# INDUSTRIAL MATHEMATICS INSTITUTE

## 2001:14

Convergence of some greedy algorithms in Banach spaces

S.J. Dilworth, D. Kutzarova and V.N. Temlyakov

# IMI

Preprint Series

# CONVERGENCE OF SOME GREEDY ALGORITHMS IN BANACH SPACES

S. J. DILWORTH, DENKA KUTZAROVA AND V. N. TEMLYAKOV

ABSTRACT. We consider some theoretical greedy algorithms for approximation in Banach spaces with respect to a general dictionary. We prove convergence of the algorithms for Banach spaces which satisfy certain smoothness assumptions. We compare the algorithms and their rates of convergence when the Banach space is $L_p(\mathrm{T}^d)$ $(1 < p < \infty)$ and the dictionary is the trigonometric system.

## CONTENTS

## 1. INTRODUCTION

We continue the investigation of nonlinear $m$-term approximation. In this paper we concentrate on studying convergence and rate of approximation of some theoretical greedy algorithms in Banach spaces.

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 ($\|g\| = 1$),

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

and $X$ is the closed linear span of $\mathcal{D}$. We study in this paper a particular case of the following *Weak Chebyshev Greedy Algorithm* with respect to $\mathcal{D}$ [17]. For an element $f \in X$ we denote by $F_f$ a norming (or 'peak') functional for $f$:

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

The existence of such a functional is guaranteed by the Hahn-Banach theorem. Let $\tau := \{t_k\}_{k=1}^{\infty}$ be a given sequence of positive numbers with $t_k \leq 1$ ($k \in \mathbb{N}$). We define the Weak Chebyshev Greedy Algorithm (WCGA), which is a generalization for Banach spaces of the Weak Orthogonal Greedy Algorithm for Hilbert spaces defined and studied in [16] (see also [4] for the Orthogonal Greedy Algorithm), as follows.

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

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

$$F_{f_{m-1}^c}(\varphi_m^c) \geq t_m \sup_{g \in \mathcal{D}} F_{f_{m-1}^c}(g).$$

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 $\Phi_m$.

3). Denote

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

We will consider the case $\tau = \{t\}$ with $0 < t \leq 1$ here. To stress this we will write $t$ instead of $\tau$ in the notation.

In Section 2 we discuss the question of convergence of the WCGA in Banach spaces which satisfy certain smoothness hypotheses. In order not to disrupt the flow of this Introduction unduly, we defer until Section 2 the precise definitions and comparisons of the various notions of smoothness which will be used.

It is known (see [17]) that the WCGA with $\tau = \{t\}$ converges for all elements $f$ in a *uniformly smooth* Banach space. In uniformly smooth Banach spaces we can even allow $t_n \to 0$ as $n \to \infty$ without losing the

convergence property [17]. It is also known [1, 5] that the assumption of *Gâteaux* smoothness is necessary to ensure that

$$\inf_{c \in \mathbf{R}, g \in \mathcal{D}} \|f - cg\| < \|f\|$$

for any $f \in X$ and any dictionary $\mathcal{D}$. This means that Gâteaux smoothness is necessary for convergence of the WCGA for an arbitrary dictionary. However, we do not know if it is a sufficient condition for convergence in the case $\tau = \{t\}$.

In Section 2 we show that the WCGA converges in a large class of *reflexive* Banach spaces satisfying some mild additional smoothness conditions. In particular, we prove the following general convergence result.

**Theorem 2.8.** *Let $X$ be a separable reflexive Banach space. Then $X$ admits an equivalent norm for which the WCGA converges for every dictionary $\mathcal{D}$ and for every $f \in X$.*

In Section 3 we discuss convergence of the WCGA in the Hardy space $H_1(\mathbf{T})$. This provides an interesting test case because $H_1(\mathbf{T})$ is nonreflexive and does not satisfy the smoothness assumptions made in Section 2. Nevertheless, we are able to prove convergence of the WCGA for a large class of dictionaries, including the trigonometric system.

In Section 4 we present some results on convergence of the *Weak Dual Greedy Algorithm* (WDGA), which is the analogue for Banach spaces of the Weak Greedy Algorithm in Hilbert space. Here is the definition of the algorithm for the special case $\tau = \{t\}$ discussed in Section 4.

**Weak Dual Greedy Algorithm (WDGA).** We define $f_0^D := f_0^{D,t} := f$. Then for each $m \geq 1$ we inductively define

1). $\varphi_m^D := \varphi_m^{D,t} \in \mathcal{D}$ is any element of $\mathcal{D}$ satisfying

$$F_{f_{m-1}^D}(\varphi_m^D) \geq t \sup_{g \in \mathcal{D}} F_{f_{m-1}^D}(g).$$

2). Define $a_m$ as follows:

$$\|f_{m-1}^D - a_m \phi_m^D\| = \min_{a \in \mathbf{R}} \|f_{m-1}^D - a\phi_m^D\|.$$

3). Denote

$$f_m^D := f_m^{D,t} := f_{m-1}^D - a_m \phi_m^D.$$

Convergence of the WDGA is not very well understood, and we can prove only a few scattered results in this direction.

In Section 5 we discuss greedy type approximation with regard to a basis. Let $X$ be a Banach space with basis $\Psi = (\psi_k)_{k=1}^{\infty}$, $\|\psi_k\| = 1$. We compare the efficiency of the WCGA with the following theoretical greedy algorithm which we call the *Thresholding Greedy Algorithm* (TGA) (see [11]).

**Thresholding Greedy Algorithm (TGA).** For a given $f \in X$, we consider the expansion

$$f = \sum_{k=1}^{\infty} c_k(f)\psi_k.$$

Suppose that $f \in X$ and that $\rho$, $\rho(j) = k_j$ $(j = 1, 2, \dots)$, is a permutation of the positive integers. We say that $\rho$ is decreasing and write $\rho \in D(f)$ if

$$|c_{k_1}(f)| \geq |c_{k_2}(f)| \geq \dots \quad .$$

In the case of strict inequalities here, $D(f)$ consists of only one permutation. We define the $m$-th greedy approximant of $f$ with regard to the basis $\Psi$ corresponding to a permutation $\rho \in D(f)$ by the formula:

$$G_m(f, \Psi) := G_m^X(f, \Psi) := G_m^X(f, \Psi, \rho) := \sum_{j=1}^{m} c_{k_j}(f)\psi_{k_j}.$$

The general conclusion is that the WCGA is more efficient than the TGA for approximation in $L_p$ for $2 < p < \infty$. We also point out in Section 5 that the WCGA provides a constructive (algorithmic) way to get best possible $m$-term trigonometric approximation for different smoothness classes of functions [3].

We use standard Banach space notation and terminology (see e.g. [12]) and basic results from functional analysis (see e.g. [6]). For clarity, however, we recall here the notation and results which will be used most heavily. Let $X$ be a Banach space. The *unit sphere* of $X$, denoted $S(X)$, is the set $\{x \in X : \|x\| = 1\}$. The *unit ball* of $X$, denoted $Ba(X)$, is the set $\{x \in X : \|x\| \leq 1\}$. The *dual space* of $X$, denoted $X^*$, is the Banach space of all continuous linear functionals $F$ equipped with the norm:

$$\|F\| = \sup\{F(x) : x \in S(X)\}.$$

The closed linear span of a set $A \subseteq X$ (resp., a sequence $(x_n)$) is denoted $[A]$ (resp. $[x_n]$). A normalized *Schauder basis* for $X$ is a sequence of unit vectors $\Psi = (\psi_k)_{k=1}^{\infty}$ such that every $f \in X$ has a *unique* expansion as a norm-convergent series

$$f = \sum_{k=1}^{\infty} c_k(f) \psi_k.$$

We say that a sequence $(x_n)$ in $X$ converges *strongly* (resp. *weakly*) to $x \in X$ if $\|x - x_n\| \to 0$ (resp. for all $F \in X^*$, $F(x - x_n) \to 0$) as $n \to \infty$. We say that a sequence $(F_n)$ in $X^*$ converges *weak-star* to $F$ if $F_n(x) \to F(x)$ as $n \to \infty$ for all $x \in X$. A Banach space is *reflexive* if $Ba(X)$ is weakly compact (equivalently, if every bounded sequence has a weakly convergent subsequence). It follows from the Banach-Alaoglu theorem (see e.g. [6, p. 424]) that if $X$ is separable, then $Ba(X^*)$ is weak-star sequentially compact (i.e. every bounded sequence in $X^*$ has a weak-star convergent subsequence).

Finally, more specialized notions from Banach space theory, including the various notions of smoothness used throughout, will be introduced as needed.

## 2. Convergence results for the WCGA

In this section we discuss convergence of the WCGA. It is known that the WCGA converges (for an arbitrary dictionary) in all uniformly smooth Banach spaces [17]. The Banach spaces which admit a uniformly smooth renorming (equivalent to the original norm) have been studied extensively: they are the so-called *super-reflexive* spaces; they coincide, in particular, with the spaces which admit an equivalent *uniformly convex* renorming [7].

The class of super-reflexive spaces is properly contained, as the term suggests, in the class of *reflexive* spaces: for example, the projective tensor product $\ell_p \hat{\otimes} \ell_q$ ($1/p + 1/q < 1$) is a separable reflexive space that is not super-reflexive. We prove below that the WCGA converges in all reflexive spaces which satisfy some mild additional smoothness hypotheses. In particular, we show that every separable reflexive space admits an equivalent norm for which the WCGA converges.

We start by recalling several notions of smoothness for Banach spaces (see e.g. [2, p. 2]). For $x, y \in X$, we define

$$\rho_{x,y}(u) = \frac{\|x + uy\| + \|x - uy\| - 2\|x\|}{2} \qquad (u \in \mathrm{R}).$$

We say that $x$ is a *point of Gâteaux smoothness* if

$$(1) \qquad \lim_{u \downarrow 0} \frac{\rho_{x,y}(u)}{u} = 0 \qquad (y \in X).$$

It is known that (1) is equivalent to the *uniqueness* of the norming functional $F_x$. If every $x \neq 0$ is a point of Gâteaux smoothness then $X$ is said to be *smooth*. The *local modulus of smoothness* at $x$ is defined as follows:

$$(2) \qquad \rho_x(u) = \sup\{\rho_{x,y}(u) \colon y \in S(X)\} \qquad (u \in \mathrm{R}).$$

We say that $x$ is a *point of Fréchet differentiability* if

$$(3) \qquad \lim_{u \downarrow 0} \frac{\rho_x(u)}{u} = 0.$$

It is known that (3) is equivalent to

$$\|x + y\| = \|x\| + F_x(y) + o(\|y\|) \qquad (y \in X).$$

If every $x \neq 0$ is a point of Fréchet differentiability then $X$ is said to be Fréchet differentiable. Finally, we define the *modulus of smoothness* of $X$:

$$(4) \qquad \rho_X(u) = \sup\{\rho_x(u) \colon x \in S(X)\} \qquad (u \in \mathrm{R}).$$

$X$ is said to be *uniformly smooth* if

$$(5) \qquad \lim_{u \downarrow 0} \frac{\rho_X(u)}{u} = 0.$$

It is known that (5) is equivalent to the following:

$$(6) \qquad \|x + y\| = \|x\| + F_x(y) + \|y\|\varepsilon(x, y),$$

where $\varepsilon(x, y) \to 0$ as $y \to 0$ uniformly for $x \in S(X)$.

Suppose that $0 \neq f \in X$ and that $(G_n^{c,t})$, $(\phi_n^{c,t})$, and $F_{f_n^{c,t}}$, are as defined in the Introduction. To simplify notation, we shall drop the superscript $t$, so that $G_n^c := G_n^{c,t}$, etc. It is clear that

$$\|f_n^c\| \downarrow \alpha \geq 0$$

as $n \to \infty$. The WCGA converges if and only if $\alpha = 0$. Note also that

$$\tag{7} \alpha = \operatorname{dist}(f, [\phi_n^c]_{n=1}^\infty).$$

**Lemma 2.1.** *Suppose that $X$ is smooth and that $\alpha > 0$. Then there exists $\beta > 0$ such that $F_{f_n^c}(\phi_{n+1}^c) > \beta$ for all sufficiently large $n$.*

*Proof.* Since $G_n^c$ is the best approximant to $f$ from $[\phi_i^c]_{i=1}^n$, it follows that there exists $F_n \in S(X^*)$ such that $F_n(\phi_i^c) = 0$ for $1 \le i \le n$ and $F_n(f_n^c) = \|f_n^c\|$. Clearly $F_n$ is a norming functional for $f_n^c$, and since $X$ is smooth it follows that $F_n = F_{f_n^c}$. Thus,

$$F_{f_n^c}(\phi_i^c) = F_n(\phi_i^c) = 0 \qquad (1 \le i \le n).$$

Let $\tilde{\mathcal{D}} = \mathcal{D} \setminus \{\pm\phi_n^c : n \ge 1\}$. Since $f \in [\mathcal{D}]$ there exist $N \ge 1$, $m \ge 1$, $f' \in [\phi_i^c]_{i=1}^N$, and

$$\tag{8} \tilde{f} = \sum_{i=1}^m a_i g_i$$

(for some $g_i \in \tilde{\mathcal{D}}$ and scalars $a_i \ne 0$ $(1 \le i \le m)$) such that

$$\|f - f' - \tilde{f}\| < \alpha/2.$$

Suppose that $n \ge N$. Then $F_{f_n^c}(\phi_i^c) = 0$ for $1 \le i \le N$, and so $F_{f_n^c}(f') = 0$. Thus,

$$F_{f_n^c}(\tilde{f}) = F_{f_n^c}(f' + \tilde{f}) \ge F_{f_n^c}(f) - \|f - f' - \tilde{f}\|$$
$$\ge \alpha - \alpha/2 = \alpha/2.$$

From (8) and the above, we obtain

$$\sum_{i=1}^m |a_i| |F_{f_n^c}(g_i)| \ge \alpha/2,$$

and hence

$$\max_{1 \le i \le m} |F_{f_n^c}(g_i)| \ge \frac{\alpha}{2m \max_{1 \le i \le m} |a_i|}.$$

Finally,

$$F_{f_n^c}(\phi_{n+1}^c) \ge t \sup_{g \in \mathcal{D}} F_{f_n^c}(g) \ge t \max_{1 \le i \le m} F_{f_n^c}(g_i).$$

So we can take $\beta = t\alpha/(2m \max_{1 \le i \le m} |a_i|)$.

$\square$

**Lemma 2.2.** *Suppose that $X$ is smooth and that $y \neq f$. If some subsequence $(G_{n_k}^c)$ converges strongly to $y$, then $y - f$ is not a point of Fréchet differentiability.*

*Proof.* To derive a contadiction, we shall assume that $y - f$ is a point of Fréchet differentiability. By Smulyan's Lemma [2, Thm. 1.4] the duality mapping $x \mapsto F_x$ from $X$ to $X^*$ is norm-to-norm continuous at $y - f$. Thus, $F_{G_{n_k}^c - f} \to F_{y-f}$ strongly as $k \to \infty$. For each $n \geq 1$, we have

$$F_{y-f}(\phi_n^c) = \lim_{k \to \infty} F_{G_{n_k}^c - f}(\phi_n^c) = 0.$$

Thus,

$$F_{G_{n_k}^c - f}(\phi_{n_k+1}^c) = (F_{G_{n_k}^c - f} - F_{y-f})(\phi_{n_k+1}^c) \leq \|F_{G_{n_k}^c - f} - F_{y-f}\| \to 0$$

as $k \to \infty$. But by Lemma 2.1, $F_{G_{n_k}^c - f}(\phi_{n_k+1}^c) > \beta > 0$, which is the desired contradiction.   □

The following result is an immediate corollary. Following the method of [17], a direct proof can also be given using the local modulus of smoothness $\rho_x(u)$ defined in (2).

**Proposition 2.3.** *Suppose that $X$ is Fréchet differentiable. Then $(G_n^c)$ converges strongly to $f$ if and only if it has a strongly convergent subsequence.*

**Lemma 2.4.** *Suppose that $(G_{n_k}^c)$ converges weakly to $y$ for some subsequence $(G_{n_k}^c)$. Then $\|y - f\| = \lim_{k \to \infty} \|G_{n_k}^c - f\| = \alpha$.*

*Proof.* Note that $y \in [\phi_n^c]_{n=1}^\infty$, since $[\phi_n^c]_{n=1}^\infty$ is weakly closed. Hence (7) gives

$$\|y - f\| \geq \text{dist}(f, [\phi_n^c]_{n=1}^\infty) = \alpha.$$

On other hand, since $(G_{n_k}^c)$ converges weakly to $f$, we have

$$\|y - f\| = F_{y-f}(y - f) = \lim_{k \to \infty} F_{y-f}(G_{n_k}^c - f) \leq \lim_{k \to \infty} \|G_{n_k}^c - f\|.$$

   □

Let us now recall the definition of the *Kadec-Klee property* in Banach spaces. (This property is also called the *Radon-Riesz property* or *Property H*.) A point $x \in X$ is said to be a Kadec-Klee point if the following holds: whenever $(x_n)$ converges weakly to $x$ and $\|x\| = \lim_{n \to \infty} \|x_n\|$,

then $(x_n)$ converges strongly to $x$. If every $x \in X$ is a Kadec-Klee point then $X$ is said to have the Kadec-Klee property. It is known that every separable Banach space admits an equivalent norm with the Kadec-Klee property (see e.g. [2, Th. 2.6]).

**Lemma 2.5.** *Suppose that some subsequence $(G_{n_k}^c)$ converges weakly to $y$, where $y$ is a Kadec-Klee point. Then $(G_{n_k}^c)$ converges strongly to $y$.*

*Proof.* This is immediate from Lemma 2.4 $\qquad\square$

The following result combines Proposition 2.3 and Lemma 2.5.

**Proposition 2.6.** *Suppose that $X$ has the Kadec-Klee property and a Fréchet differentiable norm. Then $(G_n^c)$ converges strongly to $f$ if and only if it has a weakly convergent subsequence.*

**Theorem 2.7.** *Suppose that $X$ is a reflexive Banach space which has both the Kadec-Klee property and a Fréchet differentiable norm. Then the WCGA converges for every dictionary $\mathcal{D}$ and for every $f \in X$.*

*Proof.* Since $X$ is reflexive, $(G_n^c)$ has a weakly convergent subsequence. So the result follows from Proposition 2.6. $\qquad\square$

**Theorem 2.8.** *Let $X$ be a separable reflexive Banach space and let $\varepsilon > 0$. There exists an equivalent norm $|\cdot|$ on $X$, with*

$$\|x\| \le |x| \le (1+\varepsilon)\|x\| \qquad (x \in X),$$

*for which the WCGA converges in $(X, |\cdot|)$ for every dictionary $\mathcal{D}$ and every $f \in X$.*

*Proof.* Every separable reflexive Banach space admits such a norm $|\cdot|$ which is both Fréchet differentiable and has the Kadec-Klee property (see e.g. [2, Thm. 2.6]). $\qquad\square$

Finally, we prove a result for non-reflexive Banach spaces. Suppose that $X$ is a smooth Banach space and that $y \in S(X)$. Recall that $X$ is said to be *uniformly Gâteaux differentiable in the direction $y$* if

$$(9) \qquad \lim_{u \downarrow 0} \frac{\rho_{x,y}(u)}{u} = 0$$

uniformly for $x \in S(X)$; equivalently,

$$\|x + uy\| = \|x\| + u(F_x(y) + o(u)) \qquad (u \in \mathbb{R})$$

*uniformly* for $x \in S(X)$.

We shall say that a dictionary $\mathcal{D}$ is *minimal* if $\mathcal{D}$ has no proper subdictionary. Recall that every separable Banach space contains a system $(w_n) \subset S(X)$ satisfying the following: $\{w_n : n \geq 1\}$ has dense span and no proper subset has dense span. We shall say that $(w_n)$ is *quasi-minimal*. (In fact, $(w_n)$ can be chosen to be both *minimal* and *total* [12, p. 43], although we don't need this). Clearly, $(w_n)_{n=1}^{\infty} \subset S(X)$ is quasi-minimal if and only if $\mathcal{D} = \{\pm w_n : n \geq 1\}$ is a minimal dictionary.

**Theorem 2.9.** *Suppose that $X$ is a separable smooth Banach space and that $y \in S(X)$. If $X$ is uniformly Gâteaux differentiable in the direction $y$ then there exists a minimal dictionary $\mathcal{D}$ such that the WCGA converges for every $f \in X$.*

*Proof.* Let $(w_n) \subset S(X)$ be a quasi-minimal system for $X$ with $w_1 = y$. Let $y_n = (1/n)w_{n+1} + y$ $(n \geq 1)$, so that $(y_n)$ converges strongly to $y$. Then $(y_n)$ is also quasi-minimal, and hence $\mathcal{D} = \{\pm y_n/\|y_n\| : n \geq 1\}$ is a minimal dictionary.

We show that the WCGA converges for $\mathcal{D}$. So suppose that $f \in X$, and, to derive a contradiction, that $\|f_n^c\| \downarrow \alpha > 0$. There exists a choice of signs $\epsilon_n = \pm 1$ such that $(\epsilon_n \phi_n^c)$ converges strongly to $y$. By Lemma 2.1, there exists $\beta > 0$ such that $F_{f_n^c}(\phi_{n+1}^c) > \beta$ for sufficiently large $n$. Hence

$$(10) \qquad F_{f_n^c}(\epsilon_n y) \geq \beta \quad \text{for all sufficiently large } n.$$

Since $X$ is uniformly Gâteaux differentiable in the direction $y$, it follows from (10) and (9) that there exists $s_0 > 0$ such that

$$\|f_n^c\| - \|f_n^c - \epsilon_n s_0 y\| \geq \frac{\beta s_0}{2}$$

for sufficiently large $n$. Since $(\epsilon_n \phi_n^c)$ converges strongly to $y$, we obtain

$$\|f_n^c\| - \|f_n^c - \epsilon s_0 \phi_{n+1}^c\| \geq \frac{\beta s_0}{3}$$

for all sufficiently large $n$, which implies that

$$\|f_{n+1}^c\| \leq \|f_n^c\| - \frac{\beta s_0}{3}.$$

But this contradicts our assumption that $\|f_n^c\| \downarrow \alpha$. $\qquad\square$

**Theorem 2.10.** *Let $X$ be a separable Banach space and let $\varepsilon > 0$. There exists an equivalent norm $|\cdot|$ on $X$, with*

$$\|x\| \leq |x| \leq (1 + \varepsilon)\|x\| \qquad (x \in X),$$

*such that $(X, |\cdot|)$ has a minimal dictionary $\mathcal{D}$ for which the WCGA converges for every $f \in X$.*

*Proof.* Every separable Banach space admits such an equivalent norm $|\cdot|$ which is uniformly Gâteaux differentiable in every direction [2, Cor. 6.9]. $\qquad\qquad\square$

## 3. CONVERGENCE OF THE WCGA IN $H_1(\mathrm{T})$

In this section we shall discuss convergence of the WCGA in the Hardy space $H_1(\mathrm{T})$ defined below. This space is nonreflexive and is *nowhere* Fréchet differentiable [2, Prop. 4.5], and so Theorem 2.7 does not apply. Nevertheless, we are able to prove that the WCGA converges for a large class of dictionaries, including the trigonometric system.

The WCGA was defined in the Introduction for *real* Banach spaces. But since $H_1(\mathrm{T})$ is a *complex* Banach space, we must first explain how the algorithm is to be applied in complex spaces. So suppose that $X$ is a complex Banach space with complex dual space $X^*$. Let $X^{\mathrm{R}}$ be the *real* Banach space obtained from $X$ by restricting scalar multiplication to real scalars. To apply the WCGA in $X$, we simply apply the WCGA in the real space $X^{\mathrm{R}}$. It follows from the complex version of the Hahn-Banach theorem (due to Bohnenblust and Sobczyk [6, p. 63]) that the mapping $x^* \mapsto \Re\, x^*$ defines a real linear isometry from $(X^*)^{\mathrm{R}}$ onto $(X^{\mathrm{R}})^*$. In particular, the norming functionals in $(X^{\mathrm{R}})^*$ are precisely the real parts of norming functionals in $X^*$.

Now let us recall the definition of $H_1(\mathrm{T})$. Let $\mathrm{T}$ denote the unit circle in the complex plane and let $m$ denote the normalized Lebesgue measure on $\mathrm{T}$. Let $C(\mathrm{T})$ denote the Banach space of all complex-valued continuous functions equipped with the supremum norm. The disc algebra $A(\mathrm{T})$ is the complex closed linear span of $\{z^n : n \geq 0\}$ in $C(\mathrm{T})$. Let $A_0$ denote the subspace of $A(\mathrm{T})$ consisting of the mean zero functions. The Hardy space $H_1(\mathrm{T})$ is the (complex) closed linear span of $\{z^n : n \geq 0\}$ in the complex Lebesgue space $L_1(\mathrm{T})$.

First we present an example which shows that the WCGA does not converge in general in $L_1(\mathrm{T})$ for the trigonometric dictionary $\mathcal{D} =$

$\{\pm z^n, \pm i z^n : n \geq 0\}$. (A similar example works for the usual trigono-
metric system $\{\cos n\theta\}_{n \geq 0} \cup (\sin n\theta)_{n \geq 1}$.)

**Example 3.1.** Let $f(e^{i\theta}) = 4\chi_{[3\pi/4, 5\pi/4]}(\theta)$ $(0 \leq \theta < 2\pi)$. Then, for
every $g \in \mathcal{D}$, we have

$$\|f + sg\|_1 \geq 1 + \frac{|s|}{2} \qquad (s \in \mathrm{R}).$$

Thus, $s = 0$ gives the unique best approximant to $f$ from $\mathcal{D}$. Hence
$f_1^{c,t} = f$, which of course implies that $f_n^{c,t} = f$ for all $n \geq 1$.

Before we can proceed, it is necessary to recall some facts about
$H_1(\mathrm{T})$ as a Banach space.

*Fact 3.2.* (see e.g. [9, p. 137]) $H_1(\mathrm{T})$ is the dual space of $C(\mathrm{T})/A_0(\mathrm{T})$,
with the duality given by

$$\langle f, g \rangle = \int_{\mathrm{T}} f g \, dm \qquad (f \in C(\mathrm{T}), g \in H_1(\mathrm{T})).$$

Weak-star convergence in Fact 3.4 and in the proof of Theorem 3.5
below refers to the weak-star topology on $H_1(\mathrm{T})$ induced by its predual
$C(\mathrm{T})/A_0$.

*Fact 3.3.* $H_1(\mathrm{T})$ is a smooth Banach space. For $0 \neq f \in H_1(\mathrm{T})$, the
unique (real) norming functional $F_f$ is given by

$$F_f(g) = \Re \int_{\mathrm{T}} \overline{\mathrm{sgn}\, f} g \, dm.$$

Here $\mathrm{sgn}\, 0 := 0$ and $\mathrm{sgn}\, z := z/|z|$ for $z \neq 0$.

*Fact 3.4.* [14] $H_1(\mathrm{T})$ has the *weak-star Kadec-Klee* property, i.e., if $(f_n)$
converges weak-star to $f$ and $\|f\| = \lim_{n \to \infty} \|f_n\|$, then $(f_n)$ converges
strongly to $f$.

Recall that a subset $K \subset L_1(m)$ is *uniformly integrable* if for every
$\varepsilon > 0$ there exists $\delta > 0$ such that if $E \subseteq \mathrm{T}$ satisfies $m(E) < \delta$ then

$$\int_E |f| \, dm \leq \varepsilon \qquad (f \in K).$$

**Theorem 3.5.** *Suppose that the dictionary $\mathcal{D} \subset H_1(\mathrm{T})$ satisfies the
following two conditions:*

 (a) *$\mathcal{D}$ is uniformly integrable;*

(b) $[E]$ *is weak-star closed whenever $E \subseteq \mathcal{D}$.*

*Then the WCGA converges for every $f \in H_1(\mathrm{T})$.*

*Proof.* With the notation introduced above, suppose, to derive a contradiction, that $\|f_n^c\| \downarrow \alpha > 0$. By Lemma 2.1 and Fact 3.3 there exists $\beta > 0$ such that

$$\text{(11)} \qquad F_{f_n^c}(\phi_{n+1}^c) \geq \beta \qquad (n \geq 1).$$

There is a subsequence $(G_{n_k}^c)_{k=1}^{\infty}$ which converges weak-star to some $y \in H_1(\mathrm{T})$. Let $E = \{\phi_n^c : n \geq 1\}$. Then, by assumption (b), $[E]$ is weak-star closed in $H_1(\mathrm{T})$, and so $y \in [E]$. Thus $\|y - f\| \geq \text{dist}(f, [E]) = \alpha$. On the other hand, by weak-star lower semi-continuity of $\|\cdot\|$, we have $\|y - f\| \leq \lim_{k \to \infty} \|G_{n_k}^c - f\| = \alpha$. So

$$\|y - f\| = \lim_{k \to \infty} \|G_{n_k}^c - f\| = \alpha.$$

Since $H_1(\mathrm{T})$ has the weak-star Kadec-Klee property (Fact 3.4), it follows that $(G_{n_k}^c)$ converges strongly to $y$. Since $H_1(\mathrm{T})$ is smooth (Fact 3.3), the mapping $x \mapsto F_x$ is norm-to-weak-star continuous. So

$$\text{(12)} \qquad F_{f-y}(\phi_n^c) = \lim_{k \to \infty} F_{f_{n_k}^c}(\phi_n^c) = 0 \qquad (n \geq 1).$$

Recall that $m(\{z : f(z) = 0\}) = 0$ for all $f \in H_1(\mathrm{T})$ (see e.g. [9, p. 52]). So, by passing to a subsequence if necessary, we may assume that

$$\text{sgn}(G_{n_k}^c - f) \to \text{sgn}(y - f) \qquad a.e.$$

Hence the uniform integrability of $\mathcal{D}$ (assumption (b)) gives

$$\text{(13)} \qquad \int_{\mathrm{T}} (\overline{\text{sgn}(G_{n_k}^c - f)} - \overline{\text{sgn}(y - f)}) g(z) \, dm \to 0$$

as $k \to \infty$ *uniformly* for all $g \in \mathcal{D}$. Now (11), (12), (13) and Fact 3.3 yield

$$\begin{aligned}
0 &= \lim_{k \to \infty} F_{f-y}(\phi_{n_k+1}^c) \\
&= \lim_{k \to \infty} \Re \int_{\mathrm{T}} \overline{\text{sgn}(y - f)} \phi_{n_k+1}^c(z) \, dm \\
&= \lim_{k \to \infty} \Re \int_{\mathrm{T}} \overline{\text{sgn}(G_{n_k}^c - f)} \phi_{n_k+1}^c(z) \, dm \\
&= \lim_{k \to \infty} F_{f_{n_k}^c}(\phi_{n_k+1}^c) \geq \beta > 0,
\end{aligned}$$

which is the desired contradiction. $\square$

We now show that the trigonometric system $\mathcal{D} = \{\pm z^n, \pm iz^n : n \geq 0\}$ satisfies the assumptions of Theorem 3.5. Uniform integrability is clear. The second condition is verified in the next lemma.

**Lemma 3.6.** *Suppose that $A, B \subseteq \mathrm{N}$. Let $X$ denote the real closed linear span of $\{z^n : n \in A\} \cup \{iz^n : n \in B\}$ in $H_1(\mathrm{T})$. Then $X$ is a weak-star closed subspace of $H_1(\mathrm{T})$.*

*Proof.* For $n \in \mathrm{Z}$, let $a_n(f)$ denote the $n$-th Fourier coefficient of $f$. It is clear that the weak-star closure of $X$ consists precisely of those $f \in H_1(\mathrm{T})$ such that $\Re\, a_n(f) = 0$ for all $n \in \mathrm{N} \setminus A$ and $\Im\, a_n(f) = 0$ for all $n \in \mathrm{N} \setminus B$. So the Fejer polynomials

$$p_n = \sum_{k=0}^{n} \left(1 - \frac{k}{n}\right) a_k(f) z^k \qquad (n \geq 0)$$

all belong to $X$ whenever $f$ lies in the weak-star closure of $X$. But $(p_n)$ converges strongly to $f$. Thus, $f \in X$. $\qquad\square$

**Theorem 3.7.** *In $H_1(\mathrm{T})$ the WCGA converges for the trigonometric system $\mathcal{D} = \{\pm z^n, \pm iz^n : n \geq 0\}$ and for every $f \in H_1(\mathrm{T})$.*

*Remark* 3.8. Hoffmann [10] proved that the multivariable Hardy space $H_1(\mathrm{T}^d)$ has the weak-star Kadec-Klee property with respect to its natural predual. Using this fact it is a straightforward mattter to extend Theorem 3.5 to the multivariate case. In particular, Theorem 3.7 is valid for $H_1(\mathrm{T}^d)$ and the multivariate trigonometric system.

## 4. Convergence results for the WDGA

We have only a few isolated results for convergence of the WDGA.

**Proposition 4.1.** *Let $X$ be a uniformly smooth Banach space. Then, for every increasing sequence $(n_k) \subseteq \mathrm{N}$, either $(G_{n_k}^D)$ converges weakly in $X$ to $f$ or $(F_{f_{n_k}^D})$ has a weakly null subsequence in $X^*$.*

*Proof.* Suppose that $(G_{n_k}^D)$ does not converge weakly to $f$. Since $X$ is uniformly smooth, and hence reflexive, by passing to a subsequence and relabelling we may assume that $(G_{n_k}^D)$ converges weakly to $y \neq f$ and that $(F_{f_{n_k}^D})$ converges weakly to $F$, say. Suppose, to derive a

contradiction, that $F \neq 0$. Then there exists $g \in \mathcal{D}$ such that $F(g) = \beta > 0$. By weak convergence $F_{f_{n_k}^D}(g) > \beta/2$, and hence

$$F_{f_{n_k}^D}(\phi_{n_k+1}^D) > \frac{t\beta}{2}$$

for sufficiently large $k$. Since $X$ is uniformly smooth, it follows from (6) that there exists $s_0 > 0$ such that

$$\|f_{n_k}^D - s_0\phi_{n_k+1}^D\| \leq \|f_{n_k}^D\| - \frac{s_0 F_{f_{n_k}^D}(\phi_{n_k+1}^D)}{2}$$

for all $k$. Thus, for sufficiently large $k$, we have

$$\|f_{n_k+1}^D\| \leq \|f_{n_k}^D\| - \frac{t\beta s_0}{2},$$

which contradicts the fact that $\|f_n^D\| \downarrow \alpha > 0$. Thus $F = 0$. $\qquad\square$

**Theorem 4.2.** *Suppose that $X$ is uniformly smooth and satisfies the following: $(x_n) \subset S(X)$ is weakly null whenever $(F_{x_n})$ is weakly null. Then $(G_n^D)$ converges weakly to $f$.*

*Proof.* From Proposition 4.1, the hypothesis implies that every subsequence of $(G_n^D)$ has a further subsequence which converges weakly to $f$. But this implies that $(G_n^D)$ converges weakly to $f$. $\qquad\square$

Since $\ell_p$ satisfies the hypotheses of Theorem 4.2, we get the following corollary.

**Corollary 4.3.** *The WDGA converges weakly in $\ell_p$ $(1 < p < \infty)$ for every dictionary $\mathcal{D}$ and for every $f \in \ell_p$.*

We have only one result for strong convergence of the WDGA. Say that a Schauder basis $(\psi_n)$ is *strictly supression 1-unconditional* if, for every finite $A \subset \mathbb{N}$, for every $n \in \mathbb{N} \setminus A$, and for all scalars $(a_i)_{i \in A}$, we have

$$\|\sum_{i \in A} a_i\psi_i\| < \|\sum_{i \in A} a_i\psi_i + e_n\|.$$

The standard basis of an Orlicz or Lorentz sequence space satisfies this condition.

**Theorem 4.4.** *Suppose that $X$ is Fréchet differentiable and that $\Psi = (\psi_n)$ is a strictly suppression 1-unconditional basis for $X$. Then the WDGA converges for $\mathcal{D} = \{\pm\psi_n : n \geq 1\}$ and for every $f \in X$.*

*Proof.* The fact that $\Psi$ is strictly suppression 1-unconditional implies that $G_n^D = G_n^c = P_n(f)$, where $P_n$ is the natural projection of $X$ onto $[\psi_i^c]_{i=1}^n$. In particular, $(G_n^D)$ coincides with one possible 'branch' $(G_n^c)$ for the WCGA. Moreover, since $\Psi$ is unconditional, $(P_n(f))$ converges strongly to $P(f)$, where $P$ is the natural projection onto $[\psi_i^c]_{i=1}^\infty$. By Proposition 2.6, $(G_n^c)$ (and hence $(G_n^D)$) converges strongly to $f$  □

## 5. AN APPLICATION OF THE WCGA

In this section we consider a particular case: the Banach space $X = L_p(\mathrm{T}^d)$ and the dictionary $\mathcal{D} = \mathcal{RT}$ – the real trigonometric system $\{1/2, \sin x, \cos x, \dots\}$ and its $d$-dimensional version $\mathcal{RT}^d = \mathcal{RT} \times \dots \mathcal{RT}$. It is more convenient for us to consider the real $L_p(\mathrm{T}^d)$ and the real trigonometric system because the Weak Chebyshev Greedy Algorithm is defined and studied in a real Banach space. Note that the system $\mathcal{RT}$ is not normalized in $L_p$ but only semi-normalized: $C_1 \leq \|f\|_p \leq C_2$ for any $f \in \mathcal{RT}$ with absolute constants $C_1, C_2$, $1 \leq p \leq \infty$. This is sufficient for application of the general methods developed in [17]. We will compare performance of the WCGA with performance of the Thresholding Greedy Algorithm (TGA). It is proved in [15] that in the case of the complex trigonometric system $\mathcal{T}^d = \{e^{i(k,x)}\}$, for any $f \in L_p(\mathrm{T}^d)$, we have that

$$(14) \qquad \|f - G_m(f, \mathcal{T}^d)\|_p \leq Cm^{|1/2-1/p|}\sigma_m(f, \mathcal{T}^d)_p, \quad 1 \leq p \leq \infty,$$

with an absolute constant $C$. The same proof works for $\mathcal{RT}^d$ and gives

$$(15) \quad \|f - G_m(f, \mathcal{RT}^d)\|_p \leq Cm^{|1/2-1/p|}\sigma_m(f, \mathcal{RT}^d)_p, \quad 1 \leq p \leq \infty.$$

5.1. **Convergence.** It is shown in [15, Remark 4] that $G_m(\cdot, \mathcal{T})$ may not converge in $L_p$, except $p = 2$. The same is true for $G_m(\cdot, \mathcal{RT})$. The convergence of WCGA for $1 < p < \infty$ follows from general results (see [17, Th. 2.1] cited in Introduction). The divergence of $G_m(\cdot, \mathcal{T})$ and $G_m(\cdot, \mathcal{RT})$ can be fixed by adding the "Chebyshev step" in these algorithms. We describe this in the general case of $G_m(\cdot, \Psi)$. At the step $m$ instead of taking the partial sum

$$G_m(f, \Psi) = \sum_{j=1}^m c_{k_j}\psi_{k_j}$$

we take the best approximation $B_m(f, \Psi, X)$ to $f$ from Span $\{\psi_{k_1}, \ldots, \psi_{k_m}\}$. It is easy to see that in this case we have

$$\|f - B_m(f, \Psi, X)\|_X \to 0$$

as $m \to \infty$. Thus, in the sense of convergence, the Weak Chebyshev Greedy Algorithm and the Chebyshev Thresholding Greedy Algorithm (CTGA) defined above are both effective in $L_p$, $1 < p < \infty$.

5.2. **Rate of approximation.** We will compare the rate of approximation of TGA, CTGA, and WCGA for the class

$$A_1 := A_1(\mathcal{RT}) := \{f \quad : \quad \sum_{k=0}^{\infty} |a_k(f)| + |b_k(f)| \leq 1\}$$

where $a_k, b_k$ are the corresponding Fourier coefficients. From the general results on convergence rate of the Weak Chebyshev Greedy Algorithm (see [17, Th.2.2]) we get the following lemma.

**Lemma 5.1.** *For $f \in A_1$ we have*

(16) $$\|f_m^{c,t}\|_p \leq C(p,t)m^{-1/2}, \quad 2 \leq p < \infty.$$

This estimate and (15) imply, for $f \in A_1$ and $2 < p < \infty$, that

(17) $$\|f - B_m(f, \mathcal{RT}, L_p)\|_p \leq \|f - G_m(f, \mathcal{RT})\|_p \leq C(p,t)m^{-1/p},$$

which is weaker than (16).

Let us give an example showing that (17) can not be improved in the sense of order. Consider for a given $m$

$$f := (2m)^{-1} \sum_{k=1}^{2m} (1 - \frac{k}{4m}) \cos kx \quad \in \quad A_1.$$

Then $B_m(f, \mathcal{RT}, L_p)$ is the best approximation to $f$ in $L_p$ from Span $\{\cos x, \ldots, \cos mx\}$. It is not difficult to see that

(18) $$\|f - B_m(f, \mathcal{RT}, L_p)\|_p \geq C(p)m^{-1/p}.$$

This proves that (17) cannot be improved in the sense of the order of $m$.

We will show now that the constant $C(p,t)$ in (17) can be replaced by 1. We denote $p' := p/(p-1)$ and use the Hausdorff-Young theorem

(see [18, Ch. 12, Sec. 2]): for $g \in L_{p'}$, $2 \le p < \infty$, we have

$$(19) \qquad (\sum_{k=0}^{\infty} |a_k(g)|^p + |b_k(g)|^p)^{1/p} \le \|g\|_{p'}.$$

Then we have for any $f \in A_1$

$$\|f - G_m(f, \mathcal{RT})\|_p = \sup_{g, \|g\|_{p'} \le 1} |\langle f - G_m(f, \mathcal{RT}), g \rangle|$$

$$(20)$$

$$= \sup_{g, \|g\|_{p'} \le 1} |\sum_{k \notin \Lambda_c} a_k(f) a_k(g) + \sum_{l \notin \Lambda_s} b_l(f) b_l(g)|,$$

where $\Lambda_c$ and $\Lambda_s$ are such that

$$G_m(f, \mathcal{RT}) = \sum_{k \in \Lambda_c} a_k(f) \cos kx + \sum_{l \in \Lambda_s} b_l(f) \sin kx, \quad \#\Lambda_c + \#\Lambda_s = m.$$

Using the Hölder inequality we continue (20)

$$\le \sup_{\|g\|_{p'} \le 1} (\sum_{k \notin \Lambda_c} |a_k(f)|^{p'} + \sum_{l \notin \Lambda_s} |b_l(f)|^{p'})^{1/p'} (\sum_{k \notin \Lambda_c} |a_k(g)|^p + \sum_{l \notin \Lambda_s} |b_l(g)|^p)^{1/p}.$$

By (19) we have

$$\le (\sum_{k \notin \Lambda_c} |a_k(f)|^{p'} + \sum_{l \notin \Lambda_s} |b_l(f)|^{p'})^{1/p'}$$

Using the definition of $A_1$ and $\Lambda_c, \Lambda_s$ we finish the estimate

$$\le m^{1/p'-1} = m^{-1/p}.$$

Thus we have proved that for $f \in A_1$ the inequality

$$(21) \qquad \|f - G_m(f, \mathcal{RT})\|_p \le m^{-1/p}, \quad 2 \le p < \infty.$$

holds.

The relations (16) and (18) show that the WCGA gives a better rate of approximation in $L_p$ ($2 \le p < \infty$) than the CTGA.

5.3. **Constructive approximation of function classes.** In the paper [3] two types of function classes were studied from the point of view of best $m$-term trigonometric approximation. We begin with the first class. For $0 < \alpha < \infty$ and $0 < q \le \infty$, let $\mathcal{F}_q^\alpha$ denote the class of those functions in $L_1(\mathrm{T}^d)$ such that

$$|f|_{\mathcal{F}_q^\alpha} := (\sum_{k \in \mathbf{Z}^d} (\max(1, |k_1|, \ldots, |k_d|)^{\alpha q} |\hat{f}(k)|^q)^{1/q} \le 1.$$

The following theorem has been proved in [3].

**Theorem 5.2.** *If $\alpha > 0$ and $\lambda := \alpha/d + 1/q - 1/2$, then for all $1 \leq p \leq \infty$ and all $0 < q \leq \infty$*

$$C_1 m^{-\lambda} \leq \sigma_m(\mathcal{F}_q^\alpha, \mathcal{T}^d)_p \leq C_2 m^{-\lambda}, \quad \alpha > d(1 - 1/q)_+,$$

*with $C_1, C_2 > 0$ constants depending only on $d, \alpha, q$.*

The second class is defined as follows. Let $\alpha > 0$, $0 < \tau, s \leq \infty$ and $B_s^\alpha(L_\tau)$ denote the class of functions such that there exist trigonometric polynomials $T_n$ of coordinate degree $2^n$ with the properties

$$f = \sum_{n=0}^{\infty} T_n, \qquad \|\{2^{n\alpha}\|T_n\|_\tau\}_{n=0}^\infty\|_{\ell_s(\mathbb{Z})} \leq 1.$$

The following theorem has been proved in [3] for these classes.

**Theorem 5.3.** *Let $1 \leq p \leq \infty$, $0 < \tau, s \leq \infty$; and define*

$$\alpha(p, \tau) := \begin{cases} d(1/\tau - 1/p)_+, & 0 < \tau, s \leq \infty \text{ and } 1 \leq p \leq \tau \leq \infty \\ \max(d/\tau, d/2), & otherwise \end{cases}$$

*Then for $\alpha \geq \alpha(p, \tau)$, we have*

$$C_1 m^{-\mu} \leq \sigma(B_s^\alpha(L_\tau), \mathcal{T}^d)_p \leq C_2 m^{-\mu},$$

*where*

$$\mu := \alpha/d - (1/\tau - \max(1/p, 1/2))_+$$

*and $C_1, C_2$ depend only on $\alpha, p, \tau$, and $d$.*

*Remark* 5.4. Theorems 5.2 and 5.3 hold with $\mathcal{T}^d$ replaced by $\mathcal{RT}^d$.

It was proved in [17] that in the case $1 \leq p \leq 2$ the rate of best $m$-term approximation in Theorem 5.2 can be realized by $G_m(\cdot, \mathcal{T}^d)$ for the TGA, that is by a constructive method (the same is true for $\mathcal{RT}^d$). In the same case $1 \leq p \leq 2$ the rate of $\sigma_m(B_s^\alpha(L_\tau), \mathcal{T}^d)_p$ can be realized by approximating by trigonometric polynomials of degree $m^{1/d}$ in each variable. Thus, in the case $1 \leq p \leq 2$, there exist constructive methods which provide the optimal rate in Theorems 5.2 and 5.3. However, for the case $2 < p \leq \infty$ which is the most interesting case in Theorems 5.2 and 5.3 from the point of view of upper estimates (they do not depend on $p$ in this case) there was not a constructive proof of the upper estimates. The existing methods in this case are based either on a probabilistic approach (Yu. Makovoz, [13], $2 < p < \infty$) which does not cover the most intersting case $p = \infty$ or on the geometry of

finite-dimensional Banach spaces ([3], $2 < p \leq \infty$) which covers the case $p = \infty$. Both approaches contain a nonconstructive step. In [3] this step is hidden in the following inequality (see [3, Corollary 5.1])

$$(22) \qquad \sigma_m(A_1(\mathcal{T}_n^d), \mathcal{T}^d)_\infty \leq C m^{-1/2} (1 + \ln^+ \frac{n^d}{m})^{1/2},$$

where $\mathcal{T}_n^d$ denotes the subsystem of the trigonometric system $\mathcal{T}^d$ which forms a basis for the space of trigonometric polynomials of coordinate degree $n$. The inequality (22) was proved in [3] with the help of the following theorem of Gluskin [8].

**Theorem 5.5.** *There exist absolute constants $C_1$ and $0 < \delta < 1$ such that for any finite collection $V$ of $M$ vectors from the unit Euclidean ball $B_2^N$ of $\mathrm{R}^N$, there is a vector $z \in \mathrm{R}^N$ with $|z_i| = 0, 1$, $i = 1, \ldots, N$, $\|z\|_{\ell_1^N} \geq \delta N$, and*

$$\max_{v \in V} |\langle v, z \rangle| \leq C_1 (1 + \ln^+ \frac{M}{N})^{1/2}.$$

The major purpose of this section is to note that in the case $2 < p < \infty$ the Weak Chebyshev Greedy Algorithm provides a constructive way to get an analog of (22). This follows immediately from Lemma 5.1: for $f \in A_1(\mathcal{R}\mathcal{T}_n^d)$ we have

$$(23) \qquad \|f_m^{c,t}\|_p \leq C(p, t) m^{-1/2}, \quad 2 \leq p < \infty.$$

Thus the only nonconstructive step in the proof of upper estimates in Theorems 5.2 and 5.3 can be made constructive for $p < \infty$. We do not know a constructive proof of (22).

## REFERENCES

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

[2] R. Deville, G. Godefroy and V. Zizler, *Smoothness and renormings in Banach spaces*, Longman Scientific & Technical, Harlow, 1993.

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

[4] R. A. DeVore and V. N. Temlyakov, *Some remarks on Greedy Algorithms*, Adv. Comput. Math. **5** (1996), 173–187.

[5] V. V. Dubinin, *Greedy Algorithms and Applications*, Ph.D. thesis, University of South Carolina, 1997.

[6] N. Dunford and J. Schwartz, *Linear Operators, Vol. I*, Interscience, New York, 1963.

[7] Per Enflo, *Banach spaces which can be given an equivalent uniformly convex norm*, Israel J. Math. **13** (1972), 281–288.

[8] E. D. Gluskin, *Extremal properties of orthogonal parallelpipeds and their application to the geometry of Banach Spaces*, Math. USSR-Sb. **64** (1989), 85–96.

[9] Kenneth Hoffman, *Banach Spaces of Analytic Functions*, Prentice-Hall, Eaglewood Cliffs, N.J., 1962.

[10] L. D. Hoffmann, *Pseudo-uniform convexity of $H^1$ in several variables*, Proc. Amer. Math. Soc. **26** (1970), 609–614.

[11] S. V. Konyagin and V. N. Temlyakov, *A remark on greedy approximation in Banach spaces*, East J. Approx. **5** (1999), 1–15.

[12] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces I, Sequence Spaces*, Springer-Verlag, Berlin-Heidelberg, 1977.

[13] Y. Makovoz, *On trigonometric n-widths and their generalization*, J. Approx. Theory **41** (1984), 361–366.

[14] D. J. Newman, *Pseudo-uniform convexity in $H^1$*, Proc. Amer. Math. Soc. **14** (1963), 676-679.

[15] V. N. Temlyakov, *Greedy Algorithm and m-term Trigonometric Approximation*, Constr. Approx. **14** (1998), 569–587.

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

[17] V. N. Temlyakov, *Greedy algorithms in Banach spaces*, IMI-Preprint series **10** (2000), 1–11.

[18] A. Zygmund, *Trigonometric Series*, Cambridge University Press, Cambridge, 1959.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH CAROLINA, COLUMBIA, SC 29208, U.S.A.

*E-mail address*: `dilworth@math.sc.edu`

INSTITUTE OF MATHEMATICS, BULGARIAN ACADEMY OF SCIENCES, BULGARIA

*Current address*: Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.

*E-mail address*: `denka@math.sc.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH CAROLINA, COLUMBIA, SC 29208, U.S.A.

*E-mail address*: `temlyak@math.sc.edu`