



INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus

Mathématique


Stéphane Jaffard and Hamid Krim

Regularity properties of Haar Frames

Volume 359, issue 9 (2021), p. 1107-1117

<<https://doi.org/10.5802/crmath.228>>

© Académie des sciences, Paris and the authors, 2021.
Some rights reserved.

 This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



Les Comptes Rendus. Mathématique sont membres du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org



Harmonic analysis / *Analyse harmonique*

Regularity properties of Haar Frames

Stéphane Jaffard^{*, a} and Hamid Krim^b

^a Univ Paris Est Creteil, CNRS, LAMA, F-94010 Creteil, France,
Univ Gustave Eiffel, LAMA, F-77447 Marne-la-Vallée, France.

^b Department of Electrical and Computer Engineering, North Carolina State
University, Raleigh, NC 27695 USA.

E-mails: jaffard@u-pec.fr (S. Jaffard), ahk@ncsu.edu (H. Krim)

Abstract. We prove that pointwise and global Hölder regularity can be characterized using the coefficients on the *Haar tight frame* obtained by using a finite union of shifted Haar bases, despite the fact that the elements composing the frame are discontinuous.

2020 Mathematics Subject Classification. 42B35, 42C40, 46E35, 65T60, 68T05.

Funding. The first author is supported by ANR Grant Multifrac ANR-16-CE33-0020.

Manuscript received 22nd December 2020, revised 31st May 2021, accepted 5th July 2021.

1. Introduction

If H is a separable Hilbert space, a frame is a sequence (e_n) of elements of H satisfying

$$\exists C, C' > 0, \quad \forall f \in H, \quad C \|f\|^2 \leq \sum_n |\langle f | e_n \rangle|^2 \leq C' \|f\|^2. \quad (1)$$

This condition implies that the vectors (e_n) span E , and any element $f \in H$ can be reconstructed from the e_n in a stable way: There exists a *dual frame* (g_n) such that the partial sums $\sum_{n \leq N} \langle f | g_n \rangle e_n$ converge to f in H . Wavelet frames were introduced by I. Daubechies, A. Grossmann and Y. Meyer in the seminal article [7], as a compromise between the continuous wavelet transform, which is a very flexible analysis tool (the restrictions on the shape of the wavelet are easy to meet), and orthonormal wavelet bases (the use of which being much less greedy in terms of computational costs).

In signal and image processing, frame decompositions are often a preferred and adapted tool to zoom in and unveil useful features/properties which are critical to not only uncovering algorithmic solutions to many data-driven problems, but to also simplifying the ensuing computational challenges. Properties of frame expansions have been extensively investigated, see e.g. [5, 11] and references therein for an account of the developments of frame theory, and its relevance in mathematical analysis, statistics, and in signal and image processing. Convolutional Neural Networks (CNN), see e.g. [17], broadly referred to as Deep Learning (DL) as a tool, is a

* Corresponding author.

more refined and better performing Neural Network, consisting of layers of banks of linear filters, whose respective outputs are non-linearly transformed. An alternative interpretation of CNN as a Haar wavelet multi-scale frame optimization of data (referred to as Scattering Networks), has been provided, see [4]. The computational efficiency of this systematic and optimal representation selected from a wavelet frame representation was shown to provide a near-translation-invariance and a viable inference framework. The notion of CNN was extended to data-graph structures [3]. A *graph convolution* may be defined on such a space, which yields a formalism for Graph-based DL. Similarly, a computationally more efficient Graph-based DL using Haar Convolution was constructed in [20], where a Haar decomposition is performed at a much lower computational cost. However, the main drawback of Haar decompositions is the irregularity of the basis elements, thus their inability to account for the smoothness of the data. The new regularity results for Haar systems obtained in this paper therefore present a new potential for further refining Graph DL and for additionally providing a new perspective on computational impacts, such as translation invariance which is critical to data-inherent transient changes, see [16].

Acknowledgments

The authors are grateful to Albert Cohen for pointing out reference [9], and to the referees for several suggestions of improvements.

2. Haar systems

2.1. The Haar basis

Let φ be the characteristic function of $[0, 1]$, which, more precisely, we define as

$$\begin{cases} \varphi(x) = 1 & \text{if } x \in (0, 1), \\ \varphi(0) = \varphi(1) = 1/2, \\ \varphi(x) = 0 & \text{otherwise;} \end{cases} \quad (2)$$

and let

$$\forall x \in \mathbb{R}, \quad \psi(x) = \varphi(2x) - \varphi(2x - 1). \quad (3)$$

This (slightly unusual) definition for φ (and hence ψ) at the end points of its support is motivated by the fact that we will consider pointwise values of partial sums of Haar series, and it is important that every point be a Lebesgue point of these partial sums. The Haar basis on \mathbb{R} is the orthonormal basis of $L^2(\mathbb{R})$ composed of the functions

$$\begin{cases} \varphi(x - k) & \text{for } k \in \mathbb{Z} \\ 2^{j/2} \psi(2^j x - k) & \text{for } j \geq 0 \text{ and } k \in \mathbb{Z}. \end{cases} \quad (4)$$

This system (or, more precisely, its restriction on $[0, 1]$) was introduced by A. Haar in his PhD thesis in 1909. He proved that, in contradistinction to what happens for Fourier series, if f is any continuous function, then the partial sums of the Haar expansion of f converge uniformly. However, since the Haar system is composed of discontinuous functions, it obviously cannot be a basis of function spaces of continuous functions. Let us be more precise.

Definition 1. *Let E be a separable Banach space. A sequence $(e_n)_{n \in \mathbb{N}}$ of elements of E is an unconditional basis of E if it satisfies the following requirements:*

- $\forall f \in E$, there exists a unique sequence of real numbers $(a_n)_{n \in \mathbb{N}}$ such that the partial sums $\sum_{n \leq N} a_n e_n$ converge to f , i.e.

$$\left\| \sum_{n=1}^N a_n e_n - f \right\|_E \longrightarrow 0 \quad \text{when } N \rightarrow +\infty. \quad (5)$$

- There exists $C > 0$ such that, for any sequence of real numbers $(a_n)_{n \in \mathbb{N}}$, for any sequence (ε_n) such that $|\varepsilon_n| \leq 1$, then

$$\left\| \sum \varepsilon_n a_n e_n \right\|_E \leq C \left\| \sum a_n e_n \right\|_E. \quad (6)$$

The second requirement ensures the stability of the reconstruction of f . In statistics, this key property is referred to as the *multiplier property*. Note that it is only one of the two ingredients for an unconditional basis, and it may be verified in spaces that are not separable (and thus cannot have unconditional bases). It is e.g. the case for the Hölder C^α spaces if a smooth orthonormal wavelet basis is used, i.e. a basis of the form (4) where φ and ψ are sufficiently smooth and well localized, see (10) below.

We now turn to the Haar basis. The problem of determining which function spaces the Haar system is an unconditional basis for, has been settled by G. Bourdaud in [2] for Besov spaces. Let us recall their definition. We use a smooth wavelet basis, i.e. an orthonormal basis of $L^2(\mathbb{R})$ which has the same algorithmic structure as the Haar system (4), but where φ and ψ belong to the Schwartz class, see [19]). Thus

$$\forall f \in L^2, \quad f(x) = \sum_{k \in \mathbb{Z}} c_k \varphi(x-k) + \sum_{j \geq 0} \sum_{k \in \mathbb{Z}} c_{j,k} \psi(2^j x - k), \quad (7)$$

where the wavelet coefficients of f are

$$c_{j,k} = 2^j \int_{\mathbb{R}} f(x) \psi(2^j x - k) dx \quad \text{and} \quad c_k = \int_{\mathbb{R}} f(x) \varphi(x - k) dx. \quad (8)$$

We will use the fact that convergence also holds pointwise: If $f \in L^1(\mathbb{R})$, then the partial sums of f converge almost everywhere, and in particular at Lebesgue points of f , see [24, 25].

Definition 2. Let $s \in \mathbb{R}$ and $p, q \in (0, \infty]$. A tempered distribution f belongs to the Besov space $B_p^{s,q}(\mathbb{R})$ if and only if its wavelet coefficients on a wavelet basis in the Schwartz class satisfy the two conditions: $(c_k) \in \ell^p$ and

$$\sum_{j \in \mathbb{Z}} \left(\sum_{k \in \mathbb{Z}} \left(2^{(s-1/p)j} |c_{j,k}| \right)^p \right)^{q/p} \leq C, \quad (9)$$

using the usual convention for ℓ^∞ when p or q is infinite.

In particular, the global Hölder spaces $C^\alpha(\mathbb{R}) = B_\infty^{\alpha,\infty}(\mathbb{R})$ ($\alpha \in \mathbb{R}$), sometimes referred to as *Lipschitz spaces*, are characterized by the condition

$$(c_k) \in \ell^\infty \quad \text{and} \quad \exists C, \forall j, k \quad |c_{j,k}| \leq C 2^{-\alpha j}. \quad (10)$$

The characterization supplied by (9) can roughly be interpreted as follows: The fractional derivatives of f of order s belong to L^p , see [21]. The theorem of Bourdaud states (in particular) that the Haar system is an unconditional basis of $B_p^{s,q}(\mathbb{R})$ (and (9) holds) if and only if

$$\frac{1}{p} - 1 < s < \min\left(\frac{1}{p}, 1\right).$$

This result is sharp, indeed, for larger values of s , the Haar function ψ no longer belongs to the corresponding space. Note that this trivial obstruction also prevents (9) to yield a wavelet characterization of Besov spaces in that case; indeed the Haar function (3) only has one non-vanishing coefficient on the Haar basis, and therefore its coefficients obviously satisfy (9). Thus, if this characterization held for the Haar basis, it would follow that the Haar function belongs to the corresponding Besov space, which is not the case.

Despite the limitation due to its irregularity, the Haar system has been of constant use in signal and image processing, and more recently, became a key tool in data-graph structures. A purpose of this paper is to show that this limitation can be mitigated by using a frame instead of an orthonormal basis, thus taking advantage of redundancy. In order to more precisely address

this question, we note that the information within the definition of an unconditional basis may be granularized into two different points:

- The *analysis* problem: Is it possible to characterize the fact that a function belongs to a function space by a condition on the moduli of its coefficients on the analyzing system, as given, e.g. by (9) in the case of Besov spaces?
- The *synthesis* problem: When does the partial sum reconstruction formula (5) (or (7) in the case of wavelet bases) holds in the corresponding function space?

It is clear that the second point cannot be improved upon by using a redundant system; indeed, as soon as the “building blocks” do not belong to the function space, the norms in (5) are not even defined. However, there may be some room for improvement concerning the first point. More information would be unveiled if (9) holds for the coefficient on a redundant system, and one might expect that it can be converted into some regularity information on the function allowing to go “beyond” the limitation of Bourdaud’s theorem. Following this motivation, we now introduce Haar frames.

2.2. Haar frame

The Haar frame we will consider is the union of the orthonormal Haar basis (7) together with the two orthogonal bases obtained by shifting the elements of the Haar basis by $1/3$ and $2/3$. This analysing system is composed of the

$$I_k(x) = \varphi\left(x - \frac{k}{3}\right), \text{ and } H_{j,k}(x) = \psi\left(2^j x - \frac{k}{3}\right), \quad j, k \in \mathbb{Z}. \quad (11)$$

The I_k together with the $2^{j/2}H_{j,k}$ form a union of three orthonormal bases; therefore, they constitute a *tight frame*, i.e. a frame for which the inequalities in (1) are equalities; more precisely,

$$\forall f \in L^2(\mathbb{R}), \quad \|f\|_{L^2} = \frac{1}{3} \left(\sum_k |\langle f | I_k \rangle|^2 + \sum_{j,k} |\langle f | 2^{j/2} H_{j,k} \rangle|^2 \right). \quad (12)$$

One well documented drawback of using an orthonormal wavelet basis in signal an image processing is that it does not supply a translation invariant representation, but is dependent on the particular discrete dyadic grid which is chosen. This drawback can be mitigated by oversampling this dyadic grid, thus replacing the initial orthonormal basis by a finite union of orthonormal bases, as proposed in [22, 23]. The choice supplied by (11) corresponds to an oversampling by a factor 3 (we will see in Section 3 how the results we will obtain in this setting extend to other choices of oversampling). Applying a wavelet representation in many practical scenarios such as signal detection in a Radar scenario, or reconstruction of a function/signal in noise (i.e. denoising) as well as parameter estimation in communication, is highly dependent on the translation-invariance of the transformation. As an example, a continuous wavelet transform or a wavelet frame representation guarantees that a time delay estimation or a detection of a very short transient will be successfully achieved, whereas an orthogonal wavelet representation may turn out to be unable to detect a short transient or to estimate translation-sensitive parameters. This exploitation of a redundant wavelet representation was later illustrated in [6] as what may be viewed as an “averaging procedure” for improved denoising, and referred to as cycle-spinning. We will show that this oversampling procedure also allows to turn the inherent limitations of the Haar basis for regularity analysis.

2.3. Uniform regularity

The first problem we consider is the characterization of the uniform Hölder spaces $C^\alpha(\mathbb{R})$; recall that they coincide with the spaces $B_\infty^{\alpha,\infty}(\mathbb{R})$, see [21]; if $0 < \alpha < 1$, an equivalent definition is

$$f \in L^\infty(\mathbb{R}) \quad \text{and} \quad \exists C, \forall x, y \in \mathbb{R}, \quad |f(x) - f(y)| \leq C|x - y|^\alpha.$$

Though $C^\alpha(\mathbb{R})$ cannot be characterized by a condition bearing on the Haar basis coefficients, nonetheless, Theorem 4 below shows that such a characterization is possible using the Haar frame (11). We will make the following minimal regularity assumptions on the functions of interest.

Definition 3. *Let f be a locally bounded function; f is Lebesgue-regular if every point is a Lebesgue point of f , i.e.*

$$\forall x \in \mathbb{R}, \quad f(x) = \lim_{r \rightarrow 0} \frac{1}{2r} \int_{x-r}^{x+r} f(t) dt.$$

Note that this definition implicitly makes the assumption that functions are defined “point to point” and not “except for a set of vanishing measure”. Continuous functions are of course Lebesgue regular; but this class also allows for discontinuities. For instance assume that, at every point x , f has a right and a left limit at x , and that, at every discontinuity point x_0 , f satisfies

$$f(x_0) = \frac{1}{2} \left(\lim_{x \rightarrow x_0^+} f(x) + \lim_{x \rightarrow x_0^-} f(x) \right);$$

then f is clearly Lebesgue regular. A key property that we will use is that the wavelet series of a Lebesgue-regular function f converges everywhere to f . The following result shows that the Haar frame characterization of the global C^α regularity is similar to that resulting from a smooth wavelet basis.

Theorem 4. *Let $\alpha \in (0, 1)$. Let f be a Lebesgue regular function; $f \in C^\alpha(\mathbb{R})$ if and only if its Haar frame coefficients*

$$C_k = \int f(x)I_k(x)dx \quad \text{and} \quad c_{j,k} = 2^j \int f(x)H_{j,k}(x)dx$$

satisfy

$$\exists C, \forall k, \quad |C_k| \leq C \quad \text{and} \quad \forall j, k, \quad |c_{j,k}| \leq C2^{-\alpha j}. \tag{13}$$

Proof. Assume that $f \in C^\alpha(\mathbb{R})$. Then

$$|C_k| = \left| \int f(x)I_k(x)dx \right| \leq \|I_k\|_{L^1} \|f\|_{L^\infty} \leq \|f\|_{L^\infty},$$

hence the first statement in (13) holds. Let

$$I_{j,k} = \left[\frac{k}{3 \cdot 2^j}, \frac{k}{3 \cdot 2^j} + \frac{1}{2^j} \right] \tag{14}$$

be the support of $H_{j,k}$. we denote by $I_{j,k}^-$ the left half of $I_{j,k}$ and by $I_{j,k}^+$ the right half of $I_{j,k}$. Then

$$c_{j,k} = 2^j \int_{I_{j,k}} f(x)H_{j,k}dx = 2^j \int_{I_{j,k}^-} f(x)dx - 2^j \int_{I_{j,k}^+} f(x)dx = 2^j \int_{I_{j,k}^-} (f(x) - f(x + 2^{-j}))dx,$$

so that

$$|c_{j,k}| \leq 2^j \int_{I_{j,k}^-} |f(x) - f(x + 2^{-j})| dx \leq C2^{-\alpha j},$$

hence the second statement in (13) holds.

Conversely, assume that (13) holds. Since f is Lebesgue regular, we can use the reconstruction formula for the Haar orthonormal wavelet basis only (which converges everywhere towards the pointwise value of f); thus

$$\begin{aligned} \forall x \in \mathbb{R}, \quad |f(x)| &= \sum_{k \in \mathbb{Z}} |c_k| |\varphi(x - k)| + \sum_{j \geq 0} \sum_{k \in \mathbb{Z}} |c_{j,k}| \left| \psi(2^j x - k) \right|, \\ &\leq C + C \sum_{j=0}^{\infty} 2^{-\alpha j} \end{aligned}$$

so that $f \in L^\infty$. Let us now estimate increments of f . We have three possible reconstruction formulas for f using any of the three orthonormal bases composing the tight frame; the idea of the proof is to use this extra flexibility. Let $x \neq y$ be given. Define J by

$$\frac{1}{4} 2^{-J} \leq |x - y| < \frac{1}{2} \cdot 2^{-J}.$$

Consider now the intervals $I_{j,k}$. Since these intervals are of length 2^{-j} and are deduced from each other by a shift of $\frac{1}{3} \cdot 2^{-j}$, at least one of them contains both points x and y ; we denote it by I_{j,k_j} . We now use either the Haar basis or one of its two “sisters” shifted by $1/3$, the choice being driven by the fact that the interval I_{j,k_j} that we picked, is the support of an element $H_{j,k}$ of the chosen basis. This implies that, for all generations $j < J$, either the support of an $H_{j,k}$ of this basis does not contain x and y , or x and y are in the same “half” of the support of $H_{j,k}$. Let us use the reconstruction formula using this orthonormal basis (and let us denote its elements by φ_k and $\psi_{j,k}$); since it converges everywhere towards the pointwise value of f , we get

$$\begin{aligned} f(x) - f(y) &= \sum_k C_k (\varphi_k(x) - \varphi_k(y)) + \sum_{j \leq J} \sum_k c_{j,k} (\psi_{j,k}(x) - \psi_{j,k}(y)) + \dots + \sum_{j > J} \sum_k c_{j,k} (\psi_{j,k}(x) - \psi_{j,k}(y)). \end{aligned}$$

Because of our choice of the basis, it follows that $\forall k, \varphi_k(x) = \varphi_k(y)$ and $\forall j \leq J, \forall k, \psi_{j,k}(x) = \psi_{j,k}(y)$; therefore

$$|f(x) - f(y)| \leq \sum_{j > J} \sum_k |c_{j,k}| |\psi_{j,k}(x) - \psi_{j,k}(y)|.$$

At each generation j , at most two terms bring a contribution; using (13), we get

$$|f(x) - f(y)| \leq \sum_{j > J} C 2^{-\alpha j} \leq C 2^{-\alpha J} \leq C |x - y|^\alpha. \quad \square$$

Remark 5. As pointed out to us by Albert Cohen, the proof of Theorem 4 relies on an argument which is similar to the *mixing lemma* ([9, Lemma 2.3 in Chap. 12], originally in [10]) used to determine the rate of approximation of smooth functions by sequences of splines on the interval $[0, 1]$ having knots at $T_n = \{\frac{k}{n}\}_{k=0, \dots, n}$. The fact that this family is not nested allows for a better rate than in the nested case (e.g. for dyadic subdivisions). Note however that the (piecewise constant) splines used in the proof of Theorem 4 yield much more economical system for a similar conclusion.

2.4. Pointwise regularity: The Haar basis

We now consider the problem of characterizing pointwise regularity by estimates on the Haar frame coefficients.

Definition 6. Let $x_0 \in \mathbb{R}$ and $\alpha \geq 0$. A locally bounded function $f : \mathbb{R} \rightarrow \mathbb{R}$ belongs to $C^\alpha(x_0)$ if there exist $C > 0$ and a polynomial P_{x_0} with $\deg(P_{x_0}) < \alpha$ such that, in a neighborhood of x_0 ,

$$|f(x) - P_{x_0}(x)| \leq C |x - x_0|^\alpha. \tag{15}$$

The pointwise Hölder exponent of f at x_0 is $h_f(x_0) = \sup\{\alpha : f \in C^\alpha(x_0)\}$.

The polynomial P is unique; it is called the *Taylor polynomial* of f at x_0 . When using smooth wavelets, criteria based on the wavelet coefficients on an orthonormal wavelet basis allow to recover pointwise regularity; let us recall some notations. A dyadic interval is of the form

$$\lambda (= \lambda_{j,k}) = \left[\frac{k}{2^j}, \frac{k+1}{2^j} \right).$$

Let λ be a dyadic interval; 3λ denotes the interval of same center as λ and three times wider. Wavelets coefficients can therefore be indexed by dyadic intervals: We will write $c_\lambda := c_{j,k}$.

Definition 7. Let f be a locally bounded function, and let the (φ_k) and $(\psi_{j,k})$ generate a smooth wavelet basis. The wavelet leaders of f are the quantities

$$d_\lambda = \sup_{\lambda' \subset 3\lambda} |c_{\lambda'}|.$$

Wavelet leaders allow to estimate pointwise Hölder exponents, see [12] for the initial 2-microlocal wavelet criterium and [13] for its reformulation in terms of wavelet leaders, which we now recall. We denote by $\lambda_j(x_0)$ the dyadic interval of width 2^{-j} which contains x_0 .

Theorem 8. Let $f \in C^\varepsilon(\mathbb{R})$ for an $\varepsilon > 0$. If the generating wavelets φ and ψ belong to $C^N(\mathbb{R})$ for an $N > \alpha$, then

$$\forall x_0 \in \mathbb{R} : h_f(x_0) = \liminf_{j \rightarrow +\infty} \frac{\log(d_{\lambda_j(x_0)})}{\log(2^{-j})}.$$

The same question is more difficult to answer when using an irregular wavelet basis; to our knowledge, it was only tackled in [14]. Let us first recall the results concerning pointwise regularity results obtained there for the Haar basis. Recall that a *dyadic rational* is a point of the form $k/2^j$, for $j, k \in \mathbb{Z}$.

Definition 9. Let $x \in \mathbb{R}$. The rate of approximation of x by dyadic rationals is

$$r(x) = \limsup_{j \rightarrow +\infty} \frac{\log(\text{dist}(x, 2^{-j}\mathbb{Z}))}{\log(2^{-j})}. \quad (16)$$

For any x , $r(x) \geq 1$, and almost every x satisfies $r(x) = 1$. The following result is proved in [14].

Proposition 10. Let f be a locally bounded function, and $x_0 \in \mathbb{R}$. If $f \in C^\alpha(x_0)$ for $\alpha < 1$, then its wavelet leaders on the Haar basis satisfy

$$\left| d_{\lambda_j(x_0)} \right| \leq C 2^{-\alpha j}. \quad (17)$$

Conversely, if (17) holds and if the Haar basis coefficients $c_{j,k}$ satisfy the uniform decay assumption

$$\exists \varepsilon, C > 0 : \quad |c_{j,k}| \leq C 2^{-\varepsilon j}, \quad (18)$$

then

$$h_f(x_0) \geq \frac{\alpha}{r(x_0)}.$$

This result yields the best possible pointwise regularity that can be inferred from the knowledge of the size of the Haar basis coefficients. Almost every x_0 satisfies $r(x_0) = 1$, so that, if (18) holds, then Proposition 10 yields the exact pointwise regularity at these points. Besides dyadic points (where it is clear that decay estimates on the Haar coefficients cannot allow to estimate pointwise regularity), the estimate of $h_f(x_0)$ yielded by Proposition 10 deteriorates if $r(x_0)$ is large, which means that there exists a sequence of scales such that x_0 is “very close” to dyadic points at these scales.

2.5. Pointwise regularity: The Haar frame

We will now show that, in contradistinction with Proposition 10, the Haar frame coefficients allow to recover the pointwise Hölder exponent everywhere. Let us first introduce some notations. We index the elements of the Haar frame $H_{j,k}$ by their support $\lambda = I_{j,k}$, see (14) (note that λ is not necessarily a dyadic interval). The corresponding Haar coefficient is

$$c_\lambda = 2^j \int f(x) H_{j,k}(x) dx.$$

Haar leaders are indexed by dyadic intervals and defined by

$$d_\lambda = \sup_{\lambda' \subset 3\lambda} |c_{\lambda'}|.$$

Theorem 11. *Let f be a locally bounded function, and let $\alpha > 0$. If $f \in C^\alpha(x_0)$, and if the Taylor polynomial of f at x_0 is constant, then its wavelet leaders on the Haar frame satisfy*

$$d_{\lambda_j(x_0)} \leq C 2^{-\alpha j}. \quad (19)$$

Conversely, if (19) holds and if the coefficients $c_{j,k}$ satisfy the uniform decay assumption

$$\exists \varepsilon, C > 0: \quad |c_{j,k}| \leq C 2^{-\varepsilon j}, \quad (20)$$

then

$$\exists C: \quad \text{if } |x - x_0| \leq 1/2, \quad |f(x) - f(x_0)| \leq C |x - x_0|^\alpha \log(|x - x_0|). \quad (21)$$

Remark 12. The restriction on the Taylor polynomial is automatically satisfied if $0 < \alpha < 1$. It is easy to check that it is also satisfied by a class of functions which plays an important role in multifractal analysis: The distribution functions of singular measures (with no restriction on α).

The first statement of the theorem is a classical result, and we just recall its proof for completeness, see e.g. [14], or, more recently [15] where it is used in the context of a continuous wavelet transform using the Haar basis: Its purpose was to obtain a pointwise irregularity criterium for the Brjuno function. The Haar wavelet is a natural choice in this case because integrals of the Brjuno function on certain intervals with rational ends have an (almost) explicit form, so that, for an appropriate positioning of the support of the Haar wavelet, the orders of magnitude of the Haar coefficients are known.

Proof. Assume that $f \in C^\alpha(x_0)$ for $\alpha < 1$. Let $j \geq 0$ be given, $j' \geq j$ and $\lambda' \subset 3\lambda_j(x_0)$; since $H_{j',k'}$ has a vanishing integral,

$$c_{j',k'} = 2^{j'} \int_{I_{j',k'}} f(x) H_{j',k'} dx = 2^{j'} \int_{I_{j',k'}} (f(x) - f(x_0)) H_{j',k'} dx,$$

so that

$$|c_{j',k'}| \leq 2^{j'} \int_{I_{j',k'}} |f(x) - f(x_0)| dx \leq C 2^{j'} \int_{I_{j',k'}} |x - x_0|^\alpha dx.$$

Since x and x_0 belong to the support of $I_{j',k'}$, hence to $3\lambda_j(x_0)$, $|x - x_0| \leq 2^{-j}$ and it follows that $|c_{j',k'}| \leq C 2^{-\alpha j}$; and therefore (19) holds.

Suppose now that (19) holds. Let $x \neq x_0$ be given. Since f is Lebesgue regular, the wavelet series of f converges to the pointwise value of f at x and at x_0 . Define j by

$$\frac{1}{4} 2^{-j} \leq |x - x_0| < \frac{1}{2} \cdot 2^{-j}. \quad (22)$$

At least one of the three Haar bases is such that, at the generation j , x and x_0 belong to the same interval I_{j,k_j} . As in the proof of uniform regularity, in order to estimate increments of f we use this Haar basis for the reconstruction formula:

$$f(x) - f(x_0) = \sum_k C_k (\varphi_k(x) - \varphi_k(x_0)) + \sum_{j' \leq j} \sum_{k'} c_{j',k'} (\psi_{j',k'}(x) - \psi_{j',k'}(x_0)) \\ + \cdots + \sum_{j' > j} \sum_{k'} c_{j',k'} (\psi_{j',k'}(x) - \psi_{j',k'}(x_0)).$$

The terms for $j' < j$ vanish because x and x_0 belong to the same (right or left) half of the support of $\psi_{j',k'}$ (or of φ_k). As regards the terms for $j' \geq j$, we first assume that $j' \leq [Aj]$, where A is a (large) constant, that will be fixed later. Each sum

$$\sum_{k'} c_{j',k'} (\psi_{j',k'}(x) - \psi_{j',k'}(x_0))$$

contains at most two nonvanishing terms: the ones such that x or x_0 belong to $I_{j',k'}$; but, in that case (19) implies that $|c_{j',k'}| \leq C2^{-\alpha j}$. Therefore

$$\left| \sum_{k'} c_{j',k'} (\psi_{j',k'}(x) - \psi_{j',k'}(x_0)) \right| \leq C2^{-\alpha j},$$

and

$$\sum_{j'=j}^{[Aj]} \left| \sum_{k'} c_{j',k'} (\psi_{j',k'}(x) - \psi_{j',k'}(x_0)) \right| \leq Cj2^{-\alpha j}. \quad (23)$$

Assume now that $j' > j$. Because of the localization of the Haar basis, (20) implies that

$$\left| \sum_{k'} c_{j',k'} (\psi_{j',k'}(x) - \psi_{j',k'}(x_0)) \right| \leq C2^{-\varepsilon j},$$

and

$$\sum_{j' > [Aj]} \left| \sum_{k'} c_{j',k'} (\psi_{j',k'}(x) - \psi_{j',k'}(x_0)) \right| \leq C2^{-\varepsilon Aj}. \quad (24)$$

We pick A such that $\varepsilon A = \alpha$; (22) implies that $j \leq C|\log(|x - x_0|)|$, so that (21) follows from (23) and (24). \square

3. Concluding remarks

Our choice of the regular oversampling supplied by the union of three bases shifted by $1/3$ was motivated by two reasons: Those already mentioned and developed in [6, 22, 23], i.e. using a wavelet system closer to translation invariance, but which keeps the numerical efficiency of orthonormal wavelet bases; and, on other hand, the fact that, if one wants to keep a regular sampling, then an oversampling by 3 is the smallest number for which the results stated in Theorems 4 and 11 hold.

However, variants are possible. Indeed the key argument in the proofs of Theorems 4 and 11 is that, if x and y are such that $|x - y| \sim 2^{-j}$, then one can find a translated dyadic interval of length close to 2^{-j} which contains both x and y . As pointed out by one of the referees, this is clearly possible using only the Haar basis and only one translate by a rational $r = p/(2k + 1)$. Indeed, if I is an interval of length l satisfying $2^{-j} \leq l < 2 \cdot 2^{-j}$; let m be defined by

$$\frac{p}{8(2k + 1)} \leq 2^{-m} < \frac{p}{4(2k + 1)}.$$

Then I clearly is included either in a dyadic interval of length 2^{-j+l} or in a interval of the same length obtained as a shift by r of a dyadic interval. Therefore, the proofs of Theorems 4 and 11 work in the same way for these two systems. And of course, the same conclusion also holds if we include additional translates by $r = q/(2k+1)$ for several values of q (and in particular all of them, if one is concerned by the requirement of a regular oversampling).

Similarly, Theorems 4 and 8 easily extend to the several variable version of the Haar system, which is obtained by the usual tensor product construction: For $x = (x_1, \dots, x_d)$, we define

$$\Phi(x) = \varphi(x_1) \cdots \varphi(x_d),$$

and

$$\Psi^{(i)}(x) = \psi_1(x_1) \cdots \psi_d(x_d),$$

where the ψ_l are either the one-variable functions φ or ψ , the choice $\varphi(x_1) \cdots \varphi(x_d)$ being excluded (so that there are $2^d - 1$ functions $\Psi^{(i)}$). Then the d -dimensional Haar basis is composed of the $\Phi(x - k)$ for $k \in \mathbb{Z}^d$ and the $2^{dj/2} \Psi^{(i)}(2^j x - k)$ for $j \geq 0$ and $k \in \mathbb{Z}^d$. The corresponding Haar frame is obtained by shifting this basis by the vectors $\sum \varepsilon_i e_i / 3$ where $\varepsilon_i \in \{0, 1, 2\}$ and the e_i are the elements of the canonical basis of \mathbb{R}^d . This yields 3^d bases, and the d -dimensional Haar frame is composed of the union of these bases. Theorems 4 and 11 also extend without difficulty to the d -dimensional setting. The key point is to notice that, here again, if $|x - y| \sim 2^{-j}$, then one can find a translated dyadic cube of width close to 2^{-j} which contains both x and y , which is clear; indeed it suffices to use in each direction of the canonical basis, the translation supplied by the one-dimensional case for the corresponding coordinate of the segment $[x, y]$.

Finally, note that Theorems 4 and 8 also extend to piecewise smooth wavelets, such as the “spline wavelets” constructed by G. Battle and P.-G. Lemarié, see [1, 18]; in that case, the same proofs allow to characterize global and pointwise regularity up to an order α given by the number of vanishing moments of the wavelet, which is larger than its uniform regularity (and was the natural bound for previous regularity results). This is possible because these wavelets are piecewise polynomials between integers. In contrast, an interesting open problem would be to determine if these results could be extended to e.g. Daubechies wavelets, the singularities of which are not located at integers, but on fractal sets, see [8].

References

- [1] G. Battle, “A block spin construction of ondelettes. Part II: The QFT connection”, *Commun. Math. Phys.* **114** (1988), p. 93-102.
- [2] G. Boudaud, “Ondelettes et espaces de Besov”, *Rev. Mat. Iberoam.* **11** (1995), no. 3, p. 477-512.
- [3] J. Bruna, W. Zaremba, A. Szlam, Y. Lecun, “Spectral networks and locally connected networks on graphs”, in *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings* (Y. Bengio, Y. LeCun, eds.), vol. 27, ICLR, 2016.
- [4] X. Cheng, X. Chen, S. Mallat, “Deep Haar scattering networks”, *Inf. Inference* **5** (2016), p. 105-133.
- [5] O. Christensen, *An introduction to Frames and Riesz Bases*, Applied and Numerical Harmonic Analysis, Birkhäuser, 2003.
- [6] R. R. Coifman, D. L. Donoho, “Translation-Invariant de-noising”, in *Wavelets and Statistics. Proceedings of the 15th French-Belgian meeting of statisticians, held at Villard de Lans, France, November 16-18, 1994* (O. G. Antoniadis Anestis et al., eds.), Lecture Notes in Statistics, vol. 103, Springer, 1995, p. 121-150.
- [7] I. Daubechies, A. Grossmann, Y. Meyer, “Painless nonorthogonal expansions”, *J. Math. Phys.* **27** (1986), no. 5, p. 1271-1283.
- [8] I. Daubechies, J. C. Lagarias, “On the Thermodynamic Formalism for Multifractal Functions”, *Rev. Math. Phys.* **6** (1994), p. 1033-1070.
- [9] R. A. DeVore, G. B. Lorentz, *Constructive approximation*, Grundlehren der Mathematischen Wissenschaften, vol. 303, Springer, 1993.

- [10] R. A. DeVore, F. Richards, “Saturation and inverse theorems for spline approximation”, in *Spline functions and approximation theory (Proc. Sympos., Univ. Alberta, Edmonton, Alta., 1972)* (A. Meir, A. Sharma, eds.), International Series of Numerical Mathematics, vol. 21, Birkhäuser, 1973, p. 73-82.
- [11] C. Heil, P. E. T. Jorgensen, D. R. Larson, *Wavelets, Frames, and Operator Theory. Papers from the Focused Research Group Workshop, University of Maryland, College Park, MD, USA, January 15–21, 2003*, Contemporary Mathematics, vol. 345, American Mathematical Society, 2004.
- [12] S. Jaffard, “Exposants de Hölder en des points donnés et coefficients d’ondelettes”, *C. R. Math. Acad. Sci. Paris* **308** (1989), no. 4, p. 79-81.
- [13] ———, “Wavelet techniques in multifractal analysis”, in *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot. Multifractals, probability and statistical mechanics, applications* (M. Lapidus, M. van Frankenhuijsen, eds.), Proceedings of Symposia in Pure Mathematics, vol. 72(2), American Mathematical Society, 2004, in part the proceedings of a special session held during the annual meeting of the American Mathematical Society, San Diego, CA, USA, January 2002, p. 91-152.
- [14] S. Jaffard, B. B. Mandelbrot, “Local regularity of nonsmooth wavelet expansions and application to Polya’s function”, *Adv. Math.* **120** (1996), no. 2, p. 265-282.
- [15] S. Jaffard, B. Martin, “Multifractal analysis of the Brjuno function”, *Invent. Math.* **212** (2018), no. 1, p. 109-132.
- [16] H. Krim, S. Jaffard, S. Roheda, S. Mahdizadehaghdam, A. Panahi, “On Stabilizing Generative Adversarial Networks (STGANs)”, in preparation, 2021.
- [17] Y. LeCun, Y. Bengio, G. Hinton, “Deep learning”, *Nature* **521** (2015), p. 436-444.
- [18] P.-G. Lemarié, “Ondelettes à localisation exponentielle”, *J. Math. Pures Appl.* **67** (1988), no. 3, p. 227-236.
- [19] P.-G. Lemarié, Y. F. Meyer, “Ondelettes et bases hilbertiennes”, *Rev. Mat. Iberoam.* **2** (1986), no. 1-2, p. 1-18.
- [20] M. Li, Z. Ma, Y. G. Wang, X. Zhuang, “Fast Haar Transform for Graph Neural Networks”, *Neural Networks* **128** (2020), p. 188-198.
- [21] Y. Meyer, *Ondelettes et Opérateurs. I : Ondelettes. II : Opérateurs de Calderón–Zygmund*, Actualités Mathématiques, Hermann, 1990.
- [22] J.-C. Pesquet, H. Krim, H. Carfantan, “Time-invariant orthonormal wavelet representations”, *IEEE Trans. Signal Process.* **44** (1996), no. 8, p. 1964-1970.
- [23] J.-C. Pesquet, H. Krim, H. Carfantan, J. G. Proakis, “Estimation of noisy signals using time-invariant wavelet packets”, in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, vol. 1, IEEE, 1993, p. 31-34.
- [24] T. Tao, “On the almost everywhere convergence of wavelet summation methods”, *Appl. Comput. Harmon. Anal.* **3** (1996), no. 4, p. 384-387.
- [25] G. G. Walter, “Pointwise convergence of wavelet expansions”, *J. Approx. Theory* **80** (1995), no. 1, p. 108-118.