# Approximation in the Zygmund Class

Artur Nicolau and Odí Soler i Gibert [*]

**Abstract**

We study the distance in the Zygmund class $\Lambda_*$ to the subspace I(BMO) of functions with distributional derivative with bounded mean oscillation. In particular, we describe the closure of I(BMO) in the Zygmund seminorm. We also generalise this result to Zygmund measures on $\mathbb{R}^d$. Finally, we apply the techniques developed in the article to characterise the closure of the subspace of functions in $\Lambda_*$ that are also in the classical Sobolev space $W^{1,p}$, for $1 < p < \infty$.

## 1 Introduction

A continuous real valued function $f$ on the real line belongs to the *Zygmund class* $\Lambda_*$ if

$$\|f\|_* = \sup_{\substack{x,h \in \mathbb{R} \\ h > 0}} |\Delta_2 f(x,h)| < \infty,$$

where

$$\Delta_2 f(x,h) = \frac{f(x+h) - 2f(x) + f(x-h)}{h}$$

denotes the *second divided difference* centred at $x$ with step $h$, or in other words, the second divided difference on the interval $I = (x-h, x+h)$ and denoted $\Delta_2 f(I) = \Delta_2 f(x,h)$. For a function $f \in \Lambda_*$, the quantity $\|f\|_*$ is called the *Zygmund seminorm* of $f$. The Zygmund class is the natural substitute of the space of Lipschitz functions in many different contexts as polynomial approximation, Bessel potentials, Calderón-Zygmund theory, and has been extensively studied (see for instance [Zyg45], Chapter V of [Ste70], [Mak89], [DLN14]).

For a measurable set $A \subset \mathbb{R}$, we denote by $|A|$ its Lebesgue measure, and we will denote by $\chi_A$ its indicator function. We use the standard notation $a \lesssim b$ (respectively $a \gtrsim b$) if there exists an absolute constant $C > 0$ such that $a \leq Cb$ (resp. $a \geq Cb$). We will also denote $a \simeq b$ if $a \lesssim b$ and $a \gtrsim b$.

---

1

A locally integrable function $f$ on the real line is said to have *bounded mean oscillation, $f \in$ BMO*, if

$$\|f\|_{\mathrm{BMO}} = \sup_I \left( \frac{1}{|I|} \int_I |f(x) - f_I|^2 \, dx \right)^{1/2} < \infty, \tag{1}$$

where $I$ ranges over all finite intervals in $\mathbb{R}$ and where

$$f_I = \frac{1}{|I|} \int_I f(x) \, dx,$$

is the average of $f$ on $I$. The space of continuous functions such that their derivatives, in the sense of distributions, are BMO functions is

$$\mathrm{I(BMO)} = \{ f \in \mathcal{C}(\mathbb{R}) \colon f' \in \mathrm{BMO} \}.$$

It is easy to check that $\mathrm{I(BMO)} \subsetneq \Lambda_*$. In [Str80], R. Strichartz found a characterisation for functions in I(BMO) in terms of their second divided differences. We state it below for compactly supported functions.

**Theorem A** (R. Strichartz). *A compactly supported function $f$ is in* I(BMO) *if and only if*

$$\sup_I \frac{1}{|I|} \int_I \int_0^{|I|} |\Delta_2 f(x, h)|^2 \frac{dh \, dx}{|h|} < \infty, \tag{2}$$

*where $I$ ranges over all finite intervals on $\mathbb{R}$.*

One of the main goals of this article is to give an analog of Theorem A for functions in the closure $\overline{\mathrm{I(BMO)}}$ in the Zygmund seminorm $\|\cdot\|_*$, for which we will consider the pseudometric $\mathrm{dist}(f, g) = \|f - g\|_*$ for any pair of functions $f, g \in \Lambda_*$. To this end, from now on for a given function $f \in \Lambda_*$ and $\varepsilon > 0$, consider the set

$$A(f, \varepsilon) = \{ (x, h) \in \mathbb{R}^2_+ \colon |\Delta_2 f(x, h)| > \varepsilon \},$$

where we use $\mathbb{R}^2_+$ to denote the upper halfplane $\mathbb{R}^2_+ = \{ (x, h) \colon x \in \mathbb{R}, h > 0 \}$.

**Theorem 1.** *Let $f$ be a compactly supported function in $\Lambda_*$. For each $\varepsilon > 0$, consider*

$$C(f, \varepsilon) = \sup_I \frac{1}{|I|} \int_I \int_0^{|I|} \chi_{A(f,\varepsilon)}(x, h) \frac{dh \, dx}{h},$$

*where $I$ ranges over all finite intervals. Then,*

$$\mathrm{dist}(f, \mathrm{I(BMO)}) \simeq \inf \{ \varepsilon > 0 \colon C(f, \varepsilon) < \infty \}. \tag{3}$$

We deduce the following description of $\overline{\mathrm{I(BMO)}}$.

**Corollary 1.** *Let $f$ be a compactly supported function in $\Lambda_*$. Then $f \in \overline{\mathrm{I(BMO)}}$ if and only if for every $\varepsilon > 0$ there exists a constant $C(\varepsilon) > 0$ such that*

$$\frac{1}{|I|} \int_I \int_0^{|I|} \chi_{A(f,\varepsilon)}(x,h) \frac{dh\,dx}{h} \leq C(\varepsilon),$$

*for every finite interval $I$.*

Observe that Theorem 1 is actually a local result, and in this sense it can still be applied to functions that are not compactly supported by restricting to a finite interval. Hence, these results also hold for functions defined on the unit circle. It is worth mentioning that, for functions defined on the unit circle, the closure of the trigonometric polynomials in the Zygmund seminorm is the small Zygmund class (see [Zyg45]). Observe as well that Theorem 1 also implies uniform approximation locally in the following sense. It is a well known fact (see for instance [JW84]) that for any function $f \in \Lambda_*$, and for any finite interval $I \subseteq \mathbb{R}$, there exists a polynomial $p_I$ of degree 1 such that

$$|f(x) - p_I(x)| \lesssim |I| \, \|f\|_*, \quad x \in I.$$

Thus, if $f \in \Lambda_*$ is compactly supported on an interval $I_0$, there is $g \in \mathrm{I(BMO)}$ such that for any interval $I \subseteq I_0$ there exists a linear polynomial $p_I$ with

$$|f(x) - (g + p_I)(x)| \lesssim |I| \operatorname{dist}(f, \mathrm{I(BMO)}), \quad x \in I.$$

The lower bound in (3) is easy, and the main part of the paper is devoted to prove the upper bound. We will first introduce a dyadic version of the Zygmund class, BMO and I(BMO), and the corresponding notion for dyadic martingales. Then we state and prove a discrete version of (3). Afterwards, an averaging argument of J. Garnett and P. Jones (see [GJ82]) is used to prove the continuous result from the dyadic one. To this end, certain technical estimates are needed, which we have collected in Section 2.

For $n \geq 0$, let $\mathcal{D}_n = \{[k2^{-n}, (k+1)2^{-n}) \colon k \in \mathbb{Z}\}$ be the collection of dyadic intervals of length $2^{-n}$. For $n < 0$, consider $m$ such that $n = -2m+1$ or $n = -2m$, and let $t_n = (4^m - 1)/3$. In this case, define $\mathcal{D}_n = \{[k2^{-n} - t_n, (k+1)2^{-n} - t_n) \colon k \in \mathbb{Z}\}$. Denote by $\mathcal{D} = \bigcup_{n \in \mathbb{Z}} \mathcal{D}_n$. We will call the intervals in $\mathcal{D}$ *dyadic intervals*. This definition might look unnecessarily complicated for the dyadic intervals with $n < 0$, where we add a translation by $t_n$ units with respect to the previous ones, but it will turn out to be convenient later on. The reason is that with this choice any finite interval $I \subset \mathbb{R}$ is contained in some interval of $\mathcal{D}$, which is not true if we do not include any such translations.

A locally integrable function $f$ has *dyadic bounded mean oscillation*, $f \in \mathrm{BMO}_d$, if condition (1) is required only for dyadic intervals, that is, if

$$\|f\|_{\mathrm{BMO}\,d} = \sup_{I \in \mathcal{D}} \left( \frac{1}{|I|} \int_I |f(x) - f_I|^2 \, dx \right)^{1/2} < \infty.$$

Note that $\mathrm{BMO} \subset \mathrm{BMO}_d$. The space $\mathrm{BMO}_d$ has been studied as a natural discrete substitute of BMO (see, for instance, [GJ82], [Mei03] and [Con13]). The following result is stated in [GJ82] and summarises the averaging technique previously mentioned.

**Theorem B** (J. Garnett, P. Jones). *Suppose that $\alpha \mapsto b^{(\alpha)}$ is a measurable mapping from $\mathbb{R}$ to $\mathrm{BMO}_d$ such that all $b^{(\alpha)}$ are supported on a fixed dyadic interval $I_0$, and such that for every $\alpha$, $\left\| b^{(\alpha)} \right\|_{BMOd} \leq 1$ and*

$$\int_{\mathbb{R}} b^{(\alpha)}(x) \, dx = 0.$$

*Then*

$$b_R(x) = \frac{1}{2R} \int_{-R}^{R} b^{(\alpha)}(x + \alpha) \, d\alpha$$

*is in BMO and there is a constant $C > 0$ such that $\|b_R\|_{BMO} \leq C$ for any $R \geq 1$.*

We shall need an analogous result for the Zygmund class. We say that a continuous function $f$ belongs to the *dyadic Zygmund class, $f \in \Lambda_{*d}$*, if

$$\|f\|_{*d} = \sup_{I \in \mathcal{D}} |\Delta_2 f(I)| < +\infty.$$

Observe as well that $\Lambda_* \subsetneq \Lambda_{*d}$.

**Theorem 2.** *Suppose that $\alpha \mapsto t^{(\alpha)}$ is a measurable mapping from $\mathbb{R}$ to $\Lambda_{*d}$ such that all $t^{(\alpha)}$ are supported on a fixed dyadic interval $I_0$, and such that for every $\alpha$, $\left\| t^{(\alpha)} \right\|_{*d} \leq 1$. Then, the function*

$$t_R(x) = \frac{1}{2R} \int_{-R}^{R} t^{(\alpha)}(x + \alpha) \, d\alpha, \quad x \in \mathbb{R}$$

*is in $\Lambda_*$ and there is a constant $C > 0$ such that $\|t_R\|_* \leq C$ for any $R \geq 1$.*

As an application of the techniques exposed in the article, we also show a result similar to Theorem 1 for Sobolev spaces. For $1 < p < \infty$, we consider the Sobolev space $W^{1,p}$ of functions $f \in L^p$ whose derivative $f'$ in the distributional sense is also in $L^p$. Take then the subspace of the Zygmund class $\Lambda_*^p = W^{1,p} \cap \Lambda_*$. The next theorem gives estimates for distances to this subspace. Here, for $x \in \mathbb{R}$, $\Gamma(x)$ denotes the truncated cone defined as $\Gamma(x) = \{(t, h) \in \mathbb{R}_+^2 : |x - t| < h, \, 0 < h < 1\}$.

**Theorem 3.** *Let $f$ be a compactly supported function in $\Lambda_*$. For each $\varepsilon > 0$, define the function*

$$C(f, \varepsilon)(x) = \left( \int_{\Gamma(x)} \chi_{A(f,\varepsilon)}(s, t) \frac{ds \, dt}{t^2} \right)^{1/2}, \quad x \in \mathbb{R}.$$

*Then,*

$$\mathrm{dist}(f, \Lambda_*^p) \simeq \inf\{\varepsilon > 0 : C(f, \varepsilon) \in L^p\}. \tag{4}$$

4

Finally, we find a higher dimensional analog of Theorem 1 for Zygmund measures in $\mathbb{R}^d$. Recall that a signed Borel measure $\mu$ on $\mathbb{R}^d$ is called a Zygmund measure if

$$\|\mu\|_* = \sup_Q \left| \frac{\mu(Q)}{|Q|} - \frac{\mu(Q^*)}{|Q^*|} \right| < \infty,$$

where $Q$ ranges over all finite cubes in $\mathbb{R}^d$ with edges parallel to the axis, and where $Q^*$ denotes the cube with the same centre as $Q$ but double side length. In the case $d = 1$ it is obvious that $\mu$ is a Zygmund measure if and only if its primitive $f(x) = \mu([0, x])$ is in the Zygmund class. Note that there exist Zygmund measures that are singular with respect to the Lebesgue measure, as J. P. Kahane showed [Kah69]. More information on Zygmund measures can be found in [Mak89], [AP89] and [AAN99]. We consider the space of absolutely continuous measures $\nu$ such that $d\nu(x) = f(x) \, dx$ for some $f \in \mathrm{BMO}(\mathbb{R}^d)$. We call this the space of I(BMO) measures. It is clear that a measure in I(BMO) is a Zygmund measure as well. As before, given a Zygmund measure $\mu$ on $\mathbb{R}^d$, we want to describe the distance

$$\mathrm{dist}(\mu, \mathrm{I(BMO)}) = \inf\{\|\mu - \sigma\|_* : \sigma \in \mathrm{I(BMO)}\}.$$

For $x \in \mathbb{R}^d$ and $h > 0$, let $Q(x, h)$ be the cube centred at $x$ of sidelength $h$. For a given Zygmund measure $\mu$ and for $\varepsilon > 0$, consider the set

$$A(\mu, \varepsilon) = \left\{ (x, h) \in \mathbb{R}_+^{d+1} \colon \left| \frac{\mu(Q(x,h))}{|Q(x,h)|} - \frac{\mu(Q(x,2h))}{|Q(x,2h)|} \right| > \varepsilon \right\},$$

where we use $\mathbb{R}_+^{d+1}$ to denote the upper halfspace $\mathbb{R}_+^{d+1} = \{(x, h) \colon x \in \mathbb{R}^d, h > 0\}$.

**Theorem 4.** *Let $\mu$ be a compactly supported Zygmund measure on $\mathbb{R}^d$. For each $\varepsilon > 0$, consider*

$$C(\mu, \varepsilon) = \sup_Q \frac{1}{|Q|} \int_Q \int_0^{l(Q)} \chi_{A(\mu,\varepsilon)}(x, h) \, \frac{dh \, dx}{h},$$

*where $Q$ ranges over all finite cubes and $l(Q)$ denotes the side length of $Q$. Then,*

$$\mathrm{dist}(\mu, \mathrm{I(BMO)}) \simeq \inf\{\varepsilon > 0 \colon C(\mu, \varepsilon) < \infty\}.$$

This paper is organised in the following manner. In Section 2, we expose the technical estimates that we need in order to apply the averaging argument previously mentioned. We then state and prove the dyadic analog of Theorem 1 in Section 3. In Section 4, we explain the averaging argument that yields Theorem 2 and then we use it to prove Theorem 1. Next, we explain in Section 5 the variations in the previous construction that allow us to prove

5

Theorem 4. We devote Section 6 to the application of our methods, showing Theorem 3. Finally, in Section 7 we state three open problems closely related to our results.

It is a pleasure to thank Petros Galanopoulos, Oleg Ivrii and Martí Prats for several helpful conversations and interesting comments.

## 2 Preliminaries

We need an auxiliary result that estimates the oscillation of the second divided differences when changing their centre and step size. For a continuous function $f$ we define its *first divided difference* at $x \in \mathbb{R}$ with step size $h > 0$ as
$$\Delta_1 f(x, h) = \frac{f(x+h) - f(x)}{h}.$$
For convenience, we may also denote $\Delta_1 f(x, h) = \Delta_1 f(I)$, where $I = (x, x+h)$.

**Lemma 1.** *Let $f \in \Lambda_*$ and assume that $h' > h > 0$ and $|x - t| < h'/2$. Then*

$$|\Delta_2 f(x, h) - \Delta_2 f(t, h')| \lesssim$$
$$\|f\|_* \left( \frac{h' - h}{h'} \left( 1 + \log \frac{h'}{h' - h} \right) + \frac{|x - t|}{h'} \log \left( \frac{h'}{|x - t|} + 1 \right) \right). \quad (5)$$

*Proof.* We split the proof in two steps. First, we find an estimate for the case $h' = h$ and then another one for $x = t$. We start showing that, for $h > 0$, when $|x - t| < h/2$, then

$$|\Delta_2 f(x, h) - \Delta_2 f(t, h)| \lesssim \|f\|_* \frac{|x - t|}{h} \log \left( \frac{h}{|x - t|} + 1 \right). \quad (6)$$

We claim that, if $|x - t| > h/2$, then

$$|\Delta_1 f(x, h) - \Delta_1 f(t, h)| \lesssim \|f\|_* \log \left( \frac{|x - t|}{h} + 1 \right). \quad (7)$$

Indeed, let $u$ be the harmonic extension of $f$ on the upper halfplane $\mathbb{R}^2_+$. It is a well known fact (see Chapter V of [Ste70] or [Llo02]) that

$$\left| \frac{f(x+h) - f(x)}{h} - u_x(x, h) \right| \lesssim \|f\|_*,$$

and that

$$\sup_{(x,h) \in \mathbb{R}^2_+} h |\nabla u_x(x, h)| \lesssim \|f\|_*.$$

Thus, if we denote by $\rho(a, b)$ the hyperbolic distance between two points $a, b \in \mathbb{R}^2_+$, we get

$$|u_x(x, h) - u_x(t, h)| \lesssim \|f\|_* \rho((x, h), (t, h)).$$

6

Using the estimate

$$\rho((x,h),(t,h)) \lesssim \log\left(\frac{|x-t|}{h}+1\right),$$

we get (7).

Now, assume $x > t$ without loss of generality, and $x - t < h/2$. Write

$$
\begin{aligned}
h(\Delta_2 f(x,h) - \Delta_2 f(t,h)) = {} & (f(x+h) - f(t+h)) - (f(x) - f(t)) \\
& + (f(x-h) - f(t-h)) - (f(x) - f(t))
\end{aligned}
$$

and apply (7) to the first two terms taking $x' = t + h$, $t' = t$ and $h' = x - t$, and to the last two taking $x' = t - h$, $t' = t$ and $h' = x - t$. This shows (6).

Assume now that $h' > h > 0$. We want to see that

$$|\Delta_2 f(x,h') - \Delta_2 f(x,h)| \lesssim \|f\|_* \frac{h'-h}{h'}\left(1 + \log\frac{h'}{h'-h}\right). \qquad (8)$$

First note the following identity

$$
\begin{aligned}
\Delta_2 f(x,h) - \Delta_2 f(x,h') = {} & \\
& \frac{h'-h}{h'}[\Delta_2 f(x,h) - (\Delta_1 f(x+h,h'-h) - \Delta_1 f(x-h',h'-h))].
\end{aligned}
$$

Using (7) on the last two terms, we get (8). Finally, (5) is a direct consequence of (6) and (8). $\qquad \square$

## 3 The Dyadic Results

A *dyadic rational* is a number of the form $k2^{-n}$ with $k, n \in \mathbb{Z}$. For $n \geq 0$, let $\mathcal{D}_n = \{[k2^{-n},(k+1)2^{-n}): k \in \mathbb{Z}\}$. For $n < 0$, consider $m$ such that $n = -2m + 1$ or $n = -2m$, and let $t_n = (4^m - 1)/3$. In this case, define $\mathcal{D}_n = \{[k2^{-n} - t_n, (k+1)2^{-n} - t_n): k \in \mathbb{Z}\}$. A *dyadic interval* $I$ is an interval such that $I \in \mathcal{D}_n$ for some $n \in \mathbb{Z}$, and in this case we say that $I$ is a dyadic interval *of generation $n$*. Denote by $\mathcal{D} = \bigcup_{n \in \mathbb{Z}} \mathcal{D}_n$ the set of all dyadic intervals. Note that, given $I \in \mathcal{D}_n$ for $n \in \mathbb{Z}$, there is a unique interval $I^*$ in $\mathcal{D}_{n-1}$ that contains $I$, which we call the *predecessor of $I$*. If $I_0$ is an arbitrary interval, we will use the notation $\mathcal{D}(I_0) = \{I \in \mathcal{D}: I \subseteq I_0\}$. As explained in the introduction, a continuous real valued function $f$ on $\mathbb{R}$ belongs to the *dyadic Zygmund class,* denoted $f \in \Lambda_{*d}$, if

$$\|f\|_{*d} = \sup_{I \in \mathcal{D}} |\Delta_2 f(I)| < \infty.$$

In a similar fashion, we say that a locally integrable function $f$ has *bounded dyadic mean oscillation,* $f \in \mathrm{BMO}_d$, if

$$\|f\|_{\mathrm{BMO}\,d} = \sup_{I \in \mathcal{D}} \left(\frac{1}{|I|}\int_I |f(x) - f_I|^2\,dx\right)^{1/2} < \infty,$$

and we consider the *dyadic* I(BMO) space to be the space of continuous real valued functions on $\mathbb{R}$ whose distributional derivatives belong to $\text{BMO}_d$, that is

$$I(\text{BMO})_d = \{f \in \mathcal{C}(\mathbb{R}) \colon f' \in \text{BMO}_d\}.$$

It is easy to see that each dyadic space contains its corresponding homogeneous space, that is $\text{BMO} \subseteq \text{BMO}_d$ and $\Lambda_* \subseteq \Lambda_{*d}$. It is important to remark, as well, that none of these pairs are equal. More information on the relation between BMO and $\text{BMO}_d$ can be found in [GJ82], [Mei03] and [Con13].

The spaces $\Lambda_{*d}$ and $I(\text{BMO})_d$ can be regarded as well as spaces of dyadic martingales. We say that a sequence of functions $S = \{S_n\}$ is a *dyadic martingale* if for all $n \geq 0$ the following conditions are satisfied:

(i) $S_n$ is constant on any $I \in \mathcal{D}_n$,

(ii) $S_n|I = \frac{1}{2}\left(S_{n+1}|I^{(1)} + S_{n+1}|I^{(2)}\right)$ for all $I \in \mathcal{D}_n$, where $I^{(1)}, I^{(2)}$ are the intervals in $\mathcal{D}_{n+1}$ contained in $I$.

We will denote the value of $S_n$ at $I \in \mathcal{D}_m$, $m \geq n$, by $S_n(I)$, and, if there is no ambiguity, when $I \in \mathcal{D}_n$ we will just write $S(I)$. For $x \in \mathbb{R}$ and $n \geq 0$, let $I \in \mathcal{D}_n$ be such that $x \in I$. Then, we have that $S_n(x) = S(I)$, and we will denote $S(x) = \lim_{n\to\infty} S_n(x)$ when this limit exists. For $n \geq 1$, we call *jump of $S$ at generation $n$* the function $\Delta S_n(x) = S_n(x) - S_{n-1}(x)$, and if $I \in \mathcal{D}_n$, we use the notation $\Delta S_n(I) = S_n(I) - S_{n-1}(I^*)$, where $I^*$ is the predecessor of $I$. One can easily check that for a dyadic martingale $S$ the jumps $\Delta S_j$ and $\Delta S_k$ are orthogonal in $L^2(I)$ for any $I \in \mathcal{D}_0$ when $j \neq k$.

With these concepts at hand, we can associate to each function $f \in \Lambda_{*d}$ a dyadic martingale $S$, which we shall call *the average growth martingale of $f$*, as follows. For a dyadic interval $I = [a, b) \in \mathcal{D}_n$, set

$$S_n(I) = \frac{f(b) - f(a)}{b - a} = 2^n(f(b) - f(a)). \tag{9}$$

Now, observe that the second divided difference of $f$ can be expressed in terms of the jumps of $S$; that is, for $I \in \mathcal{D}_n$, we have the relation

$$|\Delta_2 f(I^*)| = 2|\Delta S(I)|.$$

Now it is obvious that any dyadic martingale $S$ is related to a function $f \in \Lambda_{*d}$ (up to a linear term) through the relation (9) if and only if

$$\|S\|_* = \sup_{I \in \mathcal{D}} |\Delta S(I)| < \infty.$$

To get the corresponding description of martingales associated with $I(\text{BMO})_d$ functions, we will discretise (1). Note that for $f \in I(\text{BMO})_d$, with average

8

growth martingale $S$, and $I \in \mathcal{D}_N$, using that the jumps $\{\Delta S_n\}_{n \geq N}$ restricted to $I$ of the martingale $S$ are orthogonal in $L^2$, one can express

$$\int_I |f'(x) - f'_I|^2 \, dx = \int_I \sum_{n > N} |\Delta S_n(x)|^2 \, dx.$$

Thus, a martingale $S$ is related to a function $f \in I(\mathrm{BMO})_d$ through the relation (9) if and only if

$$\|S\|_{\mathrm{BMO}} = \sup_{I \in \mathcal{D}} \left( \frac{1}{|I|} \sum_{J \in \mathcal{D}(I)} |\Delta S(J)|^2 |J| \right)^{1/2} < \infty. \tag{10}$$

The analog of Theorem 1 for this setting is the following

**Theorem 5.** *Let $f$ be a compactly supported function in $\Lambda_{*d}$. For a fixed $\varepsilon > 0$, define $D(f, \varepsilon)$ by*

$$D(f, \varepsilon) = \sup_{I \in \mathcal{D}} \frac{1}{|I|} \sum_{\substack{J \in \mathcal{D}(I) \\ |\Delta_2 f(J)| > \varepsilon}} |J|. \tag{11}$$

*Then,*
$$\mathrm{dist}(f, I(\mathrm{BMO})_d) = \inf\{\varepsilon > 0 \colon D(f, \varepsilon) < \infty\}. \tag{12}$$

Note that we can rewrite this result in terms of martingales. Let $f \in \Lambda_{*d}$ be compactly supported on a dyadic interval $I_0$, and consider its average growth martingale $S$ defined by (9). In this way, $D(f, \varepsilon)$ in (11) can be expressed as

$$D(f, \varepsilon) = \sup_{I \in \mathcal{D}(I_0)} \frac{1}{|I|} \sum_{\substack{J \in \mathcal{D}(I) \\ |\Delta S(J)| > \varepsilon/2}} |J|. \tag{13}$$

*Proof of Theorem 5.* Without loss of generality, let us assume that $f$ is supported on the dyadic interval $I_0 = [0, 1]$. We need to prove that, for a given $\varepsilon > 0$, there exists a function $b \in I(\mathrm{BMO})_d$ satisfying $\|f - b\|_{*d} \leq \varepsilon$ if and only if $D(f, \varepsilon) < \infty$. Denote by $\varepsilon_0$ the infimum in the left-hand side of (12).

Given $\varepsilon > \varepsilon_0$, we will construct a function $b \in I(\mathrm{BMO})_d$ such that $\|f - b\|_{*d} \leq \varepsilon$. Consider the average growth martingale $S$ for function $f$, defined by (9). First, we approximate the martingale $S$ by a martingale $B$ related to an $I(\mathrm{BMO})_d$ function, that is satisfying (10). Take $B(I_0) = S(I_0)$ and construct $B$ by setting $\Delta B(J) = \Delta S(J)$ whenever $|\Delta S(J)| > \varepsilon/2$ and $\Delta B(J) = 0$ otherwise, for $J \in \mathcal{D}(I_0)$.

By construction, it is clear that $\|S - B\|_* \leq \varepsilon/2$. Moreover, for any $I \in \mathcal{D}$, we have

$$\sum_{J \in \mathcal{D}(I)} |\Delta B(J)|^2 |J| \leq \|S\|_*^2 \sum_{\substack{J \in \mathcal{D}(I) \\ |\Delta S(J)| > \varepsilon/2}} |J| \leq |I| \, \|S\|_*^2 \, D(f, \varepsilon),$$

9

showing that $B$ satisfies (10).

Now, using that the jumps $\Delta B_j$ and $\Delta B_k$ are orthogonal in $L^2$, we have

$$\int_{I_0} \left( \sum_{n=1}^{\infty} \Delta B_n(x) \right)^2 dx = \int_{I_0} \sum_{n=1}^{\infty} |\Delta B_n(x)|^2 \, dx = \sum_{J \in \mathcal{D}(I_0)} |\Delta B(J)|^2 |J| < \infty.$$

This gives that $\lim_{n \to \infty} B_n(x)$ exists at almost every point $x \in I_0$ and it is actually a square integrable function, so that we can integrate it to get $b(x) = \int_0^x \lim_n B_n(s) \, ds \in \mathrm{I(BMO)}_d$ such that $\|f - b\|_{*d} \leq \varepsilon$.

Finally, if $\varepsilon < \varepsilon_0$, we show that no function $b \in \mathrm{I(BMO)}_d$ satisfies $\|f - b\|_{*d} \leq \varepsilon$. Take $\varepsilon_0 > \varepsilon_1 > \varepsilon$, assume that there is $b \in \mathrm{I(BMO)}_d$ satisfying $\|f - b\|_{*d} \leq \varepsilon$, and let $S$ and $B$ be the respective average growth martingales for $f$ and $b$. For any $I \in \mathcal{D}$ such that $|\Delta S(I)| > \varepsilon_1$, we have that $|\Delta B(I)| > \varepsilon_1 - \varepsilon = \delta > 0$. Thus

$$\frac{1}{|I|} \sum_{J \in \mathcal{D}(I)} |\Delta B(J)|^2 |J| > \frac{\delta^2}{|I|} \sum_{\substack{J \in \mathcal{D}(I) \\ |\Delta S(J)| > \varepsilon_1}} |J|.$$

The supremum of this quantity when $I$ ranges over all dyadic intervals is $\delta^2 D(f, \varepsilon_1) = +\infty$. This contradicts condition (10) for martingale $B$ and, hence, that $b$ is an $\mathrm{I(BMO)}_d$ function, concluding the proof of the theorem. $\square$

# 4    From the Dyadic to the Continuous Setting

Before proving Theorem 2, let us make some observations. Consider the measurable mapping $\alpha \mapsto t^{(\alpha)}$ from $\mathbb{R}$ to $\Lambda_{*d}$ such that all $t^{(\alpha)}$ are supported on $I_0 = [0, 1]$ and such that $\left\| t^{(\alpha)} \right\|_{*d} \leq 1$, and let $R \geq 1$. We will denote by $\mathcal{D}^0 = \mathcal{D}$ the standard dyadic filtration and by $\mathcal{D}^\beta$ the translated filtration by $-\beta$ units. We also extend this notation to denote by $\mathcal{D}_n^0$ the set of intervals of size $2^{-n}$ in $\mathcal{D}^0$ and by $\mathcal{D}_n^\beta$ the set of intervals of the same size in $\mathcal{D}^\beta$. Similarly, we denote by $\Lambda_*^0$ the dyadic Zygmund class with respect to the filtration $\mathcal{D}^0$ and $\Lambda_*^\beta$ the Zygmund dyadic class with respect to $\mathcal{D}^\beta$. With this notation, if $f(x) \in \Lambda_*^0$, then $f(x + \beta) \in \Lambda_*^\beta$.

Now consider an arbitrary interval $I$ and the adjacent interval $\tilde{I} = I - |I|$ of the same size. Fix $R \geq 1$ and $\alpha \in [-R, R]$ and let $n$ be the minimum integer such that $I$ contains an interval of $\mathcal{D}_n^\alpha$, and let $\mathcal{F}_n(I)$ be the set of all such intervals. For each $m > n$, let $\mathcal{F}_m(I)$ be the set of intervals $J \in \mathcal{D}_m^\alpha$ such that $J \subset I \setminus \cup_{j=n}^{m-1} \mathcal{F}_j(I)$. Then, $\mathcal{F}(I) = \cup_{j \geq n} \mathcal{F}_j(I)$ is a covering of $I$ by intervals of $\mathcal{D}^\alpha$. The covering $\mathcal{F}(\tilde{I})$ of $\tilde{I}$ is constructed in the exact same way.

Let us say that $\mathcal{F}(I) = \{I_j\}_{j=1}^{\infty}$. We may assume that the intervals $I_j$ are ordered in the following way. Whenever $j > k$, $|I_j| \leq |I_k|$, and we may

take $I_k$ to be to the left of $I_j$ if $|I_j| = |I_k|$. That is, we order the intervals decreasing in size and left to right for those that have the same length. We consider the covering $\mathcal{F}(\tilde{I}) = \{\tilde{I}_j\}_{j=1}^{\infty}$ to be ordered in the same way.

**Lemma 2.** *Let $I \subseteq \mathbb{R}$ be a finite interval and $\mathcal{F}(I)$ its covering by intervals of $\mathcal{D}^\alpha$ constructed and ordered as previously explained. Then, the intervals of $\mathcal{F}(I)$ have disjoint interiors and, for $j \geq 1$, they satisfy that $|I_{j+2}| \leq |I_j|/2$.*

*Proof.* The intervals in $\mathcal{F}(I)$ have disjoint interiors by construction. Moreover, these intervals are maximal in the sense that if $J \in \mathcal{F}(I)$ and $J \subsetneq J' \in \mathcal{D}^\alpha$, then $J' \not\subset I$. Thus, it is clear that for each $n \geq 1$ there are at most two intervals in $\mathcal{F}(I)$ of size $2^{-n}|I|$. This yields that, for $j \geq 1$, the intervals in $\mathcal{F}(I)$ satisfy $|I_{j+2}| \leq |I_j|/2$. $\qquad\square$

When $|I| = 2^{-n}$ for some $n \in \mathbb{Z}$, the covering $\mathcal{F}(\tilde{I}) = \{\tilde{I}_j\}_{j=1}^{\infty}$ is a translation of $\mathcal{F}(I) = \{I_j\}_{j=1}^{\infty}$. More precisely, if we order both $\{\tilde{I}_j\}$ and $\{I_j\}$ as previously explained, then for each $j \geq 1$ we have that $\tilde{I}_j = I_j - |I|$ and, trivially, for every $j \geq 1$, $|\tilde{I}_j| = |I_j|$. However, for an arbitrary interval $I$, the sizes of the intervals in $\mathcal{F}(I)$ and $\mathcal{F}(\tilde{I})$ may be completely different. For instance, it could happen that for a given $j \in \mathbb{Z}$, $\mathcal{F}(I)$ had two intervals of size $2^{-j}$ while $\mathcal{F}(\tilde{I})$ had only one.

**Lemma 3.** *Let $I$ and $\tilde{I}$ be two adjacent intervals of the same length. Fix $\alpha \in \mathbb{R}$. Then there are coverings $\mathcal{G}(I) = \{J_j\}$ and $\mathcal{G}(\tilde{I}) = \{\tilde{J}_j\}$, of $I$ and $\tilde{I}$ respectively, both consisting of intervals of $\mathcal{D}^\alpha$, with $|J_j| = |\tilde{J}_j|$ for any $j$, and with $|J_{j+2}| \leq |J_j|/2$.*

*Proof.* Consider the previous coverings $\mathcal{F}(I) = \{I_j\}$ and $\mathcal{F}(\tilde{I}) = \{\tilde{I}_j\}$. If $|I_1| = |\tilde{I}_1|$, then take $J_1 = I_1$ and $\tilde{J}_1 = \tilde{I}_1$. If these sizes are different, assume $|I_1| > |\tilde{I}_1|$ (otherwise the procedure is the same), there exists an integer $k \geq 2$ such that $\sum_{j=1}^{k} |\tilde{I}_j| = |I_1|$. Note that $k$ exists because all $|I_j|$ (and also $|\tilde{I}_j|$) are dyadic rationals that add up to $|I| = |\tilde{I}|$. Then take $\tilde{J}_j = \tilde{I}_j$ for $1 \leq j \leq k$, and choose pairwise disjoint intervals $J_1, \ldots, J_k \in \mathcal{D}(I_1)$ such that $I_1 = \cup_{l=1}^{k} J_l$ and that, for each $1 \leq j \leq k$, $|J_j| = |\tilde{J}_j|$. Note that, for $1 \leq j \leq k-2$, we have that $|J_{j+2}| = |\tilde{J}_{j+2}| \leq |\tilde{J}_j|$ because of Lemma 2. Recursively, consider that we have fixed $\{J_j\}_{j=1}^{p} \in \mathcal{G}(I)$ and $\{\tilde{J}_j\}_{j=1}^{p} \in \mathcal{G}(\tilde{I})$, let $m, n$ be the smallest integers such that $I_m \subseteq I \setminus \cup_{j=1}^{p} J_j$ and $\tilde{I}_n \subseteq \tilde{I} \setminus \cup_{j=1}^{p} \tilde{J}_j$, and repeat the previous step with $I_m$ and $\tilde{I}_n$. $\qquad\square$

Given two finite intervals $I_1, I_2$, we say that their *minimal common predecessor in $\mathcal{D}^\alpha$*, denoted $P_\alpha(I_1, I_2)$, is the interval $P_\alpha(I_1, I_2) \in \mathcal{D}^\alpha$ such that $I_1 \cup I_2 \subseteq P_\alpha(I_1, I_2)$ and such that for every $J \in \mathcal{D}^\alpha$ that satisfies $I_1 \cup I_2 \subseteq J$, then $P_\alpha(I_1, I_2) \subseteq J$. If $I_1, I_2 \in \mathcal{D}^\alpha$, we define their *distance in the dyadic filtration $\mathcal{D}^\alpha$*, denoted by $\text{dist}_\alpha(I_1, I_2)$, as

$$\text{dist}_\alpha(I_1, I_2) = \log_2 \frac{|P_\alpha(I_1, I_2)|}{|I_1|} + \log_2 \frac{|P_\alpha(I_1, I_2)|}{|I_2|}.$$

Here it is necessary to specify the index $\alpha$ as one could have two intervals $I_1, I_2$ that were dyadic in two different filtrations $\mathcal{D}^\alpha$ and $\mathcal{D}^\beta$ such that the difference between both distances is as large as desired.

**Lemma 4.** *Consider $f \in \Lambda_{*d}^\alpha$ and $I, J \in \mathcal{D}^\alpha$. Then,*

$$|\Delta_1 f(I) - \Delta_1 f(J)| \le \|f\|_{*d} \operatorname{dist}_\alpha(I, J).$$

*Proof.* Consider the sequences $\{I_j\}_{j=0}^k$ and $\{J_j\}_{j=0}^l$ in $\mathcal{D}^\alpha$ such that $I_0 = I$, $J_0 = J$, $I_k = J_l = P_\alpha(I, J)$, and such that $I_j^* = I_{j+1}$ for $0 \le j < k$, and such that $J_j^* = J_{j+1}$ for $0 \le j < l$. One has that

$$|\Delta_1 f(I) - \Delta_1 f(J)| \le \sum_{j=0}^{k-1} |\Delta_1 f(I_j) - \Delta_1 f(I_{j+1})| + \sum_{j=0}^{l-1} |\Delta_1 f(J_j) - \Delta_1 f(J_{j+1})|.$$

Each term of these sums is bounded by $\|f\|_*$ and, since there are exactly $\operatorname{dist}_\alpha(I, J)$ terms, the result follows. $\square$

For future convenience, given a finite interval $I$, we will denote its midpoint by $c(I)$.

**Lemma 5.** *Fix $R \ge 1$ and let $I$ and $\tilde{I}$ be two adjacent intervals of the same length. Let $N$ be the integer such that $2^{-N-1} < |I| \le 2^{-N}$ and let $M$ be the integer such that $2^{M-1} < R \le 2^M$. Then, for each $k \ge 1$, one has that*

$$|\{\alpha \in [-R, R]\colon |P_\alpha(I, \tilde{I})| = 2^{k-N}\}| \le 2^{M+1} 2^{-k+2}.$$

*Proof.* Note that for any value of $\alpha$, one has that $|P_\alpha(I, \tilde{I})| = 2^k 2^{-N}$ for some positive integer $k$. For $k \ge 2$, the size of the minimal common predecessor in $\mathcal{D}^\alpha$ is exactly $2^{k-N}$ if and only if there is some $J \in \mathcal{D}^\alpha$, with $|J| = 2^{k-N}$, such that $c(J) \in I \cup \tilde{I}$. For the case $k = 1$, it is only true that if $J \in \mathcal{D}^\alpha$, with $|J| = 2^{1-N}$, is the minimal common predecessor, then $c(J) \in I \cup \tilde{I}$, while the reciprocal does not hold.

Consider $J \in \mathcal{D}_{N-k}$, and consider as well the translated intervals $J + \alpha$, for $\alpha \in [-R, R]$, and their midpoints $c(J + \alpha)$. The set $\{\alpha \in [-R, R]\colon c(J + \alpha) \in I \cup \tilde{I}\}$ has measure $2^{-N+1}$, which is the measure of $I \cup \tilde{I}$. Note that it is actually here that we implicitly use that $R \ge 1$, since otherwise it could be that this set had length $2^{M+1}$. Since in $[-R, R]$ there are at most $2^{M+1} 2^{N-k+1}$ intervals of length $2^{k-N}$, the result follows immediately. $\square$

*Proof of Theorem 2.* Assume without loss of generality that $I_0 = [0, 1]$ and fix $R \ge 1$. We just need to check that

$$\sup_I |\Delta_1 t_R(I) - \Delta_1 t_R(\tilde{I})| \le C < \infty,$$

where $I$ ranges over all finite intervals, with $\tilde{I} = I - |I|$, and where $C$ is independent of the value of $R$. Fix a finite interval $I$ and consider the integer $N$ such that $2^{-N-1} < |I| \le 2^{-N}$. First, we express

$$\Delta_1 t_R(I) - \Delta_1 t_R(\tilde{I}) = \frac{1}{2R} \int_{-R}^{R} \left( \Delta_1 t^{(\alpha)}(\alpha + I) - \Delta_1 t^{(\alpha)}(\alpha + \tilde{I}) \right) d\alpha.$$

Now, for a given $\alpha$, consider the coverings $\mathcal{G}^\alpha(I) = \{I_j\}_{j=1}^\infty$ and $\mathcal{G}^\alpha(\tilde{I}) = \{\tilde{I}_j\}_{j=1}^\infty$ given in Lemma 3, that satisfy $|I_j| = |\tilde{I}_j|$ for $j \ge 1$. We can express

$$|\Delta_1 t^{(\alpha)}(\alpha + I) - \Delta_1 t^{(\alpha)}(\alpha + \tilde{I})| \le \sum_{j \ge 1} \frac{|I_j|}{|I|} \left| \Delta_1 t^{(\alpha)}(\alpha + I_j) - \Delta_1 t^{(\alpha)}(\alpha + \tilde{I}_j) \right|$$

Observe that $\alpha + I_j \in \mathcal{D}^0$ and, since $t^{(\alpha)} \in \Lambda_{*d}^0$, using Lemma 2 and Lemma 4 we may bound the previous quantity by

$$\sum_{j \ge 1} \frac{|I_j|}{|I|} \left\| t^{(\alpha)} \right\|_{*d} \operatorname{dist}_\alpha(I_j, \tilde{I}_j) \lesssim \sum_{j \ge 1} 2^{-\frac{j}{2}} \log \left( 2^{N + \frac{j}{2}} |P_\alpha(I, \tilde{I})| \right),$$

where we have also used that $\left\| t^{(\alpha)} \right\|_{*d} \le 1$ for every $\alpha$. Summing over $j$, we get

$$|\Delta_1 t^{(\alpha)}(\alpha + I) - \Delta_1 t^{(\alpha)}(\alpha + \tilde{I})| \lesssim 1 + N + \log |P_\alpha(I, \tilde{I})|.$$

Averaging over $\alpha$, we have

$$|\Delta_1 t_R(I) - \Delta_1 t_R(\tilde{I})| \lesssim \frac{1}{R} \int_{-R}^{R} \left( 1 + N + \log |P_\alpha(I, \tilde{I})| \right) d\alpha.$$

Set $R_k = \{\alpha \in [-R, R] : |P_\alpha(I, \tilde{I})| = 2^{k-N}\}$ and recall that, by Lemma 5, $|R_k| \le 2^{M+1} 2^{-k+2}$, where $M$ is the integer such that $2^{M-1} < R \le 2^M$. Then, we can bound the last quantity by

$$2^{-M+1} \sum_{k \ge 1} \int_{R_k} \left( 1 + N + \log |P_\alpha(I, \tilde{I})| \right) d\alpha \lesssim \sum_{k \ge 1} 2^{-k} \left( 1 + N + \log 2^{k-N} \right),$$

which is bounded by some positive constant $C$. Note that the factors depending on $N$ and on $M$ cancel out, which means that this last constant depends neither on $R$ nor on $I$. $\qquad \square$

We are now ready to prove Theorem 1.

*Proof of Theorem 1.* Let $f \in \Lambda_*$ and let $\varepsilon_0$ be the infimum in the left-hand side of (3). First we show that, whenever $\varepsilon < \varepsilon_0$, there is no function $b \in \mathrm{I(BMO)}$ such that $\|f - b\|_* \le \varepsilon$. Indeed, assume that for a given $\varepsilon < \varepsilon_0$ there is $b \in \mathrm{I(BMO)}$ such that $\|f - b\|_* \le \varepsilon$. Take $\varepsilon < \varepsilon_1 < \varepsilon_0$ and note

that, whenever $|\Delta_2 f(x, h)| > \varepsilon_1$ we have that $|\Delta_2 b(x, h)| > \varepsilon_1 - \varepsilon = \delta > 0$. In particular, this means that $A(f, \varepsilon_1) \subseteq A(b, \delta)$. Thus,

$$\frac{1}{|I|} \int_I \int_0^{|I|} |\Delta_2 b(x, h)|^2 \frac{dh\,dx}{h} \geq \frac{\delta^2}{|I|} \int_I \int_0^{|I|} \chi_{A(f, \varepsilon_1)}(x, h) \frac{dh\,dx}{h},$$

but the supremum, with $I$ ranging over all finite intervals, of the later quantity is not finite since $\varepsilon_1 < \varepsilon_0$. By Theorem A, this contradicts that $b \in \mathrm{I(BMO)}$.

We are left with showing that there exists a universal constant $C > 0$ such that, for any $\varepsilon > \varepsilon_0$, there is $b = b(\varepsilon) \in \mathrm{I(BMO)}$ such that $\|f - b\|_* \leq C\varepsilon$. For any such $\varepsilon$, by assumption we have that

$$C(f, \varepsilon) = \sup_I \frac{1}{|I|} \int_I \int_0^{|I|} \chi_{A(f, \varepsilon)}(x, h) \frac{dh\,dx}{h} < \infty. \tag{14}$$

Assume now, without loss of generality, that $f$ has support in $I_0 = [0, 1]$. We claim that (14) implies that $D(f, \varepsilon)$ defined by (13) is finite. To see this, take $\varepsilon > \varepsilon_1 > \varepsilon_0$, and let $J \in \mathcal{D}$ be such that $|\Delta S(J)| > \varepsilon/2$, which is equivalent to say that $|\Delta_2 f(c(J^*), |J|)| > \varepsilon$. By Lemma 1, there exists $\delta > 0$ such that if $|x - c(J^*)| < \delta |J|$ and $1 - \delta < h/|J| < 1 + \delta$, then $|\Delta_2 f(x, h)| > \varepsilon_1$. Applying this to every dyadic interval $J$ with $|\Delta S(J)| > \varepsilon/2$, we find the upper bound

$$\frac{1}{|I|} \sum_{\substack{J \in \mathcal{D}(I) \\ |\Delta S(J)| > \varepsilon/2}} |J| \lesssim \frac{1}{|I|} \int_I \int_0^{|I|} \chi_{A(f, \varepsilon_1)}(x, h) \frac{dh\,dx}{h} \leq C(f, \varepsilon_1)$$

for all $I \in \mathcal{D}$. Thus,

$$D(f, \varepsilon) = \sup_{I \in \mathcal{D}} \frac{1}{|I|} \sum_{\substack{J \in \mathcal{D}(I) \\ |\Delta S(J)| > \varepsilon/2}} |J| \lesssim C(f, \varepsilon_1). \tag{15}$$

Next, for each $\alpha \in [-1, 1]$, define $f^{(\alpha)}(x) = f(x - \alpha) \in \Lambda_{*d}$. By (15) and Theorem 5, $\mathrm{dist}(f^{(\alpha)}, \mathrm{I(BMO)}_d) \leq \varepsilon$. Hence, there are $b^{(\alpha)} \in \mathrm{I(BMO)}_d$ and $t^{(\alpha)} \in \Lambda_{*d}$ such that $f^{(\alpha)} = b^{(\alpha)} + t^{(\alpha)}$, with $\|t^{(\alpha)}\|_{*d} \leq \varepsilon$ for all $\alpha \in [-1, 1]$. This allows us to express

$$f(x) = \frac{1}{2} \int_{-1}^1 f^{(\alpha)}(x + \alpha)\,d\alpha$$

$$= \frac{1}{2} \int_{-1}^1 b^{(\alpha)}(x + \alpha)\,d\alpha + \frac{1}{2} \int_{-1}^1 t^{(\alpha)}(x + \alpha)\,d\alpha.$$

By Theorem B, taking $R = 1$, the first integral yields a function $b \in \mathrm{I(BMO)}$. By Theorem 2, with $R = 1$ as well, the second integral yields a function $t \in \Lambda_*$ with $\|t\|_* \leq C\varepsilon$, where the later constant is the same that appears in Theorem 2. This completes the proof. $\qquad\square$

## 5   The Higher Dimensional Result

For a Borel set $A \subset \mathbb{R}^d$, denote by $|A|$ its Lebesgue measure. Let $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ and $h > 0$ and denote by $Q(x, h)$ the cube centred at $x$ and with side length $l(Q) = h$ and edges parallel to the axis. For a signed Borel measure $\mu$ on $\mathbb{R}^d$, we will treat its densities on cubes as *first divided differences,* and denote them by

$$\Delta_1 \mu(x, h) = \frac{\mu(Q(x, h))}{|Q(x, h)|}, \quad (x, h) \in \mathbb{R}^{d+1}_+,$$

and we also define its *second divided differences on cubes* as

$$\Delta_2 \mu(x, h) = \Delta_1 \mu(x, h) - \Delta_1 \mu(x, 2h), \quad (x, h) \in \mathbb{R}^{d+1}_+.$$

Here, $\mathbb{R}^{d+1}_+$ denotes the upper halfspace, $\mathbb{R}^{d+1}_+ = \{(x, h) \colon x \in \mathbb{R}^d, h > 0\}$. We say that a signed Borel measure $\mu$ on $\mathbb{R}^d$ is a Zygmund measure, $\mu \in \Lambda_*$, if it satisfies

$$\|\mu\|_* = \sup_{(x, h) \in \mathbb{R}^{d+1}_+} |\Delta_2 \mu(x, h)| < \infty.$$

Note that there can be Zygmund measures that are singular with respect to the Lebesgue measure (see [Kah69] and [AAN99]). Recall that a real valued function $f$ on $\mathbb{R}^d$ is said to have *bounded mean oscillation in* $\mathbb{R}^d$, $f \in \mathrm{BMO}(\mathbb{R}^d)$, if

$$\|f\|_{\mathrm{BMO}} = \sup_Q \left( \frac{1}{|Q|} \int_Q |f(x) - f_Q|^2 \, dx \right)^{1/2} < \infty,$$

where $Q$ ranges over all finite cubes in $\mathbb{R}^d$ with edges parallel to the axis and $f_Q = \frac{1}{|Q|} \int_Q f(x) \, dx$. We will say that a signed Borel measure $\nu$ on $\mathbb{R}^d$ is an I(BMO) measure, $\nu \in \mathrm{I(BMO)}$, if it is absolutely continuous with respect to the Lebesgue measure and its Radon-Nikodym derivative is

$$d\nu(x) = b(x) dx$$

for some function $b \in \mathrm{BMO}(\mathbb{R}^d)$. Using a characterisation of $\mathrm{BMO}(\mathbb{R}^d)$ functions due to R. Strichartz (see [Str80]), one can see that such a measure $\nu$ satisfies

$$\sup_Q \left( \frac{1}{|Q|} \int_Q \int_0^{l(Q)} |\Delta_2 \nu(x, h)|^2 \frac{dh \, dx}{h} \right)^{1/2} < \infty. \tag{16}$$

Conversely, whenever $\nu$ satisfies equation (16), it is an absolutely continuous measure with Radon-Nikodym derivative in $\mathrm{BMO}(\mathbb{R}^d)$ (see [DN02]).

Here we state a version of Theorem 1 for Zygmund measures in $\mathbb{R}^d$. For a given Zygmund measure $\mu$ and $\varepsilon > 0$, consider the set

$$A(\mu, \varepsilon) = \left\{ (x, h) \in \mathbb{R}^{d+1}_+ \colon |\Delta_2 \mu(x, h)| > \varepsilon \right\}.$$

**Theorem 4.** Let $\mu$ be a compactly supported Zygmund measure on $\mathbb{R}^d$. For each $\varepsilon > 0$ consider

$$C(\mu, \varepsilon) = \sup_Q \frac{1}{|Q|} \int_Q \int_0^{l(Q)} \chi_{A(\mu, \varepsilon)}(x, h) \frac{dh\, dx}{h},$$

where $Q$ ranges over all finite cubes with edges parallel to the axis. Then

$$\mathrm{dist}(\mu, \mathrm{I(BMO)}) \simeq \inf\{\varepsilon > 0 \colon C(\mu, \varepsilon) < \infty\}.$$

The proof of this result follows the same lines as that of Theorem 1. Nonetheless, one has to adapt the auxiliary results used in showing that theorem. First, we state and show the technical estimate in $\mathbb{R}^d$ which is analogous to Lemma 1. For convenience, given $y \in \mathbb{R}^{d-1}$ and $h > 0$ we will denote by $q(y, h)$ the cube in $\mathbb{R}^{d-1}$ centred at $y$, with sidelength $l(Q) = h$ and edges parallel to the axis.

**Lemma 6.** *Let $\mu \in \Lambda_*$ and assume that $h' > h > 0$ and $|x - t| < h/2$. Then*

$$|\Delta_2 \mu(x, h) - \Delta_2 \mu(t, h')|$$
$$\leq C_d \|\mu\|_* \left( \frac{h' - h}{h} \left( 1 + \log\left( \frac{h}{h' - h} + 1 \right) \right) + \frac{|x - t|}{h} \log\left( \frac{h}{|x - t|} + 1 \right) \right).$$

*Here, the constant $C_d$ does only depend on the dimension $d$.*

*Proof.* The proof is split in two steps. First, we find an estimate for the case $h = h'$ and then another one for $x = t$. We start showing that, for $h > 0$, when $|x - t| < h/2$

$$|\Delta_2 \mu(x, h) - \Delta_2 \mu(t, h)| \leq C_d \|\mu\|_* \frac{|x - t|}{h} \log\left( \frac{h}{|x - t|} + 1 \right). \tag{17}$$

First, if $|x - t| > h/2$, then

$$|\Delta_1 \mu(x, h) - \Delta_1 \mu(t, h)| \lesssim \|\mu\|_* \log\left( \frac{|x - t|}{h} + 1 \right). \tag{18}$$

The argument to show this bound is the same as in Lemma 1. The only difference is that one has to consider $u$ to be the harmonic extension of $\mu$ on the upper half-space $\mathbb{R}_+^{d+1}$, which will be itself a Bloch function, and use the well known fact (see Chapter V of [Ste70] or [Llo02]) that

$$|\Delta_1 \mu(x, h) - u(x, h)| \lesssim \|\mu\|_* .$$

To show (17), assume without loss of generality that $x = (x_1, \ldots, x_{d-1}, x_d)$ and $t = (x_1, \ldots, x_{d-1}, x'_d)$ with $x'_d < x_d$, and $|x - t| < h/2$. If we denote $y = (x_1, \ldots, x_{d-1}) \in \mathbb{R}^{d-1}$, one can see that

$$|Q(x, h)| \left( \Delta_2 \mu(x, h) - \Delta_2 \mu(t, h) \right) = \mu(l_+) - \mu(l_-) + \frac{\mu(L_-)}{2^d} - \frac{\mu(L_+)}{2^d},$$

16

where $l_+ = q(y, h) \times [x'_d + h/2, x_d + h/2)$, $l_- = q(y, h) \times [x'_d - h/2, x_d - h/2)$, $L_+ = q(y, 2h) \times [x'_d + h, x_d + h)$ and $L_- = q(y, 2h) \times [x'_d - h, x_d - h)$ are parallelepipeds at opposite sides of the cubes $Q(x, h)$ and $Q(x, 2h)$ respectively. We just show how to estimate $|\mu(l_+) - \mu(l_-)|$, as the other terms are estimated in the same way. The idea here is to cover $l_+$ with cubes $\{P_j\}$ and to use a translated cover $\{R_j\}$ for $l_-$.
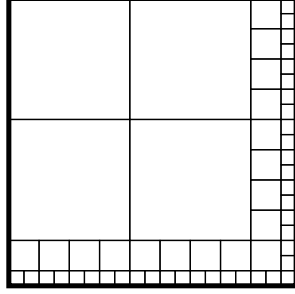


Figure 1: Parallelepiped $l_+$ seen from its base (bold square), and the distribution of the cubes $P_j^m$. Cubes of the same size belong to the same generation.

In order to cover $l_+$ with the appropriate cubes, write first $\frac{h}{|x-t|} = \sum_{n \geq 0} k_n 2^{-n}$, where $k_0 \geq 2$ and $k_n$ is 0 or 1 for $n \geq 1$ (as in a binary expansion). We construct a generation 0 placing $k_0^{d-1}$ cubes with mutually disjoint interiors of side length $|x - t|$ at one of the corners of $l_-$, forming alltogether a smaller parallelepiped with one side of length $|x - t|$ and the rest of length $k_0 |x - t|$. Let us denote by $\{P_0^m\}$ the set of cubes of generation 0. Assume we have constructed cubes up to generation $j - 1$, that is, we have chosen $\{P_i^m\}_{i=0}^{j-1}$. At generation $j$ either we do nothing if $k_j = 0$ or, when $k_j = 1$, we add a layer of cubes $\{P_j^m\}$ of side length $2^{-j}|x - t|$, such that $\{P_i^m\}_{i=0}^{j}$ have pairwise disjoint interiors, in order to get a new square based parallelepiped with one side length $|x-t|$ and the rest of $(\sum_{n=0}^{j} k_n 2^{-n})|x-t|$ (see Figure 1). Let $\{P_j^m\}$ be the cubes of generation $j$ and note that their total volume is

$$\sum_m |P_j^m| = |x - t|^d \left[ \left( \sum_{n=0}^{j} k_n 2^{-n} \right)^{d-1} - \left( \sum_{n=0}^{j-1} k_n 2^{-n} \right)^{d-1} \right].$$

If $d \geq 2$ we deduce

$$\sum_m |P_j^m| \lesssim d|x - t|^d k_j 2^{-j} \left( \sum_{n=0}^{j} k_n 2^{-n} \right)^{d-2}.$$

17

Since $\sum_{n=0}^{j} k_n 2^{-n} \le h/|x-t|$, we deduce that

$$\sum_m |P_j^m| \lesssim d|x-t|^2 h^{d-2} k_j 2^{-j} \tag{19}$$

Since the distance between the centres of $P_j^m$ and $R_j^m$ is bounded by a fixed multiple of $h$, applying equation (18) we get that

$$|\mu(P_j^m) - \mu(R_j^m)| \lesssim \|\mu\|_* |P_j^m| \log\left(\frac{h}{l(P_j^m)} + 1\right), \tag{20}$$

and using (20), we have

$$|\mu(l_+) - \mu(l_-)| \le \sum_j \sum_m |\mu(P_j^m) - \mu(R_j^m)|$$

$$\le C \|\mu\|_* \sum_j \sum_m |P_j^m| \log\left(\frac{h}{2^{-j}|x-t|} + 1\right).$$

Summing over $m$ and using (19), this is bounded by

$$C \|\mu\|_* d|x-t|^2 h^{d-2} \log\left(\frac{h}{|x-t|} + 1\right) \sum_j k_j 2^{-j} j,$$

and we deduce that

$$|\mu(l_+) - \mu(l_-)| \le Cd \|\mu\|_* |x-t| h^{d-1} \log\left(\frac{h}{|x-t|} + 1\right). \tag{21}$$

This and the analog estimate for $|\mu(L_+) - \mu(L_-)|$ yield estimate (17).

The second step is to show that, if $h' > h > 0$, then

$$|\Delta_2\mu(x,h') - \Delta_2\mu(x,h)| \le C_d \|\mu\|_* \frac{h'-h}{h} \left(1 + \log\left(\frac{h}{h'-h} + 1\right)\right). \tag{22}$$

Let $R(x,h,h') = q(y,h) \times [x_d - h'/2, x_d + h'/2)$, where $y \in \mathbb{R}^{d-1}$ is such that $x = (y, x_d)$. Note that $R(x,h,h')$ is the parallelepiped obtained from dilating the cube $Q(x,h)$ just in one direction. Denote as well

$$\Delta_2\mu(x,h,h') = \frac{\mu(R(x,h,h'))}{|R(x,h,h')|} - \frac{\mu(R(x,2h,2h'))}{|R(x,2h,2h')|}.$$

To show (22), it is enough to see that

$$|\Delta_2\mu(x,h,h') - \Delta_2\mu(x,h)| \le C_d \|\mu\|_* \frac{h'-h}{h} \left(1 + \log\left(\frac{h}{h'-h} + 1\right)\right). \tag{23}$$

Let us denote $Q = Q(x,h)$, $\tilde{Q} = Q(x,2h)$, $R = R(x,h,h')$ and $\tilde{R} = R(x,2h,2h')$. Note that we can decompose $R$ as the disjoint union $Q \cup l_+ \cup l_-$,
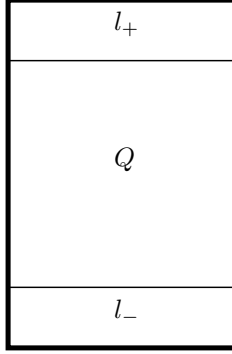
18

Figure 2: The parallelepiped $R$ can be decomposed into the cube $Q$ and the square based parallelepipeds $l_+$ and $l_-$.

where $l_+$ and $l_-$ are parallelepipeds similar to the ones we used before (see Figure 2). In the same way, decompose $\tilde{R} = \tilde{Q} \cup L_+ \cup L_-$, and note that $L_+$ (and also $L_-$) can be regarded as the union $\bigcup_{i=1}^{2^d} L_+^i$, where each $L_+^i$ is a translation of $l_+$. Now, express

$$\Delta_2\mu(x,h) - \Delta_2\mu(x,h,h') = \frac{\mu(Q)}{|Q|} - \frac{\mu(\tilde{Q})}{|\tilde{Q}|} - \frac{\mu(R)}{|R|} + \frac{\mu(\tilde{R})}{|\tilde{R}|} =$$

$$\frac{h'-h}{h'}\left(\frac{\mu(Q)}{|Q|} - \frac{\mu(\tilde{Q})}{|\tilde{Q}|}\right) - \left(\frac{\mu(l_+)}{h^{d-1}h'} - \frac{\mu(L_+)}{2^d h^{d-1}h'}\right) - \left(\frac{\mu(l_-)}{h^{d-1}h'} - \frac{\mu(L_-)}{2^d h^{d-1}h'}\right).$$

The first term is $\Delta_2\mu(x,h)(h'-h)/h'$, which is bounded by $\|\mu\|_* (h'-h)/h$. We will now show that

$$\left|\frac{\mu(l_+)}{h^{d-1}h'} - \frac{\mu(L_+)}{2^d h^{d-1}h'}\right| \leq C_d \|\mu\|_* \frac{h'-h}{h}\left(1 + \log\left(\frac{h}{h'-h} + 1\right)\right) \qquad (24)$$

The last term is estimated in a similar way. First, we use the decomposition of $L_+$ to split the difference as follows

$$\left|\frac{\mu(l_+)}{h^{d-1}h'} - \frac{\mu(L_+)}{2^d h^{d-1}h'}\right| \leq \frac{1}{2^d h^{d-1}h'} \sum_{i=1}^{2^d} |\mu(l_+) - \mu(L_+^i)|.$$

For each term in this sum, we can use the estimate in (21) for parallelepipeds, just taking into account that now the role of $h$ is taken by $Ch$ and $h'-h$ plays the role of $|x-t|$. This gives (24), which yields (23) and finishes the proof. $\qquad\square$

We also need a dyadic version of Theorem 4. We say that $Q$ is a *dyadic cube* in $\mathbb{R}^d$ if it is of the form $[k_1 2^{-n}, (k_1+1)2^{-n}) \times \ldots \times [k_d 2^{-n}, (k_d+1)2^{-n})$

where $k_1, \dots, k_d \in \mathbb{Z}$ and $n \geq 0$, or if it is of the form $[k_1 2^{-n} - t_n, (k_1 + 1)2^{-n} - t_n) \times \dots \times [k_d 2^{-n} - t_n, (k_d + 1)2^{-n} - t_n)$ where $k_1, \dots, k_d \in \mathbb{Z}$, $n < 0$ and where $t_n$ is the quantity defined in Section 3. We denote here the set of dyadic cubes in $\mathbb{R}^d$ by $\mathcal{D}$ and the set of dyadic cubes of side length $2^{-n}$ by $\mathcal{D}_n$. As we did before, if $Q_0$ is a given arbitrary cube, we may refer to the set of dyadic cubes contained in $Q_0$ by $\mathcal{D}(Q_0)$. For future convenience, given a signed Borel measure $\mu$ on $\mathbb{R}^d$, we define the *dyadic second divided difference* as

$$\Delta_2^d \mu(Q) = \Delta_1 \mu(Q) - \Delta_1 \mu(Q^*), \quad Q \in \mathcal{D},$$

where we used $Q^*$ to denote the unique dyadic cube that contains $Q$ and is such that $l(Q^*) = 2l(Q)$. We will also need the *maximal dyadic second divided difference,* defined by

$$\Delta_2^* \mu(Q) = \max_{Q'} |\Delta_1 \mu(Q') - \Delta_1 \mu(Q)|, \quad Q \in \mathcal{D},$$

where $Q'$ ranges over all dyadic cubes contained in $Q$ such that $l(Q') = l(Q)/2$. A signed Borel measure $\mu$ on $\mathbb{R}^d$ is called a *dyadic Zygmund measure,* $\mu \in \Lambda_{*d}$, if

$$\|\mu\|_{*d} = \sup_{Q \in \mathcal{D}} \Delta_2^* \mu(Q) < \infty.$$

A real valued function $f$ on $\mathbb{R}^d$ is said to have *bounded dyadic mean oscillation* on $\mathbb{R}^d$, $f \in \mathrm{BMO}_d(\mathbb{R}^d)$, if

$$\|f\|_{\mathrm{BMO}\, d} = \sup_{Q \in \mathcal{D}} \left( \frac{1}{|Q|} \int_Q |f(x) - f_Q|^2 \, dx \right)^{1/2} < \infty.$$

We will say that a signed Borel measure $\nu$ on $\mathbb{R}^d$ is a *dyadic* I(BMO) *measure,* $\nu \in \mathrm{I(BMO)}_d$, if it is absolutely continuous and its derivative is

$$d\nu = b(x) \, dx,$$

where $b \in \mathrm{BMO}_d(\mathbb{R}^d)$. It can be checked that $\nu$ is such a measure if and only if it satisfies

$$\sup_{Q \in \mathcal{D}} \left( \frac{1}{|Q|} \sum_{R \in \mathcal{D}(Q)} |\Delta_2^d \nu(R)|^2 |R| \right)^{1/2} < \infty.$$

The analog of Theorem 4 for these dyadic spaces is the following.

**Theorem 6.** *Let $\mu$ be a compactly supported measure in $\Lambda_{*d}$. For each $\varepsilon > 0$ consider*

$$D(\mu, \varepsilon) = \sup_{Q \in \mathcal{D}} \frac{1}{|Q|} \sum_{\substack{R \in \mathcal{D}(Q) \\ \Delta_2^* \mu(R^*) > \varepsilon}} |R|.$$

*Then,*

$$\mathrm{dist}(\mu, \mathrm{I(BMO)}_d) = \inf\{\varepsilon > 0 \colon D(\mu, \varepsilon) < \infty\}. \tag{25}$$

20

Note that, as we did for functions on $\mathbb{R}$, we can rewrite this result in terms of dyadic martingales on $\mathbb{R}^d$. We define a *dyadic martingale on $\mathbb{R}^d$* as a sequence of functions $S = \{S_n\}_{n=0}^{\infty}$ such that $S_n$ is constant on any cube $Q \in \mathcal{D}_n$ and such that

$$S_n | Q = \frac{1}{2^d} \sum_{\substack{Q' \in \mathcal{D}_{n+1} \\ Q' \subset Q}} S_{n+1} | Q',$$

for all $Q \in \mathcal{D}_n$, $n \geq 0$. Given a measure $\mu \in \Lambda_*$, we can define a dyadic martingale by taking

$$S_n(Q) = \Delta_1 \mu(Q), \quad Q \in \mathcal{D}_n, n \geq 0, \tag{26}$$

and then $\Delta S(Q) = S_n(Q) - S_{n-1}(Q^*) = \Delta_2^d \mu(Q)$, for $Q \in \mathcal{D}_n$, and we can rewrite Theorem 6 in terms of martingales. Following this relation between dyadic second divided differences for measures and martingale jumps, we will denote $\Delta^* S(Q) = \Delta_2^* \mu(Q)$.

*Proof of Theorem 6.* Assume that $\mu$ is supported on the unit cube $Q_0 = [0,1]^d$. We need to prove that, for a given $\varepsilon > 0$, there is a measure $\nu \in \mathrm{I(BMO)}_d$ satisfying $\|\mu - \nu\|_{*d} \leq \varepsilon$ if and only if $D(\mu, \varepsilon) < \infty$. Denote by $\varepsilon_0$ the infimum in the left-hand side of (25).

Given $\varepsilon > \varepsilon_0$, consider the martingale $S$ defined by (26). Approximate the martingale $S$ by another dyadic martingale $B$ in the following way. Start taking $B(Q_0) = S(Q_0)$. Then, for $Q \in \mathcal{D}(Q_0)$, set $\Delta B(Q) = \Delta S(Q)$ whenever $\Delta^* S(Q^*) > \varepsilon$, and set $\Delta B(Q) = 0$ otherwise. By construction, it is clear that $|\Delta S(Q) - \Delta B(Q)| \leq \varepsilon$ for any dyadic cube $Q$. Moreover, for any such cube $Q$, we have that

$$\frac{1}{|Q|} \sum_{R \in \mathcal{D}(Q)} |\Delta B(R)|^2 |R| = \frac{1}{|Q|} \sum_{\substack{R \in \mathcal{D}(Q) \\ \Delta^* S(R^*) > \varepsilon}} |\Delta S(R)|^2 |R| \lesssim \|\mu\|_* D(\mu, \varepsilon). \tag{27}$$

Define now $b(x) = \lim_n B_n(x) = \sum_{n=1}^{\infty} \Delta B_n(x)$. Using that, for any dyadic martingale, the increments $\Delta B_j$ are $L^2$ orthogonal, we get that

$$\int_{Q_0} b(x)^2 \, dx = \int_{Q_0} \sum_{n=1}^{\infty} |\Delta B_n(x)|^2 \, dx = \sum_{R \in \mathcal{D}(Q_0)} |\Delta B(R)|^2 |R| < \infty,$$

so that $b \in L^2$ and it is finite almost everywhere. Hence, the measure $\nu$ defined by

$$d\nu = b(x) \, dx,$$

is an absolutely continuous measure that, by (27), is an $\mathrm{I(BMO)}_d$ measure such that $\|\mu - \nu\|_{*d} \leq \varepsilon$.

On the other hand, if $\varepsilon < \varepsilon_0$, there exists no measure $\nu \in I(BMO)_d$ satisfying $\|\mu - \nu\|_{*d} \leq \varepsilon$. Indeed, take $\varepsilon < \varepsilon_1 < \varepsilon_0$ and assume that there is $\nu \in I(BMO)_d$ such that $\|\mu - \nu\|_{*d} \leq \varepsilon$. Then, for any $Q \in \mathcal{D}$ such that $\Delta_2^* \mu(Q^*) > \varepsilon_1$, we have that $\Delta_2^* \nu(Q^*) > \varepsilon_1 - \varepsilon = \delta > 0$. Thus

$$\frac{1}{|Q|} \sum_{R \in \mathcal{D}(Q)} |\Delta_2^d \nu(R)|^2 |R| \geq \frac{\delta^2}{|Q|} \sum_{\substack{R \in \mathcal{D}(Q) \\ \Delta_2^* \mu(R^*) > \varepsilon_1}} |R|,$$

but the supremum over $Q \in \mathcal{D}$ of this last quantity is $\delta^2 D(\mu, \varepsilon_1)$, which is infinite since $\varepsilon_1 < \varepsilon$. This contradicts that $\nu \in I(BMO)_d$. $\qquad\square$

The proof of Theorem 4 follows the same lines than the proof of Theorem 1. We just mention that the construction used to prove Theorem 2 is easily adapted to the setting of $\mathbb{R}^d$, except for the following detail. Let $Q$ be a cube in $\mathbb{R}^d$ and consider the covering $\mathcal{F}(Q) = \{R_j\}$ of $Q$ by maximal dyadic cubes, in the same sense as we did in $\mathbb{R}$. In the case $d = 1$ we could have at most two elements of the same size in $\mathcal{F}(Q)$, but this does not hold for $d \geq 2$. For $d \geq 2$, the amount of cubes $R_j$ in $\mathcal{F}(Q)$ of size $|R_j| = 2^{-kd}|Q|$, for some $k \geq 1$, is of the order of $2^{k(d-1)}$. Using this bound, one sees that the sums appearing in the estimates in the proof of Theorem 2 are convergent and bounded by a universal constant.

## 6   An Application to Sobolev Spaces

Fix $1 < p < \infty$. Consider the Sobolev space $W^{1,p}$ of functions $f \in L^p$ whose derivative $f'$ in the sense of distributions is also in $L^p$. Consider as well, in the Zygmund class, the subspace $\Lambda_*^p = W^{1,p} \cap \Lambda_*$. For $x \in \mathbb{R}$, consider the truncated cone $\Gamma(x) = \{(t, h) \in \mathbb{R}_+^2 : |x - t| < h < 1\}$. In [Nic18] it is shown that a function $f \in L^p$ is in the Sobolev space $W^{1,p}$ if and only if $C(f) \in L^p$, where

$$C(f)(x) = \left( \int_{\Gamma(x)} |\Delta_2 f(s,t)|^2 \frac{ds\,dt}{t^2} \right)^{1/2}, \quad x \in \mathbb{R}.$$

The purpose of this section is to prove Theorem 3. Following the same scheme as before, we first need a dyadic version of the previous theorem. Let us first recall some more concepts and standard results of Martingale Theory that will be useful later. The *quadratic characteristic* of a dyadic martingale $S$ is the function

$$\langle S \rangle(x) = \left( \sum_{n=1}^{\infty} |\Delta S_n(x)|^2 \right)^{1/2}, \quad x \in \mathbb{R},$$

and its *maximal function* is

$$S^*(x) = \sup_n |S_n(x) - S_0(x)|, \quad x \in \mathbb{R}.$$

Given $0 < p < \infty$ and a dyadic martingale $S$, the Burkholder-Davis-Gundy Inequality (see [BM99]) states that there exists a constant $C = C(p) > 0$ such that

$$C^{-1} \|\langle S \rangle\|_{L^p} \leq \|S^*\|_{L^p} \leq C \|\langle S \rangle\|_{L^p}. \tag{28}$$

Remember as well that the *Fatou set* of a dyadic martingale $S$, denoted by $F(S)$, is defined as

$$F(S) = \{x \in \mathbb{R} \colon \lim_n S_n(x) \text{ exists and is finite}\}.$$

It is a standard result of Martingale Theory that, for a dyadic martingale $S$ such that $\|S\|_* < \infty$, its Fatou set is $F(S) = \{x \in \mathbb{R} \colon \langle S \rangle(x) < \infty\}$, where the equality must be understood up to sets of zero measure (see [Llo02]).

Using the characterisation for the Sobolev space $W^{1,p}$ previously stated, we say that a function $b$ is in the dyadic space $\Lambda_{*d}^p$ if its average growth martingale $B$, as defined in (9), has quadratic characteristic $\langle B \rangle \in L^p$ and

$$\|B\|_* = \sup_{I \in \mathcal{D}} |\Delta B(I)| < \infty.$$

Note that, in fact, $\Lambda_{*d}^p = W^{1,p} \cap \Lambda_{*d}$. Indeed, if $b \in \Lambda_{*d}^p$, by definition $b \in \Lambda_{*d}$. Moreover, since its average growth martingale $B$ has quadratic characteristic $\langle B \rangle \in L^p$, $\langle B \rangle(x) < \infty$ for almost every $x \in \mathbb{R}$. Thus, $B(x) = \lim_n B_n(x)$ exists almost everywhere and will satisfy $b'(x) = B(x)$ in the sense of distributions. Using (28), $B^* \in L^p$ and, thus, $B \in L^p$ as well, which is the same to say that $b' \in L^p$. We now state the analogous of Theorem 3 in this context.

**Theorem 7.** *Let $f$ be a compactly supported function in $\Lambda_{*d}$ and fix $1 < p < \infty$. Let $S$ be the average growth martingale of $f$. For every $\varepsilon > 0$, define the truncated quadratic characteristic*

$$D(f, \varepsilon)(x) = (\#\{n \colon |\Delta S_n(x)| > \varepsilon\})^{1/2}.$$

*Then,*

$$\text{dist}(f, \Lambda_{*d}^p) = \inf\{\varepsilon > 0 \colon D(f, \varepsilon) \in L^p\}. \tag{29}$$

*Proof.* Let $\varepsilon_0$ be the infimum on (29). Assume $0 < \varepsilon < \varepsilon_1 < \varepsilon_0$ and that there is $b \in \Lambda_{*d}^p$ such that $\|f - b\|_{*d} \leq \varepsilon$. Let $B$ be the average growth martingale of function $b$. Whenever $|\Delta S_n(x)| > \varepsilon_1$, we have that $|\Delta B_n(x)| > \varepsilon_1 - \varepsilon = \delta > 0$. Thus,

$$\langle B \rangle^2(x) = \sum_{n=1}^{\infty} |\Delta B_n(x)|^2 \geq \sum_{|\Delta B_n(x)| > \delta} |\Delta B_n(x)|^2$$

$$\geq \frac{\delta^2}{\|f\|_{*d}^2} D^2(f, \varepsilon_1)$$

23

for all $x \in \mathbb{R}$. But, since $\varepsilon_1 < \varepsilon_0$, $D(f, \varepsilon_1) \notin L^p$ and so $\langle B \rangle \notin L^p$, getting in this way a contradiction. Hence, we see that $\operatorname{dist}(f, \Lambda_{*d}^p) \geq \varepsilon_0$.

Assume that $f$ is supported on $I_0$. Consider now $\varepsilon > \varepsilon_0$. Construct a dyadic martingale $B$ with $B(I_0) = S(I_0)$ and such that $\Delta B(I) = \Delta S(I)$ for all $I \in \mathcal{D}(I_0)$ whenever $|\Delta S(I)| > \varepsilon$, but take $\Delta B(I) = 0$ when $|\Delta S(I)| \leq \varepsilon$. Note that $\langle B \rangle \in L^p$. Therefore, using (28), we see that we can define $b'(x) = \lim_n B_n(x)$ almost everywhere with $b' \in L^p$. Taking now $b(x) = \int_0^x b'(s)\,ds$, we get $b \in \Lambda_{*d}^p$ such that $\|f - b\|_{*d} \leq \varepsilon$. This shows that $\operatorname{dist}(f, \Lambda_{*d}^p) \leq \varepsilon_0$, completing the proof. □

*Proof of Theorem 3.* Let $\varepsilon_0$ be the infimum in (4). Assume $0 < \varepsilon < \varepsilon_1 < \varepsilon_0$, take $\delta = \varepsilon_1 - \varepsilon$, and assume that there is $b \in \Lambda_*^p$ such that $\|f - b\|_* \leq \varepsilon$. The same argument used in the first part of the proof of Theorem 7 allows us to see that

$$C(b)(x) \geq \delta C(b, \delta)(x) \geq \delta C(f, \varepsilon_1)(x)$$

for $x \in \mathbb{R}$. Since $\varepsilon_1 < \varepsilon_0$, we have that $C(f, \varepsilon_1) \notin L^p$ and, thus, $C(b) \notin L^p$, contradicting that $b \in \Lambda_*^p$. Hence, $\operatorname{dist}(f, \Lambda_*^p) \geq \varepsilon_0$.

Fix $\varepsilon > \varepsilon_0$ so that $C(f, \varepsilon) \in L^p$. For $\alpha \in [-1, 1]$, consider $f^{(\alpha)} = f(x + \alpha)$. Note that $C(f^{(\alpha)}, \varepsilon) \in L^p$ as well. Using the same argument as in the proof of Theorem 1, one can see that this fact implies that $D(f^{(\alpha)}, \varepsilon) \in L^p$. Thus, for each $\alpha \in [-1, 1]$, the function $f^{(\alpha)}$ satisfies the hypothesis of Theorem 7 and may be approximated as $f^{(\alpha)} = b^{(\alpha)} + t^{(\alpha)}$, where $b^{(\alpha)} \in \Lambda_{*d}^p$ with $\|b^{(\alpha)}\|_{*d} \leq \|f\|_*$ and $\|t\|_{*d} \leq \varepsilon$. Apply now Theorem 2, with $R = 1$, both with the mapping $\alpha \mapsto b^{(\alpha)}$ and $\alpha \mapsto t^{(\alpha)}$ to obtain respectively functions $b$ and $t$ such that $f = b + t$ and such that $b \in \Lambda_*^p$ and $\|t\|_* \lesssim \varepsilon$. This completes the proof. □

# 7  Open Problems

This last section is devoted to state three open problems closely related to our results.

**1.** Observe that Theorem 4 is a generalisation of Theorem 1 for measures on $\mathbb{R}^d$ that works for any $d \geq 1$. Nonetheless, we have not been able to generalise Theorem 1 for functions on the Zygmund class on $\mathbb{R}^d$ for $d > 1$. We say that a continuous function $f : \mathbb{R}^d \to \mathbb{R}$ is in the Zygmund class $\Lambda_*(\mathbb{R}^d)$ if

$$\|f\|_* = \sup_{\substack{x, h \in \mathbb{R}^d \\ h \neq 0}} \frac{|f(x + h) - 2f(x) + f(x - h)|}{\|h\|} < +\infty.$$

A continuous function on $\mathbb{R}^d$ is in $\mathrm{I(BMO)}(\mathbb{R}^d)$ if its partial derivatives in the sense of distributions are functions in BMO. It is easy to see that $\mathrm{I(BMO)}(\mathbb{R}^d) \subset \Lambda_*(\mathbb{R}^d)$. We do not know an analog of Theorem 1 when

24

$d > 1$. Roughly speaking, the method we develop in this paper works discretising a function in the Zygmund class and modifying its average growth on certain intervals, so that we end up constructing the derivative of a function in I(BMO). Nonetheless, when applying this method in $d > 1$ variables, one ends up constructing $d$ functions that approximate the divided differences of a given function in the coordinate directions. However, this system of $d$ functions is not, in general, the gradient of an I(BMO) function.

**2.** Another related open problem is to find the closure of the space of Lipschitz functions, Lip, in the Zygmund class. It is a well known fact that singular integral operators such as the Hilbert transform are bounded in $L^p$ for $1 < p < \infty$, but they are not in $L^\infty$. Nonetheless, they are bounded from $L^\infty$ to BMO and from BMO to itself. In a similar fashion, these operators are bounded on the Hölder classes $\text{Lip}_\alpha$, for $0 < \alpha < 1$, but they are not in the Lipschitz class. However, they are actually bounded on the Zygmund class, which plays a similar role to that of BMO in the previous setting. In this sense, this problem would be related to the one solved by Garnett and Jones in [GJ78], where they find a characterisation for the closure of $L^\infty$ in the space BMO.

**3.** The open problem mentioned in the previous paragraph is analogous to the well known open problem of describing the closure of the space $\mathbb{H}^\infty$ of bounded analytic functions in the unit disk into the Bloch space $\mathcal{B}$, consisting of analytic functions $f$ in the unit disk such that $\sup_{z \in \mathbb{D}}(1 - |z|^2)|f'(z)| < \infty$ (see [ACP74]). In [GZ93] the closure of the space BMOA in $\mathcal{B}$ is described, while in [MN11] and [GMP15] the closure of the class $\mathbb{H}^p \cap \mathcal{B}$ is studied. Here $\mathbb{H}^p$ denotes the classical Hardy spaces of analytic functions in the unit disk. It is worth mentioning that in this setting, the proofs rely on reproducing formulae that analytic functions fulfill, which are not available in our real variable situation.

# References

[AAN99]   A. B. Aleksandrov, J. M. Anderson, and A. Nicolau. "Inner functions, Bloch spaces and symmetric measures". *Proc. London Math. Soc. (3)* 79.2 (1999), pp. 318–352.

[ACP74]   J. M. Anderson, J. Clunie, and C. Pommerenke. "On Bloch functions and normal functions". *J. Reine Angew. Math.* 270 (1974), pp. 12–37.

[AP89]   J. M. Anderson and L. D. Pitt. "Probabilistic behaviour of functions in the Zygmund spaces $\Lambda^*$ and $\lambda^*$". *Proc. London Math. Soc. (3)* 59.3 (1989), pp. 558–592.

[BM99]     R. Bañuelos and C. N. Moore. *Probabilistic behavior of harmonic functions*. Vol. 175. Progress in Mathematics. Birkhäuser Verlag, Basel, 1999, pp. xiv+204.

[Con13]    J. M. Conde. "A note on dyadic coverings and nondoubling Calderón-Zygmund theory". *J. Math. Anal. Appl.* 397.2 (2013), pp. 785–790.

[DLN14]    J. J. Donaire, J. G. Llorente, and A. Nicolau. "Differentiability of functions in the Zygmund class". *Proc. Lond. Math. Soc. (3)* 108.1 (2014), pp. 133–158.

[DN02]     E. Doubtsov and A. Nicolau. "Symmetric and Zygmund measures in several variables". *Ann. Inst. Fourier (Grenoble)* 52.1 (2002), pp. 153–177.

[GMP15]    P. Galanopoulos, N. Monreal Galán, and J. Pau. "Closure of Hardy spaces in the Bloch space". *J. Math. Anal. Appl.* 429.2 (2015), pp. 1214–1221.

[GJ78]     J. B. Garnett and P. W. Jones. "The distance in BMO to $L^\infty$". *Ann. of Math. (2)* 108.2 (1978), pp. 373–393.

[GJ82]     J. B. Garnett and P. W. Jones. "BMO from dyadic BMO". *Pacific J. Math.* 99.2 (1982), pp. 351–371.

[GZ93]     P. G. Ghatage and D. C. Zheng. "Analytic functions of bounded mean oscillation and the Bloch space". *Integral Equations Operator Theory* 17.4 (1993), pp. 501–515.

[JW84]     A. Jonsson and H. Wallin. "Function spaces on subsets of $\mathbf{R}^n$". *Math. Rep.* 2.1 (1984), pp. xiv+221.

[Kah69]    J.-P. Kahane. "Trois notes sur les ensembles parfaits linéaires". *Enseignement Math. (2)* 15 (1969), pp. 185–192.

[Llo02]    J. G. Llorente. *Discrete martingales and applications to analysis*. Vol. 87. Report. University of Jyväskylä Department of Mathematics and Statistics. University of Jyväskylä, Jyväskylä, 2002, pp. ii+40.

[Mak89]    N. G. Makarov. "Smooth measures and the law of the iterated logarithm". *Izv. Akad. Nauk SSSR Ser. Mat.* 53.2 (1989), pp. 439–446.

[Mei03]    T. Mei. "BMO is the intersection of two translates of dyadic BMO". *C. R. Math. Acad. Sci. Paris* 336.12 (2003), pp. 1003–1006.

[MN11]     N. Monreal Galán and A. Nicolau. "The closure of the Hardy space in the Bloch norm". *St. Petersburg Math. J.* 22.1 (2011), pp. 75–81.

[Nic18]    A. Nicolau. "Divided differences, square functions and a law of the iterated logarithm". *Real Anal. Exchange* 43.1 (2018), pp. 155–186.

[Ste70]    E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970, pp. xiv+290.

[Str80]    R. S. Strichartz. "Bounded mean oscillation and Sobolev spaces". *Indiana Univ. Math. J.* 29.4 (1980), pp. 539–558.

[Zyg45]    A. Zygmund. "Smooth functions". *Duke Math. J.* 12 (1945), pp. 47–76.

Artur Nicolau: Universitat Autònoma De Barcelona, Departament de Matemàtiques, Edifici C, 08193-Bellaterra, Catalunya
    *E-mail address*: `artur@mat.uab.cat`

Odí Soler i Gibert: Universitat Autònoma De Barcelona, Departament de Matemàtiques, Edifici C, 08193-Bellaterra, Catalunya
    *E-mail address*: `odisoler@mat.uab.cat`