Available online at www.sciencedirect.com
SciVerse ScienceDirect
Journal of Approximation Theory 164 (2012) 534–571
JOURNAL OF Approximation Theory
www.elsevier.com/locate/jat

Full length article

Approximation schemes satisfying Shapiro's Theorem

J.M. Almira[a,*], T. Oikhberg[b,c]

[a] *Departamento de Matemáticas, Universidad de Jaén, E.P.S. Linares, C/Alfonso X el Sabio, 28, 23700 Linares (Jaén), Spain*
[b] *Department of Mathematics, The University of California at Irvine, Irvine, CA 92697, United States*
[c] *Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States*

Received 25 October 2010; received in revised form 17 December 2011; accepted 20 January 2012
Available online 27 January 2012

Communicated by Frank Deutsch

Abstract

An approximation scheme is a family of homogeneous subsets (A_n) of a quasi-Banach space X, such that $A_1 \subsetneq A_2 \subsetneq \ldots \subsetneq X$, $A_n + A_n \subset A_{K(n)}$, and $\overline{\bigcup_n A_n} = X$. Continuing the line of research originating at the classical paper [8] by Bernstein, we give several characterizations of the approximation schemes with the property that, for every sequence $\{\varepsilon_n\} \searrow 0$, there exists $x \in X$ such that $\operatorname{dist}(x, A_n) \neq \mathbf{O}(\varepsilon_n)$ (in this case we say that $(X, \{A_n\})$ satisfies Shapiro's Theorem). If X is a Banach space, $x \in X$ as above exists if and only if, for every sequence $\{\delta_n\} \searrow 0$, there exists $y \in X$ such that $\operatorname{dist}(y, A_n) \geq \delta_n$. We give numerous examples of approximation schemes satisfying Shapiro's Theorem.

Keywords: Approximation scheme; Approximation error; Bernstein's Lethargy; Approximation by dictionary

1. Introduction and motivation

One of the most remarkable early results in the constructive theory of functions is Bernstein Lethargy Theorem: if $X_0 \subsetneq X_1 \subsetneq X_2 \subsetneq \cdots \subsetneq X$ is an ascending chain of finite dimensional vector subspaces of a Banach space X, and $\{\varepsilon_n\} \searrow 0$ is a non-increasing sequence of positive real numbers that converges to zero, then there exists an element $x \in X$ such that the n-th error of best approximation by elements of X_n satisfies $E(x, X_n) = \varepsilon_n$ for all $n \in \mathbb{N}$. Here and throughout the

* Corresponding author.
E-mail addresses: jmalmira@ujaen.es (J.M. Almira), toikhber@math.uci.edu (T. Oikhberg).

0021-9045/$ - see front matter
doi:10.1016/j.jat.2012.01.005

paper, we write $E(x, A) = \inf_{a \in A} \|x - a\|$ (x and A are an element and a subset of a quasi-Banach space X, respectively, (for the concept of quasi-Banach space, see Definition 1.2). Furthermore, the notation $\{\alpha_n\} \searrow 0$ means that the sequence (α_n) is non-increasing, and $\lim \alpha_n = 0$.

The result quoted above was first obtained in 1938 by Bernstein [8] for $X = C([0, 1])$ and $X_n = \Pi_n$, the vector space of real polynomials of degree $\leq n$. The case of arbitrary finite dimensional X_n is treated, for instance, in [28, p. 94ff], [61, p. 40ff], [57, Section II.5.3].

There are very few generalizations of Bernstein's result to arbitrary chains of (possibly infinite dimensional) closed subspaces $X_1 \subsetneq X_2 \subsetneq \dots$ of a Banach space X. The results due to Tjuriemskih [62] and Nikolskii [46,47] (see also [57, Section I.6.3]) assert that a sufficient (resp. necessary) condition for the existence of $x \in X$ verifying $E(x, X_n) = \varepsilon_n$ is that X is a Hilbert space (resp. X is reflexive). These results were proved independently and by other means by Almira and Luther [3,4] and Almira and Del Toro [1]. Moreover, in [2] it was shown that if X is a reflexive Banach space and $\{0\} \subset X_1 \subset X_2 \subset \cdots$ is an infinite chain of closed subspaces of X, then for every pair of sequences of positive numbers $\{\varepsilon_n\} \searrow 0$, $\{\delta_n\} \searrow 0$, there is an element $x \in X$ such that $E(x, X_n)/\varepsilon_n$ converges to zero, but $E(x, X_n)/\varepsilon_n \neq \mathbf{O}(\delta_n)$. Also, Bernstein Lethargy Theorem has been generalized to chains of finite-dimensional subspaces in non-Banach spaces (such as SF-spaces) by Lewicki [37,38]. These two approaches were successfully combined by Micherda [44].

Thanks to the work by Plesniak [54], the lethargy theorem has become a very useful tool for the theory of quasi-analytic functions of several complex variables.

In 1964 Shapiro [56] used Baire Category Theorem to prove that, for any sequence $X_1 \subsetneq X_2 \subsetneq \dots \subsetneq X$ of closed (not necessarily finite dimensional) subspaces of a Banach space X, and any sequence $\{\varepsilon_n\} \searrow 0$, there exists an $x \in X$ such that $E(x, X_n) \neq \mathbf{O}(\varepsilon_n)$. This result was strengthened by Tjuriemskih [63] who, under the very same conditions of Shapiro's Theorem, proved the existence of $x \in X$ such that $E(x, X_n) \geq \varepsilon_n$, $n = 0, 1, 2, \dots$. Moreover, Borodin [9] gave an easy proof of this result and proved that, for arbitrary infinite dimensional Banach spaces X and for sequences $\{\varepsilon_n\} \searrow 0$ satisfying $\varepsilon_n > \sum_{k=n+1}^{\infty} \varepsilon_k$, $n = 0, 1, 2, \dots$, there exists $x \in X$ such that $E(x, X_n) = \varepsilon_n$, $n = 0, 1, 2, \dots$.

However, approximation by linear subspaces of a Banach space is very restrictive. There are many other choices of approximation processes such as rational approximation, approximation by splines with of without free knots, n-term approximation with dictionaries of different kinds, and approximation of operators by operators of finite rank, just to mention a few of them. Do the results of Bernstein, Shapiro and Tjuriemskih hold in this setting, too? The following startling result of Brudnyi can be found in [12, Theorem 4.5.12]:

Theorem 1.1. *Suppose* $\{0\} = A_0 \subset A_1 \subset \cdots \subset A_n \subset$ *is an infinite chain of subsets of a Banach space X, satisfying the following conditions:* $A_n + A_m \subset A_{n+m}$ *for all* $n, m \in \mathbb{N}$*;* $\lambda A_n \subset A_n$ *for all* $n \in \mathbb{N}$ *and all scalars* λ*;* $\bigcup_{n \in \mathbb{N}} A_n$ *is dense in X; and*

$$\gamma = \inf_{n \in \mathbb{N}} \sup_{x \in A_{n+1}, \|x\| \leq 1} E(x, A_n) > 0. \tag{1.1}$$

Then here exists a constant $c = c(\gamma)$ *such that for every non-increasing convex sequence* $\{\varepsilon_n\}_{n=0}^{\infty} \searrow 0$ *there exists* $x \in X$ *such that* $E(x, A_n) \geq \varepsilon_n$ *for all* $n \in \mathbb{N}$*, and* $E(x, A_n) \leq c\varepsilon_n$ *for infinitely many values of n.*

Recall that a sequence ε_n is called *convex* if, for every $n, \varepsilon_n \leq (\varepsilon_{n-1} + \varepsilon_{n+1})/2$. By [25, pp. 113–114], for any sequence $\{\varepsilon_n\} \searrow 0$, there is a convex sequence $\{\xi_n\} \searrow 0$ such that $\xi_n \geq \varepsilon_n$ for all $n \in \mathbb{N}$. Thus, we do not need to assume the convexity of $\{\varepsilon_n\}$ to show the existence of $x \in X$ satisfying $E(x, A_n) \geq \varepsilon_n$ for any $n \in \mathbb{N}$.

In this paper, we are concerned with generalizations of results of Brudnyi and Shapiro quoted above for general approximation schemes, defined by Pietsch [51] to produce a unified approach to diverse phenomena of approximation theory. Instead of working with Banach spaces, we work with a more general class of quasi-Banach spaces.

Definition 1.2. Let X be a real or complex vector space. We say that the map $\|\cdot\| : X \to \mathbb{R}^+$ is a quasi-norm over X if it satisfies the following three properties:

(i) $\|x\| = 0$ if and only if $x = 0$.
(ii) There exists a constant $C_X \geq 1$ such that $\|x + y\| \leq C_X(\|x\| + \|y\|)$ for all $x, y \in X$.
(iii) $\|\lambda x\| = |\lambda|\, \|x\|$ for all $x \in X$ and all scalar λ.

We say that $(X, \|\cdot\|)$ is a quasi-Banach space if it is complete in the (metrizable) topology, determined by $\|\cdot\|$.

Note that, if $(X, \|\cdot\|)$ satisfies the above conditions with $C_X = 1$, then it is a Banach space. An important class of quasi-Banach spaces is formed by *p-normed spaces*, for $0 < p \leq 1$. These are spaces $(X, \|\cdot\|)$ which satisfy (i), (iii), as well as

(ii)′ For any $x, y \in X$, $\|x + y\|^p \leq \|x\|^p + \|y\|^p$.

Aoki–Rolewicz theorem states that any quasi-normed space can be equipped with an equivalent quasi-norm $|||\cdot|||$ for which there exists $p \in (0, 1]$ such that $|||x + y|||^p \leq |||x|||^p + |||y|||^p$ for any $x, y \in X$ (see e.g. [20, Theorem 2.1.1] or [35, pp. 7–8]).

Quasi-normed spaces were introduced by Hyers in 1938 [33], under the name of pseudo-normed spaces. There, it was shown that a topological linear space X is locally bounded if and only if its topology can be generated by a quasi-norm on X. The term "quasi-normed" was introduced in 1943 by Bourgin [11] (see also [52]).

The best known examples of quasi-Banach spaces are the Lebesgue spaces ℓ_p, $L_p(\Omega)$, the Hardy space $H_p(0 < p < 1)$ (these spaces are p-normed), as well as Lorentz spaces (see e.g. [12, Section 1.9] for the definition of these spaces). Quasi-Banach spaces have been widely used in functional analysis—for instance, in the study of operator ideals [53], and in interpolation theory [12] Abstract approximation spaces (introduced in [51], and studied in, for instance, [3,4], as well as in Section 5 of this paper) are quasi-Banach. We refer the reader to [34] for an up-to-date survey of quasi-Banach spaces.

Definition 1.3. Suppose X is a quasi-Banach space, and let $A_0 \subset A_1 \subset \cdots \subset X$ be an infinite chain of subsets of X, where all inclusions are strict. We say that $(X, \{A_n\})$ is an *approximation scheme* (or that (A_n) is an approximation scheme in X) if:

(i) There exists a map $K : \mathbb{N} \to \mathbb{N}$ such that $K(n) \geq n$ and $A_n + A_n \subseteq A_{K(n)}$ for all $n \in \mathbb{N}$ (we can assume that K is increasing).
(ii) $\lambda A_n \subset A_n$ for all $n \in \mathbb{N}$ and all scalars λ.
(iii) $\bigcup_{n\in\mathbb{N}} A_n$ is dense in X.

Note that part (ii) of this definition implies that, for any $n \in \mathbb{N}$ and $\lambda \neq 0$, we have $\lambda A_n = A_n$. Indeed, we already have the inclusion $\lambda A_n \subset A_n$. To prove the converse, note that $A_n = \lambda \cdot \lambda^{-1} A_n \subset \lambda A_n$, since $\lambda^{-1} A_n \subset A_n$.

One example of an approximation scheme is an increasing chain of linear subspaces of X, whose union is dense. Then we can take $K(n) = n$. Further examples of approximation schemes can be found throughout the paper.

Definition 1.4. Following e.g. [1,2], we say that the approximation scheme $(X, \{A_n\})$ *satisfies Shapiro's Theorem* if for any non-increasing sequence $\{\varepsilon_n\} \searrow 0$ there exists some $x \in X$ such that $E(x, A_n) \neq \mathbf{O}(\varepsilon_n)$. In other words, for each $c > 0$, we have $E(x, A_n) > c\varepsilon_n$ for infinitely many values of n.

Section 2 is devoted to describing approximation schemes satisfying Shapiro's Theorem (Theorems 2.2 and 2.6). In Section 3, we prove that for an approximation scheme in a Banach space X, satisfying Shapiro's Theorem is equivalent to (a weakened version of) Brudnyi's Theorem 1.1 (Theorem 3.4, Corollary 3.7). Section 4 shows some examples of "pathological" approximation schemes failing Shapiro's Theorem. Section 5 studies the relationship between approximation schemes that satisfy Shapiro's Theorem, and those verifying the abstract versions of Jackson's and Bernstein's inequalities. Section 6 contains many examples of approximation schemes which do satisfy Shapiro's Theorem. Finally, Section 7 examines the related question of controlling the rate of decay of the best approximation errors.

2. Shapiro's Theorem

Throughout this paper, we work with approximation schemes in infinite dimensional quasi-Banach spaces. The proposition below shows that a finite dimensional space cannot "host" an approximation scheme.

Proposition 2.1. *Suppose X is a finite dimensional space, and the family of its subsets $A_0 \subset A_1 \subset \cdots \subset A_n \subset \cdots \subset X$ satisfies* (i)–(iii) *of Definition* 1.3. *Then there exists $N \in \mathbb{N}$ such that $A_N = X$.*

Proof. For each n, $X_n = \mathbf{span}[A_n]$ is a closed subspace of X. Then $X_1 \subset X_2 \subset \cdots$. As $\cup_n A_n$ is dense in X, we conclude that $X_n = X$ for some n. By Caratheodory's Theorem, and by the homogeneity of the set A_n, any $x \in X$ can be represented as $x = \sum_{k=1}^{M} \alpha_k a_k$, with $a_k \in A_n$, and $\alpha_k \in \mathbb{R}$ (here $M = \dim X + 1$). Therefore, $X = A_N$, where $N = K(\cdots(K(n))\cdots)$ (M times). $\square$

Note that if $((X, \|\cdot\|), \{A_n\})$ satisfies Shapiro's Theorem, and $|||\cdot|||$ is an equivalent quasi-norm on X, then $((X, |||\cdot|||), \{A_n\})$ also satisfies Shapiro's Theorem. This remark will be particularly useful for quasi-normed spaces X, as it allows us to deploy Aoki–Rolewicz theorem (stated in Section 1): any quasi-normed space can be equipped with an equivalent quasi-norm $|||\cdot|||$ for which there exists $p \in (0, 1]$ such that $|||x + y|||^p \leq |||x|||^p + |||y|||^p$ for any $x, y \in X$.

Theorem 2.2. *Suppose (A_n) is an approximation scheme in a quasi-Banach space X. The following are equivalent:*

(a) *The approximation scheme $(X, \{A_n\})$ satisfies Shapiro's Theorem.*
(b) *There exists a constant $c > 0$ and an infinite set $\mathbb{N}_0 \subseteq \mathbb{N}$ such that for all $n \in \mathbb{N}_0$, there exists $x_n \in X \setminus \overline{A_n}$ which satisfies $E(x_n, A_n) \leq cE(x_n, A_{K(n)})$.*
(c) *There is no sequence $\{\varepsilon_n\} \searrow 0$ such that $E(x, A_n) \leq \varepsilon_n \|x\|$ for all $x \in X$ and $n \in \mathbb{N}$.*

For the proof we need:

Lemma 2.3. *Let $h : \mathbb{N} \to \mathbb{N}$ be a map such that $h(n) \geq n$ for all n, and let $\{\varepsilon_n\} \searrow 0$. Then there exists a sequence $\{\xi_n\} \searrow 0$ such that $\xi_n \geq \varepsilon_n$ and $\xi_n \leq 2\xi_{h(n)}$ for every n.*

Proof. Passing from the original function h to (say) $h'(n) = \max_{1 \leq k \leq n} h(k) + n$, we can assume that (i) $h(n) > n$ for every n, and (ii) the function h is strictly increasing. Set $m_0 = 0$, and, for $k \geq 1, m_k = h(m_{k-1})$. Set $\beta_0 = \varepsilon_1$, and $\beta_k = \max\{\varepsilon_{m_k}, \beta_{k-1}/2\}$ for $k \geq 1$. For $n \in \mathbb{N}$, find $k \geq 0$ such that $n \in [m_k, m_{k+1})$, and set $\xi_n = \beta_k$.

Then the sequence (ξ_n) has the desired properties. For $n \in [m_k, m_{k+1})$, $\xi_n = \beta_k \geq \varepsilon_{m_k} \geq \varepsilon_n$. Furthermore, as h is increasing, $h(n) \in [m_{k+1}, m_{k+2})$, hence $\xi_{h(n)} = \beta_{k+1} \geq \beta_k/2 = \xi_n/2$. It remains to show that $\lim \xi_n = 0$, or in other words, that $\lim \beta_k = 0$. If $\beta_k = \varepsilon_{m_k}$ for infinitely many values of k, then $\lim \beta_k = \lim \varepsilon_{m_k} = 0$. Otherwise, $\beta_k = \beta_{k-1}/2$ for any $k \geq k_0$. In this case, too, $\lim \beta_k = 0$. □

Proof of Theorem 2.2. As X is a quasi-Banach space, there exists a constant C_X such that $\|x + y\| \leq C_X(\|x\| + \|y\|)$ for any $x, y \in X$.

(b) $\Rightarrow$ (a): As a first step, we prove the existence of $x \in X$ satisfying $E(x, A_n) \neq \mathbf{O}(\varepsilon_n)$ under the additional assumption that $\varepsilon_n \leq 2\varepsilon_{K(n+1)-1}$ for all $n \in \mathbb{N}$. Assume, for the sake of contradiction, that $E(x, A_n) = \mathbf{O}(\varepsilon_n)$ for all $x \in X$. Then $X = \bigcup_{m=1}^{\infty} \Gamma_m$, where $\Gamma_\alpha = \{x \in X : E(x, A_n) \leq \alpha\varepsilon_n, n = 0, 1, 2, \ldots\}$ ($\alpha > 0$). The sets Γ_m are closed subsets of X. Furthermore, $E(-x, A_n) = E(x, A_n)$ for all n, hence $\Gamma_m = -\Gamma_m$ for all m. Finally,

$$\mathbf{conv}(\Gamma_m) \subset \Gamma_{2mC_X} \tag{2.1}$$

(here, $\mathbf{conv}(S)$ stands for the convex hull of a set S). Indeed, suppose $x, y \in \Gamma_m$, and $\lambda \in [0, 1]$. Recalling the inclusion $A_n + A_n \subset A_{K(n)}$, we see that, for every n,

$$\begin{aligned} E(\lambda x + (1-\lambda)y, A_{K(n)}) &= \inf_{g \in A_{K(n)}} \|\lambda x + (1-\lambda)y - g\| \\ &\leq \inf_{a,b \in A_n} \|\lambda(x-a) + (1-\lambda)(y-b)\| \\ &\leq C_X \left[\inf_{a \in A_n} \|\lambda(x-a)\| + \inf_{b \in A_n} \|(1-\lambda)(y-b)\| \right] \\ &= \lambda C_X E(x, A_n) + (1-\lambda) C_X E(y, A_n) \leq m C_X \varepsilon_n. \end{aligned}$$

For an arbitrary j, find n such that $K(n) \leq j < K(n+1)$ (for simplicity, we set $K(0) = 0$). Then

$$\begin{aligned} E(\lambda x + (1-\lambda)y, A_j) &\leq E(\lambda x + (1-\lambda)y, A_{K(n)}) \leq m C_X \varepsilon_n \leq 2m C_X \varepsilon_{K(n+1)-1} \\ &\leq 2m C_X \varepsilon_j, \end{aligned}$$

which implies $\lambda x + (1-\lambda)y \in \Gamma_{2mC_X}$, thus proving (2.1).

By Baire Category theorem, there exists some $m_0 \in \mathbb{N}$ such that Γ_{m_0} has non-empty interior. That is, there exists a ball $B(x, r) \subset \Gamma_{m_0}$ with $r > 0$. By symmetry, $-B(x, r) \subset \Gamma_{m_0}$. By (2.1),

$$B(0, r) \subset \frac{1}{2}\big(B(-x, r) + B(x, r)\big) \subset \Gamma_{2m_0C_X}.$$

Hence, $\frac{r}{\|x\|} x \in \Gamma_{2m_0C_X}$ for every $x \in X$, and the inequality

$$E(x, A_n) \leq \frac{\|x\|}{r} 2m_0 C_X \varepsilon_n$$

holds for all $x \in X$ and all $n \in \mathbb{N}$.

For $n \in \mathbb{N}_0$, find $a_n \in A_n$ such that $\|x_n - a_n\| \leq 2E(x_n, A_n)$, where $\{x_k\}_{k\in\mathbb{N}_0}$ is the sequence of elements of X given by condition (b). Take $y_n = x_n - a_n$. Then

$$\|y_n - b_n\| = \|x_n - (a_n + b_n)\| \geq E(x_n, A_{K(n)}) \geq \frac{1}{c}E(x_n, A_n) \geq \frac{1}{2c}\|y_n\|$$

for all $b_n \in A_n$. Hence

$$\frac{1}{2c}\|y_n\| \leq E(y_n, A_n) \leq \frac{\|y_n\|}{r}2m_0 C_X \varepsilon_n,$$

and consequently, $1/(2c) \leq 2m_0 C_X \varepsilon_n / r$ for all $n \in \mathbb{N}_0$. This contradicts $\varepsilon_n \to 0$. Thus, for every sequence $\{\varepsilon_n\} \searrow 0$ satisfying $\varepsilon_n \leq 2\varepsilon_{K(n+1)-1}$ $(n \in \mathbb{N})$, there exists $x \in X$ such that $E(x, A_n) \neq \mathbf{O}(\varepsilon_n)$.

Now suppose the sequence $\{\varepsilon_n\} \searrow 0$ is arbitrary. Applying Lemma 2.3 to $\{\varepsilon_n\}_{n=0}^{\infty}$ and the map $h(n) = K(n+1) - 1$, we obtain a sequence $\{\xi_n\}_{n=0}^{\infty}$ satisfying $\varepsilon_n \leq \xi_n \leq 2\xi_{K(n+1)-1}$ for all $n \in \mathbb{N}$. By the above, there exists $x \in X$ such that $E(x, A_n) \neq \mathbf{O}(\xi_n)$, which implies $E(x, A_n) \neq \mathbf{O}(\varepsilon_n)$. This ends the proof of (b) $\Rightarrow$ (a).

(a) $\Rightarrow$ (b): If $X = \cup_{n=0}^{\infty} \overline{A_n}$, then both (a) and (b) are false, since in this case, for any $x \in X$ there exists $n \in \mathbb{N}$ such that $E(x, A_n) = 0$. Suppose $X \neq \cup_{n=0}^{\infty} \overline{A_n}$, and (b) is false. Then the sequence $\{c_n\}_{n=0}^{\infty} \subset [0, \infty)$, given by

$$c_n = \inf_{x \in X \setminus \overline{A_{K(n)}}} \frac{E(x, A_n)}{E(x, A_{K(n)})},$$

has no bounded subsequences, hence $\lim_{n\to\infty} c_n = \infty$. Set $\varepsilon_k = 1/c_n$ for $K(n) \leq k < K(n+1)$ and let $\{\varepsilon_n^*\}$ denote the non-increasing rearrangement of the sequence $\{\varepsilon_n\} \in c_0(\mathbb{N})$. For any $x \in X \setminus \cup_{n=0}^{\infty} \overline{A_n}$, and any $k \in [K(n), K(n+1))$, we have

$$E(x, A_k) \leq E(x, A_{K(n)}) \leq \frac{1}{c_n}E(x, A_n) \leq \frac{1}{c_n}\|x\| = \varepsilon_k\|x\| \leq \varepsilon_k^*\|x\|, \tag{2.2}$$

hence $E(x, A_k) = \mathbf{O}(\varepsilon_k^*)$, and (a) is also false.

(a) $\Rightarrow$ (c) is clear. On the other hand, if (a) is false then (b) is also false, so that (2.2) holds true. This implies that $E(x, A_k) \leq \varepsilon_k^*\|x\|$, for the sequence $\{\varepsilon_k^*\} \searrow 0$ described above. □

Remark 2.4. It follows from Theorem 2.2 that every non-trivial linear approximation scheme (i.e., every approximation scheme verifying $K(n) = n$ and $\overline{A_n} \neq X$ for all n) satisfies Shapiro's Theorem. In particular, this extends Shapiro's result to the quasi-Banach setting.

A different proof of Theorem 2.2 was given by Almira and Del Toro in [1,2]. That proof used some general theory of approximation spaces, introduced by Almira and Luther in [3,4]. The proof presented here is self-contained, avoids the theory of generalized approximation spaces, and follows a more classical line of thinking.

One of our main tools for verifying that an approximation scheme satisfies Shapiro's Theorem is property (P).

Definition 2.5. We say that an approximation scheme $(X, \{A_n\})$ satisfies property (P) (with constants $a, b > 0$) if for every $n \in \mathbb{N}, n > 0$, there exists an element $x \in X$ with $\|x\| = 1$ such that $E(x, A_n) \geq \frac{1}{an^b}$.

Theorem 2.6. *Suppose an approximation scheme* $(X, \{A_n\})$ *satisfies property* (P)*, and there exists* $c > 1$ *such that* $A_n + A_n \subseteq A_{cn}$ *for any* $n \in \mathbb{N}$*. Then* $(X, \{A_n\})$ *satisfies Shapiro's Theorem.*

Proof. Assume, for the sake of contradiction, that $(X, \{A_n\})$ fails Shapiro's Theorem. By Theorem 2.2, for any $C > 1$ there exists $N \in \mathbb{N}$ such that $E(x, A_n) \geq CE(x, A_{cn})$ for any $x \in X$ and $n \geq N$. Pick $C > c^b$ and select k to satisfy $aN^b < \frac{C^k}{c^{bk}}$ (here, a and b are as in Definition 2.5). Take $x \in X$ with $\|x\| = 1$ and $E(x, A_{c^k N}) \geq \frac{1}{a(c^k N)^b}$. Then

$$1 = \|x\| \geq E(x, A_N) \geq CE(x, A_{cN}) \geq C^2 E(x, A_{c^2 N}) \geq \cdots \geq C^k E(x, A_{c^k N}),$$

so that

$$\frac{1}{a(c^k N)^b} \leq E(x, A_{c^k N}) \leq C^{-k},$$

hence $a(c^k N)^b \geq C^k$, which contradicts our choice of k. □

Section 6 contains several examples where Property (P) is used to show that an approximation scheme satisfies Shapiro's Theorem.

3. A comparison with Brudnyi's Theorem

To proceed, we need to introduce some notation. Recall that, for $x \in X$ and $A \subset X$, we define $E(x, A) = \inf_{a \in A} \|x - a\|$. Furthermore, for subsets A, B of X, we define $E(B, A) = \sup_{b \in B} E(b, A)$ (note that $E(A, B)$ may be different from $E(B, A)$). We denote by $S(X)$ the unit sphere of a quasi-Banach space X.

Definition 3.1. We say that the approximation scheme $(X, \{A_n\})$ *satisfies Brudnyi's condition* if

$$\gamma = \inf_{n \in \mathbb{N}} \sup_{x \in A_{n+1}, \|x\| \leq 1} E(x, A_n) > 0. \tag{3.1}$$

We say that $(X, \{A_n\})$ *satisfies weak Brudnyi's condition with constant* $c \in (0, 1]$ if $E(S(X), A_n) \geq c$ for all $n \in \mathbb{N}$.

Note that Brudnyi's condition implies the "jump condition" from Theorem 2.2(b), that is, the existence (for each $n \in \mathbb{N}$) of $x_n \in X$ satisfying $E(x_n, A_n) \leq CE(x_n, A_{K(n)})$. This implication holds for general approximation schemes, and not just for the case $K(n) = 2n$, covered by Brudnyi's Theorem. Indeed, applying (3.1) to $A_{K(n)}$, we obtain $x_n \in A_{K(n)+1}$ such that $\|x_n\| = 1$ and $E(x_n, A_{K(n)}) \geq \gamma$. Then

$$E(x_n, A_n) \leq 1 = C\gamma \leq CE(x_n, A_{K(n)}),$$

where $C = 1/\gamma$.

However, there exist approximation schemes failing Brudnyi's condition (3.1), for which one can obtain a prescribed rate of decay of $(E(x, A_n))$.

Theorem 3.2. *There exists an approximation scheme* (A_n) *in the space* c_0*, such that* $A_m + A_n \subset A_{\max\{m,n\}+1}$ *for any* $m, n \in \mathbb{N}$*, and:*

(1) *Brudnyi's condition* (3.1) *is not satisfied.*
(2) *For any* $\{\varepsilon_n\} \searrow 0$*, there exists* $x \in c_0$ *such that* $E(x, A_{2n-1}) = \varepsilon_n$ *for any* $n \in \mathbb{N}$*. Consequently,* $(c_0, \{A_n\})$ *satisfies Shapiro's Theorem.*

Proof. We introduce the sets B_n : $B_0 = \{\mathbf{0}\}$, $B_1 = \{(x_1, 0, \ldots, 0, \ldots) : x_1 \in \mathbb{R}\}$ and, for $n \geq 1$,

$$B_{n+1} = \left\{ (x_1, \ldots, x_{n+1}, 0, \ldots) : (x_1, \ldots, x_n) \in \mathbb{R}^n \text{ and } |x_{n+1}| \leq \frac{\sup_{k\leq n} |x_k|}{n+1} \right\}.$$

Let us also introduce the sets $\Pi_n = \{(x_1, \ldots, x_n, 0, \ldots) : (x_1, \ldots, x_n) \in \mathbb{R}^n\}$. Consider the approximation scheme $(X, \{A_n\}_{n=0}^{\infty})$, where $A_0 = B_0, A_1 = B_1 = \Pi_1, A_2 = B_2, A_3 = \Pi_2, A_4 = B_3, A_5 = \Pi_3, \cdots$. Clearly, $A_0 \subsetneq A_1 \subsetneq A_2 \subsetneq \cdots \subsetneq c_0$, $A_n + A_m \subset A_{\max\{n,m\}+1} \subset A_{n+m}$ for any m and n, and $\overline{\bigcup_n A_n} = c_0$. Furthermore, if $\{\varepsilon_n\} \searrow 0$, then $x = (\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots) \in c_0$ satisfies $E(x, A_{2n-1}) = \varepsilon_n$ for any n.

However, there is no $\gamma > 0$ such that $E(S(X) \cap A_{n+1}, A_n) \geq \gamma$ for every n. Indeed, it is easy to see that

$$E(S(X) \cap A_{2k}, A_{2k-1}) = E(S(X) \cap B_{k+1}, \Pi_k) = \frac{1}{k+1}.$$

Thus, the approximation scheme (A_n) has the desired properties. $\square$

The following definition is inspired by the recent papers by Deutsch and Hundal [18,19]:

Definition 3.3. Let $(X, \{A_n\})$ be an approximation scheme. We say that the distance functional $E(\cdot, A_n)$ *converges arbitrarily slowly to* 0 if for every sequence $\{\varepsilon_n\} \searrow 0$ there exists $x \in X$ such that $E(x, A_n) \geq \varepsilon_n$ for all n.

Theorem 3.4. *Suppose (A_n) is an approximation scheme in a Banach space X. Then the following claims are equivalent:*

(i) *$(X, \{A_n\})$ satisfies Shapiro's Theorem.*
(ii) *The distance functional $E(\cdot, A_n)$ converges arbitrarily slowly to* 0.

For the proof we need two lemmas. The first one will be stated for the quasi-Banach setting because we will use it later (see Corollary 3.7) to give a new characterization of approximation schemes that satisfy Shapiro's Theorem.

Lemma 3.5. *If X is a quasi-Banach space, and an approximation scheme $(X, \{A_n\})$ satisfies Shapiro's Theorem, then $E(S(X), A_n) = 1$ for $n = 0, 1, 2, \ldots$.*

Proof. Suppose otherwise. Then there exists $n \in \mathbb{N}$ such that $E(S(X), A_n) = c_1 < 1$. Find $c \in (c_1, 1)$. Then every $x \in X$ admits a decomposition $x = y_1 + z_1$ with $y_1 \in A_n$ and $\|z_1\| < c\|x\|$. Furthermore, $z_1 = y_2 + z_2$, with $y_2 \in A_n$, and $\|z_2\| < c\|z_1\| < c^2\|x\|$. Continuing in the same way, for any $k \in \mathbb{N}$ we get a decomposition $x = y_1 + y_2 + \cdots + y_k + z_k$, with $y_1, y_2, \ldots, y_k \in A_n$, and $\|z_k\| < c^k\|x\|$. Now, the sum $y_1 + y_2 + \cdots + y_k$ belongs to $A_{K^k(n)}$ (here, $K^k(n) = K(K(\ldots K(n)\ldots))$ (k times), so that $E(x, A_{K^k(n)}) \leq c^k$ for $k = 0, 1, 2, \ldots$ and $\|x\| \leq 1$. It follows that

$$E(x, A_{K^k(n)}) \leq c^k\|x\| \quad \text{for } k = 0, 1, 2, \ldots \text{ and } x \in X. \tag{3.2}$$

Now let $\varepsilon_i = c^k$ for $K^{k-1}(n) < i \leq K^k(n)$. For such i, and $x \in X$,

$$E(x, A_i) \leq E(x, A_{K^k(n)}) \leq c^k\|x\| = \varepsilon_i\|x\|.$$

As $\{\varepsilon_i\} \searrow 0$, this contradicts our assumption that $(X, \{A_n\})$ satisfies Shapiro's Theorem. $\square$

Lemma 3.6. *Suppose X is a Banach space, and (A_i) is an approximation scheme in X, satisfying Shapiro's Theorem. Then there exists a sequence of natural numbers $s_0 = 0 < s_1 < s_2 < \cdots$, such that $(X, \{A_{s_i}\})$ satisfies the hypotheses of Brudnyi's Theorem 1.1.*

Proof. Throughout, we are assuming that the function K appearing in the definition of an approximation scheme (Definition 1.3) is non-decreasing. It suffices to select $s_0 = 0 < s_1 < s_2 < \cdots$ in such a way that the sets $B_i = A_{s_i}$ satisfy (i) $B_n + B_m \subset B_{\max\{n,m\}+1}$ for all $n, m \in \mathbb{N}$, and (ii) $E(B_{n+1} \cap S(X), B_n) \geq 1/2$ for any $n \in \mathbb{N}$. Suppose $s_0 = 0 < s_1 < \cdots < s_k$ have already been selected in such a way that the (i) and (ii) are satisfied for $0 \leq m, n \leq k-1$. By Lemma 3.5, $E(S(X), B_k) = 1$. As $\overline{\cup_\ell A_\ell} = X$, there exist $\ell > K(s_k)$ and $x \in A_\ell \cap S(X)$ such that $E(x, B_k) > 1/2$. Then $s_{k+1} = \ell$ works for us. Indeed, $E(S(X) \cap B_{k+1}, B_k) > 1/2$. Furthermore,

$$B_k + B_k = A_{s_k} + A_{s_k} \subset A_{K(s_k)} \subset A_\ell = B_{k+1}.$$

Proceeding inductively, we obtain $0 = s_0 < s_1 < \cdots$ with the desired properties. $\square$

Proof of Theorem 3.4. The implication (ii) $\Rightarrow$ (i) is trivial. Let us prove that (i) $\Rightarrow$ (ii). By Edwards [25, pp. 113–114], there exists a convex sequence (δ_n), convergent to 0, such that $\delta_n \geq \varepsilon_n$ for every n. By Brudnyi's Theorem, there exists $x \in X$ such that $E(x, A_{s_i}) \geq \delta_i$ for $i = 0, 1, 2, \ldots$. But $A_i \subseteq A_{s_i}$, hence $E(x, A_i) \geq E(x, A_{s_i}) \geq \varepsilon_i$ for every i. $\square$

Corollary 3.7. *For any approximation scheme $(X, \{A_n\})$ the following are equivalent claims:*

(a) $(X, \{A_n\})$ *satisfies Shapiro's Theorem.*
(b) $(X, \{A_n\})$ *satisfies the weak Brudnyi's condition with constant c for every $c \in (0, 1]$.*
(c) $(X, \{A_n\})$ *satisfies the weak Brudnyi's condition with constant c for a certain $c \in (0, 1]$.*

Moreover, if X is a Banach space, then (a)–(c) *are equivalent to:*

(d) *The distance functional $E(\cdot, A_n)$ converges arbitrarily slowly to* 0.

Proof. (a) $\Rightarrow$ (b) follows from Lemma 3.5. (b) $\Rightarrow$ (c) is trivial. To prove (c) $\Rightarrow$ (a), assume $c \in (0, 1)$ is such that $\sup_{n\in\mathbb{N}} E(S(X), A_n) > c > 0$. Then for every $n \in \mathbb{N}$ there exists $x_n \in X$ with $\|x_n\| = 1$ and $E(x_n, A_{K(n)}) > c$, so $E(x_n, A_n) \leq \|x_n\| = 1 \leq cE(x_n, A_{K(n)})$. This, in conjunction with Theorem 2.2, implies that $(X, \{A_n\})$ satisfies Shapiro's Theorem.

Finally, the claim that (a)$\Leftrightarrow$ (d) holds for Banach spaces is just a reformulation of Theorem 3.4. $\square$

As a consequence, we show that the approximation schemes satisfying Shapiro's Theorem are stable under perturbations.

Proposition 3.8. *Suppose, for a quasi-Banach space $(X, \|\cdot\|)$, there exists $p \in (0, 1]$ for which any $x_1, x_2 \in X$ satisfy $\|x_1 + x_2\|^p \leq \|x_1\|^p + \|x_2\|^p$. Suppose the approximation schemes (A_n) and (B_n) in X are such that (A_n) satisfies Shapiro's Theorem, and* $\liminf_n E(S(X) \cap B_n, A_n) <$ 1. *Then $(X, \{B_n\})$ also satisfies Shapiro's Theorem.*

Proof. Pick $C \in (\liminf_n E(S(X) \cap B_n, A_n), 1)$. Then for any $N \in \mathbb{N}$ there exists $n \geq N$ such that $E(S(X) \cap B_n, A_n) < C$. Find $0 < c < 1$ such that $c^p + C^p(1+c^p) < 1$. By Corollary 3.7(c), it suffices to show that, for such n, $E(S(X), B_n) \geq c$, since the sequence $(E(S(X), B_k))_{k=0}^\infty$ is non-increasing.

Suppose, for the sake of contradiction, for every $x \in S(X)$ there exists $b \in B_n$ with $\|x - b\| < c$. As $b = x - (x - b)$, $\|b\| \le (\|x\|^p + \|x - b\|^p)^{1/p} < (1 + c^p)^{1/p}$ Then there exists $a \in A_n$ such that $\|b - a\| \le C(1 + c^p)^{1/p}$, hence

$$\|x - a\|^p \le \|b - a\|^p + \|x - b\|^p \le c^p + C^p(1 + c^p),$$

which contradicts Corollary 3.7(b). □

Another useful consequence of Corollary 3.7 is:

Corollary 3.9. *Let X be a quasi-Banach space and let us assume that for each $r \in \mathbb{N}$, the family $(A_{n,r})_{n=0}^{\infty}$ defines an approximation scheme in X that satisfies Shapiro's Theorem and $n_1 \le n_2, r_1 \le r_2$ imply $A_{n_1,r_1} \subseteq A_{n_2,r_2}$. Then for every pair of increasing sequences $\{n_i\} \to \infty$, $\{r_i\} \to \infty$, the approximation scheme (A_{n_i,r_i}) satisfies Shapiro's Theorem.*

Proof. Let us denote $B_i = A_{n_i,r_i}$, $i = 0, 1, 2, \ldots$. Obviously, (B_i) is an approximation scheme in X. By hypothesis and by Corollary 3.7, for each $n, r \in \mathbb{N}$ we have that $E(S(X), A_{n,r}) = 1$. Hence, for each $i \in \mathbb{N}$, we also have $E(S(X), B_i) = 1$, and the result follows as a direct application of Corollary 3.7. □

Finally, we generalize a result of Tjuriemskih, mentioned in Section 1.

Corollary 3.10. *Suppose X is a quasi-Banach space, and (A_i) is a strictly increasing sequence of closed subspaces of X, so that $\overline{\cup_i A_i} = X$. Then the approximation scheme (A_i) satisfies Shapiro's Theorem. If X is a Banach space, then, in addition, the distance functional $E(\cdot, A_n)$ converges arbitrarily slowly to* 0.

This result follows immediately from Corollary 3.7, and the following lemma.

Lemma 3.11. *Suppose Y is a subspace of a quasi-Banach space X, with $\overline{Y} \subsetneq X$. Then for every $\varepsilon > 0$ there exists $w \in X$ such that $\|w\| \le 1$, and* $\mathrm{dist}(w, Y) \ge 1 - \varepsilon$.

Proof. Take $x \in X \setminus \overline{Y}$. Then $d = E(x, Y) > 0$, and there exists $y_0 \in Y$ such that $d \le \|x - y_0\| \le \frac{1}{1-\epsilon} d$. Set $z = x - y_0$ and $w = z/\|z\| \in S(X)$. Then

$$\|w - y\| = \frac{1}{\|x - y_0\|}\|x - (y_0 + y\|z\|)\| \ge \frac{1}{\|x - y_0\|}E(x, Y) \ge (1 - \epsilon)$$

for any $y \in Y$. □

4. Approximation schemes that do not satisfy Shapiro's Theorem

Section 6 gives many examples of approximation schemes satisfying Shapiro's Theorem. In this section, we present some examples of schemes failing this condition, and explore their properties.

For an approximation scheme $(X, \{A_n\})$, define its *density sequence* $\mathbf{d}_n = \mathbf{d}_n(X, \{A_n\})$ by setting, for $n \ge 0$, $\mathbf{d}_n = E(S(X), A_n)$. Clearly, $\mathbf{d}_0 \ge \mathbf{d}_1 \ge \cdots \ge 0$. Moreover, if $A_0 = \{0\}$, then $\mathbf{d}_0 = 1$. Corollary 3.7 immediately implies:

Corollary 4.1. *An approximation scheme (A_n) in a quasi-Banach space X satisfies Shapiro's Theorem if and only if $\mathbf{d}_n = 1$ for any $n \in \mathbb{N}$.*

To analyze the behavior of a density sequence, we establish:

Proposition 4.2. *Suppose (A_n) is an approximation scheme in a quasi-Banach space X, and the function $L : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is such that $A_m + A_n \subset A_{L(m,n)}$ for any $m, n \in \mathbb{N}$ (we can take $L(n, m) = K(\max\{n, m\})$). Then $\mathbf{d}_{L(m,n)} \leq \mathbf{d}_m \mathbf{d}_n$ for any $m, n \in \mathbb{N}$.*

Proof. Consider $x \in X$. Fix $\delta > 0, m$, and n. Write $x = a + y$, with $a \in A_n$, and $\|y\| \leq (1+\delta)\mathbf{d}_n \|x\|$. Furthermore, write $y = b + z$, with $b \in A_m$ and

$$\|z\| \leq (1+\delta)\mathbf{d}_m \|y\| \leq (1+\delta)^2 \mathbf{d}_m \mathbf{d}_n \|x\|.$$

Then $x = (a + b) + z$, with $a + b \in A_{L(m,n)}$. As $\delta > 0$ is arbitrary, we are done. □

As a particular case of Proposition 4.2, consider an approximation scheme arising from a dictionary. We say that a set $\mathcal{D}$ is a *dictionary* in a quasi-Banach space X if $\overline{\mathbf{span}[\mathcal{D}]} = X$. Define the approximation scheme $(X, \Sigma_n(\mathcal{D}))$ by setting

$$\Sigma_0(\mathcal{D}) = \{0\}; \qquad \Sigma_n(\mathcal{D}) = \bigcup_{F \subset \mathcal{D}, |F| \leq n} \mathbf{span}[F] \quad \text{for } n \geq 1. \tag{4.1}$$

Then $\Sigma_n(\mathcal{D}) + \Sigma_m(\mathcal{D}) = \Sigma_{n+m}(\mathcal{D})$ for every $n, m \geq 0$ (hence we can take $L(m, n) = m + n$). We thus have:

Corollary 4.3. *Suppose an approximation scheme $(\Sigma_n(\mathcal{D}))$ is constructed as described in the previous paragraph. Then $\mathbf{d}_{m+n} \leq \mathbf{d}_m \mathbf{d}_n$ for any m and n. In particular, if $\mathbf{d}_m < 1$ for some m, then the sequence $(\mathbf{d}_n)$ decays exponentially or faster.*

In Section 6, we shall see many dictionaries (some quite redundant) for which $\mathbf{d}_n = 1$ for any n. These dictionaries cannot be "too redundant". Indeed, if a dictionary $\mathcal{D}$ is a c-net of the unit sphere $S(X)$ for some $c < 1$, then $\mathbf{d}_1 \leq c$, hence $\mathbf{d}_n \leq c^n$ for every n.

Below we consider an "extreme" case of $\mathbf{d}_n$ becoming 0 for n large enough.

Proposition 4.4. *Let $(X, \{A_n\})$ be an approximation scheme. The following are equivalent:*

(a) $\bigcup \overline{A_n} = X$ *(equivalently, for all $x \in X$ there exists $n = n(x) \in \mathbb{N}$ such that $E(x, A_n) = 0$).*
(b) $\overline{A_n} = X$ *for some $n \in \mathbb{N}$ (equivalently, $E(x, A_n) = 0$ for all $x \in X$).*

Consequently, $\bigcup \overline{A_n} \neq X$ if and only if $\mathbf{d}_n > 0$ for all n.

Proof. The implication (b) ⇒ (a) is obvious. To prove the converse, suppose $X = \cup_n \overline{A_n}$. By Baire Category Theorem, for some n, there exist $x \in X$ and $c > 0$ such that $B(x, c)$ (the ball with the center at x, and radius c) lies inside of $\overline{A_n}$. By symmetry, $B(-x, c) \subset \overline{A_n}$. Then

$$B(0, c) \subset B(x, c) + B(-x, c) \subset \overline{A_n} + \overline{A_n} \subset \overline{A_{K(n)}}.$$

But $\lambda \overline{A_{K(n)}} = \overline{A_{K(n)}}$ for any scalar λ and m, hence $\overline{A_{K(n)}} = X$.

To prove the last claim of the proposition, note that $X \neq \overline{A_n}$ if and only if $\mathbf{d}_n > 0$. □

Corollary 4.5. *Suppose $\mathcal{D}$ is a Hamel basis in a Banach space X. Then there exists $n \in \mathbb{N}$ for which $\Sigma_n(\mathcal{D})$ is dense in X.*

Note that there are no uniform bounds for the values of n with the property outlined in Proposition 4.4(b) and Corollary 4.5. Indeed, by [7], any Banach space has a dense Hamel basis $\mathcal{D}$. In particular, $\Sigma_1(\mathcal{D})$ is dense in X. On the other hand, consider a space $X = \ell_\infty^N \oplus_p Y$. If $\mathcal{H}$ is a Hamel basis in Y, then $\mathcal{D} = \{e_i \oplus h : 1 \leq i \leq N, h \in \mathcal{H}\}$ ((e_i) is the canonical basis in ℓ_∞^N) is a Hamel basis in X. Then, for any $n < N$, there exists a norm one $x \in X$ such that $E(x, \Sigma_n(\mathcal{D})) = 1$ (indeed, $(e_1 + \cdots + e_N) \oplus 0$ has this property).

Another corollary deals with Hamel bases indexed by positive reals.

Corollary 4.6. *Suppose $\mathcal{D} = \{e_i\}_{i \in [0,\infty)}$ is a Hamel basis of a separable Banach space X, and $A_n = \mathbf{span}[\{e_i\}_{i \le n}]$. Then there exists $n_0 \in \mathbb{N}$ such that A_{n_0} is dense in X. In particular, A_{n_0} is an infinite codimensional dense subspace of X.*

Next, we present an example where the "slowest possible" rate of approximation $E(x, A_n)$ is precisely controlled.

Theorem 4.7. *Suppose X is $L_\infty(0, 1)$, ℓ_∞, or $C(\Delta)$ (where Δ is the ternary Cantor set). Suppose, furthermore, that $1 \ge \varepsilon_1 \ge \varepsilon_2 \ge \cdots > 0$, and $\lim_n \varepsilon_n = 0$. Then there exists an approximation scheme (A_n) in X such that the $\mathbf{d}_n \le \varepsilon_n$ for any n, and there exists $x \in S(X)$ with the property that $E(x, A_n) \ge \varepsilon_n/(1 + \varepsilon_n) \ge \frac{\varepsilon_n}{2}$ for any n.*

The above theorem is stated for real Banach spaces. Similar results (with different constants) can also be obtained in the complex case.

Proof. We start by presenting the construction of (A_n) in the case of $X = L_\infty(0, 1)$. Find a sequence of positive integers $m(1) \le m(2) \le \cdots$, such that, for any n, $1/m(n) \le \varepsilon_n \le 1/(m(n)-1)$. Define A_n as the set of (equivalence classes of) functions in $L_\infty(0, 1)$ assuming no more than $m(n)$ different values. In other words, A_n consists of all functions $a = \sum_{i=1}^{m(n)} \alpha_i \chi_{E_i}$, where $(E_i)_{i=1}^{m(n)}$ is a partition of $(0, 1)$ into measurable sets.

(1) For a norm 1 function $x \in L_\infty(0, 1)$ and $n \in \mathbb{N}$, we shall find $a \in A_n$ such that $\|x - a\| \le 1/m(n)$. To this end, let $s_j = (2j - 1)/m(n) - 1$ ($1 \le j \le m(n)$). Let $I_1 = [-1, -1 + 2/m(n)]$, and $I_j = (-1 + 2(j - 1)/m(n), -1 + 2j/m(n)]$ for $2 \le j \le m(n)$. Note that s_j is the midpoint of I_j. For $t \in (0, 1)$, define $a(t) = s_j$ if $x(t) \in I_j$. Then a is defined almost everywhere, $a \in A_n$, and $\|x - a\| \le 1/m(n) \le \varepsilon_n$.

(2) We claim that the function $x(t) = 2t - 1$ is such that $\|x - a\| \ge 1/m(n) \ge \varepsilon_n/(1 + \varepsilon_n)$. Indeed, suppose a takes values $a_1 < a_2 < \cdots a_k$, with $k \le m(n)$, and $\|x - a\| = c < 1/m(n)$. Then $x(t) \in \cup_{j=1}^k [a_j - c, a_j + c]$ almost everywhere, which, in turn, implies $a_1 \le -1 + c$, $a_j + 2c \ge a_{j+1}$ for $1 \le j \le k - 1$, and $a_k \ge 1 - c$. This, however, is impossible.

The case of $X = \ell_\infty$ is handled the same way, with minor modifications. For $X = C(\Delta)$, consider *elementary intervals* $T_{s,k} = [\sum_{j=1}^k s_j 3^{-j}, \sum_{j=1}^k s_j 3^{-j} + 3^{-k}]$ ($k \in \mathbb{N}, s = (s_1, \ldots, s_k) \in \{0, 2\}^k$). Define A_n to be the set of functions a on Δ such that (i) a attains no more than $m(n)$ different values, and (ii) there exists $k \in \mathbb{N}$ such that the restriction of a to $T_{s,k} \cap \delta$ is constant for any $s \in \{0, 2\}^k$. To show $E(x, a) \le \|x\|/m(n)$ for any $x \in C(\Delta)$, take into account the uniform continuity of x. A version of the "Cantor ladder" gives an example of x with $E(x, A_n) \ge 1/m(n) \ge \varepsilon_n/(1 + \varepsilon_n)$ for any n. □

The theorem above implies that many Banach spaces contain an approximation scheme with controlled rate of approximation.

Corollary 4.8. *Suppose X is an infinite dimensional Banach space, and either* (1) *X is injective, or* (2) *X is separable, and contains an isomorphic copy of $C(\Delta)$. Then there exists a constant $c > 0$ such that, for every sequence $1 \ge \varepsilon_1 \ge \varepsilon_2 \ge \cdots > 0$, satisfying $\lim_n \varepsilon_n = 0$, there exists an approximation scheme (A_n) with the property that $\mathbf{d}_n \le \varepsilon_n$ for any n, and there exists $x \in S(X)$ with the property that $E(x, A_n) \ge c\varepsilon_n$ for all n.*

Proof. (1) Suppose X is injective. Then (see [39, Theorem 2.f.3]), there exists a subspace Y of X, a projection P from X onto Y, and an isomorphism $U : Y \to \ell_\infty$ with contractive inverse. By Theorem 4.7, there exists an approximation scheme (B_n) in ℓ_∞ such that $E(z, B_n) \le \delta_n \|z\|$ for any n and $z \in \ell_\infty$, where $\delta_n = \varepsilon_n/(\|U\| \|P\|)$. Furthermore, there exists $z_0 \in \ell_\infty$ with $\|z_0\| = 1$, and $E(z_0, B_n) \ge \delta_n/2$ for any n. We claim that the family $A_n = \ker P + \mathbf{span}[U^{-1}(B_n)]$ has the desired properties.

Note first that, for any $x \in X$,

$$E(x, A_n) \le E(Px, U^{-1}(B_n)) \le E(UPx, B_n) \le \delta_n \|UPx\| \le \delta_n \|U\| \|P\| \|x\| = \varepsilon_n \|x\|.$$

On the other hand, find $z_0 \in S(\ell_\infty)$ such that $E(z_0, B_n) \ge \delta_n/2$ for any n. Then $x_0 = U^{-1} z_0$ has norm not exceeding 1. To estimate $E(x_0, A_n)$, consider $b \in A_n$. Then

$$\|x_0 - b\| \ge \frac{1}{\|P\|} \|P(x_0 - b)\| = \frac{1}{\|P\|} \|x_0 - Pb\|.$$

Furthermore,

$$\|x_0 - Pb\| \ge \frac{1}{\|U\|} \|Ux_0 - UPb\| = \frac{1}{\|U\|} \|z_0 - UPb\| \ge \frac{1}{\|U\|} E(z_0, B_n) \ge \frac{\varepsilon_n}{2\|U\|}.$$

This leads to the desired estimates on $E(x_0, A_n)$.

The proof of (2) is very similar, except that now, we rely on the fact that any separable Banach space containing a copy of $C(\Delta)$, must also contain a complemented copy of the latter space (see e.g. [55]). □

Remark 4.9. A weaker version of Theorem 4.7 holds in the space c_0. More precisely, suppose $1 \ge \varepsilon_1 \ge \varepsilon_2 \ge \cdots > 0$, and $\lim_n \varepsilon_n = 0$. Then there exists an approximation scheme (A_n) in c_0, with the following properties:

(1) $\varepsilon_n \ge \mathbf{d}_n \ge \varepsilon_n/3$.
(2) For any non-increasing sequence $\{\delta_n\} \in c_0$ there exists $x \in c_0$ such that $E(x, A_n) \ge \delta_n \varepsilon_n$ for every n.

As the construction is similar to the one presented above, we do not describe it here.

5. Connection with central theorems of approximation theory

In this section we examine the connections between the so called central theorems of approximation theory – that is, the classical Jackson's (direct) and Bernstein's (inverse) results for the speed of approximation by a given approximation scheme – and Shapiro's Theorem. Throughout this section, we consider the setting where $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are quasi-Banach spaces, and j_Y is a continuous embedding (that is, there exists $C > 0$ so that $\|j_Y y\|_X \le C\|y\|_Y$ for any $y \in Y$). All embeddings are assumed to be injective. For convenience, we often omit the embedding operator j_Y, and identify $y \in Y$ with its image in X. This way, we regard Y is as a linear subspace of X, equipped with its own norm $\|\cdot\|_Y$. Note that the norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ need not be equivalent on Y, as examples below show.

Definition 5.1. Suppose X and Y are quasi-Banach spaces, and Y is continuously embedded into X. We say that an approximation scheme $(X, \{A_n\})$ satisfies (*generalized*) *Jackson's Inequality* with respect to Y if there exists a sequence (c_n) such that $\lim_{n\to\infty} c_n = +\infty$ and

$$E(y, A_n)_X = \inf_{a \in A_n} \|y - a\|_X \le \frac{1}{c_n} \|y\|_Y \quad \text{for all } y \in Y. \tag{5.1}$$

The approximation scheme $(X, \{A_n\})$ is said to satisfy (*generalized*) *Bernstein's Inequality* with respect to Y if $\bigcup_{n=0}^{\infty} A_n \subseteq Y$, and there exists a sequence (b_n) such that $\lim_{n\to\infty} b_n = +\infty$ and

$$\|x_n\|_Y \leq b_n \|x_n\|_X \quad \text{for all } x_n \in A_n. \tag{5.2}$$

To provide an illustration, consider the classical case of approximation by trigonometric polynomials on the unit circle $\mathbb{T}$. Suppose $X = C(\mathbb{T})$ is equipped with its canonical norm $\|\cdot\|_X = \|\cdot\|_\infty$. Let A_n be the subset of X, consisting of trigonometric polynomials of degree not exceeding n. Let $Y = C^r(\mathbb{T})$ be the space of r times continuously differentiable functions, equipped with the norm $\|f\|_Y = \|f\|_\infty + \|f^{(r)}\|_\infty$. Then the well-known Bernstein's Inequality states that $\|f\|_Y \leq (n^r + 1)\|f\|_X$ [20, Chapter 4]. By Jackson's Theorem, $E(f, A_n)_X \leq 1/(C_r n^r)\|f\|_Y$ for any $f \in Y$ [20, Chapter 7]. In this case, the assertions of Definition 5.1 hold, with $b_n = n^r + 1$, and $c_n = C_r n^r$.

Remark 5.2. Suppose X, Y, and Z are quasi-Banach spaces, equipped with the norms $\|\cdot\|_X$, $\|\cdot\|_Y$, and $\|\cdot\|_Z$, respectively. Suppose, furthermore, that $J_Y : Y \to X$ and $j_Z : Z \to X$ are continuous embeddings, and $J_Y(Y) \subset j_Z(Z)$. Then the embedding of Y into Z, given by $j = j_Z^{-1} j_Y$, is continuous. Indeed, by Closed Graph Theorem [35, Corollary 1.7], it suffices to show that, if $y_n \to 0$ in $\|\cdot\|_Y$, and $jy_n \to z$ in $\|\cdot\|_Z$, then $z = 0$. But, if $z \neq 0$, then $j_Z z = j_Z(\lim_n j_Z^{-1} j_Y y_n) = \lim_n j_Y y_n \neq 0$, which contradicts the continuity of the embedding j_Y.

This reasoning shows that Jackson's Inequality passes to subspaces. More precisely, suppose (A_n) is an approximation scheme in a quasi-Banach space X, and quasi-Banach spaces Y and Z are continuously embedded into X in such a way that $Y \subset Z$. If (A_n) satisfies Jackson's Inequality with respect to Z, then it also satisfies Jackson's Inequality with respect to Y. In Corollary 5.5, we show that Bernstein's Inequality passes to subspaces as well.

The following proposition demonstrates that Jackson's Inequality is satisfied for a sufficiently large space $Y \subset X$ if and only if Shapiro's Theorem fails.

Proposition 5.3. *For an approximation scheme $(X, \{A_n\})$, the following are equivalent:*

(i) $(X, \{A_n\})$ *does not satisfy Shapiro's Theorem.*
(ii) (A_n) *satisfies Jackson's Inequality with respect to some finite codimensional subspace $Y \subset X$, so that the quasi-norm of Y is equivalent to that of X.*
(iii) (A_n) *satisfies Jackson's Inequality with respect to every subspace $Y \subset X$.*

In particular, if $(X, \{A_n\})$ satisfies Shapiro's Theorem and Y is a subspace of X, equipped with a norm equivalent to that of X, such that (A_n) satisfies Jackson's Inequality with respect to Y, then Y must be of infinite codimension.

Note that, if Y is a closed subspace of X, then their quasi-norms are equivalent, by Open Mapping Theorem (see e.g. [35, Corollary 1.5]). Therefore, if (A_n) satisfies Jackson's Inequality with respect to a closed finite codimensional subspace $Y \subset X$, then assertion (ii) of the above proposition holds.

By [24, Theorem 5.6(c)], if X and Y are Banach spaces, and $T \in B(Y, X)$ is such that $T(Y)$ is of finite codimension in X, then $T(Y)$ is closed. Thus, if X and Y appearing in Proposition 5.3 are Banach spaces, and Y is embedded in X as a subspace of finite codimension, then the norms of X and Y are equivalent on Y. We do not know whether the finite codimensionality of Y implies Y being closed in a more general quasi-Banach setting.

Proof. (iii) $\Rightarrow$ (ii) is trivial.

(i) $\Rightarrow$ (iii): By Corollary 3.7, there exists a sequence $\{\varepsilon_n\} \searrow 0$ such that $E(x, A_n) \leq \varepsilon_n \|x\|_X$ for any $x \in X$. The space Y is continuously embedded into X, hence there exists a constant C such that $\|y\|_X \leq C\|y\|_Y$ for any $y \in Y$. Therefore, $E(y, A_n) \leq C\varepsilon_n\|y\|_Y$ for any $y \in Y$, which is (5.1) with $c_n = (C\varepsilon_n)^{-1}$.

(ii) $\Rightarrow$ (i): By Corollary 3.7, it suffices to find $m \in \mathbb{N}$ for which $E(S(X), A_m) < 1$. For the sake of brevity, we use the notation $\|\cdot\|$ for $\|\cdot\|_X$. By renorming Y if necessary, we may assume that $\|\cdot\|_Y = \|\cdot\|$ on Y. Let $N = \dim X/\overline{Y}$. If $N = 0$ (that is, $X = \overline{Y}$), there is nothing to prove. Otherwise, consider the quotient map $q : X \to E = X/\overline{Y}$. Find $x_1, \ldots, x_N$ in X, such that the vectors $e_i = qx_i$ form a normalized basis in E, and $\|x_i\| < 2$ for every i. Then there exists a constant $C_1 \geq 1$ such that $C_1^{-1} \max_{1\leq i\leq N} |\alpha_i| \leq \|\sum_{1\leq i\leq N} \alpha_i q x_i\|$ for any N-tuple of scalars (α_i).

Recall the existence of a constant $C_X \geq 1$ such that $\|x + y\| \leq C_X(\|x\| + \|y\|)$ for any $x, y \in X$. By induction,

$$\left\| \sum_{j=1}^{m} z_j \right\| \leq C_X^{m-1} \sum_{j=1}^{m} \|z_j\| \tag{5.3}$$

for any $z_1, \ldots, z_m \in X$. We claim that any $x \in S(X)$ has a representation

$$x = y + \sum_{i=1}^{N} \alpha_i x_i, \quad \text{with} \ \max_{1\leq i\leq N} |\alpha_i| \leq C_1, \ y \in Y, \quad \text{and} \quad \|y\| \leq C_2 = 2C_1 C_X^N. \tag{5.4}$$

Indeed, $\|qx\| \leq 1$, hence one can write $qx = \sum_{i=1}^{N} \alpha_i e_i$, with (α_i) as above. Then $y = x - \sum_{i=1}^{N} \alpha_i x_i \in Y$, and (5.3) yields the desired estimate on the norm of y.

Pick $c \in (0, 1)$, and show the existence of $m \in \mathbb{N}$ for which $E(S(X), A_m) < c$. Start by using (5.1) to find $n \in \mathbb{N}$ such that $E(y, A_n) \leq c\|y\|/(2NC_2C_X^N)$ holds for every $y \in Y$. Then find $k \geq n$ such that, for every $i \in \{1, \ldots, N\}$, there exists $a_i \in A_k$ satisfying $\|x_i - a_i\| < c/(2NC_1C_X^N)$. We claim that $E(S(X), A_m) \leq c$, where $m = K(K(\ldots(k)\ldots))$ ($N + 1$ times). Indeed, any $x \in S(X)$ can be represented as in (5.4). Find $a_0 \in A_k$ satisfying $\|y - a_0\| < c\|y\|/(2NC_2C_X^N) \leq c/(2NC_X^N)$. Then $a = a_0 + \sum_{i=1}^{N} \alpha_i a_i \in A_m$, and, by (5.3),

$$\begin{aligned}\|x - a\| &= \left\| \left(y + \sum_{i=1}^{N} \alpha_i x_i \right) - \left(a_0 + \sum_{i=1}^{N} \alpha_i a_i \right) \right\| \leq C_X^N \left(\|y - a_0\| + \sum_{i=1}^{N} |\alpha_i| \, \|a_i\| \right) \\ &< C_X^N \left(\frac{c}{2NC_X^N} + NC_1 \frac{c}{2NC_1C_X^N} \right) < c.\end{aligned}$$

Thus, $E(S(X), A_m) \leq c < 1$. An application of Corollary 3.7 completes the proof. $\square$

Now we concentrate on Bernstein's Inequality.

Theorem 5.4. *Let $(X, \{A_n\})$ be an approximation scheme that satisfies Bernstein's Inequality with respect to a certain proper quasi-Banach subspace Y of X. Then $(X, \{A_n\})$ satisfies Shapiro's Theorem.*

Proof. We show that if the approximation scheme does not satisfy Shapiro's Theorem, and it satisfies Bernstein's Inequality with respect to a certain space Y, then the norms of Y and X are equivalent. So, let us assume that $(X, \{A_n\})$ satisfies (5.2) for a certain sequence of positive real numbers (b_n), and a quasi-Banach space Y continuously included in X. By renorming X and Y if necessary (see Section 2), we may assume the existence of $p_X, p_Y \in (0, 1]$ such that: (i) for any $x_1, x_2 \in X$, $\|x_1 + x_2\|_X^{p_X} \leq \|x_1\|_X^{p_X} + \|x_2\|_X^{p_X}$, (ii) for any $y_1, y_2 \in Y$, $\|y_1 + y_2\|_Y^{p_Y} \leq \|y_1\|_Y^{p_Y} + \|y_2\|_Y^{p_Y}$. Letting $p = \min\{p_X, p_Y\}$, we see that, for any $x_1, x_2 \in X$ and $y_1, y_2 \in Y$, $\|x_1 + x_2\|_X^p \leq \|x_1\|_X^p + \|x_2\|_X^p$, and $\|y_1 + y_2\|_Y^p \leq \|y_1\|_Y^p + \|y_2\|_Y^p$. For the sake of brevity, we shall denote $\|\cdot\|_X$ simply by $\|\cdot\|$.

If $(X, \{A_n\})$ does not satisfy Shapiro's Theorem, Corollary 3.7 guarantees the existence of $n_0 \in \mathbb{N}$ for which $E(S(X), A_{n_0}) < (1/2)^{1/p}$. Therefore, for any $x \in X$, there exist $a \in A_{n_0}$ and $x' \in X$ such that $x = a + x'$, and $\|x'\| < 2^{-1/p}\|x\|$.

Now pick $x \in B(X) \setminus Y$. By the above, we can find $a_0 \in A_{n_0}$ and $x_0 \in X$ such that $x = a_0 + x_0$, with $\|x_0\| < 2^{-1/p}$. Furthermore, we can write $x_0 = a_1 + x_1$, with $a_1 \in A_{n_0}$, and $\|x_1\| < 2^{-2/p}$. Proceeding further in the same manner, we write, for each m, $x = a_0 + a_1 + \cdots + a_m + x_m$, with $a_0, a_1, \ldots \in A_{n_0}$, and $\|x_m\| < 2^{-m/p}$. Note that $a_m = x_{m-1} - x_m$, hence $\|a_m\| \leq (\|x_{m-1}\|^p + \|x_m\|^p)^{1/p} < 3^{1/p}2^{-m/p}$.

Let $z_m = x - x_m$. As $\lim_m \|x_m\| = 0$, the sequence (z_m) converges to x in the space X. We shall show that (z_m) is a Cauchy sequence in Y. Indeed, for $n > m$, $z_n - z_m = \sum_{k=m+1}^n a_k$. Furthermore, $\|a_k\|_Y \leq b_{n_0}\|a_k\|$, for each k. Therefore,

$$\|z_n - z_m\|_Y^p \leq \sum_{k=m+1}^n \|a_k\|_Y^p \leq b_{n_0}^p \sum_{k=m+1}^n \|a_k\|^p < 3b_{n_0}^p \sum_{k=m+1}^n 2^{-(k+1)} < 3b_{n_0}^p 2^{-m}.$$

As Y is a subset of X, the sequence (z_m) must converge to x in the space Y. This leads to a contradiction, since x was selected in such a way that $x \notin Y$. □

Combining Theorem 5.4 with Remark 5.2, we conclude that the property of satisfying Shapiro's Theorem is, under certain conditions, inherited by subspaces.

Corollary 5.5. *Suppose the approximation scheme $(X, \{A_n\})$ satisfies Bernstein's Inequality for a proper subspace Y of X. Suppose, furthermore, that Z is another quasi-normed subspace of X, properly containing Y, and such that $\bigcup_{n=0}^\infty A_n$ is dense in Z (in the topology determined by the norm of Z). Then $(Z, \{A_n\})$ satisfies Shapiro's Theorem.*

Section 6 contains some examples where the fact that a given approximation scheme satisfies Shapiro's Theorem is deduced from a Bernstein's Inequality.

Below we consider Bernstein's Inequality with respect to so called "smoothness spaces" (or "abstract approximation spaces"). If (A_n) is an approximation scheme in X (with $A_0 = \{0\}$), we define, for $0 < q \leq \infty$ and $0 < r < \infty$,

$$A_q^r = A_q^r(X, \{A_n\}) = \{x \in X : |x|_{A_q^r} = \|\{(n+1)^{r-1/q}E(x, A_n)\}\|_{\ell^q} < \infty\}. \tag{5.5}$$

If $A_n + A_n \subseteq A_{cn}$ for a constant $c > 1$, then A_q^r is a quasi-Banach space [4]. It was shown by DeVore and Popov (see [20, Th. 9.3, p. 236]) that A_q^r satisfies Bernstein's Inequality: $|x|_{A_q^r} \leq Cn^r\|x\|_X$ for all $x \in A_n$.

To apply Theorem 5.4 with $Y = A_q^r$, we need Y to be a proper subspace of X, which does not always hold. For instance, suppose $\mathcal{D}$ is a dictionary in a Banach space X, which is 1/2-dense in $S(X)$ (i.e., such that $\sup_{x \in S(X)} E(x, \mathcal{D}) \leq \frac{1}{2}$). Let $A_n = \Sigma_n(\mathcal{D})$. Clearly, $A_n + A_n \subset A_{2n}$. By Corollary 4.3 and the discussion following it, $E(x, A_n) \leq 2^{-n}\|x\|$ for any $x \in X$. Therefore, $A_q^r = X$ for any q, r (with equivalent norms).

On the other hand, there are many classical results in Approximation Theory devoted to the characterization of the approximation spaces A_q^r as smoothness spaces of functions (Besov, etc.), and these are always proper subspaces of the ground space X. In this setting, one can apply Theorem 5.4 to show that the corresponding approximation scheme satisfies Shapiro's Theorem. The same applies to the situation when X is a space of operators, and membership in A_q^r reflects the "degree of compactness" (see e.g. [51]).

Below, we show that the spaces A_q^r form a scale of subspaces of X if the approximation scheme (A_n) satisfies Shapiro's Theorem. We also present other results on the spaces A_q^r.

Corollary 5.6. *Let $(X, \{A_n\})$ be an approximation scheme such that $A_n + A_n \subseteq A_{cn}$ for a certain constant $c > 1$. Then the following are equivalent:*

(a) *$(X, \{A_n\})$ satisfies Bernstein's Inequality for some proper subspace Y of X.*
(b) *$(X, \{A_n\})$ satisfies Shapiro's Theorem.*
(c) *For every $r > 0$ there exists $x \in X$ such that $E(x, A_n) \neq \mathbf{O}(n^{-r})$.*
(d) *For a certain $r > 0$, there exists $x \in X$ such that $E(x, A_n) \neq \mathbf{O}(n^{-r})$.*
(e) *For any $q \in (0, \infty]$ and $r \in (0, \infty)$, A_q^r is a proper subspace of X.*
(f) *For some $q \in (0, \infty]$ and $r \in (0, \infty)$, A_q^r is a proper subspace of X.*

Moreover, if any of these conditions is satisfied, then for every $q, r > 0$, A_q^r is an infinite codimensional subspace of X.

Proof. The implication (a) $\Rightarrow$ (b) is a reformulation of Theorem 5.4. (b) $\Rightarrow$ (c) $\Rightarrow$ (d) and (c) $\Rightarrow$ (e) $\Rightarrow$ (f) are trivial.

(d) $\Rightarrow$ (a): If $E(x, A_n) \neq \mathbf{O}(n^{-r})$ for some $x \in X$ and $r > 0$, then $x \notin A_\infty^r$. Then A_∞^r is strictly contained in X, and $(X, \{A_n\})$ satisfies Bernstein's Inequality for $Y = A_\infty^r$. Theorem 5.4 yields (a).

(f) $\Rightarrow$ (c): consider $x \in X \setminus A_q^r$. By (5.5), $x \notin A_\infty^s$ for any $s > r$.

To prove the final claim, suppose, for the sake of contradiction, that (b) holds, while $Y = A_q^r$ is a finite codimensional subspace of X. By Almira and Oikhberg [5, Theorem 2.9], for every sequence $\{\varepsilon_n\} \searrow 0$ there exists $y \in Y$ so that $E(y, A_n)_X \neq \mathbf{O}(\varepsilon_n)$. On the other hand, for every $y \in Y$ and $n \in \mathbb{N}$ we have $E(y, A_n)_X \leq n^{-r}|y|_{A_q^r}$, hence $E(y, A_n)_X = \mathbf{O}(n^{-r})$. Taking $\varepsilon_n = n^{-r/2}$, we arrive at the desired contradiction. □

Proposition 5.7. *Suppose the approximation scheme $(X, \{A_n\})$ satisfies Shapiro's Theorem, and $A_n + A_n \subset A_{cn}$ for some c. Then $(B_q^r, \{A_n\})$ satisfies Shapiro's Theorem. Consequently, $A_q^{r+\varepsilon}(X, \{A_n\})$ is an infinite codimensional subspace of $A_u^r(X, \{A_n\})$ for all $\varepsilon > 0$, all $0 < q \leq \infty$ and all $0 < r, u < \infty$.*

Proof. A small modification of the proof of (a) $\Rightarrow$ (b) in Theorem 2.2 shows that there exists a constant $C > 1$ and a sequence $\{x_n\}_{n \in \mathbb{N}_0}$ (where $\mathbb{N}_0 \subseteq \mathbb{N}$ is an infinite sequence) such that $E(x_n, A_n) \leq CE(x_n, A_{K^2(n)})$ for all $n \in \mathbb{N}_0$. By the density of $\bigcup_n A_n$ in X, we can find another sequence $\{a_n\}_{n \in \mathbb{N}_0} \subset \bigcup_n A_n \subset B_q^r$ such that $E(a_n, A_n) \leq CE(a_n, A_{K^2(n)})$ for all $n \in \mathbb{N}_0$. Hence for every $n \in \mathbb{N}_0$ and $m \in \{n, n+1, \ldots, K^2(n)\}$ we have $E(a_n, A_m) \leq E(a_n, A_n) \leq CE(a_n, A_{K^2(n)})$. Furthermore, by Lemma 3.16 from [4], there exist $A, B > 0$ (depending only on X and the parameters q, r) such that, for every $n \in \mathbb{N}$,

$$A \left\| \left\{ (k+1)^{r-\frac{1}{q}} E(a_n, A_{\max\{k, K(n)\}}) \right\} \right\|_{\ell_q} \leq E(a_n, A_n)_{A_q^r}$$
$$\leq B \left\| \left\{ (k+1)^{r-\frac{1}{q}} E(a_n, A_{\max\{k, n\}}) \right\} \right\|_{\ell_q}.$$

Therefore,

$$E(a_n, A_n)_{B_q^r} \leq B \left\| \left\{ (k+1)^{r-\frac{1}{q}} E(a_n, A_{\max\{k,n\}}) \right\} \right\|_{\ell_q},$$

$$E(a_n, A_{K(n)})_{B_q^r} \geq A \left\| \left\{ (k+1)^{r-\frac{1}{q}} E(a_n, A_{\max\{k,K^2(n)\}}) \right\} \right\|_{\ell_q}.$$

It follows that

$$\begin{aligned} E(a_n, A_n)_{B_q^r} &\leq B \left\| \left\{ (k+1)^{r-\frac{1}{q}} E(a_n, A_{\max\{k,n\}}) \right\} \right\|_{\ell_q} \\ &\leq B \left(\sum_{k=0}^{K^2(n)} (k+1)^{rq-1} C^q E(a_n, A_{K^2(n)})^q \right. \\ &\quad \left. + \sum_{k=K^2(n)+1}^{\infty} (k+1)^{rq-1} E(a_n, A_k)^q \right)^{\frac{1}{q}} \\ &\leq CBA^{-1} E(a_n, A_{K(n)})_{B_q^r}. \end{aligned}$$

By Theorem 2.2(b) $\Rightarrow$ (a), $(B_q^r, \{A_n\})$ satisfies Shapiro's Theorem.

To prove the second part of our proposition, recall the reiteration theorem: if an approximation scheme satisfies $A_n + A_n \subseteq A_{cn}$ for a certain constant c, then

$$A_q^{r_2}(A_s^{r_1}(X, \{A_n\}), \{A_n\}) = A_q^{r_1+r_2}(X, \{A_n\})$$

(this is proved in [51] for the particular case of $A_n + A_m \subseteq A_{n+m}$, and in [4, Example 3.36] in full generality). Hence,

$$A_q^{r+\varepsilon}(X, \{A_n\}) = A_q^{\varepsilon}(A_u^r(X, \{A_n\}), \{A_n\}) = A_q^{\varepsilon}(A_u^r, \{A_n\}).$$

As $u < \infty$, the first part of our proposition shows that $(A_u^r, \{A_n\})$ satisfies Shapiro's Theorem. By Corollary 5.6, $A_q^{\varepsilon}(A_u^r, \{A_n\})$ is an infinite codimensional subspace of $A_u^r(X, \{A_n\})$. □

Finally, another consequence of Theorem 5.4 is the following

Corollary 5.8. *Suppose $(X, \{A_n\})$ is an approximation scheme such that, for every $n \in \mathbb{N}$, $A_n + A_n \subseteq A_{cn}$ ($c > 1$ is independent of n), and A_n is boundedly compact in X (that is, any bounded subset of A_n is relatively compact in X). Then $(X, \{A_n\})$ satisfies Shapiro's Theorem.*

Proof. If each A_n is boundedly compact in X, then, for every $r > 0$, the natural inclusion $A_\infty^r \hookrightarrow X$ is a compact operator (see [4, Theorem 3.32]). In particular, A_∞^r is a proper subspace of X. We apply Theorem 5.4 with $Y = A_\infty^r$ to complete the proof. □

6. Examples of schemes satisfying Shapiro's Theorem

In this section, we present a collection of examples of approximation schemes satisfying Shapiro's Theorem. The main tools involved are (i) Property (P), (ii) Bernstein's Inequality and (iii) the characterization of approximation schemes satisfying Shapiro's Theorem given in Corollary 3.7. Many examples involve the order of the best n-term approximation with respect to a dictionary.

6.1. Biorthogonal systems and their generalizations

Suppose X is a quasi-Banach space, I is an infinite index set, and $(X_i)_{i\in I}$ are non-trivial subspaces of X. We say that (X_i) form a *complete minimal bounded decomposition* of X (*CMBD*, for short) if $X = \overline{\mathbf{span}[X_i : i \in I]}$, and, for every $i \in I$, there exists $x \in X_i$ such that $E(x, \mathbf{span}[X_j : j \neq i]) > c\|x\|$ ($c > 0$ is independent of i).

A CMBD can be regarded as a generalization of a complete minimal system. Recall that a family $(x_i)_{i\in I}$ in a Banach space X is called *minimal* if, for any $i \in I$, x_i does not belong to the closure of $\mathbf{span}[x_j : j \in I\setminus\{i\}]$. A minimal system is called *complete* if $\mathbf{span}[x_i : i \in I]$ is dense in X. It is easy to see that a minimal system gives rise to a *biorthogonal system* (x_i, f_i), where $x_i \in X$, $f_i \in X^*$, and $\langle f_i, x_j\rangle = \delta_{ij}$ (Kronecker's delta). A biorthogonal system is *bounded* if $\sup_i \|x_i\| \|f_i\| < \infty$.

It is easy to see that, if (x_i, f_i) is a bounded complete biorthogonal system, then the family of spaces $X_i = \mathbf{span}[x_i]$ forms a CMDB. It is known that every separable Banach space has a complete bounded biorthogonal system $(x_i, f_i)_{i\in I}$ such that $\cap_{i\in I} \ker f_i = \{0\}$ [30, Theorem 1.27]. Certain non-separable spaces also possess complete bounded biorthogonal systems (see e.g. Sections 4.2 and 5.2 of [30]).

In addition to biorthogonal systems, CMBDs arise when one considers a dictionary consisting of two or more bases, possessing certain "mutual coherence". Several examples can be found in Section 4 of [29]. For instance, the union of Haar and Walsh bases works very nicely.

The following two theorems show that the approximation schemes arising from CMBDs or biorthogonal systems have Property (P). Furthermore, as the approximation schemes described there satisfy $A_n + A_n \subset A_{2n}$, both schemes satisfy Shapiro's Theorem.

Theorem 6.1. *Consider a quasi-Banach space X such that, for a certain fixed $p > 0$ and for any $x_1, \ldots, x_m \in X$,*

$$\|x_1 + \cdots + x_m\|^p \leq C^p(\|x_1\|^p + \cdots + \|x_m\|^p).$$

Suppose $(X_i)_{i\in I}$ is a complete minimal bounded decomposition of X, with $E(x, \mathbf{span}[X_j : j \neq i]) \geq c\|x\|$ for any $i \in I$, and $x \in X_i$. Suppose, furthermore, that E is a finite dimensional subspace of X, and an approximation scheme (A_n) is defined by setting, for $n \in \mathbb{N}$,

$$A_n = E + \cup_{F\subset I, |F|\leq n} \mathbf{span}[X_i : i \in F].$$

Then the approximation scheme (A_n) has Property (P), and consequently, satisfies Shapiro's Theorem.

Theorem 6.2. *For a complete minimal system $(x_i)_{i\in I}$ in a Banach space X, consider the approximation scheme $A_n = \{\sum_{i\in F} \alpha_i x_i : F \subset I, |F| \leq n\}$ $(n \geq 0)$. Then for every n there exists a norm one $y \in X$ such that (in the above notation) $E(y, A_{n-1}) > 1/(2n)$. Consequently, the approximation scheme (A_n) satisfies Shapiro's Theorem.*

Proof of Theorem 6.1. For $i \in I$, denote by $P_i : X \to X$ by setting $P_i x = x$ if $x \in X_i$, and $P_i x = 0$ if $x \in \overline{\mathbf{span}[X_j : j \in I \setminus \{i\}]}$. Then $C_0 = \sup_i \|P_i\|$ is finite. Let $m = \dim E + 1$. We shall find $y \in X$ such that $\|y\| \leq 1$, and $E(y, A_{n-1}) \geq (2C^2 C_0 m^{1/p} n^{1/p})^{-1}$.

To this end fix disjoint subsets $S_1, \ldots, S_n \in I$, of cardinality m each. For $1 \leq k \leq n$, set $Y_k = \mathbf{span}[X_i : i \in S_k]$. Then $Q_k = \sum_{i\in S_k} P_i$ is a projection onto Y_k, satisfying $Q_k \mathbf{span}[X_i : i \notin S_k] = 0$. By the assumptions about X, $\|Q_k\| \leq C(\sum_{i\in S_k} \|P_i\|^p)^{1/p} = CC_0 m^{1/p}$. Moreover, for each k, $\dim Q_k(E) < m$, while $\dim Y_k \geq m$. By Lemma 3.11, there exists a norm one $y_k \in Y_k$ such that $E(y_k, Q_k(E)) > 1/2$.

Now consider $y = (y_1 + \cdots + y_n)/(Cn^{1/p})$. Clearly, $\|y\| \le 1$. It remains to show that, for any $e \in E$, any $F \subset I$ of cardinality not exceeding $n-1$, any family of scalars $(\alpha_i)_{i\in F}$, and any family $x_i \in X_i$ (once again, $i \in F$), we have $\|y - (e + \sum_{i\in F} \alpha_i x_i)\| \ge (2C^2 C_0 m^{1/p} n^{1/p})^{-1}$. Find k such that $S_k \cap F = \emptyset$. Then

$$\|Q_k\| \left\| y - \left(e + \sum_{i\in F} \alpha_i x_i\right)\right\| \ge \left\| Q_k \left(y - \left(e + \sum_{i\in F} \alpha_i x_i\right)\right)\right\|$$
$$= \|Q_k y - Q_k e\| \ge \frac{1}{Cn^{1/p}} E(y_k, Q_k(E)) \ge \frac{1}{2Cn^{1/p}}.$$

We complete the proof by recalling that $\|Q_k\| \le CC_0 m^{1/p}$. $\square$

The following lemma (necessary for the proof of Theorem 6.2) may be known to experts, although we could not find its statement anywhere. Throughout, we use $S(X)$ and $B(X)$ to denote the unit sphere, respectively the closed unit ball, of X.

Lemma 6.3. *Suppose X is a Banach space, E is a* weak**-closed subspace of X^{**}, and Z is a subspace of X, such that* $\dim X/Z < \infty$, *and* $\dim X^{**}/E > \dim X/Z$ *(E can be of finite or infinite codimension). Then for every $c < 1$ there exists $x \in S(Z)$ such that* $\mathrm{dist}(x, E)_{X^{**}} \ge c$.

Proof. Suppose, for the sake contradiction, that the statement of the lemma is false. Then there exists $c \in (0, 1)$ with the property that, for every $x \in B(Z)$, there exists $e \in E$ such that $\|x-e\|_{X^{**}} \le c$. By the triangle inequality, $\|e\|_{X^{**}} \le 1+c$, hence $B(Z) \subset (1+c)B(E)+cB(X^{**})$. The set on the right is weak* closed (even weak* compact). Taking the weak* closure of the left hand side, we obtain

$$B(Z^{\perp\perp}) \subset (1+c)B(E) + cB(X^{**}). \tag{6.1}$$

Let $W = Z^{\perp\perp} \cap E$, and consider the quotient map $q : X^{**} \to X^{**}/W$. This map takes $Z^{\perp\perp}$ and E to $Z' = Z^{\perp\perp}/W$ and $E' = E/W$, respectively. Then $\dim E' < \infty$, and $\dim Z' > \dim E'$. By the well-known result by Krasnoselskii, Krein, and Milman (see e.g. [30, Lemma 1.19]), there exists $z' \in Z'$ such that $c < \mathrm{dist}(z', E')_{X^{**}/W} = \|z'\|_{X^{**}/W} < 1$. Find $z \in Z^{\perp\perp}$ such that $\|z\| \le 1$, and $q(z) = z'$. For every $e \in E$, we then have $\|z - e\|_{X^{**}} \ge \|q(z - e)\|_{X^{**}/W} \ge \mathrm{dist}(z', E')_{X^{**}/W} > c$, which contradicts (6.1). $\square$

Proof of Theorem 6.2. By Hahn–Banach Theorem, there exist linear functionals $f_i \in X^*$, satisfying $\langle x_i, f_j\rangle = \delta_{ij}$ for $i, j \in I$. Throughout the proof, we consider the functionals f_i as acting on X^{**}, and their kernels $\ker f_i$ as subsets of X^{**}. We also identify X with its canonical image in X^{**}.

We shall construct a sequence of finite disjoint sets $S_j \subset I$ such that for any j there exists a norm one $y_j \in \mathbf{span}[x_i : i \in S_j]$ with the property that $E(y_j, \mathbf{span}[x_i : i \notin S_j]) > 1/2$. Once this is done, let $y = (y_1 + \cdots + y_n)/n$. Clearly $\|y\| \le 1$. It remains to show that $\|y - \sum_{i\in F} \alpha_i x_i\| > 1/(2n)$ for any $F \subset I$ of cardinality less than n. As the sets S_j are disjoint, there exists j such that $S_j \cap F = \emptyset$. Then

$$\left\| y - \sum_{i\in F} \alpha_i x_i \right\| = \left\| \frac{y_j}{n} + \frac{1}{n}\sum_{k\ne j} y_k - \sum_{i\in F} \alpha_i x_i \right\| \ge \frac{1}{n} E(y_j, \mathbf{span}[x_i : i \notin S_j]) > \frac{1}{2n}.$$

We construct the sets S_j and vectors y_j inductively. Let $S_0 = \emptyset$. Suppose the sets S_j have already been obtained for all $j \le m-1$ ($m \in \mathbb{N}$). Let us construct S_m and y_m. Let $T = \cup_{j<m} S_j$. Introduce the spaces $E_0 = \cap_{i \in I} \ker f_i \hookrightarrow X^{**}$, and $E_T = \mathbf{span}[x_i : i \in T] \hookrightarrow X$. Define the projection Q_T from X^{**} onto E_T by setting $Q_T x = \sum_{i \in T} \langle f_i, x\rangle x_i$. Clearly, E_0 is weak* closed, and E_T is weak* closed due to being finite dimensional. As $E_0 \subset \ker Q_T$, we conclude that $E = E_0 + E_T$ is also weak* closed. Note that the set (f_i) is linearly independent, hence $\dim X^{**}/E_0 = \infty$.

Now set $Z = X \cap (\cap_{i \in T} \ker f_i)$. As $\dim X/Z < \infty$, Lemma 6.3 implies the existence of $z \in B(Z)$ satisfying $\mathrm{dist}(z, E)_{X^{**}} > 5/6$. As $\mathbf{span}[x_i : i \in I]$ is dense in X, there exists $z_1 \in S(\mathbf{span}[x_i : i \in I])$ such that $\|z - z_1\| < 1/(12\|Q_T\|)$, and $\mathrm{dist}(z_1, E)_{X^{**}} > 5/6$. Let $z_2 = z_1 - Q_T z_1$. Then

$$\|z_2 - z_1\| = \|Q_T z_1\| = \|Q_T(z_1 - z)\| \le \|Q_T\|\,\|z_1 - z\| < 1/12,$$

hence $\|z_2\| < 13/12$, and $\mathrm{dist}(z_2, E)_{X^{**}} > 5/6 - 1/12 = 3/4$. Letting $y = z_2/\|z_2\|$, we conclude that $\mathrm{dist}(y, E)_{X^{**}} > 2/3$.

By our construction, there exists a finite set $S \subset I \setminus T$ such that $y \in \mathbf{span}[x_i : i \in S]$. Let $I' = I \setminus (T \cup S)$, and show that there exists a finite set $F \subset I'$ such that

$$E(y, \mathbf{span}[x_i : i \in T \cup (I' \setminus F)]) > 2/3.$$

Once such a set is found, then we can take $y_m = y$, and $S_m = S \cup F$.

Suppose otherwise. Then, for every F as above, there exists $y_F \in \mathbf{span}[x_i : i \in T \cup (I' \setminus F)]$, satisfying $\|y - y_F\| \le 2/3$. Observe that the set $\mathcal{F}(I')$ of finite subsets of I' forms a net, ordered by inclusion. More precisely, for $F_1, F_2 \in \mathcal{F}(I')$, we say $F_1 \prec F_2$ if $F_1 \subset F_2$. For any $F_1, F_2 \in \mathcal{F}(I')$, there exists $F_3 \in \mathcal{F}(I')$ such that $F_1 \prec F_3$ and $F_2 \prec F_3$ (in fact, we can take $F_3 = F_1 \cup F_2$). By the triangle inequality, $\|y_F\|_{X^{**}} = \|y_F\| \le 5/3$ for each F. As the unit ball of X^{**} is weak*-compact, there exists a subnet $\mathcal{A}$ of $\mathcal{F}(I')$ such that the net $(y_F)_{F \in \mathcal{A}}$ converges weak* to some $x \in X^{**}$. Then $\|y - x\|_{X^{**}} \le \sup_F \|y - y_F\| \le 2/3$. Note that, for any $j \in F \cup S$, $\langle f_j, y_F\rangle = 0$. Moreover, for every $F \in \mathcal{F}(I')$, there exists $G \in \mathcal{A}$ containing F. Therefore, $\langle f_j, x\rangle = 0$ for any $j \in I' \cup S = I \setminus T$. Then $\langle f_j, x - Q_T x\rangle = 0$ for any $j \in I$, hence $x - Q_T x \in E_0$, and therefore, $x \in E$. This, however, contradicts $\mathrm{dist}(y, E)_{X^{**}} > 2/3$. $\square$

As an application, consider a compact set $K \subset \mathbb{C}$, such that $\Omega = \mathbf{Int}(K)$ is a Jordan domain, and $C = \partial K$ is a rectifiable Jordan curve. Define the family of *Faber polynomials* $\{F_n(z)\}_{n=0}^{\infty}$, associated with K. Let ϕ be the Riemann mapping function defined from $\mathbb{C} \setminus \overline{\mathbb{D}}$ onto $\mathbb{C} \setminus K$. Then

$$F_n(z) = \frac{1}{2\pi i} \int_{|w|=1} \frac{w^n \phi'(w)}{\phi(w) - z} dw.$$

These polynomials play a main role in complex approximation theory, so the dictionary $\mathcal{D} = \{F_n\}_{n=0}^{\infty}$ is of interest (see [60,17] for more information on Faber polynomials).

Corollary 6.4. *Let K be a closed Jordan domain of bounded boundary rotation, such that the boundary $C = \partial K$ has no external cusps. Let $\mathcal{D} = \{F_n\}_{n=0}^{\infty}$, where $F_n(z)$ denotes the n-th Faber polynomial associated to K. Then $\mathcal{D}$ satisfies Shapiro's Theorem on $A(K)$.*

Proof. We show that, for K as in the statement of the theorem, the Faber polynomials form a complete minimal system in $A(K)$. An application of Theorem 6.2 completes the proof.

On $K = \overline{\mathbb{D}}$, the Faber polynomials are the monomials e_n ($e_n(z) = z^n$). It is well known that $\mathbf{span}[e_n : n \ge 0]$ is dense in $A(\overline{\mathbb{D}})$. Moreover, the functionals $f \mapsto \widehat{f}(n)$ are biorthogonal to the e_n's. In the general case, by Gaier [27, Chapter 1, Section C], there exists a bounded injective operator $T : A(\overline{\mathbb{D}}) \to A(K)$, such that $Te_n = F_n$ for any $n \geq 0$. By Anderson and Clunie [6], the range of T coincides with $A(K)$, and $\|T^{-1}\| < \infty$. As an isomorphic image of a complete minimal system is again a complete minimal system, we are done. □

6.2. Generalized Haar schemes

In this section we introduce and investigate the class of generalized Haar families in spaces of functions (numerous examples will be given below). Suppose, for each n, A_n is a set of continuous functions on Ω. We say that the family $\{A_n\}$ is *generalized Haar* if there exists a function $\psi = \psi_{\{A_n\}} : \mathbb{N} \to \mathbb{N}$ such that no non-zero function of the form $\Re g$ ($g \in A_n$) has more than $\psi(n) - 1$ zeros on Ω. Finally, the approximation scheme $(X, \{A_n\})$ is named "generalized Haar" if $\{A_n\}$ is a generalized Haar system.

Very often, we consider the approximation schemes arising from dictionaries (see (4.1) for the definition). We say that a dictionary $\mathcal{D}$ is a *generalized Haar system* if the family $\{\Sigma_n(\mathcal{D})\}$ is Haar.

In the four examples below, we exhibit some generalized Haar dictionaries. The space X is either $C([a, b])$, or $L_p(a, b)$ ($0 < p < \infty$), and $\psi(n) = n$.

(1) The dictionary $\mathcal{D}$, consisting of the functions $f_\lambda(t) = t^\lambda$ ($\lambda \in \mathbb{R}$) on an interval $[a, b]$ with $0 < a < b$. Indeed, these functions form a generalized Haar system [10, Section 3.1]. As polynomials are dense in $C([a, b])$, **span**$[\mathcal{D}]$ is dense in X.
(2) The dictionary $\mathcal{D}$, consisting of functions $f_k(t) = t^k$ ($k \in \mathbb{N} \cup \{0\}$) on arbitrary $[a, b]$. Indeed, the family (f_k) forms a generalized Haar system on subintervals of $(0, \infty)$, and of $(-\infty, 0)$.
(3) The dictionary $\mathcal{D}$, consisting of functions $f_\lambda(t) = \exp(\lambda t)$ ($\lambda \in \mathbb{R}$), with arbitrary $[a, b]$. In this case, the density of **span**$[\mathcal{D}]$ in $C([a, b])$ can be deduced, for instance, from Stone–Weierstrass theorem. Furthermore, $\mathcal{D}$ is a generalized Haar system, by [10, Chapter 3].
(4) The dictionary $\mathcal{D}$, consisting of functions $f_k(t) = t^k$ on $\mathbb{R}$. Consider a weight W—that is, an L_1 function $W : \mathbb{R} \to [0, 1]$. Consider the measure μ, defined by $\mu(E) = \int_E W(x)\,dx$. Take X to be either $L_p(\mu)$ ($1 \leq p < \infty$), or a set of continuous functions f on $\mathbb{R}$ satisfying $\lim_{t\to\infty} f(t)W(t) = 0$. For certain weights W, **span**$[\mathcal{D}]$ is known to be dense in X. For instance, this is true for $W(x) = \exp(-|x|^\alpha)$, for any $\alpha \geq 1$. See [43] for this and other results on the density of polynomials in the weighted spaces X.

Moreover, the sets of trigonometric functions

$$\mathcal{T}_n = \mathbf{span}[\{1, \cos(t), \sin(t), \ldots, \cos(nt), \sin(nt)\}]$$

define a Haar system on $[0, 2\pi)$. A somewhat more complicated example of generalized Haar system involves rational functions. For $\Omega \subset \mathbb{C}$, denote by $R_n(\Omega)$ the set of all rational functions $p(z)/q(z)$, where the polynomials $p(z) = \sum_{k=0}^n a_k z^k$ and $q(z) = \sum_{k=0}^n b_k z^k$ have complex coefficients and degree $\leq n$, such that $q(z)$ does not vanish in Ω. We also consider the set $E_n(\Omega)$ of trigonometric rational functions of degree less than n, consisting of functions $t \mapsto p(e^{it})/q(e^{it})$, where $p(z) = \sum_{k=-n}^n a_k z^k$ and $q(z) = \sum_{k=-n}^n b_k z^k$, and $q(z) \neq 0$ for all $z \in \Omega$.

Proposition 6.5. *If $\Omega \subset \mathbb{R}$, then $\{R_n(\Omega)\}$ is a generalized Haar system. Moreover, if $\Omega \subseteq \partial\mathbb{D} = \mathbb{T}$ then $\{E_n(\Omega)\}$ is a generalized Haar family.*

Proof. We handle $\{R_n(\Omega)\}$ first. If $g = p/q \in R_n(\Omega)$, then $p = \Re p + (\Im p)i$, $q = \Re q + (\Im q)i$, and $\Re p, \Re q, \Im p, \Im q$ are polynomials of degree $\leq n$. Hence $\Re g = \Re\left(\frac{p\overline{q}}{|q|^2}\right) = \frac{\Re p \Re q + \Im p \Im q}{|q|^2}$. As $t \mapsto \Re p(t)\Re q(t) + \Im p(t)\Im q(t)$ is a polynomial of degree not exceeding $2n$, $\Re g$ must vanish if it has more than $2n$ zeros.

Now consider $g = p/q \in E_n(\Omega)$, with $\Omega \subseteq \partial\mathbb{D}$. Then $p(t) = \sum_{|j|\leq n} a_j e^{itj}$, $q(t) = \sum_{|j|\leq n} b_j e^{itj}$, and

$$g(t) = \frac{\sum_{k=-n}^{n} a_k e^{ikt}}{\sum_{k=-n}^{n} b_k e^{ikt}} = \frac{\sum_{k=0}^{2n} a_{k-n} z^k}{\sum_{k=0}^{2n} b_{k-n} z^k}; \quad (z = e^{it}), \tag{6.2}$$

so that

$$\Re g(t) = \Re\left(\frac{\left(\sum_{k=0}^{2n} a_{k-n} z^k\right)\left(\sum_{k=0}^{2n} \overline{b_{k-n}} z^{-k}\right)}{\left|\sum_{k=0}^{2n} b_{k-n} z^k\right|^2}\right) = \Re\left(\frac{\sum_{k=-2n}^{2n} c_k z^k}{\left|\sum_{k=0}^{2n} b_{k-n} z^k\right|^2}\right) \quad (z = e^{it}).$$

Representing $c_k = \alpha_k + i\beta_k$ ($\alpha_k, \beta_k \in \mathbb{R}$), we see that the zeros of $\Re g(t)$ are the zeros of

$$\begin{aligned} x(t) &= \Re\left(\sum_{k=-2n}^{2n} (\alpha_k + i\beta_k)e^{ikt}\right); \\ &= \Re\left(\sum_{k=-2n}^{2n} (\alpha_k + i\beta_k)(\cos(kt) + i\sin(kt))\right) \\ &= \sum_{k=-2n}^{2n} (\alpha_k \cos(kt) - \beta_k \sin(kt)) \\ &= \alpha_0 + \sum_{k=1}^{2n}((\alpha_k + \alpha_{-k})\cos(kt) + (\beta_{-k} - \beta_k)\sin(kt)). \end{aligned}$$

Thus, $x \in \mathcal{T}_{2n}$, and we are done, since $\{\mathcal{T}_n\}_{n=1}^{\infty}$ is a Haar family. □

The next result shows that "many" generalized Haar approximation schemes satisfy Shapiro's Theorem. Below, $C_0(I)$ denotes the closure of continuous functions with compact support in the $\|\cdot\|_\infty$ norm. In particular, $C(I) = C_0(I)$ if I is a compact set.

Theorem 6.6. *Suppose I is either a finite or infinite interval in $\mathbb{R}$, or the unit circle $\mathbb{T}$. Suppose, furthermore, that μ is a finite atomless Radon measure on I, and X is a quasi-Banach space of functions on I, satisfying $L_p(\mu) \supseteq X \supseteq C_0(I)$ with some $p > 0$. Then any generalized Haar approximation scheme $(X, \{A_n\})$ satisfies Shapiro's Theorem.*

Proof. Without loss of generality, we can assume $\mu(I) = 1$, and $\int |f|^p \, d\mu \leq \|f\|_X^p$ for any $f \in X$. By Closed Graph Theorem, there exists a constant C such that, for any $f \in C_0(I)$, we have $\|f\|_X \leq C\|f\|_\infty$.

For every $n \in \mathbb{N}$, we find a continuous function $h : I \to [-1, 1]$ with compact support, such that $\|h\|_\infty = 1$, and $\int |h - g|^p \, d\mu > 1/5$ for any $g \in A_n$. Indeed, if such an h exists, then $\|C^{-1}h\|_X \leq 1$ and $E(C^{-1}h, A_n)_X > 1/(5C)^{1/p}$. By Corollary 3.7, (A_n) satisfies Shapiro's Theorem in X.

As the measure μ is Radon, $1 = \mu(I) = \sup\{\mu([\alpha, \beta]) : [\alpha, \beta] \subset I\}$. Pick $\alpha < \beta$ in I such that $A = \mu([\alpha, \beta]) > 4/5$. Let $N = \psi(n) + 1$. Set $t_0 = \alpha, t_{4N} = \beta$. As the map $s \mapsto \mu((a, s))$ is continuous, we can find $t_0 < t_1 < \cdots < t_{4N}$ such that, for $1 \leq j \leq 4n, \mu((t_{j-1}, t_j)) = A/(4N) > 1/(5N)$. Recall that, for any $a < b, \mu((a, b))$ is the supremum of $\int \rho \, d\mu$, taken over all non-negative continuous functions ρ, supported on (a, b), and such that $\|\rho\|_\infty \leq 1$. So, for $1 \leq j \leq 4N$, we can find continuous $h_j : \mathbb{R} \to [0, 1]$, supported on (t_{j-1}, t_j), such that $\int_{t_{j-1}}^{t_j} h_j \, d\mu > 1/(5N)$.

We shall show that $h = \sum_{j=1}^{4N} (-1)^j h_j^{1/p}$ satisfies $\int |h - g|^p \, d\mu > 1/5$ for any $g \in A_n$. As the function h defined above is real-valued, it suffices to prove the inequality $\int |h - \Re g|^p \, d\mu > 1/5$. If $\Re g$ is identically 0, the desired inequality follows from the definition of h. Otherwise, denote by $\mathcal{S}$ the set of points where $\Re g$ changes sign. As $|\mathcal{S}| < N$, the set $\mathcal{F} = \{1 \leq k \leq 2N : (t_{2k-2}, t_{2k}) \cap \mathcal{S} = \emptyset\}$ has the cardinality larger than N. Note that, for $k \in \mathcal{F}$, $\int_{t_{2k-2}}^{t_{2k}} |h - \Re g|^p \, d\mu > 1/(5N)$. Indeed, if $g \leq 0$ on (t_{2k-2}, t_{2k}), then

$$\int_{t_{2k-2}}^{t_{2k}} |h - g|^p \, d\mu \geq \int_{t_{2k-1}}^{t_{2k}} |h - \Re g|^p \, d\mu \geq \int_{t_{2k-1}}^{t_{2k}} h_{2k} \, d\mu > \frac{1}{5N}.$$

The case of $g \geq 0$ is handled similarly. Thus,

$$\int |h - g|^p \, d\mu \geq \sum_{k \in \mathcal{F}} \int_{t_{2k-2}}^{t_{2k}} |h - g|^p \, d\mu > |\mathcal{F}| \cdot \frac{1}{5N} > \frac{1}{5},$$

completing the proof. $\square$

A similar result holds in the analytic case. Below, A and H_p refer to the disk algebra and to the Hardy space, respectively.

Proposition 6.7. *Suppose X is a quasi-normed space of analytic functions on $\mathbb{T}$, such that $A \subset X \subset H_p$ for some $p > 0$. Suppose $(X, \{A_n\})$ is a generalized Haar approximation scheme. Then $(X, \{A_n\})$ satisfies Shapiro's Theorem.*

Sketch of the proof. We identify $\mathbb{T}$ with $[0, 1]$. Let $N = \psi(n)$, and consider $h(t) = -i \exp(4\pi N i t)$. As in the proof of the previous theorem, it suffices to show that, for any $g \in A_n$, $\int_0^1 |\Re h(t) - \Re g(t)|^p \, dt > c/4$, where $c = \int_0^1 |\sin u|^p \, du$. Note that $\Re h(t) = \sin(4\pi N t)$. For $0 \leq j \leq 4N$, set $t_j = j/(4N)$. Then $\int_{t_{j-1}}^{t_j} |\Re h|^p = c/(4N)$ for any j. If $\Re g$ is identically 0, then

$$\int_0^1 |\Re h - \Re g|^p \, dt = \sum_{j=1}^{4N} \int_{t_{j-1}}^{t_j} |\Re h|^p = c.$$

Otherwise, denote by $\mathcal{F}$ the set of all $k \in \{1, \ldots, 2N\}$ such that $\Re g$ does not vanish on (t_{2k-2}, t_{2k}). As $|\mathcal{F}| < N$, we complete the proof as in Theorem 6.6. $\square$

Another interesting Banach space is $CBV_0(a,b) = \{f \in C([a,b]) : f(a) = 0, V_{[a,b]}(f) < \infty\}$, equipped with the norm $\|f\|_{BV} = V_{[a,b]}(f)$ (here $V_{[a,b]}(f)$ denotes the total variation of f).

Theorem 6.8. *Let $\mathcal{D}$ be a dictionary on $CBV_0(a,b)$. Suppose $\mathcal{D} \subset C^1([a,b])$, and $\mathcal{D}' = \{g' : g \in \mathcal{D}\}$ is a generalized Haar system on $[a,b]$. Then $\mathcal{D}$ satisfies Shapiro's Theorem in $CBV_0(a,b)$.*

Proof. We work with the case of $[a,b] = [0, 2\pi]$. By Corollary 3.7, we only need to prove that

$$\sup_{\|f\|_{BV}=1} E(f, \Sigma_n(\mathcal{D})) \geq \frac{1}{3} \quad \text{for } n = 0, 1, 2, \ldots.$$

Let $N = 6\psi(n)$, and consider $f(t) = (1-\cos Nt)/(4N)$. Then $\|f\|_{BV} = 4^{-1}\int_0^{2\pi} |\sin Nt|\, dt = 1$. We show that, for any $g \in \Sigma_n(\mathcal{D})$,

$$\|f - g\|_{BV} \geq \int_0^{2\pi} |f'(t) - \Re g'(t)| dt \geq \frac{1}{3}.$$

For such a g, define $\mathcal{F}$ as the set of all $\ell \in \{1, \ldots, N\}$ with the property that $\Re g'$ does not change sign on $(2\pi(\ell-1)/N, 2\pi\ell/N)$. Note that $|\mathcal{F}| \geq N - \psi(n) = 5N/6$. Furthermore, f' is positive on $(2\pi(\ell-1)/N, \pi(2\ell-1)/N)$, and negative on $(\pi(2\ell-1)/N, 2\pi\ell/N)$. One of these two intervals, $|f'| \geq |f' - \Re g|$. Furthermore,

$$\int_{2\pi(\ell-1)/N}^{\pi(2\ell-1)/N} |f'|\, dt = \int_{\pi(2\ell-1)/N}^{2\pi\ell/N} |f'|\, dt = \frac{1}{2N}.$$

Thus, for $\ell \in \mathcal{F}$,

$$\int_{2\pi(\ell-1)/N}^{2\pi\ell/N} |f' - \Re g'|\, dt \geq \frac{1}{2N},$$

and therefore,

$$\|f - g\|_{BV} \geq \int_0^{2\pi} |f'(t) - \Re g'(t)| dt \geq \sum_{\ell \in \mathcal{F}} \int_{2\pi(\ell-1)/N}^{2\pi\ell/N} |f' - \Re g'|\, dt \geq \frac{1}{2N} \cdot \frac{5N}{6} > \frac{1}{3}. \quad \square$$

6.3. Approximation by rational functions

The problem of describing the possible sequences of best rational approximations for a given function dates back at least to Dolzhenko [22]. Certain Bernstein-type results have been obtained for approximations in the uniform norm. For instance, if $\varepsilon_1 > \varepsilon_2 > \cdots$ and $\lim \varepsilon_m = 0$, then there exists $f \in C(\mathbb{T})$ such that $E(f, E_m(\mathbb{T}))_{C(\mathbb{T})} = \varepsilon_m$ for every m [58] (see also [45,49] for related results). Evidence suggests that the condition that the sequence $\{\varepsilon_m\}$ is strictly increasing can be weakened. By [59], for every sequence $\{\varepsilon_m\} \searrow 0$ there exists $f \in C[0,1]$ such that $E(f, R_{2^m-1}([0,1]))_{C[0,1]} = \varepsilon_m$ for every m. On the other hand, Bernstein's Lethargy Theorem cannot be perfectly replicated for rational approximation in L_p: by [36], for any $f \in L_p(0,1)$, the sequence $E(f, R_m([0,1]))_{L_p}$ is either strictly decreasing, or eventually null.

This section attempts to (partially) answer Dolzhenko's question by proving Shapiro's Theorem for rational approximations in a variety of function spaces.

Theorem 6.9. *Take $0 < p < \infty$. The following approximation schemes satisfy Shapiro's Theorem:*

(1) *$(X, \{R_n(I)\})$, where I is a real interval and $C_0(I) \subseteq X \subseteq L_p(I)$, and $\overline{C_0(I)}^X = X$.*

(2) *$(X, \{E_n(I)\})$, where is a real interval and $C_0(I) \subseteq X \subseteq L_p(I)$, and $\overline{C_0(I)}^X = X$.*

(3) *$(X, \{E_n(\mathbb{T})\})$, where $C(\mathbb{T}) \subseteq X \subseteq L_p(\mathbb{T})$, and $\overline{C(\mathbb{T})}^X = X$.*

(4) *$(X, \{R_n(\partial\mathbb{D})\})$, where $C(\partial\mathbb{D}) \subseteq X \subseteq L_p(\partial\mathbb{D})$, and $\overline{C(\partial\mathbb{D})}^X = X$.*

(5) *$(X, \{R_n(\overline{\mathbb{D}})\})$, where $A \subseteq X \subseteq H_p$, and $\overline{A}^X = X$.*

Proof. We start noting that the densities $\overline{C_0(I)}^X = X$, $\overline{C(\mathbb{T})}^X = X$ and $\overline{A}^X = X$ are assumed to guarantee that, if our approximation scheme (A_n) is dense, for example, in $C_0(I)$, then it is also dense in X. To prove this, take $x \in X$, $\varepsilon > 0$ arbitrarily small and $C > 0$ such that $\|\cdot\|_X \le C\|\cdot\|_\infty$. Look for $f \in C_0(I)$ such that $\|x - f\|_X^q < \frac{\varepsilon}{2}$ and $a \in \bigcup_n A_n$ such that $\|f - a\|_\infty^q \le \frac{\varepsilon}{2C^q}$. Then $\|x - a\|_X^q \le \|x - f\|_X^q + \|f - a\|_X^q \le \frac{\varepsilon}{2} + C^q\|f - a\|_\infty^q \le \varepsilon$.

Now (1)–(3) are direct consequences of Proposition 6.5 and Theorem 6.6.

To deduce (4) from (3), consider a map U, taking a function $f : \partial\mathbb{D} \to \mathbb{C}$ to $\tilde{f} : \mathbb{T} \to \mathbb{C}$, where $\tilde{f}(t) = f(e^{it})$. Clearly, U is an isometry from $C(\partial\mathbb{D})$ onto $C(\mathbb{T})$, and from $L_p(\partial\mathbb{D})$ onto $L_p(\mathbb{T})$. Hence it is clear that U maps the space X isometrically onto a space Y which satisfies $C(\mathbb{T}) \subseteq Y \subseteq L_p(\mathbb{T})$. Moreover, the equality (6.2) implies that U maps $R_{2n}(\partial\mathbb{D})$ onto $E_n(\mathbb{T})$.

To establish (5), note that the elements of A or H_p are uniquely determined by their restrictions to $\partial\mathbb{D}$ (see e.g. Appendix 3 of [42]). Thus, we identify our functions on $\mathbb{D}$ with functions on $\partial\mathbb{D}$. The density of $\cup_n R_n(\overline{\mathbb{D}})$ in X follows from the proof of Theorem 1.5.2 of [42] and the density of A in X. Identifying $\partial\mathbb{D}$ with $\mathbb{T}$, we complete the proof by applying Proposition 6.7. □

Corollary 6.10. *Suppose X is either $C(\overline{\mathbb{R}})$ (the set of continuous functions f on $\mathbb{R}$ for which $\lim_{t\to+\infty} f(t)$ and $\lim_{t\to-\infty} f(t)$ exist and are equal), or $L_p(W, \mathbb{R})$, where $0 < p < \infty$, and the weight W is given by $W(x) = 2/(1 + x^2)$. For $n \in \mathbb{N}$, denote by $R_n(\overline{\mathbb{R}})$ the set of rational functions p/q, where $\deg p \le \deg q < n$, and q has no real roots. Then the approximation scheme $(X, \{R_n(\overline{\mathbb{R}})\})$ satisfies Shapiro's Theorem.*

Proof. In this proof, we use some ideas of [42, Section 1.5]. As before, identify $\mathbb{T}$ with $[-\pi, \pi]$. Consider the map $\Phi : \mathbb{T} \to \mathbb{R} : t \mapsto \tan(t/2)$ ($-\pi \sim \pi$ is taken to ∞). The map $U_\Phi : f \mapsto f \circ \Phi$ is then an isometry from Y onto X, where Y is either $C(\mathbb{T})$ or $L_p(\mathbb{T})$.

Denote by $R'_n(\overline{\mathbb{R}})$ the set of all functions $p/q \in R_n(\overline{\mathbb{R}})$ for which all the roots of q are distinct. Similarly, let $E'_n(\mathbb{T})$ the set of all functions $p/q \in E_n(\mathbb{T})$ for which all the roots of q are distinct. A small perturbation argument shows that $R'_n(\overline{\mathbb{R}})$ ($E'_n(\mathbb{T})$) is dense in $R_n(\overline{\mathbb{R}})$ (resp. $E_n(\mathbb{T})$).

Any $f \in R'_n(\overline{\mathbb{R}})$ can be written as $f = \alpha_0\mathbf{1} + \sum_{j=1}^m \alpha_j g_{c_j}$, with $m < n$. Here, $\mathbf{1}(x) = 1$, and $g_c(x) = (1 - ix)/(x - c)$ ($c \notin \mathbb{R}$). By formula (5.13) of [42], $g_c \circ \Phi = \alpha f_z$, where $z = (i - c)/(i + c)$, α is a numerical constant, depending on z, and $f_z(t) = 1/(e^{it} - z)$. Thus, Φ implements a $1 - 1$ correspondence between $R'_n(\overline{\mathbb{R}})$ and $E'_n(\mathbb{T})$.

It is established in [42, Section 1.5] that $(Y, \{E'_n(\mathbb{T})\})$ is an approximation scheme. By Theorem 6.9, $(X, \{E'_n(\mathbb{T})\})$ satisfies Shapiro's Theorem. As U_Φ is an isometry, $(X, \{R'_n(\overline{\mathbb{R}})\})$ is also an approximation scheme, satisfying Shapiro's Theorem. The density of $R'_n(\overline{\mathbb{R}})$ in $R_n(\overline{\mathbb{R}})$ completes the proof. □

Remark 6.11. Below we outline some alternative approaches to the results of Theorem 6.9. For instance, one can show that $(C([a, b]), \{R_n([a, b])\})$ satisfies Shapiro's Theorem, one can use a Bernstein-type inequality due to Dolzhenko [23]: the total variation of $f \in R_n([a,b])$ satisfies $V_{[a,b]}(f) = \int_a^b |f'(t)|dt \leq 2n\|f\|_{C([a,b])}$. The space of continuous functions of bounded variation $Y = CBV[a,b]$, equipped with the norm $\|f\|_Y = \|f\|_{C[a,b]} + V_{[a,b]}(f) < \infty$, is a proper dense linear subspace of $C([a,b])$. An application of Bernstein's Inequality (Theorem 5.4) to $\|\cdot\|_Y$ completes the proof.

For $L_p(a,b)$, $(1 < p < \infty)$, we may use a result by Pekarskii [50] (see also Theorem 1.1 in [42, page 300]): for $k = 1, 2, \ldots, 1 < p < \infty$, and $\gamma = (r + \frac{1}{p})^{-1}$, the approximation scheme $(R_n([-1,1]))$ satisfies a Bernstein-style inequality:

$$\|f^{(k)}\|_{L_\gamma(-1,1)} \leq C(p,k)n^k\|f\|_{L_p(-1,1)} \quad (f \in R_n([-1,1])),\ n = 1, 2, \ldots.$$

Hence we can use Theorem 5.4 for $(L_p(-1,1), \{R_n([-1,1])\})$ with $Y = \{f \in L_p(-1,1) : f^{(k)} \in L_\gamma(-1,1)\}$ (with the norm $\|f\| = \|f\|_p + \|f^{(k)}\|_{L_\gamma(-1,1)}$).

One can also tackle L_p by using the strong relation between rational approximation and approximation by spline functions with free knots. By Theorem 6.8 from [42, page 340], for $1 < p < \infty$, $0 < q \leq \infty$ and $0 < \alpha < r$, the approximation spaces $\mathbb{R}^\alpha_{p,q} = \{f \in L_p(0,1) : \{n^{\alpha - \frac{1}{q}} E(f, R_n(0,1))_{L_p}\} \in \ell_q\}$ and $\mathbb{S}^\alpha_{p,q} = \{f \in L_p(0,1) : \{n^{\alpha - \frac{1}{q}} E(f, S_{n,r}(0,1))_{L_p}\} \in \ell_q\}$ are the same (with equivalent norms). Theorem 6.12 guarantees that $\mathbb{R}^1_{p,\infty} = \mathbb{S}^1_{p,\infty}$ is a strict subset of $L_p(0,1)$. Hence there exists a function $f \in L_p(0,1)$ such that $E(f, R_n(0,1))_{L_p} \neq \mathbf{O}(n^{-1})$ and the result follows from Corollary 5.6.

We next sketch an argument showing that $(C(\partial\mathbb{D}), \{R_n(\partial\mathbb{D})\}_{n=0}^\infty)$ satisfies Shapiro's Theorem. That is, by Corollary 3.7, we have to find $f \in C(\partial\mathbb{D})$ such that the sequence $E(f, R_n(\partial\mathbb{D})) \geq \alpha_n$, for a prescribed sequence $\{\alpha_n\} \searrow 0$. Having already shown that $(C([0,1]), \{R_n([0,1])\})$ satisfies Shapiro's Theorem, we conclude that there exists $h \in C([0,1])$ such that $E(h, R_n([0,1])) \geq \alpha_n$ for every n. Extend h to a bounded continuous function g on $\mathbb{R}$, for which $\lim_{t\to+\infty} g(t) = \lim_{t\to-\infty} g(t)$ exist and are equal. Clearly, $E(g, R_n(\mathbb{R})) \geq E(h, R_n([0,1])) \geq \alpha_n$ for every n. Finally, consider the linear fractional transformation $w(t) = \frac{(i-1)t+(i+1)}{(1-i)t+(i+1)}$, mapping $\mathbb{R}$ onto $\partial\mathbb{D}$. Define $f = g \circ w^{-1} \in C(\partial\mathbb{D})$. Note that $R \in R_n(\mathbb{R})$ if and only if $R \circ w^{-1} \in R_n(\partial\mathbb{D})$. Therefore,

$$E(f, R_n(\mathbb{T}))_{C(\partial\mathbb{D})} = E(f \circ w, R_n(\mathbb{R}))_{C(\mathbb{R})} \geq E(h, R_n([0,1]))_{C[0,1]} \geq \alpha_n$$

for every n. We conclude that $(C(\partial\mathbb{D}), \{R_n(\partial\mathbb{D})\})$ satisfies Shapiro's Theorem. To deduce from this that $(C(\mathbb{T}), \{E_n(\mathbb{T})\}_{n=0}^\infty)$ satisfies Shapiro's Theorem, observe that (6.2) guarantees that $E(f, E_n(\mathbb{T}))_{C(\mathbb{T})} = E(f, R_{2n}(\partial\mathbb{D}))_{C(\partial\mathbb{D})}$.

6.4. Approximation by splines

In this subsection, we show that some "very redundant" approximation systems based on splines satisfy Shapiro's Theorem. Let us denote by $\mathcal{S}_{n,r}(I)$ the set of polynomial splines of degree less than r with n free knots (nodes) on the interval I. For any pair of sequences $0 \leq r_1 \leq r_2 \leq \cdots$ and $1 \leq n_1 < n_2 < \cdots$ the sets $A_i = \mathcal{S}_{n_i,r_i}([a,b])$ form an approximation scheme in $C[a,b]$ or $L_p(a,b)$, with $1 \leq p < \infty$ (in the case of $C[a,b]$, we assume that the splines in question are continuous).

Theorem 6.12. *The approximation scheme defined above (either in $C([a,b])$, or in $L_p(a,b)$, for $0 < p < \infty$) satisfies Shapiro's Theorem.*

Proof. The case of $C([a,b])$ follows from [2, Theorem 3.1]. When working with $L_p(a,b)$ $(0 < p < \infty)$, assume with no loss of generality that $[a,b] = [0,1]$. For a fixed $r \in \mathbb{N}$, consider the approximation scheme $(L_p(0,1), \{B_{n,r}\}_{n=1}^{\infty})$, where $B_{n,r} = \mathcal{S}_{n,r}([0,1])$. Pick $0 < \alpha < \min\{r, 1/p\}$, and find $t > 0$ satisfying $1/t = \alpha + 1/p$. By Theorem 8.2 of [20, page 386], a Bernstein's Inequality holds:

$$\|f\|_{B_t^{\alpha}(L_t(0,1))} \leq Cn^{\alpha}\|f\|_{L_p(0,1)} \quad (f \in B_{n,r}),\ n = 1, 2, \ldots. \tag{6.3}$$

Here, $B_q^{\alpha}(L_p(\Omega))$ denotes the classical Besov space on $[0,1]$ (defined using the modulus of smoothness $w_r(f,t)_p$). By De Vore and Lorentz [32, Corollary 3.1], $B_t^{\alpha}(L_t(0,1))$ embeds into the classical Lorentz space $L_{p,t}(0,1)$. Furthermore, as $t < p$, $L_{p,t}$ embeds into $L_{p,p} = L_p$ (see e.g. [12, Theorem 1.9.9]). It is easy to show that the last embedding is proper. By Theorem 5.4, the approximation scheme $(L_p(0,1), \{B_{n,r}\}_{n=1}^{\infty})$ satisfies Shapiro's Theorem. We complete the proof by applying Corollary 3.9. □

6.5. *n-term approximation*

In this section we study Shapiro's Theorem for n-term approximation. More precisely, suppose $\mathcal{D}$ is a dictionary, and $\Sigma_n(\mathcal{D})$is the associated approximation scheme (defined in (4.1)). Then $\Sigma_n(\mathcal{D}) + \Sigma_n(\mathcal{D}) = \Sigma_{2n}(\mathcal{D})$, so that Theorem 2.6 is applicable in this context. Obviously the properties of the sequence of errors $E(x, \Sigma_n(\mathcal{D}))$ strongly depend on the dictionary $\mathcal{D}$. For example, if $\overline{\mathcal{D}}^X = X$, then $E(x, \Sigma_n(\mathcal{D})) = 0$ for all $n \geq 1$ and the dictionary is "too rich" to be of interest.

For the sake of brevity, we say that a dictionary $\mathcal{D}$ *satisfies Shapiro's Theorem* in a quasi-Banach space X if the approximation scheme $(X, \{\Sigma_n(\mathcal{D})\})$ satisfies Shapiro's Theorem.

Proposition 3.8 implies that the dictionaries satisfying Shapiro's Theorem are stable under small perturbations:

Corollary 6.13. *Suppose a quasi-Banach space X is such that there exists $p \in (0,1]$, for which any $x_1, x_2 \in X$ satisfy $\|x_1 + x_2\|^p \leq \|x_1\|^p + \|x_2\|^p$. Consider the dictionaries $\mathcal{D}_1 = \{u_i\}_{i \in I}$ and $\mathcal{D}_2 = \{e_i\}_{i \in I}$ in X, such that $\mathcal{D}_1$ satisfies Shapiro's Theorem. Suppose, furthermore, that there exists $\lambda \in (0,1)$ such that $\|\sum_i a_i(u_i - e_i)\| \leq \lambda \|\sum_i a_i e_i\|$ for any family $(a_i)_{i \in I}$ with finitely many non-zero entries. Then $\mathcal{D}_2$ satisfies Shapiro's Theorem. In particular, $\mathcal{D}_2$ satisfies Shapiro's Theorem in the following two situations:*

(1) $(\sum |a_i|^p)^{1/p} \leq c\|\sum a_i e_i\|$ *for arbitrary scalars a_i, and* $\sup_{i \in I}\|u_i - e_i\| < c^{-1}$.
(2) $\sup |a_i| \leq c\|\sum a_i e_i\|$ *for arbitrary scalars a_i, and* $(\sum_{i \in I}\|u_i - e_i\|^p)^{1/p} < c^{-1}$.

Note that the inequality $\sup |a_i| \leq c\|\sum a_i e_i\|$ (with an appropriate constant c) is satisfied if (e_i), or even if (e_i) arises from a bounded biorthogonal system.

Below we give several examples of redundant dictionaries satisfying Shapiro's Theorem.

Proposition 6.14. *Let $\mathcal{D} = \{\chi_{(a,b)} : 0 \leq a < b \leq 1\}$ be the set of characteristic functions of subintervals of* $[0,1]$. *Then $(L_p(0,1), \{\Sigma_n(\mathcal{D})\}_{n=0}^{\infty})$ satisfies Shapiro's Theorem for* $0 < p < \infty$.

Remark 6.15. Consider the dictionary $\mathcal{D}' \subset \mathcal{D}$, consisting of characteristic functions of binary intervals in $L_p(0,1)$. By Livshitz [41], the greedy algorithm in this setting converges "very fast", when $f \in L_p(0,1)$ is such that the sequence $(E(f, \Sigma_n(\mathcal{D}')))$ decreases in a certain controlled manner. The result above shows that, in general, $(E(f, \Sigma_n(\mathcal{D}')))$ may decrease arbitrarily slowly.

Proof. It is not difficult to prove, by induction on n, that any element of $\Sigma_n(\mathcal{D})$ can be written as a linear combination of at most $2n+1$ characteristic functions of intervals with non-empty interiors. This, in turn, implies $\Sigma_n(\mathcal{D}) \subseteq \mathcal{S}_{4n+2,1}(0,1)$, and the result follows from Theorem 6.12. □

Shapiro's Theorem also holds for the dictionary of imaginary exponentials $\mathcal{D} = \{t \mapsto \exp(i\lambda t) : \lambda \in \mathbb{R}\}$ on any interval $[a, b]$. Indeed, the theorem below deals with ridge functions, and includes these exponentials as a particular case (see e.g. [16] for an introduction to ridge functions). Suppose $\Pi = \prod_{i=1}^{N}[A_i, B_i]$ is a parallelepiped in $\mathbb{R}^N$, and the dictionary $\mathcal{D}$ consists of functions $f(t) = \exp(-i\langle\alpha, t\rangle)$, with $\alpha \in \mathbb{R}^N$ and $t \in \Pi$ ($\langle\alpha, t\rangle = \sum_{i=1}^{N} \alpha_i t_i$ denotes the usual scalar product of $\mathbb{R}^N$).

Theorem 6.16. *Suppose Π is parallelepiped, and X is a Banach space of functions on Π such that $L_1(\Pi) \supset X \supset C(\Pi)$, and that $\overline{\mathbf{span}[\mathcal{D}]} = X$. Then $\mathcal{D}$ satisfies Shapiro's Theorem.*

The spaces X with the properties described above include $L_p(\Pi)$ $(1 \le p < \infty)$ and $C(\Pi)$. Indeed, $\mathbf{span}[\mathcal{D}]$ is closed under multiplication, and separates points in Π. By Stone–Weierstrass Theorem, $\mathbf{span}[\mathcal{D}]$ is dense in $C(\Pi)$. Furthermore, $C(\Pi)$ is dense in $L_p(\Pi)$ for $1 \le p < \infty$.

Proof. By scaling, we can assume $\Pi = [0, 2\pi]^N$, and that Π is equipped with the Lebesgue measure $(2\pi)^{-N} dt_1 \dots dt_N$. Renorming X, we assume that $\|f\|_{L_1} \le \|f\|$ for any $f \in X$. Let C be a constant for which $C\|f\|_\infty \ge \|f\|$. The dictionary $\mathcal{D}$ consists of functions $f_\alpha(t) = \exp(i\langle\alpha, t\rangle)$ $(\alpha \in \mathbb{R}^N)$. We shall show that $\mathcal{D}$ has Property (P). To this end, fix $n \in \mathbb{N}$, and let $x = \sum_{k=1}^{n^2} (-1)^k f_{(k,0,\dots,0)}/n^2$ (note that $f_{(k,0,\dots,0)}(t_1, \dots, t_N) = \exp(ikt_1)$). Clearly, $\|x\| \le C$. Consider a family $\alpha^{(j)} = (\alpha_1^{(j)}, \dots, \alpha_N^{(j)}) \in \mathbb{R}^N$ $(1 \le j \le m \le n-1)$, and scalars $a_1, \dots, a_m$. Let $y = x + z$, where $z = \sum_{j=1}^{m} a_j f_{\alpha^{(j)}}$. We shall show that $\|y\| \ge 1/n^2$.

Perturbing the $\alpha^{(j)}$'s slightly, we can assume that all the quantities $\alpha_1^{(j)}$ are different, and non-integer. To estimate $\|y\|$, recall that, for a multi-index $k = (k_1, \dots, k_N) \in \mathbb{Z}^N$, and a function ϕ defined on Π, we define the Fourier coefficient

$$\hat{\phi}(k) = \langle\phi, f_k\rangle = \frac{1}{(2\pi)^N} \int_0^{2\pi} \cdots \int_0^{2\pi} \phi(t_1, \dots, t_N) \exp(-i(k_1 t_1 + \cdots + k_N t_N)) \, dt_1 \cdots dt_N.$$

We shall show that, for at least one value $k \in \{1, \dots, n^2\}$, $(-1)^k \Re(\hat{z}(k, 0, \dots, 0)) \ge 0$. Once this is done, we conclude that

$$\begin{aligned}\|y\| &\ge \|y\|_1 \ge |\hat{y}(k, 0, \dots, 0)| = |\hat{x}(k, 0, \dots, 0) + \hat{z}(k, 0, \dots, 0)| \\ &= \left| \frac{(-1)^k}{n^2} + \hat{z}(k, 0, \dots, 0) \right| \ge \left| \Re\left(\frac{(-1)^k}{n^2} + \hat{z}(k, 0, \dots, 0) \right) \right| \ge 1/n^2,\end{aligned}$$

which is what we need.

A straightforward calculation shows that, for $\alpha = (\alpha_1, \dots \alpha_N)$,

$$\hat{f}_\alpha(k, 0, \dots, 0) = \frac{c_{\alpha_1} \dots c_{\alpha_N}}{(\alpha_1 - k)\alpha_2 \dots \alpha_N}, \quad \text{where } c_\beta = \frac{\exp(i\beta) - 1}{2\pi i}.$$

Let $b_j = \Re\big(a_j c_{\alpha_1^{(j)}} \dots c_{\alpha_N^{(j)}}\big)$. Suppose, for the sake of contradiction,

$$\Re\big(\hat{z}(k, 0, \dots, 0)\big) = \sum_{j=1}^{m} b_j \frac{1}{\alpha_1^{(j)} - k} \frac{1}{\alpha_2^{(j)}} \cdots \frac{1}{\alpha_N^{(j)}}$$

has the same sign as $(-1)^{k+1}$ for every value of k. As $m < n$, there exists $L \in \{1, \dots, n(n-1)\}$ such that $[L, L+n-1] \cap \{\alpha_1^{(1)}, \dots, \alpha_1^{(m)}\} = \emptyset$. Indeed, $\{\alpha_1^{(1)}, \dots, \alpha_1^{(m)}\}$ partitions $[1, n^2]$ into no more than $m+1$ subintervals. If each of these subintervals contains less than n integer points, then the total number of integer points on $[1, n^2]$ cannot exceed $(n-1)(n+1)$, which is clearly false.

For $t \in [L, L+n-1]$ define

$$\phi(t) = \sum_{j=1}^{m} b_j \frac{1}{\alpha_1^{(j)} - t} \frac{1}{\alpha_2^{(j)}} \cdots \frac{1}{\alpha_N^{(j)}}.$$

By assumption, $\phi(k)$ is positive when k is an odd integer, and negative if k is an even integer. Therefore, for $s \in \{1, \ldots, n-1\}$ there exists $t_s \in (L+s-1, L+s)$ such that $\phi(t_s) = 0$.

Now consider the $m \times m$ matrix

$$A = \left[\left(\frac{1}{\alpha_1^{(j)} - t_s} \frac{1}{\alpha_2^{(j)}} \cdots \frac{1}{\alpha_N^{(j)}} \right)_{j,s=1}^{m} \right],$$

and the vector $b = (b_1, \ldots, b_m)^t$. Then $Ab = 0$, hence the matrix A is singular. However, by Cauchy's Lemma (see e.g. [15, p. 195]), the determinant of the matrix with entries $((x_i - y_j)^{-1})$ equals $\prod_{i<j}(x_i - x_j)(y_i - y_j)/\prod_{i,j}(x_i + y_j)$, hence A is non-singular. □

Theorem 6.16 can be connected to the problem of approximation by elements of a frame (see e.g. [13] for an introduction to the topic). By Corollary 3.10 of [31], any normalized tight frame $\mathcal{F}$ in a Hilbert space of the form $(U^n \eta)_{n \in \mathbb{Z}}$ (U is a unitary operator) is unitarily equivalent to the set $\mathcal{D}$ of the functions $t \mapsto \exp(2\pi i t)|_E$, where E is an essentially unique measurable subset of $[0, 2\pi]$. If E contains an interval, Theorem 6.16 shows that $\mathcal{D}$ (and therefore, $\mathcal{F}$) satisfies Shapiro's Theorem. We do not know whether this remains true for general sets E.

In general, a frame $\mathcal{D}$ need not satisfy Shapiro's Theorem. For instance, we can find a family of vectors $(u_i^{(j)})_{i,j \in \mathbb{N}}$, dense in $S(\ell_2)$, such that $(u_i^{(j)})_{i \in \mathbb{N}}$ is an orthonormal basis for every j. If $\sum_j |\alpha_j|^2 = 1$, then $\mathcal{D} = (\alpha_j u_i^{(j)})_{i,j \in \mathbb{N}}$ is a tight frame (that is, $\sum_{e \in \mathcal{D}} |\langle f, e \rangle|^2 = \|f\|^2$ for any $f \in \ell_2$), yet clearly $\mathcal{D}$ fails Shapiro's Theorem. Frames which are "not too rich", however, do satisfy Shapiro's Theorem. For instance, suppose a frame has *finite excess*—that is, the removal of finitely many elements turns it into a basis (see e.g. [40] for some remarkable properties of frames with finite excess). Theorem 6.1 shows that such frames satisfy Shapiro's Theorem. Another class of interest is that of *Riesz frames*—that is, of frames $(f_i)_{i \in I}$ for which there exist positive constants $A \le B$ such that, for every $J \subset I$, and every $f \in \mathbf{span}[f_i : i \in J]$, $A\|f\|^2 \le \sum_{i \in J} |\langle f_i, f \rangle|^2 \le B\|f\|^2$.

Proposition 6.17. *If a dictionary $\mathcal{D}$ is a Riesz frame in ℓ_2, then it satisfies Shapiro's Theorem.*

Proof. By Casazza [14, Theorem 2.4], we can represent $\mathcal{D}$ as a union of two disjoint subsets: an unconditional basis $(g_i)_{i \in \mathbb{N}}$, and a family $(h_i)_{i \in \Gamma}$ (Γ may be finite or infinite), such that, for every $i \in \Gamma$, there exists a set Δ_i such that $h_i \in \mathbf{span}[g_j : j \in \Delta_i]$, and $K = \sup_{i \in \Gamma} |\Delta_i| < \infty$.

By Casazza [13, Proposition 4.3], there exist $0 < C \le D$ (depending only on A and B) with the property that $C^2 \sum_j |\alpha_j|^2 \le \|\sum_j \alpha_j g_j\|^2 \le D^2 \sum_j |\alpha_j|^2$ for any $(\alpha_j) \in \ell_2$. Consider $y = \sum_{j=1}^{2nK} g_j/(D\sqrt{2nK})$ and

$$z = \sum_{i \in \mathcal{A}} \alpha_i h_i + \sum_{j \in \mathcal{B}} \beta_j g_j$$

with $|\mathcal{A}| + |\mathcal{B}| \le n$. Then $\|y\| \le 1$. We show that $\|y - z\| \ge C/(D\sqrt{2})$. As (g_j) is a basis, we can write $y - z = \sum_{j=1}^{\infty} \gamma_j g_j$. Note that, for

$$j \in \mathcal{C} = \{1, \ldots, 2nK\} \setminus \big(\mathcal{B} \cup (\cup_{i\in\mathcal{A}} \Delta_i)\big),$$

$\gamma_j = 1/(D\sqrt{2nK})$. As $|\mathcal{C}| \ge nK$, we conclude that $\|y - z\| \ge C/(D\sqrt{2})$. Since this inequality holds for any $z \in \Sigma_n(\mathcal{D})$, Corollary 3.7 completes the proof. □

Certain approximation schemes related to MRA wavelets also satisfy Shapiro's Theorem. In the exposition below, we follow the notation of [66]. Suppose ϕ is a scaling function in $L_2(\mathbb{R})$. More precisely, suppose $\|\phi\| = 1$. For $k, j \in \mathbb{Z}$, let $\phi_{k,j}(x) = 2^{k/2}\phi(2^k x - j)$. Let $V_k = \mathbf{span}[\{\phi_{k,j} : j \in \mathbb{Z}\}]$. We are assuming that $V_k \subset V_{k+1}$ for any $k \in \mathbb{Z}$, $\overline{\cup_k V_k} = L_2(\mathbb{R})$, and $\cap_k V_k = \{0\}$. Moreover, we assume that $\{\phi_{0,j}\}_{j\in\mathbb{Z}}$ is an orthonormal basis for V_0. Now we consider the dictionary $\mathcal{D} = \{\phi_{k,j} : k, j \in \mathbb{Z}\}$ in $L_p(\mathbb{R})$, for $1 < p < \infty$ and its associated n-term approximation scheme $A_n = \Sigma_n(\mathcal{D})$.

Theorem 6.18. *In the above notation, suppose the scaling function ϕ has compact support. Then:*

(i) $(L_2(\mathbb{R}), \{A_n\})$ *satisfies Shapiro's Theorem.*
(ii) *Suppose, furthermore, that $\phi \in L_\infty(\mathbb{R})$. Then $(L_p(\mathbb{R}), \{A_n\})$ satisfies Shapiro's Theorem for $1 \le p < \infty$.*

Note first that the orthogonal projection from $L_2(\mathbb{R})$ onto V_k is given by

$$P_k f = \sum_{j\in\mathbb{Z}} \phi_{k,j} \int_{\mathbb{R}} f(t)\overline{\phi}_{k,j}(t)\,dt. \tag{6.4}$$

If $\phi \in L_\infty(\mathbb{R})$, then this family of projections is also uniformly bounded on $L_p(\mathbb{R})$, for any $p \in [1, \infty)$ (see [66, Section 8.1] for the proof of this fact, and for further properties of these projections).

Define the map D_k by setting $D_k f(x) = f(2^k x)$. Then $V_k = D_k(V_0)$ for any k, and $P_k = D_k P_0 D_{-k}$.

Lemma 6.19. *Suppose ϕ is a scaling function in $L_2(\mathbb{R})$ with compact support. Then, for any $n \in \mathbb{N}$ and $\varepsilon > 0$, there exists $N \in \mathbb{N}$ with the property that, for any $S \subset \{N, N+1, \ldots,\} \times \mathbb{Z}$ of cardinality n or less, and any $f = \sum_{s\in S} \alpha_s \phi_s$, $\|P_0 f\| \le \varepsilon \|f\|$.*

Proof. Let $T = \{t \in \mathbb{R} : \phi(t) \ne 0\}$, and $T_{k,j} = 2^{-k/2}(T + j)$. By assumption, T is (up to a set of measure zero) a subset of a certain interval I, of length $|I|$. Then $T_{k,j}$ belongs to an interval of length $2^{-k/2}|I|$. Thus, there exists a constant K such that

$$|\{\ell \in \mathbb{Z} : |T_{k,j} \cap T_{0,\ell}| > 0\}| \le K$$

for any $k \ge 0$ and $j \in \mathbb{Z}$.

Consider $f = \sum_{i=1}^{n} \alpha_i \phi_{k_i,j_i}$, with $k_i \ge 0$. Let $T_f = \cup_i T_{k_i,j_i}$. Then

$$|\{\ell \in \mathbb{Z} : |T_f \cap T_{0,\ell}| > 0\}| \le Kn.$$

Find $N \in \mathbb{N}$ such that $|\langle g, \phi\rangle| \le \varepsilon\|g\|/Kn$ whenever g differs from 0 on a set of measure at most $n2^{-N}d$. Then, for any $f = \sum_{i=1}^{n} \alpha_i \phi_{k_i,j_i}$, with $k_i \ge N$, $|\langle f, \phi_{0,j}\rangle| \le \varepsilon\|f\|/(Kn)$. Moreover,

$\langle f, \phi_{0,j} \rangle \neq 0$ for at most Kn different values of j. To complete the proof, recall that, by (6.4),

$$P_0 f = \sum_{j \in \mathbb{Z}} \langle f, \phi_{0,j} \rangle \phi_{0,j}. \quad \square$$

A variant of the previous lemma (with identical proof) also holds for $1 < p < \infty$.

Lemma 6.20. *Suppose ϕ is a scaling function in $L_\infty(\mathbb{R})$ with compact support, and $1 \leq p < \infty$. Then, for any $n \in \mathbb{N}$ and $\varepsilon > 0$, there exists $N \in \mathbb{N}$ with the property that, for any $S \subset \{N, N+1, \ldots, \} \times \mathbb{Z}$ of cardinality n or less, and any $f = \sum_{s \in S} \alpha_s \phi_s$, $\|P_0 f\|_p \leq \varepsilon \|f\|_p$.*

Proof of Theorem 6.18. We prove part (i) only, since part (ii) is handled in the same manner, with minimal changes. By Theorem 2.6, it suffices to show that this approximation scheme has Property (P). To this end, fix $n \in \mathbb{N}$. Let $c = 1/(8\sqrt{n+1})$. By Lemma 6.19, there exists $N \in \mathbb{N}$ such that $\|P_0 f\| \leq c\|f\|$ for any $f = \sum_{i=1}^n \alpha_i \phi_{k_i, j_i}$ whenever $k_i \geq N - 1$ for each i. It is easy to see that, for any $m \in \mathbb{Z}$, we have $\|P_m f\| \leq c\|f\|$ for any $f = \sum_{i=1}^n \alpha_i \phi_{k_i, j_i}$ whenever $k_i \geq m + N - 1$ for each i.

Find norm 1 vectors $x_s \in V_{sN+1} \ominus V_{sN}$ $(0 \leq s \leq n)$. Let $x = (x_0 + \cdots + x_n)/\sqrt{n+1}$. We show that $E(x, A_n) \geq c$. Indeed, consider $f = \sum_{i=1}^n \alpha_i \phi_{k_i, j_i} \in A_n$, and suppose, for the sake of contradiction, that $\|x - f\| < c$. By Pigeon-Hole Principle, there exists $s \in \{0, \ldots, n\}$ with the property that no k_i belongs to $\{sN, \ldots, (s+1)N - 1\}$. Let $f_- = \sum_{k_i \leq sN} \alpha_i \phi_{k_i, j_i}$, and $f_+ = \sum_{k_i \geq (s+1)N} \alpha_i \phi_{k_i, j_i}$. Note that, by our choice of N, $\|P_m f_+\| \leq c\|f_+\|$ whenever $m \leq sN + 1$.

In this notation,

$$c > \|x - f\| \geq \|(I - P_{sN})(x - f)\| = \left\| \frac{x_s + \cdots + x_n}{\sqrt{n+1}} - (I - P_{sN}) f_+ \right\|.$$

By the triangle inequality, $\|(I - P_{sN}) f_+\| < 1 + c$. Therefore, $\|f_+\| < 2$. Indeed, otherwise we would have

$$1 + c > \|(I - P_{sN}) f_+\| \geq \|f_+\| - \|P_{sN} f_+\| \geq (1 - c)\|f_+\| \geq 2(1 - c),$$

which contradicts the fact that $c < 1/8$.

Similarly,

$$c > \|x - f\| \geq \|(I - P_{sN+1})(x - f)\| = \left\| \frac{x_{s+1} + \cdots + x_n}{\sqrt{n+1}} - (I - P_{sN+1}) f_+ \right\|.$$

Thus, by the triangle inequality,

$$\begin{aligned} 2c &> \left\| \left(\frac{x_s + \cdots + x_n}{\sqrt{n+1}} - (I - P_{sN}) f_+ \right) - \left(\frac{x_{s+1} + \cdots + x_n}{\sqrt{n+1}} - (I - P_{sN+1}) f_+ \right) \right\| \\ &= \left\| \frac{x_{s+1}}{\sqrt{n+1}} + P_{sN} f_+ - P_{sN+1} f_+ \right\| \geq \left\| \frac{x_{s+1}}{\sqrt{n+1}} \right\| - \|P_{sN} f_+\| - \|P_{sN+1} f_+\|. \end{aligned}$$

We know that $\|P_m f_+\| \leq c\|f_+\| \leq 2c$ for any $m \leq sN + 1$. Recall that $\|x_{s+1}/\sqrt{n+1}\| = 8c = 1/\sqrt{n+1}$. The previous centered inequality then implies $2c > 8c - 2c - 2c = 4c$, a contradiction. $\square$

Next we deal with the dictionaries in $L_p(\mathbb{R})$ or $C_0(\mathbb{R})$ arising from translates of a single function. More precisely, for $\phi \in L_p(\mathbb{R})$, consider the set $\mathcal{D} = \{\phi_c : c \in \mathbb{R}\}$, with $\phi_c(t) = \phi(t - c)$. It is a well known result by Wiener (see [65, pp. 97–103], or [25, Chapter 8]) that **span**$[\mathcal{D}]$ is dense in $L_1(\mathbb{R})$ if and only if the Fourier transform of ϕ does not vanish on $\mathbb{R}$, and **span**$[\mathcal{D}]$ is dense in $L_2(\mathbb{R})$ if and only if the Fourier transform of ϕ vanishes only on a measure 0 subset of $\mathbb{R}$. This condition is satisfied, for instance, if ϕ is a Gaussian function $\phi(t) = e^{-at^2/2}$, for some $a > 0$.

Theorem 6.21. *Suppose X is either $L_p(\mathbb{R})$ $(0 < p < \infty)$ or $C_0(\mathbb{R})$, and ϕ is a function in X. Denote by $\mathcal{D}$ the set of translates $\{\phi_c : c \in \mathbb{R}\}$. Then the approximation scheme $(X, \{\Sigma_n(\mathcal{D})\})$ satisfies Shapiro's Theorem in each of the following two cases:*

(1) *ϕ has compact support, and the linear span of its translates is dense in X.*
(2) *ϕ is a Gaussian function.*

Proof. (1) We consider the case of $X = L_p(\mathbb{R})$. The space $C_0(\mathbb{R})$ can be tackled in a similar fashion. By Corollary 3.7, it suffices to show that, for any $n \in \mathbb{N}$ and $\varepsilon > 0$, there exists $f \in L_p(\mathbb{R})$ such that $E(f, \Sigma_n(\mathcal{D})) > 1 - \varepsilon$. To this end, pick $m \in \mathbb{N}$ such that $n/m < \varepsilon$. Find a finite interval I such that ϕ vanishes outside of I. Set $f = m^{-1/p} \sum_{i=1}^m \chi_{[ai,ai+1]}$, where $a = |I| + 2$, and $m > n$. Consider $g = \sum_{j=1}^n \alpha_j \phi_{c_j} \in \Sigma_n(\mathcal{D})$. Then g vanishes outside $S = \cup_{j=1}^n (I + c_j)$. By definition of a, $[ai, ai + 1] \cap S = \emptyset$ for at least $m - n$ values of i. Therefore, $\|f - g\| \geq ((m - n)/m)^{1/p}$, which is what we need.

(2) By Zalik [67, Theorem 2], **span**$[\mathcal{D}]$ is dense in $C_0(\mathbb{R})$, hence also in $L_p(\mathbb{R})$ $(0 \leq p < \infty)$. A Bernstein-type inequality from [26] shows that, for any $f \in \Sigma_n(\mathcal{D})$, we have $\|f'\|_p \leq cn^{1/2}\|f\|_p$, for any $p \in (0, \infty]$. An application of Theorem 5.4 completes the proof. $\square$

Finally we consider approximation schemes in tensor products and operator ideals. Suppose X and Y are Banach spaces. A *cross-norm* α on the algebraic tensor product $X \otimes Y$ is a norm satisfying $\|x \otimes y\| = \|x\| \, \|y\|$ for any $x \in X$ and $y \in Y$. The completion of $X \otimes Y$ with respect to this norm is denoted by $X \otimes_\alpha Y$ (this is a Banach space). The reader is referred to e.g. [21,53,64] for information about tensor norms.

Proposition 6.22. *Suppose X and Y are infinite dimensional Banach spaces, and α is a cross-norm. Denote by A_n the set of sums $\sum_{j=1}^n x_j \otimes y_j$ in $Z = X \otimes_\alpha Y$. Then (A_n) is an approximation scheme in Z, satisfying Shapiro's Theorem.*

Proof. Obviously, $A_n = \Sigma_n(\mathcal{D})$, where $\mathcal{D} = X \otimes Y$ is, by definition of Z, a dense subset of Z. To tackle Shapiro's Theorem, we show that (A_n) has Property (P). Fix n. Find unit vectors $(x_i)_{i=1}^n$ and $(y_i)_{i=1}^n$ in X and Y, respectively, forming Auerbach bases in their respective linear spans. That is, for any scalars $\gamma_1, \ldots, \gamma_n$,

$$\max |\gamma_k| \leq \min \left\{ \left\| \sum \gamma_k x_k \right\|, \left\| \sum \gamma_k y_k \right\| \right\} \leq \sum |\gamma_k|.$$

Let $z = \sum_{k=1}^n x_k \otimes y_k/n$. Then $\|z\| \leq 1$. We show that $E(z, A_{n-1}) \geq 1/n^2$.

Suppose, for the sake of contradiction, that $\|z - c\| < 1/n^2$ for some $c = \sum_{j=1}^{n-1} a_j \otimes b_j \in A_{n-1}$. By Hahn–Banach Extension Theorem, there exist norm one linear functionals (f_k) and (g_k) in X^* and Y^*, respectively, which are biorthogonal to (x_k) and (y_k). For $1 \leq p, q \leq n$, $|\langle f_p \otimes g_q, z\rangle - \langle f_p \otimes g_q, c\rangle| < 1/n^2$. However, $\langle f_p \otimes g_q, z\rangle = \delta_{pq}/n$, hence $(\langle f_p \otimes g_q, z\rangle)_{p,q=1}^n = I/n$, where I is the $n \times n$ identity matrix. On the other hand, the matrix $d = (\langle f_p \otimes g_q, c\rangle)_{p,q=1}^n$ has rank less than n. Indeed, for each j, the rank of $(\langle f_p, a_j\rangle\langle g_q, b_j\rangle)_{p,q}$ does not exceed 1. As $d = \sum_{j=1}^{n-1}(\langle f_p, a_j\rangle\langle g_q, b_j\rangle)_{p,q}$, we conclude that **rank**$d < n$.

Now equip the space of $n \times n$ matrices with the Hilbert–Schmidt norm. It is well known that, for any matrix A of rank less than n, $\|I - A\|_{HS} \geq 1$. On the other hand, $\|I/n - d\|_{HS}^2 = \sum_{p,q=1}^{n} |\delta_{pq} - d_{pq}|^2 < n^2 \cdot 1/n^4 = 1/n^2$, a contradiction. $\square$

Now suppose $\mathcal{A}$ is a quasi-Banach operator ideal, equipped with the norm $\|\cdot\|_{\mathcal{A}}$ (see [21,53,64] for the definition and basic properties of operator ideals). Define the $\mathcal{A}$-*approximation numbers* by setting

$$a_n^{(\mathcal{A})}(T) = \inf_{u \in B(X,Y), \mathbf{rank}\, u < n} \|T - u\|_{\mathcal{A}}.$$

Denote by $A^{(\mathcal{A})}(X, Y)$ the set of $\mathcal{A}$-approximable operators – that is, the operators T for which $\lim_n a_n^{(\mathcal{A})}(T) = 0$ – equipped with the norm $\|\cdot\|_{\mathcal{A}}$. One can easily see this is a quasi-Banach space (a Banach space if $\mathcal{A}$ is a Banach ideal). Note that, if $\mathcal{A}$ is the ideal of bounded operators (or compact operators), with its canonical norm, we obtain the usual definitions of approximation numbers, and approximable operators, respectively.

Let $\mathcal{D}$ be the dictionary of rank 1 vectors in $A^{(\mathcal{A})}(X, Y)$, where X and Y are infinite dimensional Banach spaces. For any $T \in B(X, Y)$, we have $E(T, \Sigma_i(\mathcal{D})) = a_{i+1}^{(\mathcal{A})}(T)$.

Corollary 6.23. *In the above notation, the approximation scheme $(\Sigma_i(\mathcal{D}))$ satisfies Shapiro's Theorem.*

Proof. It is well known that $A^{(\mathcal{A})}(X, Y)$ can be identified with $X^* \otimes_\alpha Y$, for the appropriate cross-norm α. An application of Proposition 6.22 completes the proof. $\square$

For certain ideals $\mathcal{A}$, this theorem can be strengthened: it is possible to construct $T \in B(X, Y)$ for which the sequence $(a_n^{(\mathcal{A})}(T))$ "behaves like" a prescribed sequence (α_n). This result appears in the paper [48] of the second author.

7. Controlling the rate of approximation

In the previous sections of this paper, we proved that, for a number of approximation schemes (A_n), we can find and element x in the ambient space, for which the sequence $(E(x, A_n))$ decreases arbitrarily slowly. In some situations, we can go further and guarantee a prescribed behavior of $(E(x, A_n))$.

Recall that an approximation scheme (X_n) in a Banach space X is called *linear* if the sets X_n are linear subspaces of X. By a classical result of Bernstein (see Section 1), if all the X_n's are finite dimensional and $\{\varepsilon_n\} \searrow 0$, then there exists $x \in X$ such that $E(x, X_n) = \varepsilon_n$ for every $n \geq 0$. Without the finite dimensionality assumption, things are different. It was shown in [46] (see also [57, Section I.6.3]) that a Banach space X is reflexive if and only if for any finite sequence of closed subspaces $\{0\} = X_0 \subsetneq X_1 \subsetneq X_2 \ldots \subsetneq X_n \subsetneq X_{n+1} \subset X$, and for any $\varepsilon_0 \geq \varepsilon_1 \geq \cdots \geq \varepsilon_n \geq 0$, there exists $x \in X_{n+1}$ such that $E(x, X_k) = \varepsilon_k$ for any $0 \leq k \leq n$. An inspection of the proof shows the following:

Proposition 7.1. *Suppose X is a Banach space, $\{0\} = X_0 \subsetneq X_1 \subsetneq X_2 \cdots \subsetneq X_n \subsetneq X_{n+1} \subset X$ is a sequence of its closed subspaces, and $\varepsilon_0 > \varepsilon_1 > \cdots > \varepsilon_n \geq 0$. Then there exists $x \in X_{n+1}$ such that $E(x, X_k) = \varepsilon_k$ for any $0 \leq k \leq n$.*

If the chain of subspaces (X_n) is infinite, we obtain a somewhat weaker result.

Theorem 7.2. *Let $\{0\} = X_0 \subsetneq X_1 \subsetneq X_2 \dots$ be a sequence of closed subspaces of a Banach space X. Then for every $\{\varepsilon_n\} \searrow 0$ and every $\{\delta_n\} \searrow 0$ there are $x \in X$, $C > 0$ and $\{n(m)\}_{m=0}^{\infty}$ sequence of natural numbers verifying $n(m) \geq m$ for all m, such that*

$$\delta_m \varepsilon_{n(m)} \leq E(x, X_{n(m)}) \leq C\varepsilon_{n(m)}, \quad m = 0, 1, 2, \dots.$$

Moreover, there exists an strictly increasing sequence of natural numbers $\{h(m)\}_{m=0}^{\infty}$ such that

$$\delta_{h(m)} \varepsilon_{h(m)} \leq E(x, X_{h(m)}) \leq C\varepsilon_{h(m)}, \quad m = 0, 1, 2, \dots.$$

Proof. Assume, without loss of generality, that $\varepsilon_0 = 1$. Define $A_0 = A_0(\{\varepsilon_n\})$ as the set $\{x \in X : \{\frac{E(x,X_n)}{\varepsilon_n}\}_{n=0}^{\infty} \in c_0\}$, equipped with the norm $\|x\|_{A_0} = \sup_{n\in\mathbb{N}} \frac{E(x,X_n)}{\varepsilon_n}$ (see [4, Prop. 3.8, Theorems 3.12 and 3.17]). It is easy to see that A_0 is a Banach space, and the natural embedding of A_0 into X is contractive. We claim that, for any $x \in X$ and $m \geq 0$,

$$E(x, X_m)_{A_0(\{\varepsilon\})} = \sup_{n\geq m} \frac{E(x, X_n)}{\varepsilon_n}. \tag{7.1}$$

Indeed,

$$E(x, X_m)_{A_0(\{\varepsilon\})} = \inf_{y\in X_m} \|x - y\|_{A_0} = \inf_{y\in X_m} \sup_{n} \frac{E(x - y, X_n)}{\varepsilon_n}. \tag{7.2}$$

For $n < m$, we trivially have $E(x - y, X_n) \leq \|x - y\|$. For $m \geq n$, $E(x - y, X_n) = E(x, X_n)$. Taking the infimum over $y \in X_m$ in (7.2), and recalling that $\varepsilon_0 \geq \varepsilon_1 \geq \cdots$, we obtain (7.1).

Note that $\lim_m \sup_{n\geq m} \varepsilon_n^{-1} E(x, X_n) = 0$, hence (X_n) is an approximation scheme in A_0. Moreover, this scheme is non-trivial: for each n, the inclusion of $\overline{X_n}^{A_0}$ into A_0 is strict. Thus, $(A_0, \{X_n\})$ satisfies Shapiro's Theorem. By Corollary 3.7, for every $\{\delta_n\} \searrow 0$ there exists $x \in A_0(\varepsilon_n)$ such that

$$\sup_{n\geq m} \frac{E(x, X_n)}{\varepsilon_n} = E(x, X_m)_{A_0(\varepsilon_n)} \geq 2\delta_m \quad (m = 0, 1, 2, \dots).$$

In other words, for every $m \in \mathbb{N}$ there exists $n(m) \geq m$ such that $E(x, X_{n(m)}) \geq \delta_m \varepsilon_{n(m)}$. Taking $C = \|x\|_{A_0}$, we establish the first claim of this theorem. To prove the second claim, it is enough to take a strictly increasing subsequence $h(m)$ of $n(m)$, and to recall that $\{\delta_m\}$ is decreasing. $\square$

Recall the *density sequence* $\mathbf{d}_i = E(S(X), A_i)$, defined in Section 4. There, it was observed that $\{\mathbf{d}_i\}_{i=0}^{\infty}$ is non-increasing, and $(X, \{A_n\})$ satisfies Shapiro's Theorem if and only if $\mathbf{d}_i = 1$ for every i. The following result is a "mirror image" of Brudnyi's Theorem.

Theorem 7.3. *Suppose that $\{\varepsilon_i\}$ is a sequence of positive numbers converging to 0, $(X, \{A_n\})$ is an approximation scheme in a Banach space X, and $\mathbf{d}_i > 0$ for $i = 0, 1, \dots$. Then there exists $x \in X \setminus (\cup_i A_i)$ such that $0 < E(x, A_i) \leq \varepsilon_i$ for each i.*

Lemma 7.4. *Suppose $(X, \{A_n\})$ is an approximation scheme in a Banach space X. Suppose, furthermore, that $i \in \mathbb{N}$ satisfies $\mathbf{d}_{K(i)} > 0$. Then for any $c \in (0, 1)$ there exist $j > i$ and $y_0 \in A_j$, such that $\|y_0\| = 1$, and $E(x + \alpha y_0, A_i) > c|\alpha|\mathbf{d}_{K(i)}$ for any $x \in A_i$, and any scalar α.*

Proof. As $S(X) \bigcap (\bigcup_j A_j)$ is dense in $S(X)$, we can find $j \in \mathbb{N}$ and $y_0 \in A_j \cap S(X)$ in such a way that $E(y_0, A_{K(i)}) > c\mathbf{d}_{K(i)}$. Then, for any $x, z \in A_i$,

$$\|(x + \alpha y_0) - z\| = \|\alpha y_0 - (z - x)\| \geq |\alpha| E(y_0, A_{K(i)}) > c|\alpha|\mathbf{d}_{K(i)},$$

which is what we wanted to prove. $\square$

Proof of Theorem 7.3. We are going to find a "rapidly increasing" sequence $0 = i_0 < 1 = i_1 < i_2 < i_3 < \cdots$, a "rapidly decreasing" sequence $\delta_1 > \delta_2 > \cdots > 0$, and a sequence of elements $x_j \in A_{i_j}$, in such a way that the following holds for every j:

$$\begin{aligned} &\delta_j \le \min\{\varepsilon_{i_j}/2, \delta_{j-1}\mathbf{d}_{K(i_{j-2})}/4\}, \qquad \|x_j - x_{j-1}\| \le \delta_j, \\ &E(x_j, A_{i_{j-1}}) > 4\delta_j \mathbf{d}_{K(i_{j-1})}/5. \end{aligned} \tag{7.3}$$

As $\delta_j \le \delta_{j-1}/4$, $\{x_j\}$ is a Cauchy sequence in X. Let $x = \lim_j x_j$. We claim that $x \notin \cup_i A_i$ and $E(x, A_\ell) < \varepsilon_\ell$ for each ℓ. Indeed, for $i_{j-1} \le \ell < i_j$,

$$E(x, A_\ell) \le E(x, A_{i_{j-1}}) \le \sum_{k \ge j} \delta_k \le \delta_j \sum_{s=0}^{\infty} 4^{-s} < 2\delta_j \le \varepsilon_{i_j} \le \varepsilon_\ell.$$

On the other hand,

$$E(x, A_{i_j}) \ge \frac{4\mathbf{d}_{K(i_{j-1})}\delta_j}{5} - \sum_{k>j} \delta_k \ge \frac{4\mathbf{d}_{K(i_{j-1})}\delta_j}{5} - \mathbf{d}_{K(i_{j-1})}\delta_j \sum_{s=1}^{\infty} 4^{-s} > \frac{\mathbf{d}_{K(i_{j-1})}\delta_j}{3} > 0.$$

Thus, it suffices to show the existence of the sequences $\{i_j\}$, $\{x_j\}$, and $\{\delta_j\}$ with desired properties. Set $x_0 = 0$. Let $\delta_1 = \varepsilon_1/2$, and pick an arbitrary $x_1 \in A_1$ with $\|x_1\| = \delta_1$. Now suppose $x_j \in A_{i_j}$, $\delta_j > 0$, and $n_j \in \mathbb{N}$ have been defined for $j < k$, in such a way that (7.3) are satisfied. By Lemma 7.4, we can find s such that there exists $y \in A_s$ with $\|y\| = 1$, for which $E(x_{k-1} + \delta y, A_{i_{k-1}}) > 4\delta \mathbf{d}_{K(i_{k-1})}/5$ hold for any $\delta > 0$. Set $i_k = K(s)$, and $\delta_k = \min\{\varepsilon_{i_k}, \delta_{k-1}\mathbf{d}_{K(i_{k-2})}/4\}$. Then $x_k = x_{k-1} + \delta_k y \in A_{i_k}$, $\|x_k - x_{j-k}\| = \delta_k$, and $E(x_k, A_{i_{k-1}}) > 4\mathbf{d}_{K(i_{k-1})}\delta_k/5$. $\square$

Acknowledgment

We would like to thank the referee for his/her help in improving the paper, and in particular, for bringing several relevant references (such as [18,28]) to our attention.

References

[1] J.M. Almira, N. Del Toro, On negative results in approximation theory, Commun. Appl. Anal. 8 (2004) 237–250.
[2] J.M. Almira, N. Del Toro, Some remarks on negative results in approximation theory, Rend. Circ. Mat. Palermo (2) (Suppl. 68) (2002) 245–256. (Part I); Proceedings of the Fourth International Conference on Functional Analysis and Approximation Theory, Vol. I (Potenza, 2000).
[3] J.M. Almira, U. Luther, Compactness and generalized approximation spaces, Numer. Funct. Anal. Optim. 23 (2002) 1–38.
[4] J.M. Almira, U. Luther, Generalized approximation spaces and applications, Math. Nachr. 263–264 (2004) 3–35.
[5] J.M. Almira, T. Oikhberg, Shapiro's Theorem for subspaces, J. Math. Anal. Appl. 388 (1) (2012) 282–302.
[6] J.M. Anderson, J. Clunie, Isomorphisms of the disc algebra and inverse Faber sets, Math. Z. 188 (1985) 545–558.
[7] T. Bartoszynski, M. Dzamonja, L. Halbeisen, E. Murtinova, A. Plichko, On bases in Banach spaces, Studia Math. 170 (2005) 147–171.
[8] S.N. Bernstein, Sur le probleme inverse de la théorie de la meilleure approximation des functions continues, Comptes Rendus 206 (1938) 1520–1523. (See also: Ob obratnoi zadache teorii nailuchshego priblizheniya nepreryvnykh funksii, Sochineniya, vol. II (1938), 292–294).
[9] P.A. Borodin, On the existence of an element with given deviations from an expanding system of subspaces, Math. Notes 80 (5) (2006) 621–630. (Translated from Matematicheskie Zametki, 80 (5) (2006), 657–667).
[10] P. Borwein, T. Erdelyi, Polynomials and Polynomial Inequalities, in: Graduate Texts in Mathematics, Springer, 1995.
[11] D.G. Bourgin, Linear topological spaces, Amer. J. Math. 65 (1943) 637–659.

[12] Yu.A. Brudnyi, N.Ya. Krugljak, Interpolation Functors and Interpolation Spaces, Vol. I, vol. 47, North-Holland Math. Libr., 1991.
[13] P. Casazza, The art of frame theory, Taiwanese J. Math. 4 (2000) 129–201.
[14] P. Casazza, Characterizing Hilbert space frames with the subframe property, Illinois J. Math. 41 (1997) 648–666.
[15] E.W. Cheney, Introduction to Approximation Theory, second ed., Chelsea, New York, 1982.
[16] E.W. Cheney, W. Light, A Course in Approximation Theory, Amer. Math. Soc., Providence, RI, 2009.
[17] J.H. Curtiss, Faber polynomials and the Faber series, Amer. Math. Monthly 78 (1971) 577–596.
[18] F. Deutsch, H. Hundal, Slow convergence of sequences of linear operators I: almost arbitrarily slow convergence, J. Approx. Theory 162 (2010) 1701–1716.
[19] F. Deutsch, H. Hundal, Slow convergence of sequences of linear operators II: arbitrarily slow convergence, J. Approx. Theory 162 (2010) 1717–1738.
[20] R.A. DeVore, G.G. Lorentz, Constructive Approximation, Springer, 1993.
[21] J. Diestel, H. Jarschow, A. Tonge, Absolutely Summing Operators, Cambridge University Press, Cambridge, 1995.
[22] E.P. Dolzhenko, Comparison of the rates of rational and polynomial approximations, Math. Notes 1 (3) (1967) 313–320 (in Russian).
[23] E.P. Dolzhenko, Bounds for derivatives of rational functions, Izv. Akad. Nauk SSSR 27 (1963) 9–28.
[24] L. Drewnoswki, Quasi-complements in F-spaces, Studia Math. 77 (1984) 373–391.
[25] R.E. Edwards, Fourier Series. A Modern Introduction, vol. I, Springer, 1979.
[26] T. Erdelyi, Sharp Bernstein-type inequalities for linear combinations of shifted Gaussians, Bull. Lond. Math. Soc. 38 (2006) 124–138.
[27] D. Gaier, Lectures on Complex Approximation, Birkhäuser, Boston, 1987.
[28] M. Golomb, Lectures on Theory of Approximation, Argonne National Laboratory, Applied Mathematics Division, 1962.
[29] R. Gribonval, M. Nielsen, Nonlinear approximation with dictionaries. II. Direct estimates, Constr. Approx. 24 (2004) 157–173.
[30] P. Hajek, V. Montesinos, J. Vanderwerff, V. Zizler, Biorthogonal Systems in Banach Spaces, Springer-Verlag, New York, 2008.
[31] D. Han, D. Larson, Frames, Bases and Group Representations, vol. 147, Mem. Amer. Math. Soc., 2000, No. 697.
[32] D.D. Haroske, C. Schneider, Besov spaces with positive smoothness on $\mathbb{R}^n$, embeddings and growth envelopes, J. Approx. Theory 161 (2009) 723–747.
[33] D.H. Hyers, A note on linear topological spaces, Bull. Amer. Math. Soc. 44 (1938) 76–80.
[34] N.J. Kalton, Quasi-banach spaces, in: Handbook of the Geometry of Banach Spaces, vol. 2, North-Holland, Amsterdam, 2003, pp. 1099–1130.
[35] N.J. Kalton, N.T. Peck, J.W. Roberts, An F-Space Sampler, in: London Math. Soc. Lecture Note Series, vol. 89, Cambridge University Press, 1984.
[36] A. Levin, Approximation by rational functions in the complex domain, Math. Notes 9 (1971) 72–77.
[37] G. Lewicki, Bernstein's Lethargy theorem in metrizable topological linear spaces, Monatsh. Math. 113 (1992) 213–226.
[38] G. Lewicki, A theorem of Bernstein's type for linear projections, Iagel. Acta Math. 27 (1988) 23–27.
[39] J. Lindenstrauss, L. Tzafriri, Classical Banach Spaces I, Springer, 1977.
[40] R. Liu, B. Zheng, A characterization of Schauder frames which are near-Schauder bases, J. Fourier Anal. Appl. 16 (2010) 791–803.
[41] E. Livshitz, On a greedy algorithm in the space $L_p[0, 1]$, Math. Notes 85 (2009) 751–754.
[42] G.G. Lorentz, Y. Makovoz, M. von Golitschek, Constructive Approximation. Advanced Problems, Springer, 1996.
[43] D. Lubinsky, A survey of weighted polynomial approximation with exponential weights, Surv. Approx. Theory 3 (2007) 1–105.
[44] B. Micherda, Bernstein's "Lethargy" theorems in SF-spaces, Z. Anal. Anwend. 22 (2003) 3–16.
[45] M.A. Nazarenko, Existence of a function with given rational approximations in the space CA, Moscow Univ. Math. Bull. 52 (1997) 24–26.
[46] W.N. Nikolskii, On some properties of reflexive spaces, Uchen. Zap. Kalinin. Gos. Ped. Inst. 29 (1963) 121–125 (in Russian).
[47] W.N. Nikolskii, Some remarks on spaces with (B) property, Uchen. Zap. Kalinin. Gos. Ped. Inst. 39 (1964) 48–52 (in Russian).
[48] T. Oikhberg, Rate of decay of s-numbers, J. Approx. Theory 163 (3) (2011) 311–327.
[49] A.A. Pekarskii, Existence of a function with preassigned best uniform rational approximations, Izv. Akad. Nauk Belarusi, Ser. Fiz.-Mat. Nauk 1 (1994) 23–26 (in Russian).

[50] A.A. Pekarskii, Estimates of derivatives of rational functions in $L_p(-1, 1)$, Mat. Zametki 392 (1986) 388–394.
[51] A. Pietsch, Approximation spaces, J. Approx. Theory 32 (1981) 115–134.
[52] A. Pietsch, History of Banach Spaces and Linear Operators, Birkhäuser, Boston, 2007.
[53] A. Pietsch, Operator Ideals, North-Holland, Amsterdam, 1980.
[54] W. Plesniak, Quasianalytic functions in the sense of Bernstein, Dissertationes Math. 147 (1977) 1–70.
[55] H. Rosenthal, The Banach space $C(K)$, in: W. Johnson, Y. Lindenstrauss (Eds.), Handbook of Banach Spaces, vol. II, North-Holland, Amsterdam, 2003, pp. 1457–1602.
[56] H.S. Shapiro, Some negative theorems of approximation theory, Michigan Math. J. 11 (1964) 211–217.
[57] I. Singer, Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces, Springer Verlag, New York, 1970.
[58] A.P. Starovoitov, On the problem of describing sequences of best trigonometric rational approximations, Math. Notes 69 (2001) 839–844. (Translated from Matematicheskie Zameti, 69 (6) (2001), 919–924).
[59] A.P. Starovoitov, Coincidence of least uniform deviations of functions from polynomials and rational functions, Math. Notes 74 (2003) 578–582.
[60] P.K. Suetin, Fundamental properties of Faber polynomials, Russian Math. Surveys 19 (1964) 121–149.
[61] A.F. Timan, Theory of Approximation of Functions of a Real Variable, MacMillan Co., New York, 1963.
[62] I.S. Tjuriemskih, B property of Hilbert spaces, Uchen. Zap. Kalinin. Gos. Ped. Inst. 39 (1964) 53–64 (in Russian).
[63] I.S. Tjuriemskih, On a problem of S.N. Bernstein, Uchen. Zap. Kalinin. Gos. Ped. Inst. 52 (1967) 123–129 (in Russian).
[64] N. Tomczak-Jaegermann, Banach–Mazur Distances and Finite-Dimensional Operator Ideals, Longman Scientific and Technical, Harlow, 1989.
[65] N. Wiener, The Fourier Integral and Certain of its Applications, Dover, New York, 1958.
[66] P. Wojtaszczyk, A Mathematical Introduction to Wavelets, Cambridge University Press, 1997.
[67] R. Zalik, The Müntz–Szasz theorem and the closure of translates, J. Math. Anal. Appl. 82 (1981) 361–369.