# Course Notes for
# Greedy Approximations
# Math 663-601

Th. Schlumprecht

March 10, 2015

# Contents


**1 The Threshold Algorithm** **5**
- 1.1 Greedy and Quasi Greedy Bases . . . . . 5
- 1.2 The Haar basis is greedy in $L_p[0,1]$ and $L_p(\mathbb{R})$ . . . . . 15
- 1.3 Quasi greedy but not unconditional . . . . . 17

**2 Greedy Algorithms In Hilbert Space** **27**
- 2.1 Introduction . . . . . 27
- 2.2 Convergence . . . . . 30
- 2.3 Convergence Rates . . . . . 42

**3 Greedy Algorithms in general Banach Spaces** **51**
- 3.1 Introduction . . . . . 51
- 3.2 Convergence of the Weak Dual Chebyshev Greedy Algorithm . . 55
- 3.3 Weak Dual Greedy Algorithm with Relaxation . . . . . 60
- 3.4 Convergence Theorem for the Weak Dual Algorithm . . . . . 65

**4 Open Problems** **71**
- 4.1 Greedy Bases . . . . . 71
- 4.2 Greedy Algorithms . . . . . 73

**5 Appendix A: Bases in Banach spaces** **75**
- 5.1 Schauder bases . . . . . 75
- 5.2 Markushevich bases . . . . . 79

**6 Appendix B: Some facts about $L_p[0,1]$ and $L_p(\mathbb{R})$** **87**
- 6.1 The Haar basis and Wavelets . . . . . 87
- 6.2 Khintchine's inequality and Applications . . . . . 96

*Signals* or *images* are often modeled as elements of some Banach space consisting of functions, for example $C(D)$, $L_p(D)$, or more generally Sobolev spaces $W^{r,p}(D)$, for a domain $D \subset \mathbb{R}^d$. These functions need to be "processed": approximated, converted into an object which is storable, like a sequence of numbers, and then reconstructed.

This means to find an appropriate basis of the Banach space, or more generally a *dictionary* and to compute as many coordinates of the given functions with respect to this basis as necessary to satisfy the given error estimates. Now the question one needs to solve, is to find the coordinates one wants to use, given a restriction on the *budget*.

**Definition 0.0.1.** Let $X$ (always) be a separable and real Banach space. We call $\mathcal{D} \subset S_X$ a *dictionary of* $X$ if $\text{span}(\mathcal{D})$ is dense and $x \in \mathcal{D}$ implies that $-x \in \mathcal{D}$.

An *approximation algorithm* is a map

$$\mathbf{G} : X \to \text{span}(\mathcal{D})^{\mathbb{N}},\ x \mapsto \mathbf{G}(x) = (G_n(x)),$$

with the property that for $n \in \mathbb{N}$ and $x \in X$, there is a set $\Lambda_{(n,x)} \subset \mathcal{D}$ of cardinality at most $n$ so that $G_n(x) \in \text{span}(\Lambda_n)$. For $n \in \mathbb{N}$ we call $G_n(x)$ the *$n$-term approximation of $x$*.

Usually $G_n(x)$ is computed inductively by maximizing a certain value, therefore these algorithms are often called *greedy algorithms*.

**Remark.** If $X$ has a basis $(e_n)$ with biorthogonals $(e_n^*)$ and $\mathbf{G} = (P_n)$, where $P_n$ is the $n$-th canonical projection, would be an example of an approximation algorithm. Nevertheless the point is to be able to adapt the set $\Lambda_n$ to the vector $x$, and not letting it be independent. of $x$.

The main questions are

1) Does $(G_n(x))$ converge to $x$?

2) If so, how fast does it converge? How fast does it converge for *certain* $x$?

3) How does $\|x - G_n(x)\|$ compare to the *best $n$-term approximation* defined by
$$\sigma_n(x) = \sigma_n(x, \mathcal{D}) = \inf_{\Lambda \subset \mathcal{D}, \#\Lambda = n} \inf_{z \in \text{span}(\Lambda)} \|z - x\|?$$

# Chapter 1

# The Threshold Algorithm

## 1.1 Greedy and Quasi Greedy Bases

We start with the *Threshold Algorithm*:

**Definition 1.1.1.** Let $X$ be a separable Banach space with a normalized $M$-basis $\big((e_i, e_i^*) : i \in \mathbb{N}\big)$; we mean by that $\|e_i\| = 1$, for $i \in \mathbb{N}$) For $n \in \mathbb{N}$ and $x \in X$ let $\Lambda_n \subset \mathbb{N}$ so that

$$\min_{i\in\Lambda_n} |e_i^*(x)| \geq \max_{i\in\mathbb{N}\setminus\Lambda_n} |e_i^*(x)|,$$

i.e. we are reordering $(e_i^*(x))$ into $(e_{\sigma(i)}^*(x))$, so that

$$|e_{\sigma_1}^*(x)| \geq |e_{\sigma_2}^*(x)| \geq |e_{\sigma_3}^*(x)| \geq \ldots,$$

and put for $n \in \mathbb{N}$

$$\Lambda_n = \{\sigma_1, \sigma_2, \ldots \sigma_n\}.$$

Then define for $n \in \mathbb{N}$

$$G_n^T(x) = \sum_{i\in\Lambda_n} e_i^*(x) e_i.$$

$(G_n^T)$ is called the *Threshold Algorithm*.

**Definition 1.1.2.** A normalized $M$-basis $(e_i)$ is called *Quasi-Greedy*, if for all $x$

$$\text{(QG)} \qquad x = \lim_{n\to\infty} G_n^T(x).$$

A basis is called greedy if there is a constant $C$ so that

$$\text{(G)} \qquad \big\|x - G_T(x)\big\| \leq C\sigma_n(x),$$

where we define

$$\sigma_n(x) = \sigma_n\big(x, (e_j)\big) = \inf_{\Lambda\subset\mathbb{N}, \#\Lambda=n} \ \inf_{z\in\text{span}(e_j : j\in\Lambda)} \|z - x\|.$$

In that case we say that $(e_i)$ is $C$-greedy. We call the smallest constant $C$ for which (G) holds *the greedy constant of* $(e_n)$ and denote it by $C_g$.

**Remarks.** Let $\big((e_i, e_i^*) : i \in \mathbb{N}\big)$ be a normalized $M$ basis.

1. From the property that $(e_n)$ is fundamental we obtain that for every $x \in X$
$$\sigma_n(x) \to_{n\to\infty} 0,$$
it follows therefore that every greedy basis is quasi greedy.

2. If $(e_j)$ is an unconditional basis of $X$, and $x = \sum_{i=1}^\infty a_i \in X$, then
$$x = \lim_{n\to\infty} \sum_{j=1}^n a_{\pi(j)} e_{\pi(j)},$$
for any permutation $\pi : \mathbb{N} \to \mathbb{N}$ and thus, in particular, also for a *greedy permutation*, i.e. a permutation, so that
$$|a_{\pi(1)}| \geq |a_{\pi(2)}| \geq |a_{\pi(3)}| \ldots .$$
Thus, an unconditional basis is always quasi-greedy.

3. Schauder bases have a special order and might be reordered so that the cease to be basis. But
   - unconditional bases,
   - $M$ bases,
   - quasi greedy $M$-bases,
   - greedy bases

   keep their properties under any permutation, and can therefore be indexed by any countable set.

4. In order to obtain a quasi greedy $M$-Basis which is not a Schauder basis, one could take quasi greedy Schauder basis, which is not unconditional (its existence will be shown later), but admits a suitable reordering under which is not a Schauder basis anymore. Nevertheless, by the observations in (3), it will still be a quasi greedy $M$-basis. But it seems unknown whether or not there is a quasi greedy $M$-basis which cannot be reordered into a Schauder basis.

**Examples 1.1.3.** 1. If $1 \leq p < \infty$, then the unit vector basis $(e_i)$ of $\ell_p$ is 1-greedy.

2. The unit vector basis $(e_i)$ in $c_0$ is 1-greedy.
3. The summing basis $s_n$ of $c_0$ ($s_n = \sum_{j=1}^n e_j$) is not quasi greedy.
4. The unit bias of $(\ell_p \oplus \ell_q)_1$ is not greedy (but 1-unconditional and thus quasi greedy).

*Proof.* To prove (1) let $x = \sum_{j=1}^\infty x_j e_j \in \ell_p$, and let $\Lambda_n \subset \mathbb{N}$ be of cardinality $n$ so that

$$\min\{|x_j| : j \in \Lambda_n\} \geq \max\{|x_j| : j \in \mathbb{N} \setminus \Lambda_n\}$$

and $\Lambda \subset \mathbb{N}$ be any subset of cardinality $n$ and $z = \sum z_i e_i \in \ell_p$ with

$$\text{supp}(z) = \{i \in \mathbb{N} : |z_i| \neq 0\} \subset \Lambda.$$

Then

$$\begin{aligned}\|x - z\|_p^p &= \sum_{j\in\Lambda} |x_j - z_j|^p + \sum_{j\in\mathbb{N}\setminus\Lambda} |x_j|^p \\ &\geq \sum_{j\in\Lambda} |x_j - z_j|^p + \sum_{j\in\mathbb{N}\setminus\Lambda_n} |x_j|^p \\ &\geq \sum_{j\in\mathbb{N}\setminus\Lambda_n} |x_j|^p = \|G^T(x) - x\|_p.\end{aligned}$$

Thus

$$\sigma_n(x) = \inf\{\|z - x\|_p : \#\text{supp}(z) \leq n\} = \|G^T(x) - x\|_p.$$

(2) can be shown in the same way as (1).

In order to show (3) we choose sequences $(\varepsilon_j) \subset (0,1)$, $(n_j) \subset \mathbb{N}$ as follows:

$$\varepsilon_{2j} = 2^{-j} \text{ and } \varepsilon_{2j-1} = 2^{-j}\Big(1 + \frac{1}{j^3}\Big), \text{ for } j \in \mathbb{N}$$

and

$$n_j = j2^j \text{ and } N_j = \sum_{i=1}^n n_i \text{ for } i \in \mathbb{N}_0.$$

Note that the series

$$\begin{aligned}x &= \sum_{j=1}^\infty \sum_{i=N_{j-1}+1}^{N_j} (\varepsilon_{2j-1} s_{2i-1} - \varepsilon_{2j} s_{2i}) \\ &= \sum_{j=1}^\infty \sum_{i=N_{j-1}+1}^{N_j} \big((\varepsilon_{2j-1} - \varepsilon_{2j}) s_{2i-1} - \varepsilon_{2j} e_{2i}\big)\end{aligned}$$

converges, because

$$\sum_{j=1}^{\infty}\sum_{i=N_{j-1}+1}^{N_j} \varepsilon_{2j} e_{2i} \in c_0$$

and

$$\sum_{j=1}^{\infty}\sum_{i=N_{j-1}+1}^{N_j} \big\|(\varepsilon_{2j-1}-\varepsilon_{2j}) s_{2i-1}\big\| = \sum_{j=1}^{\infty} n_j(\varepsilon_{2j-1}-\varepsilon_{2j}) = \sum_{j=1}^{\infty}\frac{1}{j^2} < \infty.$$

Now we compute for $l \in \mathbb{N}_0$ the vector $x - G^T_{2N_l+n_{l+1}}(x)$:

$$x - G^T_{2N_l+n_{l+1}}(x) = -\sum_{i=N_l+1}^{N_{l+1}} \varepsilon_{2l+2} s_{2i} + \sum_{j=l+2}^{\infty}\sum_{i=N_{j-1}+1}^{N_j} (\varepsilon_{2j-1} s_{2i-1} - \varepsilon_{2j} s_{2i}).$$

From the monotonicity of $(s_i)$ we deduce that

$$\Big\|\sum_{j=l+2}^{\infty}\sum_{i=N_{j-1}+1}^{N_j} \varepsilon_{2j-1} s_{2i-1} - \varepsilon_{2j} s_{2i}\Big\| \le \|x\|.$$

However,

$$\Big\|\sum_{i=N_l+1}^{N_{l+1}} \varepsilon_{2l+2} s_{2i}\Big\| = \sum_{i=N_l+1}^{N_{l+1}} \varepsilon_{2l+2} = l+1 \to_{l\to\infty} \infty,$$

which implies that $G^T_n(x)$ is not convergent.

To show (4) assume w.l.o.g. $p < q$, and denote the unit vector basis of $\ell_p$ by $(e_i)$ and the unit vector basis of $\ell_q$ by $(f_j)$ for $n \in \mathbb{N}$ and we put

$$x(n) = \sum_{j=1}^{n} \frac{1}{2} e_j + \sum_{j=1}^{n} f_j.$$

Thus

$$G^T_n(x(n)) = \sum_{j=1}^{n} f_j, \text{ and thus } \big\|G^T_n(x(n)) - x(n)\big\| = \frac{1}{2} n^{1/p}.$$

Nevertheless

$$\Big\|x - \sum_{j=1}^{n} \frac{1}{2} e_j\Big\| = n^{1/q},$$

and since $\frac{1}{2} n^{1/p}/n^{1/q} \nearrow \infty$, for $n \nearrow \infty$, the basis $\{e_j : j \in \mathbb{N}\} \cup \{f_j : j \in \mathbb{N}\}$ cannot be greedy. □

**Remarks.** With the arguments used in (4) Examples 1.1.3 one can show that the usual bases of $\big(\oplus_{n=1}^\infty \ell_q^n\big)_{\ell_p}$ and $\ell_p(\ell_q) = \big(\oplus_{n=1}^\infty \ell_q\big)_{\ell_p}$ are also not greedy but of course unconditional.

Now in [BCLT] it was shown that $\ell_p \oplus \ell_q$ has up to permutation and up to isomorphic equivalence a unique unconditional basis, namely the one indicated above. Since, as it will be shown later, every greedy basis must be unconditional, the space does not have any greedy basis.

Due to a result in [DFOS] however $\big(\oplus_{n=1}^\infty \ell_q^n\big)_{\ell_p}$ has a greedy bases if $1 < p, q < \infty$. More precisely, the following was shown:

Let $1 \le p, q \le \infty$.

a) If $1 < q < \infty$ then the Banach space $(\oplus_{n=1}^\infty \ell_p^n)_{\ell_q}$ has a greedy basis.

b) If $q = 1$ or $q = \infty$, and $p \ne q$, then $(\oplus_{n=1}^\infty \ell_p^n)_{\ell_q}$ has not a greedy basis. Here we take $c_0$-sum if $q = \infty$.

The question whether or not $\ell_p(\ell_q)$ has a greedy basis is open and quite an interesting question.

The following result by Wojtaszczyk can be seen the analogue of the characterization of Schauder bases by the uniform boundedness of the canonical projections for quasi-greedy bases.

**Theorem 1.1.4.** [Wo2] *A bounded $M$-basis $(e_i, e_i^*)$, with $\|e_i\| = 1$, $i \in \mathbb{N}$, of a Banach space $X$ is quasi greedy if and only if there is a constant $C$ so that for any $x \in X$ and any $m \in \mathbb{N}$ it follows that*

$$\|G_m^T(x)\| \le C\|x\| \tag{1.1}$$

*We call the smallest constant so that* (1.1) *is satisfied the* Greedy Projection Constant.

**Remark.** Theorem 1.1.4 is basically a *uniform boundedness result.* Nevertheless, since the $G_m^T$ are nonlinear projections we need a direct proof.

We need first the following Lemma:

**Lemma 1.1.5.** *Assume there is no positive number $C$ so that $\|G_m^T(x)\| \le C\|x\|$ for all $x \in X$ and all $m \in \mathbb{N}$. Then the following holds:*

*For all finite $A \subset \mathbb{N}$ all $K > 0$ there is a finite $B \subset N$, which is disjoint from $A$ and a vector $x$, with $x = \sum_{j\in B} x_j e_j$, such that $\|x\| = 1$ and $\|G_m^T(x)\| \ge K$, for some $m \in \mathbb{N}$.*

*Proof.* For a finite set $F \subset \mathbb{N}$, define $P_F$ to be the coordinate projection onto $\text{span}(e_i : i \in F)$, generated by the $(e_i^*)$, i.e.

$$P_F : X \to \text{span}(e_i : i \in F),\ x \mapsto P_F(x) = \sum_{j\in F} e_j^*(x)e_j.$$

Since there are only finitely many subsets of $A$ we can put

$$M = \max_{F\subset A} \|P_F\| = \max_{F\subset A} \sup_{x\in B_X} \Big\| \sum_{j\in F} e_j^*(x) e_j \Big\| \le \sum_{j\in A} \|e_j^*\| \cdot \|e_j\| < \infty.$$

Let $K_1 > 1$ so that $(K_1 - M)/(M+1) > K$, and choose $x_1 \in S_X \cap \mathrm{span}(e_j : j \in \mathbb{N})$ and $k \in \mathbb{N}$ so that so that $\|G_k^T(x_1)\| \ge K_1$. We assume without loss of generality (after suitable small perturbation) that all the non zero numbers $|e_n^*(x_1)|$ are different from each other.

Then let $x_2 = x_1 - P_A(x_1)$, and note that $\|x_2\| \le M + 1$ and $G_k^T(x_1) = G_m^T(x_2) + P_F(x_1)$ for some $m \le k$ and $F \subset A$. Thus $\|G_m^T(x_2)\| \ge K_1 - M$, and if we define $x_3 = x_2/\|x_2\|$, we have $\|G_k^T(x_3)\| \ge (K_1 - M)/(M + 1) > K$.

It follows that the support $B$ of $x = x_3$ is disjoint from $A$ and that $\|G_m^T(x)\| > K$. □

*Proof of Theorem 1.1.4.* Let $b = \sup_i \|e_i^*\|$.
"⇒" Assume there is no positive number $C$ so that $\|G_m^T(x)\| \le C\|x\|$ for all $x \in X$ and all $m \in \mathbb{N}$.

Applying Lemma 1.1.5 we can choose recursively vectors $y_1, y_2, \ldots$ in $S_X \cap \mathrm{span}(e_j : j \in \mathbb{N})$ and numbers $m_n \in \mathbb{N}$, so that the supports of the $y_n$, which we denote by $B_n$, are pairwise disjoint, (Recall that for $z = \sum_{i=1}^\infty z_i e_i$, we call $\mathrm{supp}(z) = \{i \in \mathbb{N} : e_i^*(z) \ne 0\}$, the *support of* $z$) and so that

$$\|G_{m_n}^T(y_n)\| \ge 2^n b^n \prod_{j=1}^{n-1} \varepsilon_j^{-1}, \tag{1.2}$$

where

$$\varepsilon_j = \min\big\{2^{-j}, \min\{|e_i^*(y_j)| : i \in B_j\}\big\}.$$

Then we let

$$x = \sum_{n=1}^\infty \Big(\prod_{j=1}^{n-1} (\varepsilon_j/b)\Big) y_n,$$

(which clearly converges) and write $x$ as

$$x = \sum_{j=1}^\infty x_j e_j.$$

Since $|e_i^*(y_j)| \le b$, for $i, j \in \mathbb{N}$

$$\min\Big\{|x_i| : i \in \bigcup_{j=1}^n B_j\Big\} \ge \prod_{j=1}^{n-1} \frac{\varepsilon_j}{b} \varepsilon_n = \prod_{j=1}^n \frac{\varepsilon_j}{b} b \ge \max\Big\{|x_i| : i \in \mathbb{N} \setminus \bigcup_{j=1}^n B_j\Big\}.$$

We may assume w.l.o.g. that $m_n \le \#B_n$, for $n \in$N. Letting $k_j = m_j + \sum_{i=1}^{j-1} \#B_i$, it follows that

$$G^T_{k_j}(x) = \sum_{i=1}^{j-1} \Big(\prod_{s=1}^{i-1}(\varepsilon_s/b)\Big) y_i + G^T_{m_j}\bigg(\Big(\prod_{i=1}^{j-1}(\varepsilon_i/b)\Big) y_j\bigg).$$

and thus by (1.2)

$$\big\|G^T_{k_j}(x)\big\| \ge \bigg\|G^T_{m_j}\bigg(\Big(\prod_{i=1}^{j-1}(\varepsilon_i/b)\Big) y_j\bigg)\bigg\| - \sum_{i=1}^{j-1} \Big(\prod_{s=1}^{i-1}(\varepsilon_s/b)\Big)\|y_i\| \ge 2^j b,$$

which implies that $G^T_{k_j}$ does not converge.
"$\Leftarrow$" Let $C > 0$ such that $\|G^T_m(x)\| \le C\|x\|$ for all $m \in \mathbb{N}$ and all $x \in X$. Let $x \in X$ and assume w.l.o.g. that supp$(x)$ is infinite. For $\varepsilon > 0$ choose $x_0$ with finite support $A$ so that $\|x - x_0\| < \varepsilon$. Using small perturbations we can assume that $A \subset \text{supp}(x)$ and that $A \subset \text{supp}(x - x_0)$. We can therefore choose $m \in \mathbb{N}$ large enough so that $G^T_m(x)$ and $G^T_m(x - x_0)$ are of the form

$$G^T_m(x) = \sum_{j\in B} e^*_i(x)e_i \text{ and } G^T_m(x - x_0) = \sum_{j\in B} e^*_i(x - x_0)e_i$$

with $B \subset \mathbb{N}$ such that $A \subset B$. It follows therefore that

$$\|x - G^T_m(x)\| \le \|x - x_0\| + \|x_0 - G^T_m(x)\| = \|x - x_0\| + \|G^T_m(x_0 - x)\| \le (1 + C)\varepsilon,$$

which implies our claim by choosing $\varepsilon > 0$ to be arbitrarily small. □

**Definition 1.1.6.** An $M$ basis $(e_j, e^*_j)$ is called *unconditional for constant coefficients* if there is a positive constant $C$ so that for all finite sets $A \subset \mathbb{N}$ and all $(\sigma_n : n \in A) \subset \{\pm 1\}$ we have

$$\frac{1}{C}\Big\|\sum_{n\in A} e_n\Big\| \le \Big\|\sum_{n\in A}\sigma_n e_n\Big\| \le C\Big\|\sum_{n\in A} e_n\Big\|.$$

**Proposition 1.1.7.** *A quasi-greedy $M$ basis $(e_n, e^*_n)$ is* unconditional for constant coefficients*. Actually the constant in Definition 1.1.6 can be chosen to be equal to twice the projection constant in Theorem 1.1.4.*

**Remark.** We will show later that there are quasi-greedy bases which are not unconditional. Actually there are Banach spaces which do not contain any unconditional basic sequence, but in which every normalized weakly null sequence contains a quasi-greedy subsequence.

*Proof of Proposition 1.1.7.* Let $A \subset \mathbb{N}$ be finite and $(\sigma_n : n \in A) \subset \{\pm 1\}$. Then if we let $\delta \in (0,1)$ and put $m = \#\{j \in A : \sigma_j = +1\}$ we obtain

$$\begin{aligned}\Big\| \sum_{n\in A,\sigma_n=+1} e_n \Big\| &= \Big\| G_m^T \Big( \sum_{n\in A,\sigma_n=+1} e_n + \sum_{n\in A,\sigma_n=-1} (1-\delta) e_n \Big) \Big\| \\ &\le C \Big\| \sum_{n\in A,\sigma_n=+1} e_n + \sum_{n\in A,\sigma_n=-1} (1-\delta) e_n \Big\|.\end{aligned}$$

By taking $\delta > 0$ to be arbitrarily small, we obtain that

$$\Big\| \sum_{n\in A,\sigma_n=+1} e_n \Big\| \le C \Big\| \sum_{n\in A} e_n \Big\|.$$

Similarly we have

$$\Big\| \sum_{n\in A,\sigma_n=-1} e_n \Big\| \le C \Big\| \sum_{n\in A} e_n \Big\|,$$

and thus,

$$\Big\| \sum_{n\in A} \sigma_n e_n \Big\| \le 2C \Big\| \sum_{n\in A} e_n \Big\|.$$

□

We now present a characterization of greedy bases obtained by Konyagin and Temliakov. We need the following notation.

**Definition 1.1.8.** We call a a normalized basic sequence *democratic* if there is a constant $C$ so that for all finite $E, F \subset \mathbb{N}$, with $\#E = \#F$ it follows that

$$\Big\| \sum_{j\in E} e_j \Big\| \le C \Big\| \sum_{j\in F} e_j \Big\| \tag{1.3}$$

In that case we call the smallest constant, so that (1.3) holds, the *Constant of Democracy of* $(e_i)$ and denote it by $C_d$.

**Theorem 1.1.9.** [KT1] *A normalized basis* $(e_n)$ *is greedy if and only it is unconditional and democratic. In this case*

$$\max(C_s, C_d) \le C_g \le C_d C_s C_u^2 + C_u, \tag{1.4}$$

*where* $C_u$ *is the unconditional constant and* $C_s$ *is the suppression constant.*

**Remark.** The proof will show that the first inequality is sharp. Recently it was shown in [DOSZ1] that the second inequality is also sharp.

*Proof of Theorem 1.1.9.* "$\Leftarrow$" Let $x = \sum e_i^*(x)e_i \in X$, $n \in \mathbb{N}$ and let $\eta > 0$. Choose $\tilde{x} = \sum_{i\in\Lambda_n^*} a_i e_i$ so that $\#\Lambda_n^* = n$ which is up to $\eta$ the best $n$ term approximation to $x$ (since we allow $a_i$ to be 0, we can assume that $\#\Lambda$ is exactly $n$), i.e.

$$\|x - \tilde{x}\| \le \sigma_n(x) + \eta. \tag{1.5}$$

Let $\Lambda_n$ be a set of $n$ coordinates for which

$$b := min_{i\in\Lambda_n}|e_i^*(x)| \ge \max_{i\in\mathbb{N}\setminus\Lambda_n} |e_i^*(x)| \text{ and } G_n^T(x) = \sum_{i\in\Lambda_n} e_i^*(x)e_i.$$

We need to show that

$$\|x - G_n^T(x)\| \le (C_d C_s C_u^2 + C_u)(\sigma_n(x) + \eta).$$

Then

$$x - G_n^T(x) = \sum_{i\in\mathbb{N}\setminus\Lambda_n} e_i^*(x)e_i = \sum_{i\in\Lambda_n^*\setminus\Lambda_n} e_i^*(x)e_i + \sum_{i\in\mathbb{N}\setminus(\Lambda_n^*\cup\Lambda_n)} e_i^*(x)e_i.$$

But we also have

$$\begin{aligned}
\Big\|\sum_{i\in\Lambda_n^*\setminus\Lambda_n} e_i^*(x)e_i\Big\| &\le bC_u\Big\|\sum_{i\in\Lambda_n^*\setminus\Lambda_n} e_i\Big\| \text{ (By Proposition 5.1.11)} \\
&\le bC_uC_d\Big\|\sum_{i\in\Lambda_n\setminus\Lambda_n^*} e_i\Big\| \\
&\quad [\text{Note that } \#(\Lambda_n\setminus\Lambda_n^*) = \#(\Lambda_n^*\setminus\Lambda_n)] \\
&\le C_u^2C_d\Big\|\sum_{i\in\Lambda_n\setminus\Lambda_n^*} e_i^*(x)e_i\Big\| \\
&\quad [\text{Note that } |e_i^*(x)| \ge b \text{ if } i\in\Lambda_n\setminus\Lambda_n^*\ ] \\
&\le C_sC_u^2C_d\Big\|\sum_{i\in\Lambda_n^*}(e_i^*(x)-a_i)e_i + \sum_{i\in\mathbb{N}\setminus\Lambda_n^*} e_i^*(x)e_i\Big\| \\
&= C_sC_u^2C_d\|x-\tilde{x}\| \le C_sC_u^2C_d(\sigma_n(x)+\eta)
\end{aligned} \tag{1.6}$$

and

$$\begin{aligned}
\Big\|\sum_{i\in\mathbb{N}\setminus(\Lambda_n^*\cup\Lambda_n)} e_i^*(x)e_i\Big\| &\le C_s\Big\|\sum_{i\in\Lambda_n^*}(e_i^*(x)-a_i)e_i + \sum_{i\in\mathbb{N}\setminus\Lambda_n^*} e_i^*(x)e_i\Big\| \\
&= C_s\|x-\tilde{x}\| \le C_s(\sigma_n(x)+\eta).
\end{aligned} \tag{1.7}$$

This shows that $(e_i)$ is greedy and, since $\eta > 0$ is arbitrary, we deduce that $C_g \le C_sC_u^2C_d + C_s$.

"$\Rightarrow$" Assume that $(e_i)$ is greedy. In order to show that $(e_i)$ is democratic let $\Lambda_1, \Lambda_2 \subset \mathbb{N}$ with $\#\Lambda_1 = \#\Lambda_2$. Let $\eta > 0$ and put $m = \#(\Lambda_2 \setminus \Lambda_1)$ and

$$x = \sum_{i \in \Lambda_1} e_i + (1+\eta) \sum_{i \in \Lambda_2 \setminus \Lambda_1} e_i.$$

Then it follows

$$\begin{aligned} \Big\| \sum_{i \in \Lambda_1} e_i \Big\| &= \|x - G_m^T(x)\| \\ &\le C_g \sigma_m(x) \text{ (since } (e_i) \text{ is } C_g\text{-greedy)} \\ &\le C_g \Big\| x - \sum_{i \in \Lambda_1 \setminus \Lambda_2} e_i \Big\| \le C_g \Big\| \sum_{i \in \Lambda_1 \cap \Lambda_2} e_i + (1+\eta) \sum_{i \in \Lambda_2 \setminus \Lambda_1} e_i \Big\|. \end{aligned}$$

Since $\eta > 0$ can be taken arbitrary, we deduce that

$$\Big\| \sum_{i \in \Lambda_1} e_i \Big\| \le C_g \Big\| \sum_{i \in \Lambda_2} e_i \Big\|.$$

Thus, it follows that $(e_i)$ is democratic and $C_d \le C_g$.

In order to show that $(e_i)$ is unconditional let $x = \sum e_i^*(x) e_i \in X$ have finite support $S$. Let $\Lambda \subset S$ and put

$$y = \sum_{i \in \Lambda} e_i^*(x) e_i + b \sum_{i \in S \setminus \Lambda} e_i,$$

with $b > \max_{i \in S} |e_i^*(x)|$. For $n = \#(S \setminus \Lambda)$ it follows that

$$G_n^T(y) = b \sum_{i \in S \setminus \Lambda} e_i,$$

and since $(e_i)$ is greedy we deduce that (note that $\#\text{supp}(y - x) = n$)

$$\Big\| \sum_{i \in \Lambda} e_i^*(x) e_i \Big\| = \|y - G_n^T(y)\| \le C_g \sigma_n(y) \le C_g \|y - (y - x)\| = C_g \|x\|,$$

which implies that $(e_i)$ is unconditional with $C_s \le C_g$. $\square$

## 1.2 The Haar basis is greedy in $L_p[0,1]$ and $L_p(\mathbb{R})$

**Theorem 1.2.1.** *For $1 < p < \infty$ there are two constants $c_p \le C_p$, depending only on $p$, so that for all $n \in \mathbb{N}$ and all $A \subset T$ with $\#A = n$*

$$c_p n^{1/p} \le \Big\| \sum_{t\in A} h_t^{(p)} \Big\| \le C_p n^{1/p}.$$

*In particular $(h_t^{(p)})_{t\in T}$ is democratic in $L_p[0,1]$.*

With Theorem 1.1.9 and Theorem 6.1.1 we deduce that

**Corollary 1.2.2.** *The Haar Basis of $L_p[0,1]$, $1 < p < \infty$ is greedy.*

The proof will follow from the following three Lemmas.

**Lemma 1.2.3.** *For any $0 < q < \infty$ there is a $d_q > 0$ so that the following holds.*

*Let $n_1 < n_2 < \dots n_k$ be integers and let $E_j \subset [0,1]$ be measurable for $j = 1, \dots k$. Then we have*

$$\int_0^1 \Big(\sum_{j=1}^k 2^{n_j/q} 1_{E_j}(x)\Big)^q dx \le d_q \sum_{j=1}^k 2^{n_j} m(E_j).$$

*Proof.* Define

$$f(x) = \sum_{j=1}^k 2^{n_j/q} 1_{E_j}(x).$$

For $j = 1, \dots k$ write $E'_j = E_j \setminus \bigcup_{i=j+1}^k E_i$. It follows that for $x \in E'_j$

$$f(x) \le \sum_{i=1}^j 2^{n_i/q} \le \sum_{i=1}^{n_j} 2^{i/q} = \frac{2^{(n_j+1)/q} - 1}{2^{1/q} - 1} \le \underbrace{\frac{2^{1/q}}{2^{1/q} - 1}}_{d_q^{1/q}} 2^{n_j/q}.$$

Thus

$$\int_0^1 f(x)^q dx \le d_q \sum_{i=1}^k 2^{n_i} m(E'_i) \le d_q \sum_{j=1}^k 2^{n_j} m(E_j),$$

which finishes the proof. □

**Lemma 1.2.4.** *For $1 < p < \infty$ there is a $C_p > 0$ so that for all $n \in \mathbb{N}$, $A \subset T$ with $\#A = n$, and $(\varepsilon_t) \subset \{-1, 1\}$ it follows that*

$$\Big\| \sum_{t\in A} \varepsilon_t h_t^{(p)} \Big\|_p \le C_p n^{1/p}.$$

*Proof.* Abbreviate $h_t = h_t^{(p)}$ for $t \in T$. Let $n_1 < n_2 < \ldots < n_k$ be all the integers $n_i$ for which there is a $t \in A$ so that $m(\mathrm{supp}(h_t)) = 2^{-n_i}$. For $j = 1, \ldots k$ put

$$E_j = \bigcup_{i \in \{0,1,\ldots 2^{n_j}-1\}, (n_j,i)\in A} \mathrm{supp}(h_{(i,n_j)}).$$

Since

$$m(E_j) = 2^{-n_j} \#\{i \in \{0,1,\ldots 2^{n_j}-1\}, (n_j,i) \in A\}$$

and thus

$$\#\{i \in \{0,1,\ldots 2^{n_j}-1\}, (n_j,i) \in A\} = 2^{n_j} m(E_j).$$

It follows therefore that

$$n = \begin{cases} \sum_{j=1}^k \#\{i \in \{0,1,\ldots 2^{n_j}-1\}, (n_j,i) \in A\} = \sum_{j=1}^k 2^{n_j} m(E_j) & \text{if } 0 \notin A \\ 1 + \sum_{j=1}^k 2^{n_j} m(E_j) & \text{if } 0 \in A. \end{cases}$$

Assume without loss of generality that $0 \notin A$. It follows that

$$\Big\| \sum_{t\in A} \varepsilon_t h_t \Big\|_p = \left[ \int_0^1 \Big[ \sum_{j=1}^k 2^{n_j/p} 1_{E_j} \Big]^p dx \right]^{1/p} \le d_p^{1/p} \left[ \sum_{j=1}^k 2^{n_j} m(E_j) \right]^{1/p} = d_p^{1/p} n^{1/p}.$$

[$d_p$ as in Lemma 1.2.3]

□

**Lemma 1.2.5.** *For $1 < p < \infty$ there is a $c_p > 0$ so that for all $n \in \mathbb{N}$, $A \subset T$ with $\#A = n$, and $(\varepsilon_t) \subset \{-1, 1\}$ it follows that*

$$\Big\| \sum_{t\in A} \varepsilon_t h_t^{(p)} \Big\|_p \ge c_p n^{1/p}.$$

*Proof.* Note that for $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q}$ and $s, t \in T$ it follows that

$$\langle h_t^{(p)}, h_s^{(q)} \rangle = \delta(t, s),$$

thus the claim follows from the fact that the $h_t^{(p)}$'s are normalized in $L_p[0,1]$ and by Lemma 1.2.4 using the duality between $L_p[0,1]$ and $L_q[0,1]$. Indeed,

$$\Big\| \sum_{t\in A} \varepsilon_t h_t^{(p)} \Big\| \ge \left\langle \sum_{t\in A} \varepsilon_t h_t^{(p)}, \frac{\sum_{t\in A} \varepsilon_t h_t^{(q)}}{\big\| \sum_{t\in A} \varepsilon_t h_t^{(q)} \big\|} \right\rangle$$
$$= \frac{n}{\big\| \sum_{t\in A} \varepsilon_t h_t^{(q)} \big\|} \ge \frac{n^{1/p}}{c_q},$$

where $c_q$ is chosen like in Lemma 1.2.5. Our claim follows therefore bu letting $C_p = 1/c_q$. □

## 1.3 A quasi greedy basis of $L_p[0,1]$ which is not unconditional

In this section we make the general assumption on a separable Banach space $X$, that $X$ has a normalized basis $(e_n)$ which is *Besselian* meaning that for some constant $C_B$

$$\|x\| = \Big\|\sum_{j=1}^\infty e_j^*(x)e_j\Big\| \geq \frac{1}{C_B}\Big(\sum_{j=1}^\infty |e_j^*(x)|^2\Big)^{1/2} \text{ for all } x \in X. \tag{1.8}$$

where $(e_j^*)$ denote the coordinate functionals for $(e_j)$ We secondly assume that $(e_j)$ has a subsequence $(e_{m_j} : j \in \mathbb{N})$ which is *Hilbertian* which means that for some constant $C_H$

$$\Big\|\sum_{j=1}^\infty e_{m_j}^*(x)e_{m_j}\Big\| \leq C_H\Big(\sum_{j=1}^\infty |e_{m_j}^*(x)|^2\Big)^{1/2} \text{ for all } x \in \overline{\text{span}(e_{m_j} : j \in \mathbb{N})}. \tag{1.9}$$

**Example 1.3.1.** An example for such a basic sequence are the trigonometrically polynomial $(t_n : n \in \mathbb{Z})$ in $L_p[0,1]$ with $p > 2$. Indeed, for $(a_n : |n| \leq N) \subset \mathbb{C}$ it follows from Hölder's (or Jensen's) inequality that

$$\Big(\int_0^1 \Big|\sum_{n=-N}^N a_j e^{in\xi/2\pi}\Big|^p d\xi\Big)^{1/p} \geq \Big(\int_0^1 \Big|\sum_{n=-N}^N a_j e^{in\xi/2\pi}\Big|^2 d\xi\Big)^{1/2} = \Big(\sum_{n=-N}^N |a_j|^2\Big)^{1/2}.$$

Secondly it follows from the complex version of Khintchine' s inequality (Theorem 6.2.4) that the subsequence $(t_{2^n} : n \in \mathbb{N})$ of the trigonometric polynomials is equivalent to the $\ell_2$-unit vector basis.

We recall the $2^n$ by $2^n$ matrices $A^{(n)} = (a^{(n)}_{(i,j)} : 1 \leq i,j \leq 2^n)$, for $n\in\mathbb{N}$, which were introduced in Section 5.2. Let us recall the following two properties which we will need here:

$$A^{(n)} \text{ is unitary operator on } \ell_2^{2^n}, \text{ and} \tag{1.10}$$

$$a^{(n)}_{(j,1)} = 2^{-n/2}. \tag{1.11}$$

For $k\in\mathbb{N}$ we put $n_k = 2^{2^k}$ and $B^{(k)} = (b^{(k)}_{i,j} : 1 \leq i,j \leq n_k) = A^{(2^k)}$ (acting on $\ell_2^{n_k}$), for $k\in\mathbb{N}$ which implies that $n_{k+1} = n_k^2$. We let

$$(h_j : j \in \mathbb{N}) = (e_{m_j} : j \in \mathbb{N}) \text{ and } (f_i : i\in\mathbb{N}) = (e_s : s \in \mathbb{N} \setminus \{m_i : i\in\mathbb{N}\}),$$

so that if $f_i = e_s$ and $f_j = e_t$ then then $i < j$ if and only if $s < t$. For $k\in\mathbb{N}$ we define a family $(g_j^{(k)} : j = 1, 2, \dots n_k)$ as follows

$$g_1^{(k)} = f_k \text{ and } g_i^{(k)} = h_{S_{k-1}+i-1}, \text{ for } i = 2, 3, \dots n_k,$$

where $S_0 = 0$, and, inductively, $S_j = S_{j-1} + n_j - 1$. If we order $(g_j^{(k)} : k \in \mathbb{N}, j = 1, 2, \dots n_k)$ lexicographically we note that the sequence

$$\left(g_1^{(1)}, g_2^{(1)}, \dots g_{n_1}^{(1)}, g_1^{(2)}, \dots, g_{n_2}^{(2)}, g_1^{(3)}, , \dots\right)$$

is equal to the sequence

$$\left(f_1, h_1, h_2, \dots h_{n_1-1}, f_2, h_{n_1}, \dots h_{n_2-2}, f_3, \dots\right).$$

Then we define for $k \in \mathbb{N}$ a new system of elements $(\psi_j^{(k)} : j = 1, 2 \dots n_k)$, by

$$\begin{pmatrix} \psi_1^{(k)} \\ \psi_2^{(k)} \\ \vdots \\ \psi_{n_k}^{(k)} \end{pmatrix} = B^{(k)} \circ \begin{pmatrix} g_1^{(k)} \\ g_2^{(k)} \\ \vdots \\ g_{n_k}^{(k)} \end{pmatrix} \tag{1.12}$$

or, in other words,

$$\psi_i^{(k)} = \sum_{j=1}^{n_k} b_{(i,j)}^{(k)} g_j^{(k)} \text{ for } i = 1, 2 \dots n_k.$$

Our goal is now to prove the following result

**Theorem 1.3.2.** *Ordered lexicographically, the system $(\psi_j^{(k)} : k \in \mathbb{N}, j = 1, 2 \dots n_k)$ is a quasi-greedy basis of $X$.*

**Proposition 1.3.3.** *Ordered lexicographically, $(g_j^{(k)} : k \in \mathbb{N}, j = 1, 2 \dots n_k)$ is a Besselian basis of $X$.*

*Proof.* Given that $(g_j^{(k)} : k \in \mathbb{N}, j = 1, 2 \dots n_k)$ is a reordering of $(e_j)$, which was assumed to be a Besselian basis of $X$, we only need to show that $(g_j^{(k)} : k \in \mathbb{N}, j = 1, 2 \dots n_k)$ is a basic sequence.

To do so we need to show that there is a constant $C \geq 1$ so that for all $N \in \mathbb{N}$, all $M \in \{1, 2 \dots n_M\}$ and all $(c_j^{(k)} : k \in \mathbb{N}, j = 1, 2 \dots n_k)$, with

$$\#\{(k, j) : k \in \mathbb{N}, j = 1, 2 \dots n_k, c_j^{(k)} \neq 0\}$$

being finite, it follows

$$\Big\| \sum_{k=1}^{N-1} \sum_{j=1}^{n_k} c_j^{(k)} g_j^{(k)} + \sum_{j=1}^{M} c_j^{(N)} g_j^{(N)} \Big\| \leq C \Big\| \sum_{k=1}^{\infty} \sum_{j=1}^{n_k} c_j^{(k)} g_j^{(k)} \Big\|. \tag{1.13}$$

Since the $g_i^{(k)}$ are a reordering of the original basis $(e_j)$ we can write

$$x = \sum_{k=1}^{\infty} \sum_{j=1}^{n_k} c_j^{(k)} g_j^{(k)} \text{ as } x = \sum_{i=1}^{\infty} c_i e_i,$$

where $c_i = c_j^{(k)}$ if $e_i = g_j^{(k)}$ (and for each $i \in \mathbb{N}$ there is exactly one such choice of $k$ and $j \in \{1, 2 \dots n_k\}$). From (1.8) and (1.9) we deduce that

$$\Big\| \sum_{k=1}^{N-1} \sum_{j=2}^{n_k} c_j^{(k)} g_j^{(k)} + \sum_{j=2}^{M} c_j^{(N)} g_j^{(N)} \Big\| \le C_H \Big( \sum_{k=1}^{N-1} \sum_{j=2}^{n_k} |c_j^{(k)}|^2 + \sum_{j=2}^{M} |c_j^{(N)}|^2 \Big)^{1/2} \le C_H \Big( \sum_{j=1}^{\infty} |c_j|^2 \Big)^{1/2} \le C_H C_B \|x\|. \tag{1.14}$$

Since $g_j^{(k)} = f_j = e_{s_j}$, where $(s_j)$ which consists of the elements of $\mathbb{N} \setminus \{m_j : j \in \mathbb{N}\}$, ordered increasingly it follows that we can write

$$\sum_{k=1}^{N} c_1^{(k)} g_1^{(k)} = \sum_{j=1}^{s_N} c_j e_j - \sum_{i \in \{1,2,\dots s_N\} \setminus \{s_j : j \le N\}} c_i e_i = \sum_{(k,j) \in A} c_j^{(k)} g_j^{(k)},$$

for some set $A \subset \{(k,j) : k \in \mathbb{N}, j = 2, 3 \dots n_k\}$. If $C_e$ is the basis constant of $(e_j)$ we deduce therefore that

$$\Big\| \sum_{j=1}^{s_N} c_j e_j \Big\| \le C_e \|x\|,$$

and thus, using (1.14),

$$\Big\| \sum_{k=1}^{N} c_1^{(k)} g_1^{(k)} \Big\| \le \Big\| \sum_{j=1}^{s_N} c_j e_j \Big\| + \Big\| \sum_{(k,j) \in A} c_j^{(k)} g_j^{(k)} \Big\| \le (C_e + C_B C_H) \|x\|.$$

This implies that

$$\begin{aligned} &\Big\| \sum_{k=1}^{N-1} \sum_{j=1}^{n_k} c_j^{(k)} g_j^{(k)} + \sum_{j=1}^{M} c_j^{(N)} g_j^{(N)} \Big\| \\ &\le \Big\| \sum_{k=1}^{N} c_1^{(k)} g_1^{(k)} \Big\| + \Big\| \sum_{k=1}^{N-1} \sum_{j=2}^{n_k} c_j^{(k)} g_j^{(k)} + \sum_{j=2}^{M} c_j^{(N)} g_j^{(N)} \Big\| \\ &\le (C_e + C_B C_H) \|x\| + C_B C_H \|x\| \end{aligned}$$

which implies our claim with $C = C_e + 2C_B C_H$. □

**Proposition 1.3.4.** *Under the lexicographical order,* $(\psi_j^{(k)} : j = 1, 2 \dots n_k)$ *is a Besselian basis of* $X$ *with the same constant* $C_B$.

*Proof.* We first note that for $k \in \mathbb{N}$

$$X_k = \text{span}(\psi_j^{(k)} : j = 1, 2 \dots n_k) = \text{span}(g_j^{(k)} : j = 1, 2 \dots n_k)$$

and thus it follows that $(\psi_j^{(k)} : k \in \mathbb{N}, j = 1.2, \dots n_k)$ spans as $(g_j^{(k)} : k \in \mathbb{N}, j = 1.2, \dots n_k)$ a dense subspace of $X$.

Secondly we observe that if

$$(d_j^{(k)} : k \in \mathbb{N},\, j = 1, 2 \dots n_k) \subset \mathbb{K}$$

with

$$\#\{(k, j) : k \in \mathbb{N}, j = 1, 2 \dots n_k, d_j^{(k)} \neq 0\}$$

and let

$$x = \sum_{k=1}^{\infty} \sum_{j=1}^{n_k} d_j^{(k)} \Psi_j^{(k)}$$

or in $g$-coordinates:

$$x = \sum_{k=1}^{\infty} \sum_{j=1}^{n_k} c_j^{(k)} g_j^{(k)}.$$

We write $x = \sum_{k=1}^{\infty} x_k$ with

$$x_k = \sum_{j=1}^{n_k} d_j^{(k)} \psi_j^{(N)} = \sum_{j=1}^{n_k} c_j^{(k)} g_j^{(k)}.$$

Since

$$x_k = \sum_{i=1}^{n_k} d_i^{(k)} \psi_i^{(k)} = \sum_{i=1}^{n_k} d_i^{(k)} \sum_{j=1}^{n_k} b_{(i,j)}^{(k)} g_j^{(k)} = \sum_{j=1}^{n_k} g_j^{(k)} \sum_{i=1}^{n_k} b_{(i,j)}^{(k)} d_i^{(k)} = \sum_{j=1}^{n_k} c_j^{(k)} g_j^{(k)}$$

this means that

$$(c_i^{(k)} : j = 1, 2 \dots n_k) = (B^{(k)})^{-1} (d_i^{(k)} : j = 1, 2 \dots n_k)$$

or

$$(d_i^{(k)} : j = 1, 2 \dots n_k) = (B^{(k)}) (c_i^{(k)} : j = 1, 2 \dots n_k).$$

If we project $x$ to its first, say $L$, coordinates in the lexicographical order of $(\Psi_j^{(k)} : k \in \mathbb{N}, j = 1, \dots k_n)$, for $N \in \mathbb{N}$ and $M \leq n_N$, so that $L = \sum_{k=1}^{N-1} k_n + M$, this projected vector equals to:

$$\sum_{k=1}^{N-1} \sum_{j=1}^{n_k} d_j^{(k)} \psi_j^{(k)} + \sum_{j=1}^{M} d_j^{(N)} \psi_j^{(N)} = \sum_{k=1}^{N-1} \sum_{j=2}^{n_k} c_j^{(k)} g_j^{(k)} + \sum_{j=1}^{M} d_j^{(N)} \psi_j^{(N)}.$$

Therefore we only need to show that there is a constant $C \geq 1$ so that for all $k$ and all $M \leq n_k$

$$\Big\| \sum_{j=1}^{M} d_j^{(k)} \psi_j^{(k)} \Big\| \leq C \Big\| \sum_{j=1}^{n_k} d_j^{(k)} \psi_j^{(k)} \Big\| \tag{1.15}$$

and that $(\psi_j^{(k)})$ is Besselian.

It follows from the assumption that the matrices $B^{(k)}$ are unitary and Proposition 1.3.4 that

$$\|x\| = \Big\| \sum_{k=1}^{\infty} x_k \Big\| \geq \frac{1}{C_B} \Big( \sum_{k=1}^{\infty} \sum_{j=1}^{n_k} |c_j^{(k)}|^2 \Big)^{1/2} = \frac{1}{C_B} \Big( \sum_{k=1}^{\infty} \sum_{j=1}^{n_k} |d_j^{(k)}|^2 \Big)^{1/2},$$

which proves that $(\psi_j^{(k)})$ is Besselian. Secondly we note that (1.11) yields

$$\begin{aligned}
\Big\| \sum_{i=1}^{M} d_i^{(k)} \psi_i^{(k)} \Big\| &= \Big\| \sum_{i=1}^{M} d_i^{(k)} \sum_{j=1}^{n_k} b_{(i,j)}^{(k)} g_j^{(k)} \Big\| \\
&\leq \sum_{i=1}^{M} |d_i^{(k)}| n_k^{-1/2} \|g_1^{(k)}\| + \Big\| \sum_{i=1}^{M} d_i^{(k)} \sum_{j=2}^{n_k} b_{(i,j)}^{(k)} g_j^{(k)} \Big\| \\
&\leq \Big( \sum_{i=1}^{M} |d_i^{(k)}|^2 \Big)^{1/2} + \Big\| \sum_{j=2}^{n_k} g_j^{(k)} \sum_{i=1}^{M} d_i^{(k)} b_{(i,j)}^{(k)} \Big\| \\
&\leq \Big( \sum_{i=1}^{M} |d_i^{(k)}|^2 \Big)^{1/2} + C_H \Big( \sum_{j=2}^{n_k} \Big| \sum_{i=1}^{M} d_i^{(k)} b_{(i,j)}^{(k)} \Big|^2 \Big)^{1/2} \\
&\leq \Big( \sum_{i=1}^{n_k} |d_i^{(k)}|^2 \Big)^{1/2} + C_H \Big( \sum_{i=1}^{n_k} |d_i^{(k)}|^2 \Big)^{1/2} \text{(By (1.10))}
\end{aligned}$$

Therefore applying (1.10) and then (1.8) it follows that

$$\begin{aligned}
\Big\| \sum_{i=1}^{M} d_i^{(k)} \psi_i^{(k)} \Big\| &\leq (1 + C_H) \Big( \sum_{i=1}^{n_k} |d_i^{(k)}|^2 \Big)^{1/2} \\
&= (1 + C_H) \Big( \sum_{i=1}^{n_k} |c_i^{(k)}|^2 \Big)^{1/2} \leq (1 + C_H) C_B \|x_k\|
\end{aligned}$$

which proves our claim. □

Our last step of proving Theorem 1.3.2 is the following

**Proposition 1.3.5.** *$(\psi_j^{(k)} : j = 1, 2 \dots n_k)$ is quasi-greedy.*

*Proof.* Let
$$x = \sum_{k=1}^{\infty} \sum_{i=1}^{n_k} d_i^{(k)} \psi_i^{(k)} \in X,$$
with $\|x\| = 1$ and suppose that the $m$-th greedy approximate is given by
$$G_m^T(x, \Psi) = \sum_{k \in J} \sum_{i \in I_k} d_i^{(k)} \psi_i^{(k)},$$
where $m = \sum_{k \in J} \# I_k$. We need to show that there is a constant $C \geq 1$ (of course not dependent on $x$ and $m$) so that

$$\|G_m^T(x, \Psi)\| \leq C \|x\| \tag{1.16}$$

We write $G_m^T(x, \Psi)$ as
$$G_m^T(x, \Psi) = \underbrace{\sum_{k \in J} \sum_{i \in I_k} d_i^{(k)} (\psi_i^{(k)} - b_{(i,1)}^{(k)} f_k)}_{\Sigma_1} + \underbrace{\sum_{k \in J} \sum_{i \in I_k} d_i^{(k)} b_{(i,1)}^{(k)} f_k}_{\Sigma_2}.$$

(recall that $g_1^{(k)} = f_k$). From the definition of the $\psi_j^{(k)}$ we get that
$$\Sigma_1 = \sum_{k \in J} \sum_{i \in I_k} d_i^{(k)} \Big( \sum_{j=2}^{n_k} b_{(i,j)}^{(k)} g_j^{(k)} \Big) = \sum_{k \in J} \sum_{j=2}^{n_k} g_j^{(k)} \Big( \sum_{i \in I_k} d_i^{(k)} b_{(i,j)}^{(k)} \Big),$$
which yields by the choice of the $g_j^{(k)}$, properties (1.9), and (1.10) that

$$\begin{aligned} \big\|\Sigma_1\big\| &\leq C_H \left( \sum_{k \in J} \sum_{j=2}^{n_k} \Big| \sum_{i \in I_k} d_i^{(k)} b_{(i,j)}^{(k)} \Big|^2 \right)^{1/2} \\ &= C_H \left( \sum_{k \in J} \big\| \big[ B^{(k)} \big]^{-1} \circ (d_i^{(k)} : i \in I_k) \big\|_2^2 \right)^{1/2} \\ &= C_H \left( \sum_{k \in J} \big\| (d_i^{(k)} : i \in I_k) \big\|_2^2 \right)^{1/2} \quad \text{(By (1.10))} \\ &\leq C_H C_B \|x\| \text{ (By Proposition (1.3.4)).} \end{aligned} \tag{1.17}$$

In order to estimate $\Sigma_2$ we split $I_k$, $k \in \mathbb{N}$ into the following subsets:
$$I_k^{(1)} = \big\{ i \in I_k : |d_i^{(k)}| \leq n_k^{-1} \} \big\}$$
$$I_k^{(2)} = \big\{ i \in I_k : |d_i^{(k)}| \geq n_k^{-1/2} \big\}$$

$$I_k^{(3)} = \{ i \in I_k : n_k^{-1} < |d_i^{(k)}| < n_k^{-1/2} \}$$

and let

$$\Sigma_2^{(s)} = \sum_{k \in J} \sum_{i \in I_k^{(s)}} d_i^{(k)} b_{(i,1)}^{(k)} f_k \text{ for } s = 1, 2, 3.$$

From the definition of $I_k^{(1)}$ and 1.11 it follows that

$$\Big| \sum_{i \in I_k^{(1)}} d_i^{(k)} b_{(i,1)}^{(k)} \Big| \leq n_k^{-1/2}.$$

and thus

$$\big\| \Sigma_2^{(1)} \big\| \leq \sum_{k \in J} n_k^{-1/2} \leq 1. \tag{1.18}$$

In order to estimate $\big\| \Sigma_2^{(2)} \big\|$ we we first note that the definition of $I_k^{(2)}$ yields that

$$(\# I_k^{(2)}) n_k^{-1} \leq \sum_{i \in I_k^{(2)}} |d_i^{(k)}|^2 \leq \sum_{i=1}^{n_k} |d_i^{(k)}|^2,$$

and, thus,

$$\begin{aligned} \big\| \Sigma_2^{(2)} \big\| &= \Big\| \sum_{k \in J} \sum_{i \in I_k^{(2)}} d_i^{(k)} b_{(i,1)}^{(k)} f_k \Big\| \\ &\leq \sum_{k \in J} n_k^{-1/2} \sum_{i \in I_k^{(2)}} |d_i^{(k)}| \text{ (By (1.11))} \\ &\leq \sum_{k \in J} n_k^{-1/2} (\# I_k^{(2)})^{1/2} \Big( \sum_{i \in I_k^{(2)}} |d_i^{(k)}|^2 \Big)^{1/2} \text{ (By Hölder's inequality)} \\ &\leq \sum_{k \in J} \sum_{i \in I_k^{(2)}} |d_i^{(k)}|^2 \\ &\leq C_B^2 \|x\|^2 = C_B^2 \text{ (By Proposition (1.3.4)).} \end{aligned} \tag{1.19}$$

Finally we have to estimate $\big\| \Sigma_2^{(3)} \big\|$. Before that let us make some observations:

We first note that in the estimation of $\|\Sigma_1\|$ we did not use specific properties of the sets $I_k$. Replacing in the estimation of $\|\Sigma_1\|$ the sets $I_k$ by any set $I'_k \subset \{1, 2 \ldots n_k\}$ in (1.17) and $J$ by any set $J' \subset \mathbb{N}$ we obtain

$$\Big\| \sum_{k \in J'} \sum_{i \in I'_k} d_i^{(k)} \Big( \sum_{j=2}^{n_k} b_{(i,j)}^{(k)} g_j^{(k)} \Big) \Big\| \leq C_H C_B \|x\| \tag{1.20}$$

Taking $I'_k$ to be all of $\{1,2\dots n_k\}$ and $J'=[1,K]$ for some $K\in\mathbb{N}$ we deduce from Proposition 1.3.5 that

$$\begin{aligned}(1.21)\qquad &\Big\|\sum_{k=1}^{K}\sum_{i=1}^{n_k} d_i^{(k)} b_{(i,1)}^{(k)} f_k\Big\| \\ &\leq \Big\|\sum_{k=1}^{K}\sum_{i=1}^{n_k} d_i^{(k)}\psi_i^{(k)}\Big\| + \Big\|\sum_{k=1}^{K}\sum_{i=1}^{n_k} d_i^{(k)}(\psi_i^{(k)} - b_{(i,1)}^{(k)} f_k)\Big\| \\ &\leq C_\Psi\|x\| + C_H C_B\|x\| = C_\Psi + C_H C_B,\end{aligned}$$

where $C_\Psi$ denotes the basis constant of $(\psi_j^{(k)}; k\in\mathbb{N}, j=1,2\dots n_k)$. Secondly we note that in the estimation of $\|\Sigma_k^{(1)}\|$ we could replace the set $J$ by any subset $J'\subset\mathbb{N}$ and $I_k^{(1)}$ by any subset

$$I'_k \subset I_k^{(1)} = \big\{i\leq n_k : |d_i^{(k)}|\leq n_k^{-1}\}\big\},$$

to obtain

$$(1.22)\qquad \Big\|\sum_{k\in J'}\sum_{i\in I'_k} d_i^{(k)} b_{(i,1)}^{(k)} f_k\Big\| \leq 1.$$

Thirdly we note that in the estimation of $\big\|\Sigma_2^{(2)}\big\|$ in (1.19) we could have also replaced $J$ by any subset of $\mathbb{N}$, and for $k\in N$ the set $I_k^{(2)}$ by any subset

$$I'_k \subset I_k^{(2)} = \big\{i\in I_k : |d_i^{(k)}|\geq n_k^{-1/2}\big\}$$

to obtain

$$(1.23)\qquad \Big\|\sum_{k\in J'}\sum_{i\in I'_k} d_i^{(k)} b_{(i,1)}^{(k)} f_k\Big\| \leq C_B^2.$$

In order to estimate the $\|\Sigma_2^{(3)}\|$ we define

$$K = \max\{k\in J : I_k^{(3)} \neq \emptyset\},$$

which means that for some $i\in I_k^{(3)}$ it follows that $|d_i^{(k)}| < n_K^{-1/2}$ and note that for any $k\in[1,K-1]$ either $k\in J$ or (here we use the first time that we are dealing with the threshold algorithm) $k\notin J$, which implies

$$(1.24)\qquad |d_i^{(k)}| < n_K^{-1/2} \leq n_k^{-1} \text{ for all } i\in\{1,2,\dots n_k\}$$

(here we are using that $n_{k+1}=n_k^2$) and thus for such a $k$ the sets $I_k^{(3)}$ and $I_k^{(2)}$ are empty.

We compute now

$$\begin{aligned}\Sigma_2^{(3)} &= \sum_{i\in I_K^{(3)}} d_i^{(K)} b_{(i,1)}^{(K)} f_K + \sum_{k\in J, j<K} \sum_{i\in I_k^{(3)}} d_i^{(k)} b_{(i,1)}^{(k)} f_k \\ &= \sum_{i\in I_K^{(3)}} d_i^{(K)} b_{(i,1)}^{(K)} f_K + \sum_{k=1}^{K-1} \sum_{i=1}^{n_j} d_i^{(k)} b_{(i,1)}^{(k)} f_k \\ &\qquad - \sum_{k=1}^{K-1} \sum_{i\in I_k^{(1)}} d_i^{(k)} b_{(i,1)}^{(k)} f_k - \sum_{k\in J, k<K} \sum_{i\in I_k^{(2)}} d_i^{(k)} b_{(i,1)}^{(k)} f_k\end{aligned}$$

The first term we estimate, using Hölder's inequality:

$$\Big\| \sum_{i\in I_K^{(3)}} d_i^{(K)} b_{(i,1)}^{(K)} f_K \Big\| \le n_K^{-1/2} \sum_{j=1}^{n_K} \big|d_i^{(K)}\big| \le n_K^{-1/2} n_K^{1/2} \Big( \sum_{j=1}^{N_K} \big|d_i^{(K)}\big|^2 \Big)^{1/2} \le C_B.$$

It follows therefore from (1.21), (1.22) and (1.23)

$$\big\|\Sigma_2^{(3)}\big\| \le C_B + C_\Psi + C_H C_B + 1 + C_B^2$$

which implies our claim letting $C = C_B + C_\Psi + C_H C_B + 1 + C_B^2$. □

**Corollary 1.3.6.** *Apply Theorem 1.3.2 to the trigonometrical polynomials* $(t_n) = (e^{-i2\pi n(\cdot)} : n \in \mathbb{Z})$ *which are a basis of* $L_p[0,1]$ *and satisfy by Example 1.3.1 the assumptions if* $p > 2$*. This leads to a quasi greedy basis* $(\Psi_n : n \in \mathbb{N})$ *of* $L_p[0,1]$*.*

*Secondly note since* $(t_n)$ *is absolutely bounded by* $1$ *(in* $L_\infty[0,1]$*), and since the matrices* $B^{(k)}$*, which where used in the construction of the basis* $(\Psi_n)$ *are uniformly bounded as linear operators on* $\ell_\infty^{n_k}$*, it follows that also* $(\Psi_n : n \in \mathbb{N})$ *is bounded* $L_\infty[0,1]$*.*

*This implies by Corollary 6.2.7 that* $(\Psi_n)$ *cannot be unconditional.*

# Chapter 2

# Greedy Algorithms In Hilbert Space

## 2.1 Introduction

We will now replace in our greedy algorithms, bases by more general and possibly redundant systems.

Let $H$ (always) be a separable and real Hilbert space. Recall that $\mathcal{D} \subset S_H$ is a *dictionary of* $X$ if $\text{span}(\mathcal{D})$ is dense and $x \in \mathcal{D}$ implies that $-x \in \mathcal{D}$.

An *n-term approximation algorithm* is a map

$$\mathbf{G} : H \to \text{span}(\mathcal{D})^{\mathbb{N}}, \ x \mapsto \mathbf{G}(x) = (G_n(x)),$$

with the property that for $n \in \mathbb{N}$ and $x \in H$, there is a set $\Lambda_n \subset \mathcal{D}$ of cardinality at most $n$ so that $G_n(x) \in \text{span}(\Lambda_n)$, $G_n(x)$ is then called an *n-term approximation of* $x$.

Perhaps the first example was considered by Schmidt [Schm]:

**Example 2.1.1.** [Schm] Let $f \in L_2\big([0,1]^2\big)$, i.e. $f$ is a square integrable function in two variables. By the Theorem of Arcela and Ascoli we know that the set

$$\mathcal{D} = \Big\{ \sum_{j=1}^{n} u_j \otimes v_j : n \in \mathbb{N}, u_i, v_j \in C[0,1] \Big\}$$

is dense in $C\big([0,1]^2\big)$. Here we denote for two functions $f, g : [0,1] \to \mathbb{K}$

$$f \otimes g : [0,1]^2 \to \mathbb{K}, \quad (x,y) \mapsto f(x)g(y).$$

Since $C\big([0,1]^2\big)$ is dense in $L_2\big([0,1]^2\big)$ it follows that

$$\mathcal{D} = \Big\{ \sum_{j=1}^{n} u_j \otimes v_j : n \in \mathbb{N}, u_i, v_j \in L_2[0,1] \Big\}$$

is dense in $L_2\big([0,1]^2\big)$.

The question is now, how to find a good approximate to $f$ from $\mathcal{D}$. E. Schmidt considered the following procedure and showed that it worked:

Let $f \in L_2\big([0,1]^2\big)$ and define $f_0 = f$.

Then choose $u_1, v_1 \in L_2[0,1]$ so that

$$\|f_0 - u_1 \otimes v_1\|_2 = \inf\big\{\|f_0 - u \otimes v\|_2 : u, v \in L_2[01]\big\}.$$

Since this infimum might be hard to achieve he also considered a weaker condition, and fixed some *weakening factor* $t \in (0,1)$ and chose $u_1, v_1 \in L_2[0,1]$ so that

$$\|f_0 - u_1 \otimes v_1\|_2 \le \frac{1}{t} \inf\big\{\|f_0 - u \otimes v\|_2 : u, v \in L_2[01]\big\}.$$

Then he let

$$f_1 = f_0 - u_1 \otimes v_1.$$

After $n$ steps he obtained $u_1, v_1, u_2, v_2, \ldots u_n, v_m \in L_2[0,1]$, and let

$$f_n = f - \sum_{j=1}^{n} u_j \otimes v_j,$$

and chose $u_{n+1}$ and $v_{n+1}$ in $L_2[0,1]$ so that

$$\|f_n - u_{n+1} \otimes v_{n+1}\| = \|f_0 - \sum_{j=1}^{n+1} u_j \otimes v_j\|_2 \le \frac{1}{t} \inf\big\{\|f_n - u \otimes v\|_2 : u, v \in L_2[01]\big\}.$$

Finally he proved that $f_n$ converges in $L_2\big([0,1]^2\big)$ to 0 and thus $G_n(f) = \sum_{j=1}^{n+1} u_j \otimes v_j$ converges to $f$.

He asked whether there is some general principle behind, and how and whether this generalizes..

(PGA) The *Pure Greedy Algorithm.*

For $x \in H$ we define $G_n = G_n(x)$, for each $n \in \mathbb{N}_0$, by induction. $G_0 = 0$ and assuming that $G_0, G_1 \ldots G_{n-1}$, have been defined for some $n \in \mathbb{N}$ we proceed as follows:

1) Choose $z_n \in \mathcal{D}$ and $a_n \in \mathbb{R}$ so that

$$\|x - G_{n-1} - a_n z_n\| = \inf_{z \in \mathcal{D}, a \in \mathbb{R}} \|x - G_{n-1} - az\|.$$

2) Put $G_n = G_{n-1} + a_n z_n$.

Note that for any $x \in H$ it follows that

(2.1) $$\inf_{z \in \mathcal{D}, a \in \mathbb{R}} \|x - az\|^2$$

$$= \inf_{z\in\mathcal{D},a\in\mathbb{R}} \left[\|x\|^2 - 2a\langle x,z\rangle + a^2\|z\|^2\right]$$
$$= \inf_{z\in\mathcal{D}} \left[\|x\|^2 - \langle x,z\rangle^2\right]$$
$$[a\mapsto \|x\|^2 - 2a\langle x,z\rangle + a^2\|z\|^2 \text{ is minimal for } a=\langle x,z\rangle]$$
$$= \|x\|^2 - \sup_{z\in\mathcal{D}}\langle x,z\rangle^2.$$

So condition (1) in (PGA) can be replaced by the following condition (1')

1') Choose $z_n \in \mathcal{D}$ so that

$$\langle x - G_{n-1}, z_n\rangle = \sup_{z\in\mathcal{D}}\langle x - G_{n-1}, z\rangle$$

and (2) by

2') Put $G_n = G_{n-1} + \langle x - G_{n-1}, z_n\rangle z_n$.

As already noted in Example 2.1.1, the "sup" in (1') (PGA), respectively the "inf" in (1) might not be attained or might be hard to attain. In this case we might consider the following modification.

(WPGA) The *Weak Pure Greedy Algorithm.*
We are given a sequence $\tau = (t_n) \subset (0,1)$. For $x \in X$ we define $G_n = G_n(x)$, for each $n \in \mathbb{N}_0$, by induction.
$G_0 = 0$ and assuming that $G_0, G_1 \ldots G_{n-1}$, have been defined for some $n \in \mathbb{N}$ we proceed as follows:

1) Choose $z_n \in \mathcal{D}$, so that

$$\langle x - G_{n-1}, z_n\rangle \geq t_n \sup_{z\in\mathcal{D}}\langle x - G_{n-1}, z\rangle$$

2) Put $G_n = G_{n-1} + \langle x - G_{n-1}, z_n\rangle z_n$. For WPGA we call the sequence $(t_n)$ *the weakness factors.*

A possibly faster (but computational more laborious) algorithm is the following *Orthogonal Greedy Algorithm.*

(OGA) The *Orthogonal Greedy Algorithm.*
For $x \in H$ we define $G^o_n = G^o_n(x)$, for each $n \in \mathbb{N}_0$, by induction.
$G^o_0 = 0$ and assuming that $G^o_0, G^o_1 \ldots G^o_{n-1}$, and vectors $z_1 \ldots z_{n-1}$ have been defined for some $n \in \mathbb{N}$ we proceed as follows:

1) Choose $z_n \in \mathcal{D}$ so that

$$\langle x - G^o_{n-1}, z_n\rangle = \sup_{z\in\mathcal{D}}\langle x - G^o_{n-1}, z\rangle$$

2) Define $Z_n = \text{span}(z_1, z_2 \ldots z_n))$ and let $G^o_{n-1}$ be the best approximation of $x$ to $Z_n$, i.e.

$$\|G^o_{n-1} - x\| = \inf\left\{\|z - x\| : z \in Z_n\right\},$$

which means that $G_{n-1} = P_{Z_n}(x)$, where $P_{Z_n}$ denotes the orthonormal projection of $H$ onto $Z_n$.

(GAR) The *Greedy Algorithm with free Relaxation.*

For $x \in H$ we define $G_n^r = G_n^r(x)$, for each $n \in \mathbb{N}_0$, by induction. $G_0^r = 0$ and assuming that $G_0^r, G_1^r \dots G_{n-1}^r$, have been defined for some $n \in \mathbb{N}$ we proceed as follows:

1) Choose $z_n \in \mathcal{D}$ so that

$$\langle x - G_{n-1}^r, z_n \rangle = \sup_{z \in \mathcal{D}} \langle x - G_{n-1}^r, z \rangle$$

2) Put $G_n^r = a_n G_{n-1}^r + b_n z_n$, where $G_n^r$ is best approximation of $x$ by an element of the two dimensional space $\text{span}(G_{n-1}^r, z_n)$.

(GAFR) The *Greedy Algorithm with fixed Relaxation.*

Let $c > 0$. For $x \in H$ we define $G_n^f = G_n^f(x)$, for each $n \in \mathbb{N}_0$, by induction.
$G_0^f = 0$ and assuming that $G_0^f, G_1^f \dots G_{n-1}^f$, have been defined for some $n \in \mathbb{N}$ we proceed as follows:

1) Choose $z_n \in \mathcal{D}$ so that

$$\langle x - G_{n-1}^f, z_n \rangle = \sup_{z \in \mathcal{D}} \langle x - G_{n-1}^f, z \rangle$$

2) Put $G_n^f = c\Big(1 - \frac{1}{n}\Big) G_{n-1}^f + \frac{c}{n} z_n$,

Similar to the weak purely greedy algorithm there are also weak versions of the orthogonal greedy algorithm and the pure greedy algorithm with relaxation and We denote them by WOGA, WGAR and WGAFR.

## 2.2 Convergence

**Proposition 2.2.1.** *Assume that we consider the WPGA, WOGA or WGAR and assume for the weakness factors* $(t_n)$ *that*

$$\sum_{k \in \mathbb{N}} t_k^2 = \infty. \tag{2.2}$$

*For* $x$ *we let* $x_n = x - G_n(x)$, $x_n = x - G_n^o(x)$ *or* $x_n = x - G^r(x)$, *respectively.*

*If the sequence* $(x_n)$ *converges it converges to* $0$.

*Proof.* Assume that $x_n$ converges to some $u \in H$ and $u \neq 0$. Then, since $\mathcal{D}$ is a dictionary, there is a $d \in \mathcal{D}$ so that $\delta = \langle d, u\rangle > 0$ and thus we find a large enough $N \in \mathbb{N}$ so that $\langle d, x_n\rangle \geq \delta/2$, for all $n > N$

In the case that we consider WPGA we obtain for $n \geq N$

$$\|x_{n+1}\|^2 = \left\|x_n - \langle z_{n+1}, x_n\rangle z_{n+1}\right\|^2 = \|x_n\|^2 - \langle z_{n+1}, x_n\rangle^2 \leq \|x_n\|^2 - t_{n+1}^2\delta^2/4.$$

and thus for $k = 1, 2, 3\dots$

$$\|x_N\|^2 - \|x_{N+k}\|^2 = \sum_{j=N}^{N+k-1} \|x_j\|^2 - \|x_{j+1}\|^2 \geq \sum_{j=N}^{N+k} t_{j+1}^2\delta^2/4 \to_{N\to\infty} \infty.$$

But this is a contradiction.

In the case of the WOGA we similarly have for $n \geq N$

$$\begin{aligned}\|x_{n+1}\|^2 &= \min\Big\{\Big\|x - \sum_{j=1}^{n+1} a_j z_j\Big\| : a_1, a_2, \dots a_{n+1} \in \mathbb{R}\Big\}\\ &\leq \left\|x_n - \langle z_{n+1}, x_n\rangle z_{n+1}\right\|^2\\ &= \|x_n\|^2 - \langle z_{n+1}, x_n\rangle^2 \leq t_{n+1}^2\delta^2/4\end{aligned}$$

and we obtain a contradiction as in the WPGA case. Similarly in the case we consider the WGAR we estimate:

$$\begin{aligned}\|x_{n+1}\|^2 &= \min\Big\{\Big\|x - aG_n^r(x) - bz_{n+1}\Big\| : a, b \in \mathbb{R}\Big\}\\ &\leq \left\|x - G_n^r(x) - \langle z_{n+1}, x_n\rangle z_{n+1}\right\|^2\\ &= \|x_n\|^2 - \langle z_{n+1}, x_n\rangle^2 \leq \|x_n\|^2 - t_{n+1}^2\delta^2/4.\end{aligned}$$

□

**Theorem 2.2.2.** *Assume that condition* (2.2) *of Proposition 2.2.1 holds. Then* $(G_n^o(x) : n \in \mathbb{N})$ *(as defined in WOGA) converges for all* $x \in H$ *to* $x$.

*Proof.* Let $x \in H$. For $n \in \mathbb{N}$ let $Z_n$ be the space defined in WOGA. $G_n^o(x) = P_{Z_n}(x)$. Since $Z_1 \subset Z_2 \subset Z_3 \dots$ it follows that $G_n^o(x)$ converges to $P_Z(x)$, where $Z = \overline{\bigcup_{n\in\mathbb{N}} Z_n}$ . Thus the claim follows from Proposition 2.2.1 □

**Theorem 2.2.3.** *Assume that the sequence* $(t_k) \subset (0, 1)$ *satisfies*

$$\sum_{k\in\mathbb{N}} \frac{t_k}{k} = \infty. \tag{2.3}$$

*For* $x \in X$ *consider the WPGA* $(G_n(x))$ *with weakness factors* $(t_n)$.

*Then* $(G_n(x))$ *converges.*

**Remark.** Since by Hölder's inequality

$$\sum_{k=1}^{\infty} \frac{t_k}{k} \leq \Big( \sum_{k=1}^{\infty} t_k^2 \Big)^{1/2} \Big( \sum_{k=1}^{\infty} \frac{1}{k} \Big)^{1/2},$$

condition 2.3 implies that $\sum_{k\in\mathbb{N}} t_k^2 = \infty$.

We will need the following Lemma first

**Lemma 2.2.4.** *Assume* $y = (y_j) \in \ell_2$ *and* $(t_k) \subset (0,1)$ *satisfies* (2.3) . *Then*

$$\liminf_{n\to\infty} \frac{|y_n|}{t_n} \sum_{j=1}^{n} |y_j| = 0.$$

*Proof.* (an alternate, and shorter proof due to Sheng Zhang will be given below) We will prove the following claim:
**Claim.** If $f \in L_2[0,\infty]$ and we define

$$F(x) = \int_0^x |f(t)|\, dt$$

then

$$\int_0^\infty \frac{F^2(x)}{x^2} dx \leq 4 \int_0^\infty f^2(x)\, dx. \tag{2.4}$$

If we apply the claim to the function $f(\cdot) = \sum_{j=1}^{\infty} 1_{(j-1,j]} |y_j|$, it follows that

$$\begin{aligned} \sum_{n=1}^{\infty} \Big[ \frac{1}{n} \sum_{j=1}^{n} |y_j| \Big]^2 &\leq \sum_{n=1}^{\infty} \Big[ \frac{1}{n} \sum_{j=1}^{n-1} |y_j| \Big]^2 + \sum_{n=1}^{\infty} |y_j|^2 \\ &\leq \int_0^\infty \left[ \frac{1}{x} \int_0^x f(t)\, dt \right]^2 dx + \sum_{n=1}^{\infty} |y_j|^2 \\ &\leq 4 \int_0^\infty f^2(t)\, dt + \sum_{n=1}^{\infty} |y_j|^2 = 5 \sum_{n=1}^{\infty} |y_j|^2 \end{aligned}$$

It follows therefore from the Cauchy Schwarz inequality that

$$\sum_{n=1, t_n\neq 0}^{\infty} \frac{t_n}{n} \frac{|y_n|}{t_n} \sum_{j=1}^{n} |y_j| \leq \sum_{n=1}^{\infty} |y_n| \frac{1}{n} \sum_{j=1}^{n} |y_j| \leq \left[ \sum_{n=1}^{\infty} |y_j|^2 \right]^{1/2} \left[ \sum_{n=1}^{\infty} \Big[ \frac{1}{n} \sum_{j=1}^{n} |y_j| \Big]^2 \right]^{1/2} < \infty$$

since

$$\sum_{n=1}^{\infty} \frac{t_n}{n} = \infty$$

it follows that

$$\liminf_{n\to\infty} \frac{|y_n|}{t_n} \sum_{j=1}^{n} |y_j| = 0.$$

In order to prove the claim we can assume that $f(x)$ is a positive function, we note first that by Hölder's inequality,

$$F(x) = \int_0^x f(t)\,dt \le x^{1/2} \int_0^x f^2(t)\,dt,$$

and thus

$$\frac{F(x)}{x^{1/2}} \le \int_0^x f^2(t)\,dt \to_{x\to 0} 0 \tag{2.5}$$

For a fixed $x_0 > 0$ we also deduce from Hölder's inequality for $x > x_0$ that

$$F(x) - F(x_0) = \int_{x_0}^x f(t)\,dt \le (x - x_0)^{1/2} \int_{x_0}^x f^2(t)\,dt \le x^{1/2} \int_{x_0}^\infty f^2(t)\,dt,$$

and thus

$$\frac{F(x)}{x^{1/2}} \le \frac{F(x_0)}{x^{1/2}} + \int_{x_0}^\infty f^2(t)\,dt.$$

By choosing for a given $\varepsilon > 0$ $x_0$ far enough out so that $\int_{x_0}^\infty f^2(t)\,dt < \varepsilon/2$ and then $x_1 > x_0$ so that $x_1^{-1/2} F(x_0) < \varepsilon/2$, it follows that

$$\frac{F(x)}{x^{1/2}} < \varepsilon \text{ whenever } x > x_1,$$

and thus

$$\frac{F(x)}{x^{1/2}} \le \int_0^x f^2(t)\,dt \to_{x\to\infty} 0. \tag{2.6}$$

Using integration by parts, it follows for any $0 < a < b < \infty$ that

$$\begin{aligned}\int_a^b \frac{F^2(x)}{x^2}\,dx &= -\frac{F^2(x)}{x}\Big|_{x=a}^{b} + 2\int_a^b F(x)f(x)x^{-1}\,dx \\ &\le \left[\frac{F(a)}{a^{1/2}}\right]^2 + \left[\frac{F(b)}{b^{1/2}}\right]^2 + 2\left(\int_a^b \frac{F^2(x)}{x^2}\right)^{1/2}\left(\int_a^b f^2(x)\,dx\right)^{1/2}\end{aligned}$$

[By Hölder's inequality]

and thus, in case that $a$ is chosen small enough and $b$ large enough so that $F(x)$ does not a.e. vanish on $[a,b]$, we have

$$\left(\int_a^b \frac{F^2(x)}{x^2}\right)^{1/2} \le \left[\left[\frac{F(a)}{a^{1/2}}\right]^2 + \left[\frac{F(b)}{b^{1/2}}\right]^2\right]\left(\int_a^b \frac{F^2(x)}{x^2}\right)^{-1/2} + 2\left(\int_a^b f^2(x)\,dx\right)^{1/2}.$$

Our claim follows now by letting $a \to 0$ and $b \to \infty$

□

*Proof by Sheng Zhang.* Suppose, to the contrary that

$$\delta = \liminf_{n\to\infty} \frac{|y_n|}{t_n} \sum_{j=1}^n |y+j| > 0,$$

and, thus for some $n_0 \in \mathbb{N}$

$$\frac{|y_n|}{t_n} \sum_{j=1}^n |y+j| > \delta/2 \text{ whenever } n \ge n_0.$$

For $n \ge n_0$, we deduce fro Hölders's inequality

$$\frac{\delta}{2} < \frac{1}{t_n} \sum_{j=1}^n |y_n||y_j| \le \frac{1}{t_n}\Big(\sum_{j=1}^n |y_j|^2\Big) n|y_n|^2,$$

and thus

$$\frac{|y_n|}{\frac{t_n}{n}} \ge \frac{\delta}{2} \frac{1}{\sum_{j=1}^n |y_j|^2},$$

which yields

$$\liminf_{n\to\infty} \frac{|y_n|}{\frac{t_n}{n}} \ge \frac{\delta}{2} \frac{1}{\sum_{j=1}^\infty y_j^2} =: \varepsilon$$

Thus there is an $n_> \ge n_0$, so that for all $n \ge n_1$, $|y_n|^2 \ge \varepsilon t_n/2n$. But this contradicts the assumption that $y = (y_n) \in \ell_2$ and $\sum_{n=1}^\infty t_n/n = \infty$. □

*Proof of Theorem 2.2.3.* Let $x \in H$ and put for $n \in \mathbb{N}$, $G_n = G_n(x)$ with

$$G_n(x) = \sum_{j=1}^n \langle x - G_{j-1}, z_j\rangle z_j,$$

where $z_n \in \mathcal{D}$ satisfies

$$\langle z_n, x - G_{n-1}\rangle = \langle z_n, x_{n-1}\rangle \ge t_n \sup_{z\in\mathcal{D}} \langle z, x_{n-1}\rangle. \tag{2.7}$$

Define

$$x_n = x - G_n(x) = x - \sum_{j=1}^{n} \langle z_j, x - G_{j-1}\rangle z_j = x_{n-1} - \langle z_n, x - G_{n-1}\rangle z_n, \tag{2.8}$$

By induction we show that for every $n \in \mathbb{N}$

$$\|x_n\|^2 = \|x\|^2 - \sum_{j=1}^{n} \langle z_j, x_{j-1}\rangle^2. \tag{2.9}$$

Indeed, for $n = 1$ the claim is clear and assuming that (2.9) is true for $n \in \mathbb{N}$ we compute

$$\begin{aligned}\|x_{n+1}\|^2 &= \|x_n - \langle x_n, z_{n+1}\rangle z_{n+1}\|^2 \\ &= \|x_n\|^2 - \langle x_n, z_{n+1}\rangle \|z_{n+1}\|^2 = \|x\|^2 - \sum_{j=1}^{n+1} \langle z_j, x_{j-1}\rangle^2.\end{aligned}$$

It follows therefore from (2.9) that

$$\sum_{j=1}^{\infty} \langle z_j, x_{j-1}\rangle^2 \le \|x\|^2 \tag{2.10}$$

For $m < n$ we compute

$$\|x_n - x_m\|^2 = \|x_m\|^2 - \|x_n\|^2 - 2\langle x_m - x_n, x_n\rangle, \tag{2.11}$$

and

$$\begin{aligned}\big|\langle x_m - x_n, x_n\rangle\big| &= \Big| \sum_{j=m+1}^{n} \langle x_{j-1} - x_j, x_n\rangle \Big| \\ &\le \sum_{j=m+1}^{n} \big|\langle x_{j-1} - x_j, x_n\rangle\big| \\ &= \sum_{j=m+1}^{n} \big|\langle z_j, x_n\rangle\big| \cdot \big|\langle z_j, x_{j-1}\rangle\big| \\ &\le \frac{\big|\langle z_{n+1}, x_n\rangle\big|}{t_{n+1}} \sum_{j=m+1}^{n} \big|\langle z_j, x_{j-1}\rangle\big| \\ &\Big[\big|\langle z_j, x_n\rangle\big| \le \max_{d\in\mathcal{D}} \big|\langle d, x_n\rangle\big| \le t_{n+1}^{-1}\big|\langle z_{n+1}, x_n\rangle\big|\Big] \\ &\le \frac{\big|\langle z_{n+1}, x_n\rangle\big|}{t_{n+1}} \sum_{j=1}^{n+1} \big|\langle z_j, x_{j-1}\rangle\big|.\end{aligned}$$

We can therefore apply Lemma 2.2.4 to $(t_n)$ and $y_n = \big|\langle z_{n+1}, x_n\rangle\big|$, for $n\in\mathbb{N}$, and deduce that

$$\liminf_{n\to\infty} \max_{m<n} \big|\langle x_m - x_n, x_n\rangle\big| = 0.$$

Together with the fact that $\|x_n\|$ is decreasing and (2.11) this implies that there is subsequence $(x_{n_k})$ which converges to some $x \in H$. We claim that the whole sequence $(x_n)$ converges to that $x$, which, together with Proposition 2.2.1, would finish the proof. Note that for any $n\in\mathbb{N}$ and any $k\in\mathbb{N}$ so that $n_k > n$ we have

$$\begin{aligned}\|x_n - x\| &\le \|x_n - x_{n_k}\| + \|x_{n_k} - x\| \\ &= \big(\|x_n\|^2 - \|x_{n_k}\|^2 - 2\langle x_n - x_{n_k}, x_{n_k}\rangle\big)^{1/2} + \|x_{n_k} - x\| \\ &\le \big(\|x_n\|^2 - \|x_{n_k}\|^2\big)^{1/2} + 2\max_{m\le n_k}\big|\langle x_m - x_{n_k}, x_{n_k}\rangle\big|^{1/2} + \|x_{n_k} - x\|.\end{aligned}$$

So, given $\varepsilon > 0$ we can choose $n_0$ large enough so that $\big(\|x_n\|^2 - \|x_{n_k}\|^2\big)^{1/2} < \varepsilon/3$, for all $n \ge n_0$ and $k$ with $n_k > n$. Then we choose $k_0$ so that for all $k > k_0$, $2\max_{m\le n_k}\big|\langle x_m - x_{n_k}, x_{n_k}\rangle\big|^{1/2} < \varepsilon/3$ and $\|x_{n_k} - x\| < \varepsilon/3$. For any $n \ge n_0$, we can therefore choose $k \ge k_0$ so that also $n_k > n$, and from above inequalities we deduce that $\|x_n - x\| < \varepsilon$. □

The next Theorem proves that at least among the decreasing weakness factors $\tau$ the condition 2.3 is optimal in order to imply convergence of the WPGA.

**Theorem 2.2.5.** *In the class of monotone decreasing sequences $\tau = (t_k)$, the condition* (2.3) *is necessary for the WPGA to converge.*

*In other words, if $(t_n)$ is a decreasing sequence for which*

$$\sum_{n\in\mathbb{N}}^{\infty} \frac{t_n}{n} < \infty \tag{2.12}$$

*then there is a dictionary $\mathcal{D}$ of $H$, an $x \in H$ and sequences $(G_n) \subset H$ and $(z_n) \subset \mathcal{D}$ , with $G_0 = 0$ so that for $x_n = x - G_n$ the following is satisfied:*

$$x_n = x_{n-1} - \langle x_{n-1}, z_n\rangle z_n \tag{2.13}$$

$$\langle x_{n-1}, z_n\rangle \ge t_n \max_{z\in\mathcal{D}'}\langle x_{n-1}, z\rangle, \tag{2.14}$$

*but so that $x_n$ does not converge to* 0.

We will need the following notation:

**Definition 2.2.6.** Assume $\mathcal{D}' \subset S_H$ has the property that $z \in \mathcal{D}'$ implies that $-z \in \mathcal{D}'$ and assume that $\tau = (t_n)_{n=1}^N \subset (0,1]$, with $N \in \mathbb{N}\cup\{\infty\}$ isa finite sequence of positive numbers. A pair of sequences $(x_n)_{n=0}^N \subset H$ and $(z_n)_{n=0}^N \subset \mathcal{D}'$

and are called a *pair of WPGA-sequences with weakness factor* $\tau$ *and dictionary* $\mathcal{D}$ if $x_0 \in \text{span}(\mathcal{D}')$ and for all $n = 1, 2 \ldots N$

$$x_n = x_{n-1} - \langle x_{n-1}, z_n z_n \rangle \tag{2.15}$$

$$\langle x_{n-1}, z_n \rangle \geq t_n \max_{z \in \mathcal{D}'} \langle x_{n-1}, z \rangle. \tag{2.16}$$

**Remark.** To given sequence $\tau = (t_n)_{n=1}^{\infty} \subset (0, 1]$, satisfying (2.12) we will choose elements of a dictionary $\mathcal{D}$ as well as the elements $x_n$ and $z_n$ of a pair of WPGA-sequences with weakness factor $\tau$ and dictionary $\mathcal{D}$ recursively.

To achieve that we will choose inductively elements $x_n$, $n \geq 0$ and $z_n$ , $n \geq 1$, so that for all $n \in \mathbb{N}$

$$x_n = x_{n-1} - \langle x_{n-1}, \rangle z_n \tag{2.17}$$

$$\langle x_{n-1}, z_n \rangle \geq t_n \max\big( \sup_{i \in \mathbb{N}} |\langle x_{n-1}, e_i \rangle|, \max_{j=1,2,\ldots n-1} |\langle x_{n-1}, z_j |\big) \tag{2.18}$$

$$\langle x_j, z_{j+1} \rangle \geq t_j \langle x_j, z_n \rangle, \text{ for all } j = 0, 1, 2 \ldots n-1. \tag{2.19}$$

Here $(e_j)$ denotes an orthonormal basis of $H$.

We deduce then that $(x_n)_{n=0}^{\infty}$ and $(z_n)_{n=0}^{\infty}$ is a pair of WPGA-sequences with weakness factor $\tau$ and dictionary $\mathcal{D} = \{\pm e_j, \pm z_j : j \in \mathbb{N}\}$.

*Proof of Theorem 2.2.5.* The following procedure is the key observation towards inductively producing our example. We let $(e_j)$ be an orthonormal basis of $H$.

For given $t \in (0, 1/3]$ and $i \neq j$ in $\mathbb{N}$. We define elements $x_n \in \text{span}(e_i, e_j)$, $n \geq 0$ and $z_n \in (e_i, e_j)$, $\|z_n\| = 1$, $n \geq$, and $\alpha_n \in [0, 1]$ recursively until we stop at some $n = N$, when some criterium is satisfied, as follows:

We put $x_0 = e_i$,

Now assume that for some $n \in \mathbb{N}$, we defined $x_s = a_s e_i + b_s e_j$ and $\alpha_s \in (0, 1)$ and $z_s \in S_H$ for all $1 \leq s \leq n-1$ so that for all $1 \leq s < n$ we have

$$\langle x_{s-1}, z_s \rangle = t, \text{ as long as } s \leq N-1, \tag{2.20}$$

$$z_s = \alpha_s e_i - (1 - \alpha_s^2)^{1/2} e_j, \tag{2.21}$$

$$a_s, b_s \geq 0, \text{ and } a_s - b_s \geq \sqrt{2}, \text{ as long as } s \leq N, \tag{2.22}$$

$$x_s = x_{s-1} - \langle x_{s-1}, z_s \rangle z_s. \tag{2.23}$$

(Conditions (2.20) and (2.22) become vacuous once we defined $N$ for $s = N$) Then we first define $\tilde{z}_n$ as

$$\tilde{z}_n = \tilde{\alpha}_n e_i - (1 - \tilde{\alpha}_n^2)^{1/2} e_j$$

where $\tilde{\alpha}_n$ is defined so that $\langle \tilde{x}_{n-1}, \tilde{z}_n \rangle = t$. Secondly define $\tilde{x}_n$ to be

$$\tilde{x}_n = x_{n-1} - \langle x_{n-1}, z_n \rangle z_n$$

and write $\tilde{x}_n$ as

$$\tilde{x}_n = \tilde{a}_n e_i + \tilde{b}_n e_j.$$

**Case 1.** $\tilde{a}_n - \tilde{b}_n \geq \sqrt{2}t$ In that case we choose $\alpha_n = \tilde{\alpha}_n$ and $z_n = \alpha_n e_i - (1 - \alpha_n^2)^{1/2} e_j$. Thus (2.20) and (2.21) are satisfied for $s = n$. Then we let $x_n = \tilde{x}_n$, and have therefore satisfied (2.22) and (2.23).
**Case 2.** $\tilde{a}_n - \tilde{b}_n < \sqrt{2}t$. Then we let $N = N_t = n$ and put $\alpha_N = 1/\sqrt{2}$, $z_N = (e_i - e_j)/\sqrt{2}$ and $x_N = x_{N-1} - \langle x_{N-1}, z_N \rangle z_N$. Then (2.21),and (2.23) are satisfied while (2.20) is vacuous.

From the definition of $X_N$ in Case 2, we observe that

$$\text{(2.24)} \qquad \langle x_{N-1}, z_N \rangle = 2^{-1/2}\big(a_{N-1} - b_{N-1}\big) \geq t, \text{ and it follows therefore that}$$

$$\text{(2.25)} \qquad a_N = b_N = \frac{1}{2}(a_{N-1} + b_{N-1}).$$

In particular also (2.22) is satisfied for $n = N$, assuming that $N$ is finite, which we will see later (here the second part of (2.22) is vacuous).

Once the second case happens we finish the definition of our sequences.

We still will have to show that eventual Case 2 will happens and that $N$ is finite; for the moment we think of $N$ being an element of $\mathbb{N} \cup \{\infty\}$

We make the following observations. From (2.20) and (2.21) we deduce that

$$\text{(2.26)} \qquad a_{n+1} = a_n - t\alpha_{n+1} \text{ and } b_{n+1} = b_n + t(1 - \alpha_{n+1}^2)^{1/2}, \text{ if } n < N - 1$$

which implies that

$$\text{(2.27)} \qquad a_{n+1} - b_{n+1} = a_n - b_n - t\big(\alpha_{n+1} + (1 - \alpha_{n+1}^2)^{1/2}\big)$$
$$\begin{cases} \geq a_n - b_n - \sqrt{2}t \\ \leq a_n - b_n - t \end{cases} \quad \text{if } n < N - 1.$$

This yields

$$1 = a_0 - b_0 \geq \sum_{s=0}^{N-2} (a_s - b_s) - (a_{s+1} - b_{s+1}) \geq (N-1)t$$

and therefore we showed that $N$ is finite. Since by definition of $N$ and $\tilde{a}_N$ and $\tilde{b}_N$

$$\begin{aligned} \sqrt{2}t &> \tilde{a}_N - \tilde{b}_N \\ &= a_{N-1} - b_{N-1} - t\big(\tilde{\alpha}_N + (1 - \tilde{\alpha}_N^2)\big)^{1/2} \geq a_{N-1} - b_{N-1} - t\sqrt{2} \end{aligned}$$

it follows that

$$\text{(2.28)} \qquad \langle x_{N-1}, z_N \rangle = 2^{-1/2}\big(a_{N-1} - b_{N-1}\big) \begin{cases} \leq 2t \\ \geq t \end{cases} .$$

It follows therefore from (2.27),(2.24) and (2.25) that

$$1 = a_0 - b_0 = \sum_{s=0}^{N-1} (a_s - b_s) - (a_{s+1} - b_{s+1}) \begin{cases} \geq tN \\ \leq \sqrt{2}tN \end{cases}$$

and thus

$$\frac{1}{\sqrt{2}t} \leq N \leq \frac{1}{t}. \tag{2.29}$$

From the definition of $x_N$ and (2.28) we deduce that

$$\begin{aligned} \|x_N\|^2 &= \|x_{N-1}\|^2 - \langle x_{N-1}, z_N\rangle^2 \text{ (By (2.28))} \\ &\geq \|x_{N-1}\|^2 - 4t^2 \\ &= \|x_0\|^2 + \sum_{s=1}^{N-1} \big(\|x_s\|^2 - \|x_{s-1}\|^2\big) - 4t^2 \\ &= \|x_0\|^2 - (N-1)t^2 - 4t^2 \geq \|x_0\|^2 - t - 3t^2 \text{ (By 2.29)} \end{aligned}$$

and thus, since $t \leq 1/3$,

$$\|x_N\|^2 \geq \|x_0\|^2 - 2t \tag{2.30}$$

Finally note that the sequence $(x_n)_{n=0}^N$ is a WAGD sequence for the Dictionary $\mathcal{D}' = \{z_n : n = 1, 2 \dots N_t\} \cup \{e_i\}$ with the weakness factor $t$. We call $(x_n)_{n=0}^{N_t}$ together with the sequence $(z_n)_{n=0}^{N_t}$ *the WAGD sequence generated by $t$ and the pair $(e_i, e_j)$.*

Now we assume that $(t_n)_{n=1}^\infty$ is a sequence in $(0,1]$, so that $\sum_{j=1}^\infty t_n < \infty$. We first require the additional assumption that $\sum_{j=1}^\infty \frac{t_n}{n} < \infty < \varepsilon = \frac{3}{16}$. It follows that

$$\begin{aligned} \sum_{s=0}^\infty t_{2^s} &= t_1 + t_2 + t_4 + t_8 \dots \\ &\leq t_1 + t_2 + \frac{1}{2}(t_3 + t_4) + \frac{1}{4}(t_5 + t_6 + t_7 + t_8) + \frac{1}{8}(t_9 + t_{10} + \dots t_{16}) + \dots \\ &\leq 2\Big[t_1 + \frac{t_2}{2} + \frac{1}{4}(t_3 + t_4) + \frac{1}{8}(t_5 + t_6 + t_7 + t_8) + \frac{1}{16}(t_9 + t_{10} + \dots t_{16}) + \dots\Big] \\ &\leq 2\sum_{n=1}^\infty \frac{t_n}{n} \leq 2\varepsilon. \end{aligned}$$

We will construct recursively sequences $(x_n : n = 0, 1, 2, \dots)$ and $(z_n : n = 1, 2 \dots)$ so that $x_0 = e_1$, $x_n = x_{n-1} - \langle x_{n-1}, z_n\rangle z_n$, so that for every $n \in \mathbb{N}$

$$\langle x_{n-1}, z_n\rangle \geq t_n \max_{j=1,\dots n-1} \langle x_{n-1}, z_j\rangle \text{ and } \langle x_{n-1}, z_n\rangle \geq t_n \sup_{j\in\mathbb{N}} \langle x_{n-1}, e_j\rangle \tag{2.31}$$

and

(2.32) $\quad t_j\langle x_j, z_n\rangle \le \langle x_j, z_{j+1}\rangle$ for all $j = 0, 1, 2, \dots n-1$.

As noted in the remark before the proof, these two conditions will ensure that that for each $n$ the vector is of the form $x_n = x - G_n(x)$, where $(G_n(x) : n \in \mathbb{N}_0)$ is the result of a WPGA with weakness factors $(t_n)$ and dictionary $\mathcal{D} = \{z_n, e_n : n \in \mathbb{N}\}$.

We start with $x = x_0 = e_1$, and let $(x_n^{(1)} : n = 0, 1, 2, \dots N_{t_1})$ and $(z_n^{(1)} : n = 1, 2, \dots N_{t_1})$ be the WAGD sequence generated by $t$ and the pair $(e_1, e_2)$, then we put $x_n = x_n^{(1)}$ and $z_n = z_n^{(1)}$ for $n = 1, 2, 3 \dots N_{t_1}$.

Note that we satisfied so far our required conditions (2.31) and (2.32) since by construction $\langle x_{n-1}, z_n\rangle = t_1 \ge t_n \ge t_n\|x_{n-1}\|$ for all $n = 1, 2 \dots, N_1 = N_{t_1}$. By (2.25) $x_{N_1}$ is of the form $x_{N_1} = c_1(e_1 - e_2)$, and we deduce from (2.30) and the fact that $\|x_{N_1}\| \le 1$ that

$$c_1^2 \le 1/2, N_1 \ge 1 \text{ and } \|x_{N_1}\|^2 \ge 1 - 2t_1.$$

Then we consider let $(x_n^{(2,1)} : n = 0, 1, \dots N_{t_2})$ and $(z_n^{(2,1)} : n = 1, 2 \dots N_{t_2})$ be the WAGD sequence generated by $t_2$ and the pair $(e_1, e_3)$, and $(x_n^{(2,2)} : n = 0, 1, \dots N_{t_2})$ and $(z_n^{(2,2)} : n = 1, 2 \dots N_{t_2})$ be the WAGD sequence generated by $t_2$ and the pair $(e_2, e_4)$. We put $N_2 = N_{t_2}$ and for $n = 1, 2, \dots N_2$ we define

$$x_{N_1+n} = c_1 x_n^{(2,1)} + c_1 e_2 \text{ and } z_{N_1+n} = z_n^{(2,1)}$$
$$x_{N_1+N_2+n} = c_1 x_{N_1}^{(2,1)} + c_1 x_n^{(2,2)} \text{ and } z_{N_1+N_2+n} = z_n^{(2,2)}.$$

We observe that for $n = 1, 2 \dots N_2$

$$\langle x_{N_1+n-1}, z_{N_1+n}\rangle = c_1 t_2 \ge t_2 \max_{s\in\mathbb{N}} \langle x_{N_1+n-1}, e_s\rangle \text{ and}$$

$$\langle x_{N_1+n-1}, z_{N_1+n}\rangle = c_1 t_2 \ge t_2 \max_{s=1,2,\dots N_1} \langle x_{N_1+n-1}, z_s\rangle$$

(the first inequality follows from the fact that the coordinates of $x_{N_1+n}$, $n = 1, 2 \dots N_2$ are absolutely, not larger than $c_1$, the second inequality follows from (2.21) and the fact moreover the coordinates of $x_{N_1+n}$, $n = 1, 2 \dots N_2$ are not negative while $z_s$, $s = 1, 2 \dots N_1$, has a positive and negative coordinate). Secondly we note that for $j = 1, 2, \dots, N_1$ and $n = 1, 2, \dots, N_2 - 1$, it follows from (2.20) and (2.24) that

$$t_j\langle x_j, z_{N_1+n}\rangle \le t_1 \le \langle x_j, z_{j+1}\rangle.$$

This implies that the conditions (2.31) and (2.32) hold for all $N_1 \le n \le N_1 + N_2$. Similarly we can show that they also hold for all $N_1 + N_2 \le n \le N_1 + 2N_2$.

Finally (2.30) implies that $X_{N_1+2N_2}$ is of the form

$$X_{N_1+2N_2} = c_2(e_1 + e_2 + e_3 + e_4)$$

with

$$c_2^2 \le 1/4 \text{ and } \|x_{N_1+2N_2}\|^2 \ge \|x_{N_1}\|^2 - c_1^2 2t_2 - c_1^2 2t_2 \ge 1 - 2t_1 - 2t_2.$$

Now assume that for some $r \in \mathbb{N}$ we have chosen

$$(x_n : n = 1, 2, \dots M_r), \text{ with } M_r = \sum_{j=1}^{r} 2^{j-1} N_j \text{ and } N_j = N_{t_{2^{j-1}}}, \; j = 1, 2 \dots r,$$

and

$$(z_n : n = 1, 2, \dots M_r)$$

so that (2.31) and (2.32) hold for all $n \le M_r$, and so that

$$x_{M_r} = c_r \sum_{i=1}^{2^r} e_i$$

for some $c_r$ with $c_r^2 \le 2^{-r}$, and so that

$$\|x_{M_r}\|^2 \ge 1 - 2t_1 - 2t_2 - 2t_4 - \dots - 2t_{2^{r-1}}$$

then we let for $j = 1, 2 \dots 2^r$ $(x_n^{(r+1,j)} : n = 0, \dots N_{r+1})$ and $(z_n^{(r+1,j)} : n = 0, \dots N_{r+1})$,with $N_{r+1} = N_{t_{2^r}}$, be the WPGA sequences generated by by $t_{2^r}$ and the pair $(e_j, e_{2^r+j})$, and finally put for $i = 1, 2 \dots 2^r$ and $n = 1, 2, \dots N_r$

$$x_{M_r+(i-1)N_{r+1}+n} = \sum_{s=1}^{i-1} c_r x_{N_r}^{(r+1,s)} + c_r x_n^{(r+1,i)} + c_r \sum_{s=i+1}^{2^r} e_s, \text{ and}$$
$$z_{M_r+(i-1)N_{r+1}+n} = z_n^{(r+1,i)}.$$

We deduce as in the case $r = 1$ that the conditions (2.31) and (2.32) hold for all $n \le M_r + 2^r N_{r+1} = \sum_{s=1}^{r+1} 2^{s-1} N_s = M_{r+1}$, that $x_{M_{r+1}} = c_{r+1} \sum_{s=1}^{2^{r+1}} e_s$ , for some $c_{r+1} \le 2^{-r-1}$, and that

$$\|x_{M_{r+1}}\| \ge 1 - 2t_1 - 2t_2, \dots 2t_r.$$

This finishes the choice of the the $x_n$ and $z_n$.

Since $\|x_{M_r}\|^2 \ge 1 - 2\sum_{s=1}^{r} t_{2^s} \ge 1 - 4\varepsilon > 1 - \frac{12}{16} = \frac{1}{4}$, it follows that $(x_n)$ does not converge. We therefore proved our claim under the additional assumption that $\sum_{n=1}^{\infty} (t_n/n) < 3/16$.

In the general case we proceed as follows. We first find an $n_0$ so that

$$\sum_{s=n_0}^{\infty} t_{2^s} < 3/16,$$

and let

$$x = \sum_{j=1}^{2^{n_0}} e_j.$$

Then we choose $z_i = e_i$, $i = 1, 2 \ldots, 2^{n_0} - 1$, and thus $x_0 = x$ and recursively

$$x_n = x_{n-1} - \langle x_{n-1}, z_n \rangle = \sum j = n + 1^{2^{n_0}} e_j$$

for $n = 1, 2, \ldots 2^{n_0} - 1$. In particular $x_{2n_0-1} = e_{2^{n_0}}$ from then on we choose $x_{2n_0-1+n} = \tilde{x}_n$, $n = 1, 2 \ldots$ and $z_{2n_0-1+n} = \tilde{z}_n$, where the $\tilde{x}_n$ and $\tilde{z}_n$ are chosen like the $x_n$ and the $z_n$ in the special case, but in the Hilbertspace $\tilde{H} = \text{span}(e_j : j \geq 2^{n_0})$. □

## 2.3 Convergence Rates

Note that without any special conditions on the starting point in the Pure greedy algorithm (or others) we can not expect being able to estimate the convergence rate.

Indeed let $(\xi_n)$ be any sequence of positive numbers, which decreases to 0 and let $\mathcal{D} = \{\pm e_n : n \in \mathbb{N}\}$, where $(e_n)$ is an orthonormal basis of our Hilbert space $H$, then take

$$x = \sum_{j=1}^{\infty} \sqrt{\xi_j - \xi_{j+1}} e_j$$

then it follows for the $n$-th approximates $G_n = G_n(x)$ define as in (PGA)

$$G_n = \sum_{j=1}^{n} \sqrt{\xi_j - \xi_{j+1}} e_j$$

and thus

$$\|x - G_n\|^2 = \sum_{j=1+n}^{\infty} \xi_j - \xi_{j+1} = \xi_{n+1}.$$

Thus no matter how slow $(\xi_n)$ converges to 0, there is a $x$ so that $G_n(x)$ converges at least as slow as $(\xi_n)$.

In order to state our first result we introduce for a dictionary $\mathcal{D}$ of $H$ the following linear subspace:

$$\mathcal{A}_1 = \mathcal{A}_1(\mathcal{D}) = \Big\{ \sum_{z \in \mathcal{D}} c_z z : (c_z) \subset \mathbb{K} \text{ and } \sum_{z \in \mathcal{D}} |c_z| < \infty \Big\}. \tag{2.33}$$

For $x \in \mathcal{A}_1$ we put

$$\|x\|_{\mathcal{A}_1} = \inf \Big\{ \sum_{z \in \mathcal{D}} |c_z| : (c_z) \subset \mathbb{K} \text{ and } f = \sum_{z \in \mathcal{D}} c_z \Big\}. \tag{2.34}$$

**Theorem 2.3.1.** [DT] *Assume $\mathcal{D}$ is a dictionary of a separable Hilbert space. Let $x \in \mathcal{A}_1(\mathcal{D})$ and assume that $(G_n) = (G_n(x))$ is defined as in (PGA) and let $x_n = x - G_n$, for $n \in \mathbb{N}$. Then*

$$\|x_n\| \le \|f\|_{\mathcal{A}_1} n^{-1/6} \text{ for } n \in \mathbb{N}. \tag{2.35}$$

For the proof of Theorem 2.3.1 we need the following observation.

**Lemma 2.3.2.** *Assuming that $(\xi_m)$ is a sequence of positive numbers so that for some number $A > 0$*

$$\xi_1 \le A \text{ and } \xi_{m+1} \le \xi_m(1 - \xi_m/A), \text{ for } m \ge 1. \tag{2.36}$$

*Then*

$$\xi_m \le \frac{A}{m}, \text{ for all } m \in \mathbb{N}. \tag{2.37}$$

*Proof.* We assume $A = 1$ (pass to $\tilde{\xi}_m = \xi_m/A$) We prove the claim by induction for each $m \in \mathbb{N}$. For $m = 1$ (2.37) follows from the assumption. Assume that the claim is true for $m \in \mathbb{N}$. If $\xi_m \le \frac{1}{m+1}$ then also $\xi_{m+1} \le \frac{1}{m+1}$ since from (2.36) it follows that the sequence $(\xi_i)$ is decreasing. If $\frac{1}{m+1} < \xi_m \le \frac{1}{m}$ we deduce that

$$\xi_{m+1} \le \xi_m(1 - \xi_m) \le \frac{1}{m}\Big(1 - \frac{1}{m+1}\Big) = \frac{1}{m}\frac{m}{m+1} = \frac{1}{m+1},$$

which implies the claim for $m + 1$ and finishes the induction step. □

*Proof of Theorem 2.3.1.* For $x \in H$ we put

$$\rho(x) = \sup_{z \in \mathcal{D}} \frac{\langle x, z\rangle}{\|x\|}.$$

Note that if $x \in \mathcal{A}_1$, $\eta > 0$ and $(c_z)_{z\in\mathcal{D}} \subset \mathbb{R}_0^+$ is such that $x = \sum_{z\in\mathcal{D}} c_z z$ and $\sum c_z \le \eta + \|x\|_{\mathcal{A}_1}$ it follows that

$$\|x\|^2 = \Big\langle x, \sum_{z\in\mathcal{D}} c_z z\Big\rangle \le \sum_{z\in\mathcal{D}} c_z \sup_{z\in\mathcal{D}}\langle z, x\rangle \le \big(\|x\|_{\mathcal{A}_1} + \eta\big)\|x\|\rho(x),$$

and, thus, since $\eta > 0$ was arbitrary,

$$\rho(x) \ge \frac{\|x\|}{\|x\|_{\mathcal{A}_1}}. \tag{2.38}$$

Let $x \in \mathcal{A}_1$ and let us assume that there is a representation $x = \sum_{z\in\mathcal{D}} c_z z$ so that $\|x\|_{\mathcal{A}_1} = \sum_{z\in\mathcal{D}} c_z$ (otherwise we use arbitrary approximations). Let $(z_m)$

and $(G_m)$ be defined as in (PGA) and $x_n = x - G_n$, for $n \in \mathbb{N}$. We note that for $m \in \mathbb{N}_0$

$$\begin{aligned} \|x_{m+1}\|^2 &= \|x_m - \langle x_m, z_{m+1}\rangle z_{m+1}\|^2 \\ &= \|x_m\|^2 - \langle x_m, z_{m+1}\rangle^2 = \|x_m\|^2(1-\rho^2(x_m)). \end{aligned} \tag{2.39}$$

Putting $a_m = \|x_m\|^2$, $b_0 = \|x_0\|_{\mathcal{A}_1} = \|x\|_{\mathcal{A}_1}$ and, assuming that $b_m$ has been defined, we let $b_{m+1} = b_m + \rho(x_m)\|x_m\| = b_m + \rho(x_m)a_m^{1/2}$. First we observe that

$$\|x_m\|_{A_1} \le b_m. \tag{2.40}$$

Indeed, for $m = 0$ this simply follows from the definition of $b_0$, and assuming (2.40) holds for $m \in \mathbb{N}_0$ it follows that

$$\begin{aligned} \|x_{m+1}\|_{\mathcal{A}_1} &= \|x_m - \langle x_m, z_{m+1}\rangle z_{m+1}\|_{\mathcal{A}_1} \\ &\le \|x_m\|_{\mathcal{A}_1} + |\langle x_m, z_{m+1}\rangle| \\ &= \|x_m\|_{\mathcal{A}_1} + \rho(x_m)\|x_m\| = b_{m+1}. \end{aligned}$$

Secondly we compute using (2.39), (2.38) and (2.40)

$$a_{m+1} = \|x_{m+1}\|^2 = a_m(1-\rho^2(x_m)) \le a_m\Big(1 - \frac{\|x_m\|^2}{\|x_m\|^2_{\mathcal{A}_1}}\Big) \le a_m\Big(1 - \frac{a_m}{b_m^2}\Big)$$

and thus, since $b_{m+1} \ge b_m$

$$\frac{a_{m+1}}{b_{m+1}^2} \le \frac{a_{m+1}}{b_m^2} \le \frac{a_m}{b_m^2}\Big(1 - \frac{a_m}{b_m^2}\Big).$$

Note that $\frac{a_0}{b_0^2} = \frac{\|x\|^2}{\|x\|^2_{\mathcal{A}_1}} \le 1$. We therefore apply Lemma 2.3.2 to sequence $\xi_n$ with $\xi_n = \frac{a_{n-1}}{b_{n-1}^2}$ and deduce that

$$a_m b_m^{-2} \le \frac{1}{m}, \text{ whenever } m \in \mathbb{N}. \tag{2.41}$$

Since by the recursive definition of $(b_j)$, (2.40) and (2.38) we get

$$b_{m+1} = b_m\big(1 + \rho(x_m)a_m^{1/2}b_m^{-1}\big) \le b_m\big(1 + \rho(x_m)a_m^{1/2}\|x_m\|^{-1}_{\mathcal{A}_1}\big) \le b_m\big(1 + \rho^2(x_m)\big),$$

we obtain together with (2.39)

$$a_{m+1}b_{m+1} \le a_m b_m(1-\rho^2(x_m))(1+\rho^2(x_m)) \le a_m b_m.$$

$(a_m b_m)$ is therefore decreasing and $a_m b_m \le a_0 b_0 = \|x\|^2 \cdot \|x\|_{\mathcal{A}_1}$. Multiplying both sides of (2.41) by $a_m^2 b_m^2$ we obtain therfore

$$a_m^3 \le \frac{a_m^2 b_m^2}{m} \le \frac{\|x\|^4 \cdot \|x\|^2_{\mathcal{A}_1}}{m},$$

which implies our claim after taking on both sides the sixth root. □

The next Example due to DeVore and Temlyakov gives a lower bound for the convergence rate of (PGA)

**Example 2.3.3.** [DT] Let $H$ be a separable Hilbertspace and $(h_j)$ an orthonormal basis of $H$. We will define a dictionary $\mathcal{D} \subset H$, a vector $x \in H$ for which $\|x\|_{\mathcal{A}_1(\mathcal{D})} = 2$, and so that

$$\|x_m\| = \|x - G_m\| \geq \frac{c}{\sqrt{m}}, \text{ for } m \in \mathbb{N}.$$

Define

$$a = \left(\frac{23}{11}\right)^{1/2} \text{ and } A = \left(\frac{33}{89}\right)^{1/2}$$

and

$$z = A(h_1 + h_2) + aA \sum_{k=3}^{\infty} \big(k(k+1)\big)^{-1/2} h_k.$$

Note that

$$\begin{aligned}
\|z\|^2 &= 2A^2 + a^2A^2 \sum_{k=3}^{\infty} \frac{1}{k}\frac{1}{k+1} \\
&= 2A^2 + a^2A^2 \sum_{k=3}^{\infty} \frac{1}{k} - \frac{1}{k+1} \\
&= 2A^2 + \frac{1}{3}a^2A^2 = \frac{33}{89}\Big(2 + \frac{23}{33}\Big) = 1.
\end{aligned}$$

Put $\mathcal{D} = \{\pm g\} \cup \{\pm h_j : j \in \mathbb{N}\}$ and let $x = h_1 + h_2$ and we apply (PGA) to $f$
**Claim:** In Step 1 of (PGA) we have

$$z_1 = z \text{ and } x_1 = x - \langle x, z\rangle z = (1 - 2A^2)(h_1 + h_2) - 2aA^2 \sum_{k=3}^{\infty} \big(k(k+1)\big)^{-1/2} h_k.$$

Indeed,

$$\begin{aligned}
&\langle x, z\rangle = 2A > 1, \\
&\langle x, h_1\rangle = \langle x, h_2\rangle = A, \text{ and} \\
&\langle x, h_j\rangle = 0, \text{if}, j > 2.
\end{aligned}$$

Thus $z_1 = z$ and

$$\begin{aligned}
x_1 &= x - \langle x, z\rangle z \\
&= h_1 + h_2 - 2A\Big(A(h_1 + h_2) + aA \sum_{k=3}^{\infty} \big(k(k+1)\big)^{-1/2} h_k\Big)
\end{aligned}$$

$$= (1-2A^2)(h_1+h_2) - 2aA^2 \sum_{k=3}^{\infty} \big(k(k+1)\big)^{-1/2} h_k.$$

**Claim:** In Step 2 and Step 3 of (PGA) we have $z_2 = h_1$ and $z_3 = h_2$ or vice versa.

Indeed,

$$\begin{aligned}
&\langle x_1, z\rangle = 0 \text{ (by construction of } x_1)\\
&\langle x_1, h_1\rangle = \langle x_1, h_2\rangle = (1-2A^2) = \frac{23}{89} \text{ and}\\
&\langle x_1, h_j\rangle = 2aA^2\big(j(j+1)\big)^{-1/2} \le \frac{1}{6}aA^2 < (1-2A^2) \text{ if } j>2.
\end{aligned}$$

So take W.l.o.g $z_2 = h_1$ and thus

$$x_2 = x_1 - \langle x_1, h_1\rangle h_1 = (1-2A^2)h_2 - 2aA^2 \sum_{k=3}^{\infty} \big(k(k+1)\big)^{-1/2} h_k.$$

Then we observe that

$$\begin{aligned}
&\langle x_2, z\rangle = \langle x_1, z\rangle + \langle x_2 - x_1, z\rangle = -(1-2A^2)\langle h_1, z\rangle = -A(1-2A^2)\\
&\langle x_2, h_1\rangle = 0,\\
&\langle x_2, h_2\rangle = (1-2A^2) > |\langle x_2, z\rangle|\\
&\langle x_2, h_j\rangle = 2aA^2\big(j(j+1)\big)^{-1/2} \le \frac{1}{6}aA^2 < (1-2A^2) \text{ if } j>2.
\end{aligned}$$

which implies that $z_3 = h_2$ and that

$$x_3 = x_2 - \langle x_2, h_2\rangle h_2 = -2aA^2 \sum_{k=3}^{\infty} \big(k(k+1)\big)^{-1/2} h_k.$$

From now on we prove by induction that

$$z_m = h_{m-1} \text{ and } x_m = -2aA^2 \sum_{k=m}^{\infty} \big(k(k+1)\big)^{-1/2} h_k \text{ whenever } m \ge 3.$$

Indeed, for $m=3$ this was already shown. Assuming that our claim is true for some $m \ge 3$ we compute that

$$\begin{aligned}
&\langle x_m, z\rangle = -2a^2A^3 \sum_{k=m}^{\infty} \big(k(k+1)\big) = -\frac{2a^2A^3}{m(m+1)}\\
&\langle x_m, h_\ell\rangle = 0 \text{ for } \ell < m, \text{ and}\\
&\langle x_m, h_\ell\rangle = -2a^2A^3\big(\ell(\ell+1)\big)^{-1/2} \text{for } \ell \ge m.
\end{aligned}$$

Thus $z_{m+1} = h_m$ and $x_m = -2aA^2 \sum_{k=m+1}^{\infty} \big(k(k+1)\big)^{-1/2} h_k$, which finishes the induction step.

Finally note that for $m \geq 3$

$$\|x_m\|^2 = 4a^2A^4 \sum_{k=m}^{\infty} \frac{1}{k(k+1)} = \frac{4a^2A^4}{m}.$$

Thus $\|x\|_{\mathcal{A}_1} = 2$ and $(\|x_m\|)$ is of the order $m^{-1/2}$.

**Remark.** In [KT2] the rate of convergence in Theorem 2.3.1 was slightly improved to $Cn^{-11/62}$ where $C$ is some universal constant. And in [LT] a dictionary $\mathcal{D}$ of $H$ was constructed for which there is an $x \in \mathcal{A}_1(\mathcal{D})$ for which $\|x - G_n(x)\| \geq Cn^{-.27}$, whenever $n \in \mathbb{N}$.

It is conjectured that the rate of convergences should be of the order of $n^{-1/4}$.

**Theorem 2.3.4.** [Jo] *Consider the Greedy Greedy Algorithm* $(G^f_m(x))$, $x \in H$ *with fixed relaxation, and let* $x \in \mathcal{A}_1(D)$ *with* $\|x\|_{\mathcal{A}_1} \leq c$, *where* $c$ *is the constant in (GAFR) then*

$$\|x - G^r_m\|_{\mathcal{A}_1} \leq \frac{2c}{\sqrt{m}}, \text{ for all } m \in \mathbb{N}. \tag{2.42}$$

We first need the following elementary Lemma.

**Lemma 2.3.5.** *Let* $(a_m)$ *be a sequence of non-negative numbers, for which there is an* $A >$ *so that*

$$a_1 \leq A \text{ and } a_m \leq \Big(1 - \frac{2}{m}\Big)a_{m-1} + \frac{A}{m^2} \text{ if } m \geq 2. \tag{2.43}$$

*Then*

$$a_m \leq \frac{A}{m} \text{ for all } m \in \mathbb{N}. \tag{2.44}$$

*Proof.* We prove (2.44) by induction. For $m = 1$ (2.44) is part of our assumption, and assuming (2.47) is true for $m - 1$ we deduce from the second part of our assumption that

$$\begin{aligned} a_m &\leq \Big(1 - \frac{2}{m}\Big)a_{m-1} + \frac{A}{m^2} \\ &\leq \Big(1 - \frac{2}{m}\Big)\frac{A}{m-1} + \frac{A}{m^2} \\ &= A\Big(\frac{1}{m-1} - \frac{m+1}{m^2(m-1)}\Big) \\ &= A\Big(\frac{m^2 - m - 1}{m^2(m-1)}\Big) \end{aligned}$$

$$= \frac{A}{m}\Big(\frac{m^2-m-1}{m(m-1)}\Big) < \frac{A}{m},$$

which finishes the induction step and the proof of our claim. □

*Proof of Theorem* (2.3.4). W.l.o.g. we can assume that $c=1$. Let $G_n^f = G_n^f(x)$ and $z_n$ be as in GAFR and let $x_n = x - G_n^f$, for $n \in \mathbb{N}$ we compute:

$$\begin{aligned}
\|x_n\|^2 &= \big\|x - G_n^f\big\|^2 \\
&= \Big\|x - \Big(1-\frac{1}{n}\Big)G_{n-1}^f - \frac{1}{n}z_n\Big\|^2 \\
&= \big\|x - G_{n-1}^f\big\|^2 + \frac{2}{n}\big\langle x - G_{n-1}^f, G_{n-1}^f - z_n\big\rangle + \frac{1}{n^2}\big\|G_{n-1}^f - z_n\big\|^2 \\
&\le \big\|x - G_{n-1}^f\big\|^2 + \frac{2}{n}\big\langle x - G_{n-1}^f, G_{n-1}^f - z_n\big\rangle + \frac{4}{n^2}
\end{aligned} \tag{2.45}$$

(for the inequality notice that if $\|G_{n-1}^f\| \le \|G_{n-1}^f\|_{\mathcal{A}_1} \le 1$) and

$$\begin{aligned}
\big\langle x - G_{n-1}^f, G_{n-1}^f - z_n\big\rangle &\le \inf_{z\in\mathcal{D}} \big\langle x - G_{n-1}^f, G_{n-1}^f - z\big\rangle \\
&= \inf_{z\in\mathcal{A}_1(\mathcal{D}), \|z\|_{\mathcal{A}_1}\le 1} \big\langle x - G_{n-1}^f, G_{n-1}^f - z\big\rangle \\
&= \langle x - G_{n-1}^f, G_{n-1}^f\rangle - \sup_{z\in\mathcal{A}_1(\mathcal{D}), \|z\|_{\mathcal{A}_1}\le 1} \big\langle G_{n-1}^f, G_{n-1}^f - z\big\rangle \\
&\le \big\langle x - G_{n-1}^f, G_{n-1}^f - x\big\rangle = -\|x - G_{n-1}^f\|^2.
\end{aligned}$$

Inserting this inequality into (2.45) yields

$$\|x_n\|^2 \le \Big(1-\frac{2}{n}\Big)\big\|x - G_{n-1}^f\big\|^2 + \frac{4}{n^2} = \Big(1-\frac{2}{n}\Big)\|x_{n-1}\||^2 + \frac{4}{n^2},$$

which together with Lemma 2.3.5 yields our claim. □

**Theorem 2.3.6.** *If we consider the orthogonal greedy algorithm (OGA), then for each $x \in H$ with $\|x\|_{\mathcal{A}_1} = \|x\|_{\mathcal{A}_1(\mathcal{D})} < \infty$, it follows*

$$\|x - G_n^o(x)\| \le \frac{\|x\|_{\mathcal{A}_1(\mathcal{D})}}{\sqrt{n}}. \tag{2.46}$$

*Proof.* Assume that $\|x\|_{\mathcal{A}_1} = 1$, and $G_n^o = G_n^o(x)$, $z_n$ are given as in (OGA), i.e.

$$\langle x - G_{n_1}^o, z_n\rangle = \sup_{z\in\mathcal{D}} \langle x - G_{n-1}, z\rangle$$

and $G_n^o$ is the orthogonal projection $P^{\perp}_{Z_n}(x)$ of $x$ onto $Z_n = \text{span}(z_1, z_2, \dots z_n)$. As usual we put $x_n = x - G_n^o$. Since $G_n^0$ is the best approximation of $x$ by elements of $Z_n$, it follows that

$$\|x_n\|^2 \le \|x_{n-1} - \langle x_{n-1}, z_n\rangle z_n\|^2 \\ = \|x_{n-1}\|^2 - \langle x_{n-1}, z_n\rangle^2 = \|x_{n-1}\|^2 \left(1 - \left(\frac{\langle x_{n-1}, z_n\rangle}{\|x_{n-1}\|}\right)^2\right) \tag{2.47}$$

(W.l.o.g. $x_{n-1} \ne 0$, otherwise we would be done Write $x$ as $x = \sum_{z\in\mathcal{D}} c_z z$, with $\sum c_z = 1$, $c_z \ge 0$, for $z \in \mathcal{D}$. Then

$$\begin{aligned} \|x_{n-1}\|^2 &= \langle x - G_{n-1}^o, x - G_{n-1}^o\rangle \\ &= \langle x - G_{n-1}^o, x\rangle \text{ (Since } x - G_{n-1}^o \perp G_{n-1}^o) \\ &= \left\langle x - G_{n-1}^o, \sum_{z\in\mathcal{D}} c_z z\right\rangle \\ &\le \left\langle x - G_{n-1}^o, \sum_{z\in\mathcal{D}} c_z z_n\right\rangle \\ &= \langle x_{n-1}, z_{n-1}\rangle = \|x_{n-1}\| \frac{\langle x_{n-1}, z_n\rangle}{\|x_{n-1}\|} \end{aligned}$$

and thus by (2.47)

$$\|x_n\|^2 \le \|x_{n-1}\|^2(1 - \|x_{n-1}\|^2).$$

Our claim follows therefore again from Lemma 2.3.5. □

# Chapter 3

# Greedy Algorithms in general Banach Spaces

## 3.1 Introduction

The algorithms from Chapter 2 can be generalized to separable Banach spaces $X$. Again let $\mathcal{D} \subset S_X$ be a dictionary of $X$, i.e. $X = \overline{\text{span}(\mathcal{D})}$, and with $z \in \mathcal{D}$, it also follows that $-z \in \mathcal{D}$.

(XGA) The $X$-*Greedy Algorithm.*

For $x \in X$ we define $G_n = G_n(x)$, for each $n \in \mathbb{N}_0$, by induction. $G_0 = 0$ and assuming that $G_0, G_1 \dots G_{n-1}$, have been defined for some $n \in \mathbb{N}$ we proceed as follows:

1) Choose $z_n \in \mathcal{D}$ and $a_n \in \mathbb{R}$ so that

$$\|x - G_{n-1} - a_n z_n\| = \inf_{z \in \mathcal{D}, a \geq 0} \|x - G_{n-1} - az\|.$$

2) Put $G_n = G_{n-1} + a_n z_n$.

As in the Hibert space case, the "inf" in the above defined algorithm (XGA) may not be attained. In this case we can consider the following modification.

(WXGA) The *Weak* $X$-*Greedy Algorithm.*

We are given a sequence $\tau = (t_n) \subset (0, 1)$. For $x \in X$ we define $G_n = G_n(x)$, for each $n \in \mathbb{N}_0$, by induction. $G_0 = 0$ and assuming that $G_0, G_1 \dots G_{n-1}$, have been defined for some $n \in \mathbb{N}$ we proceed as follows:

1) Choose $z_n \in \mathcal{D}$ and $a_n \in \mathbb{R}$ so that

$$t_n \|x - G_{n-1} - a_n z_n\| \leq \inf_{z \in \mathcal{D}, a \in \mathbb{R}} \|x - G_{n-1} - az\|.$$

2) Put $G_n = G_{n-1} + a_n z_n$.

The following example shows that without any further conditions, one can "get stuck pretty easily":

**Example 3.1.1.** On $\mathbb{R}^2$ consider the $\ell_\infty^2$ norm:

$$\|(x,y)\|_\infty = \max(|x|,|y|), \text{ if } x,y \in \mathbb{R}.$$

and let $\mathcal{D} = \{\pm e_1, \pm e_2\}$ be the dictionary.

Now for vector $x = (1,1)$ we have

$$\inf_{a\geq 0, z\in\mathcal{D}} \|x - az\|_\infty = 1 = \|x\|_\infty.$$

In order to avoid cases like in Example 3.1.1 we will assume that our space $X$ is at least smooth:

**Definition 3.1.2.** A Banach space $X$ is called *smooth* if for every $x \in X$ there is a unique *support functional* $f_x \in X^*$, i.e. with $\|f_x\| = 1$ and $f_x(x) = \|x\|$.

**Remark.** As shown for example in [Schl, Theorem 4.1.3] the assumption that $X$ is smooth is equivalent with the condition that the norm is Gateaux differentiable on $X \setminus \{0\}$. In that case it follows for $x_0 i \in X \setminus \{0\}$

$$f_{x_0}(y) = \frac{1}{\|x_0\|}\frac{\partial}{\partial\lambda}\|x_0 + \lambda y\|\big|_{\lambda=0} = \frac{1}{\|x_0\|}\lim_{h\to\infty}\frac{\|x_0 + hy\| - \|x_0\|}{h},$$

for all $y \in S_X$

This implies that the $X$-greedy algorithm or the weak $X$-greedy algorithm cannot become stationary at point $x_0 \neq 0$.

Indeed, if for all $z \in \mathcal{D}$ and all $\lambda$

$$\|x_0 - \lambda z\| \geq \|x_0\|,$$

it follows that for all $z \in \mathcal{D}$

$$f_{x_0}(z) = \frac{1}{\|x_0\|}\lim_{h\to 0}\frac{\|x_0 + hz\| - \|x_0\|}{h} = 0.$$

Since $\mathrm{span}(\mathcal{D})$ is dense this would mean that $f_{x_0} = 0$ which is a contradiction.

In the Hilbert space case minimizing $\|x - az\|$ over all $z \in \mathcal{D}$ and all $a \in \mathbb{R}$ is equivalent to maximizing $\langle x, z\rangle$ over all $z \in \mathcal{D}$. Generalizing this to Banach spaces will lead to a different algorithm, i.e. to an algorithm which does not coincide with (XGA).

(DGA) The *Dual Greedy Algorithm.*
For $x \in X$ we define $G_n^d = G_n^d(x)$, for each $n \in \mathbb{N}_0$, by induction as follows. $G_0^d = 0$ and assuming $G_0^d, G_1^d \dots G_{n-1}^d$, have been defined,

1) choose $z_n \in \mathcal{D}$ so that

$$f_{x-G_{n-1}^d}(z_n) = \sup_{z \in \mathcal{D}} f_{x-G_n^d}(z),$$

2) and then $a_n$ so that

$$\|(x - G_{n-1}^d) - a_n z_n\| = \min_{a \in \mathbb{R}} \||(x - G_{n-1}^d) - az\|.$$

Then put $G_n^d = G_{n-1}^d + a_n z_n$.

Similar to XGA we can also define the weak version of (DGA) and denote it by (WDGA).

(XGDAR) The *X-Greedy Dual Algorithm with relaxation.*
We are given a sequence $\rho = (r_n) \subset [0, 1)$.
For $x \in X$ we define $G_n^r(\rho)$, for each $n \in \mathbb{N}_0$, by induction.
$G_0^r(\rho) = 0$ and assuming that $G_0^r(\rho), G_1^r(\rho) \dots G_{n-1}^r(\rho)$, have been defined for some $n \in \mathbb{N}$ we proceed as follows:

1) Choose $z_n \in \mathcal{D}$ and $a_n \in \mathbb{R}$ so that

$$\|x-(1-r_n)G_{n-1}^r(\rho)-a_n z_n\| \leq \inf_{z \in \mathcal{D}, a \in \mathbb{R}} \|x-(1-r_n)G_{n-1}^r(\rho)-az\|.$$

2) Put $G_n^r(\rho) = (1 - r_n)G_{n-1}^r(\rho) + a_n z_n$.

The following algorithm is a generalization of the the Orthogonal Greed Algorithm for Hilbert spaces.

(CDGA) The *Chebyshev Dual Greedy Algorithm.*
For $x \in X$ we define $G_n^C$, for each $n \in \mathbb{N}_0$, by induction.
$G_0^C = 0$ and assuming that $G_0^C, G_1^C \dots G_{n-1}^C$, and vectors $z_1 \dots z_{n-1}$ have been defined for some $n \in \mathbb{N}$ we proceed as follows:

1) Choose $z_n \in \mathcal{D}$ so that

$$f_{x-G_{n-1}^C}(z_n) \geq \sup_{z \in \mathcal{D}} f_{x-G_{n-1}^C}(z).$$

2) Define $Z_n = \text{span}(z_1, z_2 \dots z_n))$ and let $G_n^C$ be the (or a) best approximation of $x$ to $Z_n$.

Similar to (WXGA) there are *weak version* of (XGDAR) and (CDGA) which we denote (WXGDAR) and (WCDGA), respectively.

The following algorithm is of a different nature than the previous ones. We will assume that $(e_i)$ is a semi normalized basis of $X$.

Before discussing these algorithms and there convergence properties we will need to introduce the following strengthening of smoothness of Banach spaces.

**Definition 3.1.3.** For a Banach space $X$ define the *modulus of uniform smoothness* by

$$\rho(u) = \rho_X(u) = \sup_{x,y\in S_X} \Big(\frac{1}{2}\big(\|x+uy\| + \|x-uy\|\big) - 1\Big). \tag{3.1}$$

We say that $X$ is *uniformly smooth* if

$$\lim_{u\to 0} \frac{\rho(u)}{u} = 0. \tag{3.2}$$

**Remark.** We will use the modulus of uniform smoothness as follows. For some $x, y \in X \setminus \{0\}$ (not necessarily in $S_X$ we would like to find an upper estimate for $\|x-y\|$, and write

$$\|x-y\| = \|x\|\cdot\Big\|\frac{x}{\|x\|} - \frac{\|y\|}{\|x\|}\frac{y}{\|y\|}\Big\| \le 2\left(1+\rho\Big(\frac{\|y\|}{\|x\|}\Big)\right) - \|x+y\|.$$

**Example 3.1.4.** As was shown in [Schl], the spaces $L_p[0,1]$ are uniform smooth if $1 < p < \infty$, more precisely for $X = L_p[0,1]$ and $u \ge 0$

$$\rho_X(u) \le \begin{cases} c_p u^p & \text{if } 1 \le p \le 2 \\ (p-1)u^2/2 & \text{if } 2 \le p < \infty \end{cases} \tag{3.3}$$

**Proposition 3.1.5.** *For any Banach space $X$ $\rho$ is a convex and even function and*

$$\max(0, u-1) \le \rho(u) \le u, \ \textit{for } u \ge 0.$$

**Lemma 3.1.6.** *For $x \ne 0$. Then for $u \in \mathbb{R}$*

$$0 \le \|x+uy\| - \|x\| - f_x(y) \le 2\|x\|\rho\Big(\frac{u\|y\|}{\|x\|}\Big). \tag{3.4}$$

*Proof.* First we note that

$$\|x+uy\| \ge f_x(x+uy) = \|x\| + uf_x(y),$$

which implies the first inequality in (3.4). Secondly, from the definition of $\rho(u)$, and assuming w.l.o.g that $y \ne 0$, it follows that

$$\|x+uy\| + \|x-uy\| = \|x\|\Big\|\frac{x}{\|x\|} + \frac{u\|y\|}{\|x\|}\frac{y}{\|y\|}\Big\| + \|x\|\Big\|\frac{x}{\|x\|} - \frac{u\|y\|}{\|x\|}\frac{y}{\|y\|}\Big\|$$

$$\leq 2\|x\|\left(1+\rho\Big(\frac{u\|y\|}{\|x\|}\Big)\right)$$

and

$$\|x-uy\| \geq f_x(x-uy) = \|x\| - uf_x(y),$$

and, thus,

$$\|x+uy\| \leq 2\|x\|\left(1+\rho\Big(\frac{u\|y\|}{\|x\|}\Big)\right) - \|x-uy\| \leq 2\rho\Big(\frac{u\|y\|}{\|x\|}\Big) + uf_x(y),$$

which implies our claim. □

**Corollary 3.1.7.** *If $X$ is a uniformly smooth Banach space and $x \in X \setminus \{0\}$, then*

$$f_x(y) = \left(\frac{d}{dx}\|x+uy\|\right)(0) = \lim_{u\to 0}\frac{\|x+uy\|-\|x\|}{u}, \tag{3.5}$$

*and this convergence is uniform in $x, y \in S_Y$. The norm is therefore Fréchet differentiable*

*Proof.* Note, that by (3.4)

$$\left|\frac{\|x+uy\|-\|x\|}{u} - f_x(y)\right| \leq 2\frac{\|x\|}{u}\rho\Big(\frac{u\|y\|}{\|x\|}\Big) \to_{u\to 0} .$$

If $\|y\| = 1$ and $\|x\| > \varepsilon$ it follows therefore

$$\left|\frac{\|x+uy\|-\|x\|}{u} - f_x(y)\right| \leq 2\frac{\|x\|}{u}\|\rho\Big(\frac{u}{\|x\|}\Big) \leq 2\rho(u)/u \to_{u\to 0},$$

which implies the claimed uniform convergence. □

## 3.2 Convergence of the Weak Dual Chebyshev Greedy Algorithm

Recall the *Weak Dual Chebyshev Greedy Algorithm*:

We are given a sequence of *weakness factors* $\tau = (t_n) \subset (0,1)$ and a dictionary $\mathcal{D} \subset X$. For $x \in X$ we choose $(z_n) \subset \mathcal{D}$ and $G_n^c = G_n^c(x)$ as folllows.

$G_0^c = 0$ and assuming $G_j^c$, $j = 01, 2 \ldots n$ and $z_j$, $j = 1, 2 \ldots n$ have been chosen, we let $z_n \in \mathcal{D}$ so that

$$f^c_{x-G_n-1}(z_n) \geq t_n \sup_{z\in\mathcal{D}} f^c_{x-G_{n-1}}(z).$$

Then let $Z_n = \text{span}(z_1, z_2, \dots z_n)$ and let $G_n^c$ be the (or a) best approximation of $x$ inside $Z_n$.

The main goal of this section is to prove two results by Temliakov We will need the following technical definition first.

**Definition 3.2.1.** Let $\rho$ be an even convex function on $[-2, 2]$ with $\lim_{u\to 0} \rho(u)/u = 0$ and $\rho(2) \geq 1$, let $\tau = (t_n)$ of numbers in $(0, 1]$ and $\Theta \in (0, 1/2]$ then let $\xi_m = \xi_m(\rho, \tau, \theta) = 0$ the (by the Intermediate Value Theorem uniquely existing) number for which

$$\rho(\xi_m) = \Theta t_m \xi_m. \tag{3.6}$$

**Theorem 3.2.2.** *Let $X$ be a uniformly smooth Banach space and $\rho$ its modulus of uniform smoothness, and let $\tau = (t_n)$ be sequence of numbers in $(0, 1]$. Assume that for any $\Theta > 0$ we have*

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

*We consider the weak Chebysheev Greedy Algorithm (WCDGA). This means for $x \in X$ $G_n^c = G_n^c(x)$ and $z_n \in \mathcal{D}$ is such that*

$$f_{x-G_{n-1}^C}(z_n) \geq \sup_{z\in\mathcal{D}} f_{x-G_{n-1}^C}(z),$$

*and $G_n^c$ is the best approximation of $x$ to $Z_n = \text{span}(z_1, z_2, \dots z_n)$. Let $x_n = x - G_n^c$, for $n \in \mathbb{N}$. Then $\lim_{n\to\infty} \|x_n\| = 0$.*

**Theorem 3.2.3.** *Let $X$ be a uniformly smooth Banach space and and assume that its modulus of uniform smoothness $\rho$ satisfies $\rho(u) \leq \gamma u^q$ for some $q \in (1, 2]$ and $\gamma \geq 1$. Let $\tau = (t_n)$ be sequence of numbers in $(0, 1]$. For $x \in X$, assume that $x \in \mathcal{A}_1(\mathcal{D})$ and let $G_n^c = G_n^c(x)$be defined as in Theorem 3.2.2. Then there is constant $C(q, \gamma)$, only dependent on $q$and $\gamma$ so that for all $n \in \mathbb{N}$*

$$\left\|x - G_n^c\right\| \leq C(q, \gamma) \|x\|_{\mathcal{A}_1(\mathcal{D})} \left(1 + \sum_{k=1}^{m} t_k^p\right)^{-1/p} \tag{3.7}$$

*where $p = q/(q - 1)$.*

We will first need some Lemmas

**Lemma 3.2.4.** *Let $X$ be a uniformly smooth Banach space and let $Z \subset X$ be a finite dimensional subspace. If $y$ is the best approximate of some $x \in X \setminus Z$ from $Z$ then*

$$f_{x-y}(z) = 0, \textit{for all } z \in Z.$$

*Proof.* Assume to the contrary that there is a $z \in Z$, $\|z\| = 1$, so that $\beta = f_{x-y}(z) > 0$. By the definition of $\rho(u)$ it follows for any $\lambda$ and $z \in S_Z$ that

$$\|x - y - \lambda z\| + \|x - y + \lambda z\| \le 2\|x - y\|\left(1 + \rho\Big(\frac{\lambda}{\|x-y\|}\Big)\right) \tag{3.8}$$

and secondly

$$\|x - y + \lambda z\| \ge f_{x-y}(x - y + \lambda z) = \|x - y\| + \lambda\beta. \tag{3.9}$$

It follows therefore from (3.8) and (3.9) that

$$\begin{aligned}
\|x - y - \lambda z\| &\le 2\|x - y\|\left(1 + \rho\Big(\frac{\lambda}{\|x-y\|}\Big)\right) - \|x - y + \lambda z\| \\
&\le \|x - y\| + \rho\Big(\frac{\lambda}{\|x-y\|}\Big) - \lambda\beta \\
&= \|x - y\| - \lambda\Big(\beta - \frac{\rho(\lambda/\|x-y\|)}{\lambda/\|x-y\|}\Big).
\end{aligned}$$

Since $\rho(u)/u \to_{u\to 0}$, it follows therefore that $\|x - y + \lambda z\| < \|x - y\|$, which is a contradiction. □

**Lemma 3.2.5.** *For any $x^* \in X^*$ we have*

$$\sup_{z\in\mathcal{D}} x^*(z) = \sup_{x\in\mathcal{A}_1(\mathcal{D}),\|x\|_{\mathcal{A}_1}\le 1} x^*(x). \tag{3.10}$$

*Proof.* for $x = \sum_{z\in\mathcal{D}} c_z z$, with $c_z \ge 0$, for $z \in \mathcal{D}$ and $\sum_{z\in\mathcal{D}} c_z \le 1$, it follows that

$$x^*(x) = \sum_{z\in\mathcal{D}} c_z x^*(z) \le \sup_{z\in\mathcal{D}} x^*(z),$$

and thus

$$\sup_{z\in\mathcal{D}} x^*(z) \ge \sup_{x\in\mathcal{A}_1(\mathcal{D}),\|x\|_{\mathcal{A}_1}\le 1} x^*(x).$$

The reverse inequality is trivial. □

**Lemma 3.2.6.** *Let $X$ be a uniformly smooth Banach space and $\rho$ its modulus of uniform smoothness, and let $\tau = (t_n)$ be a sequence of numbers in $(0,1]$. For $x \in X$ let $G^c_n = G^c_n(x)$ be defined as in (WCDGA).*

*Assume that $x \in X$ and that for some $\varepsilon \ge 0$ there is a $x_\varepsilon \in X$ so that $\|x - x_\varepsilon\| \le \varepsilon$ and $x_\varepsilon \in \mathcal{A}(\mathcal{D})$. Then it follows for all $n \in \mathbb{N}$*

$$\frac{\|x - G^c_n\|}{\|x - G^c_{n-1}\|} \le \inf_{\lambda\ge 0}\left[1 - \frac{\lambda t_n}{\|x\|_{\mathcal{A}_1}}\Big(1 - \frac{\varepsilon}{\|x - G^c_{n-1}\|}\Big) + 2\rho\Big(\frac{\lambda}{\|x_{n-1}\|}\Big)\right]. \tag{3.11}$$

*Proof.* Abbreviate $A = \|x_\varepsilon\|_{\mathcal{A}_1}$. Let $z_n \in \mathcal{D}$, be chosen as in (WCDGA) and put $x_n = x - G^c_n$, for $n \in \mathbb{N}$.

From the definition of $\rho$, it follows for every $\lambda \geq 0$ that

$$\|x_{n-1} - \lambda z_n\| + \|x_{n-1} + \lambda z_n\| \leq 2\|x_{n-1}\|\Big(1 + \rho\Big(\frac{\lambda}{\|x_{n-1}\|}\Big)\Big). \tag{3.12}$$

and by the definition of (WCDGA) and Lemma 3.2.5 it follows that

$$f_{x_{n-1}}(z_n) \geq t_n \sup_{z\in\mathcal{D}} f_{x_{n-1}}(z) = t_n \sup_{z\in\mathcal{A}_1(\mathcal{D}),\|z\|_{\mathcal{A}_1}\leq 1} f_{x_{n-1}}(z) \geq \frac{t_n}{A} f_{x_{n-1}}(x_\varepsilon).$$

From Lemma 3.2.4 we deduce that

$$f_{x_{n-1}}(x_\varepsilon) = f_{x_{n-1}}(x + x_\varepsilon - x) \geq f_{x_{n-1}}(x) - \varepsilon = f_{x_{n-1}}(x_{n-1}) - \varepsilon = \|x_{n-1}\| - \varepsilon,$$

and thus

$$\begin{aligned}\|x_{n-1} + \lambda z_n\| &\geq \|x_{n-1}\| + \lambda f_{x_{n-1}}(z_n)\\ &\geq \|x_{n-1}\| + \frac{\lambda t_n}{A} f_{x_{n-1}}(x_\varepsilon) \geq \|x_{n-1}\|\Big(1 + \frac{\lambda t_n}{A}\Big) - \frac{\lambda t_n}{A}\varepsilon.\end{aligned}$$

Finally (3.12) yields

$$\begin{aligned}\|x_n\| &\leq \inf_{\lambda\geq 0} \|x_{n-1} - \lambda z_n\|\\ &\leq 2\|x_{n-1}\|\Big(1 + \rho\Big(\frac{\lambda}{\|x_{n-1}\|}\Big)\Big) - \|x_{n-1} + \lambda z_n\|\\ &\leq \|x_{n-1}\|\left(1 + 2\rho\Big(\frac{\lambda}{\|x_{n-1}\|}\Big) - \frac{\lambda t_n}{A} + \frac{\lambda t_n}{A\|x_{n-1}\|}\varepsilon\right)\\ &= \|x_{n-1}\|\left(1 + 2\rho\Big(\frac{\lambda}{\|x_{n-1}\|}\Big) - \frac{\lambda t_n}{A}\Big(1 - \frac{\varepsilon}{x_{n-1}\|}\Big)\right)\end{aligned}$$

which proves our assertion. □

*Proof of Theorem 3.2.2.* Let $x_n = x - G^c_n$, for $n \in \mathbb{N}$. By construction, the sequence $(\|x_n\| : n \in \mathbb{N})$ is decreasing and thus, $\alpha = \lim_{n\to\infty} \|x_n\|$ exists and we have to show that $\alpha = 0$.

We assume that $\alpha > 0$ and will deduce a contradiction. Let $\varepsilon = \alpha/2$. Since $\text{span}(\mathcal{D})$ is dense in $X$ we find an $x_\varepsilon \in \text{span}(\mathcal{D})$, so that $\|x - x_\varepsilon\| < \varepsilon$ and denote $A = \|x_\varepsilon\|_{\mathcal{A}_1}$.

From Lemma 3.2.6 we deduce that

$$\|x_n\| \leq \|x_{n-1}\| \inf_{\lambda\geq 0} \Big[1 + 2\rho\Big(\frac{\lambda}{\|x_{n-1}\|}\Big) - \frac{\lambda t_n}{A}\Big(1 - \frac{\varepsilon}{\|x_{n-1}\|}\Big)\Big]$$

$$\leq \|x_{n-1}\| \inf_{\lambda\geq 0}\Big[1+2\rho\Big(\frac{\lambda}{\alpha}\Big)-\frac{\lambda t_n}{2A}\Big].$$

We let $\Theta=\alpha/8A$ and take $\lambda=\alpha\xi_n(\rho,\tau,\Theta)$ (recall that this means that $\rho(\xi_n)=\Theta t_n\xi_n$) and obtain that

$$\|x_n\|\leq\|x_{n-1}\|\Big[1+2\Theta t_n\xi_n-\frac{t_n}{2A}\Big]=\|x_{n-1}\|\Big[1-2\Theta t_n\xi_n\Big].$$

and thus

$$\|x_n\|=\|x\|+\sum_{j=1}^n\|x_j\|-\|x_{j-1}\|\leq\|x\|-\sum_{j=1}^n\|x_{j-1}\|2\Theta t_j\xi_j\leq\|x\|-\alpha\sum_{j=1}^n\|2\Theta t_j\xi_j.$$

But this contradicts our assumption that $\Sigma t_n\xi_n=\infty$. □

Before proving Theorem 3.2.3 we need one more Lemma.

**Lemma 3.2.7.** *Assume that $(a_n)$ and $(s_n)$ are sequences of positive numbers satisfying for some $A>0$ the following assumption*

$$a_1<\frac{A}{1+s_1}\ \textit{and}\ a_n\leq a_{n-1}\Big(1-\frac{s_n}{A}a_{n-1}\Big). \tag{3.13}$$

*Then it follows for all $n\in\mathbb{N}$*

$$a_n\leq A\Big(1+\sum_{j=1}^n s_j\Big)^{-1} \tag{3.14}$$

*Proof.* We prove(3.14) by induction. For $n=1$ this is the assumption. Assuming our claim is true for $n-1$ we obtain

$$\begin{aligned}a_n&\leq a_{n-1}\Big(1-\frac{s_n}{A}a_{n-1}\Big)\\&\leq A\Big(1+\sum_{j=1}^{n-1}s_j\Big)^{-1}\left(1-\frac{s_n}{1+\sum_{j=1}^{n-1}s_j}\right)\leq A\Big(1+\sum_{j=1}^n s_j\Big)^{-1}\end{aligned}$$

(last inequality follows from cross multiplication). □

*Proof of Theorem 3.2.3.* W.l.o.g. we assume $\|x\|_{\mathcal{A}_1}=1$. Let the sequences $(z_n)$ and $(G_n^c)=(G_n^c(x))$ be given as in (WCDGA) and let $x_n=x-G_n^c$, for $n\in\mathbb{N}$. By Lemma 3.2.6 with $\varepsilon=0$ we obtain

$$\|x_n\|\leq\|x_{n-1}\|\inf_{\lambda\geq0}\left[1-\lambda t_n+2\gamma\Big(\frac{\lambda}{\|x_{n-1}\|}\Big)^q\right] \tag{3.15}$$

We choose $\lambda$ such that
$$\frac{1}{2}\lambda t_m = 2\gamma\Big(\frac{\lambda}{\|x_{n-1}\|}\Big)^q,$$
or
$$\lambda = \|x_{n-1}\|^{q/(q-1)}(4\gamma)^{1/(q-1)}t_n^{1/(q-1)}.$$
Abbreviating $A_q = 2(4\gamma)^{1/(q-1)}$ and $p = q/(q-1)$, and inserting the choice of $\lambda$ into (3.15), we obtain
$$\|x_n\| \le \|x_{n-1}\|\Big(1 - \frac{1}{2}\lambda t_n\Big) = \|x_{n-1}\|\Big(1 - \|x_{n-1}\|^p t_n^p / A_q\Big).$$
Taking the $p$th power on each side and using the fact that $x \ge x^p$ if $0 < x \le 1$, we obtain
$$\|x_n\|^p \le \|x_{n-1}\|^p\Big(1 - \|x_{n-1}\|^p t_n^p / A_q\Big).$$
Since $\gamma \ge 1$ and thus $A_q > 2$, and since $\|x\| \le \|x\|_{\mathcal{A}_1} = 1$ we can apply Lemma 3.2.7 to $a_n = \|x_n\|^p$ and $s_n = t_n^p$, and $A = A_q$, and deduce that for all $n \in \mathbb{N}$
$$\|x_n\|^p \le A_q\Big(1 + \sum_{j=1}^n t_n^p\Big)^{-1},$$
which implies the claim of our Theorem. □

## 3.3 Weak Dual Greedy Algorithm with Relaxation

For the Weak Chebyshev Dual Greedy Algorithm we need to approximate $x$ by an element of the $n$ dimensional space $Z_n$, which might be computationally complicated. The following algorithm is a compromise between the Chebyshev Dual Greedy Algorithm and Dual Greedy Algorithm. Here we only have to find a good approximation to a two dimensional subspace:

(WDGAFR) *The weak dual greedy algorithm with free relaxation*

As usual we are given a sequence of weakness factors $\tau = (t_n) \subset (0,1]$ and a dictionary $\mathcal{D} \subset S_X$. For $x \in X$ we define $G_n^r = G_n^r(x)$ , $n \in \mathbb{N}$, as follows: $G_0^r = 0$ and assuming $G_{n-1}^r$ has been defined we choose $z_n \in \mathcal{D}$ so that
$$f_{x-G_{n-1}^r}(z_n) \ge t_n \sup_{z\in\mathcal{D}} f_{x-G_{n-1}^r}(z),$$
and then we let $w_n$ and $\lambda_n$ so that
$$\big\|x - (1-w_n)G_{n-1}^r - \lambda_n z_n\big\| \le \inf_{\lambda,w}\big\|x - (1-w)G_{n-1}^r - \lambda z_n\big\|.,$$
and define $G_n^r q = (1-w_n)G_{n-1}^r + \lambda z_n$.

**Proposition 3.3.1.** *For all $x \in X$ $\|x - G_n^r(x)\|$ is decreasing in $n \in \mathbb{N}$.*

We will need the following analog to Lemma 3.2.6

**Lemma 3.3.2.** *Assume that $X$ is a uniformly smooth Banach space and denote its modulus of uniform smoothness by $\rho$. Let $x \in X$, $\|x\| \geq \varepsilon \geq 0$ and $x_\varepsilon \in \mathcal{A}_1$, so that $\|x - x_\varepsilon\| < \varepsilon$. For $n \in \mathbb{N}$ put $x_n = x - G_n^f$*

*Then*

$$\frac{\|x_n\|}{\|x_{n-1}\|} \leq \inf_{\lambda \geq 0} \left(1 - \frac{\lambda t_n}{\|x_\varepsilon\|_{\mathcal{A}_1}}\Big(1 - \frac{\varepsilon}{\|x_{n-1}\|}\Big) + 2\rho\Big(\frac{5\lambda}{\|x_{n-1}\|}\Big)\right). \tag{3.16}$$

*Proof.* From the definition of $\rho$ we deduce for $w \in \mathbb{R}$ and $\lambda \geq 0$ that

$$\begin{aligned} &\|x_{n-1} + wG_{n-1}^{dr} - \lambda z_n\| + \|x_{n-1} - wG_{n-1}^{dr} + \lambda z_n\| \\ &\qquad \leq 2\|x_{n-1}\|\Big(1 + \rho\Big(\frac{\|wG_{n-1}^{dr} - \lambda z_n\|}{\|x_{n-1}\|}\Big)\Big). \end{aligned} \tag{3.17}$$

For all $w \in \mathbb{R}$ and $\lambda \geq 0$ we estimate

$$\begin{aligned} \|x_{n-1} - &wG_{n-1}^{dr} + \lambda z_n\| \\ &\geq f_{x_{n-1}}(x_{n-1} - wG_{n-1}^{dr} + \lambda z_n) \\ &\geq \|x_{n-1}\| - f_{x_{n-1}}(wG_{n-1}^{dr}) + \lambda t_n \sup_{z \in \mathcal{D}} f_{x_{n-1}}(z) \\ &= \|x_{n-1}\| - f_{x_{n-1}}(wG_{n-1}^{dr}) + \lambda t_n \sup_{z \in \mathcal{A}_1, \|z\|_{\mathcal{A}_1} \leq 1} f_{x_{n-1}}(z) \\ &\text{(By Lemma 3.2.5)} \\ &\geq \|x_{n-1}\| - f_{x_{n-1}}(wG_{n-1}^{dr}) + \frac{\lambda t_n}{\|x_\varepsilon\|_{\mathcal{A}_1}} f_{x_{n-1}}(x_\varepsilon) \\ &= \|x_{n-1}\| - f_{x_{n-1}}(wG_{n-1}^{dr}) + \frac{\lambda t_n}{\|x_\varepsilon\|_{\mathcal{A}_1}} f_{x_{n-1}}(x) - \frac{\lambda t_n \varepsilon}{\|x_\varepsilon\|_{\mathcal{A}_1}}. \end{aligned} \tag{3.18}$$

Letting $w^* = \lambda t_n / \|x_\varepsilon\|_{\mathcal{A}_1}$ we deduce that

$$\begin{aligned} \|x_{n-1} - &w^* G_{n-1}^{dr} + \lambda z_n\| \\ &\geq \|x_{n-1}\| + \frac{\lambda t_n}{\|x_\varepsilon\|_{\mathcal{A}_1}} f_{x_{n-1}}(x - G_{n-1}^{dr}) - \frac{\lambda t_n \varepsilon}{\|x_\varepsilon\|_{\mathcal{A}_1}} \\ &\geq \|x_{n-1}\| + \frac{\lambda t_n}{\|x_\varepsilon\|_{\mathcal{A}_1}} \|x_{n-1}\| - \frac{\lambda t_n \varepsilon}{\|x_\varepsilon\|_{\mathcal{A}_1}}. \end{aligned} \tag{3.19}$$

Thus we obtain

$$\|x_n\| = \inf_{\lambda \geq 0, w \in \mathbb{R}} \|x_{n-1} + wG_n^{dr} - 1 - \lambda z_n\|$$

$$
\begin{aligned}
&\le \inf_{\lambda\ge 0, w\in\mathbb{R}} \Big[2\|x_{n-1}\|\Big(1+\rho\Big(\frac{\|wG^{dr}_{n-1}-\lambda z_n\|}{\|x_{n-1}}\Big)\Big) - \|x_{n-1} - wG^{dr}_{n-1} + \lambda z_n\|\Big] \\
&\le \inf_{\lambda\ge 0} \Big[2\|x_{n-1}\|\Big(1+\rho\Big(\frac{\|w^*G^{dr}_{n-1}-\lambda z_n\|}{\|x_{n-1}\|}\Big)\Big) \\
&\qquad - \Big(\|x_{n-1}\| + \frac{\lambda t_n}{\|x_\varepsilon\|_{\mathcal{A}_1}}\|x_{n-1}\| - \frac{\lambda t_n \varepsilon}{\|x_\varepsilon\|_{\mathcal{A}_1}}\Big)\Big] \\
&= \|x_{n-1}\| \inf_{\lambda\ge 0}\Bigg[1 - \frac{\lambda t_n}{\|x_\varepsilon\|_{\mathcal{A}_1}}\Big(1-\frac{\varepsilon}{\|x_{n-1}\|}\Big) + 2\rho\Big(\frac{\|w^*G^{dr}_{n-1}-\lambda z_n\|}{\|x_{n-1}\|}\Big)\Bigg]
\end{aligned}
$$

In order to achieve (3.16) we need to estimate $\|w^*G^{dr}_{n-1} - \lambda z_n\|$. First we note that

$$\|G^{dr}_{n-1}\| = \|x - x_{n-1}\| \le 2\|x\| \le 2\|x_\varepsilon\|_{\mathcal{A}_1} + 2\varepsilon \le 4\|x_\varepsilon\|_{\mathcal{A}_1}$$

and thus

$$\|w^*G^{dr}_{n-1} - \lambda z_n\| \le 4w^*\|x_\varepsilon\|_{\mathcal{A}_1} + \lambda \le 5\lambda,$$

which implies our claim since $\rho(\cdot)$ is increasing on $[0,\infty)$. □

**Remark.** Before stating the next Theorem let as note that if $\rho$ is an even and convex function on $\mathbb{R}$, with $\rho(0) = 0$, $\lim_{u\to 0}\rho(u)/u = 0$, then the function $s : u \mapsto \rho(u)/u$ is increasing on $[0,\infty)$, thus has a inverse function $s^{-1}$ which is also increasing and $s^{-1}(0) = 0$.

**Theorem 3.3.3.** *Assume that $X$ is a separable and uniformly smooth Banach space, and denote its modulus of uniform smoothness by $\rho$. Let $s^{-1}(\cdot)$ be the inverse function of $s : u \mapsto \rho(u)/u$, $u \ge 0$. We consider the (WDGAFR) with a sequence of weakness factors $\tau = (t_n) \subset (0,1]$ satisfying*

$$\sum_{n=1}^{\infty} t_n s^{-1}(\Theta t_n) = \infty \text{ for all } \Theta > 0 \tag{3.20}$$

*Then for any $x \in X$*

$$\lim_{n\to\infty} \|x - G^{dr}_n(x)\| = 0. \tag{3.21}$$

*Proof.* Let $G^{dr}_n = G^{dr}_n(x)$ and $x_n = x - G^{dr}_n$, for $n \in \mathbb{N}$. Since $\|x_n\|$ decreases

$$\beta = \lim_{n\to\infty} \|x_n\|$$

exists and we need to show that $\beta = 0$.

We assume that $\beta > 0$ and will derive a contradiction. We set $\varepsilon = \beta/2$ and choose $x_\varepsilon \in \mathcal{A}_1(\mathcal{D})$ with $\|x - x_\varepsilon\| < \varepsilon$. Note that $\|x\| \ge \beta\varepsilon$.

By Lemma 3.3.2 we have

$$\|x_n\| \le \|x_{n-1}\| \inf_{\lambda\ge 0}\Big(1 - \frac{\lambda t_n}{2A} + 2\rho\Big(\frac{5\lambda}{\beta}\Big).$$

Putting $\Theta = \beta/40A$ and $\lambda = \beta s^{-1}(\Theta t_m)/5$, we obtain

$$\begin{aligned}
\|x_n\| &\leq \|x_{n-1}\|\Big(1 - \frac{\beta t_n s^{-1}(\Theta t_n)}{10A} + 2\rho\big(s^{-1}(\Theta t_n)\big)\Big) \\
&= \|x_{n-1}\|\Big(1 - \frac{\beta t_n s^{-1}(\Theta t_n)}{10A} + 2s^{-1}(\Theta t_n)\frac{\rho\big(s^{-1}(\Theta t_n)\big)}{s^{-1}(\Theta t_n)}\Big) \\
&= \|x_{n-1}\|\Big(1 - \frac{\beta t_n s^{-1}(\Theta t_n)}{10A} + 2s^{-1}(\Theta t_n)\frac{\rho\big(s^{-1}(\Theta t_n)\big)}{s^{-1}(\Theta t_n)}\Big) \\
&= \|x_{n-1}\|\big(1 - 4\Theta t_n s^{-1}(\Theta t_n) + 2s^{-1}(\Theta t_n)\Theta t_n\big) = \|x_{n-1}\|\big(1 - 2\Theta t_n s^{-1}(\Theta t_n)\big).
\end{aligned}$$

We can iterate this inequality and obtain

$$\begin{aligned}
\|x_n\| &\leq \|x_{n-1} - \|x_{n-1}\| 2\Theta t_n s^{-1}(\Theta t_n) \\
&\leq \|x_{n-1}\| - \beta 2\Theta t_n s^{-1}(\Theta t_n) \\
&\leq \|x_{n-2}\| - \beta 2\Theta t_{n-1} s^{-1}(\Theta t_{n-1}) - \beta 2\Theta t_n s^{-1}(\Theta t_n) \\
&\vdots \\
&\leq \|x\| - \beta 2\Theta t_{n-1} \sum_{j=1}^{n} t_j s^{-1}(\Theta t_j)
\end{aligned}$$

and thus letting $n \to \infty$

$$\beta \leq \|x\| - \beta 2\Theta t_{n-1} \sum_{j=1}^{\infty} t_j s^{-1}(\Theta t_j),$$

which is the contradiction since we assumed that $\sum_{j=1}^{n} t_j s^{-1}(\Theta t_j) = \infty$. □

There is also a result on the rate of the convergence in Theorem 3.3.3 Since the proof is similar to the proof of the corresponding result for the Chebyshev Greedy Algorithm we omit a proof.

**Theorem 3.3.4.** *Let $X$ be a uniformly smooth Banach space with modulus of uniform smoothness $\rho$, which satisfies for some $q \in (1, 2]$*

$$\rho(u) \leq \gamma u^q. \tag{3.22}$$

*Then there is a number $C$ only depending on $q$ and $\gamma$ so that the following holds.*

*If $x \in X$ and if $\varepsilon > 0$ and $x_\varepsilon \in \mathcal{A}_1$ so that $\|x - x_\varepsilon\| < \varepsilon$, and if $\tau = (t_n) \subset (0, 1]$, then it follows for the (WGGAFR) $(G_n^{dr}) = (G_n^{dr}(x))$ with weakness factors $(t_n)$ that*

$$\|x - G_n^{dr}\| \leq \max\left(2\varepsilon, C(\|x_\varepsilon\|_{\mathcal{A}_1} + \varepsilon)\Big(1 + \sum_{j=1}^{n} t_j^p\Big)^{-1/p}\right) \tag{3.23}$$

*where* $p = q/(q-1)$.

**Remark.** We like to point out something about the proof of Theorem 3.3.4 which will be useful when we consider the next greedy algorithm.

In the proof of Theorem 3.3.4 It was only used that the sequence $(\|x_n\| : n \in \mathbb{N})$ is decreasing and that the inequality 3.16 holds. Of course in order to proof this inequality we needed the specific choice of $z_n$, namely in the second "$\geq$" of (3.18).

Thus if $G_n(x)$ is any algorithm so that $\|x - G_n(x)\|$ is decreasing which satisfies equation 3.16 the $G_n(x)$ converges to $x$.

Keeping that remark in mind we now turn to the *X-Greedy Algorithm with Free Relaxation.*

(XGAFR) the *X-Greedy Algorithm with Free Relaxation* As usual we are given a dictionary $\mathcal{D} \subset S_X$. For $x \in X$ we define $G^r_n = G^r_n(x)$ , $n \in \mathbb{N}$, as follows: $G^r_0 = 0$ and assuming $G^r_{n-1}$ has been defined we choose $\lambda_n \geq 0$, $w_n \in \mathbb{R}$ , $z_n \in \mathcal{D}$ so that

$$\|x - (1-w_m)G^r_{n-1} - \lambda_n z_n\| = \inf_{z\in\mathcal{D},\lambda\geq 0, w\in\mathbb{R}} \|x - (1-w)G^r_{n-1} - \lambda z\|$$

and then define

$$G^r_n q = (1 - w_n)G^r_{n-1} + \lambda z_n.$$

We notice that at each step the value of

$$\frac{\|x - G^r n(x)\|}{\|x - G^r n - 1(x)\|}$$

is at most as large as the value we would have obtained if we had computed $G^{dr}_n(x)$ form $x_{n-1}$. We deduce therefore that the conclusion of Lemma still holds and that if $0 \leq \varepsilon \leq \|x\|$ and if $x_\varepsilon \in \mathcal{A}_1$ with $\|x - x_\varepsilon\| \leq \varepsilon$ it follows that

$$\frac{\|x - G^r_n(x)\|}{\|x - G^r_{n-1}(x)\|} \leq \inf_{\lambda\geq 0}\left(1 - \frac{\lambda t_n}{\|x_\varepsilon\|_{\mathcal{A}_1}}\Big(1 - \frac{\varepsilon}{\|x_{n-1}\|}\Big) + 2\rho\Big(\frac{5\lambda}{\|x_{n-1}\|}\Big)\right). \tag{3.24}$$

From the previous made remark we deduce therefore the following convergence result.

**Theorem 3.3.5.** *Assume that* $X$ *is a separable and uniformly smooth Banach space. We consider the (XGAFR)* $(G^r_n(x) : n \in \mathbb{N})$, *for* $x \in X$

*Then for any* $x \in X$

$$\lim_{n\to\infty} \|x - G^r_n(x)\| = 0. \tag{3.25}$$

## 3.4 Convergence Theorem for the Weak Dual Algorithm

For a Banach space $X$ assume that $f_{(\cdot)}: X\setminus\{0\}\to S_{X^*}$ *support map*, i.e. for every $x\in X\setminus\{0\}$ we have $f_x(x)=\|x\|$. Recall from the the remark after Definition 3.1.2 that every $x\in X\setminus\{0\}$ has a unique support map $f_x$ if and only if the norm is Gateaux differentiable

$$f_{x_0}(y)=\frac{1}{\|x_0\|}\frac{\partial}{\partial\lambda}\|x_0+\lambda y\|\big|_{\lambda=0}=\frac{1}{\|x_0\|}\lim_{h\to\infty}\frac{\|x_0+hy\|-\|x_0\|}{h}.$$

for all $y\in S_X$

Let $\mathcal{D}\subset S_X$ be a dictionary for $X$ and put for $x\in X$:

$$\rho(x)=\rho_D(x)=\sup_{z\in\mathcal{D}}f_x(z).$$

We consider the *Weak Dual Greedy Algorithm* with fixed weakness factor as in Section 3.1 but slightly reformulated:

(WDGA) Fix $c\in(0,1)$. For $x\in X$ we choose sequences $(x_n)_{n\geq 0}$, $(z_n)_{n\geq 1}\subset\mathcal{D}$ and $(t_n)\subset[0,\infty)$ recursively as follows.

$x_0=x$ and assuming that $x_{n-1}\in X$, has been chosen we choose $z_n\in\mathcal{D}$ arbitrary and $t_n=0$ if $x_{n-1}=0$, and otherwise we choose $z_n\in\mathcal{D}$ and $t_n\geq 0$ so that

a) $f_{x_{n-1}}(z_n)\geq c\rho_{\mathcal{D}}(x_{n-1})=c\sup_{z\in\mathcal{D}}f_{x_{n-1}}(z)$,

b) $\|x_{n-1}-t_nz_n\|=\min_{t\geq 0}\|x_{n-1}-tz_n\|$,

and in both cases we finally let

c) $G_n=G_{n-1}+t_nz_n$ and $x_n=x-G_n=x_{n-1}-t_nz_n$.

We say that the weak dual greedy algorithm *converges for $\mathcal{D}$*, if for all $x\in X$,

$$\lim_{n\to\infty}x_n=0 \text{ or, equivalently, } \lim_{n\to\infty}\sum_{i=1}^n t_iz_i=x.$$

**Lemma 3.4.1.** *Let $X$ have Gateaux differentiable norm, $0<c<1$, and assume that for $x\in X$ the sequences $(x_n)$, $(t_n)$, and $(z_n)$ satisfy (a) and (c) of (WDGA), but instead of (b) the following condition*

*d)* $\dfrac{\|x_{n-1}\|-\|x_n\|}{t_n}\geq c\rho(x_{n-1})$ *for all $n\in\mathbb{N}$.*

*Then, if $\sum_{n=1}^\infty t_n=\infty$, we have $x=\sum_{n=1}^\infty t_nz_n$.*

*Proof.* Define $s_n = \sum_{i=1}^n t_i$, for $n \in \mathbb{N}$

Then

$$e^{\sum_{j=2}^n \ln((s_j-t_j)/s_j)} = \prod_{j=2}^n \frac{s_{j-1}}{s_j} = \frac{s_1}{s_n} \to_{n\to\infty} 0,$$

and thus

$$\lim_{n\to\infty} \sum_{j=2}^n \ln((s_j - t_j)/s_j) = -\infty$$

It follows that

$$\infty = -\sum_{j=2}^\infty \ln((s_j - t_j)/s_j) = -\sum_{j=2}^n \ln\Big(1 - \frac{t_j}{s_j}\Big) \le \sum_{j=2}^\infty \frac{t_j}{s_j} + \frac{t_j^2}{s_j^2},$$

and thus $\sum_{j=2}^\infty \frac{t_j}{s_j} = \infty$.

We note that if $(a_n)$ and $(b_n)$ are two positive sequences and $\sum a_n < \infty$, while $\sum b_n = \infty$, then there is a subsequence $(n_k)$ of $\mathbb{N}$, so that $\lim_{k\to\infty} a_{n_k}/b_{n_k} = 0$. Indeed, for every $k \in \mathbb{N}$ the set $N_k = \{n \in \mathbb{N} : ka_n < b_n\}$ must be infinite, and we can therefore choose $n_1 < n_2 < n_3 < \ldots$, with $a_{n_k} \in N_k$, for $k \in \mathbb{N}$.

Thus we can find $n_1 < n_2 < n_3 < \ldots$ so that

$$\lim_{k\to\infty} \frac{s_{n_k+1}(\|x_{n_k}\| - \|x_{n_k+1}\|)}{t_{n_k+1}} = 0.$$

It follows that

$$0 \le s_{n_k}\rho_D(x_{n_k}) \le \frac{1}{C}\frac{s_{n_k}(\|x_{n_k}\| - \|x_{n_k+1}\|)}{t_{n_k+1}} \le \frac{1}{C}\frac{s_{n_k+1}(\|x_{n_k}\| - \|x_{n_{k+1}}\|)}{t_{n_k+1}} \to_{k\to\infty} 0,$$

and thus, in particular,

$$\lim_{k\to\infty} \rho_D(x_{n_k}) = 0. \tag{3.26}$$

For $1 \le l \le n_k - 1$ we have:

(3.27)

$$\big|\|x_{n_k}\| - f_{x_{n_k}}(x_l)\big| = \Big|f_{x_{n_k}}\Big(\sum_{j=l+1}^{n_k} t_j z_j\Big)\Big| \le \sum_{j=1}^{n_k} t_j \rho_D(x_{n_k}) = s_{n_k}\rho_D(x_{n_k}) \to_{k\to\infty} 0.$$

Now assume that $x^*$ is a $w^*$-cluster point of the sequence $(f_{x_{n_k}})$ and let $L = \lim_{n\to\infty}\|x_n\|$.Then, by (3.27), it follows that $x^*(x_l) = L$ for all $l \in \mathbb{N}$. We now claim that this implies that $L = 0$. Indeed, other wise, $x^* \ne 0$, and thus, since $\overline{\text{span}(\mathcal{D})} = X$, and $\mathcal{D} = -\mathcal{D}$, $\theta = \sup_{z\in\mathcal{D}} x^*(z) > 0$, and thus

$$\limsup_{k\in\mathbb{N}} \sup_{z\in\mathcal{D}} f_{x_{n_k}}(z) \ge \theta,$$

which contradicts (3.26). □

**Lemma 3.4.2.** *Suppose* $1 < p < \infty$. *Then there is a* $C_p > 0$ *such that for any* $a, b \in \mathbb{R}$

$$b|a+b|^{p-1}\mathrm{sign}(a+b) - b|a|^{p-1}\mathrm{sign}(a) \le C_p\big(|a+b|^p - pb|a|^{p-1}\mathrm{sign}(a) - |a|^p\big). \tag{3.28}$$

*Proof.* First, note that replacing $a$ and $b$ simultaneously by $-a$ and $-b$, the inequality does not change. We can therefore assume that $a \ge 0$. Also note that for $b = 0$ we have equality if we let $C_p = 1$. Thus, we can assume that $a > 0$ and, since both sides of (3.28) are $p$-homogenous, we can assume that $a = 1$, and also that $b \neq 0$. We need therefore to show that

$$\phi(b) = \frac{b\big(|1+b|^{p-1}\mathrm{sign}(1+b) - 1\big)}{|1+b|^p - pb - 1}, \quad b \neq 0,$$

has an upper bound.

We note that

$$\begin{aligned}
\lim_{b\to 0}\phi(b) &= \lim_{b\to 0}\frac{b(1+b)^{p-1} - b}{(1+b)^p - pb - 1} \\
&= \lim_{b\to 0}\frac{(1+b)^{p-1} + (p-1)b(1+b)^{p-2}}{p(1+b)^{p-1} - p} \\
&= \lim_{b\to 0}\frac{(p-1)(1+b)^{p-2} + (p-1)(1+b)^{p-2} + (p-1)(p-2)b(1+b)^{p-3}}{p(p-1)(1+b)^{p-2}} = \frac{2}{p}
\end{aligned}$$

and

$$\lim_{b\to\pm\infty}\phi(b) = 1,$$

which implies the claim since $\phi$ is continuous. □

**Definition 3.4.3.** A Banach space $X$ with Gateaux differentiable norm is said to *have property* $\Gamma$ if there is a constant $0 < \gamma \le 1$ so that for $x, y \in X$ for which $f_x(y) = 0$ it follows that

$$\|x + y\| \ge \|x\| + \gamma f_{x+y}(y).$$

**Remark.** For $x, y \in L_p[0,1]$, $f_x(y) = 0$ means that

$$\int_0^1 \mathrm{sign}(x(t))|x(t)|^{p-1}y(t)\,dt = \int_0^1 \mathrm{sign}(x(t))|x(t)|^{p/q}y(t)\,dt = 0. \tag{3.29}$$

**Proposition 3.4.4.** *If* $1 < p < \infty$, *every quotient of* $L_p[0,1]$ *has property* $\Gamma$.

*Proof.* We first show that $L_p[0,1]$ itself has property $\Gamma$. So let $x, y \in L_p[0,1]$ with $f_x(y) = 0$. We can assume that $y \neq 0$, and, after dividing $x$ and $y$ by $\|y\|$, that $\|y\| = 1$

By Lemma 3.4.2,

$$\begin{aligned} &y(s)|x(s)+y(s)|^{p-1}\text{sign}(x(s)+y(s)) \\ &\quad \leq C_p\big(|x(s)+y(s)|^p - |x(s)|^p\big) + (1-pC_p)y(s)|x(s)|^{p-1}\text{sign}(x(s)). \end{aligned}$$

Integrating both sides and using (3.29), yields

$$\int_0^1 y(s)|x(s)+y(s)|^{p-1}\text{sign}(x(s)+y(s))\,ds \leq C_p\big(\|x+y\|_p^p - \|x\|_p^p\big).$$

It follows from the fact that $f_x(y)$ is a positive multiple of $\frac{d}{dt}\|x+ty\|_p$, and the convexity of the function $t \mapsto \|x+ty\|$ that $\|x\|_p \leq \|x+y\|_p$. Moreover we have $f_z = \|z\|_p^{1-p}\text{sign}(z(\cdot))|z(\cdot)|^{p-1} \in S_{L_q}$, for $z \in L_p[0,1]$, and thus

$$\begin{aligned} \|x+y\|_p^{p-1} f_{x+y}(y) &= \|x+y\|_p^{p-1}\int_0^1 y(s)|x(s)+y(s)|^{p-1}\text{sign}(x(s)+y(s))\,ds \leq \\ &\leq C_p\big(\|x+y\|_p^p - \|x\|_p^p\big) \\ &= C_p\big(\|x+y\| - \|x\|\big)\frac{d}{dt}\|x+ty\|_p^p\Big|_{t=t_0} \\ &\text{(By Taylor's Theorem for some } t_0 \in (0,1)) \\ &= C_p\big(\|x+y\| - \|x\|\big)p\|x+y\|_p^{p-1}\frac{d}{dt}\|x+ty\|_p\Big|_{t=t_0} \\ &\leq C_p\big(\|x+y\| - \|x\|\big)p\|x+y\|_p^{p-1} \\ &\Big(0 \leq \frac{d}{dt}\|x+ty\|_p\Big|_{t=t_0} \leq 1, \text{ since } \|x\|_p \leq \|x+y\|_p \text{ and since } \|y\|_p = 0\Big), \end{aligned}$$

which proves our claim if we let $\gamma = 1/pC_p$. From the following more general Proposition it will follow that every quotient of $L_p[0,1]$ has property $\Gamma$. $\square$

**Proposition 3.4.5.** *The quotient of a reflexive space $X$ with property $\Gamma$ and Gateaux differentiable norm also has property $\Gamma$ (with respect to the same constant $\gamma$).*

*Proof.* Assume that $Y = X/Z$, where $X$ is a reflexive space with property $\Gamma$ and $Z \subset X$ is a closed subspace of $X$.

Let $x, y \in X$, let $\overline{x} = x + Z$, $\overline{y} = y + Z$ be the images under the quotient map $Q : X \to Y$. Since $X$ is reflexive we can assume that $\||xb\|_{X/Z} = \inf_{\tilde{x}\in X+Z}\|\tilde{x}\|_X$ and find an element $w \in X$ so that $\|\overline{x}+\overline{y}\|_{X/Z} = \|w\|_X$.

Note that $f_x = f_{\overline{x}} \circ Q$ (since $f_{\overline{x}}(Q(x)) = f_{\overline{x}}(\overline{x}) = \|\overline{x}\|_{X/Y} = \|x\|_X$) and $f_w = f_{\overline{x}+\overline{y}} \circ Q$. Hence if $f_{\overline{x}}(\overline{y}) = 0$ it follows that $f_x(u-x) = 0$ and thus

$$\|\overline{x}+\overline{y}\| = \|w\| = \|x+(w-x)\| \geq \|x\| + \gamma f_{x+(w-x)}(w-u) = \|x\| + \gamma f_{\overline{x}+\overline{y}}(\overline{y}).$$

□

Now we are ready to show the final result which implies in particular that the (WDGA) converges in $L_p[0,1]$ for any dictionary $\mathcal{D}$.

**Theorem 3.4.6.** *Suppose $X$ is s Banach space with property $\Gamma$ and Fréchet differentiable norm. If $\mathcal{D}$ is a dictionary of $X$ and $0 < c \leq 1$ then the (WDGA) converges.*

*Proof.* By Proposition 4.2.2 (Class Notes in Functional Analysis), the map $x \mapsto f_x$ is a norm continuous map between $X \setminus \{0\}$ and $S_{X^*}$.

Let $x = x_0$ be in $X$ and let $(x_n)$, $(z_n)$ and $(t_n)$ as in (WDGA). If $t_n = 0$ for some $n \in \mathbb{N}$ then $\rho_D(x_n) = 0$, and since $\mathcal{D}$ is total it follows that $x_{n-1} = 0$ and thus $x_k = 0$, for all $k \geq n$. So we can assume without loss of generality that $t_n > 0$ for all $n \in \mathbb{N}$.

By condition (b) it follows that

$$\frac{d}{dt}\|x_{n-1} - tz_n\||_{t=t_n} = 0,$$

and thus $f_{x_n}(z_n) = 0$, which yields using property $\Gamma$, that

$$\|x_{n-1}\| = \|x_n + t_n z_n\| \geq \|x_n\| + \gamma t_n f_{x_{n-1}}(z_n)$$

and thus

$$\frac{\|x_{n-1}\| - \|x_n\|}{t_n} \geq \gamma f_{x_{n-1}}(z_n) \geq c\gamma\rho_{\mathcal{D}}(x_{n-1}).$$

Using Lemma 3.4.1 with $c\gamma$ instead of $\gamma$ we only need to show that $\lim_{n\to\infty} \|x_n\| = 0$ if $\sum_{n=1}^{\infty} t_n < \infty$.

If $\sum_{n=1}^{\infty} t_n < \infty$ then $(x_n)$ converges to some $x_\infty \in X$. To deduce a contradiction assume that $x_\infty \neq 0$ which implies that $\lim_{n\to\infty} \|f_{x_n} - f_{x_\infty}\| = 0$. Now since, as observed previously, $f_{x_n}(z_n) = 0$, we have that $\lim_{n\to\infty} f_{x_{n-1}}(z_n) = 0$, and thus by (WDGA)(a) $\lim_{n\to\infty} \rho_{\mathcal{D}}(x_{n-1}) = 0$, and thus for any $z \in \mathcal{D}$

$$f_{x_\infty}(z) = \lim_{n\to\infty} f_{x_{n-1}}(z) \leq \lim_{n\to\infty} \rho_{\mathcal{D}}(x_{n-1}) = 0$$

and similarly, since $\mathcal{D} = -\mathcal{D}$,

$$-f_{x_\infty}(z) = f_{x_\infty}(-z) = \lim_{n\to\infty} f_{x_{n-1}}(-z) \leq \lim_{n\to\infty} \rho_{\mathcal{D}}(x_{n-1}) = 0$$

which implies that $f_{x_\infty}(z) = 0$ for all $z \in \mathcal{D}$, and thus $x_\infty = 0$, which is a contradiction and proves our claim. □

**Remark.** Assume that the Banach space $X$ is Gateaux differentiable and let $x \in X$. As in (WXGA) $(x_n) \subset X$, $(z_n) \subset \mathcal{D}$ and $(t_n) \in [0,\infty)$ are chosen so that $x_0 = x$, $x_n = x_{n-1} - t_n z_n$, for $n \in \mathbb{N}$, and so that for some $c \in (0,1]$

$$\|x_{n-1}\| - \|x_n\| \geq c \sup_{g\in\mathcal{D}} \sup_{t\geq 0} \big(\|x_{n-1}\| - \|x_{n-1} - tg\|\big).$$

We note that, if $(x_n)$ has a convergent subsequence (for example if $\sum_n t_n < \infty$), then $(x_n)$ has to converge to 0.

Indeed, assume that $x_\infty = \lim_{k\to\infty} x_{n_k}$ exists for some subsequence $(n_k)$. We claim $x_\infty = 0$, and this would imply that $(x_n)$ converges to 0, since $(\|x_n\|)$ is decreasing.

Assume that $x_\infty \neq 0$. Then, since $\mathcal{D}$ is total in $X$, it follows that $\sup_{g\in\mathcal{D}} f_{x_\infty} > 0$, and thus there exists a $g \in cD$ and a $t > 0$ so that $\varepsilon = \|x_\infty\| - \|x_\infty - tg\| > 0$. But this would imply that for some $k_0 \in \mathbb{N}$

$$\|x_{n_k}\| - \|x_{n_{k+1}}\| \geq c\big(\|x_{n_k}\| - \|x_{n_k} - tg\|\big) \geq c\varepsilon/2, \text{ whenever } k \geq k_0.$$

But this is a contradiction since $(\|x_n\| - \|x_{n+1}\|)$ is a non negative and summable sequence.

# Chapter 4

# Open Problems

## 4.1 Greedy Bases

**Problem 4.1.1.** *Does every infinite dimensional Banach space contain a quasi greedy basis ?*

**Comments to Problem 4.1.1:** First of all there are separable Banach spaces which have a basis but do not have quasi greedy bases (for the whole spaces). Indeed in [DKK] it was shown that a $\mathcal{L}_\infty$-space (for example any $C(K)$, $K$ compact) which is not isomorphic to $c_0$ does not have a quasi greedy basis.

Secondly Gowers and Maurey solved the *unconditional basis problem* and showed that not every separable Banach space contains a unconditional basic sequence (which is stronger than being quasi greedy). Nevertheless, in [DKK], Dilworth, Kalton and Kutzarova proved that all known counterexamples to the unconditional basis problem actually contain quasi greedy basic sequences.

The showed the following

**Theorem 4.1.2.** *Let $(x_n)$ be a semi normalized weakly null sequence in a Banach space $X$ with spreading model $(e_n)$, and suppose that $(e_n)$ has the property that*

$$\Big\|\sum_{i=1}^n e_i\Big\| \to \infty, \text{if } n \to \infty.$$

*Then $(x_n)$ has a subsequence which is quasi greedy and whose quasi greedy constant does not exceed $3+\varepsilon$ (for given $\varepsilon > 0$).*

Here the *spreading model* of semi normalized sequences $(x_n) \subset X$ is defined as follows:

Assume that for all $k \in \mathbb{N}$ and all scalars $(a_j)_{j=1}^n$

$$\Big\|\sum_{j=1}^n a_j e_j\Big\| = \lim_{n_1\to\infty} \lim_{n_2\to\infty} \ldots\ldots \lim_{n_k\to\infty} \Big\|\sum_{j=1}^n a_j x_j\Big\|$$

exists. It is clear that $\|\!|\cdot|\!\|$ is a semi norm on $c_{00}$. Using *Ramsey's Theorem* (some kind of generalized pigeon whole principle) one can prove that every semi normalized sequence has a subsequence so that above limit exists for all $(a_j) \in c_{00}$, and that if $(x_n)$ is weakly null or basic, then $\|\!|\cdot|\!\|$ is a norm on $c_{00}$ and $(e_n)$ is a basis of the completion of $c_{00}$ with respect to $\|\!|\cdot|\!\|$. We call in this case this completion together with its basis $(e_n)$ *the spreading model of* $(x_n)$.

For more comments on the problem and its relation to the problem whether or not the *Elton number* has universal upper bound see also [DOSZ2].

**Problem 4.1.3.** *Does $\ell_p(\ell_q)$, $1 < p, q < \infty$ and $p \neq q$, have a greedy basis?*

**Comments to Problem 4.1.3** Besov spaces are function spaces defined on the real line or on some subset of it and are of importance for example in Partial Differential Equations and Approximation Theory. It can be shown that Besov spaces of function defined on $\mathbb{R}$ are isomorphic to $\ell_p(\ell_q)$, $1 < p \neq q < \infty$, where

$$\ell_p(\ell_q) = \Big\{(x_n) : x_n \in \ell_q, \text{ for } n \in \mathbb{N}, \text{ and } \sum_n \|x_n\|_q^p < \infty\Big\},$$

with the norm

$$\|(x_n)\| = \left(\sum_n \|x_n\|_q^p\right)^{1/p}, \text{ for } (x_n) \in \ell_p(\ell_q).$$

It was shown in [EW] that for the space $\ell_p \oplus \ell_q$, $p \neq q$, every unconditional basis $(x_n)$ of $\ell_p \oplus \ell_q$ *splits in a basis of* $\ell_p$ *and in a basis of* $\ell_q$, *i.e.* there is a partition of $\mathbb{N}$ into $N_1$ and $N_2$, so that $(x_n : n \in N_1)$ is a basis of $\ell_p$ and $(x_n : n \in N_2)$ is a basis of $\ell_q$. From that result it is easy to see that $\ell_p \oplus \ell_q$ cannot have a greedy basis. On the other hand the space $\big(\oplus_{n=1}^\infty \ell_q^{m_n}\big)_p$, with $1 < p < \infty$ and $1 \leq q \leq \infty$, and $m_n \to \infty$, for $n \to \infty$, which is isomorphic to Besov spaces of function defined on the torus (or any closed bounded interval in $\mathbb{R}$) has a greedy basis [DFOS]. In the case $p = 1, \infty$ it was shown in [BCLT] that $\big(\oplus_{n=1}^\infty \ell_2^{m_n}\big)_p$ has a unique unconditional basis up to permutation, and thus this space cannot have a greedy basis (the usual one is clearly not democratic). From the proof in [BCLT] we can also deduce that for general $q \in [1, \infty]$ the spaces $\big(\oplus_{n=1}^\infty \ell_q^{m_n}\big)_1$ and $\big(\oplus_{n=1}^\infty \ell_q^{m_n}\big)_{c_0}$ have no greedy bases, unless, of course, in the trivial case that $p = q = \infty$ or $p = q = 1$.

**Problem 4.1.4.** *Given any $\varepsilon > 0$, can a Banach space with normalized a greedy basis $(e_n)$ be renormed so that $(e_n)$ is $(1+\varepsilon)$-greedy?*

**Comments to 4.1.4:** First Albiac and Wojtaszczyk [AW] asked whether or not every Banach space with normalized a greedy basis $(e_n)$ can be renormed so that $(e_n)$ is 1-greedy. This was solved negatively (recall that by Theorem 1.1.9 every 1-greedy basis must be 1-democratic) in [DOSZ2] where the following was shown:

**Proposition 4.1.5.** *Assume that $X$ is a Banach space with a normalized suppression 1-unconditional basis $(e_i)$ and that there is a sequence $(\rho_n) \subset (0,1]$ with $\rho = \inf_{n\in\mathbb{N}} \rho_n > 0$ so that*

$$\Big\| \sum_{i\in E} e_i \Big\| = \rho_n n \qquad \text{whenever } n \in \mathbb{N} \text{ and } E \subset \mathbb{N} \text{ with } \#E = n\ .$$

*Then $(e_i)$ is $\frac{2}{\rho}$-equivalent to the unit vector basis of $\ell_1$.*

**Corollary 4.1.6.** 1. *Hardy space $H_1$ cannot be renormed so that the the Haar-basis in $H_1$ (which is greedy) is 1-greedy.*

2. *Tsirelson space $T_1$ cannot be renormed so that it has a 1-greedy basis.*

**Problem 4.1.7.** *Can $L_p[0,1]$, $1 < p < \infty$ be renormed so that the Haar basis becomes 1-greedy, or at least $(1+\varepsilon)$- greedy?*

**Coments to Problem 4.1.7**: In Corollary 1.2.2 it was shown that Haar basis in $L_p[0,1]$ is greedy and in [DOSZ2] it was shown that one can renorm $L_p[0,1]$ so that the Haar basis is 1-democratic and 1-unconditional. But the fact that a basis is 1-democratic and suppression 1-unconditional does only imply that it is 1- greedy. From Theorem 1.1.9 it follows that every suppression 1-unconditional and 1-democratic basis is only at least 2-greedy and in [DOSZ2] it was shown that this is optimal and that for any $\varepsilon > 0$ there is a basic sequence which is 1-democratic and suppression 1-unconditional (even 1-unconditional) which is not $2-\varepsilon$-greedy.

## 4.2 Greedy Algorithms

For Hilbertspace it is sill open to find better convergence rates of the pure greedy algorithm but for general Banach spaces the questions about the convergence of $X$-Greedy Algorithm in are quite basic and wide open.

**Problem 4.2.1.** *Find one infinite dimensional Banach space $X$, other than Hilbert space, on which ($X$-PGA) converges for any dictionary?*

**Problem 4.2.2.** *Does ($X$-PGA) converge on $\ell_p$, $1 < p < \infty$, $p \neq 2$, for any dictionary?*

**Comments on Problem 4.2.2:** In [DKSTW] at least the weak convergence of the Pure Greedy Algorithm was shown:

**Theorem 4.2.3.** [DKSTW, Theorem 3.2] *Suppose that for $n \in \mathbb{N}$ $X_n$ is a finite dimensional space whose norm is Gâtaux differentiable. Then the weak pure greedy algorithm with fixed weakness factor converges weakly.*

**Problem 4.2.4.** *Does (X-PGA) converge in* $L_p[0,1]$, $1 < p < \infty$, $p \neq 2$, *at least if one takes the Haar basis as dictionary?*

**Comments on Problem: 4.2.3:** The following finite dimensional version to Problem 4.2.4 was shown in [DOSZ3]

**Theorem 4.2.5.** *Let* $1 < p < \infty$ *and let* $h_j : j \in \mathbb{N}$ *be Haar basis of* $L_p$ *(ordered consistently with the usual partial order). For each* $m$ *there is number* $N = (N(p,m))$ *so that* $X$*- PGA terminates after* $N$ *steps, assuming that the starting point was chosen in* $\operatorname{span}(x_j : j \leq m)$.

**Problem 4.2.6.** *Are there examples of separable and uniform smooth Banach spaces* $X$ *with dictionaries for which the (weak) dual greedy algorithm does not converge?*

# Chapter 5

# Appendix A: Bases in Banach spaces

## 5.1 Schauder bases

In this section we recall some of the notions and results presented in the course on Functional Analysis in Fall 2012 [Schl]. Like every vector space a Banach space $X$ admits an *algebraic* or *Hamel basis*, i.e. a subset $B \supset X$, so that every $x \in X$ is in a unique way the (finite) linear combination of elements in $B$. This definition does not take into account that we can take infinite sums in Banach spaces and that we might want to represent elements in $X$ as converging series. Hamel bases are also not very useful for Banach spaces, since (see Exercise 1), the coordinate functionals might not be continuous.

**Definition 5.1.1.** [Schauder bases of Banach Spaces]

Let $X$ be an infinite dimensional Banach space. A sequence $(e_n) \subset X$ is called *Schauder basis of* $X$, or simply a *basis of* $X$, if for every $x \in X$, there is a unique sequence of scalars $(a_n) \subset \mathbb{K}$ so that

$$x = \sum_{n=1}^{\infty} a_n e_n.$$

**Examples 5.1.2.** For $n \in \mathbb{N}$ let

$$e_n = (\underbrace{0, \ldots 0}_{n-1 \text{ times}}, 1, 0, \ldots) \in \mathbb{K}^{\mathbb{N}}$$

Then $(e_n)$ is a basis of $\ell_p$, $1 \le p < \infty$ and $c_0$. We call $(e_n)$ *the unit vector of* $\ell_p$ and $c_0$, respectively.

**Remarks.** Assume that $X$ is a Banach space and $(e_n)$ a basis of $X$. Then

a) $(e_n)$ is linear independent.

b) $\text{span}(e_n : n\in\mathbb{N})$ is dense in $X$, in particular $X$ is separable.

c) Every element $x$ is uniquely determined by the sequence $(a_n)$ so that $x = \sum_{j=1}^{\infty} a_n e_n$. So we can identify $X$ with a space of sequences in $\mathbb{K}^{\mathbb{N}}$.

**Proposition 5.1.3.** *Let $(e_n)$ be the Schauder basis of a Banach space $X$. For $n \in \mathbb{N}$ and $x \in X$ define $e_n^*(x) \in \mathbb{K}$ to be the unique element in $\mathbb{K}$, so that*

$$x = \sum_{n=1}^{\infty} e_n^*(x) e_n.$$

*Then $e_n^* : X \to \mathbb{K}$ is linear.*

*For $n \in \mathbb{N}$ let*

$$P_n : X \to \text{span}(e_j : j \le n), x \mapsto \sum_{j=1}^{n} e_n^*(x) e_n.$$

*Then $P_n : X \to X$ are linear projections onto* $\text{span}(e_j : j \le n)$ *and the following properties hold:*

*a)* $\dim(P_n(X)) = n$,

*b)* $P_n \circ P_m = P_m \circ P_n = P_{\min(m,n)}$, *for $m, n \in \mathbb{N}$,*

*c)* $\lim_{n\to\infty} P_n(x) = x$, *for every $x \in X$.*

*$P_n$, $n \in \mathbb{N}$, are called the* Canonical Projections for $(e_n)$ *and* $(e_n^*)$ *the* Coordinate Functionals for $(e_n)$ *or* biorthogonals for $(e_n)$.

**Theorem 5.1.4.** *Let $X$ be a Banach space with a basis $(e_n)$ and let $(e_n^*)$ be the corresponding coordinate functionals and $(P_n)$ the canonical projections. Then $P_n$ is bounded for every $n \in \mathbb{N}$ and*

$$b = \sup_{n\in\mathbb{N}} ||P_n\|_{L(X,X)} < \infty,$$

*and thus $e_n^* \in X^*$ and*

$$\|e_n^*\|_{X^*} = \frac{\|P_n - P_{n-1}\|}{\|e_n\|} \le \frac{2b}{\|e_n\|}.$$

*We call $b$* the basis constant of $(e_j)$. *If $b = 1$ we say that $(e_i)$ is a monotone basis.*

*Furthermore*

$$|\!|\!| \cdot |\!|\!| : X \to \mathbb{R}_0^+, \quad \sum_{j=1}^{\infty} a_i e_i \mapsto \Big|\!\Big|\!\Big| \sum_{j=1}^{\infty} a_i e_i \Big|\!\Big|\!\Big| = \sup_{n\in\mathbb{N}} \Big\| \sum_{j=1}^{n} a_i e_i \Big\|,$$

*is an equivalent norm under which $(e_i)$ is a monotone basis.*

**Definition 5.1.5.** [Basic Sequences]
Let $X$ be a Banach space. A sequence $(x_n) \subset X \setminus \{0\}$ is called a *basic sequence* if it is a basis for $\overline{\text{span}(x_n : n \in \mathbb{N})}$.

If $(e_j)$ and $(f_j)$ are two basic sequences (in possibly two different Banach spaces $X$ and $Y$), we say that $(e_j)$ and $(f_j)$ are *isomorphically equivalent* if the map

$$T : \text{span}(e_j : j \in \mathbb{N}) \to \text{span}(f_j : j \in \mathbb{N}), \quad \sum_{j=1}^n a_j e_j \mapsto \sum_{j=1}^n a_j f_j,$$

extends to an isomorphism between the Banach spaces between $\overline{\text{span}(e_j : j \in \mathbb{N})}$ and $\overline{\text{span}(f_j : j \in \mathbb{N})}$.

Note that this is equivalent with saying that there are constants $0 < c \le C$ so that for any $n \in \mathbb{N}$ and any sequence of scalars $(\lambda_j)_{j=1}^n$ it follows that

$$c\Big\|\sum_{j=1}^n \lambda_j e_j\Big\| \le \Big\|\sum_{j=1}^n \lambda_j f_j\Big\| \le C\Big\|\sum_{j=1}^n \lambda_j e_j\Big\|.$$

**Proposition 5.1.6.** *Let $X$ be Banach space and $(x_n : n \in \mathbb{N}) \subset X \setminus \{0\}$. Then $(x_n)$ is a basic sequence if and only if there is a constant $K \ge 1$, so that for all $m < n$ and all scalars $(a_j)_{j=1}^n \subset \mathbb{K}$ we have*

$$\Big\|\sum_{i=1}^m a_i x_i\Big\| \le K\Big\|\sum_{i=1}^n a_i x_i\Big\|. \tag{5.1}$$

*In that case the basis constant is the smallest of all $K \ge 1$ so that (5.1) holds.*

**Theorem 5.1.7.** [The small Perturbation Lemma]

*Let $(x_n)$ be a basic sequence in a Banach space $X$, and let $(x_n^*)$ be the coordinate functionals (they are elements of $\overline{\text{span}(x_j : j \in \mathbb{N})}^*$) and assume that $(y_n)$ is a sequence in $X$ such that*

$$c = \sum_{n=1}^\infty \|x_n - y_n\| \cdot \|x_n^*\| < 1. \tag{5.2}$$

*Then*

*a) $(y_n)$ is also basic in $X$ and isomorphically equivalent to $(x_n)$, more precisely*

$$(1-c)\Big\|\sum_{n=1}^\infty a_n x_n\Big\| \le \Big\|\sum_{n=1}^\infty a_n y_n\Big\| \le (1+c)\Big\|\sum_{n=1}^\infty a_n x_n\Big\|,$$

*for all in $X$ converging series $x = \sum_{n\in\mathbb{N}} a_n x_n$.*

*b) If $\overline{\text{span}(x_j : j \in \mathbb{N})}$ is complemented in $X$, then so is $\overline{\text{span}(y_j : j \in \mathbb{N})}$.*

c) *If $(x_n)$ is a Schauder basis of all of $X$, then $(y_n)$ is also a Schauder basis of $X$ and it follows for the coordinate functionals $(y_n^*)$ of $(y_n)$, that $y_n^* \in \overline{\text{span}(x_j^* : j \in \mathbb{N})}$, for $n \in \mathbb{N}$.*

Now we recall the notion of *unconditional basis.* First the following Proposition.

**Proposition 5.1.8.** *For a sequence $(x_n)$ in Banach space $X$ the following statements are equivalent.*

a) *For any reordering (also called* permutation*) $\sigma$ of $\mathbb{N}$ (i.e. $\sigma : \mathbb{N} \to \mathbb{N}$ is bijective) the series $\sum_{n\in\mathbb{N}} x_{\sigma(n)}$ converges.*

b) *For any $\varepsilon > 0$ there is an $n \in \mathbb{N}$ so that whenever $M \subset \mathbb{N}$ is finite with $\min(M) > n$, then $\left\| \sum_{n\in M} x_n \right\| < \varepsilon$.*

c) *For any subsequence $(n_j)$ the series $\sum_{j\in\mathbb{N}} x_{n_j}$ converges.*

d) *For sequence $(\varepsilon_j) \subset \{\pm 1\}$ the series $\sum_{j=1}^{\infty} \varepsilon_j x_{n_j}$ converges.*

*In the case that above conditions hold we say that the* series $\sum x_n$ converges unconditionally.

**Definition 5.1.9.** A basis $(e_j)$ for a Banach space $X$ is called *unconditional*, if for every $x \in X$ the expansion $x = \sum \langle e_j^*, x\rangle e_j$ converges unconditionally, where $(e_j^*)$ are coordinate functionals of $(e_j)$.

A sequence $(x_n) \subset X$ is called an *unconditional basic sequence* if $(x_n)$ is an unconditional basis of $\overline{\text{span}(x_j : j \in \mathbb{N})}$.

**Proposition 5.1.10.** *For a sequence of non zero elements $(x_j)$ in a Banach space $X$ the following are equivalent.*

a) *$(x_j)$ is an unconditional basic sequence,*

b) *There is a constant $C$, so that for all finite $B \subset \mathbb{N}$, all scalars $(a_j)_{j\in B} \subset \mathbb{K}$, and $A \subset B$*

$$\Big\| \sum_{j\in A} a_j x_j \Big\| \le C \Big\| \sum_{j\in B} a_j x_j \Big\|. \tag{5.3}$$

c) *There is a constant $C'$, so that for all finite sets $B \subset \mathbb{N}$, all scalars $(a_j)_{j\in B} \subset \mathbb{K}$, and all $(\varepsilon_j)_{j\in B} \subset \{\pm 1\}$, if $\mathbb{K} = \mathbb{R}$, or $(\varepsilon_j)_{j\in B} \subset \{z \in \mathbb{C} : |z| = 1\}$, if $\mathbb{K} = \mathbb{C}$,*

$$\Big\| \sum_{j\in B} \varepsilon_j a_j x_j \Big\| \le C' \Big\| \sum_{j=1}^{n} a_j x_j \Big\|. \tag{5.4}$$

*In this case we call the smallest constant* $C = C_s$ *which satisfies* (5.3) *for all* $n$, $A \subset \{1, 2 \dots, n\}$ *and all scalars* $(a_j)_{j=1}^n \subset \mathbb{K}$ *the* supression-unconditional constant of $(x_n)$ *and we call the smallest constant* $C' = C_u$ *so that* (5.4) *holds for all* $n$, $(\varepsilon_j)_{j=1}^n \subset \{\pm 1\}$, *or* $(\varepsilon_j)_{j=1}^n \subset \{z \in \mathbb{C} : |z| = 1\}$, *and all scalars* $(a_j)_{j=1}^n \subset \mathbb{K}$ *the* unconditional constant of $(x_n)$.

*Moreover, it follows that*

$$C_s \le C_u \le 2C_s. \tag{5.5}$$

**Proposition 5.1.11.** *Let* $(x_n)$ *be an unconditional basic sequence. Then*

$$C_u = \sup\Big\{\Big\|\sum_{j=1}^{\infty} a_i b_i x_i\Big\| : x = \sum_{i=1}^{\infty} a_i x_i \in B_X \text{ and } |b_i| \le 1\Big\}. \tag{5.6}$$

**Remark.** While for Schauder bases it is in general important how we order them, the ordering is not relevant for unconditional bases. We can therefore index unconditional bases by any countable set.

## 5.2 Markushevich bases

Not every separable Banach space has a Schauder basis [En]. But it has at least a *bounded and norming Markushevich basis* according to a result of Ovsepian and Pełczyński [OP]. We want to present this result in this section,

**Definition 5.2.1.** A countable family $(e_n, e_n^*)_{n\in\mathbb{N}} \subset X \times X^*$ is called

- *biorthogonal,* if $e_n^*(e_m) = \delta_{(m,n)}$, for all $m, n \in \mathbb{N}$,
- *fundamental,* or *complete,* if $\text{span}(e_n : i \in \mathbb{N})$ is dense in $X$,
- *total,* if for any $x \in X$ with $e_n^*(x) = 0$, for all $n \in \mathbb{N}$, it follows that $x = 0$,
- *norming,* if for some constant $c > 0$,

  $$\sup_{x^* \in \text{span}(e_n^* : n \in \mathbb{N}) \cap B_{X^*}} |x^*(x)| \ge c\|x\|, \text{ for all } x \in X.$$

  and in that case we also say that $(e_n, e_n^*)_{n\in\mathbb{N}}$ is $c$-norming,
- *shrinking,* if $\overline{\text{span}(e_n^* : n \in \mathbb{N})} = X^*$, and
- *bounded,* or *uniformly minimal,* if $C = \sup_{n\in\mathbb{N}} \|e_n\| \cdot \|e_n^*\| < \infty$, and we say in that case that $(e_n, e_n^*)_{n\in\mathbb{N}}$ is $C$*-bounded* and call $C$ *the bound of* $(e_n, e_n^*)_{n\in\mathbb{N}}$.

A biorthogonal, fundamental and total sequence $(e_n, e_n^*)_{n\in\mathbb{N}}$ is called an Markushevich basis or simply *M-Basis.*

**Remark.** Assume $(e_n, e_n^*)$ is an $M$-basis. It follows from the totality that $\text{span}(e_n^* : n \in \mathbb{N})$ is $w^*$-dense in $X^*$ Thus in every reflexive space $M$-bases are shrinking, and shrinking $M$ bases are 1-norming.

Our goal is to prove following

**Theorem 5.2.2.** [OP] *Every separable Banach space $X$ admits a bounded, norming $M$-basis which can be chosen to be shrinking if $X^*$ is (norm) separable. Moreover, the bound of that $M$-basis can be chosen arbitrarily close to* $4(1+\sqrt{2})^2$.

**Remark.** Pełczyński [Pe] improved later the above result and showed that for all separable Banach spaces and all $\varepsilon > 0$ there exists a bounded $M$-basis, whose bound does not exceed $1+\varepsilon$.

It is an open question whether or not every separable Bach space has a 1-bounded $M$-basis. But it is not hard to show that a space $X$ with a bounded and norming $M$-basis can be renormed so that this basis becomes 1-bounded and 1 norming.

**Remark.** It might be nice to know that every separable Banach space has a bounded and norming Markushevish basis $(e_i, e_i^*)$ . Nevertheless, given $z \in X$, we do not have any (set aside a good one) procedure to approximate $z$ by finite linear combinations of the $e_i$, we only know that such an approximation exists. This is precisely the difference to Schauder bases, for which we know that the canonical projections converge point wise.

**Lemma 5.2.3.** [LT, Lemma 1.a.6] *Assume that $X$ is an infinite dimensional space and that $F \subset X$ and $G^* \subset X^*$ are finite dimensional subspaces of $X$ and $X^*$, respectively. Let $\varepsilon > 0$. Then there is an $x \in X$, $\|x\| = 1$ and an $x^* \in X^*$ so that $\|x^*\| \le 2+\varepsilon$, $x^*(x) = 1$, $z^*(x) = 0$, for all $z^* \in G^*$, and $x^*(z) = 0$ for all $z \in F$.*

*Proof.* Let $(y_i^*)_{i=1}^m \subset S_{X^*}$ be finite and $1/(1+\varepsilon)$ norming the space $F$. Pick

$$x \in {}^\perp\big(\{y_j^* : j = 1, 2, \dots m\} \cup G^*\big) \\ = \big\{z \in X : y_j^*(z) = 0,\ \ j = 1, 2 \dots m,\ \text{and}\ z^*(z) = 0,\ z^* \in G^*\big\},\ \text{with}\ \|x\| = 1.$$

It follows for all $\lambda \in \mathbb{R}$ and all $y \in F$ that

$$||y + \lambda x\| \ge \max\big|y_j^*(y + \lambda x)\big| = \max\big|y_j^*(y)\big| \ge \frac{\|y\|}{1+\varepsilon}.$$

Then define

$$u^* : \text{span}(F \cup \{x\}) \to \mathbb{R}, \quad y + \lambda x \mapsto \lambda.$$

We claim that $\|u^*\| \le 2+\varepsilon$. Indeed, let $y \in F$, $y \ne 0$, and $\lambda \in \mathbb{R}$. Then

$$\left|u^*\left(\frac{\lambda x + y}{\|\lambda x + y\|}\right)\right| = \frac{|\lambda|}{\|\lambda x + y\|}$$

$$\leq \begin{cases} 2\frac{\|y\|}{\|\lambda x+y\|} \leq 2(1+\varepsilon) & \text{if } |\lambda| \leq 2\|y\|, \\ \\ \frac{2\|y\|}{2\|y\|-\|y\|} = 2 & \text{if } |\lambda| > 2\|y\|. \end{cases}$$

Letting now $x^*$ be a Hahn Banach extension of $u^*$ onto all of $X$, our claim is proved. □

**Lemma 5.2.4.** ([Ma], see also [HMVZ, Lemma 1.21])
*Let $X$ be an infinite dimensional Banach space. Suppose that $(z_n) \subset X$ and $(z_n^*) \subset X^*$ are sequences so that $\mathrm{span}(z_n : n \in \mathbb{N})$ and $(z_n^* : n \in \mathbb{N})$ are both infinite dimensional and so that*

*(M1) $(z_n)$ separates points of $\mathrm{span}(z_n^* : n \in \mathbb{N})$,*

*(M2) $(z_n^*)$ separates points of $\mathrm{span}(z_n : n \in \mathbb{N})$.*

*Let $N \subset \mathbb{N}$ be co-infinite, and $\varepsilon > 0$.*
*Then we can choose a biorthogonal system*

$$(x_n, x_n^*) \subset \mathrm{span}(z_n : n \in \mathbb{N}) \times \mathrm{span}(z_n^* : n \in \mathbb{N})$$

*with*

$$\mathrm{span}(z_n : n \in \mathbb{N}) \subset \mathrm{span}(x_n : n \in \mathbb{N}) \text{ and } \mathrm{span}(z_n^* : n \in \mathbb{N}) \subset \mathrm{span}(x_n^* : n \in \mathbb{N}) \tag{5.7}$$

$$\sup_{n \in N} \|x_n^*\| \cdot \|x_n^*\| < 2 + \varepsilon. \tag{5.8}$$

**Remark.** Note that $(x_n, x_n^*)$ is an $M$-basis if $\mathrm{span}(z_n : n \in \mathbb{N})$ is dense in $X$, and if $(z_n^*)$ separates points of $X$. It is norming if $B_X^* \cap \mathrm{span}(z_n^*)$ is norming $X$.

*Proof.* Choose $s_1 = \min\{n \in \mathbb{N} : z_n \neq 0\}$. $x_1 = z_{s_1}/\|z_{s_1}\|$. If $1 \in N$ choose $x_1^* \in S_{X^*}$ with $x_1^*(x_1) = 1$. Otherwise choose $x_1^* = z_{t_1}^*$ with $t_1 = \min\{m \in \mathbb{N} : z_m^*(x_1) \neq 0\}$ (which exists by (M1)).

Write $\mathbb{N} \setminus N = \{k_1, k_2, \ldots\}$ and proceed by induction to choose $x_1, x_2, \ldots x_n$ and $x_1^*, \ldots x_n^*$ as follows.

Assume that $x_1, x_2, \ldots x_n$ and $x_1^*, \ldots x_n^*$ have been been chosen.
Case 1: $n+1 \in N$. Then let $F = \mathrm{span}(x_i : i \leq n)$ and $G^* = \mathrm{span}(x_i^* : i \leq n)$; and choose $x_{n+1}$ and $x_{n+1}^*$ by Lemma 5.2.3.
Case2: $n+1 = k_{2j-1} \in \mathbb{N} \setminus N$. Then let

$$s_{2j-1} = \min\left\{s : z_s \notin \mathrm{span}(x_i : i \leq n)\right\}$$

and define

$$x_{n+1} = z_{s_{2j-1}} - \sum_{i=1}^{n} x_i^*(z_{s_{2j-1}}) x_i.$$

This implies that $x_i^*(x_{n+1}) = 0$ for $i = 1, 2, \dots n$. Next choose (using (M2))

$$t_{2j-1} = \min\{t : z_t^*(x_{n+1}) \neq 0\}$$

and let

$$x_{n+1}^* = \frac{z_{t_{2j-1}}^* - \sum_{i=1}^n z_{t_{2j-1}}^*(x_i)x_i^*}{z_{t_{2j-1}}^*(x_{n+1})},$$

which yields that $x_{n+1}^*(x_i) = 0$, for $i = 1, 2 \dots n$, and $x_{n+1}^*(x_{n+1}) = 0$.
Case 3: $n + 1 = k_{2j} \in \mathbb{N} \setminus N$. Then we choose

$$t_{2j} = \min\big\{s : z_s^* \notin \text{span}(x_i^* : i \leq n)\big\}.$$

Let

$$x_{n+1}^* = z_{s_{2j}}^* - \sum_{i=1}^n z_{s_{2j-1}}(x_i)x_i^*,$$

and hence $x_{n+1}^*(x_i) = 0$, for $i = 1, 2 \dots n$, and then let (using (M1)

$$s_{2j} = \min\{s : x_{n+1}^*(x_s) \neq 0\},$$

and

$$x_{n+1} = \frac{z_{s_{2j}} - \sum_{i=1}^n x_i^*(z_{s_{2j}}))x_i}{x_{n+1}^*(z_{s_{2j}})},$$

which implies that $x_i^*(x_{n+1}) = 0$, for $i = 1, 2 \dots n$ and $x_{n+1}^*(x_{n+1}) = 1$.

We insured by this choice that $\big((x_i, x_i^*) : i \in \mathbb{N}\big)$ is a biorthogonal sequence in $X \times X^*$ which also satisfies (5.8) and, since for any $m \in \mathbb{N}$ we have $\text{span}(z_i : i \leq m) \subset \text{span}(x_{k_{2j-1}} : j \leq m)$ and $\text{span}(z_i^* : i \leq m) \subset \text{span}(x_{k_{2j}}^* : j \leq m)$, $\big((x_i, x_i^*) : i \in \mathbb{N}\big)$ it also satisfies (5.7). $\square$

For $n \in \mathbb{N}$ we consider on $\ell_2^{2^n}$ the *discrete Haar basis*

$$\{h_0\} \cup \{h_{(r,s)}, r = 0, 1, \dots, n-1, \text{ and } s = 0, 1, \dots 2^{r-1} - 1\},$$

with

$$h_0 = 2^{-n/2}\chi_{\{1,2,3\dots,2^n\}}$$
$$h_{(r,s)} = \frac{\chi_{\{2s2^{n-r-1}+1, 2s2^{n-r-1}+2, \dots (2s+1)2^{n-r-1}\}} - \chi_{\{(2s+1)2^{n-r-1}+1, (2s+1)2^{n-r-1}+2, \dots (2s+2)2^{n-r-1}\}}}{2^{(n-r)/2}}$$
$$\text{if } r = 0, 1, 2, \dots n-1 \text{ and } s = 0, 1, 2 \dots 2^r - 1.$$

The unit vector basis $(e_i)_{i=1}^{2^n}$ as well as the *Haar basis*

$$\{h_0\} \cup \{h_{(r,s)}, r = 0, 1, \dots, n-1, s = 0, 1, \dots 2^{r-1} - 1\}$$

are orthonormal bases in $\ell_2^{2^n}$. Thus the matrix $A = A^{(n)}$ with the property that $A(e_1) = h_0$ and $A(e_{2^r+s+1}) = h_{(r,s)}$ is a unitary matrix. If we write

$$A^{(n)} = (a^{(n)}_{(i,j)} : 0 \le i, j \le 2^n - 1)$$

then it follows for $k = 0, 1, 2 \dots 2^n - 1$ that $a^{(n)}_{(k,0)} = h_0(k) = 2^{-n/2}$, and if $r = 0, 1, 2 \dots n-1$ and $s = 0, 1, \dots 2^r - 1$ that

$$\begin{aligned} a^{(n)}_{(k,2^r+s)} &= h_{(r,s)}(k) \\ &= \begin{cases} 2^{-(n-r)/2} & \text{if } k \in \{2s2^{n-r-1}+1, 2s2^{n-r-1}+2, \dots (2s+1)2^{n-r-1}\}, \\ -2^{-(n-r)/2} & \text{if } \{(2s+1)2^{n-r-1}+1, (2s+1)2^{n-r-1}+2, \dots (2s+2)2^{n-r-1}\}, \\ 0 & \text{if } k \le 2s2^{n-r-1} \text{ or } k > (2s+2)2^{n-r-1}. \end{cases} \end{aligned}$$

and thus

$$A = \begin{pmatrix} 2^{-n/2} & 2^{-n/2} & 2^{-(n-1))/2} & 0 & \cdots & 2^{-1/2} & \cdots & 0 \\ 2^{-n/2} & \vdots & \vdots & \vdots & \cdots & -2^{1/2} & \cdots & \vdots \\ \vdots & \vdots & 2^{-(n-1))/2} & \vdots & \cdots & 0 & \cdots & \vdots \\ \vdots & \vdots & -2^{-(n-1))/2} & \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & 2^{-n/2} & -2^{-(n-1)/2} & 0 & \cdots & \vdots & \cdots & \vdots \\ \vdots & -2^{-n/2} & 0 & 2^{-(n-1))/2} & \cdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & 2^{(n-1))/2} & \cdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & -2^{-(n-1))/2} & \cdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & 2^{-1/2} \\ 2^{-n/2} & -2^{-n/2} & 0 & -2^{-(n-1)/2} & \cdots & 0 & \cdots & -2^{-1/2} \end{pmatrix}$$

It follows therefore that for all $k = 0, 1 \dots, 2^n - 1$ we have

$$\sum_{j=1}^{2^n-1} \left|a^{(n)}_{(k,j)}\right| = \sum_{r=0}^{n-1} 2^{-(n-r)/2} = \sum_{i=1}^{n} \Big(\frac{1}{\sqrt{2}}\Big)^i \le \frac{\sqrt{2}}{\sqrt{2}-1} = 1 + \sqrt{2}, \tag{5.9}$$

because, leaving out the first column, in each row and for each $r \in \{0, 1, 2 \dots n-1\}$ the value $2^{(n-r)/2}$ is absolutely taken exactly once.

This implies the following:

**Corollary 5.2.5.** *If* $\big((x_i,x_i^*):i=0,1,\dots 2^n-1\big)$ *is a biorthogonal sequence of length* $2^n$*, in a Banach space* $X$ *and we let*

$$e_k=\sum_{j=0}^{2^n-1} a^{(n)}_{(k,j)}x_j \text{ and} \tag{5.10}$$

$$e^*_k=\sum_{j=0}^{2^n-1} a^{(n)}_{(k,j)}x^*_j \text{ for } k=0,1,\dots 2^n-1, \tag{5.11}$$

*then*

$$\max_{0\le k<2^n}\|e_k\|<(1+\sqrt{2})\max_{0\le k<2^n}\|x_k\|+2^{-n/2}\|x_0\|, \tag{5.12}$$

$$\max_{1\le k<2^n}\|e^*_k\|<(1+\sqrt{2})\max_{0\le k<2^n}\|x^*_k\|+2^{-n/2}\|x^*_0\|, \tag{5.13}$$

$$\big((e_j,e^*_j):j=0,1,\dots,2^n-1\big) \text{ is biorthogonal} \tag{5.14}$$

$$\text{span}(e_j:0\le j<2^n)=\text{span}(x_j:0\le j<2^n) \text{ and} \tag{5.15}$$

$$\text{span}(e^*_j:0\le j<2^n)=\text{span}(x^*_j:0\le j<2^n). \tag{5.16}$$

*Proof of Theorem 5.2.2* . Let $\delta>0$ and put $M=2+\delta$. We start with a fundamental sequence $(z_i)\subset X$ and a $w^*$-dense sequence $(z^*_i)\subset B_{X^*}$ ($(B_{X^*},w^*)$ is separable if $X$ is norm separable) , which we choose norm dense if $X^*$ is norm separable. Then we use Lemma 5.2.4 to choose a norming (reps. shrinking) $M$-basis $\big((x_n,x^*_n):n\in\mathbb{N}\big)$ of $X$ which satisfies for $N$ being the odd numbers the conditions (5.7) and (5.8). Without loss of generality we assume that $\|x_n\|=1$, for $n\in\mathbb{N}$. Now we will define a reordering $\big((\tilde{x}_n,\tilde{x}^*_n):n\in\mathbb{N}\big)$ of $\big((x_n,x^*_n):n\in\mathbb{N}\big)$ as follows:

By induction we choose for $\ell\in\mathbb{N}$ a number $m_\ell\in\mathbb{N}$ and define $q_\ell=\sum_{j=1}^{\ell}2^{m_j}$ and $q_0=0$, and choose $\big(\tilde{x}_{q_{\ell-1}},\tilde{x}^*_{q_{\ell-1}}\big),\big(\tilde{x}_{q_{\ell-1}+1},\tilde{x}^*_{q_{\ell-1}+1}\big),\dots,\big(\tilde{x}_{q_\ell-1},\tilde{x}^*_{q_\ell-1}\big)$ as follows

Assume that for all $0\le r<\ell$, $\ell\ge1$, $m_r$, and $(\tilde{x}_0,\tilde{x}^*_0),(\tilde{x}_1,\tilde{x}^*_1),\dots(\tilde{x}_{q_r-1},\tilde{x}^*_{q_r-1})$ have been chosen. Put

$$s_\ell=\min\big\{s:(x_{2s},x^*_{2s})\not\in\{(\tilde{x}_t,\tilde{x}^*_t):t\le q_{\ell-1}-1\}\big\}$$

(recall that $((x_{2s},x^*_{2s}):s\in\mathbb{N})$ are the elements of $\big((x_s,x^*_s):s\in\mathbb{N}\big)$ for which we do not control the norm) and choose $m_\ell\in\mathbb{N}$, so that

$$\big[(1+\sqrt{2})+2^{-m_\ell/2}\big]\cdot\big[(1+\sqrt{2})M+\|x^*_{2s_\ell}\|2^{-m_\ell/2}\big]<(1+\sqrt{2})^2M\ +\delta.$$

Then let $(\tilde{x}_{q_{\ell-1}},\tilde{x}^*_{q_{\ell-1}})=(x_{2s_\ell},x^*_{2s_\ell})$ while $(\tilde{x}_{q_{\ell-1}+1},\tilde{x}^*_{q_{\ell-1}+1})\dots(\tilde{x}_{q_\ell-1},\tilde{x}^*_{q_\ell-1})$ consist of the elements of $\big((x_{2t-1},x^*_{2t-1}):t\in\mathbb{N}\big))$ which are not in the set $\{(\tilde{x}_t,\tilde{x}^*_t):t\le q_{\ell-1}-1\}$ and have the lowest $2^{m_\ell}-1$ indices.

By that choice we made sure that all elements of $\big((x_t, x_t^*) : t \in \mathbb{N}\big)$ appear exactly once in the sequence $\big(\tilde{x}_t, \tilde{x}_t^* : t\in\mathbb{N}\big)$.

Then we apply Corollary 5.2.5 and define for $k = 0, 1, 2, \dots 2^{m_\ell - 1}$

$$e_{q_{\ell-1}+k} = \sum_{j=0}^{2^{m_\ell}-1} a_{(k,j)}^{(m_\ell)} \tilde{x}_{q_{\ell-1}+j} \text{ and } e^*_{q_{\ell-1}+k} = \sum_{j=0}^{2^{m_\ell}-1} a_{(k,j)}^{(m_\ell)} \tilde{x}^*_{q_{\ell-1}+j}.$$

It follows then from (5.12) and (5.13) that for $k = 0, 1, 2, \dots 2^{m_\ell - 1}$

$$\|e_{q_{\ell-1}+k}\| \cdot \|e^*_{q_{\ell-1}+k}\| \le (1+\sqrt{2})^2 M + \delta.$$

Choosing $\delta > 0$ small enough we can ensure that $(1+\sqrt{2})^2 M + \delta < 2(1+\sqrt{2})^2 + \varepsilon$. Since $\big((x_n, x_n^*) : n \in \mathbb{N}\big)$ is a norming $M$-basis, it follows from (5.14), (5.15) and (5.16) that $\big((e_n, e_n^*) : n \in \mathbb{N}\big)$ is a norming $M$ basis which is shrinking if $\big((x_n, x_n^*) : n\in\mathbb{N}\big)$ is shrinking. $\square$

# Chapter 6

# Appendix B: Some facts about $L_p[0,1]$ and $L_p(\mathbb{R})$

## 6.1 The Haar basis and Wavelets

We recall the definition of the *Haar basis of* $L_p[0,1]$. Let

$$T = \{(n,j) : n \in \mathbb{N}_0, j = 0, 1 \ldots, 2^n - 1\} \cup \{0\}.$$

Let $1 \le p < \infty$ be fixed. We define *the Haar basis* $(h_t)_{t\in T}$ and the *normalized Haar basis* $(h_t^{(p)})_{t\in T}$ *in* $L_p[0,1]$ as follows.

$h_0 = h_0^{(p)} \equiv 1$ on $[0,1]$ and for $n \in \mathbb{N}_0$ and $j = 0, 1, 2 \ldots 2^n - 1$ we put

$$h_{(n,j)} = 1_{[j2^{-n},(j+\frac{1}{2})2^{-n})} - 1_{[(j+\frac{1}{2})2^{-n},(j+1)2^{-n})}.$$

and we let

$$\begin{aligned}
\Delta_{(n,j)} &= \mathrm{supp}(h_{(n,j)}) = \big[j2^{-n}, (j+1)2^{-n}\big),\\
\Delta^+_{(n,j)} &= \Big[j2^{-n}, \Big(j+\frac{1}{2}\Big)2^{-n}\Big)\\
\Delta^-_{(n,j)} &= \Big[\Big(j+\frac{1}{2}\Big)2^{-n}, (j+1)2^{-n}\Big).
\end{aligned}$$

We let $h^{(\infty)}_{(n,j)} = h_{(n,j)}$. And for $1 \le p < \infty$

$$h^{(p)}_{(n,j)} = \frac{h_{(n,j)}}{\|h_{(n,j)}\|_p} = 2^{n/p}\big(1_{[j2^{-n},(j+\frac{1}{2})2^{-n}} - 1_{[(j+\frac{1}{2})2^{-n}),(j+1)2^{-n})}\big).$$

**Theorem 6.1.1.** [Schl, Theorems 3.2.2, 5.5.1]

*We order* $(h_t^{(p)} : t \in T)$ *into as sequence* $(h_n : n \in \mathbb{N})$, *with* $h_1 = h_0^{(p)}$, *and the property that if* $2 \le m < n$, *then either* $\mathrm{supp}(h_n) \subset \mathrm{supp}(h_m)$ *or* $\mathrm{supp}(h_m) \cap \mathrm{supp}(h_n) = \emptyset$. *Then* $(h_n)$ *is a monotone basis for* $L_p[0,1]$.

*For $1 < p < \infty$, $(h_t^{(p)} : t \in T)$ is an unconditional basis for $L_p[0,1]$, but it is not unconditional for $p = 1$. In fact $L_1[0,1]$ does not embed into a Banach space with unconditional basis.*

Define $h = 1_{[0,1/2]} - 1_{(1/2,1]}$. Then we can write for $n \in \mathbb{N}_0$ and $j = 0, 1, \ldots 2^n - 1$.

$$h_{(n,j)}(t) = h(2^n t - j), \text{ for } t \in [0,1], \text{ and}$$
$$h^{(p)}_{(n,j)}(t) = 2^{n/p} h(2^n t - j), \text{ for } t \in [0,1].$$

We define now for **all** $n \in \mathbb{Z}$ and all $j \in \mathbb{Z}$ a function $h_{(n,j)}$ as follows

$$h_{(n,j)}(t) = h(2^n t - j), \text{ for } t \in \mathbb{R}, \text{ and} \tag{6.1}$$
$$h^{(p)}_{(n,j)}(t) = 2^{n/p} h(2^n t - j), \text{ for } t \in \mathbb{R}. \tag{6.2}$$

For all $n, j \in \mathbb{Z}$ we have

$$\text{supp}(h_{(n,j)}) := \{t : h_{(n,j)}(t) < 0\} = \{t : 0 \le 2^n t - j \le 1\} = [j2^{-n}, (j+1)2^{-n}) =: \Delta_{(n,j)}$$

and

$$\{h_{(n,j)} > 0\} = \Big[j2^{-n}, \Big(j + \frac{1}{2}\Big)2^{-n}\Big) =: \Delta^+_{(n,j)},$$
$$\{h_{(n,j)} < 0\} = \Big[\Big(j + \frac{1}{2}\Big)2^{-n}, j + 1)2^{-n}\Big) =: \Delta^-_{(n,j)},$$

We note for and $(m,i)$ and $(n,j)$ in $\mathbb{Z} \times \mathbb{Z}$ that

$$\text{Either } \Delta_{(m,i)} \subset \Delta_{(n,j)} \text{ or } \Delta_{(n,j)} \subset \Delta_{(m,i)} \text{ or } \Delta_{(m,i)} \cap \Delta_{(n,j)} = \emptyset. \tag{6.3}$$

**Theorem 6.1.2.** *Let $1 \le p < \infty$.*

1. $\{h^{(p)}_{(n,j)} : n \in \mathbb{N}_0, j \in \mathbb{Z}\} \cup \{1_{[j,j+1)} : j \in \mathbb{Z}\}$, *when appropriately ordered, is a monotone basis for $L_p(\mathbb{R})$, which is unconditional if $1 < p < \infty$.*

2. $\{h^{(p)}_{(n,j)} : n \in \mathbb{Z}, j \in \mathbb{Z}\}$ *is an unconditional basis for $L_p(\mathbb{R})$ if $1 < p < \infty$.*

**Remark.** Note that the second part of Theorem 6.1.2 is wrong for $p = 1$. Indeed, the *integral functional*

$$I : L_1(\mathbb{R}) \to \mathbb{R}, \qquad f \mapsto \int_{-\infty}^{\infty} f(x)\, dx,$$

is a bounded linear functional on $L_1(\mathbb{R})$ which is not identical to the zero-functional, but for all $n, j \in \mathbb{Z}$ $h^{(1)}_{(n,j)}$ is in the kernel of $I$, and thus the span of the $h^{(1)}_{(n,j)}$ cannot be dense in $L_1(\mathbb{R})$. The same argumentation is invalid for $1 < p < \infty$ since $I$ is unbounded on $L_p(\mathbb{R})$, if $p > 1$.

*Proof.* First note that $L_p(\mathbb{R})$ is isometrically isomorphic to $(\oplus_{i\in\mathbb{Z}} L_p[i,i+1])$ via the map

$$L_p(\mathbb{R}) \to (\oplus_{i\in\mathbb{Z}} L_p[i,i+1]), \quad f \mapsto (f|_{[i,i+1]} : i \in \mathbb{Z}),$$

and that for $i\in Z$, by Theorem 6.1.1, the *shifted Haar basis*

$$\begin{aligned}\big(h^{(p,i)}_{(n,j)} &: n\in\mathbb{N}_0, j=0,1,\dots,2^n-1\big)\big) \cup \{1_{[i,i+1)}\}\\ &= \big(h^{(p)}_{(n,j)} : n\in\mathbb{N}_0, j = 2^n i, 2^n i+1, 2^n i+2\dots 2^n(i+1)-1\big) \cup \{1_{[i,i+1)}\}\end{aligned}$$

is a monotone basis for $L_p[i,i+1]$, if ordered appropriately, which is unconditional if $1<p<\infty$. Since $L_p(\mathbb{R})$ is the 1-unconditional sum of the spaces $L_p[i,i+1)$, $i\in\mathbb{Z}$ the union of $\big(h^{(p,i)}_{(n,j)} : n\in\mathbb{N}_0, j=0,1,\dots,2^n-1\big)\big)\cup\{1_{[i,i+1)}$ over all $i\in\mathbb{Z}$ is a monotone basis of $L_p[0,1]$, if ordered appropriately, which is unconditional if $p>1$

In order to show (2) we assume $B\subset\mathbb{Z}\times\mathbb{Z}$ is finite and $A\subset B$ and then verify condition (5.3). Since $B$ is finite, there is $n_1,\in\mathbb{N}$ so that

$$\Delta_{(n,j)} \subset \Delta_{(-n_1,0)} = [0,2^{n_1}) \text{ for all } (n,j)\in B^+ = \{(m,i)\in\mathbb{Z}\times\mathbb{N}_0\}\cap B,$$

and there is $n_2,\in\mathbb{N}$ so that

$$\Delta_{(n,j)} \subset \Delta_{(-n_2,-1)} = [-2^{n_2},0) \text{ for all } (n,j)\in B^- = \{(m,i)\in\mathbb{Z}\times(-\mathbb{N})\}\cap B,$$

Since the $L_p$-nrm is shift invariant, it is enough to assume that $B^-=\emptyset$ and $B=B^+$.

Consider the map (rescaling)

$$\phi : \Delta_{(0,0)} = [0,1] \to \Delta_{(-n_1,0)}, \quad t\mapsto t2^{n_1}$$

and the map

$$T : L_p(\Delta_{(-n_1,0)}) \to L_p[0,1], \quad f\mapsto 2^{n_1/p} f\circ\phi,$$

which is an isometry between $L_p(\Delta_{(n_1,0)})$ and $L_p[0,1]$, mapping the family

$$\{h^{(p)}_{(n,j)} : \Delta_{(n,j)} \subset \Delta_{(n_1,0)}\} \cup \{1_{\Delta_{(n_1,0)}}\}$$

into the Haar basis of $L_p[0,1]$

$$\{h^{(p)}_{(n,j)} : \Delta_{(n,j)} \subset \Delta_{(0,0)}\} \cup \{1_{[0,1]}\}.$$

This proves with Theorem 6.1.1 that $\{h^{(p)}_{(n,j)} : \Delta_{(n,j)} \subset \Delta_{(n_1,0)}\}\cup\{1_{\Delta_{(n_1,0)}}\}$ is unconditional and therefore (5.3) is satisfied.

It is left to show that the closed linear span of $\{h^{(p)}_{(n,j)} : n\in\mathbb{Z}, j\in\mathbb{Z}\}$ is all of $L_p(\mathbb{R})$. By part (1) and using shifts, it is enough to show that $1_{[0,1]}$ is in the closed linear span of $\{h^{(p)}_{(n,j)} : n\in\mathbb{Z}, j\in\mathbb{Z}\}$.

Notice that for $N\in\mathbb{N}$ we have

$$\begin{aligned}
&\sum_{n=0}^{N} 2^{-(n+1)}\big(1_{[0,2^n)}-1_{[2^n,2^{n+1})}\big)\\
&\qquad = 1_{[0,1)}\sum_{n=0}^{N}2^{-(n+1)}\\
&\qquad\quad +1_{[1,2)}\Big(\sum_{n=1}^{N}2^{-(n+1)}-2^{-1}\Big)\\
&\qquad\quad +1_{[2,4)}\Big(\sum_{n=2}^{N}2^{-(n+1)}-2^{-2}\Big)\\
&\qquad\quad \vdots\\
&\qquad\quad +1_{[2^{N-1},2^N)}(2^{-(N+1)}-2^{-N})\\
&\qquad\quad -1_{[2^N,2^{N+1})}2^{-(N+1)}\\
&\qquad = 1_{[0,1)}(1-2^{-(N+1)})-1_{[1,2^N)}2^{-(N+1)}-1_{[2^N,2^{N+1})}2^{-(N+1)}.
\end{aligned}$$

For the last equality note that for $k=1,2,\dots N-1$

$$\sum_{n=k}^{N}2^{-n-1}-2^{-k}=2^{-k}-2^{-(N+1)}-2^{-k}=-2^{-(N+1)}.$$

Since we assumed that $p>1$, it follows that

$$L_p-\lim_{N\to\infty}\sum_{n=0}^{N}2^{-(n+1)}\big(1_{[0,2^n)}-1_{[2^n,2^{n+1})}\big)=1_{[0,1)},$$

which finishes the prove of our claim. □

**Definition 6.1.3.** A function $\Psi\in L_2(\mathbb{R})$ is called *wavelet* if the family $(\Psi_{(n,j)}: n,j\in\mathbb{Z})$, defined by

$$\Psi_{(n,j)}(t)=2^{n/2}\Psi(2^nt-j),\ \text{for } t\in\mathbb{R} \text{ and } n,j\in\mathbb{Z},$$

is an orthonormal basis of $L_2(\mathbb{R})$.

**Definition 6.1.4.** A *Multi Resolution Analysis* of $L_2(\mathbb{R})$ (MRA) is sequence of closed subspaces $(V_n: n\in\mathbb{Z})$ of $L_2(\mathbb{R})$ such that

(MRA1) $\dots V_{-2}\subset V_{-1}\subset V_0\subset V_1\subset V_2\dots,$

(MRA2) $\overline{\bigcup_{n\in\mathbb{Z}}V_n}=L_2(\mathbb{R}),$

(MRA3) $\bigcap_{n\in\mathbb{Z}} V_n = \{0\}$,

(MRA4) $V_n = \{f(2^{+n}(\cdot)) : f \in V_0\}$, for $n \in \mathbb{Z}$,

(MRA5) there is a compactly supported function $\Phi \in V_0$, so that $(\Phi((\cdot)-m) : m \in \mathbb{Z})$ is an orthonormal basis of $V_0$.

In this case we call $\Phi$ a *scaling function of the MRA* $(V_n : n\in\mathbb{Z})$.
Note that (MRA5) implies

(MRA6) $V_0$ (and thus any $V_n$) is *translation invariant* by integer shifts, i.e.

$$f \in V_0 \iff f((\cdot) - j) \in V_0, \text{ for all } j \in \mathbb{Z} \text{ and}$$

For $h \in \mathbb{R}$ and $f \in L_2(\mathbb{R})$ we put

$$T_h : L_2(\mathbb{R}) \to L_2(\mathbb{R}), \quad f \mapsto f((\cdot) - h) \text{ (Shift to there right by } h \text{ units)}$$
$$J_h : L_2(\mathbb{R}) \to L_2(\mathbb{R}), \quad f \mapsto 2^{h/2} f((\cdot)2^h) \text{ (Scaling )}$$

**Remark.** For $h \in \mathbb{R}$ the operators $T_h$ and $J_h$ are isometries and

$$T_h^{-1} = T_{-h} \text{ and } J_h^{-1} = J_{-h}. \tag{6.4}$$

We can rephrase (MRA4) and (MRA6) equivalently as follows

(MRA4') $V_n = J_n(V_0)$, for $n \in \mathbb{Z}$, and

(MRA6') $V_0 = T_n(V_0)$ for all $n \in \mathbb{Z}$.

Finally note that (MRA4), (MRA6) and (MRA5) implies that for $j \in \mathbb{Z}$

(MRA5') $\{2^{j/2}\Phi(2^j(\cdot) - k) : j, k \in \mathbb{Z}\}$ is an orthonormal basis of $V_j$.

and note that $T_h$ as well as $J_h$ are both isometries on $L_2(\mathbb{R})$.

**Example 6.1.5.** Take $\Phi = 1_{[0,1)}$, and for $(n \in \mathbb{Z})$, put

$$V_n = \overline{\text{span}(J_n \circ T_j(\phi) : j \in \mathbb{Z})} = \overline{\text{span}(1_{[j2^{-n},(j+1)2^{-n})}; j\in\mathbb{Z})}.$$

(Note $1_{[j2^{-n},(j+1)(2^{-n})}(2^{-n}(\cdot)) = 1_{[j,j+1)}(\cdot) \in V_0$).
Then $(V_n)$ is an MRA.

We now discuss how to produce a wavelet $\Psi$ starting with an MRA $(V_n : n \in \mathbb{Z})$ with scaling function $\Phi$.

We denote the orthogonal complement of $V_j$ inside $V_{j+1}$ by $W_j$; this means that any $f \in V_{j+1}$ can be written as $f = g + h$ with $g \in V_j$ and $h \in W_j$ and

$\|f\|_2 = \sqrt{\|g\|_2^2 + \|h\|_2^2}$. We write $V_{j+1} = W_j \oplus V_j$. Since $J_j$ is an unitary operator (keeping orthogonality), we deduce that

$$\begin{aligned} V_j \oplus W_j = V_{j+1} \\ & J_j(V_1) \\ & = J_j(V_0 \oplus W_0) \\ & = J_j(V_0) \oplus J_j(W_0) = V_j \oplus J_j(W_0). \end{aligned}$$

Since the orthogonal complement to a subspace $W$ of a Hilbert space is unique (namely $W^\perp = \{h \in H : \forall w \in W\ \langle h, w\rangle = 0\}$), we obtain

$$W_j = J_j(W_0) \text{ for all } j \in \mathbb{Z}. \tag{6.5}$$

Next we observe that

$$W_j \text{ and } W_i \text{ are orthonormal if } j \neq i. \tag{6.6}$$

Indeed, assume w.l.o.g. that $i < j$. Then $W_i \subset V_{i+1} \subset V_j$ and $W_j$ is orthogonal to $V_j$.

From (MRA2) we deduce that every $f \in L_2(\mathbb{R})$ can be arbitrary approximated by some $g \in V_n$ for large enough $n \in \mathbb{N}$ and (MRA3) yields that

$$\lim_{k\to\infty} P_{V_{-k}}(f) = 0$$

where $P_{V_{-k}}$ is the orthogonal projection of $L_2(\mathbb{R})$ onto $V_{-k}$. Thus, choosing $k \in \mathbb{N}$ we can arbitrarily approximate $g$ by an element $h$ in the orthogonal complement of $V_{-k}$ inside $V_n$. But since

$$V_n = W_{n-1} \oplus V_{n-1} = W_{n-1} \oplus W_{n-2} \oplus V_{n-2} = \ldots (W_{n-1} \oplus W_{n-2} \oplus W_{-k}) \oplus V_{-k},$$

it follows that

$$h \in W_{n-1} \oplus W_{n-2} \oplus \ldots W_{-k}.$$

As a consequence we deduce that $L_2(\mathbb{R})$ is the orthonormal sum of the $W_j$, $j \in \mathbb{Z}$. Together with our observations (6.5) and (6.6) this yields the following result.

**Proposition 6.1.6.** *Every $\Psi \in W_0$ for which $\{\Psi((\cdot) - j) : j \in \mathbb{Z}\}$ is an orthonormal basis of $W_0$ is a wavelet. We say in that case that $\Psi$ is the* wavelet associated to the MRA $(V_n)$.

**Example 6.1.7.** We consider the Example 6.1.5. Then

$$W_0 = \{f \in V_1 : \forall g \in V_0\ \langle g, f\rangle = 0\} = \Big\{f \in V_1 : \int_j^{j+1} f(t)\, dt = 0\Big\}.$$

Thus we could take

$$\Psi = 1_{[0,1/2)} - 1_{[1/2,1)},$$

as a wavelet associated to $(V_n)$.

We would like to explain how to construct the wavelet $\Psi$ associated to an MRA $(V_n : n \in \mathbb{Z})$ with scaling function $\Phi$.

**Theorem 6.1.8.** *Suppose that $(V_n : n \in \mathbb{Z})$ is an MRA with scaling function $\Phi$ which is integrable and*

$$\int_{-\infty}^{\infty} \Phi(t)\, dt \neq 0.$$

1. *The the following* Scaling Relation *holds:*

$$\Phi = \sum_{k \in \mathbb{Z}} p_k \Phi(2(\cdot) - k) \text{ with } p_k = 2 \int_{-\infty}^{\infty} \Phi(x) \overline{\Phi(2x - k)} dx. \tag{6.7}$$

   *More generally*

$$\Phi(2^{j-1}(\cdot) - l) = \sum_{k \in \mathbb{Z}} p_{k-2l} \Phi(2^j(\cdot) - k) \text{ for all } j, l \in \mathbb{Z}. \tag{6.8}$$

2. *The sequence $(p_k : k \in \mathbb{Z})$ satisfies*

$$\sum_{k \in \mathbb{Z}} p_{k-2l} \overline{p_k} = 2\delta_{0,l} \text{ for all } l \in \mathbb{Z} \tag{6.9}$$

$$\sum_{k \in 2\mathbb{Z}} p_k = \sum_{k \in 2\mathbb{Z}+1} p_k = 1. \tag{6.10}$$

3. *The function $\Psi$ defined by*

$$\Psi = \sum_{k \in \mathbb{Z}} (-1)^k \overline{p_{1-k}} \Phi(2(\cdot) - k) \tag{6.11}$$

   *is a wavelet associated to $(V_n : n \in \mathbb{Z})$.*

*Proof.* Since $\Phi \in V_0 \subset V_1$ and since $(2^{1/2}\Phi(2(\cdot) - k) : k \in \mathbb{Z})$ is an orthonormal basis of $V_1$, we can write $\Phi$ as in (6.7). For $j, l \in \mathbb{Z}$ it follows therefore that

$$\begin{aligned} \Phi(2^{j-1}(\cdot) - l) &= \sum_{k \in \mathbb{Z}} p_k \Phi\big(2(2^{j-1}(\cdot) - l) - k\big) \\ &= \sum_{k \in \mathbb{Z}} p_k \Phi\big(2^j(\cdot) - 2l - k\big) = \sum_{k \in \mathbb{Z}} p_{k-2l} \Phi(2^j(\cdot) - k). \end{aligned}$$

Since $(\Phi((\cdot) - l) : l \in \mathbb{Z})$ is an orthonormal sequence we obtain from (6.7) and (6.8) (with $j =$) that

$$\begin{aligned} \delta_{(0,l)} &= \langle \Phi((\cdot) - l), \Phi \rangle \\ &= \sum_{m,k \in \mathbb{Z}} p_{m-2l} \overline{p}_k \langle \Phi(2(\cdot) - m), \Phi(2(\cdot) - k) \rangle \end{aligned}$$

$$= \sum_{k\in\mathbb{Z}} p_{k-2l}\overline{p}_k \int_{-\infty}^{\infty} |\Phi(2t)|^2 dt = \frac{1}{2}\sum_{k\in\mathbb{Z}} p_{k-2l}\overline{p}_k,$$

which proves (6.9) and implies (replacing $l$ by $-l$) that

$$\begin{aligned}
2 &= 2\sum_{l\in\mathbb{Z}} \delta_{(0,-l)} \\
&= \sum_{l,k\in\mathbb{Z}} p_{k+2l}\overline{p_k} \\
&= \sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} p_{2k+2l}\overline{p_{2k}} + \sum_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} p_{2k+1+2l}\overline{p_{2k+1}} \\
&= \sum_{k\in\mathbb{Z}} \overline{p_{2k}}\Big(\sum_{l\in\mathbb{Z}} p_{2k+2l}\Big) + \sum_{k\in\mathbb{Z}} \overline{p_{2k+1}}\Big(\sum_{l\in\mathbb{Z}} p_{2k+1+2l}\Big) \\
&= \sum_{k\in\mathbb{Z}} \overline{p_{2k}}\Big(\sum_{l\in\mathbb{Z}} p_{2l}\Big) + \sum_{k\in\mathbb{Z}} \overline{p_{2k+1}}\Big(\sum_{l\in\mathbb{Z}} p_{2l+1}\Big) \\
&= \Big|\sum_{k\in\mathbb{Z}} p_{2k}\Big|^2 + \Big|\sum_{k\in\mathbb{Z}} p_{2k+1}\Big|^2 =: A^2 + B^2.
\end{aligned}$$

Moreover if we integrate the scaling relation (6.7) we obtain

$$\int_{-\infty}^{\infty} \Phi(x)dx = \frac{1}{2}\sum_{k\in\mathbb{Z}} p_k \int_{-\infty}^{\infty} \Phi(x)dx. \tag{6.12}$$

By our assumption on the integral of $\Phi$, we can cancel the integral on both sides of (6.12) and obtain that

$$A + B = \sum_{k\in\mathbb{Z}} p_k = 2.$$

The only solution of $A^2 + B^2 = 2$ and $A + B = 2$ is $A = B = 1$ which proves (6.10) (draw a picture!).

In order to prove the (3) we first note that $\Psi \in V_1$. Secondly, we note that the sequence $(\Psi((\cdot) - k) : k \in \mathbb{Z})$ is orthonormal. Indeed, it follows from (6.8) and (6.9) for $l, m \in \mathbb{Z}$ that

$$\begin{aligned}
&\langle \Psi((\cdot) - l), \Psi((\cdot) - m)\rangle \\
&\quad= \Big\langle \sum_{k=1}(-1)^k\overline{p_{1-k}}\Phi(2((\cdot) - l) - k), \sum_{k=1}(-1)^k\overline{p_{1-k}}\Phi(2((\cdot) - m) - k)\Big\rangle \\
&\quad= \Big\langle \sum_{k=1}(-1)^k\overline{p_{1-k+2l}}\Phi(2(\cdot) - k), \sum_{k=1}(-1)^k\overline{p_{1-k+2m}}\Phi(2(\cdot) - k)\Big\rangle \\
&\quad= \frac{1}{2}\sum_{k\in\mathbb{Z}} \overline{p_{1-k+2l}}p_{1-k+2m}
\end{aligned}$$

$$= \frac{1}{2}\sum_{k\in\mathbb{Z}} \overline{p_{1-k+2l-2m}} p_{1-k} = \delta_{(l,m)}.$$

Thirdly, it follows from 6.8 for any $l, m \in \mathbb{Z}$ that

$$\begin{aligned}
&\langle \Phi((\cdot)-l), \Psi((\cdot)-m)) \rangle \\
&\quad = \Big\langle \sum_{k\in\mathbb{Z}} p_{k-2l}\Phi((\cdot)-k), \sum_{k=1}(-1)^k\overline{p_{1-k}}\Phi(2((\cdot)-m)-k)\Big\rangle \\
&\quad = \Big\langle \sum_{k\in\mathbb{Z}} p_{k-2l}\Phi((\cdot)-k), \sum_{k=1}(-1)^k\overline{p_{1-k+2m}}\Phi((\cdot)-k)\Big\rangle \\
&\quad = \frac{1}{2}\sum_{k\in\mathbb{Z}}(-1)^k p_{k-2l}p_{1-k+2m} \\
&\quad = \frac{1}{2}\sum_{k\in\mathbb{Z}}(-1)^k p_k p_{1-k+r} \text{ with } r = 2m-2l \\
&\quad = \frac{1}{2}\sum_{k\in\mathbb{Z}} p_{2k}p_{1-2k+r} - \frac{1}{2}\sum_{k\in\mathbb{Z}} p_{2k+1}p_{-2k+r} \\
&\quad = \frac{1}{2}\sum_{k\in\mathbb{Z}} p_{2k}p_{1-2k+r} - \frac{1}{2}\sum_{l\in\mathbb{Z}} p_{1-2l+r}p_{2l} = 0. \\
&\quad \text{(substitute } 2l = r-2k)
\end{aligned}$$

Finally, in order to show that $(\Psi((\cdot)-k) : k \in \mathbb{Z})$ and $(\Phi((\cdot)-k) : k \in \mathbb{Z})$ span all of $V_1$ we need to show that for given $j \in \mathbb{Z}$ the projection of $\Phi(2(\cdot)-j)$ onto the space spanned by $(\Psi((\cdot)-k) : k \in \mathbb{Z})$ and $(\Phi((\cdot)-k) : k \in \mathbb{Z})$ has the same norm (namely $1/2$) as $\Phi(2(\cdot)-j)$. Let us denote the projected vector by $\tilde{\Phi}_j$. By the above shown orthonormalities of $(\Psi((\cdot)-k) : k \in \mathbb{Z})$ and $(\Phi((\cdot)-k) : k \in \mathbb{Z})$ we can write

$$\tilde{\Phi}_j = \sum_{k\in\mathbb{Z}} \big[a_k\Phi((\cdot)-k) + b_k\Psi((\cdot)-k))\big],$$

with

$$a_k = \langle \Phi(2(\cdot)-j), \Phi((\cdot)-k)\rangle = \langle \Phi(2(\cdot)+2k-j), \Phi\rangle = \frac{1}{2}\overline{p_{j-2k}}.$$

and

$$\begin{aligned}
b_k &= \langle \Phi(2(\cdot)-j), \Psi((\cdot)-k)\rangle \\
&= \sum_{l\in\mathbb{Z}}(-1)^l\langle \Phi(2(\cdot)-j), \overline{p_{1-l}}\Phi(2(\cdot)-l-2k)\rangle = \frac{1}{2}(-1)^j p_{1-j+2k}
\end{aligned}$$

for $k\in\mathbb{Z}$. It follows therefore that

$$\|\tilde{\Phi}_j\|^2 = \frac{1}{4}\sum_{k\in\mathbb{Z}}|p_{j-2k}|^2 + \frac{1}{4}\sum_{k\in\mathbb{Z}}|p_{1-j+2k}|^2 = \frac{1}{4}\sum_{k\in\mathbb{Z}}|p_k|^2 = \frac{1}{2}.$$

□

The proof of the following result goes beyond the scope of these notes, since it requires several tools from harmonic analysis

**Theorem 6.1.9.** [Wo1, Theorem 8.13] *Assume that $\Psi$ is a wavelet for $L_2(\mathbb{R})$ satisfying the following two conditions for some constant $C > 0$:*

$$|\Psi(x)| \le C(1+|x|)^{-2} \text{ for all } x \in \mathbb{R} \tag{6.13}$$

$$\Psi \text{ is differentiable and } \big|\Psi'(x)\big| \le C\big(1+|x|\big)^{-2} \text{ for all } x \in \mathbb{R} \tag{6.14}$$

*Then for every $1 < p < \infty$ the family $\Psi^{(p)}_{j,k} = (2^{j/p}\Psi(2^j(\cdot) - k) : j,k \in \mathbb{Z})$ is a basis of $L_p(\mathbb{R})$ which is isomorphically equivalent to $(h^{(p)}_{j,k}) : j,k \in \mathbb{Z})$.*

We finally, want to present without proof another basis of $L_p[0,1]$. Recall that $\big(e^{(\cdot)n/2\pi}\big)$ is an orthonormal basis of $L_2[0,1]$. A deep Theorem by M. Riesz states the following.

**Theorem 6.1.10.** (c.f. [Ka, Chapter II and III]) *The sequence of* trigonometric polynomials $(t_n : n \in \mathbb{Z})$*, with*

$$t_n(\xi) = e^{i\xi n/2\pi}, \text{ for } \xi \in [0,1]$$

*is a Schauder basis of $L_p[0,1]$, $1 < p < \infty$, when ordered as $(t_0, t_1, t_{-1}, t_2, t_{-2}, \ldots$*

In the next section we prove that for $p \neq 2$ $(t_n)$ cannot be unconditional.

## 6.2 Khintchine's inequality and Applications

**Theorem 6.2.1.** [Khintchine's Theorem, see Theorem 5.3.1 in [Schl]]
*$L_p[0,1]$, $1 \le p \le \infty$ contains subspaces isomorphic to $\ell_2$. If $1 < p < \infty$ $L_p[0,1]$, contains a complemented subspaces isomorphic to $\ell_2$.*

**Definition 6.2.2.** *The Rademacher functions* are the functions:

$$r_n : [0,1] \to \mathbb{R}, \qquad t \mapsto \text{sign}(\sin 2^n \pi t), \text{ whenever } n \in \mathbb{N}.$$

**Lemma 6.2.3.** [Khintchine inequality], see Lemma 5.3.3 in [Schl]
*For every $p \in [1,\infty)$ there are numbers $0 < A_p \le 1 \le B_p$ so that for any $m \in \mathbb{N}$ and any scalars $(a_j)_{j=1}^m$,*

$$A_p\Big(\sum_{j=1}^m |a_j|^2\Big)^{1/2} \le \Big\|\sum_{j=1}^m a_j r_j\Big\|_{L_p} \le B_p\Big(\sum_{j=1}^m |a_j|^2\Big)^{1/2}. \tag{6.15}$$

There is a *complex version of the Rademacher functions* namely the sequence $(g_n : n \in \mathbb{N})$, with

$$g_n(t) = e^{i2^n\pi t} \quad \text{for } t \in [0,1].$$

**Theorem 6.2.4.** [Complex Version of Khintchine's Theorem] *For every $p \in [1,\infty)$ there are numbers $0 < A'_p \le 1 \le B'_p$ so that for any $m \in \mathbb{N}$ and any scalars $(a_j)_{j=1}^m$,*

$$A'_p\Big(\sum_{j=1}^m |a_j|^2\Big)^{1/2} \le \Big\|\sum_{j=1}^m a_j g_j\Big\|_{L_p} \le B'_p\Big(\sum_{j=1}^m |a_j|^2\Big)^{1/2}. \tag{6.16}$$

*Moreover, if $1 < p < \infty$ then $(g_j)$ generates a copy of $\ell_2$ inside $L_p[0,1]$ which is complemented.*

**Theorem 6.2.5** (The square-function norm). *Let $1 \le p < \infty$ and let $(f_n)$ be a $\lambda$-unconditional basic sequence in $L_p[0,1]$ for some $\lambda \ge 1$.*

*Then there is a constant $C = C(p,\lambda) \ge 1$, depending only on the unconditionality constant of $(f_i)$ and the constants $A_p$ and $B_p$ in Khintchine's Inequality (Lemma 6.2.3), so that for any $g = \sum_{i=1}^\infty a_i f_i \in \overline{\text{span}(f_i : i \in \mathbb{N})}$ it follows that*

$$\frac{1}{C}\Big\|\sum_{i=1}^\infty \Big(|a_i|^2|f_i|^2\Big)^{1/2}\Big\|_p \le \|g\|_p \le C\Big\|\sum_{i=1}^\infty \Big(|a_i|^2|f_i|^2\Big)^{1/2}\Big\|_p,$$

*which means that $\|\cdot\|_p$ is on $\overline{\text{span}(f_i : i \in \mathbb{N})}$ equivalent to the norm*

$$|||f||| = \Big\|\sum_{i=1}^\infty \Big(|a_i|^2|f_i|^2\Big)^{1/2}\Big\|_p = \Big\|\sum_{i=1}^\infty |a_i|^2|f_i|^2\Big\|_{p/2}^{1/2}.$$

*Proof.* For two positive numbers $A$ and $B$ and $c > 0$ we write: $A \sim_c B$ if $\frac{1}{c}A \le B \le cA$. Let $K_p$ be the *Khintchine constant for $L_p$*, i.e the smallest number so that for the Rademacher sequence $(r_n)$

$$\Big\|\sum_{i=1}^\infty a_i r_i\Big\|_p \sim_{K_p} \Big(\sum_{i=1}^\infty |a_i|^2\Big)^{1/2} \text{ for } (a_i) \subset \mathbb{K},$$

and let $C_u$ be the unconditionality constant of $(f_i)$, i.e.

$$\Big\|\sum_{i=1}^\infty \sigma_i a_i f_i\Big\|_p \sim_{C_u} \Big\|\sum_{i=1}^\infty a_i f_i\Big\|_p \text{ for } (a_i) \subset \mathbb{K} \text{ and } (\sigma_i) \subset \{\pm 1\}.$$

We consider $L_p[0,1]$ in a natural way as subspace of $L_p[0,1]^2$, with $\tilde{f}(s,t) := f(s)$ for $f \in L_p[0,1]$. Then let $r_n(t) = r_n(s,t)$ be the $n$th Rademacher function action on the second coordinate, i.e

$$r_n(s,t) = \text{sign}(\sin(2^n\pi t)),\ (s,t) \in [0,1]^2.$$

It follows from the $C_u$-unconditionality for any $(a_j)_{j=1}^m \subset \mathbb{K}$, that

$$\Big\|\sum_{j=1}^m a_j f_j(\cdot)\Big\|_p^p \sim_{C_u^p} \Big\|\sum_{j=1}^m a_j f_j(\cdot) r_j(t)\Big\|_p^p$$

$$= \int_0^1 \Big(\sum_{j=1}^m a_j f_j(s) r_j(t)\Big)^p ds \text{ for all } t\in[0,1],$$

and integrating over all $t\in[0,1]$ implies

$$\begin{aligned}
\Big\|\sum_{j=1}^m a_j f_j(\cdot)\Big\|_p^p &\sim_{C_u^p} \int_0^1\int_0^1 \left(\sum_{j=1}^m a_j f_j(s) r_j(t)\right)^p ds\,dt \\
&= \int_0^1\int_0^1 \left(\sum_{j=1}^m a_j f_j(s) r_j(t)\right)^p dt\,ds \text{(By Theorem of Fubini)} \\
&= \int_0^1 \Big\|\sum_{j=1}^m a_j f_j(s) r_j(\cdot)\Big\|_p^p ds \\
&\sim_{K_p^p} \int_0^1 \Big(\sum_{j=1}^m |a_j f_j(s)|^2\Big)^{p/2} ds = \Big\|\Big(\sum_{j=1}^m |a_j f_j|^2\Big)^{1/2}\Big\|_p^p,
\end{aligned}$$

which proves our claim using $C = K_p C_u$. □

**Theorem 6.2.6.** *Assume that $1<p<\infty$ and assume that $(f_j)$ is a normalized $\lambda$-unconditional sequence in $L_p[0,1]$ for some $\lambda\ge 1$. Let $C = C(p,\lambda)$ as in Theorem 6.2.5.*

*Then for all scalars $(a_j)_{j=1}^n \subset \mathbb{K}$*

$$\frac{1}{C}\Big(\sum_{j=1}^n |a_j|^2\Big)^{1/2} \le \Big\|\sum_{j=1}^n a_j f_j\Big\|_p \le C\Big(\sum_{j=1}^n |a_j|^p\Big)^{1/p} \text{ if } 1\le p\le 2, \text{ and} \tag{6.17}$$

$$\frac{1}{C}\Big(\sum_{j=1}^n |a_j|^p\Big)^{1/p} \le \Big\|\sum_{j=1}^n a_j f_j\Big\|_p \le C\Big(\sum_{j=1}^n |a_j|^2\Big)^{1/2} \text{ if } 2\le p<\infty. \tag{6.18}$$

*Proof.* We will show the inequalities (6.17) and (6.18) but with their middle terms replaced by the square-function norm. Having done that the claim will follow from Theorem 6.2.5.

Let $(a_i)_{i=1}^n \subset \mathbb{K}$. First we note that since for for $1\le s<t\le\infty$ the $\ell_s$-norm of a vector $(b_j)_{j\le n}$ is at least the $\ell_t$-norm we deduce

$$\begin{aligned}
&\Big\|\Big(\sum_{j=1}^n |a_j|^2|f_j|^2\Big)^{1/2}\Big\|_p^p \\
&\qquad = \int_0^1 \Big(\sum_{j=1}^n |a_j|^2 |f_j(t)|^2\Big)^{p/2} dt
\end{aligned}$$

$$
\begin{aligned}
&\begin{cases} \leq \displaystyle\int_0^1 \sum_{j=1}^n |a_j|^p |f_j(t)|^p \, dt & \text{if } 1 \leq p \leq 2, \\ \geq \displaystyle\int_0^1 \sum_{j=1}^n |a_j|^p |f_j(t)|^p \, dt & \text{if } 2 \leq p < \infty \end{cases} \\
&= \sum_{j=1}^n |a_j|^p
\end{aligned}
$$

which implies the second inequality of (6.17) and the first of (6.18). In order to verify the other two inequalities we observe that

$$
\begin{aligned}
&\Big\| \Big( \sum_{j=1}^n |a_j|^2 |f_j|^2 \Big)^{1/2} \Big\|_p \\
&\quad = \left\| \left( \sum_{j=1}^n \frac{|a_j|^2}{\sum_{i=1}^n |a_i|^2} |f_j|^2 \right)^{1/2} \right\|_p \Big( \sum_{i=1}^n |a_i|^2 \Big)^{1/2} \\
&\quad = \left\| \left( \left( \sum_{j=1}^n \frac{|a_j|^2}{\sum_{i=1}^n |a_i|^2} |f_j|^2 \right)^{p/2} \right)^{1/p} \right\|_p \Big( \sum_{i=1}^n |a_i|^2 \Big)^{1/2} \\
&\quad \begin{cases} \geq \Big\| \Big( \displaystyle\sum_{j=1}^n \frac{|a_j|^2}{\sum_{i=1}^n |a_i|^2} |f_j|^p \Big)^{1/p} \Big\|_p \Big( \displaystyle\sum_{i=1}^n |a_i|^2 \Big)^{1/2} & \text{if } 1 \leq p \leq 2 \\ \leq \Big\| \Big( \displaystyle\sum_{j=1}^n \frac{|a_j|^2}{\sum_{i=1}^n |a_i|^2} |f_j|^p \Big)^{1/p} \Big\|_p \Big( \displaystyle\sum_{i=1}^n |a_i|^2 \Big)^{1/2} & \text{if } 2 \leq p < \infty \end{cases} \\
&\quad \big[\, \xi \to \xi^{p/2} \text{ is convex if } p > 2 \text{ and concave if } 1 \leq p < 2 \big] \\
&\quad = \left( \int_0^1 \sum_{j=1}^n \frac{|a_j|^2}{\sum_{i=1}^n |a_i|^2} |f_j|^p \, dt \right)^{1/p} \Big( \sum_{i=1}^n |a_i|^2 \Big)^{1/2} = \Big( \sum_{i=1}^n |a_i|^2 \Big)^{1/2},
\end{aligned}
$$

which implies the first inequality of (6.17) and the second of (6.18). □

**Corollary 6.2.7.** *Let $1 \leq p < \infty$. Every normalized unconditional sequence $(f_j) \subset L_p[0,1]$ which consists of uniformly bounded functions is equivalent to the $\ell_2$-unit vector basis. In particular, if $p \neq 2$, $(f_j)$ cannot span all of $L_p[0,1]$.*

We will need Jensen's inequality.

**Theorem 6.2.8.** [Jensen's Inequality] *If $f : \mathbb{R} \to \mathbb{R}$ is convex and $g : [0,1] \to \mathbb{R}$, so that $g$ and $f \circ g$ are integrable. It follows that*

$$
f\left( \int_0^1 g(\xi) \, d\xi \right) \leq \int_0^1 f(g(\xi)) \, d\xi.
$$

*(Here $[0,1]$, together with the Lebesgues measure, can be replaced by any probability space)*

*Proof of Corollary 6.2.7.* Assume that $(f_j)$ is uniformly bounded, normalized and $\lambda$-unconditional. Let $C = \sup_{j\in\mathbb{N}} \|f_j\|_{L_\infty}$. If $1 \le p \le 2$ we deduce for $(a_j) \in c_{00}$ from the proof of 6.2.6 that

$$\Big\|\Big(\sum_{j=1}^\infty |a_i f_i|^2\Big)^{1/2}\Big\|_{L_p} \begin{cases} \le C\Big(\sum_{i=1}^\infty a_i^2\Big)^{1/2} \\ \ge \Big(\sum_{i=1}^\infty a_i^2\Big)^{1/2}. \end{cases}$$

Thus our claim follows for in the case that $1 \le p \le 2$.

If $p \ge 2$ we obtain for $(a_j) \in c_{00}$ first that

$$\Big\|\Big(\sum_{i=1}^\infty |a_i f_i|^2\Big)^{1/2}\Big\|_{L_p} \le C\Big(\sum_{i=1}^\infty |a_i|^2\Big)^{1/2}.$$

Since $\|f_i\|_{L_p} = 1$ and $\|f_i\|_\infty \le C$ for $i \in \mathbb{N}$, we deduce that

$$\begin{aligned} 1 &= \|f_i\|_p^p \\ &= \int_0^1 |f_i(t)|^p\,dt \\ &\le C^p m(\{|f_i| \ge 1/2C\}) + \frac{1}{(2C)^p} m(\{|f_i| < 1/2C\}) \\ &\le C^p m(\{|f_i| \ge 1/2C\}) + \frac{1}{2}, \end{aligned}$$

and thus

$$m(\{|f_i| \ge 1/2C\}) \ge \frac{1}{2C^p}.$$

We deduce that

$$\begin{aligned} \Big\|\Big(\sum_{i=1}^\infty |a_i f_i|^2\Big)^{1/2}\Big\|_{L_p} &= \Bigg(\int_0^1 \Big(\sum_{j=1}^\infty |a_j f_j(t)|^2\Big)^{p/2} dt\Bigg)^{1/p} \\ &\ge \frac{1}{2C}\Bigg(\int_0^1 \Big(\sum_{j=1}^\infty |a_j|^2 1_{\{|f_j|\ge 1/2C\}}\Big)^{p/2} dt\Bigg)^{1/p} \\ &\ge \frac{1}{2C}\Bigg(\int_0^1 \sum_{j=1}^\infty |a_j|^2 1_{\{|f_j|\ge 1/2C\}}\,dt\Bigg)^{1/2} \quad \text{(By Jensen's inequality)} \\ &\ge \frac{1}{2C}\frac{1}{2C^p}\Big(\sum_{j=1}^\infty |a_j|^2\Big)^{1/2} \end{aligned}$$

Our claim follows therefore from the equivalence between the $L_p$-norm and the square function norm in $L_p$ (Theorem 6.2.5). □

# Bibliography


[AW] F. Albiac and P. Wojtaszczyk. *Characterization of 1-greedy bases J. Approx. Theory*, 138(1):65–86, 2006.

[EW] I. S. Edelstein and P. Wojtaszcyk, *On projections and unconditional bases in direct sums of Banach spaces.* Stud. Math. **56** (3) (1976) 263 – 276.

[An] A. D. Andrew, *On subsequences of the Haar system in* $C(\Delta)$, Israel Journal of Math, **31**, N0. 1 (1978) 85–90.

[BCLT] J. Bourgain, P.G. Casazza, J. Lindenstrauss, and L. Tzafriri, *Banach spaces with a unique unconditional basis, up to permutation*, Memoirs Amer. Math. Soc. No. 322, 1985.

[DT] R. A. DeVore and V. N. Temlyakov, *Some remarks on greedy algorithms*, Adv. in Comp. Math. **5** (1996) 173 –187.

[DFOS] S.J. Dilworth, D. Freeman and E. Odell, and Th. Schlumprecht, *Greedy bases for Besov spaces.* Constr. Approx. **34** (2011), no. 2, 281296.

[DKSTW] S. Dilworth, D. Kutzarova, K. Shuman, V. Temlyakov, and P. Wojtaszczyk, *Weak convergence of greedy algorithms in Banach spaces.* J. Fourier Anal. Appl. **14** (2008), no. 5-6, 609 - - 628.

[DOSZ1] S. Dilworth, E. Odell, Th. Schlumprecht, and A. Zsak, *Renormings and symmetry properties of one-greedy bases*, J. Approx. Theory **163** (2011), no. 9, 1049 – 1075.

[DOSZ2] S. Dilworth, E. Odell, Th. Schlumprecht, and A. Zsak,*Partial Unconditionality*

[DOSZ3] S. Dilworth, E. Odell, Th. Schlumprecht, and A. Zsak, *On the convergence of greedy algorithms for initial segments of the Haar basis.* Math. Proc. Cambridge Philos. Soc. **148** (2010), no. 3, 519529.

[1] S. Dilworth, D. Kutzarova, E. Odell, Th. Schlumprecht, and P. Wojtaszczyk, *Weak thresholding greedy algorithms in Banach spaces.* J. Funct. Anal. **263** (2012), no. 12, 3900– 3921.

[2] S. Dilworth, D. Kutzarova, E. Odell, Th. Schlumprecht, and A. Zsak, *Renorming spaces with greedy bases*, J. Approx. Theory **188** (2014), 3956.

[DKK] Dilworth, S. J.; Kalton, N. J.; Kutzarova, Denka *On the existence of almost greedy bases in Banach spaces. Dedicated to Professor Aleksander Pełczyński on the occasion of his 70th birthday.* Studia Math. **159** (2003), no. 1, 67–101.

[En] P. Enflo,, *A counterexample to the approximation problem in Banach spaces.* Acta Math. **130** (1973), 309 – 317.

[GK] M. Ganichev and N. J. Kalton, *Convergence of the weak dual greedy algorithm in $L_p$-space*, J. Approx. Theory **124** (2003) 89 – 95.

[HMVZ] P. Hájek, V. Montesinos Santalucía, J. Vanderwerff, and V. Zizler, *Biorthogonal systems in Banach spaces.* CMS Books in Mathematics/Ouvrages de Mathmatiques de la SMC, 26. Springer, New York, (2008) xviii+339 pp.

[Jo] L. K. Jones, *A simple lemma on greedy approximation in Hilbert space and convergence rates for projection pursuit regression and neural network training*, Annals of Stat. **20** (1992) 608–613.

[Ka] Y. Katznelson, *An Introduction to Harmonic Analysis*

[KT1] S. V. Konyagin and V. N. Temlyakov, *A remark on greedy approximation in Banach spaces*, East Journal on Approximation, **5** no.3 (1999), 365–379.

[KT2] S. V. Konyagin and V. N. Temlyakov, *Rate of convergence of pure greedy algorithm.* East J. Approx. **5** (1999), no. 4, 493–499.

[LT] Lindenstrauss, J. and Tzafriri, L., "Classical Banach Spaces I – Sequence Spaces," Springer-Verlag, Berlin, 1979.

[Ma] A.I. Markushevich, *On a basis in the wide sense for linear spaces,* Dokl. Akad. Nauk. **41** (1943), 241– 244.

[OP] R.I. Ovsepian and A. Pełczyński, *On the existence of a fundamental total and bounded biorthogonal sequence in every separable Banach space, and related constructions of uniformly bounded orthonormal systems in $L_2$.* Studia Math. **54** (1975), no. 2, 149 – 159.

[Pe] A. Pełczyński, *All separable Banach spaces admit for every $\varepsilon > 0$ fundamental total and bounded by $1+\varepsilon$ biorthogonal sequences.* Studia Math. **55** (1976), no. 3, 295 – 304.

[Schl] Th. Schlumprecht, Course Notes for Functional Analysis, Fall 2012, `http://www.math.tamu.edu/~schlump/course_notes_FA2012.pdf`

[Sch2] Th. Schlumprecht, *Embedding Theorem of Banach spaces into Banach spaces with bases.* Adv. Math. **274** (2015), 833 – 880

[Schm] E. Schmidt, *Zur Theoreie der linearen und nicht linearen Integralgleichungen. I*, Math. Annalen **63** (1906) , 433 – 476.

[T1] V. Temlyakov, *Nonlinear methods of approximation.* Found. Comput. Math. **3** (2003), no. 1, 33–107

[T2] V. Temlyakov, *Relaxation in greedy approximation*, preprint.

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

[Wo1] P. Wojtaszczyk, *A mathematical introduction to wavelets.* London Mathematical Society Student Texats **37** (1997).

[Wo2] P. Wojtaszczyk, *Greedy algorithm for general biorthogonal systems.* J. Approx. Theory **107** (2000), no. 2, 293 314.