Technische Universität München
Zentrum Mathematik

Characterizing Discrete-Time Function Spaces

Marie Wild

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Michael Ulbrich
Prüfer der Dissertation:	1.	Priv.-Doz. Dr. Hartmut Führ
	2.	Univ.-Prof. Dr. Rupert Lasser
	3.	Prof. Dr. Rodolfo H. Torres, University of Kansas / USA (schriftliche Beurteilung)

Die Dissertation wurde am 30. März 2006 bei der Technischen Universität eingereicht und durch die Fakultät für Mathematik am 17. Juli 2006 angenommen.

Preface

This thesis was carried out at the GSF Institute of Biomathematics and Biometry (IBB), Neuherberg, under the supervision of PD Dr. Hartmut Führ, whom I want to thank first and foremost. Without his help and encouragement, this work would not have been possible.

I want to thank all the colleagues at the IBB, especially the head of the group 'Mathematical Modelling in Ecology and the Biosciences', Prof. Dr. Gerhard Winkler, and the head of the institute, Prof. Dr. Rupert Lasser, for their support.

Last but not least many thanks to my parents and all of my friends –
Dominik, minä rakastan sinua.

Munich, March 2006 Marie Wild

From May 2002 to July 2005, the thesis was funded by the GSF and from August 2005 to March 2006 by the Munich University of Technology. Additionally funding came through the European Research Training Network HASSIP.

Contents

Introduction

This thesis deals with characterizations of spaces of discrete-time functions, with a focus on Besov spaces on $\mathbb{Z}$ and their wavelet characterization.

It is a widely accepted fact that the success of wavelets in applications is based on their ability to efficiently represent 'realistic' signals. This efficiency is twofold: **Computational efficiency** is guaranteed by fast filter bank algorithms associated to a wavelet basis, the so-called **fast wavelet transform**.
An equally important property of wavelets is their **approximation-theoretic efficiency**, that is, the ability of wavelets to capture salient features of a signal in a few large coefficients.

This property of wavelets is best exemplified by means of piecewise polynomial signals: Given a wavelet with sufficiently many vanishing moments, the nonzero wavelet coefficients will be located at the jumps of the signal.

A more elaborate (and more powerful) description of wavelet approximation theory can be formulated in terms of Besov spaces. Despite the fact that Besov spaces were conceived some 25 years prior to wavelets, it is probably fair to say that Besov space theory and wavelet approximation theory are **identical**; see Chapter 1 for an explanation of this statement.

Hence, orthonormal wavelet bases provide a class of signal transforms that are easily implemented, with fast algorithms and completely understood approximation theory, and much use has been made of these features, both for theoretical and applied purposes [5, 7, 15, 19, 24].

However, despite the fact that the computational and approximation-theoretic properties of wavelets are often used simultaneously, one should note that there is a, somewhat subtle, gap separating the two: Strictly speaking, the **computational part only applies to discrete time signals** and their decomposition by the fast wavelet transform, whereas the latter are only applicable to continuous-time signals.

This gap has been acknowledged early on, but not much has been done since to close it. In this thesis, we present a discrete-time version of wavelet approximation theory, that is specifically tuned to the fast wavelet transform. As one might expect, it is again linked to a scale of Besov spaces, previously defined by Torres [31].

To be more precise, let us define for $\alpha \in \mathbb{R}$, $0 < p, q < \infty$, the coefficient spaces $\dot{b}^{\alpha}_{p,q}(\mathbb{R})$ as the collection of all complex-valued sequences $t = (t_{j,l})_{j,l\in\mathbb{Z}}$, satisfying

$$\|t\|_{\dot{b}^{\alpha}_{p,q}(\mathbb{R})} := (\sum_{j\in\mathbb{Z}}(\sum_{l\in\mathbb{Z}}(2^{-j(\alpha+1/2-1/p)}|t_{j,l}|)^p)^{q/p})^{1/q} < \infty.$$

These norms are used to measure the decay of the expansion coefficients of a signal f in a wavelet orthonormal basis. Such a basis is a system $(\psi_{j,l})_{j,l\in\mathbb{Z}}$ of functions arising from a suitable 'mother wavelet' $\psi \in L^2(\mathbb{R})$ by translation and dilation,

$$\psi_{j,l}(x) = 2^{-j/2}\psi(2^{-j}x - l) \quad .$$

A short survey of wavelet bases in $L^2(\mathbb{R})$ and their construction can be found in Section 2.1.

Typically, wavelets fulfill additional desirable properties, besides generating an orthonormal basis, such as

- **smoothness**, i.e. $\psi \in C^M$, for $M \in \mathbb{N}$
- **vanishing moments**: $\int_{\mathbb{R}} \psi(x)x^i dx = 0$, for $i = 0, \ldots, K-1$.
- **compact support**.

It is known that if the wavelet family $(\psi_{j,l})_{j,l\in\mathbb{Z}}$ has the above properties with M and K large enough, then a function f is in a Besov space $\dot{B}^{\alpha}_{p,q}(\mathbb{R})$, if and only if $(\langle f, \psi_{j,l}\rangle)$ is in the corresponding coefficient space $\dot{b}^{\alpha}_{p,q}(\mathbb{R})$; see e.g. [17]. In addition, we have the norm equivalence

$$\|f\|_{\dot{B}^{\alpha}_{p,q}(\mathbb{R})} \asymp \|(\langle f, \psi_{j,l}\rangle)\|_{\dot{b}^{\alpha}_{p,q}(\mathbb{R})} \quad .$$

Early on, these norm equivalences have been related to the nonlinear approximation behavior of wavelet expansions and to wavelet applications: a decay of coefficients like in $\dot{b}^{\alpha}_{p,q}(\mathbb{R})$ is linked to the decay of the approximation error of wavelet expansions by $N > 0$ terms (see [7] or Section 1.2.1 below).

These results are widely used in signal and image processing: A small list of references that use the relationship between wavelets and Besov space to derive algorithms for diverse problems such as denoising, compression, deconvolution or Radon inversion, is [5, 10, 1, 2, 24, 13, 21].

In most applications however, the data under consideration are given discretely, and are processed by the fast wavelet transform. This algorithm arises naturally from a **multiresolution analysis**, which can be associated to most orthonormal wavelet bases (in particular to all smooth wavelets with compact support, see [22]).

Thus, a discrete series $(f(n))_{n\in\mathbb{Z}}$ is mapped to the family of $(d_{j,l})_{j\geq 1,l\in\mathbb{Z}}$ of discrete wavelet coefficients. Observe that by the Fischer-Riesz theorem, each coefficient $d_{j,l}$

is given by the scalar product of f with a suitable **discrete-time wavelet** $h_{j,l}$.

This suggests to consider the space of (truncated) coefficient families, $b^{\alpha}_{p,q}(\mathbb{Z})$, for $\alpha \in \mathbb{R}$, $0 < p, q < \infty$, as the collection of complex-valued sequences $s = (s_{j,l})_{j\geq 1, l\in\mathbb{Z}}$, for which

$$\|s\|_{b^{\alpha}_{p,q}(\mathbb{Z})} := (\sum_{j\geq 1}(\sum_{l\in\mathbb{Z}}(2^{-j(\alpha+1/2-1/p)}|s_{j,l}|)^p)^{q/p})^{1/q} < \infty.$$

This coefficient decay still reflects the non-linear approximation behavior of $(f(n))$. It is therefore natural to ask whether the property $(d_{j,l})_{j\geq 1, l\in\mathbb{Z}} \in b^{\alpha}_{p,q}(\mathbb{Z})$ can be characterized - in a similar satisfactory way as in the continuous case - from properties of the filter bank or, equivalently, the discrete-time wavelet family $(h_{j,l})_{j\geq 1, l\in\mathbb{Z}}$ and of the sequence f.

Somewhat surprisingly, literature so far does not seem to provide a simple answer to this question. Nonetheless, the norm equivalence in the continuous time-case is the basis of heuristics which are applied to the discrete setting, where only the truncated coefficient series are available.

The continuous theory has the following to offer: Let ϕ be the scaling function associated to the multiresolution analysis, and define the continuous-time function $F = \sum_{n\in\mathbb{Z}} f(n)\tau_n\phi$, where $\tau_n\phi$ denotes the translate of ϕ. Then the wavelet coefficients of F coincide with $(d_{j,l})_{j\geq 1, l\in Z}$ for $j \geq 1$, and vanish for scales $j \leq 0$. Hence, assuming sufficient vanishing moments, smoothness and decay of the associated continuous time wavelets,

$$(d_{j,l})_{j\geq 1, l\in\mathbb{Z}} \in b^{\alpha}_{p,q}(\mathbb{Z}) \Leftrightarrow F \in \dot{B}^{\alpha}_{p,q}(\mathbb{R}) \quad .$$

However, F is not easily accessible. The problem is presented by the scaling function ϕ: For many wavelet bases, and in particular for the compactly supported wavelets, the scaling function is only known implicitly, as the result of a limit process.

Hence, membership of F in a Besov space is not easily checked, and the equivalence is almost useless.

We have not been able to locate any result in literature dealing with this problem, the continuous-time setting does not give a conclusive answer to our question.
Hence we arrive at the following (somewhat overstated) conclusion: **An algorithm using the cascade algorithm, but derived from heuristics using the Besov space characteristics in continuous time, is not theoretically justified.**

Summary

The main purpose of this thesis is to give criteria for $(d_{j,l})_{j\geq 1, l\in\mathbb{Z}} \in b^{\alpha}_{p,q}(\mathbb{Z})$, with arguments that do not use any embedding into the continuous-time setting. In

this way, we obtain that the discrete-time wavelets associated to a multiresolution analysis are unconditional bases for a whole family of discrete-time signal spaces. It may not be too surprising that the resulting spaces are again Besov spaces, the discrete Besov spaces $B^{\alpha}_{p,q}(\mathbb{Z})$ introduced by Torres [31].
While this thesis strictly avoids any appeal to the continuous time theory, we will adopt the proof strategy and techniques from the continuous setting, as presented in the book of Frazier, Jawerth and Weiss [17] or, in a slightly more general context, by Kyriazis [14].

Chapter 1 provides a short introduction to Besov spaces on $\mathbb{R}$, and their various characterizations, using moduli of smoothness (Section 1.1.1), Littlewood-Paley theory (Section 1.1.2), φ-transforms (Section 1.1.3) or wavelets (Section 1.1.4). This is on the one hand intended as an introduction, providing a first glimpse on techniques that are used and adapted later on, and showing the versatility of Besov spaces. But it is also a precursor of the discrete-time results we establish later on: Analogs to all of these characterizations will be established in Chapter 3 for discrete-time signals.

In Section 1.2, we elaborate on our previously made statement that Besov space theory and wavelet approximation theory can be considered as identical, by discussing the relationship between Besov spaces and nonlinear wavelet approximation. With most of the central notions and definitions established, we then describe the aims of this thesis in somewhat greater detail (Section 1.3).

Chapter 2 is devoted to an introduction to wavelets. Even though the focus of this work is on discrete-time wavelets, we start out with a short rundown of the basic facts concerning wavelet orthonormal bases and multiresolution analysis in continuous time. In particular, we explain the origin of the fast wavelet transform from a continuous time multiresolution analysis (Section 2.1.1), and shortly discuss additional desirable properties of wavelets in $L^2(\mathbb{R})$ such as vanishing moments, smoothness and compact support (Section 2.1.2). We shortly comment on biorthogonal wavelets, as all of our results later on can be established without any additional effort for the biorthogonal setting (Section 2.1.3).

We then turn attention to discrete-time wavelets. Discrete-time wavelet systems arise as a byproduct of the fast wavelet transform (Section 2.2.1), but can be understood as systems of oscillatory building blocks indexed by translation parameters and (dyadic) scale parameters (Theorem 2.2.2). Of crucial importance for Chapter 3 will be additional properties of the wavelet system, such as vanishing moments, support properties and – somewhat unexpectedly – regularity (Section 2.2.2). This latter property has been introduced by Rioul [26]; it connects discrete-time wavelets to continuous-time wavelets via large scale limits (Section 2.2.3).

Chapter 3 finally deals with the central purpose of this thesis, namely the characterization of Besov spaces in discrete time. For the definition of these spaces, the starting point is the so-called φ-transform characterization of $B^{\alpha}_{p,q}(\mathbb{Z})$, established by Torres. Just as in the continuous case, this transform is in many ways quite similar to a wavelet transform, and this similarity allows a proof of the main result by studying off-diagonal decay of certain infinite matrices. This decay behavior is derived from suitable conditions regarding support, vanishing moments and smoothness of the discrete-time wavelets as provided by Section 2.2.2.

Chapter 1 showed that the Besov spaces on $\mathbb{R}$ have quite a number of different, but equivalent descriptions. Using our wavelet characterization, we can prove that the same is valid for their counterparts on $\mathbb{Z}$. We obtain further descriptions of these spaces, which can be viewed as more 'intrinsic' in contrast to the Littlewood-Paley type definition in [31], such as in terms of iterated differences and mean oscillation properties (Section 3.4).

Mimicking the definition of moduli of smoothness for continuous-time functions, we define for $1 < p < \infty$, $t \in \mathbb{R}_+$, the *r-th order modulus of smoothness of f in $l^p(\mathbb{Z})$* by

$$\omega_p^r(f,t) = \sup_{m \in \mathbb{Z}, |m| < t} \|\Delta_m^r f(\cdot)\|_p,$$

where $\Delta_m^r f(n)$ is the difference operator $\Delta_m f(n) = f(n+m) - f(n)$ iterated r times. For $\alpha > 0, 1 < p, q < \infty$, $r = \lfloor \alpha \rfloor + 1$, we define the space $B_q^\alpha(l^p(\mathbb{Z}))$ as the space of sequences $(f(n))$, such that

$$\|f\|_{B_q^\alpha(l^p(\mathbb{Z}))} := (\sum_{j \geq 1} (2^{-j\alpha} \omega_p^r(f, 2^j))^q)^{1/q} < \infty$$

We show that for $\alpha > 0, 1 < p, q < \infty$ the spaces $B_q^\alpha(l^p(\mathbb{Z}))$ coincide with the $B_{p,q}^\alpha(\mathbb{Z})$ with equivalent norms (Theorem 3.4.6).

Another characterization for the discrete-time Besov spaces we get in a straightforward manner is a description in terms of mean oscillation over intervals (Theorem 3.4.9).

For continuous-time Besov spaces, this type of characterization was given in [9] and in [8]. [30] contains a description of discrete-time in terms of mean oscillation properties of sequences for some special cases of the parameters α, p, q.
The results of Section 3.4 exhibit the usefulness of our wavelet characterization in Theorem 3.3.8, and they emphasize the view of discrete-time Besov spaces as worthy analogs of their continuous-time counterparts.

We also treat another scale of function spaces in this thesis: in **Chapter 4**, we present a discrete-time wavelet characterization for the discrete-time Triebel-Lizorkin $F_{p,q}^\alpha(\mathbb{Z})$ spaces analog to our result for $B_{p,q}^\alpha(\mathbb{Z})$. These spaces were already discussed by Q. Sun in [29] in terms of smooth atomic decompositions. Theorem 4.2.6 provides a new characterization of these spaces, not contained in Sun's results.

Preliminaries

This section is mainly thought to fix notation. It also provides a short review of bases for Hilbert and Banach spaces, but this is kept short and is just meant to to provide the language we use later on in this thesis.

For $1 \leq p < \infty$, the space $L^p(\mathbb{R})$ is defined as the space of measurable functions f on $\mathbb{R}$, for which

$$\|f\|_{L^p(\mathbb{R})} = \left(\int_{\mathbb{R}} |f(t)|^p \right)^{1/p} < \infty.$$

Equipped with the above norm, $L^p(\mathbb{R})$ is a Banach space, provided that almost everywhere agreeing functions are identified. For $p = \infty$, the integral above is replaced by ess sup in the usual way.
$\ell^p(\mathbb{Z})$ denotes the corresponding p-summable sequence space.

The **Fourier transform** on $L^1(\mathbb{R})$ is defined by

$$\widehat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-i\omega t} dt \quad ,$$

and is extended to the Plancherel transform on $L^2(\mathbb{R})$. With the chosen normalization, one has

$$\|\widehat{f}\|_{L^2(\mathbb{R})} = \frac{1}{2\pi} \|f\|_{L^2(\mathbb{R})}.$$

$\mathcal{S}(\mathbb{R})$ denotes the **Schwartz space of rapidly decreasing functions**, and $\mathcal{S}'(\mathbb{R})$ its dual, the space of **tempered distributions**.
By $\langle f, g \rangle$, we denote the standard inner product in a Hilbert space, as well as the pairing of a tempered distribution with a Schwartz function $\overline{g}$.

The **involution** for a sequence reads as $g^*(n) = \overline{g(-n)}$, where $\overline{g}$ stands for complex conjugation.

The **convolution** product of $f \in \mathcal{S}'(\mathbb{R})$ with a Schwartz function $h \in \mathcal{S}(\mathbb{R})$ is defined by

$$\langle f * h, g \rangle = \langle f, \overline{h^*} * g \rangle,$$

where for $h, g \in \mathcal{S}(\mathbb{R})$,

$$h * g(\cdot) = \int h(t) g(\cdot - t) dt.$$

The convolution product of two sequences g, h is defined by

$$g * h(n) = \sum_k g(k) h(n - k).$$

For a function f defined on $\mathbb{R}$, the **support** of f is defined as $\operatorname{supp} f = \overline{\{t : f(t) \neq 0\}}$. For a sequence $g = (g(n))_{n \in \mathbb{Z}}$, supp g will be the smallest interval containing those n for which $h(n) \neq 0$.

For $N \in \mathbb{N}$, the **upsampling operator** acting on a sequence g is defined by $\uparrow_N g(n) = g(N^{-1}n)$, if $N^{-1}n \in \mathbb{Z}$ and 0 otherwise;
the **downsampling operator** is defined by $\downarrow_N g(n) = g(N\,n)$.

To avoid cluttered notation, C denotes a constant which is allowed to change within an argument. The notion $A \asymp B$ means that there exist constants $C_1, C_2 > 0$, such that $C_1\, A \leq B \leq C_2\, B$.

Bases for Hilbert and Banach spaces

In this thesis, we frequently use different concepts of bases in Hilbert and Banach spaces. We give some of the most elementary definitions and results, taken from [4]. Let in the following $\mathcal{H}$ be a separable Hilbert space and I a countable index set. We assume I to be suitably numbered, thus introducing a summation order on I.

Definition 0.1. *A family of vectors $(e_n)_{n \in I} \subseteq \mathcal{H}$ is an **orthonormal system (ONS)**, if*

$$\langle e_n, e_m \rangle = \delta_{n,m}, \text{ for all } n, m \in I,$$

where $\delta_{n,m} = 1$ for $n = m$ and 0 otherwise.

Definition 0.2. *An ONS $(e_n)_{n \in I} \subseteq \mathcal{H}$ is an **orthonormal basis (ONB)** if it is complete in $\mathcal{H}$:*

$$\mathcal{H} = \overline{\operatorname{span}}\{e_n\}. \tag{0.0.1}$$

Equation (0.0.1) is equivalent to

$$\|f\|_{\mathcal{H}}^2 = \sum_{n \in I} |\langle f, e_n \rangle|^2 \text{ for all } f \in \mathcal{H}.$$

Definition 0.3. *A family of vectors $(f_n)_{n \in I} \subseteq \mathcal{H}$ is a **Riesz Basis** for $\mathcal{H}$, if there exists an ONB $(e_n)_{n \in I}$ for $\mathcal{H}$ and a bounded invertible mapping $T : \mathcal{H} \to \mathcal{H}$, such that $Te_n = f_n$ for all $n \in I$.*

Associated to any Riesz basis is a **dual family**, which is also a **Riesz Basis**.

Theorem 0.4. *If $(f_n)_{n\in I}$ is a Riesz Basis for $\mathcal{H}$, there exists a unique family $(\tilde{f}_n)_{n\in I} \subseteq \mathcal{H}$, such that for every $f \in \mathcal{H}$*

$$f = \sum_{n\in I} \langle f, \tilde{f}_n \rangle f_n.$$

*$(\tilde{f}_n)$ is also a Riesz basis and (f_n), $(\tilde{f}_n)$ are **biorthogonal**, i.e.*

$$\langle f_n, \tilde{f}_m \rangle = \delta_{n,m}, \text{ for all } n, m \in I.$$

In this sense, we call the families (f_n), $(\tilde{f}_n)$ **biorthogonal bases**.
For a Riesz Basis (f_n), there exist constants $A, B > 0$, such that for every $f \in \mathcal{H}$,

$$A \sum_n \|f\|^2_{\mathcal{H}} \leq \sum_n |\langle f, f_n \rangle|^2 \leq B \sum_n \|f\|^2_{\mathcal{H}}.$$

The constants therein are called **Riesz bounds**. Of course, any ONB is a Riesz basis with Riesz bounds $A = B = 1$.
Let us now consider the more general situation in a Banach space $\mathcal{B}$, where the concept of orthogonality is not applicable anymore.

Definition 0.5. *A family of vectors $(e_n)_{n\in I} \subseteq \mathcal{B}$ is a **Schauder basis** if for each $f \in \mathcal{B}$, there exist unique coefficients $(c_n)_{n\in I}$, such that*

$$f = \sum_{n\in I} c_n e_n \tag{0.0.2}$$

with convergence in $\mathcal{B}$.

The above definition depends on the order of summation: it can happen that the sum is divergent for a certain permutation of summands. We thus introduce the notion of unconditional convergence.

Definition 0.6. *A Schauder Basis $(e_n)_{n\in I} \subseteq \mathcal{B}$ is an **unconditional basis**, if the series (0.0.2) converges **unconditionally**, i.e. $\sum_{n\in I} c_{\sigma(n)} e_{\sigma(n)}$ converges for all permutations σ of I. In this case, the limit is the same, regardless of the order of summation.*

The concept of an unconditional basis in a Banach space extends that of a Riesz basis in a Hilbert space $\mathcal{H}$. In fact, any Riesz basis for $\mathcal{H}$ is an unconditional basis for $\mathcal{H}$.

Chapter 1

Besov Spaces on $\mathbb{R}$

As stated in the introduction, this thesis deals with characterizations of discrete-time function spaces with a focus on discrete-time Besov spaces.

In this chapter, we give an overview of their continuous-time counterparts, the Besov spaces on $\mathbb{R}$. Roughly spoken, the Besov class can be viewed as a generalization of classical smoothness spaces, such as Hölder or Sobolev spaces, to spaces of functions and distributions possessing smoothness of order $\alpha \in \mathbb{R}$, measured in different L^p spaces. A third parameter q allows finer distinctions.

In literature, a multitude of different but equivalent definitions of Besov spaces can be found, their description depending on the devices which are used to measure smoothness. In the first section, we recall some of these characterizations.

Furthermore, Besov spaces are related to nonlinear approximation behavior of wavelet expansions. Simplistically, approximating a function which is contained in a Besov space of order α by N terms of a wavelet series, the approximation error decreases in $O(N^{-\alpha})$. This issue, which in fact could be regarded as another characterization for certain Besov spaces, is described more precisely in Section 1.2 of this chapter.

In the next section, we switch from continuous to discrete time: In wavelet applications such as signal or image processing, the data under consideration are given discretely. Passing such a discrete-time function through a wavelet filter bank, decay properties of the arising coefficients still reflect nonlinear approximation properties of the function, whereas the continuous-time theory, though serving as a basis of heuristics, does not provide a satisfactory characterization of these signals via function spaces.

So, on the one hand, this chapter serves to motivate the study of discrete-time Besov spaces and to specify the aims of this thesis in Section 1.3. On the other hand, it is also meant as a short overview on continuous-time Besov spaces and a quick guided tour through the multitude of their different descriptions, which are spread among literature.

1.1 Besov Spaces on $\mathbb{R}$ and their Characterizations

There is no unique way to define Besov spaces. There is a large variety of descriptions, which are essentially equivalent to each other. This property of Besov spaces can be the source of confusion, in particular as notations and normalizations tend to vary between different sources, but it also reflects their status as an important class of function spaces, located at the intersection of various mathematical subdisciplines.

The characterizations given here can be grouped roughly into two categories. The first class of approaches characterizes functions and distributions by their smoothness in terms of derivatives or differences, whereas the other way to measure smoothness presented here uses Fourier analytical devices.

To clarify these notions, we start by recalling a class of well-known smoothness spaces, the class of Sobolev spaces in $L^2(\mathbb{R})$. We will introduce two types of spaces: the *homogeneous* and the *inhomogeneous* spaces.

Definition 1.1.1. *Let $k \in \mathbb{N}$. The* ***homogeneous Sobolev space*** $\dot{W}_2^k(\mathbb{R})$ *is defined as the space of tempered distributions f on $\mathbb{R}$, for which the (distributional) derivative of order k, $D^k f$, is in $L^2(\mathbb{R})$.*
For $f \in \mathcal{S}'(\mathbb{R})$, $|f|_{\dot{W}_2^k(\mathbb{R})} := \|D^k f\|_{L^2(\mathbb{R})}$ defines a semi-norm on $\dot{W}_2^k(\mathbb{R})$.

Note that these semi-norms are not norms in general: $|f|_{\dot{W}_k^2(\mathbb{R})} = 0$ for f a polynomial of order less than k. They become norms for tempered distributions modulo polynomials. We will also encounter this situation in the Besov case later on, so this remark should just serve as a first warning to the reader.

Definition 1.1.2. *Let $k \in \mathbb{N}$. The* ***inhomogeneous Sobolev space*** $W_2^k(\mathbb{R})$ *is defined as the space of tempered distributions on $\mathbb{R}$ having all their (weak) derivatives up to order k in $L^2(\mathbb{R})$.*
Equipped with the norm $\|f\|_{W_2^k(\mathbb{R})} := \sum_{i\le k} \|D^i f\|_{L^2(\mathbb{R})}$, $W_2^k(\mathbb{R})$ is a Hilbert space.

In the introduction to this chapter, we claimed that one could regard Besov spaces as certain generalizations of classical function spaces such as the Sobolev spaces. A first attempt in this direction could be to ask for spaces $W_2^\alpha(\mathbb{R})$, where α is nonintegral. H.Triebel [33] describes this as ‘filling the gaps’ betweeen the spaces $L^2(\mathbb{R}) =: W_2^0(\mathbb{R})$, $W_2^1(\mathbb{R})$, $W_2^2(\mathbb{R})$, $\ldots$.

A first ‘filling’ can be obtained via differences, inspired by the definition of Hölder spaces:

Let $B_2^\alpha(\mathbb{R})$, $0 < \alpha < 1$, be the collection of $f \in \mathcal{S}'(\mathbb{R})$, such that

$$\|f\|_{B_2^\alpha(\mathbb{R})} := \|f\|_{L^2(\mathbb{R})} + \left(\int_{\mathbb{R}\times\mathbb{R}} \left(\frac{|f(x) - f(y)|}{|x-y|^\alpha} \right)^2 \frac{dxdy}{|x-y|} \right)^{1/2} < \infty.$$

Clearly, this can be easily extended to larger values of α, using differences on the derivatives or working with higher order differences.

One could also consider to replace the $L^2(\mathbb{R})$-norms by $L^p(\mathbb{R})$-norms where $p \neq 2$. This will lead to the definition of Besov spaces via iterated differences in $L^p(\mathbb{R})$, decaying in $L^q(\mathbb{R})$ in a certain way towards small differences. This will be the topic of Subsection 1.1.1. The *Slobodeckij spaces* $B_2^\alpha(\mathbb{R})$ above will coincide with the spaces $B_2^\alpha(L^2(\mathbb{R}))$ in these terms.

For a second way to fill the gaps, let us go back to the Sobolev spaces $W_2^k(\mathbb{R})$, $k \in \mathbb{N}$. Using the fact that differentiation corresponds to pointwise multiplication with the argument on the Fourier side, it is easy to see that they admit a characterization in Fourier analytical terms. Their norm is equivalent to

$$\|f\|_{W_2^k(\mathbb{R})} \asymp \|(1 + |\omega|^2)^{k/2} \widehat{f}(\omega)\|_{L^2(\mathbb{R})}.$$

Replacing the integer exponent by a more general $\alpha \in \mathbb{R}_+$ leads to

$$\|f\|_{W_2^\alpha(\mathbb{R})} := \|(1 + |\omega|^2)^{\alpha/2} \widehat{f}(\omega)\|_{L^2(\mathbb{R})},$$

and in fact, we have $W_2^\alpha(\mathbb{R}) = B_2^\alpha(\mathbb{R})$: the two filling procedures lead to the same spaces. Also the homogeneous Sobolev space semi-norms can be described by the decay of the Fourier transform:

$$|f|_{\dot{W}_2^k(\mathbb{R})} \asymp \||\omega|^k \hat{f}(\omega)\|_{L^2(\mathbb{R})}.$$

Again, in both cases, one could think of generalizing these spaces to some $L^p(\mathbb{R})$, yielding the so called *Liouville* or *Bessel potential spaces*, see [33].

A more general approach uses *Littlewood-Paley theory*. Littlewood and Paley obtained a characterization of $L^p(\mathbb{T})$, $1 < p < \infty$ in terms of trigonometric series. Furthermore, they showed that the p-norm is equivalent to a certain norm on the Fourier coefficients. The 2π-periodic functions in L^p are thereby fully characterized by the behavior of their Fourier expansions.

So, the basic idea is to characterize function spaces by certain sets of functions which 'span' these spaces and moreover, give rise to an equivalent description in terms of the associated 'expansion coefficients'.

The precise formulation requires a number of technical assumptions:

Let $\varphi \in \mathcal{S}(\mathbb{R})$, satisfying

$$\text{supp } \widehat{\varphi} \subseteq \{\omega : \pi/4 < |\omega| < \pi\}, \tag{1.1.1}$$
$$\text{for some } C,\ \varepsilon > 0, |\widehat{\varphi}(\omega)| > C \text{ on } \{\omega : \pi/4 + \varepsilon < |\omega| < \pi - \varepsilon\}. \tag{1.1.2}$$

Further, let φ be such that for $(\varphi_\nu)_{\nu\in\mathbb{Z}} = (2^{-\nu}\varphi(2^{-\nu}\cdot))_{\nu\in\mathbb{Z}}$, we have

$$\sum_{\nu\in\mathbb{Z}} |\widehat{\varphi}_\nu(\omega)|^2 = 1 \text{ for } \omega \in \mathbb{R} \setminus \{0\}. \tag{1.1.3}$$

As the $(\hat{\varphi}_\nu)$ form a partition of unity (1.1.3), and as $|\omega|^{2k} \asymp 2^{-2\nu k}$ on supp φ_ν, this yields for the Sobolev semi-norm

$$\begin{aligned} |f|^2_{\dot{W}_2^k(\mathbb{R})} &= \|D^k f\|^2_{L^2(\mathbb{R})} \asymp \int_{\mathbb{R}} |\omega|^{2k} |\widehat{f}(\omega)|^2 d\omega \\ &= \sum_{\nu\in\mathbb{Z}} \int_{\mathbb{R}} |\omega|^{2k} |\widehat{f}(\omega)|^2 |\widehat{\varphi}_\nu(\omega)|^2 d\omega \\ &\asymp \sum_{\nu\in\mathbb{Z}} 2^{-2\nu k} \int_{\mathbb{R}} |\widehat{f}(\omega)|^2 |\widehat{\varphi}_\nu(\omega)|^2 d\omega \\ &= \sum_{\nu\in\mathbb{Z}} (2^{-\nu k} \|f * \varphi_\nu\|_{L^2(\mathbb{R})})^2. \end{aligned} \tag{1.1.4}$$

The underlying idea is that any $L^2(\mathbb{R})$-function can be decomposed as

$$f = \sum_{\nu\in\mathbb{Z}} f * \varphi_\nu * \varphi_\nu^*, \tag{1.1.5}$$

where $\varphi_\nu^*(x) = \overline{\varphi_\nu(-x)}$. (1.1.5) is called the *Calderón reproducing formula*.

The map $f \mapsto (f * \varphi_\nu)_{\nu\in\mathbb{Z}}$ can be understood as a decomposition of f into signal components $f * \varphi_\nu$ whose Fourier transforms are localized in dyadically spaced 'frequency bands'. The smoothness of f is then related to the decay of the L^2-norms of the different components. On the one hand, this definition reflects the well-known characterization of global smoothness via the Fourier transform, but the use of a smooth partition of unity in Fourier domain, resulting in a rapidly decaying window φ, allows to measure smoothness of a function locally.

One can now consider this decay in other norms than L^2, which will lead to another equivalent definition of Besov and related spaces. We describe this more detailed in Subsection 1.1.2. Of course, the meaning of (1.1.5) has then to be handled with great care. For general $f \in \mathcal{S}'(\mathbb{R})$, the expansion will converge (in the weak-$*$-sense) only 'modulo polynomials'. This issue is also a topic of Subsection 1.1.2.

The next Subsection, 1.1.3, deals with the φ-transform characterization of Besov Spaces. By discretizing Calderón's formula, Frazier and Jawerth [15], [16] obtained that for any f in a Besov space,

$$f = \sum_{\nu,k\in\mathbb{Z}} \langle f, \varphi_{\nu,k}\rangle \varphi_{\nu,k}, \tag{1.1.6}$$

where for $k \in \mathbb{Z}$, $\varphi_{\nu,k}(x) = 2^{-\nu/2}\varphi(2^{-\nu}x - k)$. Moreover, the membership of f in a Besov space is fully characterized by the membership of the coefficients $(\langle f, \varphi_{\nu,k}\rangle)_{\nu,k\in\mathbb{Z}}$ in a sequence space.

This decomposition for distributions in the discrete-time Besov space is similar to an expansion into an orthonormal wavelet basis, though non-orthogonal and with 'basis elements' that are compactly supported in Fourier domain. In Subsection 1.1.4, we shortly explain how to derive that $L^2(\mathbb{R})$-orthonormal wavelet bases - satisfying certain additional conditions - are unconditional bases for the Besov spaces as well.

Although we focus on the Littlewood-Paley description and the φ- and wavelet transform characterizations derived from it, we start with the Besov space definition based on iterated differences, which is frequently used in literature.

1.1.1 Moduli of Smoothness

In this subsection, we define Besov spaces as spaces of functions with a common order of smoothness, which is measured via iterated differences. These spaces will be denoted by $B_q^\alpha(L^p(\mathbb{R}))$ and $\dot{B}_q^\alpha(L^p(\mathbb{R}))$, respectively. As this subsection is mainly introductory, we restrict our discussion to spaces where $\alpha > 0$, $1 < p, q < \infty$ in order to avoid special case treatment. For a more detailed discussion, we refer e.g. to [25] or [34].

Let $h \in \mathbb{R}$. For a function f defined on $\mathbb{R}$, the (forward) *difference operator of step* h is given by

$$\Delta_h f(x) = f(x+h) - f(x),$$

and for $r \in \mathbb{N}_+$, define the *difference operator of order* r, step h, inductively by

$$\Delta_h^r f(x) = \Delta_h(\Delta_h^{r-1} f(x)).$$

Note that the r-th difference operator in explicit form is given by

$$\Delta_h^r f(x) = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} f(x+kh).$$

Definition 1.1.3. *For $1 < p < \infty$, $t \in \mathbb{R}_+$, the r-th order **modulus of smoothness** of f in $L^p(\mathbb{R})$ is defined by*

$$\omega_p^r(f,t) = \sup_{h\in\mathbb{R},|h|<t} \|\Delta_h^r f(\cdot)\|_p.$$

Let f, g be defined on $\mathbb{R}$. Then, for each $t \in \mathbb{R}_+$,

$$\omega_p^r(f+g,t) \le \omega_p^r(f,t) + \omega_p^r(g,t), \tag{1.1.7}$$

and for f multiplied by a scalar α,

$$\omega_p^r(\alpha f,t) \le |\alpha|\omega_p^r(f,t). \tag{1.1.8}$$

As $\omega_p^r(f,t)$ vanishes for polynomials of degree $\le r-1$, $\omega_p^r(\cdot,t)$ is a semi-norm on the set of functions for which $\omega_p^r(f,t) < \infty$.
$\omega_p^r(f,t)$ is increasing for each p and r, furthermore, for $M \in \mathbb{N}$,

$$\omega_p^r(f, M\cdot t) \le M^r \omega_p^r(f,t). \tag{1.1.9}$$

So if $\omega_p^r(f,t) < \infty$ for some $t > 0$, it is finite for all $t \in \mathbb{R}_+$.
For $f \in L^p(\mathbb{R})$, we have

$$\omega_p^r(f,t) \le 2^r \|f\|_p. \tag{1.1.10}$$

Therefore, functions in $L^p(\mathbb{R})$ have finite moduli of smoothness, but note that conversely, functions with finite moduli of smoothness are not necessarily in $L^p(\mathbb{R})$.
For $f \in L^p(\mathbb{R})$, we have $\omega_p^r(f,t) \to 0$ monotonically as $t \to 0$. Generally speaking, the faster this convergence, the smoother f.
Note that the L^p-norm in the definition of $\omega_p^r(f,t)$ allows a rather wild behavior of f, as long as the exceptional set has small measure. This property implies a certain tolerance of Besov spaces $\dot{B}_q^\alpha(L^p(\mathbb{R}))$ we define below, with respect to jumps.

Definition 1.1.4. *For $\alpha > 0, 1 < p, q < \infty$, $r = \lfloor \alpha \rfloor + 1$, a function f defined on $\mathbb{R}$ is said to be in the **homogeneous Besov space** $\dot{B}_q^\alpha(L^p(\mathbb{R}))$, if*

$$|f|_{\dot{B}_q^\alpha(l^p(\mathbb{Z}))} := (\int_0^\infty (t^{-\alpha}\omega_p^r(f,t))^q \frac{dt}{t})^{1/q} < \infty. \tag{1.1.11}$$

The $|\cdot|_{B_q^\alpha(L^p(\mathbb{R}))}$ are semi-norms in general because of the polynomial cancellation properties of the moduli of smoothness; they become norms modulo polynomials of degree $\le r-1$. Furthermore, the $\dot{B}_q^\alpha(L^p(\mathbb{R}))$-seminorms are all equivalent modulo polynomials using different moduli of smoothness $r > \alpha$ in the definition.
We define the corresponding *inhomogeneous Besov spaces* $B_q^\alpha(L^p(\mathbb{R}))$ by the norm

$$\|f\|_{B_q^\alpha(L^p(\mathbb{R}))} := \|f\|_{L^p(\mathbb{R})} + |f|_{\dot{B}_q^\alpha(L^p(\mathbb{R}))}.$$

Using the monotonicity of $\omega_p^r(f,\cdot)$ and (1.1.9), the semi-norm can be discretized, yielding an equivalent semi-norm

$$|f|_{\dot{B}_q^\alpha(L^p(\mathbb{R}))} \asymp \Big(\sum_{j\in\mathbb{Z}} (2^{-j\alpha}\omega_p^r(f,2^j))^q\Big)^{1/q}. \tag{1.1.12}$$

The above spaces $\dot{B}_q^\alpha(L^p(\mathbb{R}))$ and $B_q^\alpha(L^p(\mathbb{R}))$, respectively, generalize the considerations from the introduction to this section: choosing $p=q=2$ and $\alpha=k\in\mathbb{N}$, these spaces coincide with the homogeneous resp. inhomogeneous Sobolev spaces of order k. For nonintegral α, we have $B_2^\alpha(L^2(\mathbb{R})) = B_2^\alpha(\mathbb{R})$, the Slobodeckij spaces [33].

1.1.2 Littlewood-Paley Type Characterization

In this subsection, we give a Littlewood-Paley type definition of homogeneous and inhomogeneous smoothness spaces $\dot{B}_{p,q}^\alpha(\mathbb{R})$ and $B_{p,q}^\alpha(\mathbb{R})$, $\alpha\in\mathbb{R}$ and $0<p,q\le\infty$. These spaces will coincide with the Besov spaces $\dot{B}_q^\alpha(L^p(\mathbb{R}))$ and their inhomogeneous analogs, respectively, which we defined in the last subsection via iterated differences.

The principle behind this type of characterization of Besov spaces is to decompose distributions into series of smooth components:

Let $(\varphi_\nu)_{\nu\in\mathbb{Z}}$ be a family of rapidly decreasing functions, satisfying (1.1.1)-(1.1.3).

Recall from above that for $f\in L^2(\mathbb{R})$ (see (1.1.5)):

$$f = \sum_{\nu\in\mathbb{Z}} f*\varphi_\nu*\varphi_\nu^*,$$

with convergence in $L^2(\mathbb{R})$ (Calderón formula) and

$$\|f\|^2_{L^2(\mathbb{R})} = \sum_{\nu\in\mathbb{Z}} \|f*\varphi_\nu\|^2_{L^2(\mathbb{R})}. \tag{1.1.13}$$

Thereby, $L^2(\mathbb{R})$-functions are characterized by the size of the 'smooth parts' $f*\varphi_\nu$.

We want to generalize this type of result to other spaces (a first example was given in the beginning of this chapter concerning Sobolev spaces). As a first step in this direction, we will investigate the convergence of Calderón's formula in spaces different from $L^2(\mathbb{R})$.

For $f\in\mathcal{S}'(\mathbb{R})$, Calderón's formula does in general not converge in the distributional sense. The sum over $\nu\ge1$ may diverge, see the examples in [23] and [14].

However, it can be shown that the series

$$\sum_{\nu\in\mathbb{Z}} D^i(f*\varphi_\nu*\varphi_\nu^*) \tag{1.1.14}$$

converges for some $i \in \mathbb{N}$.

This is in fact equivalent to the existence of a sequence of polynomials $\{P_k\}_{k\geq 1}$ of degree less than i, such that

$$g := \lim_{k\to\infty} (\sum_{\nu=-\infty}^{k} f * \varphi_\nu * \varphi_\nu^* + P_k),$$

in $\mathcal{S}'(\mathbb{R})$.

Furthermore, the limit above differs from f by a polynomial, as supp $(\widehat{f} - \widehat{g}) = 0$.

Altogether, we have for $f \in \mathcal{S}'(\mathbb{R})$

$$f = \sum_{\nu\in\mathbb{Z}} f * \varphi_\nu * \varphi_\nu^*, \tag{1.1.15}$$

with convergence in $\mathcal{S}'/\mathcal{P}(\mathbb{R})$, the equivalence class of tempered distributions modulo polynomials.

We now define smoothness spaces by generalizing (1.1.13).

Definition 1.1.5. *For $\alpha \in \mathbb{R}$, $0 < p, q \leq \infty$, the space $\dot{B}^\alpha_{p,q}(\mathbb{R})$ is the collection of all $f \in \mathcal{S}'/\mathcal{P}(\mathbb{R})$, such that*

$$\|f\|_{\dot{B}^\alpha_{p,q}(\mathbb{R})} := (\sum_{\nu\in\mathbb{Z}} (2^{-\nu\alpha} \|f * \varphi_\nu\|_{L^p(\mathbb{R})})^q)^{1/q} < \infty \quad .$$

This definition is independent of the choice of the family $(\varphi_\nu)_{\nu\in\mathbb{Z}}$.

The $\dot{B}^\alpha_{p,q}(\mathbb{R})$ spaces are Banach spaces for $1 \leq p, q < \infty$ and quasi-Banach spaces otherwise.

The underlying concept of measuring smoothness as in (1.1.5) is quite different from (1.1.12), but in fact, $\dot{B}^\alpha_{p,q}(\mathbb{R})$ and $B^\alpha_q(L^p(\mathbb{R}))$ agree (modulo polynomials) at least for $\alpha > 0$, $1 < p, q < \infty$, the range of parameters for which we defined the latter. Especially we saw in (1.1.4) that $\dot{B}^k_{2,2}(\mathbb{R}) = \dot{W}^k_2(\mathbb{R})$ for $k \in \mathbb{N}$. On account of this, we will also call the $\dot{B}^\alpha_{p,q}(\mathbb{R})$ spaces *homogeneous Besov spaces*, with a now extended range of parameters.

One also can define *inhomogeneous spaces* $B^\alpha_{p,q}(\mathbb{R})$ by replacing the low frequency parts in the Calderón formula by a single function:

Let $\Phi \in \mathcal{S}(\mathbb{R})$, where supp $\widehat{\Phi} \subseteq]-\pi, \pi[$ and for some C, $\varepsilon > 0$, $|\widehat{\Phi}(\omega)| > C$ on $\{\omega : \pi + \varepsilon < |\omega| < \pi + \varepsilon\}$.

Let $(\varphi_\nu)_{\nu\in\mathbb{Z}}$ again satisfy (1.1.1) and (1.1.2), such that

$$|\widehat{\Phi}(\omega)|^2 + \sum_{\nu\leq -1} |\widehat{\varphi}_\nu(\omega)|^2 = 1.$$

We have now for $f \in \mathcal{S}'(\mathbb{R})$

$$f = f * \Phi * \Phi^* + \sum_{\nu \le -1} f * \varphi_\nu * \varphi_\nu^*. \tag{1.1.16}$$

Definition 1.1.6. *For $\alpha \in \mathbb{R}$, $0 < p, q \le \infty$, the space $B^\alpha_{p,q}(\mathbb{R})$ is the collection of all $f \in \mathcal{S}'(\mathbb{R})$, such that*

$$\|f\|_{B^\alpha_{p,q}(\mathbb{R})} := \|f * \Phi\|_{L^p(\mathbb{R})} + (\sum_{\nu \le -1} (2^{-\nu\alpha} \|f * \varphi_\nu\|_{L^p(\mathbb{R})})^q)^{1/q} < \infty \quad .$$

Again, these space agree with the Besov spaces from Subsection 1.1.1 for the parameters α, p, q, for which the latter is defined. Thereby, we will call the $B^\alpha_{p,q}(\mathbb{R})$ spaces *inhomogeneous Besov spaces* as well.

The terms *homogeneous* and *inhomogeneous* originate from the behavior of both of the spaces concerning dilation, as for the homogeneous Besov spaces $\dot{B}^\alpha_{p,q}(\mathbb{R})$, $l \in \mathbb{Z}$,

$$\|f(2^l \cdot)\|_{\dot{B}^\alpha_{p,q}(\mathbb{R})} = C 2^{l(\alpha - 1/p)} \|f\|_{\dot{B}^\alpha_{p,q}(\mathbb{R})},$$

whereas for the inhomogeneous spaces, this equality is generally not true [32].

1.1.3 φ-transform Characterization

In [15], Frazier and Jawerth introduced the so-called φ-transform, which can be viewed as a a critically sampled version of (1.1.5). The function spaces which are described by Littlewood-Paley expressions can also be characterized by the φ-transform. More precisely, the φ-transform coefficients carry all the necessary information to conclude the membership of a distribution in a Besov space. Moreover, the condition on the coefficients is just a size condition, which may simplify applications such as the study of linear operators on the $\dot{B}^\alpha_{p,q}(\mathbb{R})$ (and the $B^\alpha_{p,q}(\mathbb{R})$) and related spaces. A more detailed descripition of these results can be found in [15],[16] or [17].

Consider again a function $\varphi \in \mathcal{S}(\mathbb{R})$, satisfying (1.1.1), (1.1.2) and (1.1.3).

For $\nu, k \in \mathbb{Z}$ let

$$\varphi_{\nu,k}(x) = 2^{-\nu/2} \varphi(2^{-\nu} x - k).$$

Starting from the formula (1.1.5) and using techniques similar to Shannon sampling, Frazier and Jawerth derived that for any $f \in \mathcal{S}'(\mathbb{R})$

$$f = \sum_{\nu \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f, \varphi_{\nu,k} \rangle \varphi_{\nu,k}, \tag{1.1.17}$$

with convergence in $\mathcal{S}'/\mathcal{P}(\mathbb{R})$, which is a discretization of 1.1.15.

Furthermore, the Besov space norms can be equivalently expressed in terms of the coefficients.

Definition 1.1.7. *For $\alpha \in \mathbb{R}$, $0 < p, q \leq \infty$, let the coefficient spaces $\dot{b}^{\alpha}_{p,q}(\mathbb{R})$ be the collection of all complex-valued sequences $s = (s_{\nu,k})_{\nu,k\in\mathbb{Z}}$, satisfying*

$$\|s\|_{\dot{b}^{\alpha}_{p,q}(\mathbb{R})} := (\sum_{\nu\in\mathbb{Z}}(\sum_{k\in\mathbb{Z}}(2^{-\nu(\alpha+1/2-1/p)}|s_{\nu,k}|)^p)^{q/p})^{1/q} < \infty. \tag{1.1.18}$$

Let the φ-transform S_φ for $f \in \mathcal{S}'(\mathbb{R})$ be defined by $S_\varphi f = s = (s_{\nu,k})_{\nu,k\in\mathbb{Z}}$, where $s_{\nu,k} = \langle f, \varphi_{\nu,k}\rangle$, and for a complex-valued sequence $t = (t_{\nu,k})_{\nu,k\in\mathbb{Z}}$ define the inverse φ-transform by T_φ by $T_\varphi t = \sum_{\nu,k} t_{\nu,k}\varphi_{\nu,k}$.

In [15], Frazier and Jawerth prove the following result:

Theorem 1.1.8. *Let $\alpha \in \mathbb{R}$, $0 < p, q \leq \infty$.*
Both of the operators $S_\varphi : \dot{B}^{\alpha}_{p,q}(\mathbb{R}) \to \dot{b}^{\alpha}_{p,q}(\mathbb{R})$ and $T_\varphi : \dot{b}^{\alpha}_{p,q}(\mathbb{R}) \to \dot{B}^{\alpha}_{p,q}(\mathbb{R})$ are bounded with $\|f\|_{\dot{B}^{\alpha}_{p,q}(\mathbb{R})} \asymp \|S_\varphi f\|_{\dot{b}^{\alpha}_{p,q}(\mathbb{R})}$ and $T_\varphi \circ S_\varphi = id_{\dot{B}^{\alpha}_{p,q}(\mathbb{R})}$.

In other words, under these maps, $\dot{B}^{\alpha}_{p,q}(\mathbb{R})$ is a retract of $\dot{b}^{\alpha}_{p,q}(\mathbb{R})$, and $\dot{B}^{\alpha}_{p,q}(\mathbb{R})$ can be identified with the closed subspace $S_\varphi(\dot{B}^{\alpha}_{p,q}(\mathbb{R}))$ of $\dot{b}^{\alpha}_{p,q}(\mathbb{R})$.

Observe that the well-definedness and unconditional convergence of $T_\varphi t = \sum_{\nu,k} t_{\nu,k}\varphi_{\nu,k}$, which was not clear initially, follows from the theorem.

There is an analogous result concerning the inhomogeneous spaces, starting from the identity (1.1.16).

Let again $\Phi \in \mathcal{S}(\mathbb{R})$, with supp $\widehat{\Phi} \subseteq]-\pi, \pi[$ and for some C, $\varepsilon > 0$, $|\widehat{\Phi}(\omega)| > C$ on $\{\omega : \pi + \varepsilon < |\omega| < \pi + \varepsilon\}$ and let $(\varphi_\nu)_{\nu\in\mathbb{Z}}$ satisfy (1.1.1) and (1.1.2), such that

$$|\widehat{\Phi}(\omega)|^2 + \sum_{\nu\leq -1} |\widehat{\varphi}_\nu(\omega)|^2 = 1.$$

Write again $\varphi_{\nu,k}(x) = 2^{-\nu/2}\varphi(2^{-\nu}x - k)$ for $\nu, k \in \mathbb{Z}$ and $\Phi_k(x) = \Phi(x - k)$.

The φ-transform identity for $f \in \mathcal{S}'(\mathbb{R})$ now reads as

$$f = \sum_{k\in\mathbb{Z}}\langle f, \Phi_k\rangle\Phi_k + \sum_{\nu\leq -1}\sum_{k\in\mathbb{Z}}\langle f, \varphi_{\nu,k}\rangle\varphi_{\nu,k}, \tag{1.1.19}$$

converging in the weak-$*$ sense.

The size of the coefficients reflects the smoothness of the distribution f just like in the homogeneous case. An appropriate formulation of this size condition is given in the next definition.

Definition 1.1.9. *For $\alpha \in \mathbb{R}$, $0 < p, q \leq \infty$, let the spaces $b^{\alpha}_{p,q}(\mathbb{R})$ be the collection of all complex-valued sequences $s = (s_{\nu,k})_{\nu,k\in\mathbb{Z}}$, satisfying*

$$\|s\|_{b^{\alpha}_{p,q}(\mathbb{R})} := (\sum_{\nu\leq 0}(\sum_{k\in\mathbb{Z}}(2^{-\nu(\alpha+1/2-1/p)}|s_{\nu,k}|)^p)^{q/p})^{1/q} < \infty. \tag{1.1.20}$$

Let now the φ-transform for $f \in \mathcal{S}'(\mathbb{R})$ be defined by the mapping on the coefficients $(s_{\nu,k})_{\nu\leq 0,k\in\mathbb{Z}}$, where $s_{\nu,k} = \langle f, \varphi_{\nu,k}\rangle$ if $\nu \leq -1$ and $s_{0,k} = \langle f, \Phi_k\rangle$. Defining the inverse φ-transform in the obvious way, we have, analogously to the homogeneous spaces, that the $B^{\alpha}_{p,q}(\mathbb{R})$ spaces are retracts of $b^{\alpha}_{p,q}(\mathbb{R})$ and the coefficients satisfy

$$\|f\|_{B^{\alpha}_{p,q}(\mathbb{R})} \asymp \|(s_{\nu,k})\|_{b^{\alpha}_{p,q}(\mathbb{R})}. \tag{1.1.21}$$

These results can be used for the analysis of linear operators acting on the Besov spaces.
Let $f \in \dot{B}^{\alpha}_{p,q}(\mathbb{R})$. By Theorem 1.1.8, $f = \sum_{\nu\geq 1,k\in\mathbb{Z}} s_{\nu,k}\varphi_{\nu,k}$, where $(s_{\nu,k}) = (\langle f, \varphi_{\nu,k}\rangle)$. Applying a linear operator T leads to

$$\begin{aligned} Tf &= \sum_{\nu,k} s_{\nu,k}T\varphi_{\nu,k} \\ &= \sum_{\nu,k} s_{\nu,k}(\sum_{\mu,l}\langle T\varphi_{\nu,k}, \varphi_{\mu,l}\rangle\varphi_{\mu,l}) \\ &= \sum_{\mu,l}(\sum_{\nu,k}\langle T\varphi_{\nu,k}, \varphi_{\mu,l}\rangle s_{\nu,k})\varphi_{\mu,l}. \end{aligned} \tag{1.1.22}$$

Define the matrix A by $A := (a_{\mu,l,\nu,k})$, where $a_{\mu,l,\nu,k} = \langle T\varphi_{\nu,k}, \varphi_{\mu,l}\rangle$.
Thus, using (1.1.22)

$$Tf = \sum_{\mu,l}(As)_{\mu,l}\varphi_{\mu,l},$$

where $(As)_{\mu,l} = \sum_{\nu,k} a_{\mu,l,\nu,k}s_{\nu,k}$.
By Theorem 1.1.8, if A is bounded on the sequence space,

$$\|Tf\|_{\dot{B}^{\alpha}_{p,q}(\mathbb{R})} \leq C\|As\|_{\dot{b}^{\alpha}_{p,q}(\mathbb{R})} \leq C\|s\|_{\dot{b}^{\alpha}_{p,q}(\mathbb{R})} \leq C\|f\|_{\dot{B}^{\alpha}_{p,q}(\mathbb{R})}, \tag{1.1.23}$$

and thereby, this implies boundedness of T. Thus, the study of bounded linear operators on the Besov space reduces to studying matrices on the $\dot{b}^{\alpha}_{p,q}(\mathbb{R})$ spaces.
More formally, for a linear operator T on $\dot{B}^{\alpha}_{p,q}(\mathbb{R})$, define

$$S^*_{\varphi}(T) = S_{\varphi} \circ T \circ T_{\varphi}$$

and similarly, for A a bounded operator on $\dot{b}^{\alpha}_{p,q}(\mathbb{R})$

$$T^*_\varphi(A) = T_\varphi \circ A \circ S_\varphi.$$

As seen in (1.1.23), T^*_φ is bounded. The boundedness of S^*_φ and the fact that $T^*_\varphi \circ S^*_\varphi(T) = T_\varphi \circ S_\varphi \circ T \circ T_\varphi \circ S_\varphi = T$ imply that the space of bounded linear operators on $\dot{B}^{\alpha}_{p,q}(\mathbb{R})$ is a retract of the space of bounded linear operators on $\dot{b}^{\alpha}_{p,q}(\mathbb{R})$: the retract property is lifted to the operator level. This allows to study operators on the Besov spaces by studying properties of certain matrices [17].

One condition for matrices to be bounded on the coefficient space is the so-called 'almost diagonality property'. Roughly spoken, this condition describes how fast the size of the entries $a_{\mu,l,\nu,k}$ decays away from the diagonal, where $\mu = \nu$, $l = k$. An exact definition can be found again e.g. in [17]. We will deal with matrices of this type in more detail when dealing with the discrete-time spaces, see Lemma 3.3.7: the conditions on the sequences therein can be viewed as almost diagonal conditions.

An application of these results is the wavelet characterization of Besov spaces, which will be the issue of the upcoming subsection.

1.1.4 Wavelet Characterization

The φ-transform we discussed in the last chapter allows a characterization of Besov spaces by the decay of the associated coefficients. It is also possible to describe these spaces by other transforms, especially by orthonormal wavelet transforms.

In many respects, a wavelet orthonormal basis is quite similar to the system of functions associated to the φ-transform. Both systems arise by picking a suitable function, which is dilated by powers of two and translated by integer multiples of the dilation variable. Arbitrary functions in L^2 can be expanded in both systems, and additional properties of the φ-function (or the wavelet) guarantee that these expansions remain valid for other function spaces.

A more detailed exposition of the results cited below can be found in [17, 14]. The techniques used to prove these results provide a blueprint for our treatment of discrete-time wavelets in Chapter 3.

Let ψ be a wavelet function such that the family $(\psi_{j,l})_{j,l\in\mathbb{Z}} = (2^{-j/2}\psi(2^{-j}x - l))$ constitutes an orthonormal basis for $L^2(\mathbb{R})$ (see Chapter 2 for a quick introduction to this issue). Furthermore, let this wavelet system fulfill the additional conditions

i) ψ possesses zero moments of a certain order $N > 0$,

ii) ψ is regular of order N, e.g. ψ is N times continuously differentiable,

iii) ψ is localized in time, e.g. ψ has compact support.

We discuss these properties in more detail in 2.1.2.

In contrast to the φ-functions, which are band-limited Schwartz-functions with all of their moments vanishing, the ψ-functions are now localized in time rather than frequency and usually possess only vanishing moments and regularity of a certain finite degree.

Nevertheless, starting from Theorem 1.1.8, one can prove that the Besov spaces can be described also by the wavelet transform, where the number N above has to be chosen accordingly, depending on the order of smoothness α of the space. Note that N describes the frequency localization of ψ (see Remark 2.1.9 below), which makes the conditions on ψ in Theorem 1.1.10 intuitively plausible.

Let the wavelet transform S_ψ for $f \in \mathcal{S}'/\mathcal{P}(\mathbb{R})$ be defined by $S_\psi f = r = (r_{j,l})_{j,l\in\mathbb{Z}}$, where $r_{j,l} = \langle f, \psi_{j,l}\rangle$ and the inverse transform T_ψ accordingly by $T_\psi v = \sum_{j,l} v_{j,l}\psi_{j,l}$ for a complex-valued sequence $v = (v_{j,l})_{j,l\in\mathbb{Z}}$.

The proof of an an analog of Theorem 1.1.8 is based on estimates on how the properties of ψ, quantified by the number N, influence the **off-diagonal decay of the matrix** $A = (\langle\psi_{j,l}, \varphi_{\nu,k}\rangle)$ and to conclude boundedness of this matrix on the coeffient space. Using the retract property we mentioned in the last section, the following theorem can be established [17].

Theorem 1.1.10. *Let $\alpha \in \mathbb{R}$, $0 < p, q < \infty$.*
Provided that $N > \max\{\alpha, 1/\min\{1,p\}-1-\alpha\}$, both of the operators $S_\psi : \dot{B}^\alpha_{p,q}(\mathbb{R}) \to \dot{b}^\alpha_{p,q}(\mathbb{R})$ and $T_\psi : \dot{b}^\alpha_{p,q}(\mathbb{R}) \to \dot{B}^\alpha_{p,q}(\mathbb{R})$ are bounded with $\|f\|_{\dot{B}^\alpha_{p,q}(\mathbb{R})} \asymp \|S_\psi f\|_{\dot{b}^\alpha_{p,q}(\mathbb{R})}$.
Furthermore, $T_\psi \circ S_\psi = id_{\dot{B}^\alpha_{p,q}(\mathbb{R})}$ ***as well as*** $\mathbf{S}_\psi \circ \mathbf{T}_\psi = \mathbf{id}_{\dot{\mathbf{b}}^\alpha_{\mathbf{p},\mathbf{q}}(\mathbb{R})}$.

In the orthonormal wavelet case, the wavelet system in fact inherits more than the retract property of the φ-transform. Under the ψ-transform, $\dot{B}^\alpha_{p,q}(\mathbb{R})$ is isomorphic to the whole space $\dot{b}^\alpha_{p,q}(\mathbb{R})$. In particular, observe that

$$f = \sum_{j,l\in\mathbb{Z}} \langle f, \psi_{j,l}\rangle\psi_{j,l} \tag{1.1.24}$$

converges unconditionally in $\dot{B}^\alpha_{p,q}(\mathbb{R})$. **Wavelet systems** satisfying the sufficient conditions in the theorem are therefore **unconditional bases** for a **whole scale of Besov spaces simultaneously**.

Of course, there is a similar result for the inhomogeneous Besov spaces. These results can be found in [17], [14] or, in a slightly different context, in [20].

1.2 Nonlinear Wavelet Approximation and Besov Spaces

The results we presented in Section 1.1.4 are related to the nonlinear approximation behavior of wavelet expansions and to wavelet applications in signal and image processing, such as denoising and compression; let us mention Donoho and Johnstone's *Wavelet Shrinkage* (see e.g. [5]) or [24] for compression.

As a motivating example, we consider the problem of approximating a function $f \in L^2(\mathbb{R})$ by $N > 0$ terms of a wavelet series. The results we give in this section are borrowed from [7].

Let again $(\psi_{j,l})_{j,l\in\mathbb{Z}}$ be an orthonormal wavelet basis for $L^2(\mathbb{R})$. The task is to approximate $f \in L^2(\mathbb{R})$ by N wavelet coefficients, i.e. to pick an approximant from

$$\Sigma_N = \{g \in L^2(\mathbb{R}) : g = \sum_{(j,l)\in\Lambda} c_{j,l}\psi_{j,l},\ \ \Lambda \subset \mathbb{N},\ \mathrm{card}(\Lambda) \leq N\}.$$

This is an nonlinear problem as the spaces Σ_N are nonlinear: $\Sigma_N \neq \Sigma_N + \Sigma_N \subset \Sigma_{2N}$. Define the approximation error by

$$\sigma_N(f) = \inf_{g\in\Sigma_N} \|f - g\|_{L^2(\mathbb{R})}.$$

Heuristically, one can conceive that the 'smoother' f, the faster this error decreases. To quantify this, we ask: for which $f \in L^2(\mathbb{R})$ we have for a given $\alpha > 0$

$$\sigma_N(f) \leq MN^{-\alpha},$$

for some $M > 0$, or, slightly stronger, $f \in \mathcal{A}_p^\alpha(L^2(\mathbb{R}), \Sigma_N)$, where

$$\mathcal{A}_p^\alpha(L^2(\mathbb{R}), \Sigma_N) := \{f \in L^2(\mathbb{R}), \left(\sum_{N=1}^{\infty}(N^\alpha\sigma_N(f))^p\frac{1}{N}\right)^{1/p} < \infty\}.$$

Here the justification for describing the condition

$$\sum_{n=1}^{\infty}(N^\alpha\sigma_N(f))^p\frac{1}{N} < \infty$$

as '$\sigma_N(f)$ decays as $N^{-\alpha}$' is provided by the observation that $N^\alpha\sigma_N(f)$ has to converge to zero in order to guarantee finiteness of the sum. In fact, σ_N has to decay slightly faster. The p parameter plays the role of a fine-tuning parameter, similar to the role of q in the definition of Besov spaces.

The recipe to minimize $\|f - g\|_{L^2(\mathbb{R})}$ subject to $g \in \Sigma_N$ is to build the wavelet series up by picking the N coefficients of the largest absolute values. So let (c_n) be the permutation of the sequence $(\langle f, \psi_{j,l}\rangle)_{j,l\in\mathbb{Z}}$, such that $|c_1| \geq |c_2| \geq |c_3| \geq \ldots \geq 0$.

Then, we have $\sigma_N(f) = (\sum_{j=N+1}^{\infty} c_j^2)^{1/2}$. Estimating these coefficients and using that ℓ^p embeds continuously into ℓ^2 for $0 < p < 2$, one can show that $f \in \mathcal{A}_p^\alpha(L^2(\mathbb{R}), \Sigma_N)$ if and only if $(\langle f, \psi_{j,l}\rangle)_{j,l\in\mathbb{Z}} \in \ell^p(\mathbb{Z} \times \mathbb{Z})$, where $0 < p < 2$, $\alpha = 1/p - 1/2$.

As $\ell^p(\mathbb{Z}\times\mathbb{Z}) = \dot{b}_{p,p}^\alpha(\mathbb{Z})$ in this case, we can conclude that, assuming certain support, zero moment and smoothness properties of the wavelet system,

$$f \in \mathcal{A}_p^\alpha(L^2(\mathbb{R}), \Sigma_N) \iff f \in \dot{B}_{p,p}^\alpha(\mathbb{R}).$$

Summing up, we have that for $f \in L^2(\mathbb{R})$, a decay of the approximation error like $O(N^{-\alpha})$ corresponds to the membership of f in a Besov space.

1.2.1 Nonlinear Approximation of Discrete-Time Signals

In applications, data are usually given discretely. Usually, these data are fed straightly into Mallat's cascade algorithm (see Chapter 2), using the discrete filters arising from a continuous-time multiresolution analysis:

This way, the discrete series is mapped to a family $(d_{j,l})_{j\geq 1, l\in\mathbb{Z}}$ of discrete wavelet coefficients, which can be interpreted as expansion coefficients of the signal with respect to a discrete-time wavelet basis.

Considering the example of N-term approximation in this discrete case, say, of $f = (f(n)) \in \ell^2(\mathbb{Z})$, one can derive a similar result as in the continuous case: the discussion above is valid for any Hilbert space and not only for the $L^2(\mathbb{R})$ case we considered.

So, roughly said, we have that the ℓ^2-approximation error decreases like $O(N^{-\alpha})$ for $(d_{j,l}) \in \ell^p(\mathbb{N} \times \mathbb{Z})$, $0 < p < 2$, $\alpha = 1/p - 1/2$.

This suggests using the space of truncated coefficient families $b_{p,q}^\alpha(\mathbb{Z})$, $\alpha \in \mathbb{R}$, $0 < p, q < \infty$, as the collection of complex-valued sequences $s = (s_{j,l})_{j\geq 1, l\in\mathbb{Z}}$, for which

$$\|s\|_{b_{p,q}^\alpha(\mathbb{Z})} := (\sum_{j\geq 1} (\sum_{l\in\mathbb{Z}} (2^{-j(\alpha+1/2-1/p)}|s_{j,l}|)^p)^{q/p})^{1/q} < \infty. \tag{1.2.1}$$

Then the task arises to characterize $(d_{j,l})_{j\geq 1, l\in\mathbb{Z}} \in \dot{b}_{p,q}^\alpha(\mathbb{Z})$ in a similar way as for the continuous case from properties of the sequence f alone, and to study conditions on the filter bank rather than on the underlying continuous-time wavelets.

As we already discussed in the introduction, from the continuous-time point of view, the decay of $(d_{j,l})$ as in (1.2.1) depends on the membership of $F = \sum_{n\in\mathbb{Z}} f(n)\tau_n\phi$ in a Besov space $B_{p,q}^\alpha(\mathbb{R})$, which involves the scaling function ϕ associated to the filter bank.

At this point, there seems to be a gap in the theory existing so far: except for heuristics based on the continuous-time analogue, we could not find any literature providing a direct description for discrete data with coefficients in $b_{p,q}^\alpha(\mathbb{Z})$ in terms of membership in a space of sequences.

This observation is relevant because any argument citing results for the continuous time setting, but applying them to the coefficient computed in discrete time, implicitly uses the coefficient norm (1.2.1) on the wavelet coefficients. Existing theory does not even provide criteria for **consistency** of these norms with respect to change of wavelet basis.

1.3 Aims of This Thesis

We are now ready to formulate the main objectives of this thesis:

1. We want to **characterize wavelet coefficient decay in a purely discrete-time setting** in a way that is as complete and satisfactory as the characterization for continuous-time signals. This leads us to the study of **Besov spaces on the integers** $B^{\alpha}_{p,q}(\mathbb{Z})$, whose definition was given by Torres in [31].

   Especially, we study necessary and sufficient conditions on wavelet bases for $l^2(\mathbb{Z})$ to constitute **unconditional bases** for these spaces.

2. The spaces $B^{\alpha}_{p,q}(\mathbb{Z})$ are defined via Littlewood-Paley-theory. Another aim for us is to give **further characterizations** of these spaces such as in terms of iterated differences and moduli of smoothness, analog to the continuous-time setting as in Subsection 1.1.1. Here, the wavelet characterization will prove to be particularly useful.

3. Moreover, we want to extend our results to other scales of discrete-time function spaces, such as Triebel-Lizorkin spaces $F^{\alpha}_{p,q}(\mathbb{Z})$.

Before we start with this program, the upcoming chapter deals with wavelet analysis for $\ell^2(\mathbb{Z})$.

Chapter 2

Wavelet Analysis of Discrete-Domain Signals

This chapter is concerned with wavelet bases in discrete time. The results below are considered to be known, though the discrete-time point of view is treated less often than the continuous-time case.

Anyway, before we are ready to treat the discrete-time theory, we will deal with the continuous-time case in the first section of this chapter. We give a short compendium of wavelet bases for $L^2(\mathbb{R})$ associated to multiresolution analysis, restricting ourselves to some of the main notions and results.

There are two reasons for doing this: The first is that continuous-time wavelets (at least those that arise from a multiresolution analysis) give rise to discrete-time filter banks and associated wavelet systems in $\ell^2(\mathbb{Z})$. Moreover, desirable properties of discrete-time wavelets, vanishing moments and smoothness can be understood best by comparison to the continuous time setting, where they are natural and well-understood.

We discuss the relations between wavelet functions and filter banks in Subsection 2.1.1 and shortly describe certain useful properties of wavelet systems such as vanishing moments and regularity (2.1.2). Subsection 2.1.3 gives a short glimpse on biorthogonal wavelet bases for $L^2(\mathbb{R})$.
This first section will be kept short as it mainly serves to fix notation for the corresponding discrete-time notions and results coming up in the following section. Thus, for proofs and further results, we defer to the usual literature on wavelet theory such as the books by Daubechies [12], Meyer [20], Wojtaszczyk [34] or Mallat[19].

In the next section, 2.2, we switch from continuous to discrete time: we give an overview of wavelet systems in $\ell^2(\mathbb{Z})$ and their properties. The main sources for this are the book by Cohen [3] and the articles by Rioul [27, 26, 28].

Of the various desirable properties of wavelets, the notion of regularity is probably the least intuitive in discrete time: smoothness conditions relying on small-scale limits cannot be imposed on sequences. The notion of regularity as given by Rioul in the two latter papers we mentioned above rather deals with large scale limits. As this regularity condition will be crucial for the proofs in section 3.3, we discuss this property in an extra subsection (2.2.2).

It is also the regularity property which connects the discrete to the continuous theory: discrete time wavelet families possessing some order of regularity converge in a certain sense to continuous-time wavelets of the same regularity. This will be the issue of Subsection 2.2.3.

2.1 Wavelet Bases for $L^2(\mathbb{R})$

In this section, we will recall a few facts on wavelet bases for $L^2(\mathbb{R})$ associated to multiresolution analysis. We want to point out again that we will restrict our discussion to elementary results and refer to the sources we mentioned in the introduction to this chapter.

So, what is a wavelet?

Definition 2.1.1. *A* ***wavelet system*** *for $L^2(\mathbb{R})$ is a family of functions $(\psi_{j,l})_{j,l\in\mathbb{Z}}$ obtained from a single function $\psi \in L^2(\mathbb{R})$ by*

$$\psi_{j,l}(x) = 2^{-j/2}\psi(2^{-j}x - l). \tag{2.1.1}$$

A ***wavelet basis*** *is a wavelet system that is an orthonormal basis for $L^2(\mathbb{R})$.*

Do wavelet bases exist? The answer is yes - as a motivating example, we will treat the most elementary wavelet system, the **Haar wavelet** system.

The Haar wavelet is the function

$$H(x) = \begin{cases} 1 & \text{if } 0 \le x < 1/2; \\ -1 & \text{if } 1/2 \le x < 1; \\ 0 & \text{otherwise.} \end{cases} \tag{2.1.2}$$

We illustrate the concept of multiresolution analysis, which is presented later on in this section, by showing that the family $(H_{j,l})_{j,l\in\mathbb{Z}}$ of dilated and translated versions (see (2.1.1)) of the Haar wavelet is an orthonormal basis for $L^2(\mathbb{R})$.

The functions $H_{j,l}$ are supported on the dyadic intervals $D_{j,l} = [l2^j, (l+1)2^j[$. With this observation, it is obvious that the Haar system is orthonormal:

Consider two of the functions $H_{j,l}$, $H_{j',l'}$.

Either, we have $j = j'$ and $l \neq l'$. In this case, the supports of the functions are disjoint and thereby $\langle H_{j,l}, H_{j',l'}\rangle = 0$. Or, we have, assuming without loss of generality $j > j'$, that the support of $H_{j,l}$ is at least twice as long as the support of $H_{j',l'}$. If the supports are not disjoint, the support of $H_{j',l'}$ is contained in a constant interval of $H_{j,l}$. Again, $\langle H_{j,l}, H_{j',l'}\rangle = 0$.

We want to show that $(H_{j,l})$ is an orthonormal basis, so what is left to prove is totality, i.e. for any $f \in L^2(\mathbb{R})$, $f = \sum_{j\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\langle f, H_{j,l}\rangle H_{j,l}$ in the sense of $L^2(\mathbb{R})$:

We introduce the spaces V_j as the spaces of functions in $L^2(\mathbb{R})$ which are constant on the dyadic intervals $D_{j,l}$, $l \in \mathbb{Z}$. Due to the dyadic structure, the spaces are nested, i.e. $V_{j+1} \subset V_j$, and we have $f(x) \in V_j$ if and only if $f(2^j x) \in V_0$.
Thereby, as the translates of the function $\phi = \chi_{[0,1[}$, $(\phi_l)_{l\in\mathbb{Z}} = (\phi(\cdot - l))_{l\in\mathbb{Z}}$ obviously form an orthonormal basis for the space V_0, consisting of $L^2(\mathbb{R})$-functions constant on $[l, l+1[$, the family $(\phi_{j,l})_{l\in\mathbb{Z}} = (2^{-j/2}\phi(2^{-j}\cdot - l))_{l\in\mathbb{Z}}$ is an orthonormal basis for V_j. The whole family $(\phi_{j,l})_{j,l\in\mathbb{Z}}$ however is not an orthonormal basis for $L^2(\mathbb{R})$. However, we can derive an orthonormal basis, which will be the Haar basis, from these considerations.

For $f \in L^2(\mathbb{R})$, consider the orthogonal projection of f on V_j,

$$P_j f = \sum_{l\in\mathbb{Z}}\langle f, \phi_{j,l}\rangle\phi_{j,l} = \sum_{l\in\mathbb{Z}} f_{D_{j,l}}\chi_{D_{j,l}}, \tag{2.1.3}$$

where $f_{D_{j,l}} = 2^{-j}\int_{x\in D_{j,l}} f(x)dx$. As this projection is the best approximation of f from V_j, and as any function in $L^2(\mathbb{R})$ can be approximated arbitrarily well by functions which are piecewise constant on dyadic intervals, we have
$\lim_{j\to-\infty}\|f - P_j f\|_{L^2(\mathbb{R})} = 0$ as well as $\lim_{j\to-\infty} V_j = L^2(\mathbb{R})$.

Looking at the limit in the opposite direction, $f \in \bigcap_{j\in\mathbb{Z}} V_j$ implies that f is constant on the positive as well as on the negative real line, which for an $L^2(\mathbb{R})$-function implies $f \equiv 0$ and also, $\lim_{j\to\infty}\|P_j f\|_{L^2(\mathbb{R})} = 0$.

Using the above observations, one can write any $L^2(\mathbb{R})$ as $f = \sum_{j\in\mathbb{Z}} P_j f - P_{j+1} f$ and we denote $P_j f - P_{j+1} f$ by $Q_j f$. As P_j and P_{j+1} are orthogonal projections on the spaces V_j and V_{j+1}, respectively, Q_j is the orthogonal projection on $W_j := V_j \ominus V_{j+1}$. The spaces W_j, $j \in \mathbb{Z}$, are mutually orthogonal and we can write

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j. \tag{2.1.4}$$

It may appear that we have lost track of the initial problem, but the Haar wavelets are coming into play again right now. It is easy to see that the family $(H_l)_{l\in Z} = (H(\cdot - l))_{l\in\mathbb{Z}}$ is an orthonormal basis for W_0 and by dilation, $(H_{j,l})_{l\in\mathbb{Z}}$ is an orthonormal basis for W_j.

Therefore, we can write the projection Q_j on the W_j spaces as

$$Q_j f = \sum_{l\in\mathbb{Z}}\langle f, H_{j,l}\rangle H_{j,l}, \tag{2.1.5}$$

and together with (2.1.4), this yields

$$f = \sum_{j,l\in\mathbb{Z}} \langle f, H_{j,l}\rangle H_{j,l}, \tag{2.1.6}$$

so the Haar system is an orthonormal basis for $L^2(\mathbb{R})$.
The spaces $(V_j)_{j\in\mathbb{Z}}$ we considered in this discussion are a first example of a multiresolution analysis, which we will now formally introduce.

Definition 2.1.2. *A **multiresolution analysis (MRA)** is a sequence of closed subspaces $(V_j)_{j\in\mathbb{Z}}$ of $L^2(\mathbb{R})$, such that*

$$V_{j+1} \subset V_j, \tag{2.1.7}$$

$$\lim_{j\to\infty} V_j = \bigcap_{j=-\infty}^{\infty} V_j = \{0\}, \tag{2.1.8}$$

$$\lim_{j\to-\infty} V_j = \overline{\bigcup_{j=-\infty}^{\infty} V_j} = L^2(\mathbb{R}), \tag{2.1.9}$$

$$f(x) \in V_j \text{ if and only if } f(2^j x) \in V_0, \tag{2.1.10}$$

$$\text{there is } \phi(x) \in V_0, \text{ such that } (\phi(x-l))_{l\in\mathbb{Z}} \text{ is an orthonormal basis for } V_0. \tag{2.1.11}$$

*The function $\phi(x)$ is called the **scaling function** for (V_j).*

By (2.1.10) and (2.1.11), the system

$$(\phi_{j,l})_{l\in\mathbb{Z}} = (2^{-j/2}\phi(2^{-j}\cdot -l)))_{l\in\mathbb{Z}}$$

is an orthonormal basis for V_j.
Define the spaces W_j by

$$V_{j-1} = V_j \oplus W_j.$$

An L^2 function ψ is called **wavelet function** associated to (V_j), if $(\psi(\cdot - l))_{l\in\mathbb{Z}}$ is an orthonormal basis for W_0.
In this case,

$$(\psi_{j,l})_{l\in\mathbb{Z}} = (2^{-j/2}\psi(2^{-j}\cdot -l)))_{l\in\mathbb{Z}}$$

is an orthonormal basis for W_j.
In addition, we have

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j$$

and the system $(\psi_{j,l})_{j,l\in\mathbb{Z}}$ constitutes an orthonormal basis for $L^2(\mathbb{R})$.
Every MRA has an associated wavelet, see the next subsection.

There are at least two ways to get an MRA, yielding wavelet bases as a consequence. As a first way, one could start by defining the spaces (V_j) and try to find a scaling function ϕ, such that the translates form an orthonormal basis for V_0. Here the standard example is provided by the spline spaces (e.g. [19], Ex. 7.3),

Instead, one may also not that (2.1.10) and (2.1.11) imply that every MRA is uniquely determined by its scaling function ϕ, since V_0 is the closed span of translates of ϕ, and V_j is obtained from V_0 by dilation.

Most of the MRAs used nowadays are constructed this way, i.e., by picking a suitable ϕ, a task which is however highly nontrivial.

For $f \in L^2(\mathbb{R})$, let P_jf, Q_jf be the orthogonal projections on V_j and W_j, respectively. In particular, we can write, analogously to (2.1.3) and (2.1.5),

$$P_jf = \sum_{l\in\mathbb{Z}} \langle f, \phi_{j,l}\rangle \phi_{j,l}, \tag{2.1.12}$$

$$Q_jf = \sum_{l\in\mathbb{Z}} \langle f, \psi_{j,l}\rangle \psi_{j,l}. \tag{2.1.13}$$

Using the above projections, we can rewrite condition (2.1.8) as $\lim_{j\to\infty} \|P_jf\|_{L^2(\mathbb{R})} = 0$. Condition (2.1.9) can be replaced by $\lim_{j\to-\infty} \|f - P_jf\|_{L^2(\mathbb{R})} = 0$.

The projections on the spaces V_j can be interpreted as approximations of f at different resolutions, whereas the partial wavelet series Q_jf can be viewed as the difference between two approximation levels.

Note that our notation is different from most sources in the wavelet literature: The scale j corresponds to details of size 2^j. Thus, the spaces **decrease** as j increases. We made this change to avoid dealing with negative indices when considering discrete expansions later on: for discrete functions, resolution is obviously limited, so in our terms, there will be no scale smaller than 1 in this case.

We finish this subsection with two examples.

Example 2.1.3. : Haar MRA
As we already proved in the beginning of this paragraph, the spaces of functions in $L^2(\mathbb{R})$ which are constant on the dyadic intervals $D_{j,l}$, $l \in \mathbb{Z}$, are a multiresolution analysis with scaling function $\phi = \chi_{[0,1[}$. The corresponding wavelet basis is the Haar wavelet basis.

Example 2.1.4. : Shannon MRA
As a kind of an opposite extreme to the Haar MRA, one can define an MRA as the spaces of functions which are band-limited, i.e with their Fourier transform supported on the intervals $[-2^{-j}\pi, 2^{-j}\pi]$. The associated scaling function is the sinc function $\phi(x) = \frac{sin\pi x}{\pi x}$ and we have $\psi(x) = 2\,\mathrm{sinc}(2x) - \mathrm{sinc}(x)$.

The MRAs given above inherit different properties:

First, the Haar wavelet is compactly supported and thereby well-localized in time, whereas the Shannon wavelet, being compactly supported in frequency, has poor decay in time domain.
Second, the Haar wavelet is discontinuous, whereas the Shannon wavelet is C^∞.

Third, the Haar wavelet has precisely one vanishing moment, whereas, at least formally, all moments of the Shannon wavelet vanish.

There are certainly examples of MRAs between these extremes concerning localization in time and frequency domain. The Daubechies family, starting with the Haar MRA, can be seen as a family of MRAs that represent different compromises between the two extremes: For each N, the Daubechies construction yields a wavelet of regularity order roughly $N/5$, with support in an interval of length $2N-1$, and N vanishing moments. Thus, desirable properties (smoothness, vanishing moments) have to be 'payed for' in terms of support size. This discussion is skipped until we consider more in detail useful properties of wavelet systems in Subsection 2.1.2.

2.1.1 Wavelets and Filters

In the last subsection we described how to derive wavelet bases from multiresolution analyses. We saw that the scaling function fully determines an MRA.

There is yet another way to characterize MRAs, by use of discrete time filters: the multiresolution conditions yield a discretization we discuss in this subsection. We will see that this fact also gives rise to a fast algorithm to compute the wavelet transform.

The translates of a scaling function associated to a multiresolution analysis (V_j) form an orthonormal basis of V_0 (2.1.11), and as $2^{-1/2}\phi(x/2) \in V_1 \subset V_0$ by (2.1.10), (2.1.7), we obtain the **scaling equation**

$$2^{-1/2}\phi(x/2) = \sum_{n\in\mathbb{Z}} g(n)\phi(x-n) \tag{2.1.14}$$

with scaling coefficients $g(n) = \langle 2^{-1/2}\phi(x/2), \phi(x-n)\rangle$.

The scaling equation gives rise to a number of interesting equations fulfilled by the scaling function and its Fourier transform:

In Fourier domain, the orthonormality condition becomes

$$\sum_{l\in\mathbb{Z}} |\hat{\phi}(\omega + 2\pi l)|^2 = 1 \text{ for almost all } \omega \in \mathbb{R}. \tag{2.1.15}$$

Using the Fourier transform of (2.1.14), we have

$$\begin{aligned} 2 &= 2\sum_{l\in\mathbb{Z}} |\hat{\phi}(2\omega+2\pi l)|^2 = \sum_{l\in\mathbb{Z}} |\hat{\phi}(\omega+\pi l)|^2|\hat{g}(\omega+\pi l)|^2 \\ &= \sum_{l\in\mathbb{Z}} |\hat{\phi}(\omega+2\pi l)|^2|\hat{g}(\omega)|^2 + \sum_{l\in\mathbb{Z}} |\hat{\phi}(\omega+\pi+2\pi l)|^2|\hat{g}(\omega+\pi)|^2, \end{aligned}$$

which finally yields

$$|\hat{g}(\omega)|^2 + |\hat{g}(\omega+\pi)|^2 = 2 \quad \text{(a.e.)}. \tag{2.1.16}$$

Furthermore, say for $\hat{g}$ continuous, we have

$$\hat{g}(0) = \sqrt{2} \text{ and thereby } \hat{g}(\pi) = 0, \tag{2.1.17}$$

so we can regard $\widehat{g}$ as a **low pass filter**, as frequencies around $\omega = \pi$ are attenuated and those near $\omega = 0$ are kept by applying this function.

Conditions (2.1.16) and (2.1.17) are necessary conditions on the filter $(g(n))$, such that the function ϕ in (2.1.14) is a scaling function of an MRA. In a sense, they contain a recipe for constructing ϕ from $(g(n))$:

Iterating (2.1.14) yields

$$\hat{\phi}(\omega) = \prod_{k=1}^{n} \hat{g}(2^{-k}\omega)\hat{\phi}(2^{-n}\omega). \tag{2.1.18}$$

Hence we may start by taking a 2π-periodic $\hat{g}$, satisfying (2.1.16) and (2.1.17), and considering

$$\hat{\phi}(\omega) = \prod_{k=1}^{\infty} \hat{g}(2^{-k}\omega). \tag{2.1.19}$$

This construction has been successfully employed for constructing scaling functions. Note however that without additional assumptions on g, the above product will not necessarily converge to a scaling function associated to an MRA (see [3]).

We will return to this discussion when we deal with discrete-time MRAs in Section 2.2; there we will impose additional conditions on the filter (referred to as 'discrete-time regularity') which will ensure convergence of the scheme (2.1.19) to a (then also regular) scaling function, see especially Section 2.2.3.

Also, the wavelet functions are related to discrete time filters. For a wavelet function ψ, necessarily $2^{-1/2}\psi(x/2) \in W_1 \subset V_0$, yielding the **wavelet equation**

$$2^{-1/2}\psi(x/2) = \sum_{n\in\mathbb{Z}} h(n)\phi(x-n) \tag{2.1.20}$$

with coefficients $h(n) = \langle 2^{-1/2}\psi(x/2), \phi(x-n)\rangle$.

Similar computations to the above give that these coefficients satisfy

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega+\pi)|^2 = 2 \quad \text{(a.e.)}, \tag{2.1.21}$$

and

$$\hat{h}(0) = 0 \text{ and } \hat{h}(\pi) = \sqrt{2}, \tag{2.1.22}$$

(in the case where this pointwise statement makes sense) such that $\widehat{h}$ can be viewed as a discrete-time **high pass filter**.

The condition that the spaces W_0 and V_0 are orthogonal to each other can be expressed in terms of the filters by

$$\hat{h}(\omega)\overline{\hat{g}(\omega)} + \hat{h}(\omega+\pi)\overline{\hat{g}(\omega+\pi)} = 0. \tag{2.1.23}$$

It can be proven ([3, 18, 20]) that if ϕ is a scaling function associated to a multiresolution analysis (V_j) with corresponding g, then, the special choice

$$h(n) = (-1)^{1-n} g(1-n) \tag{2.1.24}$$

or, equivalently,

$$\hat{h}(\omega) = \overline{\hat{g}(\omega+\pi)} e^{-i\omega}, \tag{2.1.25}$$

in (2.1.20) gives a wavelet, whose translates are an orthonormal basis of W_0.

A filter pair (g, h) satisfying (2.1.16), (2.1.21) and (2.1.23) will be called **perfect reconstruction (PR) filter pair**. This name comes from the fact that the relation between wavelets and filters admits a fast algorithm for the wavelet transform. Before we discuss this algorithm in more detail, we revisit our examples:

Example 2.1.5. : Haar filters

For the multiresolution analysis with scaling function $\phi = \chi_{[0,1[}$, the corresponding low pass filter reads $g(n) = 2^{-1/2}$ for $n = 0, 1$ and 0 otherwise.

By (2.1.24), $h(0) = -g(1) = -2^{-1/2}$ and $h(1) = g(0) = 2^{-1/2}$.

Example 2.1.6. : Shannon filters

Considering the Shannon multiresolution approximation with scaling function $\hat{\phi} = \chi_{[-\pi,\pi]}$, the filters are given by $\hat{g} = 2^{1/2}\chi_{[-\pi/2,\pi/2]}$ and $\hat{h} = 2^{1/2} - \hat{g}$.

So far, MRAs have served mainly as a tool for the convenient construction of wavelet bases. Here the scaling and wavelet equation appeared as byproducts of the inclusion properties of an MRA. We have already seen, however, that the scaling equation may also serve as the starting point for the construction of an MRA, via (2.1.19). Similarly, the scaling and the wavelet equation will also serve as the source of the chief algorithmic contribution, the **fast wavelet transform**. This algorithm allows to compute coarse scale wavelet coefficients of a signal by repeated application of discrete convolution and subsampling steps.

Let $a_j = (a_j(l))_{l\in\mathbb{Z}} = (\langle f, \phi_{j,l}\rangle)_{l\in\mathbb{Z}}$. We will call the sequence a_j the **approximation coefficients** at scale $j \in \mathbb{Z}$.
Let $d_j = (d_j(l))_{l\in\mathbb{Z}} = (\langle f, \psi_{j,l}\rangle)_{l\in\mathbb{Z}}$ be the wavelet coefficients, which will also be called the **detail coefficients**, at scale j.

Proposition 2.1.7. [Mallat]
The coefficients can be recursively computed by the filtering

$$a_{j+1}(l) = (a_j * g^*)(2l) = \sum_{k\in\mathbb{Z}} a_j(k) g^*(2l-k), \tag{2.1.26}$$

$$d_{j+1}(l) = (a_j * h^*)(2l) = \sum_{k\in\mathbb{Z}} a_j(k) h^*(2l-k). \tag{2.1.27}$$

The reconstruction of a filtering step is done by

$$a_j(l) = \sum_{k\in\mathbb{Z}} a_{j+1}(k) g(l-2k) + \sum_{k\in\mathbb{Z}} d_{j+1}(k) h(l-2k) = (\uparrow_2 a_{j+1}) * g + (\uparrow_2 d_{j+1}) * h, \tag{2.1.28}$$

the sum of coefficients on a coarser scale, upsampled and convolved with the filters g, h.

Proof Remember that $\phi_{j+1,l} \in V_{j+1} \subset V_j$. Expanding $\phi_{j+1,l}$ in the orthonormal basis $(\phi_{j,k})_{k\in\mathbb{Z}}$ of V_j yields

$$\phi_{j+1,l} = \sum_{k\in\mathbb{Z}} \langle \phi_{j+1,l}, \phi_{j,k}\rangle \phi_{j,k}. \tag{2.1.29}$$

Computing the inner products $\langle \phi_{j+1,l}, \phi_{j,k}\rangle$ gives

$$\begin{aligned} \phi_{j+1,l} &= \sum_{k\in\mathbb{Z}} \langle 2^{-1/2}\phi(\cdot/2), \phi(\cdot - k + 2l)\rangle \phi_{j,k} \\ &= \sum_{k\in\mathbb{Z}} g(k-2l)\phi_{j,k}. \end{aligned}$$

Employing the inner product on both sides of (2.1.29) thereby yields

$$a_{j+1,l} = \sum_{k\in\mathbb{Z}} a_{j,k} \overline{g(k-2l)} = (a_j * g^*)(2l),$$

which is (2.1.26).
(2.1.27) follows from analogous computations.

For the reconstruction step, reconsider the fact that $V_j = V_{j+1} \oplus W_{j+1}$. Thereby, $\phi_{j,l} \in V_j$ can be expanded in the union of $(\phi_{j,k})_{k\in\mathbb{Z}}$ and $(\psi_{j,k})_{k\in\mathbb{Z}}$, which is an orthonormal basis for V_j:

$$\phi_{j,l} = \sum_{k\in\mathbb{Z}} \langle \phi_{j,l}, \phi_{j+1,k}\rangle \phi_{j+1,k} + \sum_{k\in\mathbb{Z}} \langle \phi_{j,l}, \psi_{j+1,k}\rangle \psi_{j+1,k}. \tag{2.1.30}$$

The above considerations give

$$\phi_{j,l} = \sum_{k\in\mathbb{Z}} \overline{g(l-2k)}\phi_{j+1,k} + \sum_{k\in\mathbb{Z}} \overline{h(l-2k)}\psi_{j+1,k}. \tag{2.1.31}$$

Taking the inner product on both sides,

$$a_{j,l} = \sum_{k\in\mathbb{Z}} a_{j,k} g(l-2k) + \sum_{k\in\mathbb{Z}} d_{j+1} h(l-2k).$$

■

Iterating (2.1.26), (2.1.27), the fast wavelet transform computes the map $(a_{0,l})_{l\in\mathbb{Z}} \mapsto (d_{j,l})_{j>0,l\in\mathbb{Z}}$ via the cascade

$$\begin{array}{ccccccccccc} a_0 & \rightarrow & a_1 & \rightarrow & a_2 & \rightarrow & \dots & \rightarrow & a_{j-1} & \rightarrow & a_j \\ & \searrow & & \searrow & & \searrow & & \searrow & & \searrow & \\ & & d_1 & & d_2 & & & & d_{j-1} & & d_j \end{array} \tag{2.1.32}$$

where each horizontal arrow represents the same filtering and subsampling step $a_{j+1} = \downarrow_2 (a_j * g^*)$, and similarly, $d_{j+1} = \downarrow_2 (a_j * h^*)$.

2.1.2 Properties of Wavelet Bases

In the preceding subsection, we dealt in a general way with wavelet bases and their connection to multiresolution analysis and discrete-time filters.
The above-cited results concerning characterizations of Besov spaces via wavelet coefficients (1.1.4) rely on additional properties, that is to say vanishing moments, smoothness and compact support, which we now attend to.

The first property we deal with is the notion of vanishing moments.

Definition 2.1.8. *ψ has **vanishing moments** of order $N \in \mathbb{N}$ if*

$$\int t^k \psi(t)dt = 0 \quad \textit{for} \quad 0 \le k \le N-1.$$

This property ensures that ψ is orthogonal to polynomials of order $N-1$. Consider a function f which is $k < N$ times continuously differentiable around x_0 . Then, f

can be expanded into a Taylor polynomial over an interval. As ψ cancels out this polynomial, we have for those $\psi_{j,l}$, whose supports are contained in the neighborhood around x_0, that the absolute values of coefficients $\langle f, \psi_{j,l}\rangle$ decay for small scales. Hence, local smoothness leads to a decay of coefficients, piecewise smooth functions can be well approximated by a few number of wavelet coefficients.

If f has a singularity in x_1, all the wavelets $\psi_{j,l}$ which have x_1 inside their support will feel this singularity and may have a large coefficient. Hence, another desirable property is to deal with wavelets with good **decay properties**, optimally with ones of **compact support**.

A further desirable property analyzing smooth functions is to use wavelets which also possess some **regularity** of a certain order. We say that a wavelet is regular of order $r > 0$ if it is Hölder regular of order r.

Remark 2.1.9.

1. The above properties are not independent, for example, a wavelet with a certain decay and regularity will have a certain order of zero moments.
2. There is an alternative interpretation of smoothness and vanishing moment properties for a wavelet ψ concerning the localization in Fourier domain:

   Vanishing moments of order N describes the decay of $\widehat{\psi}(\omega)$ as $\omega \to 0$: $|\widehat{\psi}(\omega)| = O(|\omega|^N)$, whereas smoothness of order r gives $|\widehat{\psi}(\omega)| = O(|\omega|^{-r})$ as $\omega \to \infty$. These observations allow to read the conditions of Theorem 1.1.10 (and also of Theorem 3.3.8) as 'wavelets provide a reasonable approximation of φ-functions'.
3. Note that the properties we discussed can usually be built into the wavelets by designing suitable filters. Daubechies [6] constructed wavelets which have vanishing moments of arbitrary order and are at the same time of minimal, compact support.

   By construction, the Daubechies wavelets with $N \in \mathbb{N}$ vanishing moments have a support length of $2N-1$ and are for large N approximately regular of order $\lfloor 0.2N \rfloor$. This will be of importance in the upcoming section (2.2), when we deal with discrete-time wavelets: at least compact support and vanishing moments carry over immediately to this setting. The connection between filters with certain properties leading to regular wavelets gets more clear when we deal with discrete-time multiresolution analysis.

We finish this subsection by looking again at our examples:

Example 2.1.10. : Properties of Haar Wavelets The Haar wavelet has compact support, possesses one vanishing moment and is in fact the Daubechies wavelet of order 1. It is not continuous and thereby non-regular.

Example 2.1.11. : Properties of Shannon Wavelets The Shannon wavelet is compactly supported in Fourier domain and thus has a poor decay in time. In contrast to the Haar wavelet, it is C^∞ and at least formally, all of its moments vanish.

2.1.3 Biorthogonal Bases

The characterization of Besov spaces can be extended to biorthogonal wavelet systems, which is why we shortly discuss these systems here. For proofs of the results cited below, as well as further details, we refer the reader to [12, 19].

A biorthogonal pair of wavelet bases is a pair of Riesz bases $(\psi_{j,l})_{j,l}$ and $(\tilde{\psi}_{j,l})_{j,l}$ arising in the usual manner from functions $\psi, \tilde{\psi}$, and fulfilling the **biorthogonality condition**

$$\langle \psi_{j,l}, \tilde{\psi}_{j,l} \rangle = \delta_{l,l'} \delta_{j,j'} \quad .$$

This relation, together with the Riesz base properties of the systems, immediately entails the expansions

$$f = \sum_{j,l} \langle f, \psi_{j,l} \rangle \tilde{\psi}_{j,l} = \sum_{j,l} \langle f, \tilde{\psi}_{j,l} \rangle \psi_{j,l} \quad .$$

Clearly, this concept generalizes wavelet orthonormal bases. It turns out that a convenient method for the construction of such bases is provided by introducing biorthogonality to multiresolution analysis: Instead of a single MRA, one constructs a pair $(V_j)_j, (\tilde{V}_j)_j$ of sequences of spaces, which have all properties of MRAs except for (2.1.11), which is replaced by functions ϕ,$\tilde{\phi}$ satisfying the requirements

$$\begin{aligned}
&(\phi(\cdot - l))_{l \in \mathbb{Z}} \text{ is a Riesz basis for } V_0 \\
&(\tilde{\phi}(\cdot - l))_{l \in \mathbb{Z}} \text{ is a Riesz basis for } \tilde{V}_0 \\
&\langle \phi(\cdot - l), \tilde{\phi}(\cdot - l') \rangle = \delta_{l,l'} \quad .
\end{aligned}$$

Defining W_j and $\tilde{W}_j$ as orthogonal complements, just as in the orthonormal wavelet case, one can prove the existence of wavelets $\psi, \tilde{\psi}$ satisfying

$$\begin{aligned}
&(\psi(\cdot - l))_{l \in \mathbb{Z}} \text{ is a Riesz basis for } W_0 \\
&(\tilde{\psi}(\cdot - l))_{l \in \mathbb{Z}} \text{ is a Riesz basis for } \tilde{W}_0 \\
&\langle \psi(\cdot - l), \tilde{\psi}(\cdot - l') \rangle = \delta_{l,l'} \quad .
\end{aligned}$$

which entails that $(\tilde{\psi}_{j,l})_{j,l}, (\psi_{j,l})_{j,l}$ are biorthogonal wavelet bases.

The drawback of biorthogonality is that the Parseval relation

$$\|f\|^2 = \sum_{j,l} |\langle f, \psi_{j,l} \rangle|^2$$

holding for orthonormal bases needs to be replaced by the norm equivalences

$$\|f\|^2 \asymp \sum_{j,l} |\langle f, \psi_{j,l} \rangle|^2 \asymp \sum_{j,l} |\langle f, \tilde{\psi}_{j,l} \rangle|^2 \quad .$$

The chief advantage of biorthogonality is higher flexibility in the choice of wavelets: For instance, one can choose symmetric wavelets (which is impossible in the orthogonal setting), or one can distribute desirable properties between ψ and $\tilde{\psi}$: ψ can be chosen with a desired number of vanishing moments (but little regularity), and $\tilde{\psi}$ with a desired degree of smoothness.

The fast wavelet transform easily adapts to the biorthogonal setting; the only change being that one now uses one filter pair g, h for the decomposition, and a different pair $\tilde{g}, \tilde{h}$ for reconstruction.

Coefficients $a_j = (a_j(l))_{l \in \mathbb{Z}} = (\langle f, \tilde{\phi}_{j,l} \rangle)_{l \in \mathbb{Z}}$ on a certain scale $j \in \mathbb{Z}$ are used to compute approximation coefficients a_{j+1} and detail coefficients $d_{j+1} = (d_{j+1}(l))_{l \in \mathbb{Z}} = (\langle f, \tilde{\psi}_{j+1,l} \rangle)_{l \in \mathbb{Z}}$ on a coarser scale by convolution with discrete-time low and high pass filters $\tilde{g}, \tilde{h}$ associated to $\tilde{\phi}$, $\tilde{\psi}$, followed by subsampling:

$$a_{j+1}(l) = (a_j * \tilde{g})(2l) = \sum_{k \in \mathbb{Z}} a_j(k) \tilde{g}(2l - k), \tag{2.1.33}$$

$$d_{j+1}(l) = (a_j * \tilde{h})(2l) = \sum_{k \in \mathbb{Z}} a_j(k) \tilde{h}(2l - k). \tag{2.1.34}$$

The reconstruction of a filtering step is done by

$$a_j(l) = \sum_{k \in \mathbb{Z}} a_{j+1}(k) g(l - 2k) + \sum_{k \in \mathbb{Z}} d_{j+1}(k) h(l - 2k), \tag{2.1.35}$$

the sum of coefficients on a coarser scale, upsampled and convolved with filters g, h dual to $\tilde{g}, \tilde{h}$.

Note that Besov spaces have a characterization in terms of biorthogonal wavelets as well: Theorem 1.1.10 can be formulated using biorthogonal bases, where vanishing moments and regularity conditions are imposed separately for the analyzing and synthesizing wavelets, see e.g. [14].

In the discrete-time case we treat in this thesis, we will also give this generalized result, see Theorem 3.3.8, which obviously includes the orthonormal case.

2.2 Discrete-Time Wavelet Bases for $\ell^2(\mathbb{Z})$

In this section, we will finally describe wavelet systems in $\ell^2(\mathbb{Z})$. These systems are associated to the filterbank algorithms which we described in 2.1.1.

Recall that the fast wavelet transform computes the map $(a_{0,l})_{l\in\mathbb{Z}} \mapsto (d_{j,l})_{j>0,l\in\mathbb{Z}}$ via the cascade (2.1.32).

The algorithm implements a unitary operator: $(a_{0,l})_{l\in\mathbb{Z}}$ and $(d_{j,l})_{j>0,l\in\mathbb{Z}}$ are the expansion coefficients of

$$F = \sum_{l\in\mathbb{Z}} a_{0,l}\phi(\cdot - l) \in V_0$$

in the ONBs $(\phi_{0,l})_{l\in\mathbb{Z}}$ and $(\psi_{j,l})_{j>0,l\in\mathbb{Z}}$ of V_0.

Discrete time wavelet bases are obtained by viewing $a_0 \in \ell^2(\mathbb{Z})$ as input to a unitary transform $W_d : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{N}\times\mathbb{Z})$. The output can then be interpreted as expansion coefficients of a_0 in the system $h_{j,l} = W_d^{-1}(\delta_{j,l})$, the preimage of the Kronecker ONB of $\ell^2(\mathbb{N}\times\mathbb{Z})$ under W_d. Hence $(h_{j,l})_{j,l}$ is an ONB of $\ell^2(\mathbb{Z})$. We intend to study bases of this kind, with the aim of describing signals with good approximation behavior.

This perspective may seem unorthodox, but it is in fact closely related to the way that wavelets are used on real-world data: We have repeatedly remarked that these data are usually given discretely and that the standard procedure feeds the discrete data $(f(k))_k$ directly into the wavelet filterbank. As a consequence, the filterbank output consists of wavelet coefficients $d_{j,l} = \langle F, \psi_{j,l}\rangle$, where $F = \sum_{l\in\mathbb{Z}} f(l)\phi(\cdot - l)$. The problem with this procedure is that it uses the scaling function ϕ, which is not known explicitly. Accordingly, the development of model assumptions on F, which could serve as a source of heuristics for signal processing algorithms, becomes a rather difficult tasks.

As a matter of fact, quite often these signal models are available for f instead of F, say, f is obtained from a measuring device with certain noise characteristics, and certain expected smoothness behaviour in the measured quantity. Having a fully discrete time theory available should thus allow to describe and analyze wavelet-based processing algorithms in a more transparent and direct way than via the embedding into $L^2(\mathbb{R})$, which is obscured by the scaling function.

In the following, we will therefore discard any reference to the continuous-time setting, and describe wavelet systems in discrete time, which arise from a pair g, h of perfect reconstruction filters, and the associated cascade (or fast wavelet transform) algorithm $\ell^2(\mathbb{Z}) \to \ell^2(\mathbb{N}\times\mathbb{Z})$, as objects of independent interest. Our exposition of discrete-time wavelets uses ideas and results from A. Cohen's book [3] and O. Rioul's papers [27, 28, 26].

2.2.1 The Discrete-Time Wavelet Transform

In this subsection, we will be concerned with the construction of operators analogous to the cascade algorithm $\ell^2(\mathbb{Z}) \to \ell^2(\mathbb{N} \times \mathbb{Z})$, but without a continuous time MRA in the background. More precisely, we will consider pairs g, h of filters, and the following objects constructed from g, h.

1. One-step decomposition operators $a_0 \mapsto (a_1, d_1) = (\downarrow_2 (a_0 * g^*), \downarrow_2 (a_0 * h^*))$ on $\ell^2(\mathbb{Z})$.

2. Full decomposition operators $a_0 \mapsto (d_j)_{j \geq 1}$ obtained by cascading the one-step decomposition. Of particular interest will be conditions on g, h making this operator unitary.

3. Provided that the fast wavelet transform $a_0 \mapsto (d_j)_{j \geq 1}$ is unitary, its output can be understood as expansion coefficients of the input signal with respect to an ONB of $\ell^2(\mathbb{Z})$. This will be the discrete-time wavelet ONB, and we are looking for explicit descriptions of this basis.

One Step Decomposition

Clearly, a necessary condition for the fast wavelet transform to be unitary is that the one-step decomposition is unitary. The following theorem gives a precise condition for this.

Theorem 2.2.1. *Given two sequences $g, h \in \ell^2(\mathbb{Z})$, consider the operator*

$$\begin{array}{rcl} S : \ell^2(\mathbb{Z}) & \to & \ell^2(\mathbb{Z}) \times \ell^2(\mathbb{Z}) \\ f & \mapsto & (\downarrow_2 (f * g^*), \downarrow_2 (f * h^*)) \end{array}$$

Then the following are equivalent:

(i) S is unitary.

(ii) The system $(g(\cdot - 2l))_{l \in \mathbb{Z}} \cup (h(\cdot - 2l))_{l \in \mathbb{Z}}$ is an ONB of $\ell^2(\mathbb{Z})$.

(iii) The Fourier transforms of g and h fulfill the perfect reconstruction (PR) conditions (2.1.16), (2.1.21), (2.1.23), i.e., for almost every $\omega \in \mathbb{R}$

$$|\hat{g}(\omega)|^2 + |\hat{g}(\omega + \pi)|^2 = 2, \tag{2.2.1}$$

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega + \pi)|^2 = 2, \tag{2.2.2}$$

$$\hat{h}(\omega)\overline{\hat{g}(\omega)} + \hat{h}(\omega + \pi)\overline{\hat{g}(\omega + \pi)} = 0. \tag{2.2.3}$$

If the equivalent conditions are fulfilled, the inverse operator is given by

$$S^* : (a,d) \mapsto (\uparrow_2 a) * g + (\uparrow_2 d) * h \quad .$$

Proof The equivalence of (i) and (ii) is obvious. For the implication $(ii) \Rightarrow (iii)$ assume that $(g(\cdot - 2l))_{l\in\mathbb{Z}} \cup (h(\cdot - 2l))_{l\in\mathbb{Z}}$ is an ONB. Then in particular

$$\begin{aligned} \delta_{0,l} &= \langle g(\cdot - 2l), g(\cdot)\rangle \\ &= \int_0^{2\pi} \hat{g}(\omega)e^{-i\omega 2l}\overline{\hat{g}(\omega)}d\omega \\ &= \int_0^{2\pi} |\hat{g}(\omega)|^2 e^{-i\omega 2l} d\omega \\ &= \int_0^{\pi} e^{-i\omega 2l}\left(|\hat{g}(\omega)|^2 + |\hat{g}(\omega+\pi)|^2\right) d\omega \quad . \end{aligned}$$

Hence the integrable function

$$\omega \mapsto |\widehat{g}(\omega)|^2 + |\widehat{g}(\omega+\pi)|^2$$

on $[0,\pi]$ has the same Fourier coefficients as the constant function $\omega \mapsto 2$, and the Fourier uniqueness theorem implies

$$|\widehat{g}(\omega)|^2 + |\widehat{g}(\omega+\pi)|^2 = 2 \quad \text{(a.e.)}.$$

Similar calculations prove

$$\begin{aligned} |\widehat{h}(\omega)|^2 + |\widehat{h}(\omega+\pi)|^2 &= 2 \qquad \text{(a.e.)} \\ \widehat{h}(\omega)\overline{\widehat{g}(\omega)} + \widehat{h}(\omega+\pi)\overline{\widehat{g}(\omega+\pi)} &= 0 \qquad \text{(a.e.)} \quad , \end{aligned}$$

which are the (PR) conditions.

The converse is shown similarly.

Note however that for $(ii) \Rightarrow (iii)$ we only needed that $(g(\cdot - 2l))_{l\in\mathbb{Z}} \cup (h(\cdot - 2l))_{l\in\mathbb{Z}}$ is orthonormal. Somewhat remarkably, orthonormality of the system already implies its completeness. ∎

In (2.1.1), we already encountered the (PR) conditions: in the construction of multiresolution analyses on $L^2(\mathbb{R})$, the perfect reconstruction property is a well-known condition. In order to properly appreciate it, recall the convolution theorem: $(f * g)^\wedge = \widehat{f} \cdot \widehat{g}$. Later on, we will choose filters h, g, such that $\widehat{h}(0) = 0$. Then (PR) entails that $\widehat{h}(\pi) = \sqrt{2}$, and consequently $\widehat{g}(0) = \sqrt{2}$ and $\widehat{g}(\pi) = 0$, which are exactly conditions (2.1.22) and (2.1.17).

Thus, using filters from a (continuous-time) MRA, we immediately obtain a discrete-time ONS.

Recall that the (PR) conditions in the continuous-time case were *necessary* conditions for the filters to be related to an MRA by the scaling and wavelet equation.

We just proved that in the discrete-time case, these conditions are also *sufficient* to yield an ONB, at least in the 'one-step' case above. In the next paragraph, we will see that (at least for finite filters) this is also the case for the 'full decomposition', leading to discrete-time orthonormal wavelet bases.

Orthonormal Wavelet Bases for $\ell^2(\mathbb{Z})$

The discrete time wavelet transform is now obtained by iterating the one step decomposition.

In the following theorem, the finite support condition can be replaced by the weaker condition that $\widehat{g}$ is infinitely differentiable [3]. In any case, the conditions for the existence of an orthonormal wavelet basis for $\ell^2(\mathbb{Z})$ are much less restrictive than for the existence of an associated wavelet system in $L^2(\mathbb{R})$; see [3] for examples of discrete-time wavelet systems that do not arise from an MRA in $L^2(\mathbb{R})$.

Theorem 2.2.2. Wavelet-ONB in $\ell^2(\mathbb{Z})$
Let $g, h \in \ell^2(\mathbb{Z})$ be given with (PR). Assume in addition that g is finitely supported. Given $f \in \ell^2(\mathbb{Z})$, define inductively

$$a_0 = f \quad , \quad a_{j+1} = \downarrow_2 (a_j * g^*) \quad , \quad d_{j+1} = \downarrow_2 (a_j * h^*) \quad .$$

(a) *The* **discrete wavelet transform**

$$W_d : f \mapsto (d_j(l))_{j \geq 1, l \in \mathbb{Z}}$$

is a unitary operator $\ell^2(\mathbb{Z}) \to \ell^2(\mathbb{N} \times \mathbb{Z})$.

(b) *The sequences d_j and a_j consist of* expansion coefficients*: $d_j(l) = \langle f, h_j(\cdot - 2^j l)\rangle$ and $a_j(l) = \langle f, g_j(\cdot - 2^j l)\rangle$, with suitable $h_j, g_j \in \ell^2(\mathbb{Z})$ (for $j \geq 1$).*

(c) *h_j and g_j can be computed recursively via*

$$\begin{aligned} g_0 &= \delta_0 \\ g_{j+1} &= g_j * (\uparrow_{2^j} g) \\ h_{j+1} &= g_j * (\uparrow_{2^j} h) \end{aligned}$$

By construction, the operator W_d computes the coefficients of f with respect to the **discrete-time wavelet basis** *$(h_{j,l}) = (h_j(\cdot - 2^j l))_{j \in \mathbb{N}, l \in \mathbb{Z}}$. This system is an ONB of $\ell^2(\mathbb{Z})$.*

Proof We first prove (b) and (c) by induction:
Noting that $a_0 = f * \delta_0$, we find in the induction step

$$\begin{aligned} a_{j+1}(k) &= \downarrow_2 (a_j * g^*)(l) \\ &= \sum_{n\in\mathbb{Z}} a_j(n) g^*(2l-n) \\ &\overset{IH}{=} \sum_{n\in\mathbb{Z}} \langle f, g_j(\cdot - 2^j n)\rangle g^*(2l-n) \\ &= \langle f, \sum_{n\in\mathbb{Z}} \overline{g^*(2l-n)} g_j(\cdot - 2^j n)\rangle \quad . \end{aligned}$$

Now we can compute

$$\begin{aligned} \sum_{n\in\mathbb{Z}} \overline{g^*(2l-n)} g_j(m-2^j n) &= \sum_{n\in\mathbb{Z}} g_j(m - 2^j(n+2l)) g(n) \\ &= g_{j+1}(m - 2^{j+1} l), \end{aligned}$$

where g_{j+1} is given as

$$\begin{aligned} g_{j+1}(m) &= \sum_{n\in\mathbb{Z}} g_j(m-2^j n) g(n) \\ &= \sum_{n\in\mathbb{Z}} g_j(m-n)(\uparrow_{2^j} g^*)(n) \\ &= (g_j * (\uparrow_{2^j} g))(m) \quad . \end{aligned}$$

Replacing g by h in the calculations yield the formula for h_j, and we have shown (b) and (c).
Now Theorem 2.2.1 implies that the mapping

$$f \mapsto (a_j, d_j, d_{j-1}, d_{j-2}, \ldots, d_1)$$

is unitary, and hence the family

$$(g_{j,l})_{l\in\mathbb{Z}} \cup (h_{i,l})_{1\le i\le j, l\in\mathbb{Z}}$$

is an ONB of $\ell^2(\mathbb{Z})$, where $g_{j,l} = g_j(\cdot - 2^j l)$. Since this holds for all $j \ge 1$, we obtain in particular that $(h_{i,l})_{1\le i\le j, l\in\mathbb{Z}}$ is an ONS in $\ell^2(\mathbb{Z})$. Hence the only missing property is totality.
For this purpose define analogously to (2.1.12)

$$P_j f := \sum_{l\in\mathbb{Z}} \langle f, g_{j,l}\rangle g_{j,l} \quad ,$$

which are the projections onto the orthogonal complement

$$\left((h_{i,l})_{1\le i\le j, l\in\mathbb{Z}}\right)^{\perp} \quad .$$

Hence we need to prove $P_j f \to 0$, for all $f \in \ell^2(\mathbb{Z})$. Since the space of finitely supported sequences is dense in $\ell^2(\mathbb{Z})$, it is enough to prove $P_j f \to 0$ for finite sequences. For this purpose we need two auxiliary statements

- $\|\widehat{g_j}\|_1 \to 0$, as $j \to \infty$. (We refer to [3], pages 31-32.)
- $|\text{supp}(g_j)| \le |\text{supp}(g)| \cdot 2^j$. This is easily proved inductively, using that

$$|\text{supp}(\uparrow_{2^j} g)| = 2^j |\text{supp}(g)| - 2^j \quad .$$

As a result,

$$\begin{aligned}
\|P_j(\delta_k)\|_2^2 &= \sum_{m\in\mathbb{Z}} |\langle \delta_k, g_j(\cdot - 2^j m)\rangle|^2 \\
&= \sum_{m\in\mathbb{Z}} |g_j(k - 2^j m)|^2 \\
&\le (|\text{supp}(g)| + 1)\, \|g_j\|_\infty \\
&\le (|\text{supp}(g)| + 1)\, \|\widehat{g_j}\|_1 \to 0 \quad ,
\end{aligned}$$

as $j \to \infty$. This concludes the proof of totality, hence (a) is shown. ■

For simplicity, we will sometimes omit the translation parameters, and call the sequence $(h_j)_{j\ge 1} \subset \ell^2(\mathbb{Z})$ a wavelet system. The associated basis is then obtained by shifting h_j by integer multiples of 2^j, just as in the theorem.

Remark 2.2.3.

1. The finiteness of the filters are only used for totality of the system. Hence any pair g, h with properties (PR) yields an ONS in $\ell^2(\mathbb{Z})$.
   Anyway, we will in the following always assume the filters to be finite.

2. Note that the filter bank properties of the DWT, i.e., g as low-pass and h as high-pass filter, enter nowhere in the proof (in fact, we could as well exchange the two). These are additional properties which we have to build into the filters.

3. The family $(P_j)_{j\ge 1}$ of projections defines a decreasing sequence $V_j = P_j(\ell^2(\mathbb{Z}))$ of closed subspaces which share many properties of an MRA in $L^2(\mathbb{R})$. Indeed, $V_j \supset V_{j+1}$ is clear by construction. $\lim_{j\to\infty} V_j = \{0\}$ has been observed in the proof of the previous theorem. We cannot expect an analog of (2.1.10), since there is no meaningful definition of dilation on $\ell^2(\mathbb{Z})$. We do however have an

analog of a scaling function, in the form of the family $(g_j)_{j>0}$: The projection onto V_j is given by

$$P_j f = \sum_{k\in\mathbb{Z}} \langle f, g_j(\cdot - 2^j k)\rangle g_j(\cdot - 2^j k) \quad .$$

4. The theorem implies that the map $f \mapsto d_j$ factors into a convolution with h_j^*, followed by a subsampling of 2^j. In particular, we can interpret the mapping

$$f \mapsto (d_1, d_2, d_3, \ldots)$$

as a (subsampled) filter bank. By the convolution theorem, $(f * h_j^*)^\wedge = \hat{f} \cdot \overline{\hat{h}_j}$, which shows that the coefficients d_j capture the part of f supported in the frequencies where $|\hat{h}_j|$ is large.

Observe that on the Fourier transform side the recursion formulae read

$$\widehat{g_{j+1}}(\omega) = \widehat{g}(2^j\omega) \cdot \widehat{g_j}(\omega) \quad , \quad \widehat{h_{j+1}}(\omega) = \widehat{h}(2^j\omega) \cdot \widehat{g_j}(\omega) \quad . \tag{2.2.4}$$

We note the similarity to the formula (2.1.19), which further emphasizes the analogy of the roles of the scaling function ϕ on the one hand, and of the g_j ($j \in \mathbb{N}$) on the other. In (2.2.3) we will see that, under suitable regularity conditions, this analogy in fact takes the form of a convergence statement:

$$2^{j/2} g_j(n) - \phi(2^{-j}n) \to 0 \quad .$$

This property (in somewhat sharper formulation) will be of crucial importance for the study of decay of discrete-time wavelet coefficients.

5. Similar observations apply to the discrete time wavelets. Observe that the elements of the wavelet basis are again indexed by a scale and a translation parameter. Instead of the dilation operator, which does not work properly on $\ell^2(\mathbb{Z})$, we now have recursively defined wavelets $(h_j)_j$ of different scales. The wavelet basis inherits the asymptotic behavior described in 4., i.e., under suitable conditions on the filters, we may think of the discrete time wavelets h_j as approximate samples of a continuous time wavelet ψ.

6. The theorem holds for biorthogonal wavelet bases as well: as long as the filters satisfy the corresponding perfect reconstruction conditions in the biorthogonal case, one will obtain **discrete-time biorthogonal wavelet bases** along the lines of 2.2.2. In the following chapter, our results will be formulated using this generalization.

Properties of Discrete-Time Wavelet Bases

As we did in the continuous case, we will describe useful properties of discrete-time wavelet systems such as *finite support*, *vanishing moments* and *regularity* properties.
A byproduct of the proof of 2.2.2 is that $|\mathrm{supp}(g_j)| \leq |\mathrm{supp}(g)| \cdot 2^j$, and likewise for h_j, if the filter h is chosen according to (2.1.24). So, starting with filters of finite length gives **finitely supported discrete-time wavelets**.
In continuous time theory, vanishing moments are necessary requirements to ensure decay of wavelet coefficients of regular signals, essentially by killing Taylor polynomials, see (2.1.2). In the discrete setting, we will encounter a similar effect. The definition carries over in a rather straightforward way:

Definition 2.2.4. *A wavelet system $(h_j)_{j\in\mathbb{N}} \subset \ell^2(\mathbb{Z})$ has N* **vanishing moments** *if*

$$\forall j \geq 1, i = 0, \ldots, N-1 \ : \ \sum_{n\in\mathbb{N}} h_j(n) n^i = 0 \quad ,$$

where the sum converges absolutely.

Again, a wavelet system having N vanishing moments kills polynomials of order $< N$: If P is any such polynomial, $j \in \mathbb{N}$ and $k \in \mathbb{Z}$, then

$$\sum_{m\in\mathbb{Z}} (h_{j,l})(m) P(m) = 0 \quad , \quad \forall (j,l) \in \mathbb{N} \times \mathbb{Z}.$$

We observe that if the h_j are finite sequences (which is the standard assumption), their Fourier transforms are trigonometric polynomials, and having N vanishing moments is equivalent to the property that the origin is a zero of order N of $\widehat{h_j}$. The following proposition shows that this property can be easily controlled by choosing the right g, via the factorization (2.2.4):

Proposition 2.2.5. *Let g, h be a perfect reconstruction pair of finite sequences, with $\widehat{g}(0) = \sqrt{2}$, and h chosen according to (2.1.24). Then $\widehat{g}(\pi) = 0$, therefore*

$$\widehat{g}(\omega) = (e^{i\omega} + 1)^N \widetilde{m}(\omega) \tag{2.2.5}$$

with $1 \leq N \leq |\mathrm{supp}(g)|$, and $\widetilde{m}$ is a trigonometric polynomial. This implies that the wavelet system $(h_j)_{j\in\mathbb{N}}$ constructed from g and h has N vanishing moments.

Proof Note that $\widehat{g}(0) = \sqrt{2}$ and (PR) imply that $\widehat{g}(\pi) = 0$. Then (2.2.5) is a standard fact about polynomials. Plugging this into (2.1.25) yields that $\widehat{h}$ has a zero of order N at 0. By (2.2.4), this zero is inherited by $\widehat{h}_{j+1}$. ∎

The notion of **regularity** is not straightforward for discrete-time sequences. Furthermore, we will see that in a sense, this property links the discrete to the continuous-time bases. We will therefore treat this property in an extra subsection.

2.2.2 Regularity of Discrete-Time Wavelets

A continuous-time function is said to be regular if it is at least continuous, or even has several continuous derivatives.

This notion of regularity does not seem to make sense in discrete time. Certainly, smoothness conditions that rely on small-scale limits cannot be adapted to the discrete time setting. Nonetheless, there is a useful (in view of later results, even crucial) notion of regularity of a wavelet system, that has to do with large scale limits.

The following notions are best understood by thinking of the sequence $(g_j)_{j\geq 1}$ as discrete approximations of a continuous-time function, with each g_j being defined on the grid $2^{-j}\mathbb{Z}$.

Regularity for discrete-time wavelets was studied by Rioul [26], mimicking Hölder-type regularity conditions for discrete-time functions.
Recall that a continuous-time function $\varphi(x)$ is *Lipschitz regular of order* α ($\varphi \in \dot{C}^\alpha$), $0 < \alpha \leq 1$, if for all $x, h \in \mathbb{R}$

$$|\varphi(x+h) - \varphi(x)| \leq C|h|^\alpha.$$

A function $\varphi(x)$ is said to be *Hölder regular of order* $r = N+\alpha$ ($\varphi \in \dot{C}^r$), $0 < \alpha \leq 1$, $N \in \mathbb{N}$, if it is N times continuously differentiable and the N-th derivative is Lipschitz of order α.

Let again $g = (g(n))_{0\leq n\leq L}$, $L \in N$, be a low pass filter of finite length and consider $(g_j)_{j\geq 1}$, obtained by the scheme in Theorem 2.2.2:

$$g_j = g^* * (\uparrow_2 g^*) * (\uparrow_4 g^*) * \cdots * (\uparrow_{2^{j-1}} g^*).$$

We will define regularity of the sequences $g_j = (g_j(n))_{n\in\mathbb{Z}}$ mimicking Lipschitz regularity. Note that the following definitions are the ones given in [26], the extra factors $2^{-j/2}$ in here arise from ℓ^2-normalization of g_j.

Definition 2.2.6. *$(g_j)_{j\geq 1}$ will be called* ***regular of order*** *α, $0 < \alpha \leq 1$, if it satisfies*

$$|g_j(n+1) - g_j(n)| \leq C 2^{-j/2} \cdot 2^{-j\alpha},$$

where C is a constant independent of j and n.

In order to extend this definition to regularity of higher orders, consider the difference operator D applied to the sequences $(g_j(n))$,

$$Dg_j(n) := (g_j(n) - g_j(n-1))/2^{-j}.$$

For $N \in \mathbb{N}$, let the sequence of N-th order differences $D^N g_j$ be the sequence obtained by applying D N times to $(g_j(n))$.

The difference operator can be seen as a discrete derivation operator, with a normalization reflecting the assumption that g_j is an approximation on the grid $2^{-j}\mathbb{Z}$.

Definition 2.2.7. *$(g_j)_{j\geq 1}$ will be called **regular of order** $r = N + \alpha$, $0 < \alpha \leq 1$, $N \in \mathbb{N}$, if it satisfies*

$$|D^N g_j(n+1) - D^N g_j(n)| \leq C 2^{-j/2} \cdot 2^{-j\alpha},$$

where C is a constant independent of j and n.

Note that regularity of $(g_j)_{j\geq 1}$ implies regularity of the family of wavelet sequences $(h_j)_{j\geq 1}$ if they are associated to (g_j) by (2.1.24).

2.2.3 Connection to Wavelet Bases for $L^2(\mathbb{R})$

In [26], definitions 2.2.6 and 2.2.7 are conceived to relate discrete-time to continuous-time wavelet transforms and their properties.

Let in the following g, h be a pair of PR filters, satisfying the requirements for Theorem 2.2.2.

Rioul [26] defines convergence of (g_j) - given by Theorem 2.2.2 - for $j \to \infty$ to a continuous-time limit function $\phi(x)$ and then relates properties of ϕ to regularity properties of the discrete-time functions g_j.

Definition 2.2.1 *The sequences (g_j) converge for $j \to \infty$ **pointwise** to a limit function $\phi(x)$ if, for any sequence of integers n_j satisfying*

$$|n_j 2^{-j} - x| \leq C 2^{-j}, \tag{2.2.6}$$

for C a constant not depending on j, we have

$$\phi(x) = \lim_{j\to\infty} 2^{j/2} g_j(n_j).$$

*Moreover, the convergence is **uniform**, if*

$$\sup_x |\phi(x) - 2^{j/2} g_j(n_j)| \to 0, \quad \text{as } j \to \infty.$$

Here we require (2.2.6) with a constant independent of x.

The above definition gives flexibility in the way interpolation of the sequences $(g_j(n))$ can be done. In particular, convergence using stepwise interpolation by $n_j = \lfloor 2^j x \rfloor$, linear interpolation or even interpolation by smoother functions such as splines are all implied by this definition.

Filters with (PR) do not necessarily give rise to a scaling function and thereby a wavelet basis in $L^2(\mathbb{R})$, but additional requirements need to be met to ensure convergence in the construction of a continuous-time scaling function ϕ. Moreover, in the study of smoothness properties of this scaling functions, conditions were needed that allowed to predict the smoothness of ϕ just from the initial discrete-time filter g. These techniques rely on the speed of convergence of the scheme (g_j). Rioul derives necessary and sufficient conditions for uniform convergence (g_j) to a limit function, which then is continuous:

Theorem 2.2.8. [26] *The collection of sequences (g_j) converges uniformly (in the sense of 2.2.1) to a limit function $\phi(x)$ if and only if*

$$\sum_n g(n) = \sqrt{2}, \tag{2.2.7}$$

$$\sum_n (-1)^n g(n) = 0 \quad \textit{and} \tag{2.2.8}$$

$$\max_n 2^{j/2} |g_j(n+1) - g_j(n)| \to 0 \ \ \textit{as } j \to \infty. \tag{2.2.9}$$

In addition, one can characterize limit functions possessing stronger regularity properties by the behavior of the (g_j).

Theorem 2.2.9. [26]

- *If g satisfies (2.2.7),(2.2.8) and, for $j \geq 1$, g_j is regular of order α for some $0 < \alpha \leq 1$, then $(g_j)_j$ will converge uniformly to a α-regular limit function φ.*

- *If the sequence of the N-th order differences $D^N g_j$ converges uniformly in the sense of (2.2.1), then φ is N times continuously differentiable. Furthermore, for $k = 0, \ldots, N$, $D^k g_j$ converges uniformly to the k-th order derivative of φ.*

- *If (2.2.7) is valid and $\sum_n (-1)^n n^i g(n) = 0$ for $i = 0, \ldots, N$ and, for $j \geq 1$, (g_j) is regular of order $r = N + \alpha$ for some $0 < \alpha \leq 1$, then the limit function $\varphi \in \dot{C}^r$, $r = N + \alpha$ and moreover, the continuous-time wavelet function ψ associated to φ by (2.1.20) possesses the same regularity.*

There is also a converse result: discrete-time finitely supported wavelet families arising from filters associated to analog scaling functions and wavelets by (2.1.14) and (2.1.20), which are Hölder regular of a certain order $N + \alpha$, $N \in \mathbb{N}_0, 0 < \alpha \leq 1$, possess the same order of discrete-time regularity, as long as the scaling function ϕ meets another condition which Rioul calls *stability* [26]: ϕ is said to be *stable* if

$$\sum_{n\in\mathbb{N}} \phi(n)e^{in\omega} \neq 0 \quad \text{for all } \omega \in \mathbb{R}.$$

All of these observations cover in particular the Daubechies families.

The Daubechies orthonormal continuous-time scaling functions and wavelets of length $L = 4$ are regular of order $\alpha \approx 0.55$, and, as the stability condition is easily checked, so are the discrete-time wavelets arising from the associated filters. Daubechies filters of length $L = 6$ give discrete-time regularity of order $r \approx 1.08$, and with further increasing filter length, regularity increases as well, for large L, the regularity is about $0.1L$ (see [26]).

To sum it up, the discrete-time notion of regularity is consistent with regularity for continuous-time functions. The regularity property will also be crucial in the next chapter.

Chapter 3

Discrete-Time Besov Spaces and their Characterizations

In Chapter 1, we introduced Besov spaces of continuous-time functions and discussed different characterizations of these spaces, as via iterated differences, Littlewood-Paley theory and in terms of the decay of wavelet coefficients.

In this chapter, we deal with analogous function spaces on the integers. In [31], R.H. Torres introduces discrete-time Besov spaces $B^{\alpha}_{p,q}(\mathbb{Z})$, $\alpha \in \mathbb{R}$, $0 < p, q < \infty$, by adapting a Littlewood-Paley type characterization for the continuous-time (homogeneous) Besov spaces $\dot{B}^{\alpha}_{p,q}(\mathbb{R})$. Roughly said, $B^{\alpha}_{p,q}(\mathbb{Z})$ are spaces of sequences, obtained by integer sampling of band-limited distributions in corresponding continuous-time Besov spaces. This will be the topic of Section 3.1.

Our main result will be that these spaces admit a characterization in terms of coefficients from discrete-time wavelet bases as described in the previous chapter:

A sequence will be in a space $B^{\alpha}_{p,q}(\mathbb{Z})$ if and only if the corresponding coefficients are in a normed space $b^{\alpha}_{p,q}(\mathbb{Z})$, describing the decay of coefficients. Note that these $b^{\alpha}_{p,q}(\mathbb{Z})$-norms are just the truncated norms defined by (1.2.1) which arose in the context of nonlinear approximation in Section 1.2.1.
Therefore, our result gives the answer to the question we posed in this section: a sequence can be approximated with a certain order if and only if it is a member in a discrete-time Besov space.

This result has been previously published in [11].

In [31], Torres also develops a discrete-time φ-transform decomposition (see Section 3.2) for $B^{\alpha}_{p,q}(\mathbb{Z})$, which will be the starting point for our considerations.

We derive our characterization of Besov spaces on $\mathbb{Z}$ in terms of discrete-time wavelet systems in Section 3.3: Just as in the continuous case, the φ-transform is in many ways quite similar to a wavelet transform, and this similarity allows a proof of the main result by studying off-diagonal decay of certain infinite matrices. This decay behavior is derived under suitable conditions regarding support, vanishing moments and smoothness of the discrete-time wavelets.

This wavelet characterization allows to obtain further descriptions of the discrete-time spaces as in Section 3.4, which can be viewed as more 'intrinsic' in contrast to the Littlewood-Paley type definition. In particular, they do not rely on the choice of auxiliary functions.

Subsection 3.4.1 contains the results on discrete-time Besov spaces in terms of discrete-time moduli of smoothness (compare to the analog continuous-time spaces in 1.1.1); another equivalent characterization in terms of oscillation over intervals is given in 3.4.2.

3.1 Littlewood-Paley Type Definition of $B^{\alpha}_{p,q}(\mathbb{Z})$

The following definition of Besov spaces on $\mathbb{Z}$ is the one employed in [31], except for a change of notation that is convenient for our purposes.
For additional background information on the corresponding spaces on $\mathbb{R}$, see Chapter 1 or e.g. [16, 17, 14, 23].

First, we need analoga of Schwartz functions and tempered distributions in terms of sequences.

We refer to [31] for basic facts concerning distributions on $\mathbb{Z}$.

Definition 3.1.1. *A complex-valued sequence $\eta = (\eta(n))_{n\in\mathbb{Z}}$ satisfying*

$$\sup_{n\in\mathbb{Z}} |\eta(n)|(1+|n|)^m < \infty \tag{3.1.1}$$

for every $m > 0$, will be called rapidly decreasing.
A sequence $f = (f(n))_{n\in\mathbb{Z}}$ will be called (tempered) distribution on $\mathbb{Z}$ *if*

$$\inf\{m \in \mathbb{N} : \sup_{n\in\mathbb{Z}} |f(n)|(1+|n|)^{-m} < \infty\} < \infty.$$

We will denote the spaces of rapidly decreasing sequences by $\mathcal{S}(\mathbb{Z})$: it is topologized by the sup-norms in (3.1.1). $\mathcal{S}'(\mathbb{Z})$ denotes the space of distributions on $\mathbb{Z}$, which is indeed the dual space of $\mathcal{S}(\mathbb{Z})$. In order to be consistent with the usual inner product notation, $\langle f, \eta\rangle = \sum_n f(n)\overline{\eta(n)}$ will stand for the pairing of a distribution f and the rapidly decreasing sequence $\overline{\eta}$. The Fourier transform of $f \in \mathcal{S}(\mathbb{Z})$ is given by $\widehat{f}(\omega) = \sum_{k\in\mathbb{Z}} f(k)e^{-ik\omega}$, extended to $\mathcal{S}'(\mathbb{Z})$ in the usual way.
A justification for this can be found again in [31].

We next define the notion of a *phi-function*, which is the basis for the Littlewood-Paley definition of discrete-time Besov spaces.

Definition 3.1.2. *A* **phi-function** *is a function* $\varphi^c \in \mathcal{S}(\mathbb{R})$ *satisfying*

$$\text{supp } \hat{\varphi}^c \subseteq \{\omega : \pi/4 < |\omega| < \pi\}, \tag{3.1.2}$$

for some $C,\ \varepsilon > 0$,

$$|\hat{\varphi}^c(\omega)| > C \text{ on } \{\omega : \pi/4 + \varepsilon < |\omega| < \pi - \varepsilon\}, \tag{3.1.3}$$

$$\hat{\varphi}^c \equiv 1 \text{ in a small neighborhood of } \{-\pi/2, \pi/2\}, \tag{3.1.4}$$

$$\sum_{\nu\in\mathbb{Z}} |\hat{\varphi}^c_\nu(\omega)|^2 = 1 \text{ for } \omega \in \mathbb{R}\setminus\{0\}. \tag{3.1.5}$$

For $\nu \in \mathbb{Z}$, set $\varphi^c_\nu(x) = 2^{-\nu+2}\varphi^c(2^{-\nu+2}x)$. The superscript c serves as a reminder that φ^c is a continuous-time function. Note that our notation differs from [31]: here, small scales correspond to small ν. We use dilation by the factor $2^{-\nu+2}$ instead of the more intuitive $2^{-\nu}$ in order to obtain a unified notation later on; this has the slightly awkward consequence that φ^c equals φ^c_2.
We will now obtain a family of rapidly decreasing sequences by sampling the functions $(\varphi_\nu)_{\nu\geq 1}$. Set $\varphi_\nu := \varphi^c_\nu|_\mathbb{Z}$ for $\nu > 1$ and $\varphi_1 := ((\chi_{[-\pi,\pi]}\hat{\varphi}^c_1)^\vee)|_\mathbb{Z}$, where the different definition for φ_1 is due to technical considerations, see [31].
Let $\mathcal{P}(\mathbb{Z})$ denote the set of polynomials on $\mathbb{R}$, sampled at the integers.

Definition 3.1.3. *Let a phi-function* $\varphi^c \in \mathcal{S}(\mathbb{R})$ *be given. For* $\alpha \in \mathbb{R}$, $0 < p, q < \infty$, *the* discrete-time Besov space $B^\alpha_{p,q}(\mathbb{Z})$ *is the collection of all* $f \in \mathcal{S}'/\mathcal{P}(\mathbb{Z})$ *(distributions on* $\mathbb{Z}$ *modulo polynomials* $\mathcal{P}(\mathbb{Z})$*), such that*

$$\|f\|_{B^\alpha_{p,q}(\mathbb{Z})} := \Big(\sum_{\nu\geq 1}(2^{-\nu\alpha}\|f * \varphi_\nu\|_{l^p(\mathbb{Z})})^q\Big)^{1/q} < \infty \quad .$$

This definition is independent of the choice of φ^c. For a distribution $f \in \mathcal{S}'(\mathbb{Z})$, we have $\|f\|_{B^\alpha_{p,q}(\mathbb{Z})} = 0$ if and only if $f * \varphi_\nu = 0$ for all $\nu \geq 1$. By the conditions on φ, this is equivalent to supp $\hat{f} = \{0\}$, or equivalently, to $f \in \mathcal{P}(\mathbb{Z})$. This is why the Besov spaces are defined as spaces of equivalence classes modulo polynomials: $\|\cdot\|_{B^\alpha_{p,q}(\mathbb{Z})}$ becomes a norm for $1 \leq p, q < \infty$ and a quasi-norm in general.
In analogy to the continuous-time case, we have a Calderón type formula for $f \in \mathcal{S}'(\mathbb{Z})$:

$$f = \sum_{\nu\geq 1} f * \varphi_\nu * \varphi^*_\nu, \tag{3.1.6}$$

with unconditional convergence in $\mathcal{S}'/\mathcal{P}(\mathbb{Z})$ [31].

3.2 φ-transform Decomposition of $B^\alpha_{p,q}(\mathbb{Z})$

A φ-transform theorem for $B^\alpha_{p,q}(\mathbb{Z})$ was derived in [31]. It can be understood as a critically sampled version of (3.1.6). We thus have at hand a decomposition for sequences in the discrete-time Besov space similar to an expansion into an orthonormal basis, though non-orthogonal and with 'basis elements' that are compactly supported in Fourier domain. Theorem 3.2.1 provides a norm equivalence that can be read as a critically sampled version of Definition 3.1.3, and which yields the characterization of $B^{\alpha}_{p,q}(\mathbb{Z})$ in terms of the membership of φ-transform coefficients in the space $b^{\alpha}_{p,q}(\mathbb{Z})$ we defined in (1.2.1).

Consider again a phi-function $\varphi^c \in \mathcal{S}(\mathbb{R})$, and for $\nu, k \in \mathbb{Z}$ let

$$\varphi^c_{\nu,k}(x) = 2^{(-\nu+2)/2}\varphi^c(2^{-\nu+2}x - k).$$

For $k \in \mathbb{Z}$, define $\varphi_{\nu,k} = \varphi^c_{\nu,k}|_{\mathbb{Z}}$ for $\nu > 1$ and $\varphi_{1,k} = \tau_k \varphi_1$.
In [31], starting from the formula (3.1.6) and following the lines of Frazier and Jawerth [15], it is derived that for any $f \in \mathcal{S}'(\mathbb{Z})$

$$f = \sum_{\nu \geq 1} \sum_{k \in \mathbb{Z}} \langle f, \varphi_{\nu,k} \rangle \varphi_{\nu,k}, \tag{3.2.1}$$

with convergence in $\mathcal{S}'/\mathcal{P}(\mathbb{Z})$. It is also well-known that condition (iv) of Definition 3.1.2 alone guarantees that for $f \in \ell^2(\mathbb{Z})$ the decomposition (3.2.1) converges in the norm.
Let the φ-transform S_φ for $f \in \mathcal{S}'(\mathbb{Z})$ be defined by $S_\varphi f = s = (s_{\nu,k})_{\nu \geq 1, k \in \mathbb{Z}}$, where $s_{\nu,k} = \langle f, \varphi_{\nu,k} \rangle$, and for a complex-valued sequence $t = (t_{\nu,k})_{\nu \geq 1, k \in \mathbb{Z}}$ define the inverse φ-transform by T_φ by $T_\varphi t = \sum_{\nu,k} t_{\nu,k} \varphi_{\nu,k}$. The convergence of the sum is guaranteed by the following result.

Theorem 3.2.1. *[[31]] Let $\alpha \in \mathbb{R}$, $0 < p, q < \infty$.*
Both of the operators $S_\varphi : B^{\alpha}_{p,q}(\mathbb{Z}) \to b^{\alpha}_{p,q}(\mathbb{Z})$ and $T_\varphi : b^{\alpha}_{p,q}(\mathbb{Z}) \to B^{\alpha}_{p,q}(\mathbb{Z})$ are bounded with $\|f\|_{B^{\alpha}_{p,q}(\mathbb{Z})} \asymp \|S_\varphi f\|_{b^{\alpha}_{p,q}(\mathbb{Z})}$ and $T_\varphi \circ S_\varphi = id_{B^{\alpha}_{p,q}(\mathbb{Z})}$.

In other words, under these maps, $B^{\alpha}_{p,q}(\mathbb{Z})$ is a retract of $b^{\alpha}_{p,q}(\mathbb{Z})$, and $B^{\alpha}_{p,q}(\mathbb{Z})$ can be identified with the closed subspace $S_\varphi(B^{\alpha}_{p,q}(\mathbb{Z}))$ of $b^{\alpha}_{p,q}(\mathbb{Z})$.
We next give a more precise statement concerning the convergence of the φ-transform decomposition, if we know a sequence to belong to a Besov space. Let for $K \in \mathbb{N}$

$$\mathcal{S}_K(\mathbb{Z}) := \{\eta \in \mathcal{S}(\mathbb{Z}) : \sum \eta(n) n^m = 0,\ \mathbb{N} \ni m \leq K\}$$

and $\mathcal{S}_\infty(\mathbb{Z}) := \{\eta \in \mathcal{S}(\mathbb{Z}) : \sum \eta(n) n^m = 0 \text{ for all } m \in \mathbb{N}\}$, $\mathcal{S}_{-1}(\mathbb{Z}) := \mathcal{S}(\mathbb{Z})$.
Note that the dual space of $\mathcal{S}_K(\mathbb{Z})$ can be identified with $\mathcal{S}'/\mathcal{P}_K(\mathbb{Z})$, the space of equivalence classes of distributions modulo polynomials of degree $\leq K$, where $(\mathcal{S}_{-1}(\mathbb{Z}))' \sim \mathcal{S}'(\mathbb{Z})$ and also $(\mathcal{S}_\infty(\mathbb{Z}))' \sim \mathcal{S}'/\mathcal{P}(\mathbb{Z})$ ([17, 14]).

Lemma 3.2.2. *Let $(\varphi_{\nu,k})_{\nu\geq 1,k\in\mathbb{Z}}$ be the family of rapidly decreasing sequences defined in Section 3.2.*
Any $f \in B^{\alpha}_{p,q}(\mathbb{Z})$ can be written as

$$f = \sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}} \langle f, \varphi_{\nu,k}\rangle \varphi_{\nu,k}$$

in the sense of $(\mathcal{S}_K(\mathbb{Z}))'$, $K = \max\{[\alpha - 1/p], -1\}$, and for any $\overline{\eta} \in \mathcal{S}_K(\mathbb{Z})$, we set

$$\langle f, \eta\rangle := \sum_{\nu,k} \langle f, \varphi_{\nu,k}\rangle \langle \varphi_{\nu,k}, \eta\rangle.$$

Proof First, notice that for $\overline{\eta} \in \mathcal{S}(\mathbb{Z})$, absolute convergence of the φ-transform decomposition holds in a pointwise sense, compare to [14]. Now, let $\overline{\eta} \in \mathcal{S}_K$. Lemma 3.3.2 below (see Remark 3.3.3) allows the estimate $|\langle \varphi_{\nu,k}, \eta\rangle| \leq C2^{-\nu(K+1+1/2)}(1 + |k|)^{-M}$ for $M > 0$. For $f \in B^{\alpha}_{p,q}(\mathbb{Z})$, $(\langle f, \varphi_{\nu,k}\rangle)_{\nu\geq 1,k\in\mathbb{Z}} \in b^{\alpha}_{p,q}(\mathbb{Z})$ by Theorem 3.2.1 and particularly for any coefficient $|\langle f, \varphi_{\nu,k}\rangle| \leq C2^{\nu(\alpha-1/p+1/2)}$.
Hence,

$$\begin{aligned}\sum_{\nu=N+1}^{\infty}\sum_{k} |\langle f, \varphi_{\nu,k}\rangle||\langle \varphi_{\nu,k}, \eta\rangle| \leq & \quad C \textstyle\sum_{\nu=N+1}^{\infty}\sum_k 2^{\nu(\alpha-1/p+1/2)}2^{-\nu(K+1+1/2)}(1 + |k|)^{-M} \\ \leq & \quad C \textstyle\sum_{\nu=N+1}^{\infty}\sum_k 2^{-\nu(K-\alpha+1/p+1)},\end{aligned}$$

which ensures convergence for $K = \max\{[\alpha - 1/p], -1\}$ as

$$\lim_{N\to\infty} \sum_{n\in\mathbb{Z}} f(n)\overline{(\eta(n) - \sum_{\nu=1}^{N} \langle \eta, \varphi_{\nu,k}\rangle \varphi_{\nu,k}(n))} = 0.$$

■

Remark 3.2.3.

1. For continuous-time functions, we decided between homogeneous and inhomogeneous spaces. In the discrete-time case, we defined a single scale of the spaces, starting from homogeneous Besov spaces. As Torres [31] notes, in discrete time, the notion of inhomogeneous spaces makes no sense: we cut off high frequencies and the discrete-time Besov norm controls the large scale behavior of the sequence. Controlling this large scale behavior as in 1.1.6 in fact results in a norm equivalent to $\|\cdot\|_{\ell^p(\mathbb{Z})}$.

2. There is a sampling theorem in [31]: it is shown that for $1 < p < \infty$, the spaces $B^{\alpha}_{p,q}(\mathbb{Z})$ correspond exactly to the spaces of samples of functions in $\dot{B}^{\alpha}_{p,q}(\mathbb{R}) \cap E_{\pi}$, where E_{π} is the set of tempered distributions whose Fourier transforms are supported on $[-\pi, \pi]$.

3.3 Wavelet Characterization of $B^\alpha_{p,q}(\mathbb{Z})$

In this section, we present the central result of this thesis. The results of this section have been published in [11].

We consider biorthogonal families $(h_j)_{j\geq 1}$ and $(\tilde{h}_j)_{j\geq 1}$ in $\ell^2(\mathbb{Z})$, and associated systems $(h_{j,l}) = (h_j(\cdot - 2^j l))$, $\tilde{h}_{j,l} = (\tilde{h}_j(\cdot - 2^j l))$. We assume that all involved $h_j, \tilde{h}_j$ have finite supports. Thus it is possible to define for $f \in S'(\mathbb{Z})$ the operator $S_h f = (\langle f, h_{j,l}\rangle)_{j\geq 1, l\in\mathbb{Z}}$, and $S_{\tilde{h}}$ likewise. Moreover, for all *finitely supported* coefficient sequences $d = (d_{j,l})_{j\geq 1, l\in\mathbb{Z}}$, define

$$T_h d = \sum_{j,l} d_{j,l} h_{j,l} \quad . \tag{3.3.1}$$

Again, let $T_{\tilde{h}}$ be defined in an analogous way. Our aim is the characterization of $f \in B^\alpha_{p,q}(\mathbb{Z})$ by use of the operators T_h and $S_{\tilde{h}}$.

In the following we want to provide criteria on $(h_j)_{j\geq 1}, (\tilde{h}_j)_{j\geq 1}$ to ensure analogs of 3.2.1, with T_h, $S_{\tilde{h}}$ replacing T_φ, S_φ. Our main result will be Theorem 3.3.8 below, which may be viewed as analogy to 3.2.1, but also to the wavelet characterization of $\dot{B}^\alpha_{p,q}(\mathbb{R})$.

In the proof we exploit the strong similarities of the biorthogonal wavelet and φ-transforms: both are based on building blocks indexed by dyadic scales 2^j, which are shifted along the grid $2^j\mathbb{Z}$.

But our result is not included in the φ-transform result: Recall that the φ-transform sequences $\varphi_{\nu,k}$ arise by sampling band-limited Schwartz functions with infinitely many vanishing moments. By contrast, the wavelet systems need only have finitely many vanishing moments, and only a finite degree of smoothness. In addition, we assume a control over the supports of the initial sequences $(h_j)_{j\geq 1}$ and $(\tilde{h}_j)_{j\geq 1}$ that can not be obtained by sampling bandlimited functions.

This setup covers the discrete-time biorthogonal wavelet bases as described in Section 2.2, but it could also be applied to even more general systems, e.g. arising from cascade algorithms where the analysis filter changes at each scale in a controlled way.

In any case, the φ-transform will be the starting point for our considerations. The blueprint for the general proof strategy is provided by the continuous time theory, as contained e.g. in [17] or [14]. However the arguments need to be adapted to properties of discrete-time wavelet families. The key to the proof is the study of the off-diagonal behavior of the *transition matrices* $A = (\langle h_{j,l}, \varphi_{\nu,k}\rangle)_{j,l,\nu,k}$ and $\tilde{A} = (\langle \varphi_{\nu,k}, \tilde{h}_{j,l}\rangle)_{\nu,k,j,l}$, and to conclude boundedness of certain associated operators acting on $b^\alpha_{p,q}(\mathbb{Z})$. For this purpose, we study how properties of $h_{j,l}$ influence the size of $|\langle h_{j,l}, \varphi_{\nu,k}\rangle|$.

First, we employ an inequality that will be used repeatedly:

Lemma 3.3.1. *Let $\alpha \leq 1$ and $M \geq N+1$. Then,*

$$\sum_{k\in\mathbb{Z}}(1+\alpha|k|)^{-M}(1+|k|)^{N} \leq C\alpha^{-(N+1)}.$$

Proof

$$\begin{aligned}
\sum_{k\in\mathbb{Z}}(1+\alpha|k|)^{-M}(1+|k|)^{N} &= \alpha^{-M}\sum_{k\in\mathbb{Z}}(\alpha^{-1}+|k|)^{-M}(1+|k|)^{N} \\
&\leq \alpha^{-M}\sum_{k\in\mathbb{Z}}(\alpha^{-1}+|k|)^{N-M} \\
&= \alpha^{-M}(\alpha^{M-N}+2\sum_{k\geq 1}(\alpha^{-1}+k)^{N-M}) \\
&\leq \alpha^{-M}(\alpha^{M-N}+2\int_{0}^{\infty}(\alpha^{-1}+x)^{N-M}dx) \\
&= \alpha^{-N}+2\alpha^{-M}\int_{\alpha^{-1}}^{\infty}x^{N-M}dx \\
&= \alpha^{-N}+2\alpha^{-M}\alpha^{M-N-1} \\
&\leq C\alpha^{-(N+1)}
\end{aligned}$$

■

Lemma 3.3.2. *Let $(h_j)_{j\geq 1}$, $(\varphi_\nu)_{\nu\geq 1}$ be discrete-time families satisfying the following conditions:*
There are $N \geq 1$ and $M > 0$ such that

$$\sum_{n\in Z} n^{i}h_j(n) = 0 \ \text{for}\ i = 0,\ldots,N-1, \tag{3.3.2}$$

$$|h_j(n)| \leq C2^{-j/2}(1+2^{-j}|n|)^{-(M+N+2)} \ \text{for}\ n \in \mathbb{Z}. \tag{3.3.3}$$

Further, assume that for each $\nu \geq 1$, $n \in \mathbb{Z}$ there is a polynomial $p_{\nu,n}$ of degree $\leq N$, and a function $\Phi_\nu : \mathbb{Z}\times\mathbb{Z} \to \mathbb{R}^{+}$ such that

$$\begin{array}{rl}
i) & |\varphi_\nu(n-k)-p_{\nu,n}(k)| \leq C2^{-\nu/2}2^{-\nu N}(1+|k|)^{N}\Phi_\nu(n,k), \\
ii) & \Phi_\nu(n,k) \leq C, \\
iii) & \Phi_\nu(n,k) \leq C(1+2^{-\nu}|n|)^{-M} \ \text{for}\ |k| < \dfrac{|n|}{2},
\end{array} \tag{3.3.4}$$

with C independent of n, ν, k.
Then, for $j \leq \nu$

$$|(h_j * \varphi_\nu)(n)| \leq C'2^{(j-\nu)(N+1/2)}(1+2^{-\nu}|n|)^{-M}, \tag{3.3.5}$$

where C' is a constant independent of ν, n, k.

Proof This can be verified analogously to the second part of the proof of Lemma 3.3 in [15]. We give this proof for the sake of completeness:
Let $j \leq \nu$.

$$\begin{aligned}
|(h_j * \varphi_\nu)(n)| &= |\sum_{k\in\mathbb{Z}} h_j(k)\varphi_\nu(n-k)| \\
&\overset{(3.3.2)}{=} |\sum_{k\in\mathbb{Z}} h_j(k)(\varphi_\nu(n-k) - p_{\nu,n}(k))| \\
&\overset{(3.3.4\ i)}{\leq} C\sum_{k\in\mathbb{Z}} |h_j(k)|2^{-\nu/2}2^{-\nu N}(1+|k|)^N\Phi_\nu(n,k) \\
&= C\Big(\sum_{|k|<\frac{|n|}{2}} |h_j(k)|2^{-\nu/2}2^{-\nu N}(1+|k|)^N\Phi_\nu(n,k) \\
&\quad + \sum_{|k|\geq\frac{|n|}{2}} |h_j(k)|2^{-\nu/2}2^{-\nu N}(1+|k|)^N\Phi_\nu(n,k)\Big) \\
&=: C(I+II)
\end{aligned}$$

Using (3.3.4 iii), the first sum can be estimated by

$$\begin{aligned}
I &\leq C2^{-(j+\nu)/2}2^{-\nu N}(1+2^{-\nu}|n|)^{-M}\sum_{|k|<|n|/2}(1+2^{-j}|n|)^{-(M+N+2)}(1+|k|)^N \\
&\leq C2^{-(j+\nu)/2}2^{-\nu N}2^{j(N+1)}(1+2^{-\nu}|n|)^{-M} \\
&= C2^{(j-\nu)(N+1/2)}(1+2^{-\nu}|n|)^{-M},
\end{aligned}$$

where in the second inequality we used Lemma 3.3.1 with $\alpha = 2^{-j}$.
From (3.3.4 ii), we get for the second sum

$$\begin{aligned}
II &\leq C2^{-(j+\nu)/2}2^{-\nu N}\sum_{|k|\geq|n|/2}(1+2^{-j}|n|)^{-(M+N+2)}(1+|k|)^N \\
&\leq C2^{-(j+\nu)/2}2^{-\nu N}(1+2^{-j-1}|n|)^{-M}\sum_{|k|>|n|/2}(1+2^{-j}|n|)^{-(N+2)}(1+|k|)^N \\
&\leq C2^{-(j+\nu)/2}2^{-\nu N}2^{j(N+1)}(1+2^{-\nu}|n|)^{-M} \\
&= C2^{(j-\nu)(N+1/2)}(1+2^{-\nu}|n|)^{-M},
\end{aligned}$$

again, we made use of Lemma 3.3.1.

■

Remark 3.3.3. As a consequence of Lemma 3.3.2, we obtain for $\eta \in \mathcal{S}_K(\mathbb{Z})$ and $(\varphi_\nu)_{\nu\geq 1}$ satisfying the requirements of the Lemma that there is $M > 0$ such that

$$|(\eta * \varphi_\nu)(n)| \leq C 2^{-\nu(K+1+1/2)}(1 + 2^{-\nu}|n|)^{-M}. \tag{3.3.6}$$

We will need a discrete-time analogon to Taylor's formula. For $f = (f(n))_{n\in\mathbb{Z}}$, consider the *forward* and *backward difference operators* $\Delta f(n) = f(n+1) - f(n)$ and $\nabla f(n) = f(n-1) - f(n)$.

Lemma 3.3.4. *Let $f = (f(n))_{n\in Z}$, $k \in \mathbb{Z}$, $N \in \mathbb{N}$.*
i) For $k \geq 0$, $k \geq N$

$$f(n-k) = \sum_{i=0}^{N-1} \binom{k}{i} \nabla^i f(n) + \sum_{m=0}^{k-N-1} \binom{k-1-m}{N-1} \nabla^N f(n-m) + \nabla^N f(n-k+N).$$

ii) For $k < 0$, $k \leq -N$

$$f(n-k) = \sum_{i=0}^{N-1} \binom{-k}{i} \Delta^i f(n) + \sum_{m=0}^{-k-N-1} \binom{-k-1-m}{N-1} \Delta^N f(n+m) + \Delta^N f(n-k-N).$$

Proof First, by direct calculation, if $k \geq 0$

$$f(n-k) = \sum_{i=0}^{k} \binom{k}{i} \nabla^i f(n)$$

and for $k < 0$

$$f(n-k) = \sum_{i=0}^{-k} \binom{-k}{i} \Delta^i f(n).$$

In the literature, this type of expansion is sometimes referred to as Newton-Gregory interpolation formula.
i) For $N = k$, we get

$$f(n-k) = \sum_{i=0}^{k} \binom{k}{i} \nabla^i f(n) = \sum_{i=0}^{k-1} \binom{k}{i} \nabla^i f(n) + \nabla^N f(n).$$

$N \to N-1$:

$$
\begin{aligned}
&f(n-k)\\
&= \sum_{i=0}^{N-1}\binom{k}{i}\nabla^i f(n) + \sum_{m=0}^{k-N-1}\binom{k-1-m}{N-1}\nabla^N f(n-m) + \nabla^N f(n-k+N)\\
&= \sum_{i=0}^{N-2}\binom{k}{i}\nabla^i f(n) + \binom{k}{N-1}\nabla^{N-1} f(n) +\\
&\quad \sum_{m=0}^{k-N-1}\binom{k-1-m}{N-1}\left(\nabla^{N-1} f(n-m-1) - \nabla^{N-1} f(n-m)\right) + \nabla^N f(n-k+N)\\
&= \sum_{i=0}^{N-2}\binom{k}{i}\nabla^i f(n) + \sum_{m=0}^{k-N}\left(\binom{k-m}{N-1} - \binom{k-1-m}{N-1}\right)\nabla^{N-1} f(n-m)\\
&\quad +\nabla^{N-1} f(n-k+N-1)\\
&= \sum_{i=0}^{N-2}\binom{k}{i}\nabla^i f(n) + \sum_{m=0}^{k-N}\binom{k-1-m}{N-2}\nabla^{N-1} f(n-m) + \nabla^{N-1} f(n-k+N-1),
\end{aligned}
$$

where in the last equation we used that for $l, m \in \mathbb{N}$ $\binom{m}{l+1} + \binom{m}{l} = \binom{m+1}{l+1}$.
The proof of ii) can be done in the same way. ∎

These difference operators are related to the operator D_j defined in Section 2.2.2:

Lemma 3.3.5. *The forward and backward difference operators Δ, ∇, defined above satisfy for $f = (f(n))_{n\in\mathbb{Z}}$*

$$
\begin{aligned}
i)\ \ \Delta^m f(n-m) &= 2^{-jm} D_j^m f(n)\\
ii)\ \ \nabla^m f(n) &= (-1)^m 2^{-jm} D_j^m f(n),
\end{aligned}
$$

where $m \in \mathbb{N}$ and D_j the difference operator defined in section 2.2.2.

Proof i)
$m = 1$:

$$2^{-j} D_j f(n) = f(n) - f(n-1) = \Delta f(n-1),$$

$m \to m+1$:

$$
\begin{aligned}
2^{-j(m+1)} D_j^{m+1} f(n) &= 2^{-j(m+1)} 2^j (D_j^m f(n) - D_j^m f(n-1))\\
&= \Delta^m f(n-m) - \Delta^m f(n-m-1)\\
&= \Delta^{m+1} f(n-(m+1)).
\end{aligned}
$$

ii)
$m = 1$:

$$-2^{-j} D_j f(n) = -f(n) + f(n-1) = \nabla f(n),$$

$m \to m+1$:

$$\begin{aligned}
(-1)^{m+1} 2^{-j(m+1)} D_j^{m+1} f(n) &= (-1)^{m+1} 2^{-j(m+1)} 2^j (D_j^m f(n) - D_j^m f(n-1)) \\
&= (-1)^m 2^{-jm} (D_j^m f(n-1) - D_j^m f(n)) \\
&= \nabla^m f(n-1) - \nabla^m f(n) \\
&= \nabla^{m+1} f(n).
\end{aligned}$$

■

For the following lemma, observe that assumptions concerning support size and vanishing moments of $(h_j)_{j\geq 1}$ carry over to the larger system $(h_{j,l})_{j\geq 1, l\in\mathbb{Z}}$. Moreover, for the case of discrete wavelet bases as constructed in Section 2.2, the support properties are trivially fulfilled, and the vanishing moments are ensured by having enough vanishing moments in the initial high-pass filter h, confer Proposition 2.2.5.

Lemma 3.3.6. *For the discrete-time families $(\varphi_{\nu,k})_{\nu\geq 1, k\in\mathbb{Z}}$, the φ-transform as defined in (3.2), and $(h_{j,l})_{j\geq 1, l\in\mathbb{Z}}$, satisfying*

$$|supp\ h_{j,l}| \leq C 2^j, \quad \|h_{j,l}\|_\infty \leq C 2^{-j/2} \tag{3.3.7}$$

$$\textstyle\sum_{n\in Z} n^i h_{j,l}(n) = 0 \ \ for\ i = 0, \ldots, N_1 - 1, \tag{3.3.8}$$

$$(h_{j,l}) \text{ regular of order } N_2 + \varepsilon \text{ in the sense of } 2.2.7, \quad 0 < \varepsilon \leq 1, \tag{3.3.9}$$

the following inequalities are valid:
There exist $C > 0$,$M_1, M_2 \in \mathbb{N}$, such that

$$|\langle h_{j,l}, \varphi_{\nu,k}\rangle| \leq C 2^{(j-\nu)(N_1+1/2)} \left(1 + \frac{|2^\nu k - 2^j l|}{2^\nu}\right)^{-M_1} \quad for\ j < \nu, \tag{3.3.10}$$

$$|\langle h_{j,l}, \varphi_{\nu,k}\rangle| \leq C 2^{(\nu-j)(N_2+1/2)} \left(1 + \frac{|2^\nu k - 2^j l|}{2^j}\right)^{-M_2} \quad for\ j \geq \nu. \tag{3.3.11}$$

Proof We first show that for any $j \geq 1$ and $\nu > 1$, the sequences $h_j = (h_{j,0}(n))_{n\in\mathbb{Z}}$ and $\varphi_\nu = (\varphi_{\nu,0}(n))_{n\in\mathbb{Z}}$, satisfy the conditions in Lemma 3.3.2. Setting $n = 2^\nu k - 2^j l$ in (3.3.5) will give the first of the above inequalities.
Let $j \geq 1$. By assumption (3.3.8), h_j satisfies (3.3.2) with $N = N_1$ and by assumptions (3.3.7), for $n \in$ supp h_j we have $(1 + 2^{-j}|n|) \leq C$, which gives

$$|h_j(n)| \leq C 2^{-j/2} (1 + 2^{-j}|n|)^{-M} \ \text{ for any } \ M > 0.$$

Choosing $M_1 > 0$ and setting $M = M_1 + N_1 + 2$ gives (3.3.3).
For $\nu > 1$, the sequence $(\varphi_{\nu,0}(n))_{n\in Z}$ arises by sampling from a continuous-time function $\varphi^c_{\nu,0} \in \mathcal{S}(\mathbb{R})$, where $\varphi^c_{\nu,0}(x) = 2^{(-\nu+2)/2}\varphi^c(2^{-\nu+2}x)$. By definition, φ^c corresponds to φ^c_2, so we can rewrite the above equation for convenience:
$\varphi_{\nu,0}(n-k) = 2^{-\nu/2}\varphi^c_0(2^{-\nu}(n-k))$.
Let P_x be the Taylor polynomial of φ^c_0 of degree $N_1 - 1$ in $x \in \mathbb{R}$:

$$P_x(\cdot) = \sum_{m=0}^{N_1-1} \frac{(\varphi^c_0)^{(m)}(x)}{m!}(\cdot - x)^m.$$

For $y \in \mathbb{R}$, there is ξ between x and y, such that $\varphi^c_0(y) = P_x(y) + \frac{(\varphi^c_0)^{(N_1)}(\xi)}{N_1!}$. Setting $p_{\nu,n}(k) = 2^{-\nu/2}P_{2^{-\nu}n}(2^{-\nu}(n-k))$ and $\Phi_\nu(n,k) = \sup\{\frac{(\varphi^c_0(\xi))^{(N_1)}}{N_1!},\ \xi \text{ between } x \text{ and } y\}$, we have (3.3.4 i), and (3.3.4 ii) as $\|(\varphi^c_0)^{(N_1)}\|_\infty < \infty$.
Let $|k| < \frac{|n|}{2}$. In this case, for ξ between $2^{-\nu}n$ and $2^{-\nu}(n-k)$, $|\xi| \geq 2^{-\nu}\min(|n|, |n-k|) \geq 2^{-\nu}|n|/2$. Since $\varphi^c_0 \in \mathcal{S}(\mathbb{R})$, we obtain (3.3.4 iii), as $(\varphi^c_0)^{(N_1)}(\xi) \leq C(1 + 2^{-\nu-1}|n|)^{-M_1} \leq C(1 + 2^{-\nu}|n|)^{-M_1}$.
In order to prove the second inequality, we make again use of Lemma 3.3.2, exchanging the roles of φ_ν and h_j. We will therefore need to check that the system (φ_ν) fulfills the requirements imposed on (h_j) in Lemma 3.3.2, and vice versa.

For $\nu \geq 1$, $\varphi_\nu \in \mathcal{S}(\mathbb{Z})$, which gives (3.3.3) for any $N \in \mathbb{N}$. Since the moments of any order of φ_ν vanish (3.3.2) is valid. Hence it remains to check the condition (3.3.4) with (h_j) replacing (φ_ν). Once this is achieved, Lemma 3.3.2 provides the estimate

$$|(h^*_j * \varphi_\nu)(n)| \leq C2^{(\nu-j)(N_2+1/2)}(1 + 2^{-j}|n|)^{-M},$$

and setting $n = 2^j l - 2^\nu k$ will finish the proof.
As we chose the filters to be of finite length, supp h^*_j is finite. By Lemma 3.3.4, for n and $n-k$ in supp h^*_j, for $k > 0$

$$h^*_j(n-k) = \sum_{i=0}^{k} \binom{k}{i} \nabla^i h^*_j(n)$$

and for $k < 0$

$$h^*_j(n-k) = \sum_{i=0}^{-k} \binom{-k}{i} \Delta^i h^*_j(n).$$

For n outside the support of h^*_j, one has to expand the upper series in the right resp. the left endpoint of the supporting interval.
In the case $k > 0$, set

$$p^1_{j,n}(k) = \sum_{i=0}^{N_2-1} \binom{k}{i} \nabla^i h^*_j(n)$$

and for $k<0$

$$p^2_{j,n}(k)=\sum_{i=0}^{N_2-1}\binom{-k}{i}\Delta^i h^*_j(n).$$

Using Lemmata 3.3.4 and 3.3.5 gives for $k\geq N_2$

$$\begin{aligned}
&|h^*_j(n-k)-p^1_{j,n}(k)|\\
&= \ |\sum_{m=0}^{k-N_2-1}\binom{k-1-m}{N_2-1}\nabla^{N_2}h^*_j(n-m)+\nabla^{N_2}h^*_j(n-k+N_2)|\\
&\leq \left(\sum_{m=0}^{k-N_2-1}\binom{k-1-m}{N_2-1}+\binom{N_2}{N_2}\right)\cdot\sup_{m\in\{0,\dots,k-N_2\}}|\nabla^{N_2}h^*_j(n-m)|\\
&\leq \binom{k}{N_2}\cdot\sup_{m\in\{0,\dots,k\}}|(-1)^{N_2}2^{-jN_2}D^{N_2}_j h^*_j(n-m)|\\
&\leq C2^{-jN_2}(1+|k|)^{N_2}\cdot\sup_{l\in\{n-k,\dots,n\}}|D^{N_2}h^*_j(l)|.
\end{aligned}$$

The regularity condition on h^*_j gives

$$\sup_{l\in\{n-k,\dots,n\}}|D^{N_2}h^*_j(l)|\leq C2^{-j/2}$$

and due to the finite support of h^*_j, for $k<|n|/2$,

$$\begin{aligned}
\sup_{l\in\{n-k,\dots,n\}}|D^{N_2}h^*_j(l)| &\leq C2^{-j/2}(1+2^{-j}|l|)^{-M}\\
&\leq C2^{-j/2}(1+2^{-j-1}|n|)^{-M}\\
&\leq C2^{-j/2}(1+2^{-j}|n|)^{-M}
\end{aligned}$$

for any $M>0$.
Likewise, for $k<-N_2$,

$$\begin{aligned}
|h^*_j(n-k)-p^2_{j,n}(k)| &\leq C(1+|k|)^{N_2}\cdot\sup_{m\in\{0,\dots,-k-N_2\}}|\Delta^{N_2}h^*_j(n+m)|\\
&\leq C2^{-jN_2}(1+|k|)^{N_2}\sup_{l\in\{n,\dots,n-k\}}|D^{N_2}h^*_j(l)|.
\end{aligned}$$

■

Lemma 3.3.7. *Let $\alpha \in \mathbb{R}$, $0 < p,q < \infty$ and let $(\varphi_{\nu,k})_{\nu\geq 1,k\in\mathbb{Z}}$ be the family of sequences defined in (3.2).*
If $(h_{j,l})_{j\geq 1,l\in\mathbb{Z}}$ satisfies the conditions of Lemma 3.3.6 with $N_1 > 1/(\min(1,p))-1-\alpha$, $N_2 > \alpha$, then the matrix $A := (\langle h_{j,l},\varphi_{\nu,k}\rangle)_{j,l,\nu,k}$, defines a bounded operator on $b^\alpha_{p,q}(\mathbb{Z})$, where $As = (\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}\langle h_{j,l},\varphi_{\nu,k}\rangle s_{j,l})_{\nu,k}$ for $s \in b^\alpha_{p,q}(\mathbb{Z})$.
Also, for $(\tilde{h}_{j,l})_{j\geq 1,l\in\mathbb{Z}}$, satisfying the conditions of Lemma 3.3.6, where now $N_1 > \alpha$, $N_2 > 1/(\min(1,p)) - 1 - \alpha$, the matrix $\tilde{A} := (\langle\varphi_{\nu,k},\tilde{h}_{j,l}\rangle)_{\nu,k,j,l}$, where $\tilde{A}s = (\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}\langle\varphi_{\nu,k},\tilde{h}_{j,l}\rangle s_{\nu,k})_{j,l}$, is bounded on $b^\alpha_{p,q}(\mathbb{Z})$ as well.

Proof The proof is quite close to the continuous case treated in [15, 14]. We reproduce it here for the sake of convenience.
First, we show boundedness of A:

$$\begin{aligned}
\|As\|^q_{b^\alpha_{p,q}(\mathbb{Z})} &= \sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\left(2^{-\nu(\alpha-1/p+1/2)}|\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}\langle h_{j,l},\varphi_{\nu,k}\rangle s_{j,l}|\right)^p\right]^{q/p}\\
&\leq \sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\left(\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}2^{(j-\nu)(\alpha-1/p+1/2)}|\langle h_{j,l},\varphi_{\nu,k}\rangle|2^{-j(\alpha-1/p+1/2)}|s_{j,l}|\right)^p\right]^{q/p}\\
&\leq C\left\{\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\left(\sum_{j<\nu}\sum_{l\in\mathbb{Z}}2^{(j-\nu)(\alpha-1/p+1/2)}|\langle h_{j,l},\varphi_{\nu,k}\rangle|2^{-j(\alpha-1/p+1/2)}|s_{j,l}|\right)^p\right]^{q/p}\right.\\
&\quad\left.+\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\left(\sum_{j\geq\nu}\sum_{l\in\mathbb{Z}}2^{(j-\nu)(\alpha-1/p+1/2)}|\langle h_{j,l},\varphi_{\nu,k}\rangle|2^{-j(\alpha-1/p+1/2)}|s_{j,l}|\right)^p\right]^{q/p}\right\}\\
&=: C\Big\{I^q + II^q\Big\}
\end{aligned}$$

In the case $1 < p < \infty$, we have from Lemma 3.3.6 and Minkowski's inequality for the first term

$$\begin{aligned}
I^q &\leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\left(\sum_{j<\nu,l\in\mathbb{Z}}2^{(j-\nu)(\alpha-1/p+1/2+N_1+1/2)}\left(1+\frac{|2^\nu k-2^j l|}{2^\nu}\right)^{-M_1}2^{-j(\alpha-1/p+1/2)}|s_{j,l}|\right)^p\right]^{q/p}\\
&\leq C\sum_{\nu\geq 1}\left[\sum_{j<\nu}2^{(j-\nu)(\alpha-1/p+1+N_1)}\left(\sum_{k\in\mathbb{Z}}(\sum_{l\in\mathbb{Z}}\left(1+\frac{|2^\nu k-2^j l|}{2^\nu}\right)^{-M_1}2^{-j(\alpha-1/p+1/2)}|s_{j,l}|)^p\right)^{1/p}\right]^{q/p}
\end{aligned}$$

By Hölder's inequality, where $1/p + 1/p' = 1$, the inner p-sum can be estimated by

$$
\begin{aligned}
& \left[\sum_{k\in\mathbb{Z}}\left(\sum_{l\in\mathbb{Z}}(1+|k-2^{j-\nu}l|)^{-M_1}2^{-j(\alpha-1/p+1/2)}|s_{j,l}|\right)^p\right]^{1/p} \\
= & \left[\sum_{k\in\mathbb{Z}}\left(\sum_{l\in\mathbb{Z}}(1+|k-2^{j-\nu}l|)^{-M_1/p'}(1+|k-2^{j-\nu}l|)^{-M_1/p}2^{-j(\alpha-1/p+1/2)}|s_{j,l}|\right)^p\right]^{1/p} \\
\leq & \left[\sum_{k\in\mathbb{Z}}\left\{\left(\sum_{l\in\mathbb{Z}}(1+|k-2^{j-\nu}l|)^{-M_1}\right)^{1/p'}\cdot\right.\right. \\
& \left.\left.\left(\sum_{l\in\mathbb{Z}}(1+|k-2^{j-\nu}l|)^{-M_1}2^{-jp(\alpha-1/p+1/2)}|s_{j,l}|^p\right)^{1/p}\right\}^p\right]^{1/p} \\
\leq & \; C2^{(\nu-j)/p'}\left(\sum_{l\in\mathbb{Z}}(2^{-j(\alpha-1/p+1/2)}|s_{j,l}|)^p\right)^{1/p},
\end{aligned}
$$

where the last inequality follows from Lemma 3.3.1.
Inserting this estimate into I^q and using the condition on N_1, we find

$$
\begin{aligned}
I^q & \leq C\sum_{\nu\geq 1}\left[\sum_{j<\nu}2^{(j-\nu)(\alpha+N_1)}\left(\sum_{l\in\mathbb{Z}}(2^{-j(\alpha-1/p+1/2)}|s_{j,l}|)^p\right)^{1/p}\right]^{q/p} \\
& \leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}(2^{-\nu(\alpha-1/p+1/2)}|s_{\nu,k}|)^p\right]^{q/p}.
\end{aligned}
$$

Similarly, for the second term

$$
\begin{aligned}
II^q & \leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\left(\sum_{j\geq\nu}\sum_{l\in\mathbb{Z}}2^{(j-\nu)(\alpha-1/p-N_2)}\left(1+\frac{|2^\nu k-2^j l|}{2^j}\right)^{-M_2}2^{-j(\alpha-1/p+1/2)}|s_{j,l}|\right)^p\right]^{q/p} \\
& \leq C\sum_{\nu\geq 1}\left[\sum_{j\geq\nu}2^{(j-\nu)(\alpha-1/p-N_2)}\left(\sum_{k\in\mathbb{Z}}(\sum_{l\in\mathbb{Z}}\left(1+\frac{|2^\nu k-2^j l|}{2^j}\right)^{-M_2}2^{-j(\alpha-1/p+1/2)}|s_{j,l}|)^p\right)^{1/p}\right]^{q/p} \\
& \leq C\sum_{\nu\geq 1}\left[\sum_{j\geq\nu}2^{(j-\nu)(\alpha-N_2)}\left(\sum_{l\in\mathbb{Z}}(2^{-j(\alpha-1/p+1/2)}|s_{j,l}|)^p\right)^{1/p}\right]^{q/p} \\
& \leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}(2^{-\nu(\alpha-1/p+1/2)}|s_{\nu,k}|)^p\right]^{q/p}.
\end{aligned}
$$

Now consider the case $0 < p \leq 1$:

$$\begin{aligned}
I^q &\leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\left(\sum_{j<\nu,l\in\mathbb{Z}} 2^{(j-\nu)(\alpha-1/p+1/2+N_1+1/2)}\left(1+\frac{|2^\nu k-2^j l|}{2^\nu}\right)^{-M_1} 2^{-j(\alpha-1/p+1/2)}|s_{j,l}|\right)^p\right]^{q/p}\\
&\leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\sum_{j<\nu}\sum_{l\in\mathbb{Z}} 2^{p(j-\nu)(\alpha-1/p+N_1+1)}\left(1+\frac{|2^\nu k-2^j l|}{2^\nu}\right)^{-M_1 p} 2^{-jp(\alpha-1/p+1/2)}|s_{j,l}|^p\right]^{q/p}\\
&\leq C\sum_{\nu\geq 1}\left[\sum_{j<\nu}\sum_{l\in\mathbb{Z}} 2^{p(j-\nu)(\alpha-1/p+N_1+1)} 2^{-jp(\alpha-1/p+1/2)}|s_{j,l}|^p\right]^{q/p}\\
&\leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}(2^{-\nu(\alpha-1/p+1/2)}|s_{\nu,k}|)^p\right]^{q/p}.
\end{aligned}$$

Along the same lines, for the second term in case $0 < p \leq 1$

$$\begin{aligned}
II^q &\leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\left(\sum_{j\geq\nu}\sum_{l\in\mathbb{Z}} 2^{(j-\nu)(\alpha-1/p-N_2)}\left(1+\frac{|2^\nu k-2^j l|}{2^j}\right)^{-M_2} 2^{-j(\alpha-1/p+1/2)}|s_{j,l}|\right)^p\right]^{q/p}\\
&\leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}\sum_{j\geq\nu}\sum_{l\in\mathbb{Z}} 2^{p(j-\nu)(\alpha-1/p-N_2)}\left(1+\frac{|2^\nu k-2^j l|}{2^j}\right)^{-M_2 p} 2^{-jp(\alpha-1/p+1/2)}|s_{j,l}|^p\right]^{q/p}\\
&\leq C\sum_{\nu\geq 1}\left[\sum_{j\geq\nu}\sum_{l\in\mathbb{Z}} 2^{p(j-\nu)(\alpha-1/p-N_2)} 2^{(j-\nu)} 2^{-jp(\alpha-1/p+1/2)}|s_{j,l}|^p\right]^{q/p}\\
&= C\sum_{\nu\geq 1}\left[\sum_{j\geq\nu}\sum_{l\in\mathbb{Z}} 2^{p(j-\nu)(\alpha-N_2)} 2^{-jp(\alpha-1/p+1/2)}|s_{j,l}|^p\right]^{q/p}\\
&\leq C\sum_{\nu\geq 1}\left[\sum_{k\in\mathbb{Z}}(2^{-\nu(\alpha-1/p+1/2)}|s_{\nu,k}|)^p\right]^{q/p}.
\end{aligned}$$

Reversing the roles of j and ν in the above proof gives boundedness of $\tilde{A}$:

$$\begin{aligned}
\|\tilde{A}s\|^q_{b^{\alpha}_{p,q}(\mathbb{Z})} &= \sum_{j\geq 1}\left[\sum_{l\in\mathbb{Z}}\left(2^{-j(\alpha-1/p+1/2)}\Big|\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}\langle\varphi_{\nu,k},\tilde{h}_{j,l}\rangle s_{\nu,k}\Big|\right)^p\right]^{q/p}\\
&\leq C\left\{\sum_{j\geq 1}\left[\sum_{l\in\mathbb{Z}}\left(\sum_{j\geq\nu}\sum_{k\in\mathbb{Z}} 2^{(\nu-j)(\alpha-1/p+1/2)}|\langle\varphi_{\nu,k},\tilde{h}_{j,l}\rangle| 2^{-\nu(\alpha-1/p+1/2)}|s_{\nu,k}|\right)^p\right]^{q/p}\right.\\
&\quad\left.+\sum_{j\geq 1}\left[\sum_{l\in\mathbb{Z}}\left(\sum_{j<\nu}\sum_{k\in\mathbb{Z}} 2^{(\nu-j)(\alpha-1/p+1/2)}|\langle\varphi_{\nu,k},\tilde{h}_{j,l}\rangle| 2^{-\nu(\alpha-1/p+1/2)}|s_{\nu,k}|\right)^p\right]^{q/p}\right\}.
\end{aligned}$$

By Lemma 3.3.6, the size of $|\langle\tilde{h}_{j,l},\varphi_{\nu,k}\rangle| = |\overline{\langle\tilde{h}_{j,l},\varphi_{\nu,k}\rangle}| = |\langle\varphi_{\nu,k},\tilde{h}_{j,l}\rangle|$ again can be estimated by (3.3.10), (3.3.11) respectively, such that for $1 < p < \infty$, we can estimate the first term in the above sum by

$$C\sum_{j\geq 1}\left[\sum_{\nu\leq j}2^{(\nu-j)(\alpha+N_2)}\left(\sum_{k\in\mathbb{Z}}(2^{-\nu(\alpha-1/p+1/2)}|s_{\nu,k}|)^p\right)^{1/p}\right]^{q/p},$$

and the second one by

$$C\sum_{j\geq 1}\left[\sum_{\nu>j}2^{(\nu-j)(\alpha-N_1)}\left(\sum_{k\in\mathbb{Z}}(2^{-\nu(\alpha-1/p+1/2)}|s_{\nu,k}|)^p\right)^{1/p}\right]^{q/p}.$$

As in this case, we chose $N_1 > \alpha$, $N_2 > 1/(\min(1,p)) - 1 - \alpha$, $\tilde{A}$ is bounded and the case $0 < p \leq 1$ follows using the same arguments as above. ■

We are now ready to state our main result: under appropriate support, moment and regularity conditions on the biorthogonal wavelet families, the membership of a distribution $f = (f(n))_{n\in\mathbb{Z}}$ in a discrete-time Besov space is fully characterized by the decay of coefficients $(\langle f, \tilde{h}_{j,l}\rangle)_{j\geq 1, l\in\mathbb{Z}}$. Moreover, the associated analysis and synthesis operators are isomorphisms onto the full coefficient spaces $b^{\alpha}_{p,q}(\mathbb{Z})$, not just onto certain closed subspaces.

Theorem 3.3.8. *Let $\alpha \in \mathbb{R}$, $0 < p, q < \infty$, $N > \alpha$, $\tilde{N} > 1/(\min\{1,p\}) - 1 - \alpha$. Suppose that $(h_{j,l})$, $(\tilde{h}_{j,l})$, $j \geq 1, l \in \mathbb{Z}$, are biorthogonal wavelet bases for $\ell^2(\mathbb{Z})$, satisfying*

$$|supp\ h_j|, |supp\ \tilde{h}_j| \leq C2^j, \tag{3.3.12}$$
$$\|h_j\|_\infty \leq C2^{-j/2},\ \ \|\tilde{h}_j\|_\infty \leq C2^{-j/2} \tag{3.3.13}$$
$$\textstyle\sum_{n\in Z} n^i h_j(n) = 0 \ \ for\ i = 0, \ldots, \tilde{N}-1, \tag{3.3.14}$$
$$\textstyle\sum_{n\in Z} n^i \tilde{h}_j(n) = 0 \ \ for\ i = 0, \ldots, N-1, \tag{3.3.15}$$
$$(h_j)_{j\geq 1}\ regular\ of\ order\ N+\varepsilon\ (in\ the\ sense\ of\ 2.2.7),\ 0<\varepsilon\leq 1, \tag{3.3.16}$$
$$(\tilde{h}_j)_{j\geq 1}\ regular\ of\ order\ \tilde{N}+\tilde{\varepsilon},\ 0<\tilde{\varepsilon}\leq 1. \tag{3.3.17}$$

Then the following statements hold:

(a) The analysis operator $S_{\tilde{h}} : B^{\alpha}_{p,q}(\mathbb{Z}) \to b^{\alpha}_{p,q}(\mathbb{Z})$ is well-defined and continuous.

(b) The synthesis operator extends uniquely to a bounded operator $b^{\alpha}_{p,q}(\mathbb{Z}) \to B^{\alpha}_{p,q}(\mathbb{Z})$, also denoted by T_h. For arbitrary $(d_{j,l})_{j\geq 1, l\in\mathbb{Z}} \in b^{\alpha}_{p,q}(\mathbb{Z})$,

$$T_h\left((d_{j,l})_{j\geq 1,l\in\mathbb{Z}}\right) = \sum_{j,l} d_{j,l}h_{j,l}\ \ , \tag{3.3.18}$$

with unconditional convergence in the Besov space norm.

(c) $T_h \circ S_{\tilde{h}} = id_{B^\alpha_{p,q}(\mathbb{Z})}$, *and* $S_{\tilde{h}} \circ T_h = id_{b^\alpha_{p,q}(\mathbb{Z})}$. *Thus,* $B^\alpha_{p,q}(\mathbb{Z})$ *can be identified with* $b^\alpha_{p,q}(\mathbb{Z})$ *under the maps* $S_{\tilde{h}}$ *and* T_h.

(d) *We have the norm equivalence* $\|f\|_{B^\alpha_{p,q}(\mathbb{Z})} \asymp \|S_{\tilde{h}} f\|_{b^\alpha_{p,q}(\mathbb{Z})}$. *Moreover, the wavelet expansion*

$$f = \sum_{j,l} \langle f, \tilde{h}_{j,l} \rangle h_{j,l}$$

holds with unconditional convergence in the Besov space norm.

Proof (Compare to [14, 17].) For part (a), let $f \in B^\alpha_{p,q}(\mathbb{Z})$, $j \geq 1, l \in \mathbb{Z}$ and $K = \max\{[\alpha - 1/p], -1\}$. Then $K + 1 \geq N$, and thus

$$f = \sum_{\nu,k} \langle f, \varphi_{\nu,k} \rangle \varphi_{\nu,k}$$

holds in $(\mathcal{S}_K(\mathbb{Z}))'$, by Lemma 3.2.2. But $\tilde{h}_{j,l} \in \mathcal{S}_K(\mathbb{Z})$, and therefore

$$\langle f, \tilde{h}_{j,l} \rangle = \sum_{\nu,k} \langle f, \varphi_{\nu,k} \rangle \langle \varphi_{\nu,k}, \tilde{h}_{j,l} \rangle = (\tilde{A}((\langle f, \varphi_{\nu,k} \rangle)))(j,l) \quad .$$

Here we used the operator defined by the matrix $\tilde{A} := ((\varphi_{\nu,k}, \tilde{h}_{j,l}))_{\nu,k,j,l}$. In short, $S_{\tilde{h}} = \tilde{A} \circ S_\varphi$. By the support, size and moment conditions on $(\tilde{h}_{j,l})$, (3.3.12),(3.3.13),(3.3.15), Lemma 3.3.7 yields that $\tilde{A}$ is bounded on $b^\alpha_{p,q}(\mathbb{Z})$, whereas Theorem 3.2.1 contributes boundedness of S_φ. This proves part (a).

For the operator T_h, we consider the matrix $A = (\langle h_{j,l}, \varphi_{\nu,k} \rangle)_{j,l,\nu,k}$. Recall that $h_{j,l} = \sum_{\nu,k} \langle h_{j,l}, \varphi_{\nu,k} \rangle \varphi_{\nu,k}$ holds in $l^2(\mathbb{Z})$. Then for finitely supported sequences d,

$$\begin{aligned} T_h d &= \sum_{j,l} d_{j,l} h_{j,l} \\ &= \sum_{j,l} d_{j,l} \sum_{\nu,k} \langle h_{j,l}, \varphi_{\nu,k} \rangle \varphi_{\nu,k} \\ &= \sum_{\nu,k} \left(\sum_{j,l} d_{j,l} \langle h_{j,l}, \varphi_{\nu,k} \rangle \right) \varphi_{\nu,k} \\ &= (T_\varphi \circ A)(d) \quad . \end{aligned}$$

By the assumptions on the system $(h_j)_{j \geq 1}$, the operator A is bounded on $b^\alpha_{p,q}(\mathbb{Z})$. Hence on the finitely supported coefficient sequences - which are dense in $b^\alpha_{p,q}(\mathbb{Z})$ - T_h coincides with the bounded operator $T_\varphi \circ A$. But then T_h has a bounded extension to the whole space, and (3.3.18) in fact converges unconditionally for all $d \in b^\alpha_{p,q}(\mathbb{Z})$: The net of restrictions of d to finite subsets of $\mathbb{N} \times \mathbb{Z}$ converges in the norm on $b^{\alpha}_{p,q}(\mathbb{Z})$. Then boundedness of T_h implies that the net of finite partial sums in (3.3.18) converges also, which is unconditional convergence.

For the proof of part (c), consider $f \in B^{\alpha}_{p,q}(\mathbb{Z})$. Parts (a) and (b) imply that $T_h S_{\tilde{h}} f = \sum_{j,l} \langle f, \tilde{h}_{j,l} \rangle h_{j,l}$ with unconditional convergence in $B^{\alpha}_{p,q}(\mathbb{Z})$; therefore it remains to prove that f is the limit of the expansion. For this purpose we first prove that $f = \sum_{j,l} \langle f, \tilde{h}_{j,l} \rangle h_{j,l}$ holds in $(\mathcal{S}_K(\mathbb{Z}))'$.
Hence let $\overline{\eta} \in \mathcal{S}_K$. Then,

$$\begin{aligned}
\langle f, \eta \rangle &= \langle \sum_{\nu,k} \langle f, \varphi_{\nu,k} \rangle \varphi_{\nu,k}, \eta \rangle = \sum_{\nu,k} \langle f, \varphi_{\nu,k} \rangle \langle \varphi_{\nu,k}, \eta \rangle \\
&= \sum_{\nu,k} \sum_{j,l} \langle f, \varphi_{\nu,k} \rangle \langle \varphi_{\nu,k}, \tilde{h}_{j,l} \rangle \langle h_{j,l}, \eta \rangle \\
&\overset{(*)}{=} \sum_{j,l} \sum_{\nu,k} \langle f, \varphi_{\nu,k} \rangle \langle \varphi_{\nu,k}, \tilde{h}_{j,l} \rangle \langle h_{j,l}, \eta \rangle \\
&= \sum_{j,l} \langle f, \tilde{h}_{j,l} \rangle \langle h_{j,l}, \eta \rangle \\
&= \langle \sum_{j,l} \langle f, \tilde{h}_{j,l} \rangle h_{j,l}, \eta \rangle.
\end{aligned}$$

The order of summation in $(*)$ can be interchanged, because the series converges absolutely:
$(\langle f, \varphi_{\nu,k} \rangle)_{\nu,k} \in b^{\alpha}_{p,q}(\mathbb{Z})$ and the matrix $\tilde{A} := (\langle \varphi_{\nu,k}, \tilde{h}_{j,l} \rangle)_{\nu,k,j,l}$ is bounded on $b^{\alpha}_{p,q}(\mathbb{Z})$ by Lemma 3.3.7.
This yields

$$(\sum_{\nu,k} |\langle f, \varphi_{\nu,k} \rangle| |\langle \varphi_{\nu,k}, \tilde{h}_{j,l} \rangle|)_{j,l} \in b^{\alpha}_{p,q}(\mathbb{Z}),$$

and in particular

$$\sum_{\nu,k} |\langle f, \varphi_{\nu,k} \rangle| |\langle \varphi_{\nu,k}, \tilde{h}_{j,l} \rangle| \leq C 2^{j(\alpha - 1/p + 1/2)}.$$

We noted in the proof of Lemma 3.3.6 that (h_j) fulfills the requirements imposed on the family (φ_ν) in Lemma 3.3.2. Hence (3.3.3) implies

$$|\langle h_{j,l}, \eta \rangle| \leq C 2^{-j(K+1+1/2)} (1 + |l|)^{-M},$$

for $M > 0$, as $\overline{\eta} \in \mathcal{S}_K(\mathbb{Z})$.

Overall, this gives

$$\begin{aligned}\sum_{j,l}\sum_{\nu,k}|\langle f,\varphi_{\nu,k}\rangle\langle\varphi_{\nu,k},\tilde{h}_{j,l}\rangle\langle h_{j,l},\eta\rangle| &\leq \sum_{j,l}\sum_{\nu,k}|\langle f,\varphi_{\nu,k}\rangle||\langle\varphi_{\nu,k},\tilde{h}_{j,l}\rangle||\langle h_{j,l},\eta\rangle| \\ &\leq C\sum_{j,l}2^{j(\alpha-1/p+1/2)}2^{-j(K+1+1/2)}(1+|l|)^{-M} \\ &\leq C\sum_{j}2^{-j(K+1-\alpha+1/p)}<\infty.\end{aligned}$$

Hence $f=\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}\langle f,\tilde{h}_{j,l}\rangle h_{j,l}$ in $(\mathcal{S}_K(\mathbb{Z}))'$.
In order to prove $T_hS_{\tilde{h}}f=f$ in the Besov norm, note that in particular $\varphi_{\nu,k}\in\mathcal{S}_K(\mathbb{Z})$, which, together with Theorem 3.2.1, leads to

$$\begin{aligned}&\|f-\sum_{j=1}^{J}\sum_{l\in\mathbb{Z}}\langle f,\tilde{h}_{j,l}\rangle h_{j,l}\|^q_{B^{\alpha}_{p,q}(\mathbb{Z})} \\ \leq\quad & C\|(\langle f-\sum_{j=1}^{J}\sum_{l\in\mathbb{Z}}\langle f,\tilde{h}_{j,l}\rangle h_{j,l},\varphi_{\nu,k}\rangle)_{\nu,k}\|^q_{b^{\alpha}_{p,q}(\mathbb{Z})} \\ =\quad & C\|(\sum_{j=J+1}^{\infty}\sum_{l\in\mathbb{Z}}\langle f,\tilde{h}_{j,l}\rangle\langle h_{j,l},\varphi_{\nu,k}\rangle)_{\nu,k}\|^q_{b^{\alpha}_{p,q}(\mathbb{Z})} \\ \leq\quad & C\Big\{\sum_{\nu\geq 1}\Big[\sum_{k\in\mathbb{Z}}\Big(\sum_{j>J,j<\nu}\sum_{l\in\mathbb{Z}}2^{-\nu(\alpha-1/p+1/2)}|\langle f,\tilde{h}_{j,l}\rangle\langle h_{j,l},\varphi_{\nu,k}\rangle|\Big)^p\Big]^{q/p} \\ & +\sum_{\nu\geq 1}\Big[\sum_{k\in\mathbb{Z}}\Big(\sum_{j>J,j\geq\nu}\sum_{l\in\mathbb{Z}}2^{-\nu(\alpha-1/p+1/2)}|\langle f,\tilde{h}_{j,l}\rangle\langle h_{j,l},\varphi_{\nu,k}\rangle|\Big)^p\Big]^{q/p}\Big\} \\ =:\quad & C\Big\{I^q+II^q\Big\}.\end{aligned}$$

For $1<p<\infty$, we can proceed analogously to the proof of Lemma 3.3.7 and estimate the first term by

$$I^q\leq C\sum_{\nu\geq 1}\Big[\sum_{j>J,\,j<\nu}2^{(j-\nu)(\alpha+\tilde{N})}\Big(\sum_{l\in\mathbb{Z}}(2^{-j(\alpha-1/p+1/2)}|\langle f,\tilde{h}_{j,l}\rangle|)^p\Big)^{1/p}\Big]^{q/p},$$

and

$$II^q\leq C\sum_{\nu\geq 1}\Big[\sum_{j>J,\,j\geq\nu}2^{(j-\nu)(\alpha-N)}\Big(\sum_{l\in\mathbb{Z}}(2^{-j(\alpha-1/p+1/2)}|\langle f,\tilde{h}_{j,l}\rangle|)^p\Big)^{1/p}\Big]^{q/p}.$$

This yields

$$\lim_{J\to\infty}\|f-\sum_{j=1}^{J}\sum_{l\in\mathbb{Z}}\langle f,\tilde{h}_{j,l}\rangle h_{j,l}\|_{B^{\alpha}_{p,q}(\mathbb{Z})}=0. \tag{3.3.19}$$

The case $0 < p \leq 1$ follows along the same lines, using the estimates established in Lemma 3.3.7.

Furthermore, by the biorthogonality of $h_{j,l}$, $\tilde{h}_{j,l}$, any $s = (s_{i,m})_{i\geq 1, m\in\mathbb{Z}} \in b^{\alpha}_{p,q}(\mathbb{Z})$ can be written as

$$(s_{i,m})_{i,m} = (\langle \sum_{j,l} s_{j,l} h_{j,\,l}, \tilde{h}_{i,m}\rangle)_{i,m}\,,$$

which gives $S_{\tilde{h}} \circ T_h = id_{b^{\alpha}_{p,q}(\mathbb{Z})}$.
Part (d) is immediate from (c). ∎

Remark 3.3.9. Torres [31] characterizes the spaces $B^{\alpha}_{p,q}(\mathbb{Z})$ as spaces of sequences obtained by sampling band-limited distributions in $\dot{B}^{\alpha}_{p,q}(\mathbb{R})$.
Theorem 3.3.8 provides another relation between the $B^{\alpha}_{p,q}(\mathbb{Z})$ and the $\dot{B}^{\alpha}_{p,q}(\mathbb{R})$ spaces more in terms of multiresolution analysis.
Let $F \in \dot{B}^{\alpha}_{p,q}(\mathbb{R})$, $F = \sum_{n\in\mathbb{Z}} a_n \tau_n \varphi$, where φ a scaling function associated to a multiresolution analysis (i.e. $F \in V_0$ in MRA language).
Recalling that we have $F \in \dot{B}^{\alpha}_{p,q}(\mathbb{R})$ if and only if the discrete wavelet coefficients $(d_{j,l})_{j\geq 1, l\in\mathbb{Z}} \in b^{\alpha}_{p,q}(\mathbb{Z})$. By our theorem, this is in fact equivalent to $(a_n)_{n\in\mathbb{Z}}$ to be in $B^{\alpha}_{p,q}(\mathbb{Z})$.

3.4 'Intrinsic' Characterizations of $B^{\alpha}_{p,q}(\mathbb{Z})$

There are other, more 'intrinsic' possibilities to describe Besov spaces in discrete time than using Littlewood-Paley theory. E.g. [30] contains a description in terms of mean oscillation properties of sequences for some special cases of the parameters α, p, q. Using our wavelet characterization result, it is easy to extend this kind of description to the whole parameter family. This will be the issue of paragraph 3.4.2.

Before we give this result, however, we can give another characterization of the discrete-time Besov spaces in terms of iterated differences. This result is a consequence of Theorem 3.3.8.

3.4.1 Discrete-Time Moduli of Smoothness

Analogously to the continuous-time function spaces, the discrete-time Besov spaces possess a description via differences. We adapt the notion *modulus of smoothness* to functions given in discrete time, and show that for a certain range of parameters α, p, q, the arising spaces coincide with the $B^{\alpha}_{p,q}(\mathbb{Z})$-spaces defined via Littlewood-Paley theory (3.1.3).

For the corresponding theory on $\mathbb{R}$, see Chapter 1, or [25] or [34].

Let $m \in \mathbb{Z}$. For a sequence $f = (f(n))_{n\in\mathbb{Z}}$, define the (forward) *difference operator of step* m by

$$\Delta_m f(n) = f(n+m) - f(n),$$

and for $r \in \mathbb{N}_+$, define the *difference operator of order* r, step m, inductively by

$$\Delta_m^r f(n) = \Delta_m(\Delta_m^{r-1} f(n)).$$

Note that the $r-$th difference operator in explicit form is given by

$$\Delta_m^r f(n) = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} f(n+km).$$

Definition 3.4.1. *For* $1 < p < \infty$, $t \in \mathbb{R}_+$, *the* r*-th order* ***modulus of smoothness*** *of* f *in* $l^p(\mathbb{Z})$ *is defined by*

$$\omega_p^r(f,t) = \sup_{m\in\mathbb{Z}, |m|<t} \|\Delta_m^r f(\cdot)\|_p.$$

The $l^p(\mathbb{Z})-$moduli of smoothness share properties of their $L^p(\mathbb{R})-$analogs, see [25, 34]. In the following, we will list some of them which will be needed further on.

1. $\omega_p^r(f,t)$ is an increasing function of t.
2. For $1 \le s \le r$ and each $t \in \mathbb{R}_+$
$$\omega_p^r(f,t) \le 2^{r-s}\omega_p^s(f,t), \tag{3.4.1}$$
and moreover, if $f \in l^p(\mathbb{Z})$
$$\omega_p^r(f,t) \le 2^r \|f\|_p. \tag{3.4.2}$$
3. Let f, g be defined on $\mathbb{Z}$. Then, for each $t \in \mathbb{R}_+$,
$$\omega_p^r(f+g,t) \le \omega_p^r(f,t) + \omega_p^r(g,t), \tag{3.4.3}$$
and for f multiplied by a scalar α,
$$\omega_p^r(\alpha f,t) \le |\alpha|\omega_p^r(f,t). \tag{3.4.4}$$
As $\omega_p^r(f,t)$ vanishes for polynomials on $\mathbb{Z}$ of degree $\le r-1$, $\omega_p^r(\cdot,t)$ is a semi-norm on the set of sequences for which $\omega_p^r(f,t) < \infty$ for all $t \in \mathbb{R}_+$.

4. For $M \in \mathbb{N}$,
$$\omega_p^r(f, M \cdot t) \leq M^r \omega_p^r(f,t). \tag{3.4.5}$$
So if $\omega_p^r(f,t) < \infty$ for some $t > 0$, it is finite for all $t \in \mathbb{R}_+$.

Definition 3.4.2. *For $\alpha > 0, 1 < p,q < \infty$, $r = \lfloor \alpha \rfloor + 1$, the sequence f is said to be in $B^{\alpha}_q(l^p(\mathbb{Z}))$ if*

$$\|f\|_{B^{\alpha}_q(l^p(\mathbb{Z}))} := (\sum_{j\geq 1}(2^{-j\alpha}\omega_p^r(f,2^j))^q)^{1/q} < \infty. \tag{3.4.6}$$

The $\|\cdot\|_{B^{\alpha}_q(l^p(\mathbb{Z}))}$ are semi-norms in general because of the polynomial cancellation properties of the moduli of smoothness; they become norms modulo polynomials on $\mathbb{Z}$ of degree $\leq r-1$. Furthermore, the $B^{\alpha}_q(l^p(\mathbb{Z}))$-norms are all equivalent modulo polynomials using different moduli of smoothness $r > \alpha$ in the definition.

Our aim is to show that the $B^{\alpha}_q(l^p(\mathbb{Z}))$-spaces coincide with the discrete-time Besov spaces, at least for the range of parameters given in Definition 3.4.2.

We start our preparations for this by considering an orthonormal discrete-time wavelet basis for $\ell^2(\mathbb{Z})$, $(h_{j,l})_{j\geq 1, l\in\mathbb{Z}}$ with associated scaling sequences $(g_{j,l})_{j\geq 1, l\in\mathbb{Z}}$, satisfying

$$|\operatorname{supp} g_{j,l}|,\ |\operatorname{supp} h_{j,l}| \leq C2^j, , \tag{3.4.7}$$
$$\|g_{j,l}\|_\infty \leq C2^{-j/2},\ \ \|h_{j,l}\|_\infty \leq C2^{-j/2} \tag{3.4.8}$$
$$. \tag{3.4.9}$$

In Remark 2.2.3, we noted that the family of projections $(P_j)_{j\geq 1}$, $P_j f = \sum_{l\in\mathbb{Z}}\langle f, g_{j,l}\rangle g_{j,l}$ defines a decreasing sequence $V_j = P_j(\ell^2(\mathbb{Z}))$ of closed subspaces which share many properties of an MRA in $L^2(\mathbb{R})$. Let the spaces W_j be defined likewise, using the projections $Q_j f = \sum_{l\in\mathbb{Z}}\langle f, h_{j,l}\rangle h_{j,l}$.

Obviously, we have for $F_j \in V_j$, $F_j = \sum_{l\in\mathbb{Z}} a_{j,l} g_{j,l}$,

$$\|F_j\|^2_{\ell^2(\mathbb{Z})} = \sum_{l\in\mathbb{Z}} |a_{j,l}|^2$$

and analogously for $G_j \in W_j$, $G_j = \sum_{l\in\mathbb{Z}} d_{j,l} h_{j,l}$

$$\|G_j\|^2_{\ell^2(\mathbb{Z})} = \sum_{l\in\mathbb{Z}} |d_{j,l}|^2.$$

We will now investigate the behavior of the projections in spaces $\ell^p(\mathbb{Z})$, for $p \neq 2$. The following Lemma relates the p-norm of functions in V_j, W_j to the p-norms of their coefficients. This type of result is sometimes called 'p-stability'; for analogous results in continuous time, see e.g [34], Section 8.1.

Lemma 3.4.3. *Let $1 < p < \infty$, $(g_{j,l})$, $(h_{j,l})$ satisfying (3.4.7), (3.4.8). Then, for $F_j \in V_j$, $F_j = \sum_{l\in\mathbb{Z}} a_{j,l} g_{j,l}$, $j \geq 1$,*

$$\|F_j\|_{\ell^p(\mathbb{Z})} \asymp 2^{-j/2} 2^{j/p} (\sum_{l\in\mathbb{Z}} |a_{j,l}|^p)^{1/p}, \tag{3.4.10}$$

as well as for $G_j \in W_j$, $G_j = \sum_{l\in\mathbb{Z}} d_{j,l} h_{j,l}$, $j \geq 1$,

$$\|G_j\|_{\ell^p(\mathbb{Z})} \asymp 2^{-j/2} 2^{j/p} (\sum_{l\in\mathbb{Z}} |d_{j,l}|^p)^{1/p}. \tag{3.4.11}$$

Proof As $(g_{j,l})$ satisfies (3.4.7), (3.4.8), we have especially for any $M > 0$ that

$$|g_{j,l}(n)| \leq C 2^{-j/2} (1 + 2^{-j}|n - 2^j l|)^{-M} \text{ for any } M > 0,$$

see the argument in the proof of Lemma 3.3.6. Using this together with Hölder's inequality gives for $1/p + 1/p' = 1$

$$\begin{aligned}
\|F_j\|^p_{\ell^p(\mathbb{Z})} &= \|\sum_{l\in\mathbb{Z}} a_{j,l} g_{j,l}(\cdot)\|^p_{\ell^p(\mathbb{Z})} \\
&\leq \sum_{n\in\mathbb{Z}} (\sum_{l\in\mathbb{Z}} |a_{j,l}||g_{j,l}(n)|^{1/p}|g_{j,l}(n)|^{1/p'})^p \\
&\leq C \cdot \sum_{n\in\mathbb{Z}} (\sum_{l\in\mathbb{Z}} |a_{j,l}| 2^{-j/2p} (1 + 2^{-j}|n - 2^j l|)^{-M/p} 2^{-j/2p'} (1 + 2^{-j}|n - 2^j l|)^{-M/p'})^p \\
&\leq C \cdot \sum_{n\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} |a_{j,l}|^p 2^{-j/2} (1 + 2^{-j}|n - 2^j l|)^{-M} (\sum_{j\in\mathbb{Z}} 2^{-j/2} (1 + 2^{-j}|n - 2^j l|)^{-M})^{p/p'}.
\end{aligned}$$

With the help of Lemma 3.3.1, we know that

$$\sum_{l\in\mathbb{Z}} (1 + 2^{-j}|n - 2^j l|)^{-M} \leq C,$$

as well as

$$\sum_{n\in\mathbb{Z}} (1 + 2^{-j}|n - 2^j l|)^{-M} \leq C \cdot 2^j,$$

such that

$$\|F_j\|^p_{\ell^p(\mathbb{Z})} \leq C \cdot 2^{-j/2} 2^j 2^{-jp/2p'} \sum_{l\in\mathbb{Z}} |a_{j,l}|^p = C \cdot 2^j 2^{-jp/2} \sum_{l\in\mathbb{Z}} |a_{j,l}|^p.$$

The converse inequality follows with the same arguments, see e.g. Proposition 8.1 in [34]. (3.4.11) follows immediately from the above discussion. ■

Let now $(h_{j,l})$, $(g_{j,l})$ satisfy

$$|\operatorname{supp} g_{j,l}|,\ |\operatorname{supp} h_{j,l}| \leq C2^j, \tag{3.4.12}$$

$$\|g_{j,l}\|_\infty \leq C2^{-j/2},\ \ \|h_{j,l}\|_\infty \leq C2^{-j/2} \tag{3.4.13}$$

$$\textstyle\sum_{n\in Z} n^i h_{j,l}(n) = 0 \ \text{ for } i = 0, \ldots, N-1, \tag{3.4.14}$$

$$(g_{j,l}),\ (h_{j,l}) \text{ regular of order } N+\varepsilon,\ 0 < \varepsilon \leq 1. \tag{3.4.15}$$

Below, it will be useful to express Theorem 3.3.8 in terms of the projections P_j, Q_j. We obtain by Lemma 3.4.3 that for $N > \alpha$, the following conditions are equivalent:

$$f \in B^{\alpha}_{p,q}(\mathbb{Z}) \tag{3.4.16}$$

$$(\langle f, h_{j,l}\rangle)_{j\geq 1, l\in\mathbb{Z}} \in b^{\alpha}_{p,q}(\mathbb{Z}) \tag{3.4.17}$$

$$\textstyle(\sum_{j\geq 1}(2^{-j\alpha}\|Q_j f\|_p)^q)^{1/q} < \infty, \tag{3.4.18}$$

$$\textstyle(\sum_{j\geq 1}(2^{-j\alpha}\|f - P_j f\|_p)^q)^{1/q} < \infty, \tag{3.4.19}$$

where the equivalence of (3.4.18) and (3.4.19) easily follows from the fact that

$$\|f - P_j f\|_p \leq \sum_{i=1}^{j} \|Q_j f\|_p$$

and by summing up the geometric series.

The next step towards our intended result is the following Lemma, which in literature often is called an **inequality of Bernstein-type**:

Lemma 3.4.4. *Let $1 < p < \infty$, $r \in \mathbb{N}_+$ and let $(g_{j,l})_{j\geq 1, l\in\mathbb{Z}}$ satisfy (3.4.12), (3.4.13) and (3.4.15) with $N \geq r$.*

For $F_i \in V_i$, $i \geq 1$, we have for any $j \geq 1$

$$\omega^r_p(F_i, 2^j) \leq C \min(2^{(j-i)r}, 1)\|F_i\|_p. \tag{3.4.20}$$

Proof Let $F_i \in V_i \cap l^p(\mathbb{Z})$. By Lemma 3.4.3, $F_i = \sum_{l\in\mathbb{Z}} a_{i,l} g_{i,l}$ with $\|F_i\|_p \asymp 2^{-i/2} \cdot 2^{i/p}\|(a_{i,l})_{l\in\mathbb{Z}}\|_p$.

In the case $j > i$, we have immediately $\omega^r_p(F_i, 2^j) \leq 2^r\|F_i\|_p$ by (3.4.2), so we only

treat the case $j \le i$.
Consider the r-th order difference operator of step 1:

$$\begin{aligned}\|\Delta_1^r F_i(\cdot)\|_p^p &= \sum_{n\in\mathbb{Z}} |\sum_{l\in\mathbb{Z}} a_{i,l}\Delta_1^r g_{i,l}(n)|^p \\ &\le \sum_{n\in\mathbb{Z}} (\sum_{l\in\mathbb{Z}} |a_{i,l}||\Delta_1^r g_{i,l}(n)|^{1/p}|\Delta_1^r g_{i,l}(n)|^{1/p'})^p,\end{aligned}$$

where $1/p + 1/p' = 1$.
As $N \ge r$, we have by the regularity assumption (3.4.15) that $|\Delta_1^r g_{i,l}(n)| = 2^{-ir}|D_i^r g_{i,l}(n)| \le C 2^{-ir}2^{-i/2}$. Together with Hölder's inequality, this gives

$$\begin{aligned}\|\Delta_1^r F_i(\cdot)\|_p^p &\le \sum_{n\in\mathbb{Z}} (\sum_{l\in\mathbb{Z}} |a_{i,l}|^p|\Delta_1^r g_{i,l}(n)|)(\sum_{l\in\mathbb{Z}} |\Delta_1^r g_{i,l}(n)|)^{p/p'} \\ &\le C\sum_{l\in\mathbb{Z}} |a_{i,l}|^p 2^i \cdot 2^{-i(r+1/2)} \cdot 2^{-i(r+1/2)(p/p')} \\ &\le C \cdot 2^{-irp} \cdot 2^{i(1-p/2)} \sum_{l\in\mathbb{Z}} |a_{i,l}|^p \\ &\le C \cdot 2^{-irp} \|F_i\|_p^p,\end{aligned}$$

where the last inequality follows from Lemma 3.4.3.
This yields the result, as by (3.4.5) and (3.4.2)

$$\omega_p^r(F_i, 2^j) \le C \cdot 2^{jr}\omega_p^r(F_i, 2) \le C \cdot 2^{(j-i)r}\|F_i\|_p.$$

■

Next, we relate the size of coefficients at a given scale to the modulus of smoothness, deriving an **inequality of Jackson-type**.

Lemma 3.4.5. *Let $1 < p < \infty$, $r \in N_+$, and let $(h_{j,l})_{j\ge 1, l\in\mathbb{Z}}$ be an orthonormal wavelet basis for $\ell^2(\mathbb{Z})$, satisfying (3.4.12), (3.4.13) and (3.4.14) for some $N \ge r$. For any $j \ge 1$,*

$$\|(\langle f, h_{j,l}\rangle)_{l\in\mathbb{Z}}\|_p \le C \cdot 2^{j/2}2^{-j/p}\omega_p^r(f, 2^j). \tag{3.4.21}$$

Proof Let C_h be the smallest *integer* such that $|\text{supp } h_{j,l}| \le C_h \cdot 2^j$. Without loss of generality, we assume h to be causal, i.e. supp h_j = supp $h_{j,0} \subseteq [0, C_h \cdot 2^j[$.
Let $N \in \mathbb{N}$, $0 < -k < N$. Then (see 3.3.4),

$$f(n-k) = \sum_{i=0}^{N} \binom{N}{i} \Delta^i f(n-k-N).$$

Set

$$p_n(k) = \sum_{i=0}^{N-1} \binom{N}{i} \Delta^i f(n-k-N).$$

Write

$$\begin{aligned}\langle f, h_{j,l}\rangle &= \sum_{n\in \text{ supp } h_{j,l}} f(n)\overline{h_{j,l}(n)} = \sum_{n\in \text{ supp } h_{j,l}} f(n)h^*_j(2^j l - n)\\ &= \sum_{k=-C_h\cdot 2^j+1}^{0} h^*_j(k) f(2^j l - k),\end{aligned}$$

where $h^*_j(n) = \overline{h_j(-n)}$.

Due to the vanishing moment condition (3.4.14) on $(h_{j,l})$,

$$\begin{aligned}|\langle f, h_{j,l}\rangle|^p &= |\sum_{k=-C_h\cdot 2^j+1}^{0} h^*_j(k)(f(2^j l - k) - p_{2^j l}(k))|^p\\ &\le (\sum_{k=-C_h\cdot 2^j+1}^{0} |h^*_j(k)|^{p'})^{p/p'} \cdot \sum_{k=-C_h\cdot 2^j+1}^{0} |\Delta^N f(2^j l - k - N)|^p,\end{aligned}$$

where $1/p + 1/p' = 1$.

By Lemma 3.4.3, $\|h_j\|_{p'} \asymp 2^{-j/2}2^{j/p'} = 2^{j/2}2^{-j/p}$. Hence, summing over l, and observing that each $m \in \mathbb{Z}$ is in the support of at most C_h shifts of h_j, we obtain

$$\begin{aligned}\left(\sum_{l\in\mathbb{Z}} |\langle f, h_{j,l}\rangle|^p\right)^{1/p} &\le C\cdot 2^{j/2}2^{-j/p}\left(\sum_{l\in\mathbb{Z}}\sum_{k=-C_h\cdot 2^j+1}^{0} |\Delta^N f(2^j l - k - N)|^p\right)^{1/p}\\ &\le C\cdot 2^{j/2}2^{-j/p}\left(\sum_{m\in\mathbb{Z}} |\Delta^N f(m)|^p\right)^{1/p}\\ &\le C\cdot 2^{j/2}2^{-j/p}\omega^N_p(f,2)\\ &\le C\cdot 2^{j/2}2^{-j/p}\omega^r_p(f,2^j)\ ,\end{aligned}$$

where we used the monotonicity properties of ω^r_p, specifically (3.4.1), and $N \ge r$.

■

We will make use of both of the above inequalities in order to show the equivalence of the semi-norm (3.4.6) and the Besov semi-norm.

Theorem 3.4.6. *Let $\alpha > 0, 1 < p, q < \infty$.*
The spaces $B^{\alpha}_{q}(l^p(\mathbb{Z}))$ and $B^{\alpha}_{p,q}(\mathbb{Z})$ coincide (modulo polynomials) and moreover,

$$\|\cdot\|_{B^{\alpha}_{q}(l^p(\mathbb{Z}))} \asymp \|\cdot\|_{B^{\alpha}_{p,q}(\mathbb{Z})}. \tag{3.4.22}$$

Proof (Compare to [34].) Let $(V_j)_{j\geq 1}$ a $N-$regular multiresolution analysis with $N > \alpha$. Let $P_j f$ be the orthogonal projection of $f = (f(n))_{n\in\mathbb{Z}}$ onto V_j.
Then, $f = f - P_j f + \sum_{i\geq j} P_i f - P_{i+1} f$. As $P_i f - P_{i+1} f \in V_i$ and $N \geq r = \lfloor\alpha\rfloor + 1$ we can employ (3.4.20) which, together with (3.4.2) and the triangle inequality, gives

$$\begin{aligned} \omega^r_p(f, 2^j) &\leq \omega^r_p(f - P_j f, 2^j) + \sum_{i\geq j} \omega^r_p(P_i f - P_{i+1} f, 2^j) \\ &\leq 2^r \|f - P_j f\|_p + C \sum_{i\geq j} 2^{j-i} \|P_i f - P_{i+1} f\|_p \\ &\leq C \sum_{i\geq j} 2^{j-i} \|f - P_i f\|_p. \end{aligned}$$

So, using Minkowski's inequality and (3.4.19), we get

$$\begin{aligned} \|f\|_{B^{\alpha}_{q}(l^p(\mathbb{Z}))} &= (\sum_{j\geq 1} (2^{-j\alpha} \omega^r_p(f, 2^j))^q)^{1/q} \\ &\leq C(\sum_{j\geq 1} (2^{-j\alpha} \sum_{i\geq j} 2^{j-i} \|f - P_i f\|_p)^q)^{1/q} \\ &\leq C(\sum_{j\geq 1} (2^{-j\alpha} \|f - P_j f\|_p)^q)^{1/q} \leq C\|f\|_{B^{\alpha}_{p,q}(\mathbb{Z})}. \end{aligned}$$

The converse inequality easily follows from (3.4.21):

$$\begin{aligned} \|f\|_{B^{\alpha}_{p,q}(\mathbb{Z})} &\leq C(\sum_{j\geq 1} (\sum_{l\in\mathbb{Z}} (2^{-j(\alpha+1/2-1/p)} |\langle f, h_{j,l}\rangle|)^p)^{q/p})^{1/q} \\ &\leq C(\sum_{j\geq 1} (2^{-j\alpha} \omega^r_p(f, 2^j))^q)^{1/q} = C\|f\|_{B^{\alpha}_{q}(l^p(\mathbb{Z}))}. \end{aligned}$$

■

Theorem 3.3.8 is a first example how the wavelet characterization can be employed for further analysis of discrete-time Besov spaces. In particular, the finite support of the wavelets has greatly facilitated the proof.

Another such characterization, also a direct consequence of Theorem 3.3.8, is given in the next subsection.

3.4.2 Mean Oscillation Characterization of $B^{\alpha}_{p,q}(\mathbb{Z})$

In this section, we derive a description of the discrete-time Besov spaces in terms of oscillations, similarly to the mean oscillation characterization of their continuous-time counterparts, see [8, 9].

For the special case $B^{1/p}_{p,p}(\mathbb{Z})$, $1 < p < \infty$, R.H. Torres [30] showed that the corresponding norm is equivalent to the $B_p(\mathbb{Z})$-norm, defined by

$$\|f\|_{B_p(\mathbb{Z})} := (\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}(\frac{1}{|I_{j,l}|}\sum_{n\in I_{j,l}}|f(n) - f_{I_{j,l}}|)^p)^{1/p}, \tag{3.4.23}$$

where $I_{j,l} := [2^j l, 2^j(l+4)]$ and $f_{I_{j,l}}$ the average of f on $I_{j,l}$.

This defines a semi-norm in general and a norm modulo constants.

Here, we extend this result to Besov spaces where $\alpha > 0$, $1 < p, q < \infty$, using Theorem 3.3.8. We use a slightly different notation compared to the articles cited above, in particular the $B_p(\mathbb{Z})$-spaces will correspond to $MO^{1/p}_{p,p}(\mathbb{Z})$ defined below.

Definition 3.4.7. *A family of intervals $(I_{j,l})_{j\geq 1, l\in\mathbb{Z}}$ is called a* family of admissible coverings *if*

1. *Any of the intervals is of the form $I_{j,l} = [2^j l, 2^j l + L_j[$, where*
2. *for $j \to \infty$, $2^{-j}\cdot L_j \to C$, where $C > 1$.*

For any $j \geq 1$, the family $J_j := (I_{j,l})_{l\in\mathbb{Z}}$ is the family of *enlarged* dyadic intervals, satisfying $\bigcup_{l\in\mathbb{Z}} I_{j,l} \cap \mathbb{Z} = \mathbb{Z}$.

Let $f = (f(n))_{n\in\mathbb{Z}}$, I an arbitrary interval and $m \in \mathbb{N}_0$. By $f^{(m)}_I(n)$, we denote the (unique) polynomial on $\mathbb{Z}$ of degree smaller or equal to m, such that

$$\sum_{n\in I}(f(n) - f^{(m)}_I(n))n^k = 0 \text{ for } k = 0, 1, \ldots, m. \tag{3.4.24}$$

By definition, $f^{(m)}_I$ is the best approximation of f on I by a polynomial of degree m, measured in the $l^2(\mathbb{Z})$-norm.

The mean oscillation norm is now defined by the norms of the residuals.

Definition 3.4.8. *Let $1 < p, q < \infty$, $m \in \mathbb{N}_+$ and let $(I_{j,l})_{j\geq 1, l\in\mathbb{Z}}$ be a family of admissible coverings, where $J_j = (I_{j,l})_{l\in\mathbb{Z}}$ the covering of the line at scale j. Define for $f = (f(n))_{n\in\mathbb{Z}}$*

$$osc_{p,m}(f, J_j) = (\sum_{l\in\mathbb{Z}}(\frac{1}{|I_{j,l}|}\sum_{n\in I_{j,l}}|f(n) - f^{(m)}_{I_{j,l}}(n))|)^p)^{1/p}. \tag{3.4.25}$$

Let $\alpha > 0$, $m = \lfloor \alpha \rfloor$. A sequence f is in $MO^{\alpha}_{p,q}(\mathbb{Z})$ if

$$\|f\|_{MO^{\alpha}_{p,q}(\mathbb{Z})} := (\sum_{j\geq 1}(2^{-j(\alpha-1/p)} osc_{p,m}(f, J_j))^q)^{1/q} < \infty. \tag{3.4.26}$$

The semi-norms in (3.4.26) are norms modulo discrete-time polynomials.

Using a different family of admissible coverings or a different $m > \lfloor \alpha \rfloor$ results in equivalent norms, see Section 9 in [9].

Just like in the previous section, we use Theorem 3.3.8 to show that the Besov and mean oscillation norms are in fact equivalent.

So, let in the following again $(h_{j,l})_{j\geq 1, l\in\mathbb{Z}}$ be an orthogonal wavelet basis for $\ell^2(\mathbb{Z})$, satisfying (3.4.12), (3.4.13), (3.4.14) and (3.4.15) for some $N > \alpha$

Theorem 3.4.9. *Let $\alpha > 0, 1 < p, q < \infty$.*
The spaces $MO^{\alpha}_{p,q}(\mathbb{Z})$ and $B^{\alpha}_{p,q}(\mathbb{Z})$ coincide (modulo polynomials) and moreover,

$$\|\cdot\|_{MO^{\alpha}_{p,q}(\mathbb{Z})} \asymp \|\cdot\|_{B^{\alpha}_{p,q}(\mathbb{Z})}. \tag{3.4.27}$$

Proof First, let $I^h_{j,l} :=$ supp $h_{j,l}$.
Due to the moment condition (3.4.14) on $h_{j,l}$ and as $N > m = \lfloor \alpha \rfloor$,

$$\begin{aligned}
|\langle f, h_{j,l}\rangle| &= |\langle f - f^{(m)}_{I^h_{j,l}}, h_{j,l}\rangle| \\
&\leq \|h_{j,l}\|_\infty \sum_{n\in I^h_{j,l}} |f(n) - f^{(m)}_{I^h_{j,l}}(n)| \\
&\leq C2^{-j/2} \sum_{n\in I^h_{j,l}} |f(n) - f^{(m)}_{I^h_{j,l}}(n)| \\
&\leq C2^{j/2} \frac{1}{|I^h_{j,l}|} \sum_{n\in I^h_{j,l}} |f(n) - f^{(m)}_{I^h_{j,l}}(n)|.
\end{aligned}$$

Using theorem 3.3.8,

$$\begin{aligned}
\|f\|_{B^{\alpha}_{p,q}(\mathbb{Z})} &\leq C\|(\langle f, h_{j,l}\rangle)\|_{b^{\alpha}_{p,q}(\mathbb{Z})} \\
&= C(\sum_{j\geq 1}(\sum_{l\in\mathbb{Z}}(2^{-j(\alpha+1/2-1/p)}|\langle f, h_{j,l}\rangle|)^p)^{q/p})^{1/q} \\
&\leq C(\sum_{j\geq 1}(2^{-j(\alpha-1/p)}(\sum_{l\in\mathbb{Z}}(\frac{1}{|I^h_{j,l}|}\sum_{n\in I^h_{j,l}} |f(n) - f^{(m)}_{I^h_{j,l}}(n))|)^p))^{q/p})^{1/q},
\end{aligned}$$

where the last term is equivalent to the $MO^{\alpha}_{p,q}(\mathbb{Z})$-norm, as $(I^h_{j,l})_{j,l}$ is a family of admissible coverings (eventually, the indexing has to be changed in order to be conform with the first postulation of Definition 3.4.7, as the $h_{j,l}$ are not necessarily causal).

In order to prove the converse inequality, note that for any polynomial p of degree less or equal to m and a given interval I (see [8], proof of Theorem 1)

$$\frac{1}{|I|}\sum_{n\in I}|f(n)-f^{(m)}_I(n)| \le C\frac{1}{|I|}\sum_{n\in I}|f(n)-p(n)|. \tag{3.4.28}$$

Let $(I_{j,l})$ be a family of admissible coverings. Define $p^m(n) := \sum_{i=0}^m \Delta^i f(n-(m+1))$. As $f(n)-p^m(n) = \Delta^{m+1}f(n-(m+1))$ and by (3.4.28),

$$\begin{aligned}
\frac{1}{|I_{j,l}|}\sum_{n\in I_{j,l}}|f(n)-f^{(m)}_{I_{j,l}}(n))| &\le C\frac{1}{|I_{j,l}|}\sum_{n\in I_{j,l}}|f(n)-p^m(n)| \\
&\le C\frac{1}{|I_{j,l}|}\sum_{n\in I_{j,l}}|\Delta^{m+1}f(n-(m+1))|,
\end{aligned}$$

which gives

$$\begin{aligned}
\|f\|_{MO^{\alpha}_{p,q}(\mathbb{Z})} &\le C(\sum_{j\ge 1}(2^{-j(\alpha-1/p)}(\sum_{l\in\mathbb{Z}}(\frac{1}{|I_{j,l}|}\sum_{n\in I_{j,l}}|\Delta^{m+1}f(n-(m+1))|)^p))^{q/p})^{1/q} \\
&\le C(\sum_{j\ge 1}(2^{-j(\alpha-1/p)}2^{-j/p}(\sum_{n\in I_{j,l}}|\Delta^{m+1}f(n-m+1)|^p)^{1/p})^q)^{1/q} \\
&\le C(\sum_{j\ge 1}(2^{-j\alpha}\omega_p^{m+1}(f,2^j))^q)^{1/q} \\
&\le C\|f\|_{B^{\alpha}_{p,q}(\mathbb{Z})},
\end{aligned}$$

using (3.4.22) as $r = m+1$ by definition. ∎

Remark 3.4.10.

The usage of enlarged dyadic intervals in the definition of the $MO^{\alpha}_{p,q}$-spaces is crucial. Consider the usual non-overlapping dyadic interval family $(I^*_{j,l})_{j\ge 1, l\in\mathbb{Z}}$, where $I^*_{j,l} = [2^j l, 2^j(l+1)[$. This family is not admissible in terms of our definition.

Define the spaces $B^*_p(\mathbb{Z})$, $1<p<\infty$ as the collection of sequences (modulo constants) for which

$$\|f\|_{B^*_p(\mathbb{Z})} := (\sum_{j\ge 1}\sum_{l\in\mathbb{Z}}(\frac{1}{|I^*_{j,l}|}\sum_{n\in I^*_{j,l}}|f(n)-f_{I^*_{j,l}}|)^p)^{1/p}$$

is finite.

Let $(f(n))_{n\in\mathbb{Z}}$ be the sequence defined by $f(n) = 0$ for $n < 0$ and $f(n) = 1$, $n \geq 0$. Obviously, $f \in B^*_p(\mathbb{Z})$, as the $B^*_p(\mathbb{Z})$-norm will not feel the 'discontinuity' at 0, whereas the $B_p(\mathbb{Z})$-norm (using admissible families) will:
Consider the interval family $(I_{j,l})_{j\geq 1, l\in\mathbb{Z}}$, $I_{j,l} = [2^j l, 2^j(l+2)[$, which is admissible.

$$\begin{aligned}\|f\|_{B_p(\mathbb{Z})} = \|f\|_{MO^{1/p}_{p,p}(\mathbb{Z})} &= (\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}(\frac{1}{|I_{j,l}|}\sum_{n\in I_{j,l}}|f(n) - f_{I_{j,l}}|)^p)^{1/p} \\ &= (\sum_{j\geq 1}(2^{-(j+1)}(2^{j+1}\cdot\frac{1}{2}))^p)^{1/p} = (\sum_{j\geq 1}2^{-p})^{1/p},\end{aligned}$$

which gives $f \notin B_p(\mathbb{Z})$.

This also indicates that the regularity condition is crucial for the wavelet characterization of the spaces $MO^{\alpha}_{p,q}(\mathbb{Z})$ and thus for $B^{\alpha}_{p,q}(\mathbb{Z})$:
Consider the discrete-time Haar filters (2.1.5).
The discrete-time Haar wavelet bases and scaling sequences read as

$$G_j(n) = \begin{cases} 2^{-j/2} & \text{for } n = 0, \dots, 2^j - 1; \\ 0 & \text{otherwise} \end{cases} \tag{3.4.29}$$

and

$$H_j(n) = \begin{cases} -2^{-j/2} & \text{for } n = 0, \dots, 2^{j-1} - 1; \\ 2^{-j/2} & \text{for } n = 2^{j-1}, \dots, 2^j - 1; \\ 0 & \text{otherwise} \end{cases} \tag{3.4.30}$$

with the usual translation $G_{j,l} = G_j(\cdot - 2^j l)$, $H_{j,l} = H_j(\cdot - 2^j l)$.
The discrete-time Haar system $(H_{j,l})_{j\geq 1, l\in\mathbb{Z}}$ is an orthonormal basis for $\ell^2(\mathbb{Z})$, possessing one vanishing moment but not being regular in the sense of (2.2.7). It is easy to show that due to the moment condition, for $f \in B_p(\mathbb{Z})$,

$$\|(\langle f, H_{j,l}\rangle)\|_{b^{1/p}_{p,p}(\mathbb{Z})} \leq C\|f\|_{B_p(\mathbb{Z})}.$$

But the converse inequality cannot be obtained.
In fact, one can adapt the above arguments to show that $\|(\langle f, H_{j,l}\rangle)\|_{b^{1/p}_{p,p}(\mathbb{Z})}$ is equivalent to the $B^*_p(\mathbb{Z})$-norm, which is *not the same as* $B_p(\mathbb{Z})$.
In a sense, the required overlap of admissible intervals is related to the support size, and thus to the regularity of the discrete time wavelets (recall the correlation between support size and regularity for the Daubechies family.)

Chapter 4

Discrete-Time Triebel-Lizorkin Spaces

For the reader familiar with function spaces, it will not be surprising that the wavelet-based treatment of discrete-time Besov spaces can be extended to discrete-time versions of Triebel-Lizorkin spaces.

These spaces of sequences were studied by Q. Sun in [29], more in terms of smooth atomic decompositions. With our notions at hand, we can give a description in terms of discrete-time bases.

In the first section, we give the definition of the Triebel-Lizorkin spaces in discrete time and their φ-transform characterization.

In Subsection 4.2, we establish a result analogous to Theorem 3.3.8 for this type of spaces. The techniques used will be mostly the same as for the Besov spaces.

For background information concerning Triebel-Lizorkin spaces in continuous time, we refer to the usual literature as e.g. [32], [17] or [16].

4.1 Definition and φ-transform Decomposition

Let φ^c a phi-function (see 3.1.2) and set $\varphi^c_\nu(x) = 2^{-\nu+2}\varphi^c(2^{-\nu+2}x)$ for $\nu \in \mathbb{Z}$.
Consider again the family $(\varphi_\nu)_{\nu\geq 1}$, obtained by $\varphi_\nu := \varphi^c_\nu|_{\mathbb{Z}}$ for $\nu > 1$ and $\varphi_1 := ((\chi_{[-\pi,\pi]}\hat{\varphi}^c_1)^\vee)|_{\mathbb{Z}}$. Recall that for $f \in \mathcal{S}'(\mathbb{Z})$, we have $f = \sum_{\nu\geq 1} f * \varphi_\nu * \varphi^*_\nu$ with unconditional convergence in $\mathcal{S}'/\mathcal{P}(\mathbb{Z})$ (3.1.6).

Definition 4.1.1. *For $\alpha \in \mathbb{R}$, $0 < p, q < \infty$, the* ***discrete-time Triebel-Lizorkin space*** *$F^\alpha_{p,q}(\mathbb{Z})$ is the collection of all $f \in \mathcal{S}'/\mathcal{P}(\mathbb{Z})$, such that*

$$\|f\|_{F^\alpha_{p,q}(\mathbb{Z})} := \|(\sum_{\nu\geq 1}(2^{-\nu\alpha}|f * \varphi_\nu|)^q)^{1/q}\|_p < \infty.$$

This definition is independent of the choice of φ^c [29].

Definition 4.1.2. *Let the space of (truncated) coefficient families $f^{\alpha}_{p,q}(\mathbb{Z})$, for $\alpha \in \mathbb{R}$, $0 < p, q < \infty$, be the collection of complex-valued sequences $s = (s_{j,l})_{j\geq 1, l\in\mathbb{Z}}$, for which*

$$\|s\|_{f^{\alpha}_{p,q}(\mathbb{Z})} := \|(\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}(2^{-j\alpha}|s_{j,l}|\tilde{\chi}_{j,l})^q)^{1/q}\|_p < \infty,$$

where $\tilde{\chi}_{j,l}(n) = \begin{cases} 2^{-j/2}, & \text{if } 2^j l \leq n < 2^j(l+1); \\ 0, & \text{otherwise.} \end{cases}$

We note that the definition of $f^{\alpha}_{p,q}(\mathbb{Z})$ is similar to that of $b^{\alpha}_{p,q}(\mathbb{Z})$, but uses a different summation order (first over scales, then over positions) and the characteristic functions $\tilde{\chi}_{j,l}$. Hence the coefficient space $b^{\alpha}_{p,q}(\mathbb{Z})$ is somewhat easier to handle by comparison.

In [29], Q. Sun relates the above spaces via the $\varphi-$transform, similarly to the result for the discrete-time Besov and $b^{\alpha}_{p,q}$ spaces in section 3.2.

For $\nu, k \in \mathbb{Z}$ let again

$$\varphi^c_{\nu,k}(x) = 2^{(-\nu+2)/2}\varphi^c(2^{-\nu+2}x - k),$$

and for $k \in \mathbb{Z}$, define $\varphi_{\nu,k} = \varphi^c_{\nu,k}|_{\mathbb{Z}}$ for $\nu > 1$ and $\varphi_{1,k} = \tau_k\varphi_1$.

Recall that for any $f \in \mathcal{S}'(\mathbb{Z})$ (3.2.1)

$$f = \sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}\langle f, \varphi_{\nu,k}\rangle\varphi_{\nu,k}, \tag{4.1.1}$$

with convergence in $\mathcal{S}'/\mathcal{P}(\mathbb{Z})$.

Let the φ-transform S_φ for $f \in \mathcal{S}'(\mathbb{Z})$ be defined by $S_\varphi f = s = (s_{\nu,k})_{\nu\geq 1, k\in\mathbb{Z}}$, where $s_{\nu,k} = \langle f, \varphi_{\nu,k}\rangle$, and for a complex-valued sequence $t = (t_{\nu,k})_{\nu\geq 1,k\in\mathbb{Z}}$ define the inverse φ-transform by T_φ by $T_\varphi t = \sum_{\nu,k} t_{\nu,k}\varphi_{\nu,k}$. The convergence of the sum is justified by the following result.

Theorem 4.1.3. *([29]) Let $\alpha \in \mathbb{R}$, $0 < p, q < \infty$.*
Both of the operators $S_\varphi : F^{\alpha}_{p,q}(\mathbb{Z}) \to f^{\alpha}_{p,q}(\mathbb{Z})$ and $T_\varphi : f^{\alpha}_{p,q}(\mathbb{Z}) \to F^{\alpha}_{p,q}(\mathbb{Z})$ are bounded with $\|f\|_{F^{\alpha}_{p,q}(\mathbb{Z})} \asymp \|S_\varphi f\|_{f^{\alpha}_{p,q}(\mathbb{Z})}$ and $T_\varphi \circ S_\varphi = id_{B^{\alpha}_{p,q}(\mathbb{Z})}$.

Remark 4.1.4. Considering the special choice $\alpha = 0$, $1 < p < \infty$, $q = 2$,

$$\|f\|_{F^0_{p,2}(\mathbb{Z})} = \|(\sum_{\nu\geq 1}(|f * \varphi_\nu|)^2)^{1/2}\|_p,$$

one can see that this is exactly the Littlewood-Paley type definition of the $\ell^p(\mathbb{Z})$-norm (see again [29]).

Thus, our result in the upcoming section includes the discrete-time wavelet characterization of the spaces $\ell^p(\mathbb{Z})$, $1 < p < \infty$.

4.2 Wavelet Characterization of Discrete-Time Triebel-Lizorkin Spaces

For our wavelet description of the spaces $F^{\alpha}_{p,q}(\mathbb{Z})$, we need an additional tool, and certain inequalities related to it.

This will be the issue of the next definition and the Lemmata 4.2.2, 4.2.3 and 4.2.4.

Definition 4.2.1. *Let* $f = (f(n))_{n\in\mathbb{Z}}$.
Define the ***Hardy-Littlewood maximal operator on*** $\mathbb{Z}$ *by*

$$Mf(k) := \sup_{a\leq k<b,\ a,b\in\mathbb{Z}} \frac{1}{b-a} \sum_{a\leq n<b} |f(n)|. \tag{4.2.1}$$

The next Lemma is taken from [29] and can be viewed as a discrete-time version of the *Fefferman-Stein maximal inequality*:

Lemma 4.2.2. *Let* $1 < p, q < \infty$. *For any family of sequences* $(f_i)_{i\in\mathbb{Z}}$

$$\|(\sum_{i\in\mathbb{Z}} |Mf_i|^q)^{1/q}\|_p \leq C\|(\sum_{i\in\mathbb{Z}} |f_i|^q)^{1/q}\|_p. \tag{4.2.2}$$

A similar result holds for $0 < p, q \leq 1$.

In this case, replacing the maximal operator M by M_r, $0 < r < \min(p, q)$, in Lemma (4.2.2), defined by

$$M_r f(k) := (\sup_{a\leq k<b,\ a,b\in\mathbb{Z}} \frac{1}{b-a} \sum_{a\leq n<b} |f(n)|^r)^{1/r}, \tag{4.2.3}$$

yields

Lemma 4.2.3.

$$\|(\sum_{i\in\mathbb{Z}} |M_r f_i|^q)^{1/q}\|_p \leq C\|(\sum_{i\in\mathbb{Z}} |f_i|^q)^{1/q}\|_p. \tag{4.2.4}$$

We will also need the following inequalities:

The first one can be found in [14] and is easily adapted to the discrete case.

Lemma 4.2.4. *Let $(s_{j,l})_{j\geq 1, l\in\mathbb{Z}} \subset \mathbb{C}$. Fix $\nu, j \geq 1$, $k \in \mathbb{Z}$ and $2^\nu k \leq n \leq 2^\nu(k+1)$. For $M_1 > 1$*

$$\sum_{l\in\mathbb{Z}} |s_{j,l}|(1 + \frac{|2^\nu k - 2^j l|}{2^\nu})^{-M_1} \leq C\, 2^{(\nu-j)} M(\sum_{l\in\mathbb{Z}} |s_{j,l}|\chi_{j,l})(n), \ \text{if } j < \nu, \tag{4.2.5}$$

$$\sum_{l\in\mathbb{Z}} |s_{j,l}|(1 + \frac{|2^\nu k - 2^j l|}{2^j})^{-M_1} \leq C\, M(\sum_{l\in\mathbb{Z}} |s_{j,l}|\chi_{j,l})(n), \ \text{if } j \geq \nu. \tag{4.2.6}$$

Let $0 < r < 1$, $M_2 > 1/r$. Then,

$$\sum_{l\in\mathbb{Z}} |s_{j,l}|(1 + \frac{|2^\nu k - 2^j l|}{2^\nu})^{-M_2} \leq C\, 2^{\frac{\nu-j}{r}} M_r(\sum_{l\in\mathbb{Z}} |s_{j,l}|\chi_{j,l})(n), \ \text{if } j < \nu, \tag{4.2.7}$$

$$\sum_{l\in\mathbb{Z}} |s_{j,l}|(1 + \frac{|2^\nu k - 2^j l|}{2^j})^{-M_2} \leq C\, M_r(\sum_{l\in\mathbb{Z}} |s_{j,l}|\chi_{j,l})(n), \ \text{if } j \geq \nu. \tag{4.2.8}$$

With these results at hand, we are able to establish a similar boundedness result as in Lemma 3.3.7:

Lemma 4.2.5. *Let $\alpha \in \mathbb{R}$, $0 < p, q < \infty$ and let $(\varphi_{\nu,k})_{\nu\geq 1, k\in\mathbb{Z}}$ be the $\varphi-$transform family.*
If $(h_{j,l})_{j\geq 1, l\in\mathbb{Z}}$ satisfies the conditions of Lemma 3.3.6 with $N_1 > 1/(\min(1,p,q)) - 1 - \alpha$, $N_2 > \alpha$, then the matrix $A := (\langle h_{j,l}, \varphi_{\nu,k}\rangle)_{j,l,\nu,k}$, defines a bounded operator on $f^\alpha_{p,q}(\mathbb{Z})$, where $As = (\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}\langle h_{j,l}, \varphi_{\nu,k}\rangle s_{j,l})_{\nu,k}$ for $s \in f^\alpha_{p,q}(\mathbb{Z})$.

Also, for $(\tilde{h}_{j,l})_{j\geq 1, l\in\mathbb{Z}}$, satisfying the conditions of Lemma 3.3.6, where now $N_1 > \alpha$, $N_2 > 1/(\min(1,p,q)) - 1 - \alpha$, the matrix $\tilde{A} := (\langle \varphi_{\nu,k}, \tilde{h}_{j,l}\rangle)_{\nu,k,j,l}$, where $\tilde{A}s = (\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}\langle\varphi_{\nu,k}, \tilde{h}_{j,l}\rangle s_{\nu,k})_{j,l}$, is bounded on $f^\alpha_{p,q}(\mathbb{Z})$ as well.

Proof

$$\begin{aligned}
\|As\|_{f^\alpha_{p,q}} &= \|(\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(2^{-\nu\alpha}|\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}\langle h_{j,l}, \varphi_{\nu,k}\rangle s_{j,l}|\tilde{\chi}_{\nu,k})^q)^{1/q}\|_p \\
&\leq \|(\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(2^{-\nu\alpha}\sum_{j\geq 1}\sum_{l\in\mathbb{Z}}|\langle h_{j,l}, \varphi_{\nu,k}\rangle||s_{j,l}|\tilde{\chi}_{\nu,k})^q)^{1/q}\|_p \\
&\leq C\ \Big\{\|(\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(2^{-\nu\alpha}\sum_{j<\nu}\sum_{l\in\mathbb{Z}}|\langle h_{j,l}, \varphi_{\nu,k}\rangle||s_{j,l}|\tilde{\chi}_{\nu,k})^q)^{1/q}\|_p \\
&+ \|(\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(2^{-\nu\alpha}\sum_{j\geq\nu}\sum_{l\in\mathbb{Z}}|\langle h_{j,l}, \varphi_{\nu,k}\rangle||s_{j,l}|\tilde{\chi}_{\nu,k})^q)^{1/q}\|_p\Big\} \\
&=: C\Big\{I + II\Big\}.
\end{aligned}$$

Consider the case $1 < p, q < \infty$.
Using Lemma 3.3.6 and (4.2.6), we can estimate the first term by

$$\begin{aligned} I &\leq C \ \|(\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(\sum_{j<\nu} 2^{(j-\nu)(N_1+1/2)} \sum_{l\in\mathbb{Z}}(1+\frac{|2^\nu k - 2^j l|}{2^\nu})^{-M_1}|s_{j,l}|2^{-\nu\alpha}\tilde{\chi}_{\nu,k})^q)^{1/q}\|_p \\ &\leq C \ \|((\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(\sum_{j<\nu} 2^{(j-\nu)(N_1+1/2-1)}2^{-\nu\alpha}M(\sum_{l\in\mathbb{Z}}|s_{j,l}|\chi_{j,l})(n)\tilde{\chi}_{\nu,k}(n))^q)^{1/q})_{n\in\mathbb{Z}}\|_p \\ &= C \ \|((\sum_{\nu\geq 1}(\sum_{j<\nu} 2^{(j-\nu)(N_1+1/2-1+\alpha+1/2)}M(\sum_{l\in\mathbb{Z}} 2^{-j\alpha}|s_{j,l}|\tilde{\chi}_{j,l})(n))^q)^{1/q})_{n\in\mathbb{Z}}\|_p \\ &\leq C \ \|(\sum_{\nu\geq 1}(M(\sum_{k\in\mathbb{Z}} 2^{-\nu\alpha}|s_{\nu,k}|\tilde{\chi}_{\nu,k}))^q)^{1/q}\|_p, \end{aligned}$$

where we sum up a geometric series the in last inequality, as $N_1 > -\alpha$ by assumption. The inequality in Lemma 4.2.2 yields

$$I \leq C \ \|(\sum_{\nu\geq 1}(\sum_{k\in\mathbb{Z}} 2^{-\nu\alpha}|s_{\nu,k}|\tilde{\chi}_{\nu,k})^q)^{1/q}\|_p \leq C \ \|s\|_{f^\alpha_{p,q}(\mathbb{Z})}. \tag{4.2.9}$$

Along the same lines, we can estimate the second term by

$$\begin{aligned} II &\leq C \ \|(\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(\sum_{j\geq\nu} 2^{(j-\nu)(N_1+1/2)} \sum_{l\in\mathbb{Z}}(1+\frac{|2^\nu k - 2^j l|}{2^j})^{-M_2}|s_{j,l}|2^{-\nu\alpha}\tilde{\chi}_{\nu,k})^q)^{1/q}\|_p \\ &\leq C \ \|((\sum_{\nu\geq 1}(\sum_{j\geq\nu} 2^{(\nu-j)(N_2+1/2-\alpha-1/2)}M(\sum_{l\in\mathbb{Z}} 2^{-j\alpha}|s_{j,l}|\tilde{\chi}_{j,l})(n))^q)^{1/q})_{n\in\mathbb{Z}}\|_p \\ &\leq C \ \|s\|_{f^\alpha_{p,q}(\mathbb{Z})}, \end{aligned}$$

as we assumed $N_2 > \alpha$.

Now to the case $0 < p, q \leq 1$. Let $0 < r < \min(p, q)$. By Lemmata 3.3.6, 4.2.3 and (4.2.8)

$$\begin{aligned} I &\leq C \ \|(\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(\sum_{j<\nu} 2^{(j-\nu)(N_1-1/r+\alpha+1)}M_r(\sum_{l\in\mathbb{Z}} 2^{-j\alpha}|s_{j,l}|\tilde{\chi}_{j,l}))^q)^{1/q}\|_p \\ &\leq C \ \|s\|_{f^\alpha_{p,q}(\mathbb{Z})}, \end{aligned}$$

as $N_1 > 1/r - 1 - \alpha$ by assumption.

The other term can be treated analogously. ■

Once we established the above result, we immediately get an analogon to Theorem 3.3.8:

Let $(h_{j,l})$, $(\tilde{h}_{j,l})$, $j \geq 1, l \in \mathbb{Z}$, are biorthogonal wavelet bases for $\ell^2(\mathbb{Z})$. Recall the operators S_h, T_h: for $f \in S'(\mathbb{Z})$, let $S_h f = (\langle f, h_{j,l}\rangle)_{j\geq 1, l\in\mathbb{Z}}$. For finitely supported coefficient sequences $d = (d_{j,l})_{j\geq 1, l\in\mathbb{Z}}$, define

$$T_h d = \sum_{j,l} d_{j,l} h_{j,l}. \tag{4.2.10}$$

Let $S_{\tilde{h}}$ and $T_{\tilde{h}}$ be defined likewise.

Theorem 4.2.6. *Let $\alpha \in \mathbb{R}$, $0 < p, q < \infty$, $N > \alpha$, $\tilde{N} > 1/(\min\{1, p, q\}) - 1 - \alpha$. Suppose that $(h_{j,l})$, $(\tilde{h}_{j,l})$, $j \geq 1, l \in \mathbb{Z}$, are biorthogonal wavelet bases for $\ell^2(\mathbb{Z})$, satisfying*

$$\begin{gathered}
|supp\ h_j|, |supp\ \tilde{h}_j| \leq C2^j, \\
\|h_j\|_\infty \leq C2^{-j/2}, \quad \|\tilde{h}_j\|_\infty \leq C2^{-j/2}, \\
\textstyle\sum_{n \in Z} n^i h_j(n) = 0 \ \ for\ i = 0, \ldots, \tilde{N} - 1, \\
\textstyle\sum_{n \in Z} n^i \tilde{h}_j(n) = 0 \ \ for\ i = 0, \ldots, N - 1, \\
(h_j)_{j \geq 1} \text{ regular of order } N + \varepsilon \text{ (in the sense of 2.2.7)}, \ 0 < \varepsilon \leq 1, \\
(\tilde{h}_j)_{j \geq 1} \text{ regular of order } \tilde{N} + \tilde{\varepsilon}, \ 0 < \tilde{\varepsilon} \leq 1.
\end{gathered}$$

Then the following statements hold:

(a) The analysis operator $S_{\tilde{h}} : F^\alpha_{p,q}(\mathbb{Z}) \to f^\alpha_{p,q}(\mathbb{Z})$ is well-defined and continuous.

(b) The synthesis operator extends uniquely to a bounded operator $f^\alpha_{p,q}(\mathbb{Z}) \to F^\alpha_{p,q}(\mathbb{Z})$, also denoted by T_h. For arbitrary $(d_{j,l})_{j \geq 1, l \in \mathbb{Z}} \in f^\alpha_{p,q}(\mathbb{Z})$,

$$T_h\left((d_{j,l})_{j \geq 1, l \in \mathbb{Z}}\right) = \sum_{j,l} d_{j,l} h_{j,l} \quad , \tag{4.2.11}$$

with unconditional convergence in the Triebel-Lizorkin space norm.

(c) $T_h \circ S_{\tilde{h}} = id_{F^\alpha_{p,q}(\mathbb{Z})}$, and $S_{\tilde{h}} \circ T_h = id_{f^\alpha_{p,q}(\mathbb{Z})}$. Thus, $F^\alpha_{p,q}(\mathbb{Z})$ can be identified with $f^\alpha_{p,q}(\mathbb{Z})$ under the maps $S_{\tilde{h}}$ and T_h.

(d) We have the norm equivalence $\|f\|_{F^\alpha_{p,q}(\mathbb{Z})} \asymp \|S_{\tilde{h}} f\|_{f^\alpha_{p,q}(\mathbb{Z})}$. Moreover, the wavelet expansion

$$f = \sum_{j,l} \langle f, \tilde{h}_{j,l} \rangle h_{j,l}$$

holds with unconditional convergence in the Triebel-Lizorkin space norm.

Proof The structure of the proof is the same as for Theorem 3.3.8.

As Lemma 3.2.2 immediately carries over to the discrete-time Triebel-Lizorkin spaces, we have that for $f \in F^\alpha_{p,q}(\mathbb{Z})$, $j \geq 1, l \in \mathbb{Z}$ and $K = \max\{[\alpha - 1/p], -1\}$,

$$f = \sum_{\nu,k} \langle f, \varphi_{\nu,k} \rangle \varphi_{\nu,k}$$

holds in $(\mathcal{S}_K(\mathbb{Z}))'$. We have again $T_h = T_\varphi \circ A$ and $S_h = \tilde{A} \circ S_\varphi$, hence boundedness of A and $\tilde{A}$ (as provided by Lemma 4.2.5) yields (a), and the first part of (b).

For unconditional convergence, we note by applying the dominated convergence theorem, it can be shown that every $(d_{j,l}) \in f^{\alpha}_{p,q}(\mathbb{Z})$ is the limit of the net of its finite restrictions. This and boundedness of T_h yields unconditional convergence of (4.2.11).

Again, we have $(T_h \circ S_h)f = f \in (\mathcal{S}_K(\mathbb{Z}))'$ and also in the norm. In analogy to (3.3.19), the tail $\sum_{j=J+1}^{\infty}\sum_{l\in\mathbb{Z}}\langle f, \tilde{h}_{j,l}\rangle h_{j,l}$ converges again strongly to 0 as $J \to \infty$:

$$\begin{aligned}
&\|f - \sum_{j=1}^{J}\sum_{l\in\mathbb{Z}}\langle f, \tilde{h}_{j,l}\rangle h_{j,l}\|_{F^{\alpha}_{p,q}(\mathbb{Z})} \\
\leq\ & C\|(\langle f - \sum_{j=1}^{J}\sum_{l\in\mathbb{Z}}\langle f, \tilde{h}_{j,l}\rangle\langle h_{j,l}, \varphi_{\nu,k}\rangle)_{\nu,k}\|_{f^{\alpha}_{p,q}(\mathbb{Z})} \\
=\ & C\|(\sum_{j=J+1}^{\infty}\sum_{l\in\mathbb{Z}}\langle f, \tilde{h}_{j,l}\rangle\langle h_{j,l}, \varphi_{\nu,k}\rangle)_{\nu,k}\|_{f^{\alpha}_{p,q}(\mathbb{Z})} \\
\leq\ & C\ \Big\{\|(\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(2^{-\nu\alpha}\sum_{j>J,j<\nu}\sum_{l\in\mathbb{Z}}|\langle f, \tilde{h}_{j,l}\rangle\langle h_{j,l}, \varphi_{\nu,k}\rangle|\tilde{\chi}_{\nu,k})^{q})^{1/q}\|_{p} \\
+\ & \|(\sum_{\nu\geq 1}\sum_{k\in\mathbb{Z}}(2^{-\nu\alpha}\sum_{j>J,j\geq\nu}\sum_{l\in\mathbb{Z}}|\langle f, \tilde{h}_{j,l}\rangle\langle h_{j,l}, \varphi_{\nu,k}\rangle|\tilde{\chi}_{\nu,k})^{q})^{1/q}\|_{p}\Big\} \\
=:\ & C\Big\{I + II\Big\}
\end{aligned}$$

Let $1 < p < \infty$. We can use the proof of Lemma 4.2.5 to estimate the truncated series.

For the first term,

$$I \leq C\|(\sum_{\nu\geq 1}(\sum_{j>J,j<\nu}2^{(j-\nu)(\tilde{N}+\alpha)}M(\sum_{l\in\mathbb{Z}}2^{-j\alpha}|\langle f, \tilde{h}_{j,l}\rangle|\tilde{\chi}_{j,l}))^{q})^{1/q}\|_{p},$$

and

$$II \leq C\|(\sum_{\nu\geq 1}(\sum_{j>J,j\geq\nu}2^{(\nu-j)(N-\alpha)}M(\sum_{l\in\mathbb{Z}}2^{-j\alpha}|\langle f, \tilde{h}_{j,l}\rangle|\tilde{\chi}_{j,l}))^{q})^{1/q}\|_{p}$$

Summing up geometric series and applying Lemma 4.2.2 yields

$$\begin{aligned}
I\ \leq\ & C\ \|(\sum_{\nu>J}(M(\sum_{k\in\mathbb{Z}}2^{-\nu\alpha}|\langle f, \tilde{h}_{\nu,k}\rangle|\tilde{\chi}_{\nu,k}))^{q})^{1/q}\|_{p} \\
\leq\ & C\ \|(\sum_{\nu>J}(\sum_{k\in\mathbb{Z}}2^{-\nu\alpha}|\langle f, \tilde{h}_{\nu,k}\rangle|\tilde{\chi}_{\nu,k})^{q})^{1/q}\|_{p} \\
=\ & C\ \|(\langle f, \tilde{h}_{\nu,k}\rangle)\chi_{\nu>J}\|_{f^{\alpha}_{p,q}(\mathbb{Z})},
\end{aligned}$$

where $\chi_{\nu>J}$ is the characteristic function of $\{(\nu,k) : \nu > J\}$. This shows that $I \to 0$ for $J \to \infty$. The same arguments apply to II and the case $p \leq 1$ follows in the same fashion from the according estimates established in Lemma 4.2.5.

This yields

$$\lim_{J\to\infty} \|f - \sum_{j=1}^{J} \sum_{l\in\mathbb{Z}} \langle f, \tilde{h}_{j,l}\rangle h_{j,l}\|_{F^{\alpha}_{p,q}(\mathbb{Z})} = 0.$$

Again, biorthogonality of $h_{j,l}$, $\tilde{h}_{j,l}$ gives $S_{\tilde{h}} \circ T_h = id_{b^{\alpha}_{p,q}(\mathbb{Z})}$ and part (d) is immediate from the above.

■

Discussion And Outlook

We finish this thesis with a discussion of our results and an outlook on possible further work.

The central result in this thesis is the study of necessary and sufficient conditions on wavelet bases for $l^2(\mathbb{Z})$ to constitute unconditional bases for discrete-time Besov spaces $B^\alpha_{p,q}(\mathbb{Z})$ in Theorem 3.3.8.

Thus, the heuristics from continuous-time theory are substantiated by our result: Discrete-time wavelet coefficient decay can be characterized in terms of membership in a suitable sequence space.

More precisely, the arising coefficients are in $b^\alpha_{p,q}(\mathbb{Z})$, if and only if the sequence is in the space $B^\alpha_{p,q}(\mathbb{Z})$.

The results of Chapter 3 have provided a variety of new characterizations of Besov spaces in discrete time, showing that these spaces are worthwhile objects of study. Several possible ways of exploiting these characterizations suggest themselves, some of them inspired by existing results for the continuous domain. Specifically, we mention (roughly in the order of importance) the following list of problems:

- *Extending the results to higher dimensions:* In particular for applications in image processing, a two-dimensional result would be desirable. For the continuous domain case, extensions of the wavelet characterization to arbitrary dimensions have been obtained, and we expect that analogous results should hold for $\mathbb{Z}^d$. However, a proof of such a result will have to deal with even more involved notation. Also, the precise choice of a Littlewood-Paley type characterization of Besov spaces in higher dimension can be expected to have a strong influence, perhaps not on the spaces it characterizes, but on the effort necessary to prove characterizations in terms of wavelet systems obtained from tensor products of one-dimensional wavelets and scaling functions.

- *Developing discrete time heuristics for signal processing algorithms:* An obvious task in further work is to throw light on applications such as compression or denoising from a fully discrete-time viewpoint, which our result provides. For the sake of concreteness, let us only mention denoising. The results of Donoho and Johnstone rely on the embedding in the continuous time setting.

It should be possible to derive analogues of these results which rely on Theorem 3.3.8 instead, without assuming a 'true' underlying continuous-time function.

- *Extending the characterization to operators:* The study of continuity properties of suitable operators with respect to various Besov norms is a natural application of the wavelet characterization (see e.g. [17], Chapter 8). The operators could be discrete-time Calderón-Zygmund operators, or non-linear operators describing histogram equalization.

- *Studying further relations between discrete time and continuous time Besov spaces:* In [30], Torres proves a sampling theorem for Besov spaces showing that a Besov function in discrete time can be understood as restriction of a bandlimited Besov function in continuous time. This gives rise to an embedding of discrete time into continuous time Besov space via the sinc function.

  On the other hand, we have the often cited embedding of discrete time Besov space into its continuous time analog via the scaling function of an MRA with suitable smoothness and vanishing moment properties, which are not fulfilled by the sinc function. Thus there exist two fundamentally different embeddings.

  A description of the continuous time theory as asymptotic case of the discrete time theory (in a suitable sense) could close this gap, and provide additional insight into the discrete and continuous-time spaces. Note that also the relation between discrete-time and continuous-time wavelet systems via large scale limits, as described in Section 2.2.3, seems to indicate such a connection.

  These results could also shed additional light on the relationship between algorithm heuristics derived from discrete time and continuous time Besov space theory, respectively.

Bibliography

[1] A. Chambolle, R.A. DeVore, N.-Y. Lee and B.J. Lucier. Nonlinear Wavelet Image Processing: Variational Problems, Compression, and Noise Removal Through Wavelet Shrinkage. *IEEE Trans. Image Process.*, 7:319–335, 1998.

[2] A. Cohen, I. Daubechies, Ingrid, O.G. Guleryuz, Onur G. and M.T. Orchard. On the Importance of Combining Wavelet-Based Nonlinear Approximation with Coding Strategies. *IEEE Trans. Inform. Theory*, 48:1895–1921, 2002.

[3] A. Cohen, R.D. Ryan. *Wavelets and Multiscale Signal Processing*. Chapman and Hall, 1995.

[4] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhäuser Boston, 2003.

[5] D. Donoho, I. Johnstone. Minimax Estimation Via Wavelet Shrinkage. *Ann. Stat.*, 26(3):879–921, 1998.

[6] I. Daubechies. Orthonormal Bases Of Compactly Supported Wavelets. *Commun. on Pure and Appl. Math.*, 41:909–996, 1988.

[7] R. DeVore. Nonlinear Approximation. *Acta Numerica*, pages 51–150, 1998.

[8] J. Dorronsoro. Mean Oscillation and Besov Spaces. *Canad. Math. Bull*, 28(4):474–480, 1985.

[9] F. Ricci, M. Taibleson. Boundary Values of Harmonic Functions in Mixed Norm Spaces and Their Atomic Structure. *Ann. Sc. Norm. Super. Pisa*, IV(10):1–54, 1983.

[10] H. Choi, R.G. Baraniuk. Multiple Basis Wavelet Denoising Using Besov Projections. *ICIP*, 1:595–599, 1999.

[11] H. Führ, M. Wild. Characterizing Wavelet Coefficient Decay of Discrete-Time Signals. *Appl. and Comp. Harm. Anal.*, 20(2):184–201, March 2006.

[12] I. Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.

[13] J. Fan and J.-Y. Koo. Wavelet Deconvolution. *IEEE Trans. Inform. Theory*, 48:734–747, 2002.

[14] G.C. Kyriazis. Decomposition Systems of Function Spaces. *Studia Math.*, 157(2):133–169, 2003.

[15] M. Frazier, B. Jawerth. Decomposition of Besov Spaces. *Indiana Univ. Math. J.*, 34:777–799, 1985.

[16] M. Frazier, B. Jawerth. A Discrete Transform and Decompositions of Distribution Spaces. *J. Funct. Anal.*, 93:34–170, 1990.

[17] M. Frazier, B. Jawerth, G.L. Weiss. *Littlewood-Paley Theory and the Study of Function Spaces*. CBMS Regional Conference Series, 1991.

[18] S. Mallat. Multiresolution Approximations and Wavelet Orthonormal Bases in $L^2(\mathbb{R})$. *Trans. Amer. Math. Soc.*, 315:69–87, September 1989.

[19] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[20] Y. Meyer. *Wavelets and Operators*. Cambridge University Press, Cambridge, England, 1992.

[21] N.-Y. Lee and B.J. Lucier. Wavelet Methods For Inverting the Radon Transform with Noisy Data. *IEEE Trans. Image Process.*, 10(1):79–94, 2001.

[22] P. Auscher. Solution of Two Problems on Wavelets. *J. Geom. Anal*, 5(2):181–236, 1995.

[23] J. Peetre. *New Thoughts on Besov Spaces*. Duke Univ. Math. Series, 1976.

[24] R. DeVore, B. Jawerth, B. Lucier. Image Compression through Wavelet Transform Coding. *IEEE Trans. Inform. Theory*, 38:719–746, 1992.

[25] R. DeVore, G. Lorentz. *Constructive Approximation*. Springer-Verlag, 1993.

[26] O. Rioul. Simple Regularity Criteria for Subdivision Schemes. *SIAM J. Math. Anal.*, 23(12), November 1992.

[27] O. Rioul. A Discrete-Time Multiresolution Theory. *IEEE Trans. Signal Process.*, 41(8), August 1993.

[28] O. Rioul. Regular Wavelets: A Discrete-Time Approach. *IEEE Trans. Signal Process.*, 41(12), December 1993.

[29] Q. Sun. Sequence Spaces and Stability of Integer Translates. *Z. Anal. Anwendungen*, 12:567–584, 1993.

[30] R. Torres. Mean Oscillation of Functions and the Paley-Wiener Space. *J. Fourier Anal. and Appl.*, 4:283–297, 1998.

[31] R. H. Torres. Spaces of sequences, sampling theorem and functions of exponential type. *Studia Math.*, 100:51–74, 1991.

[32] H. Triebel. *Theory of Function Spaces.* Birkhäuser Basel, 1983.

[33] H. Triebel. *Theory of Function Spaces II.* Birkhäuser Basel, 1992.

[34] P. Wojtaszczyk. *A Mathematical Introduction to Wavelets.* Cambridge University Press, 1997.