The Annals of Probability
1994, Vol. 22, No. 3, 1331–1354

THE SMALL BALL PROBLEM FOR THE BROWNIAN SHEET[1]

BY MICHEL TALAGRAND

University of Paris VI and Ohio State University

We show that the logarithm of the probability that the Brownian sheet has a supremum at most ε over $[0,1]^2$ is of order $\varepsilon^{-2}(\log(1/\varepsilon))^3$.

1. Introduction. The Brownian sheet is the centered Gaussian process $B_{s,t}$ $(s,t \in \mathbb{R}^+)$ such that

$$EB_{s,t}B_{s',t'} = \min(s,s')\min(t,t').$$

Our main result is as follows.

THEOREM 1.1. *For some universal constant C and all $\varepsilon < 1/2$, we have*

$$\exp\left(-\frac{C}{\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^3\right) \le P\left(\sup_{0\le s,t\le 1}|B_{s,t}| \le \varepsilon\right) \le \exp\left(-\frac{1}{C\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^3\right). \tag{1.1}$$

The lower bound in (1.1) was obtained by Bass [2] and was the starting point of this work.

As the third power of the logarithmic term in (1.1) might indicate, there are some unexpectedly subtle phenomena that occur in the present setting. Some of these are reflected in the contrast between Theorem 1.1 and the next result. There, we denote by λ Lebesgue measure on $[0,1]^2$; by $\|\cdot\|_2$ the norm in $L_2(\lambda)$; and, for $\alpha \ge 2$, by $\|\cdot\|_{\psi_\alpha}$ the Orlicz norm

$$\|f\|_{\psi_\alpha} = \inf\left\{c > 0;\ \int \exp\left(\frac{|f|^\alpha}{c^\alpha}\right) d\lambda \le 2\right\}.$$

THEOREM 1.2. (a) *For some universal constant K and all $0 < \varepsilon \le 1/2$, we have*

$$P(\|B_{s,t}\|_2 \le \varepsilon) \le \exp\left(-\frac{K}{\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^2\right). \tag{1.2}$$

Received January 1993.

[1]Partially supported by an NSF grant.

AMS 1991 *subject classifications.* Primary 60E15, 60G17; secondary 60B11, 28C20, 46A35, 46B20.

Key words and phrases. Probabilities for suprema, probabilities for norms.

(b) *Given $2 \le \alpha < \infty$, there exists a constant $K(\alpha)$ depending on α only such that, for $0 < \varepsilon \le 1/2$, we have*

$$\exp\left(-\frac{K(\alpha)}{\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^{3-2/\alpha}\right) \le P(\|B_{s,t}\|_{\psi_\alpha} \le \varepsilon) \le \exp\left(-\frac{1}{K(\alpha)\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^{3-2/\alpha}\right). \tag{1.3}$$

The fact that the behavior of the probability of balls of small radius is different whether the ball is considered for the supremum norm or (say) for the $\|\cdot\|_{\psi_2}$-norm can heuristically be interpreted by saying that the set of points (s,t) for which $|B_{s,t}|$ is of order $\sup_{s,t}|B_{s,t}|$ is very small. (A consequence of this is that it is apparently impossible to exhibit these points by a purely probabilistic argument.)

A version of Chung's law of the iterated logarithm follows easily from Theorem 1.1.

THEOREM 1.3. *Almost surely, we have*

$$0 < \liminf_{u\to\infty} \frac{(\log\log u)^{1/2}}{u(\log\log\log u)^{3/2}} \sup_{0\le s,t\le u} |B_{s,t}| < \infty.$$

We conclude this section with a few remarks. It is also of interest to prove estimates of the type given by Theorem 1.1 when the Brownian sheet is replaced by the "tied-down" Brownian sheet. It is, however, a matter of routine to deduce these from Theorem 1.1. [Roughly speaking, the difference between two suitable versions of these processes consists of a process that creates lower-order terms in the bounds of (1.1).] Also of interest is the question of getting upper and lower bounds for $P(\sup_{s,t}|B_{s,t} - f(s,t)| \le \varepsilon)$ for a function f on $[0,1]^2$. Bounds for these quantities are deduced from (1.1) using the Cameron–Martin formula, in a manner that is better described in the general setting of a centered Gaussian measure μ on a separable Banach space X. If $\|\cdot\|_\mu$ denotes the natural norm on the reproducing kernel H of μ (which, in the case of the Brownian sheet, will be described in Section 2), let us consider the quantity

$$I(f,\varepsilon) = \inf\{\|g\|_\mu;\ \|f-g\| \le \varepsilon\}.$$

Then it is easy to prove the inequalities

$$\mu\left(\{\|x-f\| \le \varepsilon\}\right) \ge \mu\left(\left\{\|x\| \le \frac{\varepsilon}{2}\right\}\right) \exp\left(-\frac{I(f,\varepsilon/2)^2}{2}\right), \tag{1.4}$$

$$\mu\left(\{\|x-f\| \le \varepsilon\}\right) \le \mu\left(\{\|x\| \le 2\varepsilon\}\right) \exp\left(-\frac{I(f,\varepsilon)^2}{2}\right). \tag{1.5}$$

Also, we should mention that we do *not* know how to generalize the upper bound of (1.1) to the case of the Brownian sheet on $[0,1]^d$, $d \geq 3$. This appears to be considerably more difficult.

The paper is organized as follows. In Section 2 we give the probabilistic arguments of the proofs of Theorems 1.1 and 1.2. In Section 3 we prove the combinatorial statement on which these proofs rely. In Section 4 we discuss the situation of the Brownian sheet constructed on more general measures and the problem of the measure of the small balls of the Brownian motion built on a general Gaussian measure.

2. Probability. Probably the easiest way to understand the Brownian sheet (or, more accurately, its restriction to $[0,1]^2$) is as follows. Recalling that λ denotes Lebesgue measure on $[0,1]^2$, consider the operator T from $L^2(\lambda)$ to $C([0,1]^2)$ given by

$$T(f)(s,t) = \int\int_{[0,s]\times[0,t]} f\, d\lambda.$$

Then the law of the Brownian sheet is the Gaussian measure μ on $C([0,1]^2)$ which is the image by T of the canonical cylindrical Gaussian measure on $L^2(\lambda)$. Consequently, well-understood general principles show that, given any complete orthonormal system (C.O.S) $(f_k)_{k\geq 1}$ of $L^2(\lambda)$, then, if $(g_k)_{k\geq 1}$ denotes an independent sequence of standard Gaussian random variables (r.v.'s), the series $\sum_{k\geq 1} g_k T(f_k)$ converges a.s. in $C([0,1]^2)$ and has law μ. Proper choice of the C.O.S. (f_k) will be essential. Let us mention that the space H_μ (the reproducing kernel of μ) is the image of $L^2(\lambda)$ by T, and that its unit ball K [for the norm $\|\cdot\|_\mu$ transported from $L^2(\lambda)$] is (of course) the image under T of the unit ball of $L^2(\lambda)$. Before we start any computation, we must mention that Kuelbs and Li [3] have recently discovered the striking fact that (in the general setting of a Gaussian measure μ on a Banach space X) there is a tight relationship between the measure of the small balls and the covering numbers $N(K,\varepsilon)$ (i.e., the smallest number of balls of X of radius ε centered on K needed to cover K). It certainly would be immoral to deprive the reader of a discussion of this beautifully simple fact (that once again demonstrates the power of abstract methods). Let us set

$$\varphi(\varepsilon) = -\log \mu\Big(\big\{\|x\| \leq \varepsilon\big\}\Big)$$

and let us denote by B the unit ball of X. Consider first a number $a > 0$ and points $(t_i)_{i\leq N}$ of aK such that the balls $t_i + \varepsilon B$ are disjoint. The Cameron–Martin formula and Jensen's inequality easily imply that

$$\mu(t_i + \varepsilon B) \geq \exp\left(-\frac{a^2}{2}\right)\mu(\varepsilon B) = \exp\left(-\varphi(\varepsilon) - \frac{a^2}{2}\right). \tag{2.1}$$

Thus, since μ has mass 1, we have $N \leq \exp(\varphi(\varepsilon) + a^2/2)$.

If we take N maximal, we then get

$$N\left(K, \frac{2\varepsilon}{a}\right) = N(aK, 2\varepsilon) \le \exp\left(\varphi(\varepsilon) + \frac{a^2}{2}\right), \tag{2.2}$$

so that

$$N\left(K, \frac{\sqrt{2}\varepsilon}{\sqrt{\varphi(\varepsilon)}}\right) \le \exp\big(2\varphi(\varepsilon)\big). \tag{2.3}$$

Conversely, we recall that, by a result of Anderson [1] for any centrally symmetric convex body D, we have

$$\mu(x + D) \le \mu(D). \tag{2.4}$$

Consider again $a > 0$ and, for $N = N(aK, \varepsilon)$, consider points $(t_i)_{i \le N}$ such that $aK \subset \cup_{i \le N}(t_i + \varepsilon B)$. Then

$$aK + \varepsilon B \subset \bigcup_{i \le N} (t_i + 2\varepsilon B)$$

and using (2.4) for $D = 2\varepsilon B$ yield

$$\mu(aK + \varepsilon B) \le N\left(K, \frac{\varepsilon}{a}\right)\mu(2\varepsilon B),$$

so that

$$N\left(K, \frac{\varepsilon}{a}\right) \ge \mu(aK + \varepsilon B)\exp\big(\varphi(2\varepsilon)\big).$$

By the Gaussian isoperimetric inequality, for $a = \sqrt{\varphi(\varepsilon)/2}$, we have $\mu(aK + \varepsilon B) \ge 1/2$, so that

$$N\left(K, \frac{\sqrt{2}\varepsilon}{\sqrt{\varphi(\varepsilon)}}\right) \ge \frac{1}{2}\exp\big(\varphi(2\varepsilon)\big). \tag{2.5}$$

In the usual cases, for ε small, we have $\varphi(\varepsilon) \le C\varphi(2\varepsilon)$ for a constant C independent of ε, and thus (2.3) and (2.5) show that the order of φ is determined by the function $N(K, \cdot)$. (See [3] for the details.)

The drawback of the result of Kuelbs and Li [3] is that the numbers $N(K, \eta)$ are very hard to estimate. In the case of the Brownian sheet, to obtain the correct lower bound for $\varphi(\varepsilon)$ (which is the hard part), it would suffice to obtain the correct lower bound for $\log N(K, \eta)$. This, in principle, can be done by exhibiting a large 2η-separated subset of K. While this can be done using Theorem 3.1, it is apparently, at least in the case of the Brownian sheet, of equivalent difficulty to proceed directly (which is what we will do).

We now go back to the discussion of the Brownian sheet. An important observation is that if $(f_k)_{k \ge 0}$ denotes a C.O.S. of $L^2([0, 1])$, then the family $(f_{k,l})_{k,l \ge 0}$ is a C.O.S. of $L_2(\lambda)$, where $f_{k,l} = f_k \otimes f_l$ is given by

$$f_{k,l}(s, t) = f_k(s)f_l(t).$$

Such C.O.S. systems of $L_2(\lambda)$ demonstrate explicitly the structure of T. Indeed, if U is the operator from $L^2([0,1])$ to $C([0,1])$ given by

$$U(f)(t) = \int_0^t f(x)\,dx,$$

then

$$T(f_k \otimes f_l) = U(f_k) \otimes U(f_l).$$

We now start the proof of (1.2) (let us mention that this inequality was obtained independently by Kuelbs and Li). For this we use the trigonometric system as a basis for $L^2([0,1])$; that is, $f_0(x) = 1$ and, for $k \geq 1$,

$$f_{2k}(x) = \cos 2k\pi x; \qquad f_{2k+1}(x) = \sin 2k\pi x.$$

Thus, by the previous considerations, if $(g_{k,l})_{k,l>1}$ denotes an independent family of standard normal r.v.'s, the sum of the series $\sum_{k,l\geq 0} g_{k,l} U(f_k) \otimes U(f_l)$ has the same distribution as $B_{s,t}$ in $C([0,1]^2)$. The following simple observation comes in handy in the proof of upper bounds.

LEMMA 2.1. *If Z_1, Z_2 are two independent Gaussian random vectors valued in a Banach space X, then for a centrally symmetric convex body D, we have*

$$P(Z_1 + Z_2 \in D) \leq P(Z_1 \in D).$$

PROOF. Use (2.4) conditionally on Z_2. □

Consider an integer n and set

$$A_n = \left\{(2k, 2l);\ 2^n \leq kl \leq 2^{n+1}\right\}$$

Thus, it follows from Lemma 2.1 that

$$P\left(\|B_{s,t}\|_2 \leq \varepsilon\right) \leq P\left(\|Z\|_2 \leq \varepsilon\right),$$

where

$$Z = \sum_{(i,j)\in A_n} g_{i,j} U(f_i) \otimes U(f_j).$$

Now, since

$$U(f_{2k})(x) = \frac{1}{2k\pi}\sin(2k\pi x),$$

the functions $(U(f_i) \otimes U(f_j))_{(i,j)\in A_n}$ are orthogonal in $L^2(\lambda)$. Moreover, their L^2-norm is greater than or equal to $2^{-n}/C$, where C is universal. Thus

$$\|Z\|_2^2 \geq \frac{2^{-2n}}{C^2} \sum_{(i,j)\in A_n} g_{i,j}^2$$

and hence

$$P(\|B_{s,t}\|_2 \le \varepsilon) \le P\left(2^{-2n} \sum_{(i,j)\in A_n} g_{i,j}^2 \le C^2\varepsilon^2\right).$$

Now, it is well known and elementary (see, e.g., [4], proof of (15.1), page 428) that if $m_n = \operatorname{card} A_n$, then

$$P\left(\sum_{(i,j)\in A_n} g_{i,j}^2 \le \frac{m_n}{e^2}\right) \le e^{-m_n/2}.$$

Thus, if

$$m_n \ge C^2 2^{2n} e^2 \varepsilon^2, \tag{2.6}$$

we have

$$P(\|B_{s,t}\|_2 \le \varepsilon) \le e^{-m_n/2}.$$

We take for n the largest for which (2.6) holds. Since m_n is of order $n2^n$, $n2^{-n}$ is of order ε^2, so that m_n is of order $n2^n = n^2/(n2^{-n})$, that is of order n^2/ε^2, that is of order $(\log(1/\varepsilon))^2/\varepsilon^2$. This completes the proof of (1.2).

We now turn to the lower bounds in (1.1) and (1.3). In Section 4 we will prove a statement that considerably generalizes the lower bound of (1.1). The following more special argument, however, pinpoints exactly where the basic difficulty lies. It is more convenient now (following [2]) to use the Haar basis of $L^2([0,1])$. Let us recall that this basis $(h_{m,l})$ consists of the function $h_{-1,0} = 1$, and, for $m \ge 0$, of the functions

$$h_{m,l}(x) = 2^{m/2} h(2^m(x - l2^{-m}))$$

for $0 \le l < 2^m$, where

$$h(x) = \begin{cases} 0, & \text{if } x < 0 \text{ or } x \ge 1,\\ 1, & \text{if } 0 \le x < 1/2,\\ -1, & \text{if } 1/2 \le x < 1.\end{cases}$$

For simplicity, let us denote by H the set of indexes $\{(-1,0)\} \cup \{(m,l);\, m \ge 0,\ 0 \le l < 2^m\}$.

Thus, by previous considerations, with obvious notation, the series

$$S = \sum_{H\times H} g_{m,l,m',l'} U(h_{m,l}) \otimes U(h_{m,l'})$$

has the same law in $C([0,1]^2)$ as $B_{s,t}$.

Consider the function $u_{m,l} = U(h_{m,l})$. Let us note the following elementary facts:

(2.7) The support of $u_{m,l}$ is contained in the interval $I_{l,m} = [l2^{-m}, (l+1)2^{-m}]$ (with the convention that $I_{0,-1} = [0,1]$).

$$\sup|u_{m,l}| \le 2^{-m/2}. \tag{2.8}$$

Let us fix an integer n and set $a_k = 2^{(k-n)/4}$ for $k > n$ and $a_k = 2^{-(n-k)}$ for $k \le n$. Consider the event

$$\Omega_n = \{\forall\,(m,l) \in H,\ \forall\,(m',l') \in H,\ |g_{m,l,m',l'}| \le a_{m+m'}\}. \tag{2.9}$$

Thus, if g is a standard normal r.v. and if, for $k \ge -2$, we set

$$n(k) = \operatorname{card}\{(m,l) \in H;\ (m',l') \in H,\ m + m' = k\},$$

we have

$$P(\Omega_n) = \prod_{k \ge -2} P(|g| \le a_k)^{n(k)}.$$

We use standard estimates

$$u \le 1 \Rightarrow P(|g| \le u) \ge \frac{2u}{\sqrt{\pi e}},$$

$$u \ge 1 \Rightarrow P(|g| \le u) \ge 1 - e^{-u^2/2} \ge \exp(-2e^{-u^2/2}).$$

Since $n(k) \le (k+3)2^{k+1}$, we get

$$P(\Omega_n) \ge \exp\Biggl(-\sum_{-2 \le k \le n} (k+3)2^{k+1} \log\left(\frac{\sqrt{\pi e}}{2^{(k-n)+1}}\right) - \sum_{k > n} 2(k+3)2^{k+1} e^{-2^{(k-n)/2-1}}\Biggr).$$

Elementary estimates then show that

$$P(\Omega_n) \ge \exp(-Cn2^n), \tag{2.10}$$

where C is universal.

To obtain the lower bound in (1.1), we bound S on Ω_n. We write $S = \sum_{k \ge -2} S_k$, where $S_k = \sum_{m+m'=k} S_{m,m'}$ and

$$S_{m,m'} = \sum_{l,l'} g_{m,l,m',l'} u_{m,l} \otimes u_{m',l'}. \tag{2.11}$$

We observe that, by (2.7), the functions $u_{m,l} \otimes u_{m',l'}$ involved in the summation of (2.11) have supports with disjoint interiors. Thus (denoting by $\|\cdot\|$ the supremum norm) we get on Ω_n that, by (2.8),

$$\|S_{m,m'}\| \le 2^{-(m+m')/2} a_{m+m'}. \tag{2.12}$$

To bound S_k, we use the triangle inequality

$$\|S_k\| \le \sum_{m+m'=k} \|S_{m,m'}\|. \tag{2.13}$$

Certainly, this is brutal, and, to prove that the left-hand-side inequality in (1.1) can be reversed, we will have at some point to prove a statement to the effect that (2.13) is more or less the best that can be done. (Such is the role of Theorem 3.1.) Thus, combining (2.12) and (2.13), we get, on Ω_n,

$$\|S_k\| \le (k+3)2^{-k/2}a_k. \tag{2.14}$$

Thus, on Ω_n,

$$\|S\| \le \sum_{k\ge -2} \|S_k\| \le \sum_{k\ge -2} (k+3)2^{-k/2}a_k. \tag{2.15}$$

[It should be noted that in (2.15) use of the triangle inequality is rather sharp, since the term for $k = n$ dominates.] Elementary computations from (2.15) show that, on Ω_n,

$$\|S\| \le Cn2^{-n/2}. \tag{2.16}$$

There, as in the sequel, C denotes a universal constant, not necessarily the same at each occurrence. Now, given ε, if one takes the smallest n such that $Cn2^{-n/2} \le \varepsilon$, since $n2^n = n^3/(n^2 2^{-n})$ and since n is of order $\log(1/\varepsilon)$, the left-hand inequality of (1.1) follows from (2.10) and (2.16).

We now start the proof of the left-hand side of (1.3). The important case is the case $\alpha = 2$, since the general case will follow by interpolation with the left-hand side of (1.1).

LEMMA 2.2. *On Ω_n, we have*

$$\forall\, s,t \in [0,1], \qquad \sum_{(m,l)\in H,\,(m',l')\in H} \big(g_{m,l,m',l'}u_{m,l}(s)u_{m'l'}(t)\big)^2 \le Cn2^{-n}. \tag{2.17}$$

PROOF. Arguing as in the proof of (2.12), we see first that, given m, m',

$$\sum_{l,l'} \big(g_{m,l,m',l'}u_{m,l}(s)u_{m',l'}(t)\big)^2 \le a_{m+m'}^2 2^{-(m+m')}. \tag{2.18}$$

Summation over m, m' yields that the left-hand side of (2.17) is bounded by

$$\sum_{k\ge -2} (k+2)a_k^2 2^{-k},$$

from which the result follows by straightforward estimates. □

The second ingredient is the following well-known (and elementary) fact.

LEMMA 2.3. *Consider an independent Bernoulli sequence $(\varepsilon_i)_{i\ge 0}$ [that is, $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = 1/2$]. Then, for numbers $(b_i)_{i\ge 0}$ and $\lambda > 0$, we have*

$$E\exp\lambda\sum_{i\ge 0}\varepsilon_i b_i \le \exp\frac{\lambda^2}{2}\sum_{i\ge 0} b_i^2. \tag{2.19}$$

Consider an independent Bernoulli sequence $(\varepsilon_{m,l,m',l'})$ that is independent of the sequence $g_{m,l,m',l'}$. For clarity, we will assume that these sequences are constructed on a product space $(\Omega \times \Omega', P \otimes P')$ and that, for $(\omega, \omega') \in \Omega \times \Omega'$, $g_{m,l,m',l'}(\omega, \omega')$ depends only on ω and $\varepsilon_{m,l,m',l'}(\omega, \omega')$ depends only on ω'. Thus the set Ω_n of (2.9) depends only on ω.

If we combine (2.17) and (2.19) and integrate over s, t, we have that, for any $\omega \in \Omega_n$,

$$\int\int_{[0,1]^2} \exp\big(\lambda S(\omega, \omega', s, t)\big)\, dP'(\omega')\, ds\, dt \le \exp C\lambda^2 n 2^{-n}, \tag{2.20}$$

where

$$S(\omega, \omega', s, t) = \sum_{m,l,m',l'} \varepsilon_{m,l,m',l'}(\omega') g_{m,l,m',l'}(\omega) u_{m,l}(s) u_{m',l'}(t).$$

Use of Chebyshev's exponential inequality

$$P(Z \ge t) \le (\exp -\lambda t)\, E \exp \lambda Z$$

in the space $[0,1]^2 \times \Omega'$ and a straightforward computation using (2.20) show that

$$\int\int \exp \frac{S^2(\omega, \omega', s, t)}{Cn2^{-n}} dP'(\omega')\, ds\, dt \le \frac{3}{2},$$

so that, by the Fubini theorem for $\omega \in \Omega_n$, we have

$$P'\Big(\big\{\omega';\ \|S(\omega, \omega', \cdot, \cdot)\|_{\psi_2} \le C\sqrt{n} 2^{-n/2}\big\}\Big) \ge \tfrac{1}{4}.$$

In summary, if

$$\Omega'_n = \big\{(\omega, \omega');\ \omega \in \Omega_n;\ \|S(\omega, \omega', \cdot, \cdot)\|_{\psi_2} \le C\sqrt{n} 2^{-n/2}\big\},$$

we have shown that $P \otimes P'(\Omega'_n) \ge \frac{1}{4} \exp(-Cn2^n)$.

We now observe that (since $\varepsilon|g|$ is distributed like g) $S(\omega, \omega', s, t)$ is distributed like $B_{s,t}$. Thus we have shown [using (2.16)] that

$$P\Big(\big\{\|B_{s,t}\|_\infty \le Cn2^{-n/2};\ \|B_{s,t}\|_{\psi_2} \le C\sqrt{n}2^{-n/2}\big\}\Big) \ge \tfrac{1}{4}\exp(-Cn2^n). \tag{2.21}$$

We now observe the elementary interpolation formula, for $\alpha \ge 2$,

$$\|f\|_{\psi_\alpha} \le \|f\|_{\psi_2}^{2/\alpha} \|f\|_\infty^{1-2/\alpha}.$$

Combining this with (2.21) shows that

$$P\Big(\big\{\|B_{s,t}\|_{\psi_\alpha} \le Cn^{1-1/\alpha}2^{-n/2}\big\}\Big) \ge \tfrac{1}{4}\exp(-Cn2^n).$$

Now, if n is the smallest such that $Cn^{1-1/\alpha}2^{-n/2} \le \varepsilon$, since

$$n2^n = n^{3-2/\alpha}\big(n^{1-1/\alpha}2^{-n/2}\big)^{-2}$$

and since n is of order $\log(1/\varepsilon)$, the right-hand-side inequality of (1.3) follows.

We now turn to the meat of this paper, that is, the upper bounds in (1.1) and (1.3). It will be convenient to use a special system of functions. Central to this system is the function ξ on $[0,1]$ given by

$$\xi(s) = \begin{cases} 1, & \text{if } 0 \le s < \frac{1}{4} \text{ or } \frac{3}{4} < s \le 1, \\ -1, & \text{if } \frac{1}{4} \le s \le \frac{3}{4}. \end{cases}$$

The idea is that, while ξ somewhat resembles a Haar function, the function $\eta(s) = \int_0^s \xi(t)\,dt$ also resembles a Haar function. We consider a parameter $q \ge 1$ to be fixed later on (actually $q = 9$ works, but using that specific value now might be distracting), and, for $m \ge 1$, $0 \le l < 2^{qm}$, consider the function

$$\xi_{m,l}(s) = 2^{qm/2}\xi\big(2^{qm}(s - l2^{-qm})\big).$$

It is routine to check that the functions $(\xi_{m,l})$ form an orthogonal system in $L^2([0,1])$. Thus, given n, the functions $\xi_{m,l} \otimes \xi_{n-m,l'}$ for

$$(m,l,l') \in T_n = \big\{(m,l,l');\ 0 \le m \le n-1;\ 0 \le l < 2^{qm};\ 0 \le l' < 2^{q(n-m)}\big\}$$

form an orthogonal system in $L^2(\lambda)$.

Thus, setting $\eta_{m,l} = U(\xi_{m,l})$, it follows from our discussion and Lemma 2.1 that

$$P\big(\|B_{s,t}\| \le \varepsilon\big) \le P\left(\left\| \sum_{(m,l,l') \in T_n} g_{m,l,l'}\eta_{m,l} \otimes \eta_{n-m,l'} \right\| \le \varepsilon\right), \tag{2.22}$$

where, of course, the r.v.'s $g_{m,l,l'}$ are independent standard normal. The main property of the functions $\eta_{m,l} \otimes \eta_{n-m,l'}$ is as follows.

PROPOSITION 2.4. *One can choose q such that for all $n \ge 1$ and all families of numbers $(\alpha_{m,l,l'})_{(m,l,l') \in T_n}$ one has*

$$\begin{aligned} \sup_{0 \le s,t \le 1} & \sum_{(m,l,l') \in T_n} \alpha_{m,l,l'}\eta_{m,l}(s)\eta_{n-m,l'}(t) \\ & \ge 2^{-3qn/2-7} \sum_{(m,l,l') \in T_n} |\alpha_{m,l,l'}|. \end{aligned} \tag{2.23}$$

This will be proved in Section 3. We now prove the right-hand side of (1.1). If we combine (2.22) and (2.23), we get

$$P\big(\|B_{s,t}\| \le \varepsilon\big) \le P\left(\sum_{(m,l,l') \in T_n} 2^{-3qn/2}|g_{m,l,l'}| \le 2^7\varepsilon \right).$$

We observe that $\operatorname{card} T_n = n2^{qn}$. Given independent standard normal r.v.'s $(g_i)_{i \le N}$, we observe that for a universal constant C_0, we have

$$P\left(\sum_{i \le N} |g_i| \le \frac{N}{C_0}\right) \le \exp\left(-\frac{N}{C_0}\right). \tag{2.24}$$

Indeed, by Chebyshev's exponential inequality, we have

$$P\left(\sum_{i \le N} |g_i| \le u\right) \le \exp u \, E \exp\left(-\sum_{i \le N} |g_i|\right),$$

so (2.24) follows, for example, by taking

$$u = -\frac{N}{2} \log E \exp\left(-|g|\right).$$

Using (2.24) with $N = \operatorname{card} T_n = n2^{qn}$, we see that if $2^{3qn/2-7}\varepsilon \le n2^{qn}/C_0$, that is, if

$$2^7 \varepsilon \le \frac{n2^{-qn/2}}{C_0}, \tag{2.25}$$

then

$$P\left(\|B_{s,t}\| \le \varepsilon\right) \le \exp\left(-\frac{n2^{qn}}{C_0}\right).$$

Now

$$\frac{n2^{qn}}{C_0} = \frac{n^3}{C_0^3\left(n2^{-qn/2}/C_0\right)^2}.$$

Thus, if one takes for n the largest possible value in (2.25), $n2^{qn}$ is indeed of order $\varepsilon^{-2}(\log(1/\varepsilon))^3$. This completes the proof of (1.1).

We now turn to the proof of the right-hand side of (1.3). Consider a sequence $\tau = (\alpha_{m,l,l'})_{(m,l,l') \in T_{n'}}$, and the function

$$\varphi_\tau(s,t) = \sum_{(m,l,l') \in T_n} \alpha_{m,l,l'} \eta_{m,l}(s) \eta_{n-m,l'}(t).$$

Set $|\tau| = \sum_{(m,l,l') \in T_n} |\alpha_{m,l,l'}|$. We observe that since $|\xi_{m,l}| \le 2^{qm/2}$, for $m \le n$, we have

$$|\eta_{m,l}(s) - \eta_{m,l}(s')| \le 2^{qm/2} |s - s'|.$$

It follows easily that

$$|\varphi_\tau(s,t) - \varphi_\tau(s',t')| \le |\tau| 2^{qn/2}\left(|s - s'| + |t - t'|\right).$$

Combining with (2.23), we see that

$$\lambda\left(\left\{(s,t) \colon \varphi_\tau(s,t) \ge \tfrac{1}{16} 2^{-3qn/2} |\tau|\right\}\right) \ge 2^{-4qn-10}. \tag{2.26}$$

For a function φ and $a > 0$, we have

$$\int\int_{[0,1]^2} \exp\left(\frac{\varphi}{c}\right)^{\alpha} d\lambda \geq \exp\left(\frac{a}{c}\right)^{\alpha} \lambda(\{\varphi \geq a\}).$$

Thus

$$\|\varphi\|_{\psi_\alpha} \geq a\left(\log \frac{2}{\lambda(\{\varphi \geq a\})}\right)^{-1/\alpha},$$

so that, by (2.26), we have

$$\|\varphi_\tau\|_{\psi_\alpha} \geq \frac{1}{Cn^{1/\alpha}} 2^{-3qn/2}|\tau| \tag{2.27}$$

for some C universal. We now replace $\alpha_{m,l,l'}$ by $g_{m,l,l'}$. Then, as already observed,

$$P\left(|\tau| \leq \frac{n2^{qn}}{C_0}\right) \leq \exp\left(-\frac{n2^{nq}}{C_0}\right),$$

so that

$$P\left(\|\varphi_\tau\|_{\psi_\alpha} \leq \frac{1}{CC_0} n^{1-1/\alpha} 2^{-qn/2}\right) \leq \exp\left(-\frac{n2^{nq}}{C_0}\right),$$

and the conclusion follows from the fact that, if $\varepsilon = n^{1-1/\alpha}2^{-qn/2}$, then

$$n2^{nq} = \frac{n^{3-2/\alpha}}{\left(n^{1-1/\alpha}2^{-qn/2}\right)^2}$$

is of order $\varepsilon^{-2}(\log(1/\varepsilon))^{3-2/\alpha}$. □

3. Combinatorics. Given $p \geq 1$, $0 \leq i < 2^p$, we denote by $J_{p,i}$ the dyadic interval

$$J_{p,i} = \left]i2^{-p}, (i+1)2^{-p}\right[,$$

and we denote by D_p the collection of the intervals $J_{p,i}$, for $0 \leq i < 2^p$. Consider the function u on $\mathbb{R}$ given by

$$u(x) = \begin{cases} 0, & \text{if } x \notin]0,1[, \\ x, & \text{if } 0 \leq x \leq \frac{1}{4}, \\ \frac{1}{2} - x, & \text{if } \frac{1}{4} \leq x \leq \frac{3}{4}, \\ x - 1, & \text{if } \frac{3}{4} \leq x \leq 1. \end{cases}$$

For $m \geq 1$, $0 \leq l < 2^{mq}$, we set

$$u_{m,l}(s) = u\left(2^{mq}(s - l2^{-mq})\right).$$

Thus $u_{m,l}$ is supported by $J_{qm,l}$. To prove Proposition 2.4, it certainly suffices (since $u_{m,l} = 2^{qm/2}\eta_{m,l}$) to prove the following.

THEOREM 3.1. *If $q = 9$, then, for each $n \geq 1$ and each family of numbers $(\alpha_{m,l,l'})_{(m,l,l') \in T_n}$, where*

$$T_n = \{(m, l, l');\ 0 \leq m \leq n-1,\ 0 \leq l < 2^{qm},\ 0 \leq l' < 2^{q(n-m)}\}, \tag{3.1}$$

we have

$$\sup_{0 \leq s,t \leq 1} \sum_{(m,l,l') \in T_n} \alpha_{m,l,l'} u_{m,l}(s) u_{n-m,l'}(t) \geq 2^{-qn-7} \sum_{(m,l,l') \in T_n} |\alpha_{m,l,l'}|. \tag{3.2}$$

The main idea of the proof of Theorem 3.1 is rather simple. Unfortunately, it is obscured by a number of technical problems. So, in order to make the idea transparent, we will first prove a theorem of the same nature (whose only purpose is to help the reader to penetrate Theorem 3.1) where the technical problems do not exist. Consider the function h on $\mathbb{R}$ given by

$$h(x) = \begin{cases} 0, & \text{if } x \notin [0,1[, \\ 1, & \text{if } x \in \left[0, \frac{1}{2}\right[, \\ -1, & \text{if } x \notin \left[\frac{1}{2}, 1\right[\end{cases}$$

and define $h_{m,l} = h(2^m(s - l2^{-m}))$.

PROPOSITION 3.2. *If $q = 1$, then, for each $n \geq 1$ and each $(\alpha_{m,l,l'})_{(m,l,l') \in T_n}$, we have*

$$\sup_{0 \leq s,t \leq 1} \sum_{(m,l,l') \in T_n} \alpha_{m,l,l'} h_{m,l}(s) h_{n-m,l'}(t) \geq 2^{-n-1} \sum_{(m,l,l') \in T_n} |\alpha_{m,l,l'}|. \tag{3.3}$$

PROOF. We denote by S the left-hand side of (3.3). For simplicity, we set, for $m < n$,

$$w_m = \sum_{0 \leq l < 2^m;\ 0 \leq l' < 2^{n-m}} \alpha_{m,l,l'} h_{m,l} \otimes h_{n-m,l'}.$$

For a function w on $[0,1]^2$, an interval I of $[0,1]$ and $t \in [0,1]$, we set

$$M(I,t)(w) = \frac{1}{|I|} \int_I w(s,t)\, ds,$$

where $|I|$ is the length of I.

Consider m such that $0 \leq m \leq n$. Consider an interval $I \in D_m$ and, for $0 \leq j < 2^{n-m+1}$, consider a number $t_j \in J_{n-m+1,j}$. Consider the following statement:

(H_m): For all possible choices of I in D_m and of points t_j, we have

$$\sum_{0 \le j < 2^{n-m+1}} M(I, t_j)\left(\sum_{0 \le r < m} w_r + \sum_{m \le r < n} |w_r|\right) \le 2^{n-m+1} S. \tag{3.4}$$

We note that, for $m = n$, this is an immediate consequence of the definition of S, by integration. We also note that, for $m = 0$, (H_m) reduces to

$$\sum_{j \le 2^{n+1}} M([0,1], t_j)\left(\sum_{r \ge 0} |w_r|\right) \le 2^{n+1} S, \tag{3.5}$$

from which (3.3) follows by integration over all possible choices of t_j.

Thus our task reduces to proving (H_m) by decreasing induction over m. Assuming (H_{m+1}), we prove (H_m). Consider $I \in D_m$ and, for $j < 2^{n-m+1}$, consider $t_j \in J_{n-m+1, j}$. We set $I = I_0 \cup I_1$, where $I_0 \in D_{m+1}$, $I_1 \in D_{m+1}$, in such a way that, if $I = [2^{-m}\tau, 2^{-m}(\tau+1)[$, we have

$$I_0 = \left[2^{-m}\tau,\ 2^{-m}\left(\tau + \tfrac{1}{2}\right)\right[, \qquad I_1 = \left[2^{-m}\left(\tau + \tfrac{1}{2}\right), 2^{-m}(\tau+1)\right[.$$

For $i < 2^{n-m}$, we set

$$x_i^0 = t_{2i}, \qquad x_i^1 = t_{2i+1}$$

if $\alpha_{m,\tau,j} \ge 0$, and

$$x_i^0 = t_{2i+1}, \qquad x_i^1 = t_{2i}$$

if $\alpha_{m,\tau,j} < 0$. The point of this choice is that $w_m \ge 0$ on $I_\varepsilon \times \{x_i^\varepsilon\}$, so that

$$M(I_\varepsilon, x_i^\varepsilon)(w_m) = M(I_\varepsilon, x_i^\varepsilon)(|w_m|). \tag{3.6}$$

By the induction hypothesis, (H_{m+1}) holds, so we have, for $\varepsilon \in \{0,1\}$,

$$\sum_{0 \le i < 2^{n-m}} M(I_\varepsilon, x_i^\varepsilon)\left(\sum_{0 \le r \le m} w_r + \sum_{r \ge m+1} |w_r|\right) \le 2^{n-m} S.$$

Thus, by (3.6),

$$\sum_{0 \le i < 2^{n-m}} M(I_\varepsilon, x_i^\varepsilon)\left(\sum_{0 \le r < m} w_r + \sum_{r \ge m} |w_r|\right) \le 2^{n-m} S. \tag{3.7}$$

We now observe the following. First, for all $t \in [0,1]$, $r < m$, we have

$$M(I_0, t)(w_r) = M(I_1, t)(w_r). \tag{3.8}$$

(A statement of the same nature will be detailed in Lemma 3.3.) Thus

$$M(I, x_i^\varepsilon)(w_r) = M(I_\varepsilon, x_i^\varepsilon)(w_r)\left(= \tfrac{1}{2}\left(M(I_0, x_i^\varepsilon)(w_r) + M(I_1, x_i^\varepsilon)(w_r)\right)\right). \tag{3.9}$$

Next, we note that, if $r \geq m$, we have, for $\varepsilon \in \{0, 1\}$,

$$M(I_\varepsilon, x_i^0)(|w_r|) = M(I_\varepsilon, x_i^1)(|w_r|), \tag{3.10}$$

so that [since $2M(I, x) = M(I_0, x) + M(I_1, x)$], for $\varepsilon \in \{0, 1\}$,

$$M(I, x_i^\varepsilon)(|w_r|) = M(I_\varepsilon, x_i^\varepsilon)(|w_r|). \tag{3.11}$$

Thus, if we sum (3.7) for $\varepsilon = 0$, $\varepsilon = 1$ and use (3.9) and (3.11), we have completed the induction. □

PROOF OF THEOREM 3.1. We now set, for $m < n$,

$$w_m = \sum_{0 \leq l < 2^m;\ 0 \leq l' < 2^{n-m}} \alpha_{m,l,l'} u_{m,l} \otimes u_{n-m,l'}. \tag{3.12}$$

The main obstacle in carrying out the method of Proposition 3.2 is that (3.9) and (3.11) are now only approximately true (the error being a decreasing function of $|r - m|$). Thus we will have to modify (H_m) in order to introduce error terms. The role of the parameter q is to ensure that the contribution of these error terms is sufficiently small.

We denote by $v_{m,l,l'}$ the indicator of the set $J_{qm,l} \times J_{(n-m)q,l'}$, and we set

$$\theta_m = \sum_{0 \leq l < 2^{mq};\ 0 \leq l' < 2^{(n-m)q}} |\alpha_{m,l,l'}| v_{m,l,l'}.$$

We set

$$b_{r,m} = \sum_{i \neq r;\ i \geq m} 2^{-q|r-i|}.$$

We consider, for $0 \leq m \leq n$, the following statement, where

$$S = \sup_{0 \leq s,t \leq 1} \sum_{(m,l,l') \in T_n} \alpha_{m,l,l'} u_{m,l}(s) u_{n-m,l'}(t).$$

(H_m) Given $I \in D_{qm}$ and, for $0 \leq j < 2^{(n-m)q+1}$ numbers $t_j \in J_{(n-m)q+1,j}$, we have

$$\sum_{0 \leq j < 2^{(n-m)q+1}} M(I, t_j)\left(\sum_{r<m} w_r + \sum_{r \geq m} |w_r|\right)$$
$$\leq 2^{(n-m)q+1} S + \sum_{1 \leq r \leq n} b_{r,m} \sum_{0 \leq j < 2^{(n-m)q+1}} M(I, t_j)(\theta_r).$$

For $m = 0$, since $b_{r,0} \leq 2^{-q+1}/(1 - 2^{-q}) \leq 2^{-q+2}$, this implies

$$\sum_j M([0,1], t_j)\left(\sum_{r \geq 0} |w_r|\right) \leq 2^{nq+1} S + 2^{-q+2} \sum_{r \geq 0} \sum_j M([0,1], t_j)(\theta_r).$$

We now average this inequality over all possible values of t_j. We observe that the average value of $\sum_j M([0,1]t_j)(f)$ is $2^{nq+1}\iint_{[0,1]^2} f\,d\lambda$. We also observe that $\int_0^1 |u(t)|\,dt = 2^{-3}$, so that

$$\iint_{[0,1]^2} |u_{m,l} \otimes u_{n-m,l'}|\,d\lambda = 2^{-nq-6},$$

and also $\iint_{[0,1]^2} v_{m,l,l'}\,d\lambda = 2^{-nq}$, so that we get

$$2^{-t} \sum_{(r,l,l')\in T_n} |\alpha_{r,l,l'}| \le 2^{qn+1}S + 2^{-q+3} \sum_{(r,l,l')\in T_n} |\alpha_{r,l,l'}|.$$

Thus, taking $q = 9$, so that $2^{-q+3} = 2^{-6}$, we indeed get (3.2).

For $m = n$ (since there is no term $|w_r|$ for $r \ge n$), it is obvious that (H_n) holds. Thus our task is again to prove (H_m) by decreasing induction over m.

Assuming that (H_{m+1}) holds, we prove (H_m). Consider $I \in D_{qm}$ and, for $0 < j \le 2^{(n-m)q+1}$, consider $t_j \in J_{(n-m)q+1,j}$. Consider τ such that $J = J_{mq,\tau}$ and write $I_0 = J_{mq+1,2\tau}$, $I_1 = J_{mq+1,2\tau+1}$, so that $I = I_0 \cup I_1$.

For $i < 2^{(n-m)q}$, we define

$$x_i^0 = t_{2i}, \qquad x_i^1 = t_{2i+1} \quad \text{if } \alpha_{m,\tau,i} \ge 0,$$
$$x_i^0 = t_{2i+1}, \qquad x_i^1 = t_{2i} \quad \text{if } \alpha_{m,\tau,i} < 0.$$

We note that

$$M(I_\varepsilon, x_i^\varepsilon)(w_m) = M(I_\varepsilon, x_i^\varepsilon)\big(|w_m|\big), \tag{3.13}$$

since $w_m \ge 0$ on $I_\varepsilon \times \{x_i^\varepsilon\}$.

LEMMA 3.3. *For $\varepsilon \in \{0,1\}$, we have*

$$\sum_{i<2^{(n-m)q+1}} M(I_\varepsilon, x_i^\varepsilon)\left(\sum_{r<m} w_r + \sum_{r\ge m} |w_r|\right) \le 2^{(n-m)q}S + \sum_{1\le r<n} b_{r,m+1} \sum_{i<2^{(n-m)q+1}} M(I_\varepsilon, x_i^\varepsilon)(\theta_r). \tag{3.14}$$

PROOF. Observe first that, in view of (3.13), it suffices to prove this when $\sum_{r<m} w_r + \sum_{r\ge m}|w_r|$ is replaced by $\sum_{r<m+1} w_r + \sum_{r\ge m+1}|w_r|$. Also, it suffices to prove (3.14) when I_ε is replaced by an interval $I' \subset I_\varepsilon$, $I' \in D_{q(m+1)}$. [Inequality (3.14) then follows by averaging over all possible choices of I'.]

Given $0 \le i < 2^{(n-m-1)q+1}$ and $0 \le k < 2^{q-1}$, we define

$$y_i^k = x^\varepsilon_{2^{q-1}i+k} \in J_{q(n-m),2^{q-1}i+k} \subset J_{q(n-m-1)+1,i}.$$

For each k, we then apply (H_{m+1}) to the interval I' and the points $(y_i^k)_{i<2^{q(n-m-1)+1}}$. Summation over k concludes the proof. □

We now sum the relations (3.14) for $\varepsilon = 0$ and $\varepsilon = 1$, to get

$$
\begin{aligned}
&\sum_{j < 2^{(n-m)q+1}} \Big(M\big(I_0, x_j^0\big) + M\big(I_1, x_j^1\big)\Big)\Bigg(\sum_{r<m} w_r + \sum_{r \geq m} |w_r|\Bigg) \\
&\quad \leq 2^{(n-m)q+1} S + \sum_{1 \leq r < n} b_{r,m+1} \\
&\qquad \times \sum_{j < 2^{(n-m)q+1}} \Big(M\big(I_0, x_j^0\big)(\theta_r) + M\big(I_1, x_j^1\big)(\theta_r)\Big).
\end{aligned} \tag{3.15}
$$

The rest of the proof consists of deducing (H_m) from (3.15).

LEMMA 3.4. *For all* $j < 2^{q(n-m)+1}$, *we have*

$$
\begin{aligned}
M\big(I_0, x_j^0\big)(\theta_r) + M\big(I_1, x_j^1\big)(\theta_r) &= M\big(I, x_j^0\big)(\theta_r) + M\big(I, x_j^1\big)(\theta_r) \\
&= M\big(I, t_{2j}\big)(\theta_r) + M\big(I, t_{2j+1}\big)(\theta_r).
\end{aligned} \tag{3.16}
$$

PROOF. We give separate arguments for $r \leq m$ and $r \geq m$ (both arguments work for $r = m$).

Case 1. $r \leq m$. For each t, the function $s \to \theta_r(s,t)$ is D_{qr}-measurable, so it is D_{qm}-measurable. Thus

$$M(I_\varepsilon, t)(\theta_r) = M(I, t)(\theta_r)$$

for each $\varepsilon \in \{0, 1\}$.

Case 2. $r \geq m$. For each s, the function $t \to \theta_r(s,t)$ is $D_{(n-r)q}$-measurable, so it is $D_{(n-m)q}$-measurable. Thus is still true of the function $t \to M(I_0, t)(\theta_r)$. Since both x_j^0 and x_j^1 belong to $J_{(n-m)q,j}$, we have

$$M\big(I_0, x_j^0\big)(\theta_r) = M\big(I_0, x_j^1\big)(\theta_r).$$

In a similar way, we have

$$M\big(I_1, x_j^0\big)(\theta_r) = M\big(I_1, x_j^1\big)(\theta_r),$$

from which the result follows since

$$M(I, t) = \tfrac{1}{2}\big(M(I_0, t) + M(I_1, t)\big). \qquad \square$$

LEMMA 3.5. *For* $\varepsilon \in \{0, 1\}$,

$$M\big(I_\varepsilon, x_j^\varepsilon\big)\big(|w_m|\big) = M\big(I, x_j^\varepsilon\big)\big(|w_m|\big).$$

PROOF. This is a simple consequence of the fact that $|u(x + 1/2)| = |u(x)|$. $\square$

LEMMA 3.6. *If $r < m$, we have*

$$M(I_\varepsilon, x_j^\varepsilon)(w_r) \geq M(I, x_j^\varepsilon)(w_r) - 2^{-q(m-r)-2} M(I, x_j^\varepsilon)(\theta_r).$$

PROOF. Since $M(I,t) = \frac{1}{2}(M(I_0,t) + M(I_1,t))$ for all t, we have

$$\begin{aligned}
&|M(I_\varepsilon, x_j^\varepsilon)(w_r) - M(I, x_j^\varepsilon)(w_r)| \\
&\quad \leq \frac{1}{2}|M(I_0, x_j^\varepsilon)(w_r) - M(I_1, x_j^\varepsilon)(w_r)| \\
&\quad = \frac{1}{2}\frac{1}{|I_0|}\left| \int \left(w_r(s, x_j^\varepsilon) - w_r(s + 2^{-qm-1}, x_j^\varepsilon)\right) ds \right| \qquad (3.17)\\
&\quad \leq \frac{1}{2} \sup_{s \in I_0} |w_r(s, x_j^\varepsilon) - w_r(s + 2^{-qm-1}, x_j^\varepsilon)|.
\end{aligned}$$

Consider l such that $I \subset J_{r,l}$. For $s \in I$, we have $w_r(s, x_j^\varepsilon) = \alpha_{r,l,j} u_{r,l}(s) u_{r,j}(x_j^\varepsilon)$. Since $|u| \leq 1$, $|u(s) - u(t)| \leq |s - t|$, we see that

$$|M(I_\varepsilon, x_j^\varepsilon)(w_r) - M(I, x_j^\varepsilon)(w_r)| \leq \tfrac{1}{4} 2^{q(r-m)} |\alpha_{r,l,j}|,$$

and the conclusion follows from the fact that $|\alpha_{r,l,j}| = M(I, x_j^\varepsilon)(\theta_r)$. □

LEMMA 3.7. *If $r > m$, we have*

$$\begin{aligned}
M(I, x_j^0)(|w_r|) + M(I, x_j^1)(|w_r|) &\leq M(I_0, x_j^0)(|w_r|) + M(I_1, x_j^1)(|w_r|) \\
&\quad + 2^{-q(r-m)}\left(M(I, x_j^0)(\theta_r) + M(I, x_j^1)(\theta_r)\right).
\end{aligned}$$

PROOF. We note that

$$\begin{aligned}
&|M(I, x_j^0)(|w_r|) + M(I, x_j^1)(|w_r|) - M(I_0, x_j^0)(|w_r|) - M(I_1, x_j^1)(|w_r|)| \\
&\quad \leq \frac{1}{2}|M(I_1, x_j^0)(|w_r|) - M(I_1, x_j^1)(|w_r|)| \\
&\quad\quad + \frac{1}{2}|M(I_0, x_j^0)(|w_r|) - M(I_0, x_j^1)(|w_r|)| \\
&\quad \leq \frac{1}{2}\frac{1}{|I_1|}\int_{I_1} |w_r(s, x_j^0) - w_r(s, x_j^1)|\, ds + \frac{1}{2}\frac{1}{|I_0|}\int_{I_0} |w_r(s, x_j^0) - w_r(s, x_j^1)|\, ds.
\end{aligned}$$

Now, if $s \in J_{r,p(s)}$, we have

$$\begin{aligned}
|w_r(s, x_j^0) - w_r(s, x_j^1)| &\leq |\alpha_{r,p(s),j}| 2^{(n-r)q} |x_j^0 - x_j^1| \\
&\leq |\alpha_{r,p(s),j}| 2^{(n-r)q} 2^{-(n-m)q} \\
&\leq |\alpha_{r,p(s),j}| 2^{-(r-m)q}.
\end{aligned}$$

Since

$$|\alpha_{r,p(s),j}| = \theta_r(s, x_j^0) = \theta_r(s, x_j^1),$$

the conclusion follows easily. □

We now prove (H_m) and conclude the proof. Combining (3.15) and Lemmas 3.4 to 3.7 and setting

$$\Delta_{j,r} = M(I, x_j^0)(\theta_r) + M(I, x_j^1)(\theta_r),$$

we see the left-hand side of (3.15) is at most

$$2^{(n-m)q+1}S + \sum_{1 \le r < n} b_{r,m+1} \sum_{j < 2^{(n-m)q+1}} \Delta_{j,r}$$
$$+ \sum_{1 \le r < n,\ r \ne m} 2^{-q|r-m|} \sum_{j < 2^{(n-m)q+1}} \Delta_{j,r}.$$

But this is exactly the right-hand side of (3.14). □

4. Further results. Consider a positive measure μ on $[0,1]^2$. We can consider the Gaussian process $B_{s,t} = B^{\mu}_{s,t}$ on $[0,1]^2$ such that

$$E(B_{s,t}B_{s',t'}) = \mu\Big([0, \min(s,s')] \times [0, \min(t,t')]\Big).$$

Equivalently, the law of $B_{s,t}$ in $C([0,1]^2)$ is the image of the canonical cylindrical Gaussian measure on $L^2(\mu)$ under T.

PROPOSITION 4.1. *For some universal constant C, have, for all $\varepsilon \le 1/2$,*

$$P\Big(\sup_{s,t \le 1} |B^{\mu}_{s,t}| \le \varepsilon \|\mu\|^{1/2}\Big) \ge \exp\bigg(-\frac{C}{\varepsilon^2}\Big(\log\frac{1}{\varepsilon}\Big)^3\bigg). \tag{4.1}$$

PROOF.

Step 1. To prove (4.1), it suffices to prove it when the supremum over $s, t \le 1$ is replaced by the supremum for $(s,t) \in F$, where F is an arbitrary finite subset of $[0,1]^2$. The law of $(B_{s,t})_{(s,t)\in F}$ depends only on the numbers $\mu([0,s] \times [0,t])$ for $s, t \in G$, where G is a certain finite set. It is then simple to see that there is no loss of generality to assume that both marginals of μ are atomless. This assumption will simplify the notation.

Step 2. For $k \ge 1$, $0 \le l \le 2^k$, we consider the point $a(k,l)$ such that

$$\mu\Big([0, a(k,l)] \times [0,1]\Big) = l2^{-k}\|\mu\|.$$

For $k \geq 1$, $0 \leq r < 2^{k-1}$, we consider the r.v.

$$Z_{k,r}(t) = B_{a(2r+1,k),t} - B_{a(2r,k),t}. \tag{4.2}$$

We claim that

$$\sup_{s,t} |B_{s,t}| \leq \sum_k \sup_{0 \leq r < 2^{k-1}} \sup_{0 \leq t \leq 1} |Z_{k,r}(t)|. \tag{4.3}$$

This is a consequence of the fact that

$$|B_{s,t}| \leq \sum_{k \in I(s)} |Z_{k,r_k}(t)|,$$

where, setting $a = \mu([0,s] \times [0,1])$, we have

$$I(s) = \left\{k;\ 2^{-k}[2^k a] > 2^{-k+1}\left[2^{k-1}a\right]\right\}$$

and $r_k = 2^{-k+1}[2^{k-1}a]$ for $k \in I(s)$.

Step 3. We fix $n \geq 1$, and we set $\varepsilon_k = 2^{-n}\|\mu\|^{1/2}$ for $k \leq 2n$, $\varepsilon_k = 2^{-(k-2n)/4}\varepsilon_{2n}$ for $k > 2n$. For $0 \leq r < 2^{k-1}$, consider the events

$$\Omega_{n,k,r} = \left\{\sup_{t \leq 1} |Z_{k,r}(t)| \leq \varepsilon_k\right\},$$

$$\Omega_{n,k} = \bigcap_{0 \leq r < 2^{k-1}} \Omega_{n,k,r}, \qquad \Omega_n = \bigcap_{k \geq 1} \Omega_{n,k}.$$

Using (4.3), we see that on Ω_n we have $\sup_{s,t \leq 1} |B_{s,t}| \leq Cn2^{-n}\|\mu\|^{1/2}$. Since $n2^{2n} = n^3(n2^{-n})^{-2}$, it suffices to show that

$$P(\Omega_n) \geq \exp\left(-Cn2^{2n}\right). \tag{4.4}$$

Let us fix k, r and set

$$\varphi(t) = \mu\left(\left[a(2r,k), a(2r+1,k)\right] \times [0,t]\right).$$

We note that $Z_{k,r(t)}$ is a Gaussian process that satisfies

$$E\left(Z_{k,r}(t)Z_{k,r}(t')\right) = \varphi\left(\min(t,t')\right).$$

Thus if B_t denotes standard Brownian motion, the process $(Z_{k,r}(t))_{t \leq 1}$ is distributed like $(B_{\varphi(t)})_{t \leq 1}$.

Thus we can use the well-known estimates on $P(\sup_{t \leq b} |B_t| \leq \varepsilon)$ to estimate $P(\Omega_{n,k,r})$. The events $(\Omega_{n,k,r})_{r < 2^{k-1}}$ are independent; but unfortunately the events $(\Omega_{n,k})_k$ are *not* independent. Thus a little detour is necessary. The idea is to construct Gaussian variables $(X_{n,k,r,l})_{l\geq 1}$ and numbers $\beta_{n,k,l}$ with the following properties:

$$\bigcap_{l\geq 1}\{|X_{n,k,r,l}|\leq \beta_{n,k,l}\}\subset \Omega_{n,k,r}, \tag{4.5}$$

$$\prod_{k,l\geq 1,\, r<2^{k-1}} P(|X_{n,k,r,l}|\leq \beta_{n,k,l})\geq \exp(-Cn2^{2n}). \tag{4.6}$$

It then follows from (4.5) that

$$\bigcap_{k,l\geq 1,\, r<2^{k-1}}\{|X_{n,k,r,l}|\leq \beta_{n,k,l}\}\subset \Omega_n. \tag{4.7}$$

A theorem of Šidák [5] implies that the probability of the left-hand side of (4.7) is larger than the left-hand side of (4.6). Thus (4.4) follows from (4.6).

Step 4. For clarity, given $\varepsilon > 0$, $b > 0$, we show the existence of Gaussian random variables $(X_l)_{l\geq 1}$ and numbers δ_l such that

$$\begin{aligned} &\forall\, l\geq 1,\ |X_l|\leq \delta_l \Rightarrow \sup_{t\leq b}|B_t|\leq \varepsilon\sqrt{b},\\ &\prod_{l\geq 1}P(|X_l|\leq \delta_l)\geq h(\varepsilon), \end{aligned}\tag{4.8}$$

where $h(\varepsilon) = \exp(-C/\varepsilon^2)$ if $\varepsilon \leq 1$, $h(\varepsilon) = 1-\exp(-\varepsilon^2/C)$ if $\varepsilon\geq 1$. The variables $X_{n,k,r,l}$ are then obtained by applying the same procedure to the process $Z_{k,r}$, for $b = \|\mu\|2^{-k}$, $\varepsilon\sqrt{b} = \varepsilon_k$, and we take $\beta_{n,k,l} = \delta_n$. To prove (4.6), we observe that by (4.8) the left-hand side of (4.6) is bounded below by

$$\prod_{k\leq 2n}\left(\exp\left(-\frac{C2^{2n}}{2^k}\right)\right)^{2^{k-1}}\prod_{k>2n}\left(1-\exp\left(-\frac{2^{(k-2n)/2}}{C}\right)\right)^{2^{k-1}},$$

from which (4.6) follows by a simple computation.

Step 5. The procedure to construct the variables (X_l) uses the same idea as step 2. The variables (X_l) are an enumeration of the variables $W_{p,q} = B_{2^{-p}(q+1/2)b} - B_{2^{-p}qb}$. When $X_l = W_{p,q}$, then δ_l is given as follows. If $\varepsilon < 1$, consider p_0 with $2^{-p_0/2}\sim\varepsilon$ and set $\delta_l = 2^{-p_0/2-|p-p_0|/4}b^{1/2}/C$ (C universal large enough). If $\varepsilon\geq 1$, set $\delta_l = \varepsilon 2^{-l/4}/C$. The details of the computations are left to the reader. □

We now extend the upper bound of Theorem 1.1.

THEOREM 4.2. *Assume that μ is not singular. Then, for all $\varepsilon\leq 1$, we have*

$$P\Big(\sup_{s,t\leq 1}|B^{\mu}_{s,t}|\leq \varepsilon C(\mu)\Big)\leq \exp\left[-\frac{1}{C(\mu)\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^3\right],$$

where $C(\mu)$ depends only on μ.

PROOF.

Step 1. Consider a parameter $0 < \alpha < 1$ to be determined later. Since we assume that μ is not singular, we can find $\beta > 0$, a square $C = [a_0, b_0] \times [a_1, b_1]$ in $[0,1]^2$ and a subset A of C such that $\mu = \mu_1 + \mu_2$, where μ_2 is positive and μ_1 has a density $\beta 1_A$ with respect to λ and where

$$\lambda(C \backslash A) \leq \alpha \lambda(C). \tag{4.9}$$

Step 2. We let the reader show that if μ_1, μ_2 are two positive measures on $[0,1]^2$, we have

$$\left(B_{s,t}^{\mu_1+\mu_2}\right)_{s,t} \overset{\mathcal{D}}{=} \left(B_{s,t}^{\mu_1} + B_{s,t}^{\mu_2}\right)_{s,t}, \tag{4.10}$$

where the two processes on the right are independent. Thus we can assume by Lemma 2.1 that $\mu = \mu_1$, where μ_1 was described in step 1.

Step 3. It will simplify the notation and not reduce the generality to assume that $\beta = 1$, $C = [0,1]^2$. Consider the measure μ_3 of density $1_{[0,1]^2 \backslash A}$ with respect to λ. Thus $\lambda = \mu_1 + \mu_3$. Using (4.10) for $\mu_1, \mu_3, \lambda = \mu_1 + \mu_3$, we see that

$$P\left(\sup_{s,t \leq 1} |B_{s,t}^{\mu_1}| \leq \varepsilon\right) P\left(\sup_{s,t \leq 1} |B_{s,t}^{\mu_3}| \leq \varepsilon\right) \leq P\left(\sup_{s,t \leq 1} |B_{s,t}| \leq 2\varepsilon\right).$$

Thus, using (1.1) and Proposition 4.1, we have

$$P\left(\sup_{s,t \leq 1} |B_{s,t}^{\mu_1}| \leq \varepsilon\right) \leq \exp\left(-\frac{1}{C\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^3\right) \exp\left(\frac{C_1}{\eta^2}\left(\log\frac{1}{\eta}\right)^3\right), \tag{4.11}$$

where $\eta = \varepsilon \alpha^{-1/2}$. Thus we see that if we have chosen α such that $\alpha = (2CC_1)^{-1}$, the right-hand side of (4.11) is bounded by

$$\exp\left(-\frac{1}{C'\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^3\right)$$

for ε small enough. □

Consider a centered Gaussian measure μ on a Banach space X. This measure can be used to define a "Brownian motion" $B_{\mu,t}$ on $C([0,1], X)$. For our purposes, the best way to do this is by the series representation

$$B_{\mu,t} = \sum_{k,l} 2^{-k/2} w_{k,l}(t) Y_{k,l}.$$

There, $k \geq 0$, $0 \leq l \leq 2^{k-1}$, the r.v.'s $Y_{k,l}$ are valued in X, independent, of law μ, $w_{0,0}(t) = t$,

$$w_{k,l}(t) = w\big(2^k(t - l2^{-k})\big)$$

and

$$w(t) = \begin{cases} 0, & \text{if } t \notin [0,1], \\ t, & \text{if } 0 \leq t \leq \frac{1}{2}, \\ 1-t, & \text{if } \frac{1}{2} \leq t \leq 1. \end{cases}$$

When μ is the canonical Gaussian measure on $\mathbb{R}, B_\mu$ is the usual Brownian motion. When μ is the usual Brownian motion, B_μ identifies the Brownian sheet. A natural question is to ask about the behavior of $P(\sup_t \|B_{\mu,t}\| \leq \varepsilon)$. Certainly this behavior is related to the behavior of $\mu(\{x;\ \|x\| \leq \varepsilon\})$; but, unfortunately, it is not determined by it. An instructive example is when $X = c_0$ and where μ is the law of $\sum_{n \geq 1} n^{-1/2} g_n e_n$, where e_n denotes the canonical basis and where the sequence g_n is independent standard normal. Straightforward estimates (in the spirit of those presented before) show that, for $\varepsilon \leq \frac{1}{2}$,

$$\frac{1}{C\varepsilon^2} \leq -\log \mu\Big(\big\{\|x\| \leq \varepsilon\big\}\Big) \leq \frac{C}{\varepsilon^2}. \tag{4.12}$$

Thus, with respect to the measure of the small balls, this is the same behavior as Brownian motion. However, in that case it is simple to show that

$$\frac{1}{C\varepsilon^2}\log\frac{1}{\varepsilon} \leq -\log P\Big(\sup_t \|B_{\mu,t}\| \leq \varepsilon\Big) \leq \frac{C}{\varepsilon^2}\log\frac{1}{\varepsilon},$$

a result to be contrasted with (1.3). A natural question is then to ask about the possible behaviors of $P(\sup_t \|B_{\mu,t}\| \leq \varepsilon)$, given the behavior of $\mu(\{x;\ \|x\| \leq \varepsilon\})$. It turns out that the most interesting case is where (4.12) holds. In that case, it can be shown that, for a constant C' depending on C only, we have

$$\frac{1}{C'\varepsilon^2}\log\frac{1}{\varepsilon} \leq -\log P\Big(\sup_{t \leq 1} \|B_{\mu,t}\| \leq \varepsilon\Big) \leq \frac{C'}{\varepsilon^2}\Big(\log\frac{1}{\varepsilon}\Big)^3,$$

a result that extends the estimates of Bass [2] to a considerably more general setting.

Acknowledgment. I am much indebted to M. Ledoux who attracted my attention to this very enjoyable problem.

REFERENCES

[1] ANDERSON, T. W. (1955). The integral of symmetric unimodular functions over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6** 170–176.

[2] BASS, R. (1988). Probability estimates for multiparameter Brownian processes. *Ann Probab.* **16** 251–264.

[3] KUELBS, J. and LI, W. (1993). Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.* **116** 133–157.

[4] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in a Banach Space.* Springer, New York.

[5] ŠIDÁK, Z. (1966). On multivariate normal probability of rectangles: their dependence on correlation. *Ann. Math. Statist.* **39** 1425–1434.

EQUIPE D'ANALYSE-TOUR 46
U.A. AU C.N.R.S. NO. 754
UNIVERSITÉ PARIS VI
4 PL JUSSIEU
75230 PARIS CEDEX 05
FRANCE

DEPARTMENT OF MATHEMATICS
OHIO STATE UNIVERSITY
231 WEST 18TH AVENUE
COLUMBUS, OHIO 43210