convergence for $f \in A_1(\mathcal{D})$ - the closure (in X) of the convex hull of $\mathcal{D}$.

1. Introduction

Let X be a Banach space with norm $\|\cdot\|$. We say that a set of elements (functions) $\mathcal{D}$ from X is a dictionary if each $g \in \mathcal{D}$ has norm one ($\|g\| = 1$), and $\overline{\operatorname{span}} \mathcal{D} = X$. In addition we assume for convenience that

$$g \in \mathcal{D} \quad \text{implies} \quad -g \in \mathcal{D}.$$

Thus from the definition of a dictionary it follows that any element $f \in X$ can be approximated arbitrarily well by finite linear combinations of the dictionary

[1]This research was supported by the National Science Foundation Grant DMS 0200187 and by ONR Grant N00014-91-J1343

elements. The primary goal of this paper is to study representations of an element $f \in X$ by a series

$$f \sim \sum_{j=1}^{\infty} c_j(f) g_j(f), \quad g_j(f) \in \mathcal{D}, \quad c_j(f) > 0, \quad j = 1, 2, \dots. \tag{1.1}$$

In building a representation (1.1) we should construct two sequences: $\{g_j(f)\}_{j=1}^{\infty}$ and $\{c_j(f)\}_{j=1}^{\infty}$. In this paper the construction of $\{g_j(f)\}_{j=1}^{\infty}$ will be based on ideas used in greedy-type nonlinear approximation (greedy-type algorithms). This justifies the use of the term *greedy expansion* for (1.1) considered in the paper. The construction of $\{g_j(f)\}_{j=1}^{\infty}$ is, clearly, the most important and difficult part in building a represenration (1.1). On the base of contemporary theory of nonlinear approximation with regard to redundant dictionaries we may conclude that the method of using a norming functional in greedy steps of an algorithm is the most productive in approximation in Banach spaces. This method has been utilized in the Weak Chebyshev Greedy Algorithm and in the Weak Dual Greedy Algorithm (see below). We use this same method in new algorithms considered in the paper. A new qualitative result of the paper establishes that we have a lot of flexibility in constructing a sequence of coefficients $\{c_j(f)\}_{j=1}^{\infty}$. For instance, in Section 3 we make on observation that at each step of the Pure Greedy Algorithm (see below) we can choose a fixed fraction of the optimal coefficient for that step instead of the optimal coefficient itself. Surprisingly, this leads to better upper estimates than those known for the Pure Greedy Algorithm (see Section 4 for details).

We will study in this paper greedy algorithms with regard to $\mathcal{D}$ that provide greedy expansions. For a nonzero element $f \in X$ we denote by F_f a norming (peak) functional for f:

$$\|F_f\| = 1, \qquad F_f(f) = \|f\|.$$

The existence of such a functional is guaranteed by Hahn-Banach theorem. Denote

$$r_{\mathcal{D}}(f) := \sup_{F_f} \sup_{g \in \mathcal{D}} F_f(g).$$

We note that in general a norming functional F_f is not unique. This is why we take $\sup_{F_f}$ over all norming functionals of f in the definition of $r_{\mathcal{D}}(f)$. It is known that in the case of uniformly smooth Banach spaces (our primary object here) the norming functional F_f is unique. In such a case we do not need $\sup_{F_f}$ in the definition of $r_{\mathcal{D}}(f)$. We begin our discussion with the Weak Chebyshev Greedy Algorithm (WCGA) that is the best studied. This algorithm does not provide an expansion. However, the results on converegence (rate of convergence) for this algorithm may serve as a bench mark in the study of convergence properties of expantion (1.1). Then we proceed to the Pure Greedy Algorithm (PGA) and its generalization the Weak Greedy Algorithm (WGA) that provide an expansion in Hilbert spaces. We complete our discussion of known results by presenting definitions of and some results for natural generalizations of the PGA (or WGA) to the case of Banach spaces: the X-Greedy Algorithm and the Weak Dual Greedy Algorithm (WDGA). These algorithms are not as well studied as either the WGA in Hilbert spaces or the WCGA in Banach spaces. Finally, we turn to a new modification of the WDGA. This modification, similarly to the WDGA, provides an expansion. In this paper we prove convergence results for the above mentioned modification of the WDGA. We note that these results are more general than the known results for the WDGA. We also obtain results on the rate of convergence for elements from the set $A_1(\mathcal{D})$ that is the closure (in X) of the convex hull of $\mathcal{D}$. We are not aware of any general results on the rate of convergence of the WDGA for elements from $A_1(\mathcal{D})$. We also note that the new modification considered here is even simpler than the WDGA from the point of view of algorithmical implementation.

Let $\tau := \{t_k\}_{k=1}^\infty$ be a given sequence of nonnegative numbers $t_k \le 1$, $k = 1, \dots$. We define first (see [T3]) the Weak Chebyshev Greedy Algorithm (WCGA) that is a generalization for Banach spaces of Weak Orthogonal Greedy Algorithm defined and studied in [T1] (see also [DT] for Orthogonal Greedy Algorithm).

Weak Chebyshev Greedy Algorithm (WCGA). We define $f_0^c := f_0^{c,\tau} := f$. Then for each $m \ge 1$ we inductively define

1). $\varphi_m^c := \varphi_m^{c,\tau} \in \mathcal{D}$ is any satisfying

$$F_{f_{m-1}^c}(\varphi_m^c) \ge t_m r_{\mathcal{D}}(f_{m-1}^c).$$

2). Define

$$\Phi_m := \Phi_m^\tau := \operatorname{span}\{\varphi_j^c\}_{j=1}^m,$$

and define $G_m^c := G_m^{c,\tau}$ to be the best approximant to f from Φ_m.

3). Denote

$$f_m^c := f_m^{c,\tau} := f - G_m^c.$$

We consider here approximation in uniformly smooth Banach spaces. For a Banach space X we define the modulus of smoothness

$$\rho(u) := \sup_{\|x\|=\|y\|=1} \left(\frac{1}{2}(\|x+uy\| + \|x-uy\|) - 1\right).$$

A uniformly smooth Banach space is one with the property

$$\lim_{u\to 0} \rho(u)/u = 0.$$

It is easy to see that for any Banach space X its modulus of smoothness $\rho(u)$ is an even convex function satisfying the inequalities

$$\max(0, u-1) \le \rho(u) \le u, \quad u \in (0,\infty). \tag{1.2}$$

It is well known (see for instance [DGDS], Lemma B.1) that in the case $X = L_p$, $1 \le p < \infty$ we have

$$\rho(u) \le \begin{cases} u^p/p & \text{if} \quad 1 \le p \le 2, \\ (p-1)u^2/2 & \text{if} \quad 2 \le p < \infty. \end{cases} \tag{1.3}$$

It is also known (see [LT], p.63) that for any X with $\dim X = \infty$ one has

$$\rho(u) \geq (1+u^2)^{1/2} - 1$$

and for every X, $\dim X \geq 2$,

$$\rho(u) \geq Cu^2, \quad C > 0.$$

This limits power type modulus of smoothness of nontrivial Banach spaces to the case $1 \leq q \leq 2$.

We begin a discussion of known results with a theorem on convergence of WCGA [T3]. In the formulation of this theorem we need a special sequence which is defined for a given modulus of smoothness $\rho(u)$ and a given $\tau = \{t_k\}_{k=1}^\infty$.

Definition 1.1. *Let $\rho(u)$ be an even convex function on $(-\infty, \infty)$ with the property: $\rho(2) \geq 1$ and*

$$\lim_{u \to 0} \rho(u)/u = 0.$$

For any $\tau = \{t_k\}_{k=1}^\infty$, $0 \leq t_k \leq 1$, and $0 < \theta \leq 1/2$ we define $\xi_m := \xi_m(\rho, \tau, \theta)$ as a number u satisfying the equation

$$\rho(u) = \theta t_m u. \tag{1.4}$$

Remark 1.1. *Assumptions on $\rho(u)$ imply that the function*

$$\epsilon(u) := \rho(u)/u, \quad u \neq 0, \quad \epsilon(0) = 0,$$

is a continuous increasing on $[0, \infty)$ function with $\epsilon(2) \geq 1/2$. Thus (1.4) has a unique solution $0 \leq \xi_m \leq 2$.

The following theorem and a corollary have been proved in [T3].

Theorem 1.1. *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u)$. Assume that a sequence $\tau := \{t_k\}_{k=1}^\infty$ satisfies the condition: for any $\theta > 0$ we have*

$$\sum_{m=1}^{\infty} t_m \xi_m(\rho, \tau, \theta) = \infty. \tag{1.5}$$

Then for any $f \in X$ we have

$$\lim_{m \to \infty} \|f_m^{c,\tau}\| = 0.$$

Corollary 1.1. *Let a Banach space X have modulus of smoothness $\rho(u)$ of power type $1 < q \leq 2$; ($\rho(u) \leq \gamma u^q$). Assume that*

$$\sum_{m=1}^{\infty} t_m^p = \infty, \quad p = \frac{q}{q-1}.$$

Then WCGA converges for any $f \in X$.

The following theorem has been proved in [T3] for the WCGA.

Theorem 1.2. *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u) \le \gamma u^q$, $1 < q \le 2$. Then for a sequence $\tau := \{t_k\}_{k=1}^\infty$, $t_k \le 1$, $k = 1, 2, \dots$, we have for any $f \in A_1(\mathcal{D})$ that*

$$\|f_m^{c,\tau}\| \le C(q,\gamma)(1 + \sum_{k=1}^m t_k^p)^{-1/p}, \quad p := \frac{q}{q-1},$$

with a constant $C(q,\gamma)$ which may depend only on q and γ.

Theorems 1.1 and 1.2 provide convergence results in a very general situation: X is any uniformly smooth Banach space, τ is any satisfying (1.5) (in particular $\tau = \{t\}$, $t \in (0,1]$). However, as we already mentioned above the WCGA does not provide an expansion (1.1). We now proceed to algorithms that provide an expansion (1.1). Unfortunately, these algorithms are not as good as the WCGA in the sence of convergence.

Let us begin this discussion in the special case of a Hilbert space with the inner product $\langle \cdot, \cdot \rangle$. We define first the Pure Greedy Algorithm (PGA) in Hilbert space H. We describe this algorithm for a general dictionary $\mathcal{D}$. If $f \in H$, we let $g(f) \in \mathcal{D}$ be an element from $\mathcal{D}$ which maximizes $\langle f, g \rangle$. We will assume for simplicity that such a maximizer exists; if not suitable modifications are necessary (see Weak Greedy Algorithm below) in the algorithm that follows. We define

$$G(f,\mathcal{D}) := \langle f, g(f) \rangle g(f)$$

and

$$R(f,\mathcal{D}) := f - G(f,\mathcal{D}).$$

Pure Greedy Algorithm (PGA). *We define $R_0(f,\mathcal{D}) := f$ and $G_0(f,\mathcal{D}) := 0$. Then, for each $m \ge 1$, we inductively define*

$$G_m(f,\mathcal{D}) := G_{m-1}(f,\mathcal{D}) + G(R_{m-1}(f,\mathcal{D}),\mathcal{D})$$

$$R_m(f,\mathcal{D}) := f - G_m(f,\mathcal{D}) = R(R_{m-1}(f,\mathcal{D}),\mathcal{D}).$$

Let a sequence $\tau = \{t_k\}_{k=1}^\infty$, $0 \le t_k \le 1$, be given. Following [T1] we define the Weak Greedy Algorithm.

Weak Greedy Algorithm (WGA). *We define $f_0^\tau := f$. Then for each $m \ge 1$, we inductively define:*

1). $\varphi_m^\tau \in \mathcal{D}$ is any satisfying

$$\langle f_{m-1}^\tau, \varphi_m^\tau \rangle \ge t_m \sup_{g \in \mathcal{D}} \langle f_{m-1}^\tau, g \rangle;$$

2).

$$f_m^\tau := f_{m-1}^\tau - \langle f_{m-1}^\tau, \varphi_m^\tau \rangle \varphi_m^\tau;$$

3).

$$G^\tau_m(f,\mathcal{D}) := \sum_{j=1}^m \langle f^\tau_{j-1},\varphi^\tau_j\rangle\varphi^\tau_j.$$

We note that in a particular case $t_k = t$, $k = 1,2,\dots$, this algorithm was considered in [J]. Thus, the WGA is a generalization of the PGA in the direction of making it easier to construct an element φ^τ_m at the mth greedy step. We point out that the WGA contains, in addition to the first (greedy) step, the second step (see 2), 3) in the above definition) where we update an approximant by adding an orthogonal projection of the residual f^τ_{m-1} onto φ^τ_m. Therefore, the WGA provides for each $f\in H$ an expansion into a series (greedy expansion)

$$f \sim \sum_{j=1}^\infty c_j(f)\varphi^\tau_j, \quad c_j(f) := \langle f^\tau_{j-1},\varphi^\tau_j\rangle.$$

In general it is not an expansion into an orthogonal series but it has some similar properties. The coefficients $c_j(f)$ of an expansion are obtained by the Fourier formulas with f replaced by the residuals f^τ_{j-1}. It is easy to see that

$$\|f^\tau_m\|^2 = \|f^\tau_{m-1}\|^2 - |c_m(f)|^2.$$

In the case of convergent greedy expansion (see, for instance, Theorem 1.3 below) we get for this expansion an analog of the Parseval formula for orthogonal expansions:

$$\|f\|^2 = \sum_{j=1}^\infty |c_j(f)|^2.$$

We proved in [T2] a criterion on τ for convergence of WGA. To explain this we need some notation.

We define by $\mathcal{V}$ the class of sequences $x = \{x_k\}_{k=1}^\infty$, $x_k \ge 0$, $k = 1,2,\dots$, with the following property: there exists a sequence $0 = q_0 < q_1 < \dots$ such that

$$\sum_{s=1}^\infty \frac{2^s}{\Delta q_s} < \infty;$$

and

$$\sum_{s=1}^\infty 2^{-s}\sum_{k=1}^{q_s} x_k^2 < \infty,$$

where $\Delta q_s := q_s - q_{s-1}$.

Theorem 1.3. *The condition $\tau \notin \mathcal{V}$ is necessary and sufficient for convergence of Weak Greedy Algorithm with weakness sequence τ for each f and all Hilbert spaces H and dictionaries $\mathcal{D}$.*

For a general dictionary $\mathcal{D}$ we define the class of functions

$$\mathcal{A}_1^o(\mathcal{D},M) := \{f\in H : f = \sum_{k\in\Lambda} c_k w_k, \quad w_k\in\mathcal{D},\ \#\Lambda<\infty \text{ and } \sum_{k\in\Lambda}|c_k| \le M\}$$

and we define $\mathcal{A}_1(\mathcal{D}, M)$ as the closure (in X) of $\mathcal{A}_1^o(\mathcal{D}, M)$. Furthermore, we define $\mathcal{A}_1(\mathcal{D})$ as the union of the classes $\mathcal{A}_1(\mathcal{D}, M)$ over all $M > 0$. For $f \in \mathcal{A}_1(\mathcal{D})$, we define the norm

$$|f|_{\mathcal{A}_1(\mathcal{D})}$$

as the smallest M such that $f \in \mathcal{A}_1(\mathcal{D}, M)$.

It was proved in [DT] that for a general dictionary $\mathcal{D}$ the Pure Greedy Algorithm provides the following estimate

$$\|f - G_m(f, \mathcal{D})\| \leq |f|_{\mathcal{A}_1(\mathcal{D})} m^{-1/6}. \tag{1.6}$$

(In this and similar estimates we consider that the inequality holds for all possible choices of $\{G_m\}$.) We proved in [KT] an estimate

$$\|f - G_m(f, \mathcal{D})\| \leq 4|f|_{\mathcal{A}_1(\mathcal{D})} m^{-11/62} \tag{1.7}$$

which improves a little the original one (see (1.6)).

Recently, we proved in [LiT] that there exist a dictionary $\mathcal{D}$ and an element $f \in H$, $f \neq 0$, such that

$$\|f - G_m(f, \mathcal{D})\| \geq C m^{-0.27} |f|_{\mathcal{A}_1(\mathcal{D})} \tag{1.8}$$

with a positive constant C.

Much less is known about greedy expansions with regard to general redundant dictionaries in the case of a general Banach space X. We discuss next two versions of generalization of PGA from a Hilbert space H to a Banach space X. The first one is a straightforward generalization of PGA. We call it Pure Greedy Algorithm or X-Greedy Algorithm when we want to indicate a Banach space. For a given X and $\mathcal{D}$ we define $G(f, \mathcal{D}, X) := \alpha(f) g(f)$ where $\alpha(f) \in \mathbb{R}$ and $g(f) \in \mathcal{D}$ satisfy (we assume existence) the relation

$$\min_{\alpha \in \mathbb{R}, g \in \mathcal{D}} \|f - \alpha g\| = \|f - \alpha(f) g(f)\|.$$

X-Greedy Algorithm. *We define $R_0(f, \mathcal{D}, X) := f$ and $G_0(f, \mathcal{D}, X) := 0$. Then, for each $m \geq 1$, we inductively define*

$$R_m(f) := R_m(f, \mathcal{D}, X) := R_{m-1}(f) - G(R_{m-1}(f), \mathcal{D}, X)$$

$$G_m(f, \mathcal{D}, X) := G_{m-1}(f, \mathcal{D}, X) + G(R_{m-1}(f), \mathcal{D}, X).$$

The second version of PGA in a Banach space is based on the concept of norming functional. We call it the Dual Greedy Algorithm (DGA). Let a dictionary $\mathcal{D}$ in X be given. Take an element $f \in X$ and find a norming functional F_f. Now the basic step of PGA is modified to the following. Assume that there exists $g_f \in \mathcal{D}$ such that

$$F_f(g_f) = \max_{g \in \mathcal{D}} F_f(g).$$

We take this g_f and solve one more optimization problem: find a number a such that

$$\|f - ag_f\|_X = \min_b \|f - bg_f\|_X.$$

We put

$$G^D(f,\mathcal{D}) := ag_f, \quad R^D(f,\mathcal{D}) := f - ag_f.$$

Repeating this step m times we get $G^D_m(f,\mathcal{D})$ as an approximant and $R^D_m(f,\mathcal{D})$ as a residual.

Let us give a definition of the Weak Dual Greedy Algorithm (see [T4,p.66]).

Weak Dual Greedy Algorithm (WDGA). *We define $f^D_0 := f^{D,\tau}_0 := f$. Then for each $m \ge 1$ we inductively define*

1). $\varphi^D_m := \varphi^{D,\tau}_m \in \mathcal{D}$ is any satisfying

$$F_{f^D_{m-1}}(\varphi^D_m) \ge t_m r_{\mathcal{D}}(f^D_{m-1}). \tag{1.9}$$

2). Define a_m as

$$\|f^D_{m-1} - a_m\varphi^D_m\| = \min_{a\in\mathbb{R}} \|f^D_{m-1} - a\varphi^D_m\|.$$

3). Denote

$$f^D_m := f^{D,\tau}_m := f^D_{m-1} - a_m\varphi^D_m.$$

It is clear that in the case $\tau = \{1\}$ the WDGA coincides with the DGA defined above. The following conjecture has been formulated in [T4,p.73, Open problem 4.3]: the Dual Greedy Algorithm converges for all dictionaries $\mathcal{D}$ and each element $f \in X$ in uniformly smooth Banach spaces X with modulus of smoothness of fixed power type q, $1 < q \le 2$, ($\rho(u) \le \gamma u^q$).

Recently, M. Ganichev and N.J. Kalton [GK] have proved the following very interesting result.

Theorem 1.4. *Let $\tau = \{t\}$, $t \in (0,1]$ and $X = L_p$, $1 < p < \infty$. Then the WDGA converges for any dictionary $\mathcal{D}$ for all functions $f \in L_p$.*

We consider here a modification of the WDGA that is a little simpler than the WDGA. One more good feature of this new modification is that we can prove its convergence in every uniformly smooth Banach space. In addition to this we obtain good rate of convergence of it for elements from $A_1(\mathcal{D})$. We note that other greedy-type algorithms in Banach spaces have been recently introduced and studied by E.D. Livshitz [L].

We begin with a description of a general scheme that provides an expansion for a given element f. Later, specifying this general scheme, we will obtain different methods of expansion.

Dual Based Expansion (DBE). *Let $t \in (0,1]$ and $f \neq 0$. Denote $f_0 := f$. Assume $\{f_j\}_{j=0}^{m-1} \subset X$, $\{\varphi_j\}_{j=1}^{m-1} \subset \mathcal{D}$ and a set of coefficients $\{c_j\}_{j=1}^{m-1}$ of expansion have already been constructed. If $f_{m-1} = 0$ then we stop (set $c_j = 0$, $j = m, m+1, \dots$ in the expansion) and get $f = \sum_{j=1}^{m-1} c_j \varphi_j$. If $f_{m-1} \neq 0$ then we*

1). choose $\varphi_m \in \mathcal{D}$ such that

$$F_{f_{m-1}}(\varphi_m) \geq tr_{\mathcal{D}}(f_{m-1});$$

2). define

$$f_m := f_{m-1} - c_m \varphi_m,$$

where $c_m > 0$ is a coefficient either prescribed in advance or chosen from a concrete approximation procedure.

We call the series

$$f \sim \sum_{j=1}^{\infty} c_j \varphi_j \tag{1.10}$$

the Dual Based Expansion of f with coefficients $c_j(f) := c_j$, $j = 1, 2, \dots$ with regard to $\mathcal{D}$.

Denote

$$S_m(f, \mathcal{D}) := \sum_{j=1}^{m} c_j \varphi_j.$$

Then it is clear that

$$f_m = f - S_m(f, \mathcal{D}).$$

We prove some convergence results for the DBE in Sections 2 and 3. In Section 3 we consider a variant of the Dual Based Expansion with coefficients chosen by a certain simple rule. The rule depends on two numerical parameters $t \in (0,1]$ (the weakness parameter from the definition of the DBE) and $b \in (0,1)$ (the tuning parameter of the approximation method). The rule also depends on a majorant μ of the modulus of smoothness of the Banach space X.

Dual Greedy Algorithm with parameters (t, b, μ) **(DGA**(t, b, μ)**).** *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u)$ and let $\mu(u)$ be a majorant of $\rho(u)$: $\rho(u) \leq \mu(u)$, $u \in [0, \infty)$. For parameters $t \in (0,1]$, $b \in (0,1)$ we define sequences $\{f_m\}_{m=0}^{\infty}$, $\{\varphi_m\}_{m=1}^{\infty}$, $\{c_m\}_{m=1}^{\infty}$ inductively. Let $f_0 := f$. If for $m \geq 1$ $f_{m-1} = 0$ then we set $f_j = 0$ for $j \geq m$ and stop. If $f_{m-1} \neq 0$ then we conduct the following three steps:*

1). take any $\varphi_m \in \mathcal{D}$ such that

$$F_{f_{m-1}}(\varphi_m) \geq tr_{\mathcal{D}}(f_{m-1}); \tag{1.11}$$

2). choose $c_m > 0$ from the equation

$$\|f_{m-1}\| \mu(c_m / \|f_{m-1}\|) = \frac{tb}{2} c_m r_{\mathcal{D}}(f_{m-1}); \tag{1.12}$$

3). define

$$f_m := f_{m-1} - c_m \varphi_m. \tag{1.13}$$

In Section 3 we prove the following convergence result.

Theorem 1.5. *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u)$ and let $\mu(u)$ be a majorant of $\rho(u)$ with the property $\mu(u)/u \downarrow 0$ as $u \to +0$. Then for any $t \in (0,1]$ and $b \in (0,1)$ the DGA(t,b,μ) converges for each dictionary $\mathcal{D}$ and all $f \in X$.*

The following result from Section 3 gives the rate of convergence.

Theorem 1.6. *Assume X has a modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$. Denote $\mu(u) = \gamma u^q$. Then for any dictionary $\mathcal{D}$ and any $f \in A_1(\mathcal{D})$ the rate of convergence of the DGA(t,b,μ) is given by*

$$\|f_m\| \le C(t,b,\gamma,q) m^{-\frac{t(1-b)}{p(1+t(1-b))}}, \quad p := \frac{q}{q-1}.$$

2. Convergence of the Dual Based Expansion

We begin with the following lemma.

Lemma 2.1. *Let $f \in X$. Assume that the coefficients $\{c_j\}_{j=1}^{\infty}$ of the expansion*

$$f \sim \sum_{j=1}^{\infty} c_j \varphi_j, \qquad f_m := f - \sum_{j=1}^{m} c_j \varphi_j$$

satisfy the following two conditions

$$\sum_{j=1}^{\infty} c_j r_{\mathcal{D}}(f_{j-1}) < \infty, \tag{2.1}$$

$$\sum_{j=1}^{\infty} c_j = \infty. \tag{2.2}$$

Then

$$\liminf_{m \to \infty} \|f_m\| = 0. \tag{2.3}$$

Proof. The proof of this lemma is similar to the proof of Lemma 1 from [GK]. Denote $s_n := \sum_{j=1}^{n} c_j$. Then (2.2) implies (see [B,p.904]) that

$$\sum_{n=1}^{\infty} c_n/s_n = \infty. \tag{2.4}$$

Using (2.1) we get

$$\sum_{n=1}^{\infty} s_n r_{\mathcal{D}}(f_{n-1}) c_n/s_n = \sum_{n=1}^{\infty} c_n r_{\mathcal{D}}(f_{n-1}) < \infty.$$

Thus by (2.4)

$$\liminf_{n\to\infty} s_n r_{\mathcal{D}}(f_{n-1}) = 0$$

and also $(s_{n-1} \le s_n)$

$$\liminf_{n\to\infty} s_n r_{\mathcal{D}}(f_n) = 0.$$

Let

$$\lim_{k\to\infty} s_{n_k} r_{\mathcal{D}}(f_{n_k}) = 0. \tag{2.5}$$

Consider $\{F_{f_{n_k}}\}$. The unit sphere in the dual X^* is weakly* compact (see [HHZ,p.45]). Let $\{F_i\}_{i=1}^{\infty}$, $F_i := F_{f_{n_{k_i}}}$ be a w^*-convergent subsequence. Denote

$$F := w^* - \lim_{i\to\infty} F_i.$$

We will complete the proof of Lemma 2.1 by contradiction. We assume that (2.3) does not hold: $\exists \alpha > 0$ and $\exists N \in \mathbb{N}$:

$$\|f_m\| \ge \alpha, \quad m \ge N, \tag{2.6}$$

and will get a contradiction.

We begin by deriving from (2.6) that $F \ne 0$. Indeed, we have

$$F(f) = \lim_{i\to\infty} F_i(f) \tag{2.7}$$

and

$$F_i(f) = F_i(f_{n_{k_i}} + \sum_{j=1}^{n_{k_i}} c_j \varphi_j) = \|f_{n_{k_i}}\| + \sum_{j=1}^{n_{k_i}} c_j F_i(\varphi_j) \ge \alpha - s_{n_{k_i}} r_{\mathcal{D}}(f_{n_{k_i}}) \tag{2.8}$$

for big i. Relations (2.7), (2.8), and (2.5) imply that $F(f) \ge \alpha$ and, hence $F \ne 0$. This implies that $\exists g \in \mathcal{D} : F(g) > 0$. However,

$$F(g) = \lim_{i\to\infty} F_i(g) \le \lim_{i\to\infty} r_{\mathcal{D}}(f_{n_{k_i}}) = 0.$$

We got a contradiction that completes the proof of Lemma 2.1.

Let us now consider a variant of the Dual Based Expansion when the coefficient sequence $\mathcal{C} = \{c_j\}_{j=1}^{\infty}$ is prescribed in advance and does not depend on f. We will call such a procedure the Weak Dual Greedy Approximation with coefficients $\mathcal{C}$ (WDGA($\mathcal{C}$)).

Theorem 2.1. *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u)$. Assume $\mathcal{C} = \{c_j\}_{j=1}^{\infty}$ is such that*

$$\sum_{j=1}^{\infty} c_j = \infty,$$

and for any $y > 0$

$$\sum_{j=1}^{\infty} \rho(yc_j) < \infty. \tag{2.9}$$

Then for the Dual Based Expansion of any $f \in X$ with coefficients $\mathcal{C}$ with regard to any dictionary $\mathcal{D}$ we have

$$\liminf_{m\to\infty} \|f_m\| = 0. \tag{2.10}$$

Proof. The proof is by contradiction. Assume (2.10) does not hold. Then $\exists \alpha > 0$ and $\exists N \in \mathbb{N}$ such that for all $m \geq N$

$$\|f_m\| \geq \alpha > 0.$$

From the definition of the modulus of smoothness we have

$$\|f_{n-1} - c_n\varphi_n\| + \|f_{n-1} + c_n\varphi_n\| \leq 2\|f_{n-1}\|(1 + \rho(c_n/\|f_{n-1}\|)). \tag{2.11}$$

Using the definition of φ_n:

$$F_{f_{n-1}}(\varphi_n) \geq tr_{\mathcal{D}}(f_{n-1}) \tag{2.12}$$

we get

$$\begin{aligned} \|f_{n-1} + c_n\varphi_n\| &\geq F_{f_{n-1}}(f_{n-1} + c_n\varphi_n) \\ &= \|f_{n-1}\| + c_nF_{f_{n-1}}(\varphi_n) \geq \|f_{n-1}\| + c_n tr_{\mathcal{D}}(f_{n-1}). \end{aligned} \tag{2.13}$$

Combining (2.11) and (2.13) we get

$$\|f_n\| = \|f_{n-1} - c_n\varphi_n\| \leq \|f_{n-1}\|(1 + 2\rho(c_n/\|f_{n-1}\|)) - c_n tr_{\mathcal{D}}(f_{n-1}). \tag{2.14}$$

We note that by Remark 1.1

$$\|f_{n-1}\|\rho(c_n/\|f_{n-1}\|) \leq \alpha\rho(c_n/\alpha), \quad n > N.$$

Therefore, by the assumption (2.9)

$$\sum_{n=1}^{\infty} \|f_{n-1}\|\rho(c_n/\|f_{n-1}\|) < \infty. \tag{2.15}$$

This and (2.14) imply

$$\sum_{n=1}^{\infty} c_n r_{\mathcal{D}}(f_{n-1}) \leq t^{-1}(\|f\| + 2\sum_{n=1}^{\infty} \|f_{n-1}\|\rho(c_n/\|f_{n-1}\|)) < \infty.$$

It remains to apply Lemma 2.1 to complete the proof.

3. A modification of the Weak Dual Greedy Algorithm

We begin this section with a proof of Theorem 1.5. We give a definition of the DGA(τ,b,μ), $\tau=\{t_k\}_{k=1}^\infty$, $t_k\in(0,1]$ here.

Dual Greedy Algorithm with parameters (τ,b,μ) **(DGA(τ,b,μ)).** *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u)$ and let $\mu(u)$ be a majorant of $\rho(u)$: $\rho(u)\le\mu(u)$, $u\in[0,\infty)$. For a sequence $\tau=\{t_k\}_{k=1}^\infty$, $t_k\in(0,1]$ and a parameter $b\in(0,1)$ we define sequences $\{f_m\}_{m=0}^\infty$, $\{\varphi_m\}_{m=1}^\infty$, $\{c_m\}_{m=1}^\infty$ inductively. Let $f_0:=f$. If for $m\ge1$ $f_{m-1}=0$ then we set $f_j=0$ for $j\ge m$ and stop. If $f_{m-1}\ne0$ then we conduct the following three steps:*

1). take any $\varphi_m\in\mathcal{D}$ such that

$$F_{f_{m-1}}(\varphi_m)\ge t_m r_{\mathcal{D}}(f_{m-1}); \tag{3.1}$$

2). choose $c_m>0$ from the equation

$$\|f_{m-1}\|\mu(c_m/\|f_{m-1}\|)=\frac{t_m b}{2}c_m r_{\mathcal{D}}(f_{m-1}); \tag{3.2}$$

3). define

$$f_m:=f_{m-1}-c_m\varphi_m. \tag{3.3}$$

Proof of Theorem 1.5. In this case $\tau=\{t\}$, $t\in(0,1]$. We have by (2.14)

$$\|f_m\|=\|f_{m-1}-c_m\varphi_m\|\le\|f_{m-1}\|(1+2\rho(c_m/\|f_{m-1}\|))-c_m t r_{\mathcal{D}}(f_{m-1}). \tag{3.4}$$

Using the choice of c_m we get from here

$$\|f_m\|\le\|f_{m-1}\|-t(1-b)c_m r_{\mathcal{D}}(f_{m-1}). \tag{3.5}$$

In particular, (3.5) implies that $\{\|f_m\|\}$ is a monotone decreasing and

$$t(1-b)c_m r_{\mathcal{D}}(f_{m-1})\le\|f_{m-1}\|-\|f_m\|.$$

Thus

$$\sum_{m=1}^\infty c_m r_{\mathcal{D}}(f_{m-1})<\infty. \tag{3.6}$$

We have the following two cases:

$$(I)\quad \sum_{m=1}^\infty c_m=\infty,\qquad (II)\quad \sum_{m=1}^\infty c_m<\infty.$$

In the first case by Lemma 2.1 we obtain

$$\liminf_{m\to\infty} \|f_m\| = 0 \quad \Rightarrow \quad \lim_{m\to\infty} \|f_m\| = 0.$$

It remains to consider the case (II). We prove convergence in this case by contradiction. Assume

$$\lim_{m\to\infty} \|f_m\| = \alpha > 0. \tag{3.7}$$

By (II) we have $f_m \to f_\infty \neq 0$ as $m\to\infty$. We note that by uniform smoothness of X we get

$$\lim_{m\to\infty} \|F_{f_m} - F_{f_\infty}\| = 0.$$

We have $F_{f_\infty} \neq 0$ and therefore there is a $g \in \mathcal{D}$ such that $F_{f_\infty}(g) > 0$. However,

$$F_{f_\infty}(g) = \lim_{m\to\infty} F_{f_m}(g) \le \lim_{m\to\infty} r_{\mathcal{D}}(f_m) = 0. \tag{3.8}$$

Indeed, by (3.2) and (3.7) we get

$$r_{\mathcal{D}}(f_{m-1}) \le \alpha c_m^{-1} \mu(c_m/\alpha)\frac{2}{tb} \to 0$$

as $m\to\infty$.

Theorem 1.5 is proved.

We proceed to studying the rate of convergence of the DGA(τ, b, μ) in the uniformly smooth Banach spaces with the power type majorant of the modulus of smoothness: $\rho(u) \le \mu(u) = \gamma u^q$, $1 < q \le 2$. We now prove a statement more general that Theorem 1.6.

Theorem 3.1. *Let $\tau := \{t_k\}_{k=1}^\infty$ be a nonincreasing sequence $1 \ge t_1 \ge t_2 \dots > 0$ and $b \in (0,1)$. Assume X has a modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$. Denote $\mu(u) = \gamma u^q$. Then for any dictionary $\mathcal{D}$ and any $f \in A_1(\mathcal{D})$ the rate of convergence of the DGA(τ, b, μ) is given by*

$$\|f_m\| \le C(b,\gamma,q)\Big(1 + \sum_{k=1}^m t_k^p\Big)^{-\frac{t_m(1-b)}{p(1+t_m(1-b))}}, \quad p := \frac{q}{q-1}.$$

Proof. Similar to (3.5) we get

$$\|f_m\| \le \|f_{m-1}\| - t_m(1-b)c_m r_{\mathcal{D}}(f_{m-1}). \tag{3.9}$$

Thus we need to estimate from below $c_m r_{\mathcal{D}}(f_{m-1})$. It is clear that

$$\|f_{m-1}\|_{\mathcal{A}_1(\mathcal{D})} = \|f - \sum_{j=1}^{m-1} c_j\varphi_j\|_{\mathcal{A}_1(\mathcal{D})} \le \|f\|_{\mathcal{A}_1(\mathcal{D})} + \sum_{j=1}^{m-1} c_j. \tag{3.10}$$

Denote $b_n := 1 + \sum_{j=1}^n c_j$. Then by (3.10) we get

$$\|f_{m-1}\|_{\mathcal{A}_1(\mathcal{D})} \le b_{m-1}.$$

Next, by Lemma 2.2 from [T3] we get

$$r_{\mathcal{D}}(f_{m-1}) = \sup_{g\in\mathcal{D}} F_{f_{m-1}}(g) = \sup_{\varphi\in A_1(\mathcal{D})} F_{f_{m-1}}(\varphi) \tag{3.11}$$

$$\ge \|f_{m-1}\|^{-1}_{\mathcal{A}_1(\mathcal{D})} F_{f_{m-1}}(f_{m-1}) \ge \|f_{m-1}\|/b_{m-1}.$$

Substituting (3.11) into (3.9) we get

$$\|f_m\| \le \|f_{m-1}\|(1 - t_m(1-b)c_m/b_{m-1}). \tag{3.12}$$

From the definition of b_m we find

$$b_m = b_{m-1} + c_m = b_{m-1}(1 + c_m/b_{m-1}).$$

Using the inequality

$$(1+x)^\alpha \le 1 + \alpha x, \quad 0 \le \alpha \le 1, \quad x \ge 0,$$

we obtain

$$b_m^{t_m(1-b)} \le b_{m-1}^{t_m(1-b)}(1 + t_m(1-b)c_m/b_{m-1}). \tag{3.13}$$

Multiplying (3.12) and (3.13) and using that $t_m \le t_{m-1}$ we get

$$\|f_m\| b_m^{t_m(1-b)} \le \|f_{m-1}\| b_{m-1}^{t_{m-1}(1-b)} \le \|f\| \le 1. \tag{3.14}$$

The function $\mu(u)/u = \gamma u^{q-1}$ is increasing on $[0,\infty)$. Therefore the c_m from (3.2) is greater than or equal to c'_m from (see (3.11))

$$\gamma\|f_{m-1}\|(c'_m/\|f_{m-1}\|)^q = \frac{t_m b}{2} c'_m \|f_{m-1}\|/b_{m-1}, \tag{3.15}$$

$$c'_m = (\frac{t_m b}{2\gamma})^{\frac{1}{q-1}} \frac{\|f_{m-1}\|^{\frac{q}{q-1}}}{b_{m-1}^{\frac{1}{q-1}}}. \tag{3.16}$$

Using notations

$$p := \frac{q}{q-1}, \qquad A^{-1} := (1-b)(\frac{b}{2\gamma})^{\frac{1}{q-1}} \le 1/2,$$

we get from (3.9), (3.11), (3.16)

$$\|f_m\| \le \|f_{m-1}\|(1-\frac{t_m^p}{A}\frac{\|f_{m-1}\|^p}{b_{m-1}^p}). \tag{3.17}$$

Noting that $b_m \ge b_{m-1}$ we derive from (3.17) that

$$(\|f_m\|/b_m)^p \le (\|f_{m-1}\|/b_{m-1})^p(1-A^{-1}t_m^p(\|f_{m-1}\|/b_{m-1})^p). \tag{3.18}$$

Taking into account that $\|f\| \le 1 < A$ we obtain from (3.18) by Lemma 3.1 from [T1]

$$(\|f_m\|/b_m)^p \le A(1+\sum_{k=1}^m t_k^p)^{-1}. \tag{3.19}$$

Combining (3.14) and (3.19) we get

$$\|f_m\| \le C(b,\gamma,q)(1+\sum_{k=1}^m t_k^p)^{-\frac{t_m(1-b)}{p(1+t_m(1-b))}}, \quad p := \frac{q}{q-1}.$$

This completes the proof of Theorem 3.1.

In the case $\tau = \{t\}$, $t \in (0,1]$ we get Theorem 1.6 from Theorem 3.1.

It follows from the proof of Theorem 3.1 that it holds for a modification of the DGA(τ,b,μ) where we replace in the definition the quantity $r_{\mathcal{D}}(f_{m-1})$ by its lower estimate (see (3.11)) $\|f_{m-1}\|/b_{m-1}$ with $b_{m-1} := 1+\sum_{j=1}^{m-1} c_j$. Clearly, this modification is more ready for practical implementation than the DGA(τ,b,μ). We formulate the above remark as a separate result.

Modified Dual Greedy Algorithm (τ,b,μ) **(MDGA**(τ,b,μ)**).** *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u)$ and let $\mu(u)$ be a majorant of $\rho(u)$: $\rho(u) \le \mu(u)$, $u \in [0,\infty)$. For a sequence $\tau = \{t_k\}_{k=1}^\infty$, $t_k \in (0,1]$ and a parameter $b \in (0,1)$ we define for $f \in A_1(\mathcal{D})$ sequences $\{f_m\}_{m=0}^\infty$, $\{\varphi_m\}_{m=1}^\infty$, $\{c_m\}_{m=1}^\infty$ inductively. Let $f_0 := f$. If for $m \ge 1$ $f_{m-1} = 0$ then we set $f_j = 0$ for $j \ge m$ and stop. If $f_{m-1} \ne 0$ then we conduct the following three steps:*

1). take any $\varphi_m \in \mathcal{D}$ such that

$$F_{f_{m-1}}(\varphi_m) \ge t_m\|f_{m-1}\|(1+\sum_{j=1}^{m-1} c_j)^{-1};$$

2). choose $c_m > 0$ from the equation

$$\mu(c_m/\|f_{m-1}\|) = \frac{t_m b}{2} c_m (1+\sum_{j=1}^{m-1} c_j)^{-1};$$

3). define

$$f_m := f_{m-1} - c_m\varphi_m.$$

Theorem 3.2. *Let $\tau := \{t_k\}_{k=1}^{\infty}$ be a nonincreasing sequence $1 \ge t_1 \ge t_2 \dots > 0$ and $b \in (0,1)$. Assume X has a modulus of smoothness $\rho(u) \le \gamma u^q$, $q \in (1,2]$. Denote $\mu(u) = \gamma u^q$. Then for any dictionary $\mathcal{D}$ and any $f \in A_1(\mathcal{D})$ the rate of convergence of the MDGA(τ, b, μ) is given by*

$$\|f_m\| \le C(b,\gamma,q)(1 + \sum_{k=1}^{m} t_k^p)^{-\frac{t_m(1-b)}{p(1+t_m(1-b))}}, \quad p := \frac{q}{q-1}.$$

Let us discuss an application of Theorem 1.6 in the case of Hilbert space. It is well known and easy to check that for a Hilbert space H one has

$$\rho(u) \le (1+u^2)^{1/2} - 1 \le u^2/2.$$

Therefore, by Theorem 1.6 with $\mu(u) = u^2/2$ the DGA(t, b, μ) provides the following error estimate

$$\|f_m\| \le C(t,b) m^{-\frac{t(1-b)}{2(1+t(1-b))}} \quad \text{for} \quad f \in A_1(\mathcal{D}). \tag{3.20}$$

The estimate (3.20) with $t = 1$ gives

$$\|f_m\| \le C(b) m^{-\frac{1-b}{2(2-b)}} \quad \text{for} \quad f \in A_1(\mathcal{D}). \tag{3.21}$$

The exponent $\frac{1-b}{2(2-b)}$ in this estimate is approaching $1/4$ with b approaching 0. Comparing (3.21) with the upper estimates (1.6) and (1.7) for the PGA we observe that the DGA$(1, b, u^2/2)$ with small b has better upper estimate for the rate of convergence than the known estimates for the PGA. We note also that (1.8) indicates that the exponent in the power rate of decay of error for the PGA is less than 0.27.

Let us figure out how the DGA$(1, b, u^2/2)$ works in Hilbert space. Consider the mth step of it. Let $\varphi_m \in \mathcal{D}$ be from (1.11). Then it is clear that φ_m maximizes the $\langle f_{m-1}, g\rangle$ over the dictionary $\mathcal{D}$ and

$$\langle f_{m-1}, g\rangle = \|f_{m-1}\| r_{\mathcal{D}}(f_{m-1}).$$

The PGA would use φ_m with the coefficient $\langle f_{m-1}, g\rangle$ at this step. The DGA$(1, b, u^2/2)$ uses the same φ_m and only a fraction of $\langle f_{m-1}, g\rangle$:

$$c_m = b\|f_{m-1}\| r_{\mathcal{D}}(f_{m-1}). \tag{3.22}$$

Thus the choice $b = 1$ in (3.22) corresponds to the PGA. However, it is clear from the above considerations that our technique, designed for general Banach spaces, does not work in the case $b = 1$. The above discussion brings us the following surprising observation. The use of a small fraction ($c_m = b\langle f_{m-1}, g\rangle$) of an optimal coefficient results in improvement of the upper estimate for the rate of convergence.

We present some more results in this direction in the next section.

4. The Weak Greedy Algorithm with parameter b

Motivated by results of Section 3 we consider here a generalization of the WGA obtained by introducing to it a tuning parameter $b \in (0,1]$. Using specifics of Hilbert space structure we will prove more precise estimates that those obtained from the general theory developed in Section 3. Let a sequence $\tau = \{t_k\}_{k=1}^\infty$, $0 \le t_k \le 1$ and a parameter $b \in (0,1]$ be given. We define the Weak Greedy Algorithm with parameter b.

Weak Greedy Algorithm with parameter b (WGA(b)). *We define $f_0^{\tau,b} := f$. Then for each $m \ge 1$, we inductively define:*

1). $\varphi_m^{\tau,b} \in \mathcal{D}$ is any satisfying

$$\langle f_{m-1}^{\tau,b}, \varphi_m^{\tau,b} \rangle \ge t_m \sup_{g\in\mathcal{D}} \langle f_{m-1}^{\tau,b}, g\rangle;$$

2).

$$f_m^{\tau,b} := f_{m-1}^{\tau,b} - b\langle f_{m-1}^{\tau,b}, \varphi_m^{\tau,b}\rangle \varphi_m^{\tau,b};$$

3).

$$G_m^{\tau,b}(f,\mathcal{D}) := b\sum_{j=1}^m \langle f_{j-1}^{\tau,b}, \varphi_j^{\tau,b}\rangle \varphi_j^{\tau,b}.$$

Theorem 4.1. *Let $\tau \notin \mathcal{V}$. Then the WGA(b) with $b \in (0,1]$ converges for each f and all Hilbert spaces H and dictionaries $\mathcal{D}$.*

Proof. In the case $b=1$ the statement is proved in [T2] (see Theorem 1.3 from the Introduction of this paper). In the case $b \in (0,1)$ the proof repeats the proof from [T2] and [T1] with the relation

$$\|f_m^\tau\|^2 = \|f_{m-1}^\tau\|^2 - \langle f_{m-1}^\tau, \varphi_m^\tau\rangle^2$$

replaced by

$$\|f_m^{\tau,b}\|^2 = \|f_{m-1}^{\tau,b}\|^2 - b(2-b)\langle f_{m-1}^{\tau,b}, \varphi_m^{\tau,b}\rangle^2.$$

We will not carry it out here.

We proceed to the rate of convergence.

Theorem 4.2. *Let $\mathcal{D}$ be an arbitrary dictionary in H. Assume $\tau := \{t_k\}_{k=1}^\infty$ is a nonincreasing sequence and $b \in (0,1]$. Then for $f \in A_1(\mathcal{D})$ we have*

$$\|f - G_m^{\tau,b}(f,\mathcal{D})\| \le (1 + b(2-b)\sum_{k=1}^m t_k^2)^{-(2-b)t_m/2(2+(2-b)t_m)}. \tag{4.1}$$

Proof. We introduce some notations:

$$a_m := \|f_m^{\tau,b}\|^2, \quad y_m := \langle f_{m-1}^{\tau,b}, \varphi_m^{\tau,b}\rangle, \quad m = 1,2,\dots,$$

and consider the sequence $\{b_n\}$ defined as follows

$$b_0 := 1, \quad b_m := b_{m-1} + by_m, \quad m = 1, 2, \dots.$$

It is clear that $f_n^{\tau,b} \in A_1(\mathcal{D}, b_n)$. By Lemma 3.5 from [DT] we get

$$\sup_{g\in\mathcal{D}} \langle f_{m-1}^{\tau,b}, g\rangle \geq \|f_{m-1}^{\tau,b}\|^2/b_{m-1}. \tag{4.2}$$

From here and from the equality

$$\|f_m^{\tau,b}\|^2 = \|f_{m-1}^{\tau,b}\|^2 - b(2-b)\langle f_{m-1}^{\tau,b}, \varphi_m^{\tau,b}\rangle^2$$

we obtain the following relations

$$a_m = a_{m-1} - b(2-b)y_m^2, \tag{4.3}$$

$$b_m = b_{m-1} + by_m, \tag{4.4}$$

$$y_m \geq t_m a_{m-1}/b_{m-1}. \tag{4.5}$$

From (4.3) and (4.5) we get

$$a_m \leq a_{m-1}(1 - b(2-b)t_m^2 a_{m-1} b_{m-1}^{-2}).$$

Using that $b_{m-1} \leq b_m$ we derive from here

$$a_m b_m^{-2} \leq a_{m-1} b_{m-1}^{-2}(1 - b(2-b)t_m^2 a_{m-1} b_{m-1}^{-2}).$$

By Lemma 3.1 from [T1] with $A = 1$ we obtain

$$a_m b_m^{-2} \leq (1 + b(2-b)\sum_{k=1}^{m} t_k^2)^{-1}. \tag{4.6}$$

The relations (4.3) and (4.5) imply

$$a_m \leq a_{m-1} - b(2-b)y_m t_m a_{m-1}/b_{m-1} = a_{m-1}(1 - b(2-b)t_m y_m/b_{m-1}). \tag{4.7}$$

Using the inequality $(1-x)^{1/2} \leq 1 - x/2$ we get from (4.7)

$$a_m^{1/2} \leq a_{m-1}^{1/2}(1 - b(1-b/2)t_m y_m/b_{m-1}). \tag{4.8}$$

Rewriting (4.4) in the form

$$b_m = b_{m-1}(1 + by_m/b_{m-1}), \tag{4.9}$$

and using the inequality

$$(1+x)^\alpha \le 1+\alpha x, \quad 0\le \alpha \le 1, \quad x\ge 0,$$

we get from (4.9) that

$$b_m^{(1-b/2)t_m} \le b_{m-1}^{(1-b/2)t_m}(1+b(1-b/2)t_m y_m/b_{m-1}), \tag{4.10}$$

Multiplying (4.8) and (4.10) we obtain

$$a_m^{1/2} b_m^{(1-b/2)t_m} \le a_{m-1}^{1/2} b_{m-1}^{(1-b/2)t_m}.$$

Next, $b_{m-1}\ge 1$ and $t_m \le t_{m-1}$. Therefore

$$b_{m-1}^{(1-b/2)t_m} \le b_{m-1}^{(1-b/2)t_{m-1}}$$

and

$$a_m^{1/2} b_m^{(1-b/2)t_m} \le a_{m-1}^{1/2} b_{m-1}^{(1-b/2)t_{m-1}} \le \dots \le a_0^{1/2} \le 1. \tag{4.11}$$

Combining (4.6) and (4.11) we obtain

$$a_m^{1+(1-b/2)t_m} \le (1+b(2-b)\sum_{k=1}^m t_k^2)^{-(1-b/2)t_m},$$

what completes the proof.

Remark 4.1. *It follows from the proof of Theorem 4.1 that it holds for a modification of the WGA(b) obtained by replacing in its definition* $\sup_{g\in\mathcal{D}}\langle f_{m-1}^{\tau,b}, g\rangle$ *by* $\|f_{m-1}^{\tau,b}\|^2/b_{m-1}$*, where* $b_{m-1} := 1+b\sum_{j=1}^{m-1}\langle f_{j-1}^{\tau,b}, \varphi_j^{\tau,b}\rangle$.

References

[B] N.K. Bary, *Trigonometric series*, Nauka, Moscow, 1961; English transl. in Pergamon Press, Oxford, 1964.

[D] R.A. DeVore, *Nonlinear Approximation*, Acta Numerica (1998), 51–150.

[DT] R.A. DeVore and V.N. Temlyakov, *Some remarks on Greedy Algorithms*, Advances in comp. Math. **5** (1996), 173–187.

[DGDS] M. Donahue, L. Gurvits, C. Darken, E. Sontag, *Rate of convex approximation in non-Hilbert spaces*, Constr. Approx. **13** (1997), 187–220.

[GK] M. Ganichev and N.J. Kalton, *Convergence of the Weak Dual Greedy Algorithm in* L_p*-spaces*, manuscript (2002), 1–5.

[HHZ] P. Habala, P. Hájek, and V. Zizler, *Introduction to Banach spaces [I]*, Matfyzpress, Univerzity Karlovy, 1996.

[J] L. Jones, *On a conjecture of Huber concerning the convergence of projection pursuit regression*, Annals of Stat. **15** (1987), 880–882.

[KT] S.V. Konyagin and V.N. Temlyakov, *Rate of convergence of Pure Greedy Algorithm*, East J. on Approx. **5** (1999), 493–499.

[L] E.D. Livshitz, *On convergence of greedy algorithms in Banach spaces*, Matem. Zametki **73** (2003), 371-389.

[LT] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces II*, Springer-Verlag, Berlin, 1979.

[LiT] E.D. Livshitz and V.N. Temlyakov, *Two lower estimates in greedy approximation*, Constr. Approximation **19** (2003), ??.

[T1] V.N. Temlyakov, *Weak greedy algorithms*, Advances in Comput. Math. **12** (2000), 213–227.

[T2] V.N. Temlyakov, *A criterion for convergence of Weak Greedy Algorithms*, Advances in Comput. Math. **17** (2002), 269–280.

[T3] V.N. Temlyakov, *Greedy algorithms in Banach spaces*, Advances in Comput. Math. **14** (2001), 277–292.

[T4] V.N. Temlyakov, *Nonlinear Methods of Approximation*, Found. Comput. Math. **3** (2003), 33–107.