CONSTRUCTIVE
APPROXIMATION

Path Regularity of the Brownian Motion and the Brownian Sheet

H. Kempka[1] · C. Schneider[2] · J. Vybiral[3]

© The Author(s) 2023

Abstract

By the work of P. Lévy, the sample paths of the Brownian motion are known to satisfy a certain Hölder regularity condition almost surely. This was later improved by Ciesielski, who studied the regularity of these paths in Besov and Besov-Orlicz spaces. We review these results and propose new function spaces of Besov type, strictly smaller than those of Ciesielski and Lévy, in which the sample paths of the Brownian motion almost surely lie. In the same spirit, we review and extend the work of Kamont, who investigated the same question for the multivariate Brownian sheet and function spaces of dominating mixed smoothness.

Keywords Brownian motion · Brownian sheet · Besov spaces · Faber bases · Path regularity · Dominating mixed smoothness

---

✉ C. Schneider
cornelia.schneider@math.fau.de

H. Kempka
henning.kempka@eah-jena.de

J. Vybiral
jan.vybiral@fjfi.cvut.cz

[1] Ernst-Abbe-Hochschule Jena, University of Applied Sciences, FB Grundlagenwissenschaften, Carl-Zeiss-Promenade 2, 07745 Jena, Germany

[2] Applied Mathematics III, Friedrich-Alexander University Erlangen-Nuremberg, Cauerstr. 11, 91058 Erlangen, Germany

[3] Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Trojanova 13, 12000 Praha, Czech Republic

Mathematics Subject Classification 60G17 · 60G15 · 46E35 · 60G60

Contents

1 Introduction

Already in 1937, Paul Lévy showed [36, Section 52], that the sample paths of the
Wiener process $W = (W_t)_{t \geq 0}$ satisfy almost surely the Hölder condition

$$|W_{t'} - W_t| \leq c \cdot \sqrt{2|t' - t| \log\left(\frac{1}{|t' - t|}\right)} \tag{1}$$

for every $c > 1$ and $|t' - t|$ small enough. In general, one can define for a positive
function g on $[0, 1]$ the Hölder space $\mathcal{C}^g([0, 1])$ as the collection of all functions f on
$[0, 1]$ such that

$$|f(s) - f(t)| \leq c\, g(|s - t|) \quad \text{for all} \quad 0 \leq s, t \leq 1.$$

Then the result of Lévy shows that the paths of W almost surely lie in the Hölder
space $\mathcal{C}^g([0, 1])$ with $g(r) = |r \log r|^{1/2}$ for $r > 0$ small. Furthermore, this result is
known to be optimal and this space is the smallest one in the scale of Hölder spaces,
in which the paths of W almost surely lie in [11, 37]. This shows, in particular, that

the log-factor in (1) is necessary and that the smoothness regularity $1/2$, which is $\mathcal{C}^{1/2}([0, 1])$, i.e., the space above with $g(r) = |r|^{1/2}$, cannot be achieved in the scale of Hölder spaces.

Later on, Ciesielski proved in [8], that one can actually obtain smoothness of the order $1/2$ if one gives up slightly on the integrability. Namely, [8] shows that the paths of W lie almost surely in the Besov space $B_{p,\infty}^{1/2}([0, 1])$ for $1 \leq p < \infty$. The excluded endpoint space is again $B_{\infty,\infty}^{1/2}([0, 1]) = \mathcal{C}^{1/2}([0, 1])$. Shortly after, Ciesielski and his co-authors [9, 10] refined the analysis of [8] and discovered, that almost all paths of W lie in the Besov-Orlicz space $B_{\Phi_2,\infty}^{1/2}([0, 1])$, which combines the technique of Besov spaces together with the Orlicz space generated by the Orlicz function $\Phi_2(t) = \exp(t^2) - 1$. This space is (properly) included both in the Hölder space $\mathcal{C}^g([0, 1])$ discovered by Lévy as well as in the Besov spaces $B_{p,\infty}^{1/2}([0, 1])$ for $1 \leq p < \infty$. As such, $B_{\Phi_2,\infty}^{1/2}([0, 1])$ represents currently the smallest space in which the sample paths of the Wiener process are almost surely known to lie. On the other hand, its definition is certainly more involved than the Hölder condition of Lévy (1).

The results on the regularity of sample Wiener paths were later complemented, generalized, and applied in several different ways. The optimality of the result of Ciesielski in the scale of Besov spaces was studied in [47] and [3], where the latter reference studies the topic in the frame of modulation spaces and Wiener amalgam spaces. Path regularity of more general processes was investigated in [21, 50, 51] and we refer to [62] for results on the torus. The approach was also generalized to Wiener processes with values in Banach spaces in [23], and applied to regularity properties of stochastic differential equations [43, 44].

Let us also mention that many other properties of sample paths of the Wiener process, Brownian sheet, and other random processes were studied extensively. They include different dimensions of the graph set, small ball probabilities, hitting probabilities or the law of iterated logarithm. We refer in this context to [30, 31, 40, 61, 66, 67] and the references given therein.

The first aim of our work is to present an essentially self-contained proof of the results of Lévy and Ciesielski, which should be easily accessible to readers familiar with the theory of function spaces. This will be done by first deriving the decomposition of sample Wiener paths into a series of Faber splines with independent standard Gaussian random coefficients. Afterwards, the proof that almost all paths of the Wiener process lie in a Hölder space of Lévy or in the Besov or Besov-Orlicz spaces of Ciesielski, reduces to rather straightforward concentration inequalities for independent Gaussian variables. For this purpose we collect basic facts on Gaussian variables (and some other related random variables), that are needed throughout the manuscript, in Sect. 4.

Let us briefly summarize the main steps of this approach. It is essentially based on two very well-known properties of the Faber system. This is a system of shifted and dilated hat functions $v_{j,m}$, cf. (3), where $j \in \mathbb{N}_0$ and $0 \leq m \leq 2^j - 1$, which are concentrated on the dyadic intervals $I_{j,m} = [m \cdot 2^{-j}, (m + 1)2^{-j}]$. The first property of this system is described in detail in Theorem 3. It states, that if $\{\xi_{j,m} : j \in \mathbb{N}_0, 0 \leq m \leq 2^j - 1\}$ are independent standard Gaussian variables, then the series

$$\sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} 2^{-(j+2)/2} \xi_{j,m} v_{j,m}(t) \quad \text{for } t \in [0, 1]$$

converges almost surely uniformly on [0, 1] and its limit coincides with the Wiener process $W = (W_t)_{t \geq 0}$.

The second key property of the Faber system is that it can be used to describe several classical function spaces of Besov-type. To be more precise, a function f representable (in some sense) by the series

$$f = \sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} \mu_{j,m} 2^{-js} v_{j,m}$$

belongs to such a Besov-type space if, and only if, the sequence of coefficients $\{\mu_{j,m}\}_{j,m}$ satisfies some summability and/or integrability condition. Naturally, these conditions differ from one space to another, but usually they can be rewritten in the language of the step functions

$$f_j = \sum_{m=0}^{2^j-1} \mu_{j,m} \chi_{j,m}, \tag{2}$$

where $\chi_{j,m}$ is the characteristic function of $I_{j,m}$. For example, the proof that Wiener paths almost surely lie in the Besov-Orlicz space $B_{\Phi_2,\infty}^{1/2}([0, 1])$, reduces by this technique to the statement, that $\|f_j\|_{\Phi_2}$ is finite and uniformly bounded over $j \in \mathbb{N}_0$ if we replace the $\mu_{j,m}$'s in (2) by independent standard Gaussian variables $\xi_{j,m}$. We use this approach to re-prove the results of Lévy and Ciesielski and, in the named case of the Besov-Orlicz space $B_{\Phi_2,\infty}^{1/2}([0, 1])$, we provide an alternative proof, based on a characterization of the Orlicz space $L_{\Phi_2}([0, 1])$ in terms of non-increasing rearrangements.

The second aim of this paper is to show that this procedure can be stepped up and that (based on some knowledge about Gaussian variables) one can produce even smaller function spaces, which still contain the sample Wiener paths almost surely. The price to pay in this context is that the new spaces do not fall into any standard scale of function spaces. Let us again briefly sketch the main idea and the main results. First, we observe that if we use independent Gaussian variables as the coefficients in (2), then the Orlicz space $L_{\Phi_2}([0, 1])$ measures very effectively the size of the f_j's among the function spaces invariant with respect to the rearrangement of a function. But it does not take any effort to describe the position of large values of f_j. Indeed, if $\xi_j = (\xi_{j,0}, \ldots, \xi_{j,2^j-1})$ are independent Gaussian variables, then the maximum of $|\xi_{j,m}|$ over m is known to behave asymptotically like $\sqrt{j}$ with high probability. But these large values are unlikely to appear close to each other in ξ_j. Therefore, we expect that the averages of randomly constructed f_j would be of much smaller size than the f_j's themselves. This is indeed the case, as is shown in Theorem 11, where we prove that $\|A_k f_j\|_{\Phi_2}$ behaves (up to a polynomial factor) as $2^{(k-j)/2}$ for $0 \leq k \leq j$. Here,

$A_k g$ is the average of a function g over the dyadic intervals $I_{k,l}$, cf. (32). In Theorem 13 and Theorem 15 we provide three more function spaces of this kind, including certain function spaces based on some sort of ball means of differences.

We study also the generalization of the previous approach to the multivariate setting. The high-dimensional analogue of the Wiener process is known as the Brownian sheet, cf. Definition 16. For the sake of brevity, we restrict ourselves to $d = 2$ and the Brownian sheet defined on the unit square $[0, 1]^2$ of $\mathbb{R}^2$, but higher dimensions could be treated in the same way with only minor modifications. The known results in this area go essentially back to the work of Anna Kamont [27, 28] (whose Ph.D. supervisor was Zbigniew Ciesielski). Note however that in general, certain properties of Brownian sheets, especially at the end-points of the function space scales ($p \leq 1$ or $p = \infty$), become much more complicated for dimensions $d \geq 3$ than for $d = 2$ (e.g., small deviation/small ball probabilities, cf. [4] and [5]).

Similarly to the one-dimensional case, one first obtains a decomposition of the paths of the Brownian sheet in a suitable basis, the so-called multivariate Faber system. This is nothing else than the tensor products of the hat functions of the one-dimensional Faber system. The coefficients in this decomposition are again independent standard Gaussian variables. The corresponding function spaces, the spaces of dominating mixed smoothness, are very well-known in the field of approximation theory of functions of several variables. Unfortunately, in [27] Kamont called these spaces *anisotropic Hölder classes*, which might explain why her work went essentially unnoticed by the community of researchers investigating function spaces of dominating mixed smoothness.

Also in this part we re-prove the known results in a way which we hope will be easily accessible for readers with a background in the theory of function spaces. Again, we employ a number of different scales of function spaces to describe the path regularity of the Brownian sheet. These include Besov and Besov-Orlicz spaces of dominating mixed smoothness, as well as Besov spaces of logarithmic dominating mixed smoothness (which surprisingly differ from those introduced by Triebel [58]). Finally, we also propose new function spaces, which are strictly smaller than the best known spaces so far, in which the paths of the Brownian sheet almost surely belong to.

The structure of the paper is as follows. Section 2 treats the univariate Wiener process. We first present Lévy's decomposition of its paths into the Faber system (Theorem 3). Then (in Sect. 2.2) we review the necessary notation from the area of function spaces. In Sect. 2.3 we merge these two subjects and re-prove the results of Lévy and Ciesielski, giving an alternative proof for the Besov-Orlicz space $B^{1/2}_{\Phi_2,\infty}([0, 1])$ in Sect. 2.4. The new function spaces, where the paths of the Wiener process lie in almost surely, are then investigated in Sect. 2.5. Section 3 studies the regularity of Brownian sheets and follows essentially the same pattern. After reviewing the necessary tools in the multivariate setting (Lévy's decomposition, multivariate Faber systems, function spaces of dominating mixed smoothness) we re-prove the results of [27] and sketch the new function spaces, in which one can find almost all paths of the Brownian sheet. Finally, to make the exposition self-contained, Sect. 4 collects basic properties of ran-

dom variables (including Gaussian variables, their absolute values and the integrated absolute Wiener process). We also collect some facts about Orlicz spaces.

2 Regularity of Brownian Paths

In this section we discuss the regularity of the sample paths of the classical Wiener process. Our approach is based on the decomposition, which can be traced back to Lévy [36]. Essentially, it gives a decomposition of the Wiener paths into the Faber system of shifted and dilated hat functions, with the coefficients given by independent Gaussian variables. This, together with characterizations of various function spaces in terms of the Faber system, will allow us to re-prove the classical results of Lévy and Ciesielski, as well as to define new function spaces, where the sample paths of the Wiener process almost surely belong to.

In our work (as it is common in the literature) the notions of Wiener process and Brownian motion are used as synonyms, which both refer to the following definition.

Definition 1 A real-valued random process $W = (W_t)_{t \geq 0}$ is called a Wiener process (or a Brownian motion) if it satisfies

(1) $W_0 = 0$;
(2) W has almost surely continuous paths, i.e., W_t is almost surely continuous in t;
(3) W has independent increments, i.e., if $0 \leq t_0 < t_1 < \cdots < t_n$, then $W_{t_n} - W_{t_{n-1}}, W_{t_{n-1}} - W_{t_{n-2}}, \ldots, W_{t_1} - W_{t_0}$ are independent random variables;
(4) W has Gaussian increments, i.e., $W_t - W_s \sim \mathcal{N}(0, t - s)$ for $0 \leq s \leq t$.

Let the random variables $(W_t)_{t \geq 0}$ be defined on the common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, for every $\omega \in \Omega$ fixed, we call the mapping $t \to W_t(\omega)$ a Brownian path (or a Wiener path).

2.1 Lévy's Decomposition of Brownian Paths

We now present Lévy's representation of Brownian motion, which is essentially a dyadic decomposition of the paths of Brownian motion into a series of piecewise linear functions with random coefficients. Although much of this idea applies to general continuous functions on any closed interval, we restrict ourselves to $(W_t)_{t \in I}$, where $I = [0, 1]$.
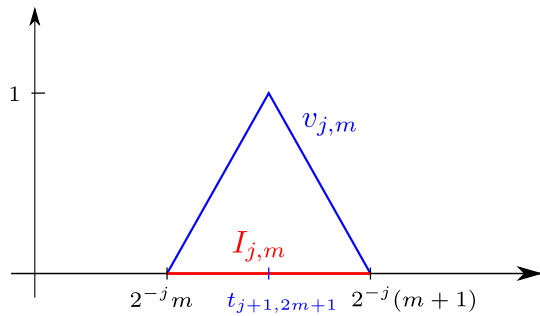
For every $j \in \mathbb{N}_0$ we construct a (random) continuous function $W_j(t)$, which is piecewise linear on all dyadic intervals

$$I_{j,m} = \left[\frac{m}{2^j}, \frac{m+1}{2^j} \right], \quad m \in \{0, \ldots, 2^j - 1\}$$

and which coincides with a given path W_t at their endpoints

$$t_{j,m} = \frac{m}{2^j}, \quad m \in \{0, \ldots, 2^j\}.$$

Fig. 1 Hat function $v_{j,m}$

For $j = 0$, we put $W_0(t) = W_1 \cdot t$ and observe that $W_0(0) = W_0 = 0$ and $W_0(1) = W_1$, i.e., $W_0(t)$ coincides with W_t for $t = t_{0,0} = 0$ and $t = t_{0,1} = 1$.

For $j = 1$, we are looking for a continuous function $W_1(t)$, which would coincide with W_t not only in $t_{1,0} = t_{0,0} = 0$ and $t_{1,2} = t_{0,1} = 1$, but also in $t_{1,1} = 1/2$. For this sake, we add to $W_0(t)$ a suitable multiple of a continuous function $v(t)$, which vanishes at $t = 0$ and $t = 1$ and is linear on $I_{1,0} = [0, 1/2]$ as well as on $I_{1,1} = [1/2, 1]$. Therefore, $v(t)$ is the usual hat function supported in I, i.e.,

$$v(t) = \begin{cases} 2t & \text{if } 0 \le t < \frac{1}{2}, \\ 2(1-t) & \text{if } \frac{1}{2} \le t < 1, \\ 0 & \text{otherwise} \end{cases}$$

and we put

$$W_1(t) = W_0(t) + (W_{1/2} - W_0(1/2))v(t), \quad t \in [0, 1].$$

We proceed further inductively. Let $j \in \mathbb{N}$ be fixed and let us assume that $W_0(t), \ldots, W_j(t)$ were already constructed. Then $W_t - W_j(t)$ vanishes at $t_{j,m}$ for all $m \in \{0, 1, \ldots, 2^j\}$. For $m \in \{0, 1, \ldots, 2^j - 1\}$, we define $W_{j+1}(t)$ for $t \in I_{j,m}$ by adding to $W_j(t)$ a continuous piecewise linear function with support in $I_{j,m}$ to ensure that $W_{j+1}(t) = W_t$ also in the middle point of $I_{j,m}$, i.e., in $t_{j+1,2m+1} = \frac{2m+1}{2^{j+1}} = \frac{m}{2^j} + \frac{1}{2^{j+1}}$. Hence, we need to add to $W_j(t)$ a multiple of the hat function $v_{j,m}(t) = v(2^j(t - t_{j,m}))$ with support in $I_{j,m}$ (see Fig. 1)

$$v_{j,m}(t) = \begin{cases} 2^{j+1}(t - 2^{-j}m) & \text{if } 2^{-j}m \le t < 2^{-j}m + 2^{-j-1}, \\ 2^{j+1}(2^{-j}(m+1) - t) & \text{if } 2^{-j}m + 2^{-j-1} \le t < 2^{-j}(m+1), \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

We repeat this procedure for every $m \in \{0, 1, \ldots, 2^j - 1\}$ and obtain (Fig. 2)

$$W_{j+1}(t) := W_j(t) + \sum_{m=0}^{2^j-1} \left\{ W_{t_{j+1,2m+1}} - W_j(t_{j+1,2m+1}) \right\} v_{j,m}(t), \quad t \in [0, 1]. \tag{4}$$
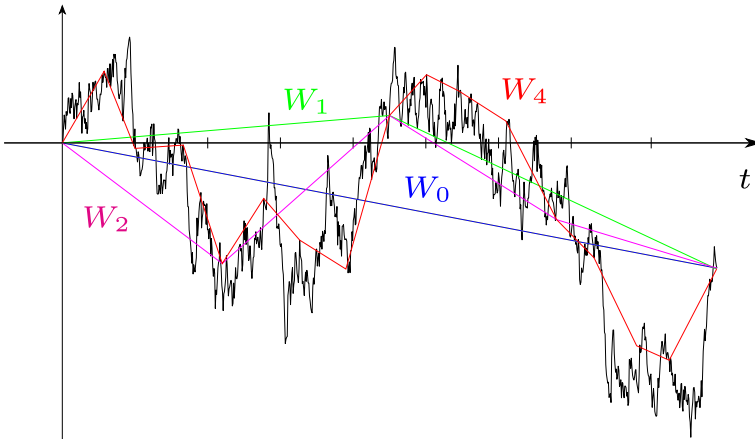
Fig. 2 Piecewise linear functions $W_j(t)$ approximating W_t

The main disadvantage of (4) is that the coefficients in the sum over m involve both the values of the Wiener path W_t as well as the values of its approximation $W_j(t)$. Therefore, we rewrite it in such a form, that only the values of W_t at dyadic points get used.

By its construction, $W_j(t)$ is linear on every $I_{j,m}$ and coincides with W_t at its endpoints and, therefore, we may also write it as

$$W_j(t) = W_{\frac{m}{2^j}} + \left(t - \frac{m}{2^j}\right) \cdot 2^j \cdot \left[W_{\frac{m+1}{2^j}} - W_{\frac{m}{2^j}}\right], \quad t \in I_{j,m}.$$

This allows us to rewrite the coefficients of (4) as

$$
\begin{aligned}
W_{t_{j+1,2m+1}} - W_j(t_{j+1,2m+1}) &= W_{\frac{2m+1}{2^{j+1}}} - W_j\left(\frac{2m+1}{2^{j+1}}\right) \\
&= W_{\frac{2m+1}{2^{j+1}}} - \left(W_{\frac{m}{2^j}} + \frac{1}{2^{j+1}} \cdot 2^j \cdot \left[W_{\frac{m+1}{2^j}} - W_{\frac{m}{2^j}}\right]\right) \\
&= -\frac{1}{2}\left(W_{\frac{2m+2}{2^{j+1}}} - 2W_{\frac{2m+1}{2^{j+1}}} + W_{\frac{2m}{2^{j+1}}}\right) \\
&= -\frac{1}{2}(\Delta^2_{2^{-j-1}}W)\left(\frac{2m}{2^{j+1}}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
(\Delta^2_h f)(x) = f(x+2h) - 2f(x+h) + f(x) &= \Big(f(x+2h) - f(x+h)\Big) \\
&\quad - \Big(f(x+h) - f(x)\Big)
\end{aligned}
$$

are the second order differences of a function f. Together with (4) this leads to

$$W_{j+1}(t) := W_j(t) - \frac{1}{2} \sum_{m=0}^{2^j-1} (\Delta^2_{2^{-j-1}} W)\left(\frac{2m}{2^{j+1}}\right) v_{j,m}(t), \quad t \in [0,1]. \tag{5}$$

The reader may notice that all what we did so far, including (5), applies to general continuous functions on I. We summarize this in the following theorem (and refer to [58, Theorem 2.1] for a detailed proof and to [16, 18] for historic sources).

Theorem 2 *Let $f \in C(I)$. Then*

$$f(t) = f(0) \cdot (1-t) + f(1) \cdot t - \frac{1}{2} \sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} (\Delta^2_{2^{-j-1}} f)(2^{-j}m) v_{j,m}(t) \tag{6}$$

for every $0 \le t \le 1$ and the series converges uniformly on I.

To transform (5) into a series representation of W_t, we note that the variables $W_{\frac{2m+2}{2^{j+1}}} - W_{\frac{2m+1}{2^{j+1}}}$ and $W_{\frac{2m+1}{2^{j+1}}} - W_{\frac{2m}{2^{j+1}}}$ are independent and have by Definition 1 the distribution $\mathcal{N}(0, 2^{-(j+1)})$. Hence, by the 2-stability of the normal distribution, cf. Lemma 26, $W_{t_{j+1,2m+1}} - W_j(t_{j+1,2m+1}) = -\frac{1}{2}(\Delta^2_{2^{-j-1}} W)\left(\frac{2m}{2^{j+1}}\right)$ is normally distributed with mean zero and variance $2^{-(j+2)}$.

We can therefore rewrite (5) as

$$W_{j+1}(t) = W_j(t) + \sum_{m=0}^{2^j-1} 2^{-(j+2)/2} \xi_{j,m} v_{j,m}(t), \quad t \in [0,1],$$

where $\xi_{j,m}$ are standard normal variables. An explicit formula for W_t can then be obtained by noting that the series

$$\sum_{j=0}^{\infty} \left(W_{j+1}(t) - W_j(t) \right) \tag{7}$$

converges almost surely uniformly to

$$W_t - W_0(t) = W_t - W_1 \cdot t \quad \text{with} \quad W_1 = \xi_{-1} \sim \mathcal{N}(0,1).$$

This follows by the tail bound for normal variables, cf. Lemma 27, and a straightforward union bound, which give for every real $A > 1$ that

$$\mathbb{P}\left(\exists j \in \mathbb{N}_0 : \|W_{j+1} - W_j\|_\infty > A \cdot 2^{-j/4} \right) \le \sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} \mathbb{P}(|\xi_{j,m}| > 2A \cdot 2^{j/4})$$

$$\le \sum_{j=0}^{\infty} 2^j \exp(-2A^2 \cdot 2^{j/2}).$$

If A goes to infinity, the last sum tends to zero and, therefore, the probability that $\|W_{j+1} - W_j\|_\infty \leq A \cdot 2^{-j/4}$ for all $j \in \mathbb{N}_0$ grows to one. This ensures that (7) converges uniformly almost surely.

This yields that we have almost surely

$$W_t = \xi_{-1} \cdot t + \sum_{j=0}^{\infty} \sum_{m=0}^{2^j - 1} 2^{-(j+2)/2} \xi_{j,m} v_{j,m}(t), \quad t \in [0, 1]. \tag{8}$$

As the last step, we need to complement (8) by the crucial observation that the random variables $\{\xi_{-1}\} \cup \{\xi_{j,m}, j \in \mathbb{N}_0, 0 \leq m \leq 2^j - 1\}$ are independent. For that sake, let $\xi_{j_1,m_1}, \ldots, \xi_{j_N,m_N}$ be fixed and let us put

$$J = \max(j_1, \ldots, j_N).$$

We collect the independent Gaussian variables

$$W^l = W_{\frac{l+1}{2^{J+1}}} - W_{\frac{l}{2^{J+1}}}, \quad l = 0, 1, \ldots, 2^{J+1} - 1,$$

into a vector $\widetilde{W}^J = (W^0, \ldots, W^{2^{J+1}-1})^T$. Using this notation, we observe that

$$\xi_{-1} = W_1 - W_0 = \sum_{l=0}^{2^{J+1}-1} W^l = \langle (1, \ldots, 1)^T, \widetilde{W}^J \rangle.$$

Similarly, for every $0 \leq j \leq J$ and $m \in \{0, \ldots, 2^j - 1\}$, we get

$$-2^{-j/2} \xi_{j,m} = \left(W_{\frac{2m+2}{2^{j+1}}} - W_{\frac{2m+1}{2^{j+1}}} \right) - \left(W_{\frac{2m+1}{2^{j+1}}} - W_{\frac{2m}{2^{j+1}}} \right)$$

$$= \sum_{l=0}^{2^{J-j}-1} W^{(2m+1)2^{J-j}+l} - \sum_{l=0}^{2^{J-j}-1} W^{2m \cdot 2^{J-j}+l} = \langle h_{j,m}^J, \widetilde{W}^J \rangle,$$

where

$$(h_{j,m}^J)_l = \begin{cases} +1 & \text{if } (2m+1)2^{J-j} \leq l < (2m+2)2^{J-j}, \\ -1 & \text{if } 2m \cdot 2^{J-j} \leq l < (2m+1)2^{J-j}, \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

for $0 \leq l \leq 2^{J+1} - 1$ are the discrete version of the usual Haar functions. The independence of $\xi_{j_1,m_1}, \ldots, \xi_{j_N,m_N}$ now follows from the orthogonality of the vectors $(h_{j_i,m_i}^J)_{i=1}^N$, cf. Lemma 26. This leads to the following representation.

Theorem 3 *Let $(W_t)_{t \in I}$ be the Brownian motion according to Definition 1. If $v_{j,m}$ denotes the Faber system according to* (3)*, then almost surely it holds*

$$W_t = \xi_{-1} \cdot t + \sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} 2^{-(j+2)/2} \xi_{j,m} v_{j,m}(t) \quad for \ t \in [0,1],$$

where $\{\xi_{-1}\} \cup \{\xi_{j,m} : j \in \mathbb{N}_0, 0 \le m \le 2^j - 1\}$ are independent $\mathcal{N}(0,1)$ random variables and the series converges uniformly on I.

2.2 Function Spaces and Faber Systems

As already explained in Sect. 1, the description of the regularity of the paths of the Brownian motion will be given in different scales of function spaces of Besov, Hölder, and Orlicz type. In the sequel, we try to give in brief the basic definitions and characterizations of these spaces.

We assume, that the reader is familiar with the spaces of complex-valued continuous functions $C(\mathbb{R})$ and $C(I)$ as well as with the Lebesgue spaces of integrable functions $L_p(\mathbb{R})$ and $L_p(I)$. For Lebesgue spaces we simplify the notation by writing the norms

$$\|f\|_p := \|f \,|\, L_p\|. \tag{10}$$

The domain I or $\mathbb{R}$ of the function f in (10) should always be clear from the context. For any $0 < p \le \infty$ and any $f \in L_p(\mathbb{R})$, we denote by

$$(\Delta_h^1 f)(x) = f(x+h) - f(x), \qquad (\Delta_h^{M+1} f) = \Delta_h^1(\Delta_h^M f), \tag{11}$$

the usual first-order and higher-order differences (as already briefly mentioned in the previous section), where $x \in \mathbb{R}$, $h \in \mathbb{R}$ and $M \in \mathbb{N}$. We start with the definition of Besov spaces. These spaces have a long history and many of their aspects were studied in the last decades, cf. [45, 48, 57]. In particular, there are many alternative definitions of Besov spaces to be found in the literature (e.g., through the Fourier transform, Littlewood-Paley-type decompositions, wavelets or atoms), which under certain restrictions of the parameters coincide (i.e., they result in the same space with an equivalent (quasi-)norm). For our purposes the approach through finite differences is the most natural one.

Definition 4 (i) Let $s > 0$ and $0 < p, q \le \infty$. Then the Besov space $B_{p,q}^s(\mathbb{R})$ is the collection of all $f \in L_p(\mathbb{R})$ such that for $M = \lfloor s \rfloor + 1$,

$$\|f \,|\, B_{p,q}^s(\mathbb{R})\| = \|f\|_p + \left(\int_0^1 t^{-sq} \sup_{|h| \le t} \|\Delta_h^M f\|_p^q \frac{dt}{t} \right)^{1/q} < \infty.$$

Here, $\lfloor s \rfloor$ is the greatest integer less than or equal to s.

(ii) Besov spaces on the interval $I = [0, 1] \subset \mathbb{R}$ are defined via restriction, i.e.,

$$B_{p,q}^s(I) := \left\{ f \in L_p(I) : \ f = g\big|_I \ \text{for some} \ g \in B_{p,q}^s(\mathbb{R}) \right\}, \qquad (12)$$

normed by

$$\|f|B_{p,q}^s(I)\| = \inf \|g|B_{p,q}^s(\mathbb{R})\|,$$

where the infimum is taken over all $g \in B_{p,q}^s(\mathbb{R})$ with $g\big|_I = f$.

Our approach to the regularity of the Brownian paths is based on the close connection between Besov spaces (and other function spaces) and the decompositions in the Faber system. The Faber system on the interval $I = [0, 1]$ is the collection of functions

$$\{v_0, v_1, v_{j,m} : j \in \mathbb{N}_0, m = 0, \dots, 2^j - 1\},$$

where

$$v_0(x) = 1 - x, \quad v_1(x) = x, \quad x \in I$$

and $v_{j,m}$ is defined by (3) for $j \geq 0$. Let us note that the Faber functions are (possibly up to normalization) essentially the antiderivatives of the Haar functions, which we encountered in their discrete version already in (9). Then Theorem 2.1 of [58] (cf. Theorem 2) shows that the Faber system is a (conditional) basis of $C(I)$ and that every $f \in C(I)$ can be written as

$$f(x) = f(0)v_0(x) + f(1)v_1(x) - \frac{1}{2} \sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} (\Delta_{2^{-j-1}}^2 f)(2^{-j}m)v_{j,m}(x), \quad x \in I.$$

Furthermore, concerning the decomposition of Besov spaces $B_{p,q}^s(I)$ with the above Faber system, we recall Theorem 3.1 in [58, p. 126].

Theorem 5 *Let $0 < p, q \leq \infty$ and*

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

be the admissible range for s as illustrated in the figure aside. Then the sum

$$f = \mu_0 v_0 + \mu_1 v_1 + \sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} \mu_{j,m} 2^{-js} v_{j,m} \qquad (13)$$

with $\mu_0 = \mu_0(f) = f(0)$, $\mu_1 = \mu_1(f) = f(1)$ and

$$\mu_{j,m} = \mu_{j,m}(f) = -2^{js-1}(\Delta^2_{2^{-j-1}}f)(2^{-j}m)$$
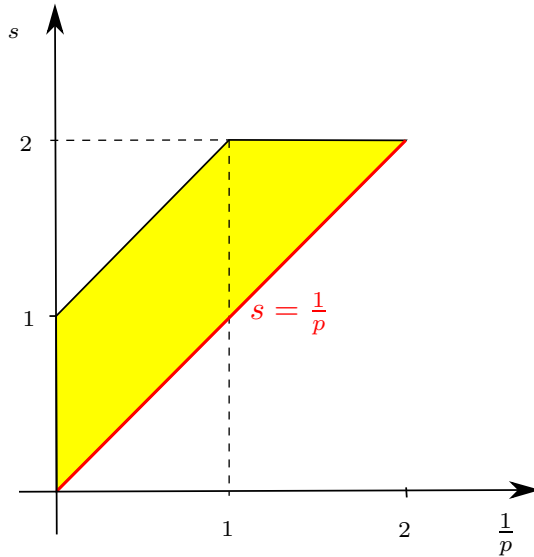
lies in $B^s_{p,q}(I)$ if, and only if,

$$\|\mu|b^+_{p,q}(I)\| := |\mu_0| + |\mu_1| + \left(\sum_{j=0}^{\infty}\left(\sum_{m=0}^{2^j-1} 2^{-j}|\mu_{j,m}|^p\right)^{q/p}\right)^{1/q} < \infty.$$

2.2.1 Function Spaces of Logarithmic Smoothness

As pointed out already in the Introduction, one can obtain finer descriptions of the regularity properties of the Brownian paths, if one uses different (and more sophisticated) scales of function spaces.

Therefore, we now introduce the so-called function spaces of logarithmic smoothness, cf. [13, 17, 26, 41]. Here again we rely on the exposition and results from [58]. In particular, function spaces of logarithmic smoothness are a special case of Besov and Triebel-Lizorkin spaces of generalized smoothness, where the smoothness factor

$$2^{js} \qquad \text{gets replaced by} \qquad 2^{js}(1+j)^{-\alpha}$$

with $s, \alpha \in \mathbb{R}$. Following [58, Proposition 1.7.4] we can define function spaces of logarithmic smoothness by differences as follows.

Definition 6 Let $0 < p, q \le \infty$, $s, \alpha \in \mathbb{R}$, and $M = \lfloor s \rfloor + 1$ with

$$s > \max\left(\frac{1}{p}, 1\right) - 1.$$

Then the logarithmic space $B_{p,q}^{s,\alpha}(\mathbb{R})$ contains all functions $f \in L_p(\mathbb{R})$ with

$$\left\| f \,|\, B_{p,q}^{s,\alpha}(\mathbb{R}) \right\| = \|f\|_p + \left(\int_0^1 t^{-sq} (1 + |\log t|)^{-\alpha q} \sup_{0 < h < t} \|\Delta_h^M f\|_p^q \frac{dt}{t} \right)^{1/q} < \infty. \tag{14}$$

Note that in comparison with [58] we have replaced α by $-\alpha$ in the definition of the logarithmic space $B_{p,q}^{s,\alpha}(\mathbb{R})$ leading to some minor adaptations in the theorem below. If $p = \infty$ and/or $q = \infty$, then (14) has to be interpreted accordingly. Especially, if $p = q = \infty$, then

$$\left\| f \,|\, B_{\infty,\infty}^{s,\alpha}(\mathbb{R}) \right\| = \|f\|_\infty + \sup_{0 < t < 1} \frac{\omega_M(f,t)_\infty}{t^s (1 + |\log t|)^\alpha},$$

where $\omega_M(f,t)_\infty = \sup_{0 < h < t} \|\Delta_h^M f\|_\infty$ denotes the corresponding modulus of continuity. As a consequence, the spaces $B_{\infty,\infty}^{s,\alpha}(\mathbb{R})$ coincide with the spaces $\mathcal{C}^{s,\alpha}(\mathbb{R})$ of Hölder-Zygmund type, which are usually defined by

$$\|f|\mathcal{C}^{s,\alpha}(\mathbb{R})\| := \|f\|_\infty + \sup_{0 < t < 1/2} \frac{\omega_M(f,t)_\infty}{t^s |\log t|^\alpha},$$

cf. [20, Rem. 9] or [42, Rem. 2.8.(ii)]. Following [58, Theorem 3.30] we have also a characterization in terms of the Faber system of the spaces $B_{p,q}^{s,\alpha}(I)$ in which we are interested here. The restriction from $\mathbb{R}$ to I is done in the same way as described in Sect. 2.2, cf. (12).

Theorem 7 *Let* $0 < p, q \leq \infty$ *and* $s, \alpha \in \mathbb{R}$ *with*

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

Then $f \in L_p(I)$ *belongs to* $B_{p,q}^{s,\alpha}(I)$ *if, and only if, it can be represented as*

$$f = \mu_0 v_0 + \mu_1 v_1 + \sum_{j=0}^{\infty} \sum_{m=0}^{2^j - 1} \mu_{j,m} 2^{-js} v_{j,m}, \tag{15}$$

where $\mu_{j,m} \in b_{p,q}^{+,\alpha}(I)$, *i.e.,*

$$\|\mu|b_{p,q}^{+,\alpha}(I)\| := |\mu_0| + |\mu_1| + \left(\sum_{j=0}^{\infty} (1+j)^{-\alpha q} \left(\sum_{m=0}^{2^j - 1} 2^{-j} |\mu_{j,m}|^p \right)^{q/p} \right)^{1/q} < \infty.$$

Here the sum in (15) converges unconditionally in $B_{p,q}^{\sigma}(I)$ *for* $\sigma < s$ *and in* $C(I)$. *Moreover, the representation is unique with*

$$\mu_0 = \mu_0(f) = f(0), \ \mu_1 = \mu_1(f) = f(1) \ and$$

$$\mu_{j,m} = \mu_{j,m}(f) = -2^{js-1}(\Delta^2_{2^{-j-1}} f)(2^{-j}m)$$

where $j \in \mathbb{N}_0$ and $m = 0, \dots, 2^j - 1$.

Remark 1 In the theory of function spaces there are two different approaches on how to deal with the normalization factors appearing in characterizations with building blocks such as atoms, wavelets or Faber functions (as it is our case here). The first approach tries to take the same building blocks for any function space $B^s_{p,q}$ or $B^{s,\alpha}_{p,q}$ independent from the chosen smoothness parameters s, α. This results in the adaption of the corresponding sequence spaces in $b^s_{p,q}$ or $b^{s,\alpha}_{p,q}$, where now these parameters play a role.

The second approach changes the definition of the building blocks according to the smoothness parameters s and α. The consequence is that then the corresponding sequence spaces $b_{p,q}$ are independent on s and α. In this work, we always use the same decomposition of the function, which corresponds to Lévy's decomposition (8). For that reason, we prefer to work with (13) and (15) in the above theorems. This results in Theorem 5, where the sequence spaces $b^+_{p,q}$ corresponding to $B^s_{p,q}(I)$ are independent of the chosen smoothness s. Furthermore in Theorem 7 we only include the logarithmic smoothness parameter within the sequence space norm. To that end, now $b^{+,\alpha}_{p,q}(I)$ corresponds to $B^{s,\alpha}_{p,q}(I)$, which is still independent on s but depends on α.

2.2.2 Besov-Orlicz Spaces

We replace the Lebesgue norm $\|\cdot\|_\infty$ in the interesting boundary case $p = \infty$ in the definition of the Besov spaces by the Orlicz-norm $\|\cdot\|_{\Phi_2}$, which is given by the following Young function

$$\Phi_2(t) = \exp(t^2) - 1 \quad \text{for} \quad t > 0. \tag{16}$$

In particular, we define for dimension $d \in \mathbb{N}$ the Orlicz space $L_{\Phi_2}([0,1]^d)$ as the collection of all measurable functions on $[0,1]^d$ with

$$\|f\|_{\Phi_2} := \inf\left\{\lambda > 0 : \int_{[0,1]^d} \Phi_2\left(\frac{|f(t)|}{\lambda}\right) dt \le 1\right\} < \infty. \tag{17}$$

This Orlicz norm and its equivalent expressions play a fundamental role in the characterization of sub-gaussian random variables [25, 63].

By Theorems 31 and 32, $\|f\|_{\Phi_2}$ is also equivalent to

$$\sup_{0<t<1} \frac{f^*(t)}{\sqrt{\log(1/t)+1}} \quad \text{and} \quad \sup_{p\ge 1} \frac{\|f\|_p}{\sqrt{p}},$$

where

$$f^*(t) = \inf \left\{ s \in [0,1] : |\{r \in [0,1]^d : |f(r)| > s\}| \leq t \right\}$$

is the non-increasing rearrangement of f.

Now we are going to define the Besov-Orlicz spaces $B^{1/2}_{\Phi_2,\infty}(I)$ directly via the decompositions with the Faber system. The sequence space norm is a direct adaptation from the sequence spaces $b^+_{p,q}(I)$ from Theorem 5, where the $L_p(I)$ is now replaced by the Orlicz norm $L_{\Phi_2}(I)$, and the characterization from Theorem 32. We also refer to [10, Theorem III.8], where a characterization of these spaces with the help of the Faber system is given and to [46], where one can find an alternative approach to Besov-Orlicz spaces.

Definition 8 (i) The sequence space $b^+_{\Phi_2,\infty}$ is the collection of all sequences

$$\{\mu = (\mu_{j,m}) : j \in \mathbb{N}_0, \ m = 0, \ldots, 2^j - 1\}$$

such that

$$\left\| \mu \,|\, b^+_{\Phi_2,\infty}(I) \right\| := \sup_{j \in \mathbb{N}_0} \left\| \sum_{m=0}^{2^j-1} \mu_{j,m} \chi_{j,m} \right\|_{\Phi_2} \approx \sup_{j \in \mathbb{N}_0} \sup_{p \geq 1} \frac{1}{\sqrt{p}} \left\| \sum_{m=0}^{2^j-1} \mu_{j,m} \chi_{j,m} \right\|_p$$

$$= \sup_{j \in \mathbb{N}_0} \sup_{p \geq 1} \frac{1}{\sqrt{p}} \left(\sum_{m=0}^{2^j-1} 2^{-j} |\mu_{j,m}|^p \right)^{1/p} < \infty,$$

where $\chi_{j,m}$ is the characteristic function of $I_{j,m}$.

(ii) The function space $B^{1/2}_{\Phi_2,\infty}(I)$ is the collection of all $f \in C(I)$ such that the coefficients of its representation

$$f(x) = \lambda_0 v_0(x) + \lambda_1 v_1(x) + \sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} 2^{-j/2} \lambda_{j,m} v_{j,m}(x), \quad x \in I, \qquad (18)$$

satisfy $\left\| \lambda \,|\, b^+_{\Phi_2,\infty}(I) \right\| < \infty$.

2.3 Results of Lévy and Ciesielski

In this section we present proofs of the results of Lévy [36] and Ciesielski [9] concerning the regularity of Wiener paths in certain function spaces. Our main aim is to show that they both follow quite directly from Lévy's decomposition of Wiener paths into the Faber system (8) combined with the characterization of the corresponding function spaces via the Faber system. For that sake, we summarize Theorems 5, 7, and Definition 8, which state the conditions on the coefficients guaranteeing that a function

belongs to the function spaces considered by Lévy and Ciesielski. In particular, we choose a formulation, which corresponds directly to (8).

Theorem 9 *Consider a function* $f \in C(I)$ *with the representation*

$$
f(x) = \lambda_0 v_0(x) + \lambda_1 v_1(x) + \sum_{j=0}^{\infty} \sum_{m=0}^{2^j-1} 2^{-\frac{j+2}{2}} \lambda_{j,m} v_{j,m}(x), \quad x \in I,
$$

where $\{v_0, v_1, v_{j,m} : j \in \mathbb{N}_0, m = 0, \dots, 2^j - 1\}$ *denotes the Faber system on the interval* $I = [0, 1]$.

(i) f *belongs to* $B_{\infty,\infty}^{1/2,1/2}(I)$ *if, and only if*

$$
\sup_{j \in \mathbb{N}} \frac{1}{\sqrt{j}} \sup_{m=0,\dots,2^j-1} |\lambda_{j,m}| < \infty. \tag{19}
$$

(ii) Let $1 \le p < \infty$. *Then* f *belongs to* $B_{p,\infty}^{1/2}(I)$ *if, and only if*

$$
\sup_{j \in \mathbb{N}} \Big(\sum_{m=0}^{2^j-1} 2^{-j} |\lambda_{j,m}|^p \Big)^{1/p} < \infty. \tag{20}
$$

(iii) f *belongs to* $B_{\Phi_2,\infty}^{1/2}(I)$ *if, and only if*

$$
\sup_{j \in \mathbb{N}} \sup_{p \ge 1} \frac{1}{\sqrt{p}} \Big(\sum_{m=0}^{2^j-1} 2^{-j} |\lambda_{j,m}|^p \Big)^{1/p} < \infty. \tag{21}
$$

1. **Wiener paths belong to** $B_{\infty,\infty}^{1/2,1/2}(I)$ **almost surely (Lévy** [36]) Comparing (8) with (19), it is enough to show that

$$
\sup_{j \in \mathbb{N}} \frac{1}{\sqrt{j}} \sup_{m=0,\dots,2^j-1} |\xi_{j,m}| < \infty \quad \text{almost surely}, \tag{22}
$$

where $\{\xi_{j,m} : j \in \mathbb{N}, m = 0, \dots, 2^j - 1\}$ are independent standard Gaussian variables.

To prove (22), we denote by A_j^N the event when $\sup_{m=0,\dots,2^j-1} |\xi_{j,m}| \ge N\sqrt{j}$ and estimate

$$
\mathbb{P}(A_j^N) \le 2^j \mathbb{P}(|\omega| \ge N\sqrt{j}) \le 2^j e^{-N^2 j/2}, \tag{23}
$$

where we used the estimate from Lemma 27 (i). Using an estimate, which resembles the approach of the Borel-Cantelli lemma and which we shall use frequently

later on, we obtain for every positive integer N_0,

$$\mathbb{P}\Big(\sup_{j\in\mathbb{N}}\frac{1}{\sqrt{j}}\sup_{m=0,\dots,2^j-1}|\xi_{j,m}|=\infty\Big)=\mathbb{P}\Big(\bigcap_{N=1}^{\infty}\bigcup_{j=1}^{\infty}A_j^N\Big)\le\mathbb{P}\Big(\bigcup_{j=1}^{\infty}A_j^{N_0}\Big)$$

$$\le\sum_{j=1}^{\infty}\mathbb{P}(A_j^{N_0})\le\sum_{j=1}^{\infty}2^je^{-N_0^2j/2}.\quad(24)$$

As the last expression tends to zero if $N_0\to\infty$, this finishes the proof of (22).

2. **Wiener paths belong to $B_{p,\infty}^{1/2}(I)$ almost surely (Ciesielski [8])**

We restrict ourselves to $2<p<\infty$, which allows us to use Theorem 5. The smaller values of p are then covered by the monotonicity of Besov spaces on domains with respect to the integrability parameter p. Again, by (8) and (20), it is enough to prove that

$$\sup_{j\in\mathbb{N}}\Big(\sum_{m=0}^{2^j-1}2^{-j}|\xi_{j,m}|^p\Big)^{1/p}<\infty\quad\text{almost surely for every }\ 2<p<\infty.\quad(25)$$

Fix $2<p<\infty$ and let $\mu_p=\mathbb{E}\,|\omega|^p$ be the pth absolute moment of a standard Gaussian variable ω. This time, we denote for every $t>0$ and $j\in\mathbb{N}$ by A_j^t the event that

$$\frac{1}{2^j}\sum_{m=0}^{2^j-1}|\xi_{j,m}|^p-\mu_p\ge t.$$

Then, by Markov's inequality,

$$t^2\mathbb{P}(A_j^t)=t^2\mathbb{P}\bigg(\Big(\frac{1}{2^j}\sum_{m=0}^{2^j-1}|\xi_{j,m}|^p-\mu_p\Big)^2\ge t^2\bigg)\le\mathbb{E}\Big(\frac{1}{2^j}\sum_{m=0}^{2^j-1}|\xi_{j,m}|^p-\mu_p\Big)^2$$

$$=\mathbb{E}\frac{1}{2^{2j}}\sum_{m=0}^{2^j-1}|\xi_{j,m}|^{2p}+\mathbb{E}\frac{1}{2^{2j}}\sum_{m\ne n=0}^{2^j-1}|\xi_{j,m}|^p|\xi_{j,n}|^p-\frac{2\mu_p}{2^j}\mathbb{E}\sum_{m=0}^{2^j-1}|\xi_{j,m}|^p+\mu_p^2$$

$$=\frac{\mu_{2p}}{2^j}+\frac{2^j(2^j-1)}{2^{2j}}\mu_p^2-\mu_p^2=\frac{\mu_{2p}-\mu_p^2}{2^j}.\quad(26)$$

Similarly to (24), we conclude, that for every $N_0\in\mathbb{N}$ it holds

$$\mathbb{P}\bigg(\sup_{j\in\mathbb{N}_0}\frac{1}{2^j}\sum_{m=0}^{2^j-1}|\xi_{j,m}|^p=\infty\bigg)=\mathbb{P}\Big(\bigcap_{N=1}^{\infty}\bigcup_{j=0}^{\infty}A_j^N\Big)\le\mathbb{P}\Big(\bigcup_{j=0}^{\infty}A_j^{N_0}\Big)\le\sum_{j=0}^{\infty}\mathbb{P}(A_j^{N_0})$$

$$\le\sum_{j=0}^{\infty}\frac{\mu_{2p}-\mu_p^2}{2^jN_0^2}=\frac{2(\mu_{2p}-\mu_p^2)}{N_0^2}.$$

The last expression tends to zero if $N_0\to\infty$, which renders (25).

3. **Wiener paths belong to $B^{1/2}_{\Phi_2,\infty}(I)$ almost surely (Ciesielski [9])**
   This time, we need to show that

$$\sup_{j\in\mathbb{N}} \sup_{p\geq 1} \frac{1}{\sqrt{p}} \Big(\frac{1}{2^j} \sum_{m=0}^{2^j-1} |\xi_{j,m}|^p\Big)^{1/p} < \infty \quad \text{almost surely.} \tag{27}$$

By monotonicity, it is enough to restrict the supremum over p to the integer values $p \in \mathbb{N}$. Furthermore, we only need to refine the analysis done before. Indeed, it follows directly from (26) that

$$\mathbb{P}\Big(\frac{1}{\sqrt{p}} \Big(\frac{1}{2^j} \sum_{m=0}^{2^j-1} |\xi_{j,m}|^p\Big)^{1/p} \geq \frac{(2t)^{1/p}}{\sqrt{p}}\Big) \leq \mathbb{P}(A_j^t) \leq \frac{\mu_{2p}-\mu_p^2}{2^j t^2} \quad \text{for every} \quad t > \mu_p.$$

Hence, if $N > (2\mu_p)^{1/p}/\sqrt{p}$, and if $A_j^{N,p}$ denotes the event when

$$\frac{1}{\sqrt{p}} \Big(\frac{1}{2^j} \sum_{m=0}^{2^j-1} |\xi_{j,m}|^p\Big)^{1/p} \geq N,$$

then

$$\mathbb{P}\big(A_j^{N,p}\big) \leq \frac{\mu_{2p}-\mu_p^2}{2^{j-2}(N\sqrt{p})^{2p}} \leq \frac{\mu_{2p}}{2^{j-2}(N\sqrt{p})^{2p}}.$$

Since $\mu_{2p} = \dfrac{2^p \Gamma\big(p+\frac{1}{2}\big)}{\sqrt{\pi}} \leq 2^p \cdot \Gamma(p+1) = 2^p \cdot p! \leq (2p)^p$, we obtain for every N_0 large enough

$$\mathbb{P}\Big(\sup_{j\in\mathbb{N}} \sup_{p\in\mathbb{N}} \frac{1}{\sqrt{p}} \Big(\frac{1}{2^j} \sum_{m=0}^{2^j-1} |\xi_{j,m}|^p\Big)^{1/p} = \infty\Big) = \mathbb{P}\Big(\bigcap_{N=1}^{\infty} \bigcup_{j,p=1}^{\infty} A_j^{N,p}\Big) \leq \mathbb{P}\Big(\bigcup_{j,p=1}^{\infty} A_j^{N_0,p}\Big)$$

$$\leq \sum_{j,p=1}^{\infty} \mathbb{P}\big(A_j^{N_0,p}\big) \leq \sum_{j,p=1}^{\infty} \frac{\mu_{2p}}{2^{j-2}(N_0\sqrt{p})^{2p}} \lesssim \sum_{p=1}^{\infty} \frac{2^p}{N_0^{2p}}.$$

As the last expression tends to zero if $N_0 \to \infty$, we obtain (27).

Remark 2 Let us have a closer look at the relations of the different spaces we are dealing with in terms of their embeddings. We rely on the characterizations (19), (20), and (21) collected in Theorem 9. Furthermore, one can show in the same way that a function f with (18) belongs to the Hölder-Zygmund space $B^{1/2}_{\infty,\infty}(I)$ if, and only if,

$$\sup_{j\in\mathbb{N}_0} \sup_{m=0,\ldots,2^j-1} |\lambda_{j,m}| < \infty. \tag{28}$$

One can use (28) to recover the well known fact that the Wiener paths almost surely do not belong to $B_{\infty,\infty}^{1/2}(I)$. Indeed, the supremum of 2^j independent standard Gaussian variables grows asymptotically as $\sqrt{j}$, cf. also Lemma 27(ii), which violates (28). Furthermore, (20), (21), and (28) show that

$$B_{\infty,\infty}^{1/2}(I) \hookrightarrow B_{\Phi_2,\infty}^{1/2}(I) \hookrightarrow B_{p,\infty}^{1/2}(I).$$

To compare the spaces $B_{\Phi_2,\infty}^{1/2}(I)$ and $B_{\infty,\infty}^{1/2,1/2}(I)$ we estimate for every $j \in \mathbb{N}$

$$\frac{1}{\sqrt{j}} \cdot \sup_{m=0,\ldots,2^j-1} |\lambda_{j,m}| = \frac{1}{\sqrt{j}} \cdot \|\lambda_{j,\cdot}\|_\infty \le \frac{1}{\sqrt{j}} \cdot \|\lambda_{j,\cdot}\|_j = \frac{2}{\sqrt{j}} \left(\sum_{m=0}^{2^j-1} 2^{-j} |\lambda_{j,m}|^j \right)^{1/j}$$

$$\le 2 \sup_{p \ge 1} \frac{1}{\sqrt{p}} \left(\sum_{m=0}^{2^j-1} 2^{-j} |\lambda_{j,m}|^p \right)^{1/p}.$$

Therefore, $B_{\Phi_2,\infty}^{1/2}(I) \hookrightarrow B_{\infty,\infty}^{1/2,1/2}(I)$ and the result of Ciesielski is an improvement over the result of P. Lévy, providing a strictly smaller space, which contains almost all Wiener paths.

Finally, let us note that $B_{\infty,\infty}^{1/2,1/2}(I)$ and $B_{p,\infty}^{1/2}(I)$ with $1 \le p < \infty$ are incomparable. This can again be easily seen by looking at the sequence space characterization (19) and (20). First, the special sequence $\lambda_{j,m}^{(1)} = \sqrt{j}$ for all $j \in \mathbb{N}$ and all $m = 0, \ldots, 2^j - 1$ belongs to $B_{\infty,\infty}^{1/2,1/2}(I)$ and not to $B_{p,\infty}^{1/2}(I)$ for any $1 \le p < \infty$. Second, the sequence

$$\lambda_{j,m}^{(2)} = \begin{cases} j, & \text{for } j \in \mathbb{N} \text{ and } m = 0, \\ 0, & \text{for } j \in \mathbb{N} \text{ and } m \ge 1 \end{cases}$$

belongs to $B_{p,\infty}^{1/2}(I)$ for all $1 \le p < \infty$ but not to $B_{\infty,\infty}^{1/2,1/2}(I)$.

2.4 An Alternative Proof for the Besov-Orlicz Space $B_{\Phi_2,\infty}^{1/2}(I)$

We use the characterization given in Theorem 31 to re-prove the result of Ciesielski [9], i.e., to show that the Wiener paths almost surely lie in the Besov-Orlicz space $B_{\Phi_2,\infty}^{1/2}(I)$. Comparing (8) with (18) we observe, that it is enough to show that

$$\left\| \xi | b_{\Phi_2,\infty}^+(I) \right\| = \sup_{j \in \mathbb{N}_0} \left\| \sum_{m=0}^{2^j-1} \xi_{j,m} \chi_{j,m}(\cdot) \right\|_{\Phi_2} < \infty$$

almost surely. Here, $\xi = \{\xi_{j,m}, j \in \mathbb{N}_0, 0 \le m \le 2^j - 1\}$ is again a sequence of i.i.d. standard Gaussian variables and $\chi_{j,m}$ denotes the characteristic function of the interval $I_{j,m}$.

Therefore, we put for every integer $j \geq 0$

$$f_j(t) = \sum_{m=0}^{2^j - 1} \xi_{j,m} \chi_{j,m}(t), \quad 0 < t < 1, \tag{29}$$

and observe that its non-increasing rearrangement is given by

$$f_j^*(t) = \sum_{m=0}^{2^j - 1} \left(\xi_j\right)_{m+1}^* \chi_{j,m}(t), \quad 0 < t < 1,$$

where $\left(\left(\xi_j\right)_m^*\right)_{m=1}^{2^j}$ is the non-increasing rearrangement of the sequence $\xi_j = (|\xi_{j,m}|)_{m=0}^{2^j - 1}$.

This allows us to calculate for $j \in \mathbb{N}$ fixed

$$\sup_{0 < t < 1} \frac{f_j^*(t)}{\sqrt{\log(1/t) + 1}} = \sup_{m=1,\ldots,2^j} \frac{f_j^*(m 2^{-j})}{\sqrt{\log(2^j/m) + 1}}$$

$$= \sup_{k=0,\ldots,j-1} \sup_{2^k \leq m \leq 2^{k+1}} \frac{f_j^*(m 2^{-j})}{\sqrt{\log(2^j/m) + 1}}$$

$$\leq \sup_{k=0,\ldots,j-1} \frac{f_j^*(2^{k-j})}{\sqrt{\log(2^{j-k-1}) + 1}} \leq c \sup_{0 \leq k < j} \frac{f_j^*(2^{k-j})}{\sqrt{j-k}}$$

$$= c \sup_{0 \leq k < j} \frac{\left(\xi_j\right)_{2^k}^*}{\sqrt{j-k}}. \tag{30}$$

Next, we denote by $A_{j,k}^K$ the event that $\left(\xi_j\right)_{2^k}^* \geq K\sqrt{j-k}$ and use Lemma 27 to estimate $\mathbb{P}(A_{j,k}^K)$. We conclude, that for every $K_0 \in \mathbb{N}$ large enough such that $16 e^{-K_0^2/2} < 1$ it holds

$$\mathbb{P}\left(\sup_{0 < t < 1} \sup_{j \in \mathbb{N}_0} \frac{f_j^*(t)}{\sqrt{\log(1/t) + 1}} = \infty\right) = \mathbb{P}\left(\bigcap_{K=1}^{\infty} \bigcup_{0 \leq k < j < \infty} A_{j,k}^K\right) \leq \mathbb{P}\left(\bigcup_{0 \leq k < j < \infty} A_{j,k}^{K_0}\right)$$

$$\leq \sum_{0 \leq k < j < \infty} \mathbb{P}\left(A_{j,k}^{K_0}\right) \leq \sum_{0 \leq k < j < \infty} \left(2 e^{-K_0^2/2}\right)^{(j-k) 2^k} \cdot e^{2^k}$$

$$= \sum_{k=0}^{\infty} \sum_{l=1}^{\infty} \left(2 e^{-K_0^2/2}\right)^{l 2^k} \cdot e^{2^k} = \sum_{k=0}^{\infty} e^{2^k} \sum_{l=1}^{\infty} \left(\left(2 e^{-K_0^2/2}\right)^{2^k}\right)^l$$

$$\leq c \sum_{k=0}^{\infty} e^{2^k} \left(2 e^{-K_0^2/2}\right)^{2^k} \leq c \sum_{k=0}^{\infty} \left(8 e^{-K_0^2/2}\right)^{k+1} \leq 2c \cdot 8 e^{-K_0^2/2}.$$

The last expression tends to zero if $K_0 \to \infty$, which yields the desired result.

2.5 New Function Spaces

The approach to regularity of Wiener paths presented in Sects. 2.3 and 2.4 follows actually a rather straightforward pattern. If a certain function space under consideration allows for an equivalent characterization in terms of the Faber system, then this might be combined directly with Theorem 3. The proof that almost all Wiener paths lie in this space then reduces to a statement about independent standard Gaussian variables.

In this section we show, that this approach can be further developed to introduce even smaller spaces than $B_{\Phi_2,\infty}^{1/2}(I)$, where the Wiener paths almost surely lie.

2.5.1 Spaces of Besov Type: Discrete Averages of Differences

By Definition 8 and Theorem 31, a continuous function $f \in C(I)$ lies in the Besov-Orlicz space $B_{\Phi_2,\infty}^{1/2}(I)$ if, and only if, its decomposition into the Faber system (18) satisfies

$$\sup_{j \in \mathbb{N}_0} \sup_{0<t<1} \frac{f_j^*(t)}{\sqrt{\log(1/t)+1}} < \infty,$$

where

$$f_j(t) = \sum_{m=0}^{2^j-1} \lambda_{j,m} \chi_{j,m}(t), \quad 0 < t < 1. \tag{31}$$

Here $\chi_{j,m}$ stands for the characteristic function of the dyadic interval $I_{j,m} = [m \cdot 2^{-j}, (m+1) \cdot 2^{-j}]$. This condition quantifies very precisely the possible size of the components f_j, but (due to the use of the rearrangements f_j^*) it fails to describe the distribution of large values of f_j on $[0,1]$. For example, it does not exclude the possibility that the large values of the components f_j appear close to each other. But this is actually unlikely for the Wiener process because in that case, the $\lambda_{j,m}$'s get replaced by independent Gaussian variables.

Therefore, we introduce new function spaces, where we measure the size of the averages

$$\frac{1}{t-s} \int_s^t f_j(u)du \quad \text{or} \quad \frac{1}{t-s} \int_s^t |f_j(u)|du$$

and we expect them to be much smaller (in L_∞ or the L_{Φ_2}-norm) than f_j itself. To describe this idea mathematically, we define for an integrable function g on I the averaging operators

$$(A_k g)(x) = \sum_{l=0}^{2^k-1} 2^k \int_{l2^{-k}}^{(l+1)2^{-k}} g(t)dt \cdot \chi_{k,l}(x) \quad \text{and} \ (\widetilde{A}_k g)(x) = (A_k(|g|))(x), \ x \in I. \tag{32}$$

By what we outlined so far, we expect the kth dyadic average A_k of the jth dyadic level of Lévy's decomposition (8) to be much smaller in the L_{Φ_2}-norm than the jth dyadic level itself. This paves the way to the following definition.

Definition 10 Let $\varepsilon > 0$. Then the space $A(\varepsilon)$ is the collection of all $f \in C(I)$, which satisfy the following condition. If (18) is the decomposition of f into the Faber system and f_j, $j \in \mathbb{N}_0$, is defined by (31), then

$$\|f\|_{A(\varepsilon)} := \sup_{j \in \mathbb{N}_0} \sup_{0 \le k \le j} \sup_{0 < t < 1} \frac{2^{(j-k)/2}}{(j-k+1)^\varepsilon} \cdot \frac{(A_k f_j)^*(t)}{\sqrt{\log(1/t)+1}}$$

$$\approx \sup_{j \in \mathbb{N}_0} \sup_{0 \le k \le j} \frac{2^{(j-k)/2}}{(j-k+1)^\varepsilon} \cdot \|A_k f_j\|_{\Phi_2} < \infty.$$

With this definition we are now able to prove the following.

Theorem 11 *Let $\varepsilon > 0$. Then $\|W.\|_{A(\varepsilon)} < \infty$ almost surely.*

Proof To estimate the norm of the Brownian paths in $A(\varepsilon)$, we use the decomposition of W_t into the Faber system (8). We therefore replace (31) by

$$f_j(t) = \sum_{m=0}^{2^j-1} \xi_{j,m} \chi_{j,m}(t), \quad 0 < t < 1,$$

where $\xi_j = (\xi_{j,m})_{m=0}^{2^j-1}$ is again a vector of independent standard Gaussian variables.
    Let $0 \le k \le j$ and $x \in I$. Then

$$2^{(j-k)/2} A_k f_j(x) = 2^{(j-k)/2} \sum_{l=0}^{2^k-1} 2^k \int_{l \cdot 2^{-k}}^{(l+1)2^{-k}} \left(\sum_{m=0}^{2^j-1} \xi_{j,m} \chi_{j,m}(t) \right) dt \cdot \chi_{k,l}(x)$$

$$= \sum_{l=0}^{2^k-1} 2^{(j+k)/2} \left(\sum_{m=0}^{2^j-1} \xi_{j,m} |I_{k,l} \cap I_{j,m}| \right) \chi_{k,l}(x)$$

$$= \sum_{l=0}^{2^k-1} 2^{(k-j)/2} \left(\sum_{m=l \cdot 2^{j-k}}^{(l+1) \cdot 2^{j-k}-1} \xi_{j,m} \right) \chi_{k,l}(x).$$

By the 2-stability of Gaussian variables, cf. Lemma 26, it follows that $2^{(j-k)/2} A_k f_j$ is equidistributed with f_k. Therefore,

$$\mathbb{P}(\|W.\|_{A(\varepsilon)} \ge K_0) \le \sum_{0 \le k \le j} \mathbb{P}\left(\frac{2^{(j-k)/2}}{(j-k+1)^\varepsilon} \cdot \sup_{0 < t < 1} \frac{(A_k f_j)^*(t)}{\sqrt{\log(1/t)+1}} \ge K_0 \right)$$

$$= \sum_{0 \le k \le j} \mathbb{P}\left(\sup_{0 < t < 1} \frac{(f_k)^*(t)}{\sqrt{\log(1/t)+1}} \ge K_0 \cdot (j-k+1)^\varepsilon \right) = I_1 + I_2,$$

where I_1 collects the terms with $0 = k \le j$ and I_2 includes the terms with $1 \le k \le j$.

The estimate of I_1 is rather straightforward

$$I_1 \leq \sum_{j=0}^{\infty} \mathbb{P}\Big(f_0^*(1) \geq K_0 \cdot (j+1)^{\varepsilon}\Big) \leq \sum_{j=0}^{\infty} \exp\Big(-\frac{K_0^2(j+1)^{2\varepsilon}}{2}\Big)$$

and this expression tends to zero if K_0 grows to infinity.

Using (30) and Lemma 27, we may estimate I_2 as follows

$$
\begin{aligned}
I_2 &\leq \sum_{0 \leq m < k \leq j} \mathbb{P}\left(\frac{(\xi_k)_{2^m}^*}{\sqrt{k-m}} \geq c_1 K_0 \cdot (j-k+1)^{\varepsilon}\right) \\
&\leq \sum_{0 \leq m < k \leq j} c\, e^{2^m} \Big(2 \exp(-c_1^2 K_0^2 (j-k+1)^{2\varepsilon})\Big)^{(k-m)2^m} \\
&= \sum_{0 \leq m < k} c\, e^{2^m} 2^{(k-m)2^m} \sum_{l=1}^{\infty} \Big(\exp(-c_1^2 K_0^2 (k-m)2^m)\Big)^{l^{2\varepsilon}}.
\end{aligned}
$$

We assume that K_0 is large enough to ensure $8 \exp(-c_1^2 K_0^2) < 1/2$ and obtain

$$
\begin{aligned}
I_2 &\leq c_{\varepsilon} \sum_{0 \leq m < k} e^{2^m} 2^{(k-m)2^m} \exp(-c_1^2 K_0^2 (k-m)2^m) \\
&= c_{\varepsilon} \sum_{m=0}^{\infty} e^{2^m} \sum_{\nu=1}^{\infty} \Big(2^{2^m} \exp(-c_1^2 K_0^2 2^m)\Big)^{\nu} \leq 2 c_{\varepsilon} \sum_{m=0}^{\infty} 8^{2^m} \exp(-c_1^2 K_0^2 2^m) \\
&\leq 32 c_{\varepsilon} \exp(-c_1^2 K_0^2),
\end{aligned}
$$

which tends again to zero if K_0 grows to infinity. $\qquad\square$

Remark 3 In general, one can take also $\varepsilon = 0$ in Definition 10 and obtain the space $A(0)$. Nevertheless, it is quite easy to see that Theorem 11 fails for $A(0)$ and that the Wiener paths do almost surely not belong to $A(0)$. Indeed, observe that

$$\|f\|_{A(0)} = \sup_{j \in \mathbb{N}_0} \sup_{0 \leq k \leq j} \sup_{0 < t < 1} 2^{(j-k)/2} \cdot \frac{(A_k f_j)^*(t)}{\sqrt{\log(1/t) + 1}} \geq \sup_{j \in \mathbb{N}_0} 2^{j/2} \cdot (A_0 f_j)^*(1).$$

Similarly as in the proof of Theorem 11 we therefore obtain

$$
\begin{aligned}
\mathbb{P}(\|W.\|_{A(0)} \geq K_0) &\geq \mathbb{P}\left(\sup_{j \in \mathbb{N}_0} 2^{j/2} \left|\int_0^1 \Big(\sum_{m=0}^{2^j - 1} \xi_{j,m} \chi_{j,m}(t)\Big) dt\right| \geq K_0\right) \\
&= \mathbb{P}\left(\sup_{j \in \mathbb{N}_0} 2^{-j/2} \left|\sum_{m=0}^{2^j - 1} \xi_{j,m}\right| \geq K_0\right) = \mathbb{P}\big(\sup_{j \in \mathbb{N}_0} |\omega_j| \geq K_0\big) = 1,
\end{aligned}
$$

where $K_0 \geq 0$ is arbitrarily large and $(\omega_j)_{j=0}^{\infty}$ are independent standard Gaussian variables. Note that we have again used the 2-stability of Gaussian variables from Lemma 26.

A similar result can be obtained if we use the absolute averaging operators $\widetilde{A}_k$ instead of A_k. The decay of the averages $\widetilde{A}_k(f_j)$ will now be described by the Orlicz spaces $L_{\Phi_{2,A}}$, which are defined in (67), also cf. Theorem 33.

Definition 12 The space $\widetilde{A}$ is the collection of all $f \in C(I)$, which satisfy the following condition. If (18) is the decomposition of f into the Faber system and f_j, $j \in \mathbb{N}_0$, is defined by (31), then

$$\|f\|_{\widetilde{A}} := \sup_{j \in \mathbb{N}_0} \sup_{0 \leq k \leq j} \sup_{0 < t < 1} \frac{(\widetilde{A}_k f_j)^*(t)}{\sqrt{2^{k-j} \log(1/t) + 1}} \approx \sup_{j \in \mathbb{N}_0} \sup_{0 \leq k \leq j} \|\widetilde{A}_k f_j\|_{2,2^{k-j}} < \infty.$$

With this we can now state and prove the following statement.

Theorem 13 *It holds that* $\|W.\|_{\widetilde{A}} < \infty$ *almost surely.*

Proof By its construction,

$$\widetilde{A}_k f_j(x) = \sum_{m=0}^{2^k-1} v_{k,m} \chi_{k,m}(x), \quad x \in [0,1],$$

where $v_k = (v_{k,0}, \ldots, v_{k,2^k-1})$ is a vector of independent variables from $\mathcal{G}_{2^{j-k}}$, see Definition 28. Similarly to (30), we obtain

$$\sup_{0 < t < 1} \frac{(\widetilde{A}_k f_j)^*(t)}{\sqrt{2^{k-j} \log(1/t) + 1}} = \sup_{m=1,\ldots,2^k} \frac{(\widetilde{A}_k f_j)^*(m2^{-k})}{\sqrt{2^{k-j} \log(2^k/m) + 1}}$$

$$\leq c \sup_{z=0,\ldots,k-1} \frac{(\widetilde{A}_k f_j)^*(2^{z-k})}{\sqrt{2^{k-j}(k-z-1) + 1}} = c \sup_{z=0,\ldots,k-1} \frac{(v_k^*)_{2^z}}{\sqrt{2^{k-j}(k-z-1) + 1}}.$$

Then, for K_0 large enough using the estimate above and (58), we obtain

$$\mathbb{P}(\|W.\|_{\widetilde{A}} \geq K_0) \leq \sum_{0 \leq k \leq j} \mathbb{P}\left(\sup_{0 < t < 1} \frac{(\widetilde{A}_k f_j)^*(t)}{\sqrt{2^{k-j} \log(1/t) + 1}} \geq K_0\right)$$

$$\leq \sum_{0 \leq z < k \leq j} \mathbb{P}\left(\frac{(v_k^*)_{2^z}}{\sqrt{2^{k-j}(k-z-1) + 1}} \geq cK_0\right)$$

$$\leq \sum_{0 \leq z < k \leq j} \left[e \cdot 2^{k-z} \exp(-(2^{k-j}(k-z-1) + 1)c^2 K_0^2 2^{j-k}/4)\right]^{2^z}$$

$$= \sum_{z=0}^{\infty} e^{2^z} \sum_{k=z+1}^{\infty} 2^{(k-z)2^z} \exp\left[-(k-z-1)c^2 K_0^2/4 \cdot 2^z\right]$$

$$\times \sum_{j=k}^{\infty} \exp\left[-2^{j-k}c^2 K_0^2/4 \cdot 2^z\right]$$

$$\leq 2 \sum_{z=0}^{\infty} e^{2^z} \sum_{k=z+1}^{\infty} 2^{(k-z)2^z}$$

$$\times \exp\left[-(k-z-1)c^2 K_0^2/4 \cdot 2^z\right] \exp\left[-c^2 K_0^2/4 \cdot 2^z\right]$$

$$= 2 \sum_{z=0}^{\infty} e^{2^z} \sum_{k=z+1}^{\infty} \left(2^{2^z} \exp\left[-c^2 K_0^2/4 \cdot 2^z\right]\right)^{k-z}$$

$$\leq 4 \sum_{z=0}^{\infty} e^{2^z} \cdot 2^{2^z} \exp\left[-c^2 K_0^2/4 \cdot 2^z\right]$$

$$\leq 4 \sum_{z=0}^{\infty} \left(8 e^{-c^2 K_0^2/4}\right)^{2^z} \leq 8 \cdot 8 e^{-c^2 K_0^2/4}.$$

As the last expression tends to zero if $K_0 \to \infty$, this finishes the proof.   □

Remark 4 We discuss the relation between the Besov-Orlicz space $B_{\Phi_2,\infty}^{1/2}(I)$ and the new function spaces $A(\varepsilon)$ and $\widetilde{A}$. We show in several steps that, for all $\varepsilon > 0$, both $A(\varepsilon)$ and $\widetilde{A}$ are strictly smaller than $B_{\Phi_2,\infty}^{1/2}(I)$ and that $A(\varepsilon)$ and $\widetilde{A}$ are mutually incomparable.

(i) Setting $j = k$ in the definition of $A(\varepsilon)$ or $\widetilde{A}$ and observing that $A_j f_j = f_j$ and $\widetilde{A}_j f_j = |f_j|$, we conclude that $A(\varepsilon)$ and $\widetilde{A}$ are both subsets of the space $B_{\Phi_2,\infty}^{1/2}(I)$ considered by Ciesielski.
    On the other hand, if we put

$$\lambda_{j,m} = \sqrt{\log\left(\frac{2^j}{m+1}\right) + 1}, \quad j \in \mathbb{N}_0, \quad m = 0, \ldots, 2^j - 1,$$

and define f_j by (31), then it follows from $(A_k f_j)(t) = (\widetilde{A}_k f_j)(t) \geq f_k(t)$ for $0 \leq k \leq j$ and $t \in [0, 1]$ that $A(\varepsilon)$ and $\widetilde{A}$ are proper subsets of $B_{\Phi_2,\infty}^{1/2}(I)$.

(ii) To show that $\widetilde{A}$ is not a subset of $A(\varepsilon)$ it is enough to set $\lambda_{j,m} = 1$ for all $j \in \mathbb{N}_0$ and $0 \leq m \leq 2^j - 1$ and define f_j again by (31). Then $A_k f_j = \widetilde{A}_k f_j = 1$ on $[0, 1]$ for all $0 \leq k \leq j$ and for the corresponding f defined by (18) we obtain for every $\varepsilon > 0$

$$\|f\|_{\widetilde{A}} = \sup_{0 \leq k \leq j} (\widetilde{A}_k f_j)^*(1) = 1 \quad \text{and} \quad \|f\|_{A(\varepsilon)}$$

$$= \sup_{0 \leq k \leq j} \frac{2^{(j-k)/2}}{(j-k+1)^\varepsilon} (A_k f_j)^*(1) = +\infty.$$

(iii) Finally, we show that $A(\varepsilon)$ is not a subset of $\widetilde{A}$ for any $\varepsilon > 0$. We put for $j \geq 1$ and $0 \leq m \leq 2^j - 1$

$$
\lambda_{j,m} = \begin{cases} \sqrt{\log\left(\dfrac{2^j}{m+1}\right) + 1} & \text{if } m \text{ is even,} \\[4mm] (-1) \cdot \sqrt{\log\left(\dfrac{2^j}{m}\right) + 1} & \text{if } m \text{ is odd.} \end{cases}
$$

Again, we define f_j by (31) and f by (18). We observe, that due to the cancellation property $A_k f_j = 0$ for all $0 \leq k < j$ and, therefore, f lies in $A(\varepsilon)$. On the other hand, $(\widetilde{A}_k f_j)(m \cdot 2^{-k}) \geq f_j^*(m \cdot 2^{-k})$ for $0 \leq k < j$ and $0 \leq m < 2^k$ by monotonicity and we obtain

$$
\|f\|_{\widetilde{A}} \geq \sup_{k \in \mathbb{N}} \sup_{0 < t < 1} \frac{(\widetilde{A}_k f_{2^k})^*(t)}{\sqrt{2^{-k} \log(1/t) + 1}} \geq \sup_{k \in \mathbb{N}} \frac{f_{2^k}^*(2^{-k})}{\sqrt{2^{-k} \log(2^k) + 1}} = +\infty.
$$

2.5.2 Spaces of Besov Type: Continuous Averages of Differences

The function space $\tilde{A}$ introduced in Definition 12 is somehow difficult to handle. In order to decide whether a continuous function f belongs to $\tilde{A}$, we first have to construct its Faber decomposition (18) and the sequence $\{f_j\}_{j=0}^\infty$, cf. (31). Afterwards, we need to apply the averaging operators $\widetilde{A}_k$ of (32) and, finally, we have to measure the size of $\widetilde{A}_k f_j$ in the corresponding Orlicz space $L_{\Phi_2}(I)$.

Therefore, we investigate if $\widetilde{A}$ could be possibly replaced by a space which is defined more directly, without the detour through the Faber system decomposition. For this, we first note that by Theorem 33

$$
\|f\|_{\widetilde{A}} \approx \sup_{j \in \mathbb{N}_0} \sup_{0 \leq k \leq j} \|\widetilde{A}_k f_j\|_{2, 2^{k-j}},
$$

where $\|\cdot\|_{2, 2^{k-j}}$ is the Orlicz norm introduced in (68). To avoid the use of f_j and $\widetilde{A}_k$, we observe that

$$
f_j(t) = \sum_{m=0}^{2^j-1} \lambda_{j,m} \chi_{j,m}(t) \quad \text{and} \quad \lambda_{j,m} = -2^{j/2} (\Delta_{2^{-j-1}}^2 f)(m \cdot 2^{-j})
$$

gives for $0 \leq k \leq j$

$$
\widetilde{A}_k f_j(x) = \sum_{l=0}^{2^k-1} 2^k \int_{I_{k,l}} |f_j(t)| dt \cdot \chi_{k,l}(x)
$$

$$
= \sum_{l=0}^{2^k-1} \chi_{k,l}(x) \sum_{m=0}^{2^j-1} 2^k \cdot |\lambda_{j,m}| \cdot |I_{k,l} \cap I_{j,m}|
$$

$$= 2^{j/2} \sum_{l=0}^{2^k-1} \chi_{k,l}(x) \cdot \frac{1}{2^{j-k}}$$

$$\times \sum_{m=l\cdot 2^{j-k}}^{(l+1)2^{j-k}-1} |(\Delta^2_{2^{-j-1}}f)(m \cdot 2^{-j})|.$$

We now replace the discrete averages of second order differences by the continuous averages. Before we come to that, we need to complement (11) by the differences restricted to I and set

$$\Delta^2_h f(x) = \begin{cases} f(x+2h) - 2f(x+h) + f(x) & \text{if } \{x, x+h, x+2h\} \subset I, \\ 0 & \text{otherwise.} \end{cases} \tag{33}$$

This paves the way for the following defintion.

Definition 14 Let $f \in C(I)$.

1. Then we define for every $0 \le k \le j$ and every $x \in I$

$$D^2_{j,k}f(x) := \sum_{l=0}^{2^k-1} \chi_{k,l}(x) \frac{1}{2^{-k}} \int_{l\cdot 2^{-k}}^{(l+1)2^{-k}} |\Delta^2_{2^{-j-1}}f(t)|dt,$$

and

$$\mathfrak{D}^2_{j,k}f(x) := \frac{1}{2^{-k}} \int_{x-2^{-k-1}}^{x+2^{-k-1}} |\Delta^2_{2^{-j-1}}f(t)|dt.$$

2. We define

$$\|f\|_D = \sup_{j\in\mathbb{N}_0} \sup_{0\le k\le j} \left\|2^{j/2}D^2_{j,k}f\right\|_{2,2^{k-j}} \tag{34}$$

and

$$\|f\|_{\mathfrak{D}} = \sup_{j\in\mathbb{N}_0} \sup_{0\le k\le j} \left\|2^{j/2}\mathfrak{D}^2_{j,k}f\right\|_{2,2^{k-j}}. \tag{35}$$

Now we are in a position to state and prove the following.

Theorem 15 *(i) Let $f \in C(I)$. Then $\|f\|_D \approx \|f\|_{\mathfrak{D}}$.*
(ii) $\|W.\|_D < \infty$ almost surely.
(iii) $\|W.\|_{\mathfrak{D}} < \infty$ almost surely.

Proof Step 1. Let $x \in I_{k,l}$ for some $l \in \{0, \dots, 2^k-1\}$. Then $I_{k,l} \subset (x-2^{-k}, x+2^{-k})$, which gives $D^2_{j,k}f(x) \le 2\mathfrak{D}^2_{j,k-1}f(x)$. This allows to estimate the terms with $0 < k \le j$ in (34) from above by $\|f\|_{\mathfrak{D}}$.

To estimate also the terms with $0 = k \le j$, we put $g(t) = |\Delta^2_{2^{-j-1}}f(t)|$ and calculate

$$\left\| D_{j,0}^2 f \right\|_{2,2^{-j}} = \left\| \int_0^1 g(t)dt \cdot \chi_{[0,1]} \right\|_{2,2^{-j}}$$

$$= \left\| \int_0^{1/2} g(t)dt \cdot \chi_{[0,1]} \right\|_{2,2^{-j}} + \left\| \int_{1/2}^1 g(t)dt \cdot \chi_{[0,1]} \right\|_{2,2^{-j}}$$

$$\leq C \left\{ \left\| \int_0^{1/2} g(t)dt \cdot \chi_{[0,1/2]} \right\|_{2,2^{-j}} + \left\| \int_{1/2}^1 g(t)dt \cdot \chi_{[1/2,1]} \right\|_{2,2^{-j}} \right\}$$

$$\leq C \left\{ \left\| \mathfrak{D}_{j,0}^2 f(x) \cdot \chi_{[0,1/2]}(x) \right\|_{2,2^{-j}} + \left\| \mathfrak{D}_{j,0}^2 f(x) \cdot \chi_{[1/2,1]}(x) \right\|_{2,2^{-j}} \right\}$$

$$\leq 2C \left\| \mathfrak{D}_{j,0}^2 f \right\|_{2,2^{-j}},$$

which gives that $\|f\|_D \lesssim \|f\|_{\mathfrak{D}}$.

Step 2. Fix $0 \leq k \leq j$ and set again $g(t) = |\Delta_{2^{-j-1}}^2 f(t)|$. If $x \in I_{k,l}$, then $(x - 2^{-k-1}, x + 2^{-k-1}) \subset I_{k,l-1} \cup I_{k,l} \cup I_{k,l+1}$. Therefore,

$$\mathfrak{D}_{j,k}^2 f(x) = \sum_{l=0}^{2^k-1} \mathfrak{D}_{j,k}^2 f(x)\chi_{k,l}(x) \leq \sum_{l=0}^{2^k-1} \frac{1}{2^{-k}} \left\{ \int_{I_{k,l-1}} |g(t)|dt \right.$$

$$\left. + \int_{I_{k,l}} |g(t)|dt + \int_{I_{k,l+1}} |g(t)|dt \right\} \chi_{k,l}(x)$$

We now apply the shift-invariance of the space $L_{2,2^{k-j}}$ and obtain that $\|f\|_{\mathfrak{D}} \lesssim \|f\|_D$.

Step 3. It is clear that (iii) follows from (i) and (ii). Hence, it is enough to prove (ii), i.e., we show $\|W.\|_D < \infty$ almost surely. Although the proof resembles the proof of Theorem 13, we will omit the use of the Faber system in this case. Moreover, in order to avoid technicalities, we first make the following observation. When we use the values of $(W_t)_{t\geq 0}$ also for $t > 1$, we can assume that

$$\Delta_h^2 W(x) = W(x + 2h) - 2W(x + h) + W(x)$$

for every $x, h > 0$. This means that we do not make use of the restriction to I as it appeared in (33), which can make $\|W.\|_D$ only larger.

Furthermore, we distinguish again between $k = 0$ and $k \geq 1$. If $k = 0$, then

$$D_{j,0}^2 W(x) = \chi_I(x) \cdot \int_0^1 |(\Delta_{2^{-j-1}}^2 W)(t)|dt$$

and, using Theorem 33, we see that

$$\sup_{j\geq 0} 2^{j/2} \|D_{j,0}^2 W\|_{2,2^{-j}} = \sup_{j\geq 0} 2^{j/2} \int_0^1 |(\Delta_{2^{-j-1}}^2 W)(t)|dt \cdot \|\chi_I(x)\|_{2,2^{-j}}$$

$$\approx \sup_{j\geq 0} 2^{j/2} \int_0^1 |(\Delta_{2^{-j-1}}^2 W)(t)|dt \qquad (36)$$

$$= \sup_{j \geq 0} 2^{j/2} \sum_{m=0}^{2^{j+1}-1} \int_{m \cdot 2^{-j-1}}^{(m+1)2^{-j-1}} |(\Delta_{2^{-j-1}}^2 W)(t)| dt.$$

If $m \cdot 2^{-j-1} \leq t \leq (m+1)2^{-j-1}$, then we estimate

$$
\begin{aligned}
|(\Delta_{2^{-j-1}}^2 W)(t)| &= |W(t+2^{-j}) - 2W(t+2^{-j-1}) + W(t)| \\
&\leq |W(t+2^{-j}) - W((m+2)2^{-j-1})| + |W((m+2)2^{-j-1}) - W(t+2^{-j-1})| \\
&\quad + |W((m+1)2^{-j-1}) - W(t+2^{-j-1})| + |W(t) - W((m+1)2^{-j-1})|. \quad (37)
\end{aligned}
$$

If we plug this estimate into (36), we obtain four (very similar) terms. We only estimate the first term, since the others can be handled in the same manner. If we set

$$\alpha_m^j = \frac{1}{2^{-j-1}} \int_{m \cdot 2^{-j-1}}^{(m+1)2^{-j-1}} |W(t) - W(m \cdot 2^{-j-1})| dt,$$

we can see that

$$
\begin{aligned}
\sup_{j \geq 0} 2^{j/2} &\sum_{m=0}^{2^{j+1}-1} \int_{m \cdot 2^{-j-1}}^{(m+1)2^{-j-1}} |W(t+2^{-j}) - W((m+2)2^{-j-1})| dt \\
&= \sup_{j \geq 0} 2^{-j/2-1} \sum_{m=0}^{2^{j+1}-1} \alpha_{m+2}^j.
\end{aligned}
$$

By their definition, $\{\alpha_m^j\}_{m \geq 0}$ are independent random variables, all equidistributed with

$$\frac{1}{2^{-j-1}} \int_0^{2^{-j-1}} |W(s)| ds$$

and therefore have the same distribution as $2^{-(j+1)/2} \mathcal{W}$, where $\mathcal{W} = \int_0^1 |W(s)| ds$ is the integrated absolute Wiener process. For the convenience of the reader we recall a few of its basic properties in Sect. 4.3. In particular, Lemma 30 (ii) allows us to conclude for every $K \geq 1$ that

$$
\begin{aligned}
\mathbb{P}\left(\sup_{j \geq 0} 2^{-j/2-1} \sum_{m=0}^{2^{j+1}-1} \alpha_{m+2}^j = \infty \right) &\leq \sum_{j=0}^{\infty} \mathbb{P}\left(\frac{1}{2^{j+1}} \sum_{m=0}^{2^{j+1}-1} 2^{\frac{j+1}{2}} \alpha_m^j > K \right) \\
&\leq \sum_{j=0}^{\infty} \exp(1 - c2^{j+1} K^2),
\end{aligned}
$$

which goes to zero if $K \to \infty$.

Step 4. Next, we estimate the terms with $0 < k \le j$. The argument is quite similar to the previous step, which allows us to leave out some technical details. First for every $k \ge 1$ we rewrite

$$D^2_{j,k}W(x) = \sum_{l=0}^{2^k-1} \chi_{k,l}(x) \cdot \frac{1}{2^{-k}} \int_{l\cdot 2^{-k}}^{(l+1)2^{-k}} |\Delta^2_{2^{-j-1}} W(t)| dt$$

$$= \sum_{l=0}^{2^k-1} \chi_{k,l}(x) \cdot \frac{1}{2^{-k}} \sum_{m=0}^{2^{j-k+1}-1} \int_{l\cdot 2^{-k}+m\cdot 2^{-j-1}}^{l\cdot 2^{-k}+(m+1)2^{-j-1}} |\Delta^2_{2^{-j-1}} W(t)| dt. \quad (38)$$

If $0 \le t - l \cdot 2^{-k} - m \cdot 2^{-j-1} \le 2^{-j-1}$, then we obtain (37) with $m \cdot 2^{-j-1}$ replaced by $\tau := l \cdot 2^{-k} + m \cdot 2^{-j-1}$. We insert this into (38) and derive again an estimate of $D^2_{j,k}W(x)$ invoking a sum of four terms, which we denote by $D^{2,i}_{j,k}W(x)$ with $i \in \{1, 2, 3, 4\}$. Again, we estimate only one of these terms (say, the one with $i = 4$), since the others are very similar to deal with. We put

$$\alpha^{j,k}_{l,m} = \frac{1}{2^{-j-1}} \int_{l\cdot 2^{-k}+m\cdot 2^{-j-1}}^{l\cdot 2^{-k}+(m+1)2^{-j-1}} |W(t) - W(l \cdot 2^{-k} + m \cdot 2^{-j-1} + 2^{-j-1})| dt$$

and obtain

$$D^{2,4}_{j,k}W(x) = \sum_{l=0}^{2^k-1} \chi_{k,l}(x) \cdot \frac{1}{2^{j-k+1}} \sum_{m=0}^{2^{j-k+1}-1} \alpha^{j,k}_{l,m}.$$

As $V_t = W_1 - W_{1-t}$ with $0 \le t \le 1$ is equidistributed with W_t, we observe that

$$\{\alpha^{l,m}_{j,k} : l = 0, \ldots, 2^k - 1, m = 0, \ldots, 2^{j-k+1} - 1\}$$

are independent random variables distributed like $2^{-(j+1)/2}\mathcal{W}$. Hence, if we set

$$B^{j,k}_l = \frac{1}{2^{j-k+1}} \sum_{m=0}^{2^{j-k+1}-1} 2^{(j+1)/2} \alpha^{j,k}_{l,m},$$

then $\{B^{j,k}_l\}_{l=0}^{2^k-1}$ are independent and each of them is distributed as the average of 2^{j-k+1} variables from $\mathcal{W}$. Finally, we observe that $\|2^{j/2} D^{2,4}_{j,k}W\|_{2,2^{k-j}}$ has the same distribution as

$$\frac{1}{\sqrt{2}} \left\| \sum_{l=0}^{2^k-1} \chi_{k,l}(x) B^{j,k}_l \right\|_{2,2^{k-j}}.$$

The proof is then finished in the same manner as the proof of Theorem 13 by using (63) instead of (58). □

We close this section with the following remark and three open problems.

Remark 5 A natural question at this place is, if the method and the results presented so far could also be applied to other processes. The first natural candidate is the *fractional Brownian motion*, which is a Gaussian process $B_H = (B_H(t))_{t \geq 0}$ with $B_H(0) = 0$, $\mathbb{E}B_H(t) = 0$ for all $t \geq 0$ and

$$\mathbb{E}[B_H(t)B_H(s)] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t - s|^{2H}) \quad \text{for all } s, t \geq 0.$$

Here, $H \in (0, 1)$ is the so-called Hurst index. For $H = 1/2$, one recovers the standard Brownian motion with independent increments, but the increments are no longer independent if $H \neq 1/2$.

One can again recover the decomposition of the paths of B_H into the Faber system, but (due to loss of independence of the increments) the random coefficients are no longer independent, see [10, Lemma IV.2]. Although this obstacle can be overcome, the proofs are technically much more involved. That is why we decided to concentrate on the standard Brownian motion in this paper.

Open Problem 1 *All the regularity results obtained so far used Besov spaces and their numerous variants. The other well-known scale of Fourier-analytic function spaces, the so-called* Triebel-Lizorkin spaces, *did not play any important role up to now. For these spaces a characterization by the Faber system is also available and actually quite similar to Theorem 5. The main difference to Besov spaces is that one first applies some sequence space norm to $(f_j(t))_{j \geq 0}$, cf. (29), and only afterwards some function space norm. The analysis of the regularity of Brownian paths in the frame of Triebel-Lizorkin spaces could be based on the observation, that the functions $f_j(t)$ are extremely unlikely to be large for the same value of t.*

Open Problem 2 *It is very well known, that Besov (and Triebel-Lizorkin) spaces can be characterized by differences in several different ways, including also the so-called* ball means of differences, *cf. [57]. Note however, that the usual ball means of differences differ essentially from $D^2_{j,k}f(x)$ and $\mathfrak{D}^2_{j,k}f(x)$ introduced in Definition 14. It would be of some interest to know, if one could build new scales of function spaces in the spirit of (34) and (35) and investigate their relation with the already known function spaces.*

Open Problem 3 *Theorem 15 shows that the norms of a continuous function f in D and $\mathfrak{D}$ are equivalent. It would be also interesting to compare the norm of f in D with its norm in A or $A(\varepsilon)$. As their definitions differ substantially, we leave it as an open problem for future research.*

3 Path Regularity of Brownian Sheets

The aim of this section is to show that the methods presented in Sect. 2 for the univariate Wiener process can be quite easily generalized to the multivariate setting. First, let

us introduce the multivariate analogue of the Wiener process, the so-called Brownian sheet.

Definition 16 A continuous Gaussian process $B = (B(t))_{t \in \mathbb{R}_+^d}$ is called a Brownian sheet, if $\mathbb{E}B(t) = 0$ for every $t = (t_1, \ldots, t_d) \in \mathbb{R}_+^d$ and

$$\text{Cov}(B(s), B(t)) = \prod_{i=1}^{d} \min(s_i, t_i) \quad \text{for all} \quad s, t \in \mathbb{R}_+^d. \tag{39}$$

The study of the Brownian sheet goes back to the 1950s. Since then, many of its properties (including the properties of sample paths) were studied in great detail. For an overview we refer to [30] and [65] and the references given therein. The relation of the Brownian sheet with approximation theory also attracted a lot of attention in connection with the probability estimates of small balls, cf. [4, 5, 14, 32–34, 38, 54]. However, this problem in its full generality remains unsolved up to now.

3.1 Lévy's Decomposition of the Brownian Sheet

We now present the decomposition of the paths of the Brownian sheet into the corresponding Faber system. For this sake, we need to develop the multivariate analogues of the tools of Sect. 2.1, i.e., multivariate Faber systems and second order differences of functions of several variables. We restrict ourselves to the two-dimensional setting, i.e., to the case $d = 2$. This simplifies to some extent the notation used, nevertheless, the general case $d \geq 2$ can be treated in the same way. Hence, we now focus on the Brownian sheet $B = (B(x))_{x \in Q}$ on $Q = [0, 1]^2$.

Similar to (8), we show that the paths of the Brownian sheet can be decomposed into the multivariate Faber system and that the coefficients of this decomposition are again independent Gaussian variables. In order to do this, we follow [58, Section 3.2] and first describe the decomposition of (continuous) functions into the two-dimensional Faber system.

3.1.1 Multivariate Faber System and Second Order Differences

The *multivariate Faber system* is obtained by considering the tensor products of the functions from the univariate Faber system on [0, 1], which was introduced in Sect. 2.2 as

$$\{v_0(t) = 1 - t, v_1(t) = t, v_{j,m}(t) : j \in \mathbb{N}_0, m = 0, \ldots, 2^j - 1\}, \quad t \in [0, 1]. \tag{40}$$

In the sequel we use the notation $\mathbb{N}_{-1} := \mathbb{N}_0 \cup \{-1\} = \{-1, 0, 1, \ldots\}$ and put $v_{-1,0}(t) = v_0(t)$ and $v_{-1,1}(t) = v_1(t)$. We consider the tensor products of the functions in (40)

$$v_{k,m}(t_1, t_2) = v_{k_1,m_1}(t_1) \cdot v_{k_2,m_2}(t_2), \quad (t_1, t_2) \in [0, 1]^2,$$

where $k = (k_1, k_2)$ with $k_i \in \mathbb{N}_{-1}$ and $m = (m_1, m_2)$ with $m_i \in \{0, 1\}$ if $k_i = -1$ and $m_i \in \{0, 1, \dots, 2^{k_i} - 1\}$ if $k_i \in \mathbb{N}_0$ for $i = 1, 2$. Moreover, by $\mathcal{P}_k^F$ we denote the admissible set of m's for given k. Finally, the system

$$\{v_{k,m} : k \in \mathbb{N}_{-1}^2, m \in \mathcal{P}_k^F\}$$

is the Faber system on $[0, 1]^2$.

Similarly to (6), the coefficients of the decomposition of a continuous $f \in C(Q)$ will be the *second order differences* of f. These are defined in a rather straightforward manner.

Definition 17 1. If f is a continuous function on $\mathbb{R}^2$, we define second order differences

$$\Delta_{h,1}^2 f(t_1, t_2) := f(t_1 + 2h, t_2) - 2f(t_1 + h, t_2) + f(t_1, t_2)$$

$$= \Delta_{h,1}^1 (\Delta_{h,1}^1 f)(t_1, t_2) = \sum_{i=0}^{2} (-1)^i \binom{2}{i} f(t_1 + ih, t_2)$$

and similarly for $\Delta_{h,2}^2 f$. The second order mixed differences are defined as

$$\Delta_{h_1, h_2}^{2,2} f(t_1, t_2) := \Delta_{h_2, 2}^2 (\Delta_{h_1, 1}^2 f)(t_1, t_2)$$

$$= \sum_{i,j=1}^{2} (-1)^{i+j} \binom{2}{i} \binom{2}{j} f(t_1 + ih_1, t_2 + ih_2).$$

2. For $m \in \mathcal{P}_k^F$, we put

$$d_{k,m}^2(f) := \begin{cases} f(m_1, m_2), & \text{if } k = (-1, -1), \\ -\frac{1}{2} \Delta_{2^{-k_2-1}, 2}^2 f(m_1, 2^{-k_2} m_2), & \text{if } k = (-1, k_2), \ k_2 \in \mathbb{N}_0, \\ -\frac{1}{2} \Delta_{2^{-k_1-1}, 1}^2 f(2^{-k_1} m_1, m_2), & \text{if } k = (k_1, -1), \ k_1 \in \mathbb{N}_0, \\ \frac{1}{4} \Delta_{2^{-k_1-1}, 2^{-k_2-1}}^{2,2} f(2^{-k_1} m_1, 2^{-k_2} m_2), & \text{if } k = (k_1, k_2), \ k_1, k_2 \in \mathbb{N}_0. \end{cases} \tag{41}$$

The two-dimensional analogue of Theorem 2 then reads as follows.

Theorem 18 ([58, Thm. 3.10]) *For $f \in C([0, 1]^2)$ it holds*

$$f(t) = \sum_{k \in \mathbb{N}_{-1}^2} \sum_{m \in \mathcal{P}_k^F} d_{k,m}^2(f) v_{k,m}(t)$$

$$= \lim_{K \to \infty} \sum_{k \in \{-1, 0, \dots, K\}^2} \sum_{m \in \mathcal{P}_k^F} d_{k,m}^2(f) v_{k,m}(t), \quad t \in [0, 1]^2, \tag{42}$$

where the limit is taken in the uniform norm.

3.1.2 Decomposition of Paths of the Brownian Sheet

Similarly to (8) we can apply (42) to paths of the Brownian sheet $B = (B(x))_{x \in Q}$ if we replace the scalars $d_{k,m}^2(f)$ by random variables $d_{k,m}^2(B)$. First, observe that (39) ensures that $B(0, t) = B(s, 0) = 0$ almost surely for every $s, t \geq 0$, which implies that almost surely we also have

$$d_{k,m}^2(B) = 0 \quad \text{if} \quad \begin{cases} k = (-1, -1) \quad \text{and} \quad m = (m_1, m_2) \in \{(0, 0), (0, 1), (1, 0)\}, \\ k = (-1, k_2) \quad \text{if} \quad k_2 \in \mathbb{N}_0 \quad \text{and} \quad m_1 = 0, \\ k = (k_1, -1) \quad \text{if} \quad k_1 \in \mathbb{N}_0 \quad \text{and} \quad m_2 = 0. \end{cases}$$
(43)

Furthermore, all the random variables $d_{k,m}^2(B)$ are Gaussian with mean zero. If $k_1 = -1$ or $k_2 = -1$, then their variance can be computed directly as

$$\operatorname{var} d_{(-1,-1),(1,1)}^2(B) = \operatorname{var} B(1, 1) = 1,$$

$$\begin{aligned} \operatorname{var} d_{(-1,k_2),(1,m_2)}^2(B) &= \operatorname{var}\left(-\frac{1}{2}\Delta_{2^{-k_2-1},2}^2 B(1, 2^{-k_2}m_2)\right) \\ &= \frac{1}{4}\operatorname{var}\Big(B(1, 2^{-k_2}(m_2 + 1)) - 2B(1, 2^{-k_2}(m_2 + 1/2)) \\ &\quad + B(1, 2^{-k_2}m_2)\Big) \\ &= 2^{-k_2-2}, \end{aligned}$$

and similarly, we obtain $\operatorname{var} d_{(k_1,-1),(m_1,1)}^2(B) = 2^{-k_1-2}$.

In order to calculate the variance of $d_{k,m}^2(B)$ for $k_1, k_2 \in \mathbb{N}_0$, we introduce some further notation. For a cube $\tilde{Q} = [s_1, s_2] \times [t_1, t_2]$, where $0 \leq s_1 \leq s_2 \leq 1$ and $0 \leq t_1 \leq t_2 \leq 1$, we put

$$\begin{aligned} B(\tilde{Q}) &:= B(s_2, t_2) - B(s_1, t_2) - B(s_2, t_1) + B(s_1, t_1) \\ &= \Delta_{s_2-s_1,1}^1(\Delta_{t_2-t_1,2}^1 B)(s_1, t_1) = \sum_{i=1}^{2}\sum_{j=1}^{2}(-1)^{i+j} B(s_i, t_j). \end{aligned}$$
(44)

Then $B(\tilde{Q})$ is a centered Gaussian variable with variance

$$\begin{aligned} \operatorname{var} B(\tilde{Q}) &= \mathbb{E}B(\tilde{Q})^2 = \sum_{i,j,k,l=1}^{2}(-1)^{i+j+k+l}\mathbb{E}\left[B(s_i, t_j)B(s_k, t_l)\right] \\ &= \sum_{i,j,k,l=1}^{2}(-1)^{i+j+k+l}\min(s_i, s_k)\min(t_j, t_l) \\ &= \sum_{i,k=1}^{2}(-1)^{i+k}\min(s_i, s_k)\sum_{j,l=1}^{2}(-1)^{j+l}\min(t_j, t_l) = (s_2 - s_1)(t_2 - t_1), \end{aligned}$$

where in the third step we used (39). Furthermore, if $\tilde{Q}_1 = [s_1, s_2] \times [t_1, t_2]$ and $\tilde{Q}_2 = [\sigma_1, \sigma_2] \times [\tau_1, \tau_2]$ are disjoint, then we obtain in the same way

$$\mathbb{E}B(\tilde{Q}_1)B(\tilde{Q}_2) = \sum_{i,k=1}^{2} (-1)^{i+k} \min(s_i, \sigma_k) \sum_{j,l=1}^{2} (-1)^{j+l} \min(t_j, \tau_l) = 0.$$

Actually, $B(\tilde{Q}_1)$ and $B(\tilde{Q}_2)$ are not only uncorrelated but also independent, cf. [12, Sect. 2.4, Prop. 1]. For $k \in \mathbb{N}_0^2$ and $m \in \mathcal{P}_k^F$, we define

$$Q_{k,m} = I_{k_1,m_1} \times I_{k_2,m_2} = [2^{-k_1}m_1, 2^{-k_1}(m_1 + 1)] \times [2^{-k_2}m_2, 2^{-k_2}(m_2 + 1)]$$

and using this notation, we compute for $k_1, k_2 \geq 0$,

$$\begin{aligned}
4d_{k,m}^2(B) &= \Delta_{2^{-k_1-1}, 2^{-k_2-1}}^{2,2} B(2^{-k_1}m_1, 2^{-k_2}m_2) \\
&= B(Q_{(k_1+1,k_2+1),(2m_1+1,2m_2+1)}) - B(Q_{(k_1+1,k_2+1),(2m_1+1,2m_2)}) \\
&\quad - B(Q_{(k_1+1,k_2+1),(2m_1,2m_2+1)}) + B(Q_{(k_1+1,k_2+1),(2m_1,2m_2)}). \qquad (45)
\end{aligned}$$

Since the four summands on the right hand side are independent Gaussian variables with variance $2^{-(k_1+k_2)-2}$, we obtain that $d_{k,m}^2(B)$ is a Gaussian variable with variance $2^{-(k_1+k_2+4)}$.

Merging all what we said about $d_{k,m}^2(B)$ so far, we arrive at the decomposition of the paths of the Brownian sheet into the multivariate Faber system. The discussion of independence is postponed to Sect. 3.1.3.

Theorem 19 *For the Brownian sheet on $[0, 1]^2$ it holds almost surely that*

$$\begin{aligned}
B(t_1, t_2) &= \xi_{(-1,-1),(1,1)}t_1t_2 + \sum_{k_1=0}^{\infty} \sum_{m_1=0}^{2^{k_1}-1} 2^{-(k_1+2)/2}\xi_{(k_1,-1)(m_1,1)}v_{k_1,m_1}(t_1)t_2 \\
&\quad + \sum_{k_2=0}^{\infty} \sum_{m_2=0}^{2^{k_2}-1} 2^{-(k_2+2)/2}\xi_{(-1,k_2)(1,m_2)}t_1v_{k_2,m_2}(t_2) \\
&\quad + \sum_{k\in\mathbb{N}_0^2} \sum_{m_1=0}^{2^{k_1}-1} \sum_{m_2=0}^{2^{k_2}-1} 2^{-(k_1+k_2+4)/2}\xi_{k,m}v_{k,m}(t_1, t_2), \qquad (46)
\end{aligned}$$

where $\{\xi_{k,m} : k \in \mathbb{N}_{-1}^2, m \in \mathcal{P}_k^F\}$ are independent standard Gaussian variables and the series converges uniformly on $[0, 1]^2$.

Remark 6 If $k \in \mathbb{N}^2_{-1}$ and $m \in \mathcal{P}^F_k$, then we denote $\gamma_{k,m} = \gamma_{(k_1,k_2),(m_1,m_2)} = \gamma_{k_1,m_1} \cdot \gamma_{k_2,m_2}$, where

$$\gamma_{j,l} = \begin{cases} 0 & \text{if } j = -1 \text{ and } l = 0, \\ 1 & \text{if } j = -1 \text{ and } l = 1, \\ 2^{-(j+2)/2} & \text{if } j \geq 0. \end{cases}$$

This allows us to reformulate (46) as

$$B(t_1, t_2) = \sum_{k \in \mathbb{N}^2_{-1}} \sum_{m \in \mathcal{P}^F_k} \gamma_{k,m} \xi_{k,m} v_{k,m}(t_1, t_2). \tag{47}$$

3.1.3 Independence

Next, we show that the random variables $\{d^2_{k,m}(B) : k \in \mathbb{N}^2_{-1}, m \in \mathcal{P}^F_k\}$, which appear in Theorem 19, are indeed independent. The argument is a tensor product variant of the proof given for the univariate Wiener process, cf. Theorem 3. For that sake, let us consider mutually different $(k^1, m^1), \ldots, (k^N, m^N)$ with $k^i = (k^i_1, k^i_2)$ and $m^i = (m^i_1, m^i_2)$. Let

$$K = (K_1, K_2), \quad \text{where } K_j = \max\{k^1_j, \ldots, k^N_j\} \text{ for } j \in \{1, 2\},$$

and let us consider the array of Gaussian variables

$$\widetilde{B}^K = (B(Q_{K+1,m}) : m = (m_1, m_2) \text{ and } 0 \leq m_j \leq 2^{K_j} - 1 \text{ for } j \in \{1, 2\}),$$

where $B(\tilde{Q})$ was defined in (44) for a closed cube $\tilde{Q} \subset [0, 1]^2$ with sides parallel to the coordinate axes.

Moreover, let $I, J \subset [0, 1]$ be two (closed) intervals, where $I = I_1 \cup I_2$ and $J = J_1 \cup J_2$ are decompositions of I and J into two intervals I_1, I_2 and J_1, J_2, respectively, which intersect only at one point. Then a straightforward calculation shows that

$$B(I \times J) = B((I_1 \cup I_2) \times J) = B(I_1 \times J) + B(I_2 \times J) \tag{48}$$

$$B(I \times J) = B(I \times (J_1 \cup J_2)) = B(I \times J_1) + B(I \times J_2). \tag{49}$$

Furthermore, we denote

$$\left(h^{(K_1,K_2)}_{(k_1,k_2),(m_1,m_2)}\right)_{(l_1,l_2)} = \left(h^{K_1}_{k_1,m_1}\right)_{l_1} \cdot \left(h^{K_2}_{k_2,m_2}\right)_{l_2}, \tag{50}$$

where the vectors on the right-hand side were defined in (9) for $k_1, k_2 \geq 0$ and we complement this definition by putting $h^{K_j}_{-1,1} = (1, \ldots, 1)^T$ for $j = 1, 2$.

If $k = (k_1, k_2) = (-1, -1)$, we observe that by (43) it is enough to consider $m = (1, 1)$. Then we apply (41), (48) and (49) together with (50) and obtain

$$d^2_{(-1,-1),(1,1)}(B) = B(Q) = \sum_{l_1=0}^{2^{K_1+1}-1} \sum_{l_2=0}^{2^{K_2+1}-1} B(Q_{K+1,l}) = \langle h^K_{k,m}, \widetilde{B}^K \rangle.$$

If $k = (k_1, k_2) = (-1, k_2)$ with $0 \le k_2 \le K_2$, we assume by (43) that $m = (1, m_2)$, where $0 \le m_2 \le 2^{k_2} - 1$. Then we combine (41), (48), (49), and (50) and obtain

$$-2d^2_{(-1,k_2),(1,m_2)}(B) = B([0, 1] \times I_{k_2+1,2m_2+1}) - B([0, 1] \times I_{k_2+1,2m_2}) = \langle h^K_{k,m}, \widetilde{B}^K \rangle.$$

The case of $k = (k_1, -1)$ is treated similarly. And finally, if $k = (k_1, k_2) \in \mathbb{N}_0^2$ with $0 \le k_1 \le K_1$ and $0 \le k_2 \le K_2$, we employ (45) and observe that

$$4d^2_{k,m} = \langle h^K_{k,m}, \widetilde{B}^K \rangle.$$

The independence of $\left(d^2_{k^i,m^i}(B)\right)_{i=1}^N$ now follows again by the orthogonality of $\left(h^K_{k^i,m^i}\right)_{i=1}^N$, which in turn is a consequence of (50).

3.2 Function Spaces

As already mentioned before, the appropriate function spaces, which best capture the regularity of the paths of the Brownian sheet, are the function spaces of dominating mixed smoothness. They appeared for the first time in the work of Babenko [1], but they are also known to play an important role in approximation theory and numerics of PDE's [6, 15, 52, 55]. In this context we also refer to [19, 49, 59, 60, 64] (and the references given therein) for a comprehensive treatment.

Similarly to the univariate case, we shall need several different variants of function spaces of dominating mixed smoothness. As before, we start with the spaces of Besov type.

Definition 20 Let $0 < r < l \in \mathbb{N}$ and $1 \le p, q \le \infty$. Then $S^r_{pq}B(\mathbb{R}^2)$ is the collection of all $f \in L_p(\mathbb{R}^2)$ such that

$$\|f|L_p(\mathbb{R}^2)\| + \left(\int_0^1 t^{-rq} \sup_{|h_1| \le t} \|\Delta^l_{h_1,1} f|L_p(\mathbb{R}^2)\|^2 \frac{dt}{t}\right)^{1/q}$$

$$+ \left(\int_0^1 t^{-rq} \sup_{|h_2| \le t} \|\Delta^l_{h_2,2} f|L_p(\mathbb{R}^2)\|^2 \frac{dt}{t}\right)^{1/q}$$

$$+ \left(\int_0^1 \int_0^1 (t_1 t_2)^{-rq} \sup_{|h_1| \le t_1, |h_2| \le t_2} \|\Delta^{l,l}_{h_1,h_2} f|L_p(\mathbb{R}^2)\|^q \frac{dt_1 dt_2}{t_1 t_2}\right)^{1/q}$$

is finite. Here we use the usual mixed version of differences resulting from Definition 17, i.e.,

$$\Delta_{h_1,h_2}^{l,l} f(t_1, t_2) = \Delta_{h_2,2}^{l}\left(\Delta_{h_1,1}^{l} f\right)(t_1, t_2) \quad \text{and} \quad \Delta_{h_i,i}^{l+1} f(x) = \Delta_{h_i,i}^{1}\left(\Delta_{h_i,i}^{l} f\right)(x).$$

The restriction of the spaces from $\mathbb{R}^2$ to $Q = [0, 1]^2$ is done in the same way as described in Sect. 2.2, cf. (12).

We now provide a decomposition of functions f from $S_{pq}^r B(Q)$ via the higher dimensional Faber system introduced in Sect. 3.1. This follows from Theorem 3.16 of [58] and its extension [7, Thm. 4.25], which ensure the following two-dimensional counterpart of Theorem 5, cf. also [27, Theorem A].

Theorem 21 *Let $0 < p, q \leq \infty$, $p > \frac{1}{2}$, and*

$$\frac{1}{p} < r < 1 + \frac{1}{p}$$

be the admissible range for r. Then $f \in L_1(Q)$ lies in $S_{p,q}^r B(Q)$ if, and only if, it can be represented (with convergence in $L_1(Q)$) as

$$f = \sum_{k \in \mathbb{N}_{-1}^2} \sum_{m \in \mathcal{P}_k^F} \lambda_{k,m} 2^{-(k_1+k_2)r} v_{k,m}, \tag{51}$$

with

$$\|\lambda | s_{p,q}^F b(Q)\| = \left(\sum_{k \in \mathbb{N}_{-1}^2} \left(\sum_{m \in \mathcal{P}_k^F} 2^{-(k_1+k_2)} |\lambda_{k,m}|^p \right)^{q/p} \right)^{1/q} < \infty. \tag{52}$$

Furthermore, the representation (51) is unique with

$$\lambda_{k,m} = \lambda_{k,m}(f) = 2^{(k_1+k_2)r} d_{k,m}^2(f), \quad k \in \mathbb{N}_{-1}^2, \, m \in \mathcal{P}_k^F.$$

The dominating mixed smoothness counterpart of the one dimensional spaces $B_{p,q}^{s,\alpha}(I)$ with logarithmic smoothness has been studied in [58]. Unfortunately, we can not rely exclusively on the results of [58] for two reasons. First, the characterization with the Faber system presented in Theorem 3.35 in [58] does not include the case $q = \infty$, which is important for our considerations. Furthermore, it will turn out later, that the way we introduce the logarithmic smoothness differs from [58]. Indeed, the factor $(k_1 \cdot k_2)^{-\alpha}$ used in this reference, gets replaced by $(k_1 + k_2)^{-\alpha}$, which is strictly larger for $k = (k_1, k_2) \in \mathbb{N}^2$ and $\alpha > 0$.

However, the function spaces of dominating logarithmic smoothness, which we introduce in Definition 22, coincide with the ones used in [28]. There these spaces were defined by differences and moduli of smoothness. Moreover, in [28, Lemma 3.1] (based [27]) the author further obtained isomorphisms between these function spaces and corresponding sequence spaces. An equivalent Fourier-analytic characterization

of these spaces still seems to be missing. In order to avoid the technicalities we *define* the function spaces by posing a condition on the coefficients appearing in the Faber system expansion (51).

Essentially the same comment applies to Besov-Orlicz spaces of dominating mixed smoothness. Except [28], we are not aware of any existing work, which introduces and systematically studies these function spaces. Nevertheless, comparing the one dimensional case with Theorems 21 and 3.35 in [58], the following definition seems to be well-motivated and a natural generalization. Again we restrict ourselves to the case of parameters which we need later on, i.e. $q = \infty$ and $r = 1/2$.

Let us recall, that $\chi_{j,l}$ was defined as the characteristic function of $I_{j,l}$ for $j \geq 0$. We complement this notation by $I_{-1,0} = I_{-1,1} = [0, 1]$ and put

$$\chi_{k,m}(t_1, t_2) = \chi_{k_1,m_1}(t_1) \cdot \chi_{k_2,m_2}(t_2), \quad k = (k_1, k_2) \in \mathbb{N}_{-1}^2, \ t = (t_1, t_2) \in Q.$$

Definition 22 1. Let $0 < p \leq \infty$ and $\alpha \in \mathbb{R}$. The function space $S_{p,\infty}^{1/2,\alpha} B(Q)$ is the collection of all $f \in C(Q)$ which can be represented by (51) with $r = 1/2$ and

$$\|\lambda|s_{p,\infty}^{F,\alpha} b(Q)\| = \sup_{k \in \mathbb{N}_{-1}^2} \max(1, k_1 + k_2)^{-\alpha} \Big(\sum_{m \in \mathcal{P}_k^F} 2^{-(k_1+k_2)} |\lambda_{k,m}|^p \Big)^{1/p} < \infty. \tag{53}$$

2. The function space $S_{\Phi_2,\infty}^{1/2} B(Q)$ is the collection of all $f \in C(Q)$ which can be represented by (51) with $r = 1/2$ and

$$\|\lambda|s_{\Phi_2,\infty}^F b(Q)\| = \sup_{k \in \mathbb{N}_{-1}^2} \Big\| \sum_{m \in \mathcal{P}_k^F} \lambda_{k,m} \chi_{k,m}(\cdot) \Big\|_{\Phi_2} < \infty. \tag{54}$$

Remark 7 If in (53) we had used the usual logarithmic Hölder spaces from [58, Theorem 3.35] the sequence space norm would be

$$\sup_{k \in \mathbb{N}_{-1}^2} (2 + k_1)^{-\alpha} (2 + k_2)^{-\alpha} \Big(\sum_{m \in \mathcal{P}_k^F} 2^{-(k_1+k_2)} |\lambda_{k,m}|^p \Big)^{1/p} < \infty.$$

Compared to this the advantage of our approach in (53) is, that the sequence spaces are strictly smaller for $\alpha > 0$ and we therefore obtain better regularity results. The disadvantage on the other hand is, that the smoothness weights in (53) do not have any tensor product structure.

Open Problem 4 *For decades the role of function spaces of dominating mixed smoothness $S_{p,q}^r B(Q)$ in approximation theory has been studied intensively. We refer to [15] for a recent overview on various results and a list of open problems in this field. Compared to that, much less seems to be known regarding function spaces with logarithmic*

smoothness and Besov-Orlicz spaces as introduced before. In our opinion, it would be worth investigating these spaces together with their embeddings, which in turn might shed new light on (some of) the open problems in [15].

3.3 Results of Kamont

We now recover the results of Kamont [28] (cf. also [22]) on the regularity of the sample paths of Brownian sheets. We combine the representation in (47) (or (46)) with Theorem 21 and Definition 22 in order to show that the sample paths almost surely lie in $S_{\infty,\infty}^{1/2,1/2}B(Q)$, $S_{p,\infty}^{1/2}B(Q)$ for all $1 \le p < \infty$, and $S_{\Phi_2,\infty}^{1/2}B(Q)$. Similar to the method used in Sect. 2.3, we just need to verify that the condition on the coefficients in the Faber system decomposition is fulfiled almost surely, if we replace λ in (52), (53), and (54) by a sequence of independent normal variables ξ.

For the terms with $k_1 = -1$ or $k_2 = -1$, the conditions (52), (53), and (54) reduce to their one-dimensional counterparts, which were already discussed in Sect. 2.3, cf. also (43). Furthermore, the same is true for the terms with $k_1 = 0$ or $k_2 = 0$. Therefore, it will be enough to handle the terms with $k = (k_1, k_2) \in \mathbb{N}^2$.

1. Paths of the Brownian sheet almost surely lie in $S_{\infty,\infty}^{1/2,1/2}B(Q)$

In view of Definition 22, we need to show that

$$\sup_{k\in\mathbb{N}^2} \frac{1}{\sqrt{k_1 + k_2}} \sup_{m_1=0,\ldots,2^{k_1}-1} \sup_{m_2=0,\ldots,2^{k_2}-1} |\xi_{k,m}| < \infty \quad \text{almost surely},$$

where $\{\xi_{k,m} : k \in \mathbb{N}^2, 0 \le m_1 \le 2^{k_1}-1, 0 \le m_2 \le 2^{k_2}-1\}$ are independent standard Gaussian variables. For this sake, we define the event A_k^N as

$$\sup_{m_1=0,\ldots,2^{k_1}-1} \sup_{m_2=0,\ldots,2^{k_2}-1} |\xi_{k,m}| > N\sqrt{k_1 + k_2}.$$

Similar to (23), we obtain that

$$\mathbb{P}(A_k^N) \le 2^{k_1+k_2} e^{-N^2(k_1+k_2)/2}, \quad k \in \mathbb{N}^2,$$

and following (24), for every $N_0 \ge 1$ we get

$$\mathbb{P}\left(\sup_{k\in\mathbb{N}^2} \frac{1}{\sqrt{k_1 + k_2}} \sup_{m_1=0,\ldots,2^{k_1}-1} \sup_{m_2=0,\ldots,2^{k_2}-1} |\xi_{k,m}| = \infty\right) = \mathbb{P}\left(\bigcap_{N=1}^{\infty} \bigcup_{k\in\mathbb{N}^2} A_k^N\right)$$

$$\le \mathbb{P}\left(\bigcup_{k\in\mathbb{N}^2} A_k^{N_0}\right) \le \sum_{k\in\mathbb{N}^2} \mathbb{P}(A_k^{N_0}) \le \sum_{k\in\mathbb{N}^2} 2^{k_1+k_2} e^{-N_0^2(k_1+k_2)/2}.$$

As the last sum again tends to zero if $N_0 \to \infty$, the proof is finished.

2. Paths of the Brownian sheet almost surely lie in $S_{p,\infty}^{1/2}B(Q)$ for every $1 \le p < \infty$

In order to be able to apply Theorem 21 for $r = 1/2$, we assume that $2 < p < \infty$. Then the smaller values of p follow easily by monotonicity of the function spaces with

dominating mixed smoothness on domains with respect to the integrability parameter p. In this case (52) for $k \in \mathbb{N}^2$ reduces to

$$\sup_{k \in \mathbb{N}^2} \frac{1}{2^{k_1+k_2}} \sum_{m_1=0}^{2^{k_1}-1} \sum_{m_2=0}^{2^{k_2}-1} |\xi_{k,m}|^p < \infty \quad \text{almost surely.}$$

We denote again by μ_p the p absolute moment of a standard Gaussian variable. Furthermore, for $t > 0$ and $k \in \mathbb{N}^2$, we denote by A_k^t the event

$$\frac{1}{2^{k_1+k_2}} \sum_{m_1=0}^{2^{k_1}-1} \sum_{m_2=0}^{2^{k_2}-1} |\xi_{k,m}|^p - \mu_p \geq t.$$

By (26) applied to $j = k_1 + k_2$, we observe that

$$\mathbb{P}(A_k^t) \leq \frac{1}{t^2} \cdot \frac{\mu_{2p} - \mu_p^2}{2^{k_1+k_2}}.$$

Finally, we conclude, that for every $N_0 \in \mathbb{N}$ it holds

$$\mathbb{P}\left(\sup_{k \in \mathbb{N}^2} \frac{1}{2^{k_1+k_2}} \sum_{m_1=0}^{2^{k_1}-1} \sum_{m_2=0}^{2^{k_2}-1} |\xi_{k,m}|^p = \infty \right)$$

$$= \mathbb{P}\left(\bigcap_{N=1}^{\infty} \bigcup_{k \in \mathbb{N}^2} A_k^N \right) \leq \mathbb{P}\left(\bigcup_{k \in \mathbb{N}^2} A_k^{N_0} \right) \leq \sum_{k \in \mathbb{N}^2} \mathbb{P}(A_k^{N_0})$$

$$\leq \frac{\mu_{2p} - \mu_p^2}{N_0^2} \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \frac{1}{2^{k_1+k_2}} = \frac{4(\mu_{2p} - \mu_p^2)}{N_0^2}.$$

The last expression tends to zero if $N_0 \to \infty$, which renders the result.

3. Paths of the Brownian sheet almost surely lie in $S_{\Phi_2,\infty}^{1/2} B(Q)$

Again, it is enough to show that (54) is finite almost surely if we restrict ourselves to $k \in \mathbb{N}^2$ and replace $\lambda_{k,m}$ by independent normal variables $\xi_{k,m}$. Therefore, we put

$$f_k(t) = \sum_{m_1=0}^{2^{k_1}-1} \sum_{m_2=0}^{2^{k_2}-1} \xi_{k,m} \chi_{k,m}(t), \quad k = (k_1, k_2) \in \mathbb{N}^2 \quad \text{and} \quad t \in Q \qquad (55)$$

and show that

$$\sup_{k \in \mathbb{N}^2} \|f_k\|_{\Phi_2} < \infty \quad \text{almost surely.}$$

Again we use the characterization of $L_{\Phi_2}(Q)$ with the non-increasing rearrangement from Theorem 31. As in Sect. 2.4 we have

$$f_k^*(s) = \sum_{m=0}^{2^{k_1+k_2}-1} (\xi_k)_{m+1}^* \chi_{k_1+k_2,m}(s) \quad \text{for } 0 < s < 1 \tag{56}$$

and using (30) we can estimate

$$\|f_k\|_{\Phi_2} \leq c \sup_{0 \leq u < k_1+k_2} \frac{f_k^*(2^{u-(k_1+k_2)})}{\sqrt{k_1+k_2-u}} = c \sup_{0 \leq u < k_1+k_2} \frac{(\xi_k)_{2^u}^*}{\sqrt{k_1+k_2-u}},$$

where $\xi_k = (\xi_{k,m} : 0 \leq m_1 \leq 2^{k_1}-1, 0 \leq m_2 \leq 2^{k_2}-1)$ and $\left((\xi_k)_m^*\right)_{m=1}^{2^{k_1+k_2}}$ is its non-increasing rearrangement. Then, by Lemma 27 we get for every K_0 large enough

$$\mathbb{P}\left(\sup_{k \in \mathbb{N}^2} \|f_k\|_{\Phi_2} = \infty\right) \leq \sum_{0 \leq u < k_1+k_2 < \infty} \left(2e^{-K_0^2/2}\right)^{(k_1+k_2-u)2^u} \cdot e^{2^u}$$

$$\leq c \sum_{u=0}^{\infty} e^{2^u} \sum_{l=u+1}^{\infty} l \cdot \left(2e^{-K_0^2/2}\right)^{(l-u)\cdot 2^u} \leq c \sum_{u=0}^{\infty} (u+1)e^{2^u} \cdot \left(2e^{-K_0^2/2}\right)^{2^u},$$

which again goes to zero if $K_0 \to \infty$.

Remark 8 At this point, we can observe how versatile the method with the non-increasing rearrangement is performing. As input we have given a two dimensional function in (55) and using the rearrangement we obtain in (56) a one-dimensional function similar to the one from Sect. 2.4. Now the remaining estimates done above are an analogue of the estimates of the case $d = 1$ with the only difference that k is replaced by $k_1 + k_2$.

Clearly, a generalization to $d > 2$ follows directly from the considerations above - as long as a characterization in terms of a Faber basis as in Theorem 19 is given.

3.4 New Function Spaces

We also can carry over the one-dimensional approach of averaging operators to the two-dimensional setting. To that end, we add to (32) the extra case

$$(A_{-1}g)(x) = g(x) \quad \text{for all } x \in I.$$

Now, for an integrable function g on Q we can introduce the averaging operators $A_k g$ for all $k = (k_1, k_2) \in \mathbb{N}_{-1}^2$ by

$$(A_k g)(x) := A_{k_1}^1\left(A_{k_2}^2 g(\cdot, x_2)\right)(x_1) \quad \text{and} \quad (\tilde{A}_k g)(x) = (A_k(|g|))(x), \quad x \in Q,$$

where the upper index stands for the dimension in which the one-dimensional averaging operator is applied. In the non-borderline case when $k \in \mathbb{N}_0^2$, the operator is given by

$$(A_k g)(x) = \sum_{l_1=0}^{2^{k_1}-1} \sum_{l_2=0}^{2^{k_2}-1} 2^{k_1+k_2} \int_{Q_{k,l}} g(t)dt \cdot \chi_{k,l}(x),$$

where $k \in \mathbb{N}_0^2$, $Q_{k,l} = I_{k_1,l_1} \times I_{k_2,l_2}$, and $\chi_{k,l}(t_1, t_2) = \chi_{k_1,l_1}(t_1)\chi_{k_2,l_2}(t_2)$.

With these preparations we can define the spaces $A_\varepsilon(Q)$ and $\widetilde{A}(Q)$ as follows.

Definition 23 Let $f \in C(Q)$ be represented in the Faber system by

$$f = \sum_{k \in \mathbb{N}_{-1}^2} \sum_{m \in \mathcal{P}_k^F} \lambda_{k,m} 2^{-(k_1+k_2)\frac{1}{2}} v_{k,m},$$

and define f_j by

$$f_j(t) = \sum_{m \in \mathcal{P}_j^F} \lambda_{j,m} \chi_{j,m}(t), \quad j = (j_1, j_2) \in \mathbb{N}_{-1}^2 \quad \text{and} \quad t \in Q.$$

1. For $\varepsilon > 0$ the space $A_\varepsilon(Q)$ is the collection of all $f \in C(Q)$ with

$$\|f\|_{A_\varepsilon} := \sup_{j \in \mathbb{N}_{-1}^2} \sup_{-1 \le k_1 \le j_1} \sup_{-1 \le k_2 \le j_2} \sup_{0 < t < 1} \frac{2^{(j_1+j_2-(k_1+k_2))/2}}{(j_1 + j_2 - (k_1 + k_2) + 1)^\varepsilon} \cdot \frac{(A_k f_j)^*(t)}{\sqrt{\log(1/t) + 1}}$$

$$\approx \sup_{j \in \mathbb{N}_{-1}^2} \sup_{-1 \le k_i \le j_i} \frac{2^{(j_1+j_2-(k_1+k_2))/2}}{(j_1 + j_2 - (k_1 + k_2) + 1)^\varepsilon} \cdot \|A_k f_j\|_{\Phi_2} < \infty.$$

2. The space $\widetilde{A}(Q)$ is the collection of all $f \in C(Q)$ with

$$\|f\|_{\widetilde{A}} := \sup_{j \in \mathbb{N}_{-1}^2} \sup_{-1 \le k_1 \le j_1} \sup_{-1 \le k_2 \le j_2} \sup_{0 < t < 1} \frac{(\widetilde{A}_k f_j)^*(t)}{\sqrt{2^{k_1+k_2-(j_1+j_2)} \log(1/t) + 1}}$$

$$\approx \sup_{j \in \mathbb{N}_{-1}^2} \sup_{-1 \le k_i \le j_i} \|\widetilde{A}_k f_j\|_{2, 2^{k_1+k_2-(j_1+j_2)}} < \infty.$$

Using this definition we are now able to prove the following result.

Theorem 24 *Let B be the Brownian sheet according to Definition 16 and $\varepsilon > 0$.*

(i) It holds that $\|B(\cdot)\|_{A_\varepsilon} < \infty$ almost surely.
(ii) It holds that $\|B(\cdot)\|_{\widetilde{A}} < \infty$ almost surely.

Proof For both assertions we use that according to Theorem 19 we have the representation needed for the Brownian sheet. Furthermore, we only have to treat the cases where

$j = (j_1, j_2) \in \mathbb{N}_0^2$ since the other cases follow directly from the one-dimensional case. For $j = (j_1, j_2) \in \mathbb{N}_0^2$ and $t = (t_1, t_2) \in Q$ we set

$$f_j(t) = \sum_{m_1=0}^{2^{j_1}-1} \sum_{m_2=0}^{2^{j_2}-1} \xi_{j,m} \chi_{j,m}(t),$$

with independent normal variables $\xi_{j,m}$.

In order to establish the first assertion it is enough to observe that $2^{[j_1+j_2-(k_1+k_2)]\frac{1}{2}}$ $A_k f_j(t)$ is equidistributed as $f_k(t)$, which follows again by the 2-stability, cf. Lemma 26. Now the proof of this assertion follows directly from (56) and repeating the same arguments as in the proof of Theorem 11 with $k_1 + k_2$ and $j_1 + j_2$ replacing k and j there.

For the second assertion of the theorem we use

$$(\widetilde{A}_k f_j)(t) = \sum_{l_1=0}^{2^{k_1}-1} \sum_{l_2=0}^{2^{k_2}-1} \nu_{k,l} \chi_{k,l}(t),$$

where $\nu_k = (\nu_{k,l} : l_i \in \{0, \ldots, 2^{k_i} - 1\})$ is a vector of independent variables from $\mathcal{G}_{2^{j_1+j_2-(k_1+k_2)}}$, see Definition 28.

Similar to the proof of Theorem 13, using (56) we estimate with the rearrangement

$$\sup_{0<t<1} \frac{(\widetilde{A}_k f_j)^*(t)}{\sqrt{2^{k_1+k_2-(j_1+j_2)} \log(1/t) + 1}} \le c \sup_{z=0,\ldots,k_1+k_2-1} \frac{(\nu_k^*)_{2^z}}{\sqrt{2^{k_1+k_2-(j_1+j_2)}(k_1 + k_2 - z - 1) + 1}},$$

where $\left((\nu_k^*)_m\right)_{m=1}^{2^{k_1+k_2}}$ is the rearrangement of ν_k above. Now, we have reformulated again the two dimensional problem into a one dimensional one with the help of the non-increasing rearrangement and we can use the arguments of the proof of Theorem 13. □

Remark 9 The proposed analogy with the one-dimensional setting could be investigated even further by defining the analogues of the function spaces D and $\mathfrak{D}$ from Definition 14. Although the general direction seems to be quite obvious, we do not pursue it in this paper in order to avoid further technicalities.

4 A Few Facts About Random Variables and Function Spaces

4.1 Gaussian Variables

For the sake of completeness, we present the definition of Gaussian random variables and recall some of their basic properties.

Definition 25 We say that the random variable ξ has standard normal distribution (or standard Gaussian distribution) and write $\xi \sim \mathcal{N}(0, 1)$, if its density function is given

by

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}.$$

We will need the following two properties of Gaussian variables (and refer to [56, Section 3.2] for details).

Lemma 26 *Let $k \in \mathbb{N}$ and let $\xi = (\xi_1, \ldots, \xi_k)$ be a vector of independent standard normal random variables.*

(i) **(2-stability of normal distribution)** *Let $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$. Then the random variable $\langle \lambda, \xi \rangle = \lambda_1 \xi_1 + \cdots + \lambda_k \xi_k$ is a normal variable with mean zero and variance $\sum_{i=1}^{k} \lambda_i^2$.*
(ii) *Let $1 \leq j \leq k$ be positive integers and let $u^1, \ldots, u^j \in \mathbb{R}^k$ be orthogonal. Then the random variables $\langle u^1, \xi \rangle, \ldots, \langle u^j, \xi \rangle$ are independent.*

We shall also make use of the following tail bounds.

Lemma 27 (**Concentration inequalities for standard Gaussian variables**)

(i) *Let ω be a standard Gaussian variable. Then*

$$\mathbb{P}(|\omega| \geq x) \leq \frac{2 \exp(-x^2/2)}{\sqrt{2\pi} x}, \quad x > 0. \tag{57}$$

(ii) *Let $0 \leq k < j$ be integers and let $\xi = (\xi_0, \ldots, \xi_{2^j - 1})$ be a vector of independent standard normal random variables. Then, for every $K \geq 1$,*

$$\mathbb{P}\left(\xi_{2^k}^* \geq K\sqrt{j-k}\right) \leq \left(2e^{-K^2/2}\right)^{(j-k)2^k} \cdot e^{2^k}.$$

Here, $\xi^* = (\xi_1^*, \ldots, \xi_{2^j}^*)$ is the non-increasing rearrangement of ξ.

Proof (i) The proof follows from the elementary bound

$$\int_x^\infty e^{-u^2/2} du \leq \frac{1}{x} \int_x^\infty u e^{-u^2/2} du = \frac{e^{-x^2/2}}{x}.$$

(ii) Since $\xi_m \sim \mathcal{N}(0,1)$ are independent, (57) together with the estimate $\binom{n}{k} < \left(\frac{en}{k}\right)^k$ for the binomial coefficients yields

$$\mathbb{P}\left(\xi_{2^k}^* \geq K\sqrt{j-k}\right) \leq \binom{2^j}{2^k} \mathbb{P}\left(|\omega| \geq K\sqrt{j-k}\right)^{2^k} \leq \left(\frac{e2^j}{2^k}\right)^{2^k} \left(\frac{2e^{-K^2(j-k)/2}}{\sqrt{2\pi K^2(j-k)}}\right)^{2^k}$$

$$\leq e^{2^k} \left(2e^{-K^2/2}\right)^{(j-k)2^k}.$$

$\square$

4.2 Absolute Values of Gaussian Variables

In this section we consider averages of absolute values of Gaussian variables.

Definition 28 Let $N \geq 1$ be a positive integer and let $\xi_1, \ldots, \xi_N$ be independent standard normal variables. Let ν be a random variable. We write $\nu \sim \mathcal{G}_N$ if ν has the same distribution as

$$\frac{1}{N} \sum_{m=1}^{N} |\xi_m|.$$

In terms of concentration inequalities we have the following result.

Lemma 29 (Concentration inequalities for ν)

(i) Let $N \geq 1$, $\nu \sim \mathcal{G}_N$, and $\omega \sim \mathcal{N}(0, 1)$. Then

$$\mathbb{P}(\nu \geq t) \leq 2^N \mathbb{P}(\omega \geq \sqrt{N} t), \quad t > 0.$$

Moreover, if $t \geq 2\sqrt{\ln 2}$, then

$$\mathbb{P}(\nu \geq t) \leq \exp(-N t^2 / 4).$$

(ii) Let $0 \leq k < j$ be integers and let $v = (v_0, \ldots, v_{2^j - 1})$ be a vector of independent $\mathcal{G}_N$ variables. Then

$$\mathbb{P}(v_{2^k}^* \geq t) \leq \left[e \cdot 2^{j-k} \exp(-N t^2 / 4) \right]^{2^k}, \quad t \geq 2\sqrt{\ln 2}. \tag{58}$$

Proof (i) We use the 2-stability of normal variables from Lemma 26 and estimate

$$\mathbb{P}(\nu \geq t) = \mathbb{P}\Big(\sum_{m=1}^{N} |\xi_m| \geq N t \Big) = \mathbb{P}\Big(\exists \varepsilon \in \{-1, +1\}^N : \sum_{m=1}^{N} \varepsilon_m \xi_m \geq N t \Big)$$

$$\leq 2^N \mathbb{P}\Big(\sum_{m=1}^{N} \xi_m \geq N t \Big) = 2^N \mathbb{P}(\omega \geq \sqrt{N} t).$$

The second statement follows by (57).

(ii) The proof resembles very much the proof of Lemma 27 (ii). We estimate

$$\mathbb{P}(v_{2^k}^* \geq t) \leq \binom{2^j}{2^k} \mathbb{P}(v_1 \geq t)^{2^k} \leq \left(\frac{e\, 2^j}{2^k} \right)^{2^k} \exp(-N t^2 / 4 \cdot 2^k) = \left[e \cdot 2^{j-k} \exp(-N t^2 / 4) \right]^{2^k}.$$

$\square$

4.3 Integrated Absolute Wiener Process

Let

$$\mathcal{W} = \int_0^1 |W_s| \, ds \tag{59}$$

be the integral of the absolute value of the Wiener process. The distribution of the random variable $\mathcal{W}$ is rather complicated, cf. [24, 53]. For our purposes, it will be sufficient to obtain concentration inequalities for $\mathcal{W}$ similar to those for Gaussian and absolute values of Gaussian variables as given in Lemmas 27 and 29. We use as a tool the integral of the square of the Wiener process

$$\mathcal{S} = \left(\int_0^1 |W_s|^2 \, ds \right)^{1/2} \tag{60}$$

in order to get the tail bounds for (59).

Lemma 30 (Concentration inequalities for $\mathcal{W}$)

(i) Let $t \geq 1$. Then

$$\mathbb{P}\left(\mathcal{W} \geq \frac{1}{2} + t \right) \leq \exp(-t^2). \tag{61}$$

(ii) There is a constant $c > 0$ such that for every positive integer $N \geq 1$ and $(\mathcal{W}_j)_{j=1}^N$ i.i.d. as in (59) it holds

$$\mathbb{P}\left(\frac{1}{N} \sum_{j=1}^N \mathcal{W}_j > t \right) \leq \exp(1 - c \, N t^2), \quad t > 1. \tag{62}$$

(iii) Let $0 \leq k < j$ be two integers and let $v = (v_0, \ldots, v_{2^j-1})$ be a vector of independent random variables, each equidistributed with $\frac{1}{N} \sum_{j=1}^N \mathcal{W}_j$. Then

$$\mathbb{P}(v_{2^k}^* \geq t) \leq \left[e^2 \cdot 2^{j-k} \exp(-cNt^2) \right]^{2^k}, \quad t > 1. \tag{63}$$

Proof *Step 1.* We use the Karhunen–Loève expansion of the Wiener process (see [29] or [39, Chapter XI]), i.e.,

$$W_t = \sqrt{2} \sum_{k=1}^\infty Z_k \frac{\sin((k - 1/2)\pi t)}{(k - 1/2)\pi}, \quad t \in [0, 1], \tag{64}$$

where $(Z_k)_{k=1}^\infty$ is a sequence of independent standard Gaussian variables and the series converges in L_2 uniformly over $t \in [0, 1]$. We insert (64) into (60) and obtain

$$\mathcal{W}^2 \leq \mathcal{S}^2 = \int_0^1 |W_s|^2 ds = \sum_{k=1}^\infty \frac{Z_k^2}{(k - 1/2)^2 \pi^2} = \sum_{k=1}^\infty \alpha_k Z_k^2,$$

where $\alpha_k = \dfrac{1}{(k-1/2)^2\pi^2}$. By [35, Lemma 1],

$$\mathbb{P}\left(\sum_{k=1}^{\infty}\alpha_k(Z_k^2-1)\geq 2\|\alpha\|_2\sqrt{x}+2\|\alpha\|_\infty x\right)\leq \exp(-x)$$

for every $x > 0$. Using that

$$\|\alpha\|_1=\frac{1}{2},\quad \|\alpha\|_2=\frac{1}{\sqrt{6}},\quad \text{and}\quad \|\alpha\|_\infty=\frac{4}{\pi^2},$$

we obtain

$$\mathbb{P}\left(\mathcal{W}\geq\frac{1}{2}+t\right)\leq\mathbb{P}\left(\mathcal{S}\geq\frac{1}{2}+t\right)\leq\exp(-t^2),\quad t\geq 1,$$

which gives (61).

Step 2. Using the properties of the Brownian motion, cf. Definition 1, we obtain

$$\mathbb{E}\mathcal{W}=\int_0^1\mathbb{E}|W_s|ds=\int_0^1\mathbb{E}|W_s-W_0|ds=\sqrt{\frac{2}{\pi}}\int_0^1\sqrt{s}ds=\sqrt{\frac{2}{\pi}}\cdot\frac{2}{3}\in\left(\frac{1}{2},1\right).$$

Put $X=\mathcal{W}-\mathbb{E}\mathcal{W}$. Then $\mathbb{P}(|X|>t)\leq 1$ for $t\leq\sqrt{2}$ and

$$\mathbb{P}(|X|>t)=\mathbb{P}(\mathcal{W}>t+\mathbb{E}\mathcal{W})\leq\mathbb{P}\left(\mathcal{W}>t+\frac{1}{2}\right)\leq\exp(-t^2)$$

for $t>\sqrt{2}$. We conclude that $\mathbb{P}(|X|>t)\leq\exp(1-t^2/2)$ for all $t>0$, i.e., X is a centered subgaussian variable, cf. [63, Definition 5.7]. Therefore, by the Hoeffding-type inequality, cf. [63, Proposition 5.10], there is a constant $c_1>0$, such that

$$\mathbb{P}\left(\frac{1}{N}\sum_{j=1}^{N}\mathcal{W}_j>t+\mathbb{E}\mathcal{W}\right)\leq\exp(1-c_1Nt^2),\quad t>0,$$

which, in turn, implies (62) for $t>1$ and $c>0$ small enough.

Step 3. Finally, we conclude that if $v=(v_0,\ldots,v_{2^j-1})$ is a vector of independent random variables, each equidistributed with $\frac{1}{N}\sum_{j=1}^{N}\mathcal{W}_j$, then

$$\mathbb{P}(v_{2^k}^*\geq t)\leq\left[e^2\cdot 2^{j-k}\exp(-c_2Nt^2)\right]^{2^k}$$

by an argument quite similar to the proof of Lemma 29 (ii).   □

4.4 Orlicz Spaces

Let us recall, that the Orlicz function Φ_2 was defined in (16) and the corresponding Orlicz space L_{Φ_2} was introduced in (17). The following characterization is a special case of [2, Theorem 10.3] adapted to the domain $[0,1]^d$. We include its short proof to make our presentation self-contained.

Theorem 31 *A measurable function f on $[0,1]^d$ belongs to $L_{\Phi_2}([0,1]^d)$ if, and only if, there exists $c > 0$ such that*

$$f^*(t) \le c\sqrt{\log(1/t) + 1}, \quad 0 < t < 1. \tag{65}$$

Moreover, the expression

$$\|f\|_{\Phi_2^*} := \sup_{0 < t < 1} \frac{f^*(t)}{\sqrt{\log(1/t) + 1}}$$

is equivalent to $\|f\|_{\Phi_2}$.

Proof If f satisfies (65) for some c, then we obtain for $\lambda = 2c$

$$
\begin{aligned}
\int_{[0,1]^d}\Big[\exp(f(x)^2/\lambda^2) - 1\Big]dx &= \int_0^1\Big[\exp(f^*(t)^2/\lambda^2) - 1\Big]dt \\
&\le \int_0^1\Big[\exp\Big(\frac{c^2(\log(1/t) + 1)}{\lambda^2}\Big) - 1\Big]dt \\
&= \int_0^1\Big[\exp\Big(\frac{\log(1/t) + 1}{4}\Big) - 1\Big]dt < 1.
\end{aligned}
$$

Hence, $f \in L_{\Phi_2}([0,1]^d)$ and $\|f\|_{\Phi_2} \le 2\|f\|_{\Phi_2^*}$.

If, on the other hand, $f \in L_{\Phi_2}([0,1]^d)$ with $\|f\|_{\Phi_2} \le 1$, then

$$
\begin{aligned}
1 &\ge \int_{[0,1]^d}\Phi_2(|f(x)|)dx = \int_0^1\Phi_2(f^*(s))ds \ge \int_0^t\Big[\exp(f^*(s)^2) - 1\Big]ds \\
&\ge t\Big[\exp(f^*(t)^2) - 1\Big],
\end{aligned}
$$

i.e., $f^*(t)^2 \le \log(1 + 1/t) \le \log(e/t) = 1 + \log(1/t)$ for every $0 < t < 1$ and (65) follows. $\square$

We need also another characterization of the norm of $L_{\Phi_2}([0,1]^d)$. Its proof can be found in [9, Theorem 3.4] but again we include it for the reader's convenience.

Theorem 32 *A measurable function f on $[0,1]^d$ belongs to $L_{\Phi_2}([0,1]^d)$ if, and only if, there exists $c > 0$ such that*

$$\|f\|_p \le c\sqrt{p} \ \ holds\,for\,all \ \ p \ge 1.$$

Moreover,

$$\|f\|_{(\Phi_2)} := \sup_{p \geq 1} \frac{\|f\|_p}{\sqrt{p}} \tag{66}$$

is an equivalent norm on $L_{\Phi_2}([0,1]^d)$.

Proof First of all we show $\|f\|_{(\Phi_2)} \leq \|f\|_{\Phi_2}$. To that end, let $f \in L_{\Phi_2}([0,1]^d)$ be given with $\|f\|_{\Phi_2} \leq 1$. Then by the power series of the exponential function we estimate for any $n \in \mathbb{N}$

$$1 \geq \int_{[0,1]^d} \Phi_2(|f(x)|)dx = \int_{[0,1]^d} \sum_{k=1}^{\infty} \frac{|f(x)|^{2k}}{k!} dx \geq \frac{1}{n!} \|f\|_{2n}^{2n}.$$

Now using $n! \leq n^n$ we obtain

$$\frac{\|f\|_{2n}^{2n}}{n^n} \leq 1 \quad \text{which is equivalent to} \quad \frac{\|f\|_{2n}}{\sqrt{2n}} \leq \frac{1}{\sqrt{2}}.$$

If $1 \leq p < 2$, we obtain

$$\frac{\|f\|_p}{\sqrt{p}} \leq \|f\|_2 \leq 1.$$

If $2 < p < \infty$, we choose the unique $n \in \mathbb{N}$ with $n \geq 2$ such that $2(n-1) < p \leq 2n$ and obtain

$$\frac{\|f\|_p}{\sqrt{p}} \leq \frac{\|f\|_{2n}}{\sqrt{2(n-1)}} \leq \frac{\sqrt{n}}{\sqrt{2(n-1)}} \leq 1,$$

which finishes the first step.

In the second step, we are going to show $\|f\|_{\Phi_2} \leq C\|f\|_{(\Phi_2)}$. To that end, we choose f such that (66) is finite. Using Stirling's formula we can fix an $\lambda_0 > 1$ such that $\lambda_0 n! \geq (n/e)^n$ holds for all $n \in \mathbb{N}$. We estimate with the help of the power series of the exponential function

$$\int_{[0,1]^d} \Phi_2\left(\frac{|f(x)|}{\lambda}\right) dx = \int_{[0,1]^d} \sum_{n=1}^{\infty} \frac{|f(x)|^{2n}}{\lambda^{2n} n!} dx \leq \lambda_0 \sum_{n=1}^{\infty} \frac{(2e)^n}{\lambda^{2n}} \frac{\|f\|_{2n}^{2n}}{(2n)^n}$$

and now choosing $\lambda = \sqrt{2e(1+\lambda_0)}\|f\|_{(\Phi_2)}$ gives

$$\int_{[0,1]^d} \Phi_2\left(\frac{|f(x)|}{\lambda}\right) dx \leq \lambda_0 \sum_{n=1}^{\infty} (1+\lambda_0)^{-n} = 1.$$

This shows $\|f\|_{\Phi_2} \leq \lambda = \sqrt{2e(1+\lambda_0)}\|f\|_{(\Phi_2)}$ and finishes the proof. $\qquad\square$

In the refined analysis concerning the regularity of Brownian paths, we need also a generalization of Theorem 31. First, we define a scale of Orlicz functions $\Phi_{2,A}$, where $0 < A \leq 1$ is a real parameter, via

$$
\Phi_{2,A}(u) = \begin{cases} u^2, & 0 < u \leq 1, \\ \exp\left(\dfrac{u^2 - 1}{A}\right), & 1 < u < \infty. \end{cases} \tag{67}
$$

It is easy to see that this scale of Orlicz functions fulfills the following estimates for all $u > 0$ and all $0 < A \leq 1$

$$
\Phi_2\left(\frac{u}{\sqrt{2}}\right) \leq \Phi_{2,1}(u) \leq \Phi_{2,A}(u) \leq \Phi_2\left(\frac{u}{\sqrt{A}}\right).
$$

Therefore the Orlicz space associated to $\Phi_{2,A}$ coincides with L_{Φ_2} for every $0 < A \leq 1$. Nevertheless, the equivalence constants in the respective norms will depend on A. It is quite interesting (and of a crucial importance for us) that the following simple expression

$$
\|f\|_{2,A}^{(1)} := \sup_{0 < t < 1} \frac{f^*(t)}{\sqrt{A \log(1/t) + 1}} \tag{68}
$$

is equivalent to the Orlicz norm associated with $\Phi_{2,A}$ (which we denote by $\|f\|_{2,A}$) and that the equivalence constants are independent on the parameter $A \in (0, 1]$.

Theorem 33 *Let $0 < A \leq 1$ and let f be a measurable function on $[0, 1]^d$. Then*

$$
\|f\|_{2,A}^{(1)} \leq \|f\|_{2,A} \leq 4 \|f\|_{2,A}^{(1)}.
$$

Proof Let $\|f\|_{2,A} \leq 1$ and let $0 < t < 1$. If $f^*(t) \leq 1$, then also $f^*(t) \leq \sqrt{A \log(1/t) + 1}$ and there is nothing to prove. If $f^*(t) > 1$, then we estimate

$$
1 \geq \int_{[0,1]^d} \Phi_{2,A}(|f(s)|)ds \geq \int_0^t \Phi_{2,A}(f^*(s))ds \geq t\, \Phi_{2,A}(f^*(t))
$$
$$
= t \exp\left(\frac{f^*(t)^2 - 1}{A}\right).
$$

By simple algebraic manipulations, it follows that

$$
f^*(t) \leq \sqrt{A \log(1/t) + 1}, \quad 0 < t < 1.
$$

Let, on the other hand, $\|f\|_{2,A}^{(1)} = c$. Then, putting $\lambda = 4c$ we obtain $\dfrac{f^*(t)^2}{\lambda^2} \leq \dfrac{A \log(1/t)+1}{16}$ and estimate

$$\int_{[0,1]^d} \Phi_{2,A}\left(\frac{|f(x)|}{\lambda}\right)dx = \int_0^1 \Phi_{2,A}\left(\frac{f^*(t)}{\lambda}\right)dt \le \int_0^1 \exp\left(\frac{f^*(t)^2/\lambda^2 - 1}{A}\right)dt + \int_0^1 \frac{f^*(t)^2}{\lambda^2}dt$$

$$\le \int_0^1 \exp\left(\frac{\log(1/t)}{16} - \frac{15}{16A}\right)dt + \int_0^1 \frac{A\log(1/t)+1}{16}dt$$

$$\le \exp\left(-\frac{15}{16}\right)\int_0^1 \exp\left(\frac{\log(1/t)}{16}\right)dt + \frac{1}{16}\int_0^1 \left(\log(1/t)+1\right)dt$$

$$= \exp\left(-\frac{15}{16}\right)\cdot\frac{16}{15} + \frac{1}{8} \le 1,$$

which implies that $\|f\|_{2,A} \le \lambda = 4c$. $\qquad\square$

References

1. Babenko, K.I.: On the approximation of a certain class of periodic functions of several variables by trigonometric polynomials. Dokl. Akad. Nauk USSR **132**, 982–985 (1960). (**English transl. in Soviet Math. Dokl., 1 (1960)**))
2. Bennett, C., Rudnick, K.: On Lorentz–Zygmund spaces. Dissertationes Math. (Rozprawy Mat.) **175**, 67 (1980)
3. Bényi, Á., Oh, T.: Modulation spaces, Wiener amalgam spaces, and Brownian motions. Adv. Math. **228**(5), 2943–2981 (2011)
4. Bilyk, D., Lacey, M.T.: On the small ball inequality in three dimensions. Duke Math. J. **143**(1), 81–115 (2008)
5. Bilyk, D., Lacey, M.T., Vagharshakyan, A.: On the small ball inequality in all dimensions. J. Funct. Anal. **254**(9), 2470–2502 (2008)
6. Bungartz, H.-J., Griebel, M.: Sparse grids. Acta Numer. **13**, 147–269 (2004)
7. Byrenheid, G.: Sparse representation of multivariate functions based on discrete point evaluations, PhD Thesis, Bonn (2018)
8. Ciesielski, Z.: Modulus of smoothness of the Brownian paths in the L_p norm. Constructive theory of functions (Varna, Bulgaria, 1991), pp. 71–75 (1992)
9. Ciesielski, Z.: Orlicz spaces, spline systems, and Brownian motion. Constr. Approx. **9**, 191–208 (1993)
10. Ciesielski, Z., Kerkyacharian, G., Roynette, B.: Quelques espaces fonctionnels associés à des processus Gaussiens. Stud. Math. **107**(2), 171–204 (1993)
11. Csörgő, M., Révész, P.: Strong Approximations in Probability and Statistics. Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York (1981)
12. Dalang, R.C.: Level Sets and Excursions of the Brownian Sheet, Topics in Spatial Stochastic Processes. Lecture Notes in Mathematics, vol. 1802. Springer, Berlin (2003)
13. Domínguez, O., Tikhonov, S.: Function spaces of logarithmic smoothness: embeddings and characterizations. Mem. Am. Math. Soc. **282**, 1393 (2023)

14. Dunker, T.: Estimates for the small ball probabilities of the fractional Brownian sheet. J. Theor. Probab. **13**(2), 357–382 (2000)
15. Dung, D., Temlyakov, V., Ullrich, T.: Hyperbolic Cross Approximation. Springer, Berlin (2018)
16. Faber, G.: Über stetige Funktionen. Math. Ann. **66**, 81–94 (1909)
17. Farkas, W., Leopold, H.-G.: Characterisations of function spaces of generalised smoothness. Ann. Mat. Pura Appl. **185**(1), 1–62 (2006)
18. Haar, A.: Zur Theorie der orthogonalen Funktionensysteme. Math. Ann. **69**, 331–371 (1910)
19. Hansen, M.: Nonlinear approximation and function spaces of dominating mixed smoothness, PhD Thesis, Jena (2010)
20. Haroske, D.D.: On more general Lipschitz spaces. Z. Anal. Anwendungen **19**(3), 781–800 (2000)
21. Herren, V.: Lévy-type processes and Besov spaces. Potential Anal. **7**(3), 689–704 (1997)
22. Hummel, F.: Sample paths of white noise in spaces with dominating mixed smoothness. Banach J. Math. Anal. **15**(3), 54, 38 (2021)
23. Hytönen, T.P., Veraar, M.C.: On Besov regularity of Brownian motions in infinite dimensions. Probab. Math. Stat. **28**, 143–162 (2008)
24. Kac, M.: On the average of a certain Wiener functional and a related limit theorem in calculus of probability. Trans. Am. Math. Soc. **59**, 401–414 (1946)
25. Kahane, J.P.: Some Random Series of Functions, 2nd edn. Cambridge University Press, London (1985)
26. Kalyabin, G.A., Lizorkin, P.I.: Spaces of functions of generalized smoothness. Math. Nachr. **133**, 7–32 (1987)
27. Kamont, A.: Isomorphism of some anisotropic Besov and sequence spaces. Stud. Math. **110**(2), 169–189 (1994)
28. Kamont, A.: On the fractional anisotropic Wiener field. Probab. Math. Stat. **16**(1), 85–98 (1996)
29. Karhunen, K.: Über lineare Methoden in der Wahrscheinlichkeitsrechnung. Ann. Acad. Sci. Fennicae Ser. A. I. Math.-Phys. **37**, 79 (1947)
30. Khoshnevisan, D.: Multiparameter Processes. An Introduction to Random Fields, Springer Monographs in Mathematics. Springer, New York (2002)
31. Khoshnevisan, D., Xiao, Y.: Lévy processes: capacity and Hausdorff dimension. Ann. Probab. **33**(3), 841–878 (2005)
32. Kühn, T., Linde, W.: Optimal series representation of fractional Brownian sheets. Bernoulli **8**, 669–696 (2002)
33. Kuelbs, J., Li, W.V.: Metric entropy and the small ball problem for Gaussian measures. J. Funct. Anal. **116**, 133–157 (1993)
34. Kuelbs, J., Li, W.: Small ball estimates for Brownian motion and the Brownian sheet. J. Theor. Probab. **6**(3), 547–577 (1993)
35. Laurent, B., Massart, P.: Adaptive estimation of a quadratic functional by model selection. Ann. Stat. **28**(5), 1302–1338 (2000)
36. Lévy, P.: Théorie de l'addition des variables aléatoires. Monographies des Probabilités; calcul des probabilités et ses applications **1**, Paris (1937)
37. Lévy, P.: Processus Stochastiques et Mouvement Brownien. Gauthier-Villars & Cie, Paris (1965)
38. Li, W., Linde, W.: Approximation, metric entropy and small ball estimates for Gaussian measures. Ann. Probab. **27**(3), 1556–1578 (1999)
39. Loève, M.: Probability Theory II. Graduate Texts in Mathematics, vol. 46. Springer, New York (1978)
40. Mörters, P., Peres, Y.: Brownian Motion, Cambridge Series in Statistical and Probabilistic Mathematics 30. Cambridge University Press, Cambridge (2010)
41. Moura, S.D.: Function spaces of generalized smoothness. Dissertationes Math. **398**, 88 (2001)
42. Neves, J.S.: Extrapolation results on general Besov–Hölder–Lipschitz spaces. Math. Nachr. **230**, 117–141 (2001)
43. Ondreját, M., Šimon, P., Kupsa, M.: Support of solutions of stochastic differential equations in exponential Besov-Orlicz spaces. Stoch. Anal. Appl. **36**(6), 1037–1052 (2018)
44. Ondreját, M., Veraar, M.: On temporal regularity of stochastic convolutions in 2-smooth Banach spaces. Ann. Inst. Henri Poincaré Probab. Stat. **56**(3), 1792–1808 (2020)
45. Peetre, J.: New Thoughts on Besov Spaces, Duke University Mathematics Series 1. Duke University, Durham (1976)
46. Pick, L., Sickel, W.: Several types of intermediate Besov-Orlicz spaces. Math. Nachr. **164**, 141–165 (1993)
47. Roynette, B.: Mouvement Brownien et espaces de Besov. Stoch. Stoch. Rep. **43**, 221–260 (1993)

48. Runst, T., Sickel, W.: Sobolev Spaces of Fractional Order, Nemytskij Operators, and Nonlinear Partial Differential Equations, De Gruyter Series in Nonlinear Analysis and Applications 3. Walter de Gruyter & Co., Berlin (1996)
49. Schmeisser, H.-J., Triebel, H.: Topics in Fourier Analysis and Function Spaces. Wiley, Chichester (1987)
50. Schilling, R.L.: On Feller processes with sample paths in Besov spaces. Math. Ann. **309**(4), 663–675 (1997)
51. Schilling, R.L.: Function spaces as path spaces of Feller processes. Math. Nachr. **217**, 147–174 (2000)
52. Sickel, W., Ullrich, T.: Tensor products of Sobolev–Besov spaces and applications to approximation from the hyperbolic cross. J. Approx. Theor. **161**, 748–786 (2009)
53. Takács, L.: On the distribution of the integral of the absolute value of the Brownian motion. Ann. Appl. Probab. **3**(1), 186–197 (1993)
54. Talagrand, M.: The small ball problem for the Brownian sheet. Ann. Probab. **22**, 1331–1354 (1994)
55. Temlyakov, V.N.: Approximation of Periodic Functions. Nova Science Publishes Inc, New York (1993)
56. Tong, Y.L.: The Multivariate Normal Distribution, Springer Series in Statistics. Springer, New York (1990)
57. Triebel, H.: Theory of Function Spaces, Monographs in Mathematics 78. Birkhäuser Verlag, Basel (1983)
58. Triebel, H.: Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration, EMS Tracts in Mathematics, 11. European Mathematical Society, Zürich (2010)
59. Triebel, H.: Faber Systems and Their use in Sampling, Discrepancy, Numerical Integration. European Mathematical Society, Zürich (2012)
60. Triebel, H.: Function Spaces with Dominating Mixed Smoothness. European Mathematical Society, Zürich (2019)
61. Tudor, C.A., Xiao, Y.: Sample path properties of bifractional Brownian motion. Bernoulli **13**(4), 1023–1052 (2007)
62. Veraar, M.C.: Regularity of Gaussian white noise on the d-dimensional torus. In: Nawrocki, M., Wnuk, W. (eds.) Marcinkiewicz Centenary Volume 95, Banach Center Publication, pp. 385–398. Institute of Mathematics of the Polish Academy of Sciences, Warsaw (2011)
63. Vershynin, R.: Introduction to the non-asymptotic analysis of random matrices. In: Compressed sensing, Cambridge Univerity Press, Cambridge, pp. 210–268 (2012)
64. Vybíral, J.: Function spaces with dominating mixed smoothness. Dissertationes Math. **436**, 73 (2006)
65. Walsh, J.B.: An Introduction to Stochastic Partial Differential Equations, École d'été de Probabilités de Saint-Flour XIV-1984. Lecture Notes in Mathematics, vol. 1180, pp. 265–439. Springer, New York (1986)
66. Wang, W.: Almost-sure path properties of fractional brownian sheet. Ann. de l'I.H.P. Probabilités et Stat. Tome **43**(5), 619–631 (2007)
67. Xiao, Y.: Sample Path Properties of Anisotropic Gaussian Random Fields, A Minicourse on Stochastic Partial Differential Equations, 145–212. Lecture Notes in Math, vol. 1962. Springer, Berlin (2009)