The Annals of Probability
1990, Vol. 18, No. 1, 1–49

SAMPLE BOUNDEDNESS OF STOCHASTIC PROCESSES UNDER INCREMENT CONDITIONS[1]

BY MICHEL TALAGRAND

University of Paris VI and The Ohio State University

Let (T, d) be a compact metric space of diameter D, and $\|\cdot\|_\Phi$ be an Orlicz norm. When is it true that all (separable) processes $(X_t)_{t\in T}$ that satisfy the increment condition $\|X_t - X_s\|_\Phi \le d(t, s)$ for all s, t in T are sample bounded? We give optimal necessary conditions and optimal sufficient conditions in terms of the existence of a probability measure m on T that satisfies an integral condition $\int_0^D f(\varepsilon, m(B(x, \varepsilon)))\, d\varepsilon \le K$ for each x in T, where f is a function suitably related to Φ. When T is a compact group and d is translation invariant, we are able to compute the necessary and sufficient condition in several cases.

Table of contents

1. Introduction. Consider a metric space (T, d). A stochastic process $(X_t)_{t\in T}$ is a collection of random variables (r.v.) indexed by T defined on a probability space (Ω, Σ, P). A rather important problem can be stated, somewhat imprecisely, as follows. Suppose that we have some control on the increments $X_t - X_u$, that is, we have some smallness property of these r.v.'s depending on $d(t, u)$. What can we say about the regularity of the trajectories of the process? This has of course received considerable attention. The fundamental results of Kolmogorov are still much used. Along the same line as Kolmogorov's conditions is the following optimal result, due to Ibragimov [7], Klass and Hahn [8] and Kôno [9]. If $T = [0, 1]$, and the distance d is given by $d(t, u) = \eta(|t - u|)$ for some increasing concave function η, then we have Theorem 1.1.

THEOREM 1.1. *Let $p > 1$. All processes satisfying $\|X_t - X_u\|_p \le d(t, u)$ are sample bounded if and only if the integral $\int_0^1 \eta(\varepsilon)/\varepsilon^{1+1/p}\, d\varepsilon$ is finite.*

While this is an important result, its true nature is obscured by the use of the special properties of the index set. Much understanding has been gained by

Received September 1987; revised January 1989.
[1]Work partially supported by an NSF grant.
AMS 1980 *subject classifications*. Primary 60G17; secondary 28A99, 46E30.
Key words and phrases. Sample boundedness, moment conditions, majorizing measures.

looking at the problem in an abstract way. An important tool has been the use of the covering numbers $N(\varepsilon) = N(\varepsilon, T, d)$ of T, that is, the smallest number of open balls of radius ε needed to cover T. Consider now a Young function Φ, that is, a positive, symmetric, convex function with $\Phi(0) = 0$. For a random variable X, we define its Orlicz norm $\|X\|_\Phi$ by

$$\|X\|_\Phi = \operatorname{Inf}\left\{\lambda > 0;\ E\left(\Phi\left(\frac{X}{\lambda}\right)\right) \le 1\right\}$$

(if no such λ exists, $\|X\|_\Phi = \infty$). The corresponding space L_Φ is the set of r.v.'s X for which $\|X\|_\Phi < \infty$. Consider the increment condition

$$(*) \qquad \forall\, t, u \in T, \qquad \|X_t - X_u\|_\Phi \le d(t, u)$$

or, equivalently,

$$\forall\, t, u \in T, \qquad E\Phi\left(\frac{X_t - X_u}{d(t, u)}\right) \le 1.$$

We denote by $D = D(T)$ the diameter of T. We then have the following important result.

THEOREM 1.2. *For any process satisfying* $(*)$, *we have*

$$E\Big(\operatorname*{Sup}_{t, u \in T}(X_t - X_u)\Big) \le K\int_0^D \Phi^{-1}(N(\varepsilon))\, d\varepsilon.$$

Moreover, if the right-hand side is finite, $(X_t)_{t \in T}$ *has a version that is sample continuous.*

By $\operatorname{Sup}_{t, u \in T}(X_t - X_u)$ we mean throughout the paper the essential supremum of the (in general uncountable) collection of r.v. $X_t - X_u$. This quantity is natural to introduce; since we have only information on the increments $X_t - X_u$, we know nothing about the variables X_t themselves. We could also use the quantity $\operatorname{Sup}_{t \in T}(X_t - X_{t_0})$ for a fixed t_0 in T. In the sub-Gaussian case $[\Phi(x) = \exp x^2) - 1]$, Theorem 1.2 is due to Dudley [3], and plays an important role in the development of the theory of Gaussian processes. The general case is more delicate, and was obtained (with minor differences) independently by Kôno [10] and Pisier [12, 13]. Theorem 1.2 has been generalized by Fernique [5] and Weber [15], who replaced condition $(*)$ by more subtle conditions on the tails of $X_t - X_u$; this is definitely more appropriate in some situations (e.g., the study of p-stable processes). Their conditions however are also based on the covering numbers $N(\varepsilon, T, d)$. Unfortunately, covering numbers are not a very precise way to describe a metric space; in particular, they give too much importance to the parts of the space that are sparse. So, while Theorem 1.2 is fairly sharp, the finiteness of the integral $\int_0^D \Phi^{-1}(N(\varepsilon))\, d\varepsilon$ is by no means necessary to ensure a good behavior of the processes satisfying $(*)$. (The typical situation where this condition is too strong is when T consists of a converging sequence and its limit.) The purpose of this paper is to investigate the use of a more precise tool, the existence of special probability measures on T that are called majorizing measures. Typically in this paper we call a probability m a majorizing measure if it satisfies a condition of the type

$$\operatorname{Sup}_{x \in T} \int_0^D f(\varepsilon, m(B(x, \varepsilon)))\, d\varepsilon < \infty$$

for some function $f: \mathbb{R}^{+2} \to \mathbb{R}^+$, where $B(x, \varepsilon)$ is the open ball of center x and radius ε. The strength of this condition is determined by the rate of growth of the function $f(\varepsilon, t)$ at 0. The point of this approach is that the global integral condition $\int_0^D \Phi^{-1}(N(\varepsilon))\, d\varepsilon$ is replaced by a family of local integral conditions, reflecting the local variations of the structure of T. Majorizing measures have brought a complete understanding of sample continuity of general Gaussian processes. We refer the reader to the introduction of [14] for more history. The present paper will try to demonstrate that the use of majorizing measures allows definite progress in the study of more general processes. It is based on very few and (originally at least) simple ideas. In order not to obscure them, we have chosen not to consider the more subtle increment conditions as in [5] and [15], but to consider only condition (*), which already has a significant degree of generality.

The central problem we study is as follows: For which metric spaces (T, d) and which Young functions Φ is it true that all processes satisfying (*) also satisfy $\operatorname{Sup}_{t,u}(X_t - X_u) < \infty$ a.s.? (Some attention will also be paid to sample continuity, but the results are less complete there.) As we will see, this implies the existence of a constant A, independent of the process, for which $E \operatorname{Sup}_{t,u}(X_t - X_u) \leq A$. The smallest such constant will be denoted by $S = S(T, d, \Phi)$.

In Section 2, we relate $S(T, d, \Phi)$ to certain (finitely additive) measures on $T \times T$. Since we are interested in understanding $S(T, d, \Phi)$ in terms of the geometry of T, but not of $T \times T$, the characterizations there are not satisfactory. They are however the foundations for much of the rest of the work. The main step here is due to P. Assouad. I am most indebted to G. Pisier, from whom I learned it, since this made this paper possible.

In Section 3, we investigate the consequences of the hypothesis $S = S(T, d, \Phi) < \infty$ on the structure of T. Denote by ϕ the derivative of Φ. The norm $\|\cdot\|_\Phi$ depends only, within equivalence, on the behavior of Φ at ∞; so there is no essential loss of generality to assume that ϕ is strictly increasing (which avoids some inessential technicalities). We set $\psi = \phi^{-1}$. We assume from now on (T, d) compact (we will see that this does not decrease the generality). The following result shows how majorizing measures come naturally into the problem.

THEOREM 1.3. *There exists a probability measure m on T such that*

$$\forall x \in T, \quad \int_0^D \psi\left(\frac{\varepsilon}{8Sm(B(x, \varepsilon))}\right) d\varepsilon \leq 4S. \tag{1.1}$$

We should mention at this point that no efforts are made to find sharp numerical constants.

Let us say that a metric space (T, d) is ultrametric if the following condition holds:

$$\forall s, t, u \in T, \qquad d(s, u) \le \max(d(s, t), d(t, u)).$$

In an ultrametric space, two balls of the same radius are either disjoint or identical; this makes the structure of these spaces rather simple. Ultrametric spaces play a natural role in the study of Gaussian processes [14], and will also be important here. When (T, d) is ultrametric, we can reinforce Theorem 1.3.

THEOREM 1.4. *If (T, d) is ultrametric compact, there exists a probability measure m on T such that*

$$\forall x \in T, \qquad \int_0^D \Phi^{-1}\left(\frac{1}{m(B(x, \varepsilon))}\right) d\varepsilon \le 8S. \tag{1.2}$$

One can show that (1.2) is always stronger than (1.1), but in general it is strictly stronger, e.g., if $\Phi(x) = |x|^p$, $p > 1$.

In Section 4, we investigate the converse problem. Given a probability measure m on T, is it true that

$$S(T, d, \Phi) \le K \operatorname*{Sup}_{x \in T} \int_0^D \Phi^{-1}\left(\frac{1}{m(B(x, \varepsilon))}\right) d\varepsilon?$$

We prove that this is the case provided Φ satisfies a growth condition that essentially means that it grows faster than $|x|^p$ for some $p > 1$; but the proof is unexpectedly hard. We show that this result is a natural generalization of the Pisier–Kôno Theorem 1.2. In the case where Φ increases fast enough at ∞ (essentially faster than $x^{\alpha \log \log x}$ for some $\alpha > 0$), we show that the convergence of the integral in (1.1) and (1.2) are equivalent, so our results provide a complete understanding of the condition $S(T, d, \Phi) < \infty$ in that case.

As already mentioned, the basic superiority of majorizing measures over covering numbers lies in their ability to take into account the lack of homogeneity of the index space. However, in order to fully justify the theory, it is wise to show that the underlying ideas also bring clarification in the more classical case where homogeneity is not an issue. A typical case is when T is a compact group (or a compact subset of a nonempty interior of a locally compact group), and d is translation invariant. In that case, all the points of T play the same role, and majorizing measure conditions are equivalent to integral conditions on the covering numbers. In Section 5, we investigate three specific examples. We first deal with the cube $[0,1]^k$, provided with the usual distance. We show that $S(T, d, \Phi) < \infty$ is equivalent to the condition of Theorem 1.3, so in that case Theorem 1.3 is optimal. Next, we investigate the case of $T = [0,1]$, provided with a distance as in Theorem 1.1, and we show how to compute $S(T, d, \Phi)$ for any Orlicz function Φ, thus generalizing Theorem 1.1. Finally, we investigate a genuinely complex case, where $\Phi(x) = |x|^p$, and where $T = U^q \times H$, where U^q is the q-dimensional torus, H is ultrametric and T is provided with the product distance. The appropriate integral condition there lies strictly between those of Theorems 1.3 and 1.4. There is no doubt that the exact conditions could be found in many more situations but the energy of the author was exhausted long before the power of the method.

Marcus and Pisier [11] and Weber [16] have investigated when stochastic processes satisfying condition $(*)$ have their sample paths in exponential-type Orlicz spaces pertaining to measures on T. Our methods allow us to settle this problem; this is the object of Section 6.

2. Preliminaries; conditions on $T \times T$. Our first task is to spell out what we will need concerning Young functions. It is not more difficult (and often much clearer) to work with general Young functions than with specific choices; but the reader who still refuses to use them can set $\Phi(t) = t^p/p$, $\Psi(t) = t^q/q$ $(p > 1, 1/p + 1/q = 1)$ and $\xi(t) = t^{-1/p}q^{-1/q}$ throughout the paper. Most of our results are not any easier in that special case.

Consider $\phi: \mathbb{R}^+ \to \mathbb{R}^+$ such that $\phi(0) = 0$, ϕ is strictly increasing continuous; denote $\psi = \phi^{-1}$. Set, for $x > 0$, $\Phi(x) = \int_0^x \phi(t)\,dt$, $\Psi(x) = \int_0^x \psi(t)\,dt$. Then Φ and Ψ are called conjugate Young functions. The following obvious inequalities are very useful:

$$\Phi(x) \leq x\phi(x), \qquad \Psi(x) \leq x\psi(x). \tag{2.1}$$

For $u, v \geq 0$, we have Young's inequality [17],

$$uv \leq \Phi(u) + \Psi(v). \tag{2.2}$$

It will be most useful when written as

$$u \leq b\Phi\left(\frac{u}{a}\right) + b\Psi\left(\frac{a}{b}\right), \tag{2.3}$$

for $u \geq 0$, $a, b > 0$. We extend Φ to $\mathbb{R}$ by $\Phi(x) = \Phi(|x|)$.

In Theorem 1.4, we have met the function $\Phi^{-1}(1/x)$. For technical reasons, it is not convenient to work with this function (in particular, it need not be convex) so we introduce for $x > 0$ the function $\xi(x) = 1/(x\Psi^{-1}(1/x))$. The next lemma relates ξ and $\Phi^{-1}(1/x)$.

LEMMA 2.1. *For $x > 0$, we have $\frac{1}{2}\Phi^{-1}(1/x) \leq \xi(x) \leq \Phi^{-1}(1/x)$.*

PROOF. Using (2.1), we have, for $t > 0$,

$$\Phi\left(\frac{\Psi(t)}{t}\right) \leq \frac{\Psi(t)}{t}\phi\left(\frac{\Psi(t)}{t}\right) \leq \frac{\Psi(t)}{t}\phi(\psi(t)) = \Psi(t).$$

Letting $t = \Psi^{-1}(1/x)$ yields $\Phi(\xi(x)) \leq 1/x$, so $\xi(x) \leq \Phi^{-1}(1/x)$. Using (2.2), we have

$$2\Psi(t) = \frac{2\Psi(t)}{t} \cdot t \leq \Phi\left(\frac{2\Psi(t)}{t}\right) + \Psi(t)$$

so $\Psi(t) \le \Phi(2\Psi(t)/t)$. Setting $t = \Psi^{-1}(1/x)$ yields $1/x \le \Phi(2\xi(x))$, i.e., $\frac{1}{2}\Phi^{-1}(1/x) \le \xi(x)$. □

LEMMA 2.2. (a) *$x\Phi(1/x)$ is convex decreasing.*
(b) *$x\Phi^{-1}(1/x)$ is concave increasing.*
(c) *$\xi(x)$ is convex decreasing.*

PROOF. (a) Let $f(x) = x\Phi(1/x)$, so $f'(x) = \Phi(1/x) - (1/x)\phi(1/x)$ is less than or equal to 0 by (2.1). To show that f' increases, it suffices to show that $h(t) = \Phi(t) - t\phi(t)$ decreases; but for $u \ge t$, we have, since $\Phi(u) \le \Phi(t) + (u - t)\phi(u)$,

$$h(u) \le \Phi(t) + (u - t)\phi(u) - u\phi(u) = \Phi(t) - t\phi(u) \le h(t).$$

(b) Let $f(x) = x\Phi^{-1}(1/x)$, so $f'(x) = \Phi^{-1}(1/x) - (1/x)[1/\phi(\Phi^{-1}(1/x))]$. Set $h(t) = t - [\Phi(t)/\phi(t)]$, so $f'(x) = h(\Phi^{-1}(1/x))$ and $h(t) \ge 0$ by (2.1). Since $x \to \Phi^{-1}(1/x)$ decreases, to show that f' decreases, it is enough to show that h increases; but for $u > t$, we have

$$h(t) \le t - \frac{\Phi(t)}{\phi(u)} = u - \frac{(u - t)\phi(u) + \Phi(t)}{\phi(u)} \le h(u).$$

(c) By (b), $f(x) = x\Psi^{-1}(1/x)$ is concave increasing, so $\xi(x) = 1/f(x)$ is decreasing and $\xi'(x) = -[f'(x)/f^2(x)]$ is increasing since $f' \ge 0$ decreases and $f(x)$ increases. □

We should mention that our definition of Φ (always finite) excludes the case where $\|\cdot\|_\Phi$ is the supremum norm $\|\cdot\|_\infty$, but this case is completely uninteresting since $S(T, d, \|\cdot\|_\infty) < \infty$ if and only if T has a finite diameter.

We denote by Δ the diagonal of $T \times T$; and by G the space of continuous bounded functions on $T \times T \setminus \Delta$, provided with the supremum norm. The main result of this section is as follows.

THEOREM 2.3. *Let (T, d) be a metric space (we do not assume T to be compact). Then the following are equivalent:*

(a) *For any process $(X_t)_{t\in T}$ that satisfies $(*)$, we have $P(\operatorname{Sup}_{t,u}(X_t - X_u) < \infty) > 0$.*

(b) *For each $\varepsilon > 0$, there is $A > 0$, such that for each process $(X_t)_{t\in T}$ that satisfies $(*)$, we have $P(\operatorname{Sup}_{t,u}(X_t - X_u) \ge A) \le \varepsilon$.*

(c) *There exists a constant S such that for each process $(X_t)_{t\in T}$ that satisfies $(*)$, we have $E(\operatorname{Sup}_{t,u}(X_t - X_u)) \le S$.*

(d) *There exists a constant M, a positive linear functional θ on G, with $\theta(1) = 1$, such that for any Lipschitz function f on T we have*

$$\theta\left(\Phi\left(\frac{f(t) - f(u)}{d(t,u)}\right)\right) \le 1 \Rightarrow \operatorname*{Sup}_{t,u\in T}(f(t) - f(u)) \le M. \tag{2.4}$$

Moreover, these conditions imply that T is totally bounded, and if S, M are chosen minimal, we have $M \le S \le 2M$.

We note that in (d), we assume f Lipschitz in order to ensure that the function $\Phi([f(t) - f(u)]/d(t,u))$ that is defined on $T \times T \setminus \Delta$ belongs to G; the Lipschitz constant of f, which can be arbitrarily large, is irrelevant.

The main contribution to Theorem 2.3 is due to Assouad [1] (after much work by others); the rest is essentially routine (see, e.g., [5], Theorem 5.2).

PROOF. (a) $\Rightarrow$ (b). Assume that (b) fails. Then for some $\varepsilon > 0$, and for each $n > 0$, there is a process $(X_{t,n})_{t \in T}$ that satisfies $(*)$ and for which

$$P\Big(\operatorname*{Sup}_{t,u}(X_{t,n} - X_{u,n}) \ge 2^{2n+1}\Big) \ge \varepsilon.$$

Let us fix v in T. We can replace the processes $(X_{t,n})$ by $(X_{t,n} - X_{v,n})_{t \in T}$, so there is no loss of generality to assume $X_{v,n} = 0$. The two processes $X_{t,n}^+$ and $X_{t,n}^-$ still satisfy $(*)$, while

$$X_{t,n} - X_{u,n} \le X_{t,n}^+ + X_{u,n}^-.$$

So for each n, by choosing either $X_{t,n}^+$ or $X_{t,n}^-$ we can find a process $(Y_{t,n})$, with $Y_{t,n} \ge 0$ and $Y_{v,n} = 0$ that satisfies $(*)$ and for which

$$P\Big(\operatorname*{Sup}_{t} Y_{t,n} \ge 2^{2n}\Big) \ge \varepsilon/2. \tag{2.5}$$

We can assume that the processes $(Y_{t,n})_{t \in T}$ are independent. For each t, set $Y_t = \sum_{n \ge 1} 2^{-n} Y_{t,n}$. [Since $\|Y_{t,n}\|_\Phi \le d(t,v)$, the series converges in L_Φ.] The convexity of Φ shows that Y_t satisfies $(*)$, since

$$E\left(\Phi\left(\frac{\sum_{n\ge1} 2^{-n}Y_{t,n} - \sum_{n\ge1} 2^{-n}Y_{u,n}}{d(t,u)}\right)\right) \le E\left(\sum_{n\ge1} 2^{-n}\Phi\left(\frac{Y_{t,n} - Y_{u,n}}{d(t,u)}\right)\right) \le 1.$$

By independence and (2.5), we have

$$P\Big(\operatorname*{Sup}_{n}\operatorname*{Sup}_{t} 2^{-n}Y_{t,n} \ge 2^n\Big) = 1,$$

so $P(\operatorname{Sup}_t Y_t = \infty) = 1$ since $Y_{t,n} \ge 0$. Since $Y_v = 0$, we have shown that (a) fails.

(b) $\Rightarrow$ (d). Let M be large enough that for any process that satisfies $(*)$, we have

$$P\Big(\operatorname*{Sup}_{t,u} X_t - X_u \ge M\Big) < 1. \tag{2.6}$$

Consider the subset $\mathscr{C}$ of G that consists of functions of the type

$$\sum_{i \in I} \alpha_i \Phi([f_i(t) - f_i(u)]/d(t,u)),$$

where I is finite, $\alpha_i \ge 0$, $\sum_{i \in I}\alpha_i = 1$, f_i is a Lipschitz function on T and $\operatorname{Sup}_{t,u \in T}(f_i(t) - f_i(u)) > M$. We show that any function of $\mathscr{C}$ takes values greater than 1. Indeed, if we provide the set I with the probability P given by

$P(\{i\}) = \alpha_i$, we can define a process $(X_t)_{t\in T}$ on this probability space by $X_t(i) = f_i(t)$. We have $P(\mathrm{Sup}_{t,u}(X_t - X_u) > M) = 1$, so by the choice of M, the process must fail condition $(*)$; so, for some t, u in T, we have

$$\sum_{i\in I} \alpha_i \Phi\left(\frac{f_i(t) - f_i(u)}{d(t,u)}\right) = E\Phi\left(\frac{X_t - X_u}{d(t,u)}\right) > 1,$$

which proves the claim.

By definition, $\mathscr{C}$ is convex. Denote by $\mathscr{C}'$ the set of functions g in G such that $g < 1$. Then $\mathscr{C}'$ is open convex, and we have shown that $\mathscr{C} \cap \mathscr{C}' = \varnothing$. It follows from the geometric form of the Hahn–Banach theorem that there is a linear functional θ on G and $\alpha \in \mathbb{R}$ such that $\theta(g) < \alpha$ for $g \in \mathscr{C}'$ and $\theta(g) \geq \alpha$ for $g \in \mathscr{C}$. Since $\mathscr{C}'$ contains 0, we have $\alpha > 0$. Since $\theta(\lambda g) < \alpha$ whenever $g \leq 0$, $\lambda \geq 0$ we have $\theta(g) \leq 0$, so θ is positive. Also $\theta(1) \leq \alpha$; replacing θ by $\theta/\theta(1)$, we can assume that $\theta(1) = 1$, and that we have $\theta(g) \geq 1$ whenever $g \in \mathscr{C}$. If f is a Lipschitz function on T, such that $\mathrm{Sup}_{t,u}(f(t) - f(u)) > M$, by definition $\mathscr{C}$ contains $\Phi([f(t) - f(u)]/d(t,u))$, so $\theta(\Phi([f(t) - f(u)]/d(t,u))) \geq 1$. It follows that

$$\theta\left(\Phi\left(\frac{f(t) - f(u)}{d(t,u)}\right)\right) < 1 \Rightarrow \mathop{\mathrm{Sup}}_{t,u}(f(t) - f(u)) \leq M.$$

If we replace f by λf for $\lambda < 1$ and let λ go to 1, we get (2.4).

(d) $\Rightarrow$ (c). We are going to show that whenever a process satisfies $(*)$, we have

$$E \mathop{\mathrm{Sup}}_{t,u}(X_t - X_u) \leq 2M.$$

We first suppose that the basic probability space (Ω, Σ, P) is such that Σ is finite. For simplicity of the notation, we identify points in each atom of Σ, so we can assume that Ω is finite, and that $P(\{\omega\}) > 0$ for each ω in Ω. Consider a process $(X_t)_{t\in T}$ defined on Ω that satisfies $(*)$. For ω in Ω, t, u in T, $t \neq u$, we set

$$g_\omega(t,u) = \Phi\left(\frac{X_t(\omega) - X_u(\omega)}{d(t,u)}\right).$$

Condition $(*)$ means that for each $t \neq u$, we have

$$\sum_{\omega\in\Omega} P(\{\omega\}) g_\omega(t,u) \leq 1. \tag{2.7}$$

In particular, for each ω, we have

$$|X_t(\omega) - X_u(\omega)| \leq d(t,u)\Phi^{-1}(1/P(\{\omega\})),$$

so $t \to X_t(\omega)$ is Lipschitz and hence g_ω belongs to G. Since $\theta \geq 0$, $\theta(1) = 1$, we get from (2.7) that

$$\sum_{\omega\in\Omega} P(\{\omega\})\theta(g_\omega) = \theta\left(\sum_{\Omega} P(\{\omega\}) g_\omega\right) \leq 1. \tag{2.8}$$

For ω in Ω, set $a_\omega = \max(1, \theta(g_\omega))$. The convexity of Φ shows that for any

$t \neq u$, we have

$$\Phi\left(\frac{X_t(\omega) - X_u(\omega)}{a_\omega d(t,u)}\right) \leq \frac{g_\omega(t,u)}{a_\omega},$$

so

$$\theta\left(\Phi\left(\frac{X_t(\omega) - X_u(\omega)}{a_\omega d(t,u)}\right)\right) \leq 1$$

and (2.4) implies that

$$\operatorname*{Sup}_{t,u}\left(X_t(\omega) - X_u(\omega)\right) \leq Ma_\omega.$$

It follows from (2.8) that

$$E\operatorname*{Sup}_{t,u}\left(X_t(\omega) - X_u(\omega)\right) \leq M\sum_{\omega\in\Omega} P(\{\omega\})a_\omega \leq 2M.$$

We now turn to the general case. It is enough to show that for any countable subset D of T, we have

$$E\operatorname*{Sup}_{t,u\in D}\left(X_t - X_u\right) \leq 2M. \tag{2.9}$$

Since D is countable, we can assume the σ-field Σ to be countably generated, so there exists an increasing sequence Σ_n of finite σ-fields whose union generates Σ. Denote by X_t^n the conditional expectation $E(X_t|\Sigma_n)$. Since Φ is convex, and since $E\Phi((X_t - X_u)/d(t,u)) \leq 1$ for each t, u, Jensen's inequality shows that $E\Phi((X_t^n - X_u^n)/d(t,u)) \leq 1$, so the processes X_t^n satisfy condition $(*)$. Since Σ_n is finite, we have seen that $E\operatorname{Sup}_{t,u\in T}(X_t^n - X_u^n) \leq 2M$. So (2.9) follows from the fact that $X_t^n \to X_t$ P-a.s.

Since (c) $\Rightarrow$ (b) trivially, we have proved the equivalence of (a)–(d). To prove that one can take $M \leq S$, it is enough to note that (2.6) holds whenever $M > S$, and to use a compactness argument. It remains to prove that T is totally bounded, and we actually prove the fact, due to Pisier [12], that the covering numbers $N(\varepsilon)$ of T satisfy $N(2\varepsilon) \leq \Phi(2S/\varepsilon)$, where S is as in (d). Let $t_1, \ldots, t_n$ be points of T such that $d(t_i, t_j) \geq \varepsilon$ for $i < j$. Let $S' > S$. For $i \leq n$, define

$$f_i(t) = S'\max(0, 1 - 2d(t, t_i)/\varepsilon),$$

so f_i is Lipschitz of Lipschitz constant less than or equal to $2S'/\varepsilon$. Take $\Omega = \{1, \ldots, n\}$, with the uniform probability. Define $X_t(i) = f_i(t)$. Since for $t \neq u$, there are at most two indexes $i \leq n$ such that $f_i(t) \neq 0$ or $f_i(u) \neq 0$, we have

$$E\Phi\left(\frac{X_t - X_u}{d(t,u)}\right) \leq \frac{2}{n}\Phi\left(\frac{2S'}{\varepsilon}\right).$$

On the other hand, $E\operatorname{Sup}(X_t - X_u) = S' > S$, so condition $(*)$ must fail, and this implies $n \leq 2\Phi(2S'/\varepsilon)$; this finishes the proof. □

A process $(X_t)_{t \in T}$ that satisfies $(*)$ is a 1-Lipschitz map from T to L_Φ. This map has a unique extension to the completion $\hat{T}$ of T. When $S(T, d, \Phi) < \infty$, Theorem 2.4 shows that $\hat{T}$ is compact, and it is clear that $S(\hat{T}, d, \Phi) = S(T, d, \Phi)$, so there is actually no loss of generality to assume that (T, d) is compact. We make this hypothesis throughout the rest of the paper; it often allows simpler statements.

For application to questions of sample continuity, we note the following statement. Its proof is similar to the proof of the equivalence of (c) and (d) in Theorem 2.3, and therefore is left to the reader.

THEOREM 2.4. *Let $\delta > 0$. The following are equivalent.*

(a) *There exists a constant S_δ such that for each process satisfying $(*)$, we have*

$$E \sup_{d(t,u) \leq \delta} (X_t - X_u) \leq S_\delta.$$

(b) *There exists a constant M_δ and a positive linear functional θ on G such that for each Lipschitz function f on T, we have*

$$\theta\left(\Phi\left(\frac{f(t) - f(u)}{d(t,u)}\right)\right) \leq 1 \Rightarrow \sup_{d(t,u) \leq \delta} (f(t) - f(u)) \leq M_\delta.$$

Moreover, the best possible choices of S_δ and M_δ are related by $M_\delta \leq S_\delta \leq 2M_\delta$.

PROPOSITION 2.5. *Assume that there is a probability ν on $T \times T$ such that for any Lipschitz function f on T, we have*

$$\iint_{T \times T \setminus \Delta} \Phi\left(\frac{f(t) - f(u)}{d(t,u)}\right) d\nu(t,u) \leq 1 \Rightarrow \sup_{t,u} (f(t) - f(u)) \leq M.$$

Then $S(T, d, \Phi) \leq 2M$.

PROOF. Consider the function θ on G given by

$$\theta(g) = \iint_{T \times T \setminus \Delta} g \, d\nu \tag{2.10}$$

and use Theorem 2.4. □

It is a natural question whether the linear functional θ of Theorem 2.3 can be induced by a measure on $T \times T$ as in (2.10). The question does not seem to be of great importance in the present perspective, although it might be related to more important open problems that we will mention later. It is a mere exercise to see that a linear functional θ on G is given by (2.10) for some probability ν on $T \times T$ if and only if there exists a function h from $\mathbb{R}^+$ to $\mathbb{R}^+$, with $\lim_{t \to 0} h(t) = \infty$, such that $\theta(g) \leq 1$ whenever $g(t,u) \leq h(d(t,u))$. Once this is noted, one can prove as in Theorem 2.3 the following.

PROPOSITION 2.6. *The following are equivalent.*

(a) *There is a probability measure ν on $T \times T$, and $K \geq 0$ such that for any Lipschitz function f on T, we have*

$$\iint_{T\times T\setminus\Delta} \Phi\left(\frac{f(t)-f(u)}{d(t,u)}\right) d\nu(t,u) \leq 1 \Rightarrow \operatorname*{Sup}_{t,u}\left(f(t)-f(u)\right) \leq K.$$

(b) *There is a function h from $\mathbb{R}^+$ to $\mathbb{R}^+$ with $\lim_{t\to 0} h(t) = \infty$, and $L \geq 0$, such that for any process $(X_t)_{t\in T}$ that satisfies*

$$\forall\, t,u \in T, \qquad E\Phi\left(\frac{X_t - X_u}{d(t,u)}\right) \leq h(d(t,u)),$$

we have $E\,\mathrm{Sup}(X_t - X_u) \leq L$.

(It is also possible, as in Theorem 2.4, to give other equivalent conditions.) When $\lim_{x\to\infty}[\Phi(x)/x] = \infty$, we do not have any example where $S(T,d,\Phi) < \infty$, but where the conditions of Proposition 2.6 fail. When $\Phi(x) = |x|$, a simple but still instructive example is given by $T = [0,1]$, with the usual distance. Then if $(X_t)_{t\in[0,1]}$ satisfies $(*)$, for each sequence $0 \leq u_0 < u_1 < \cdots, u_n \leq 1$ we have

$$E\left(\sum_{1\leq i\leq n} |X_{u_i} - X_{u_{i-1}}|\right) \leq \sum_{1\leq i\leq n} |u_i - u_{i-1}| \leq 1,$$

so if (X_t) is separable, the expected total variation of the trajectories is less than or equal to 1, and, in particular, $S([0,1],d,|\cdot|) \leq 1$.

For v in $[0,1]$, $\varepsilon > 0$, consider the function f_v^ε given by $f_v^\varepsilon(t) = \max(0, 1 - |t-v|/\varepsilon)$. We have $|f_v^\varepsilon(t) - f_v^\varepsilon(u)| \leq |t-u|/\varepsilon$ for any $t,u,v \leq 1$ if $|t-v| \leq \varepsilon$ or $|u-v| \leq \varepsilon$ and 0 otherwise. It is a simple exercise to show that for any $A > 0$ and any probability ν on $T\times T$, one can find $\varepsilon > 0$, v in $[0,1]$ with

$$\iint_{T\times T\setminus D} A\left|\frac{f_v^\varepsilon(t) - f_v^\varepsilon(u)}{d(t,u)}\right| d\nu(t,u) \leq 1,$$

while $\mathrm{Sup}_{t,u}(Af_v^\varepsilon(t) - Af_v^\varepsilon(u)) = A$; so the conditions of Proposition 2.6 do not hold.

An interesting feature of Theorem 2.3 is the automatic integrability of $\mathrm{Sup}_{t,u}(X_t - X_u)$. It is a natural question to ask whether one can say more. The following proposition covers a number of cases.

PROPOSITION 2.7. *Assume that there is Orlicz function Γ and $0 < \alpha$ such that*

$$a \geq \Phi^{-1}(\tfrac{1}{2}), \qquad b \geq 1 \Rightarrow \Phi(ab) \geq \alpha\Phi(a)\Gamma(b). \tag{2.11}$$

Then the conditions of Theorem 2.3 imply that for any process that satisfies $(*)$, *we have*

$$\left\| \operatorname{Sup}_{t,u} (X_t - X_u) \right\|_\Gamma \le KM/\alpha,$$

where K depends on Γ *only, and M is as in Theorem* 2.3 *and* 2.4.

Proof. As in the proof of Theorem 2.3, (a) $\Rightarrow$ (b), it is enough to assume Ω finite. For each ω, we define $h(\omega)$ as the infimum of the $\lambda > 0$ for which $\theta(\Phi([X_t(\omega) - X_u(\omega)]/\lambda d(t,u))) \le 1$. So, if $h(\omega) \ne 0$, we have

$$\theta\left(\Phi\left(\frac{X_t(\omega) - X_u(\omega)}{h(\omega)d(t,u)}\right)\right) = 1. \tag{2.12}$$

Define

$$A_\omega = \left\{(t,u); |X_t(\omega) - X_u(\omega)| \ge \Phi^{-1}\left(\tfrac{1}{2}\right) d(t,u)\right\},$$
$$B_\omega = T \times T \setminus (\Delta \cup A_\omega).$$

If $h(\omega) \ge 1$, we have, from (2.11),

$$\Phi\left(\frac{X_t(\omega) - X_u(\omega)}{d(t,u)}\right) = \Phi\left(\frac{|X_t(\omega) - X_u(\omega)|}{h(\omega)d(t,u)} h(\omega)\right)$$
$$\ge \alpha\Phi\left(\frac{X_t(\omega) - X_u(\omega)}{h(\omega)d(t,u)}\right)\Gamma(h(\omega))1_{A_\omega}(t,u).$$

Since $\Phi([X_t(\omega) - X_u(\omega)]/h(\omega)d(t,u))1_{B_\omega}(t,u) \le \frac{1}{2}$, from (2.12) we have $\theta(\Phi([X_t(\omega) - X_u(\omega)]/d(t,u))1_{A_\omega}(t,u)) \ge \frac{1}{2}$. It follows that

$$\Gamma(h(\omega)) \le \frac{2}{\alpha}\theta\left(\Phi\left(\frac{X_t(\omega) - X_u(\omega)}{d(t,u)}\right)\right).$$

If $h(\omega) \le 1$, we have $\Gamma(h(\omega)) \le \Gamma(1)$; so we have

$$E\Gamma(h(\omega)) \le \Gamma(1) + \frac{2}{\alpha}\theta\left(\sum_{\omega\in\Omega} P(\{\omega\})\Phi\left(\frac{X_t(\omega) - X_u(\omega)}{d(t,u)}\right)\right).$$

So

$$E\Gamma(h(\omega)) \le \Gamma(1) + \frac{2}{\alpha} \le \frac{2}{\alpha}[1 + \Gamma(1)]$$

and, in particular, $\|h(\omega)\|_\Gamma \le (2/\alpha)[1 + \Gamma(1)]$. Since by condition (d) of Theorem 2.3 we have $\operatorname{Sup}_{t,u}(X_t(\omega) - X_u(\omega)) \le Mh(\omega)$, the proof is complete. □

Proposition 2.7 applies in particular to the case where $\Phi(x) = x^p$, $p > 1$; we have taken care in the statement of (2.11) to eliminate the values of a, b near 0 since it is the behavior of Φ, Γ at ∞ that really matters. The following example shows that in general one does not have $\|\operatorname{Sup}(X_t - X_u)\|_\Phi < \infty$.

Proposition 2.8. *Let* $\Phi(x) = x\log(x+1)$. *Let* $\Gamma: \mathbb{R}^+ \to \mathbb{R}^+$ *be such that* $\lim_{x\to\infty} \Gamma(x)/x = \infty$. *Then there exists a space* (T, d) *that satisfies the conditions of Theorem* 2.3, *and a process* $(X_t)_{t\in T}$ *that satisfies* $(*)$, *but for which* $E\Gamma(\operatorname{Sup}_{t,u}(X_t - X_u)) = \infty$.

Proof. The hypothesis on Γ shows that there is a sequence $(n_i)_{i\geq 1}$ such that $\Sigma 2^{-i} n_i/i < \infty$ but $\Sigma\Gamma(2^i)2^{-2i}n_i/i = \infty$. We can obviously assume $n_i \geq 2^{i/2}$. We set $T = \bigcup_{1\leq i\leq\infty} T_i$, where $T_\infty = \{\infty\}$, and $\operatorname{card} T_i = n_i$ if $i < \infty$. For t in T_i, u in T_j, $u \neq t$, we set $d(t,u) = \max(2^{-i}, 2^{-j})$. It follows, e.g., from Theorem 1.2 (or the generalizations we will give later on) that $S(T, d, \Phi) < \infty$. For $i \geq 1$, consider disjoint sets $(A_{t,i})_{t\in T_i}$ in Ω, with $P(A_{t,i}) = 2^{-2i-1}/(3i\log 2)$. For $i \neq j$, we can assume independence of the sequences $(A_{t,i})_{t\in T_i}$ and $(A_{t,j})_{t\in T_j}$. For $t \in T_i$, define $X_t = 2^i 1_{A_{t,i}}$, so easily $\|X_t\|_\Phi \leq 2^{-i-1}$. Define $X_\infty = 0$. The definition of d shows that (X_t) satisfies $(*)$. But $\operatorname{Sup}_{t\in T_i} X_t = 2^i$ on $A_i = \bigcup_{t\in T_i} A_{t,i}$, and $P(A_i) = 2^{-2i-1}n_i/3i\log 2$. Since the sets A_i are independent, and since $\Sigma\Gamma(2^i)P(A_i) = \infty$, we have

$$E\Gamma\Big(\operatorname*{Sup}_{t,u}(X_t - X_u)\Big) \geq E\Gamma\Big(\operatorname*{Sup}_{t} X_t\Big) = \infty. \qquad \square$$

We say that Φ satisfies the Δ_2 condition if for some constant C, and all $x \geq 1$, we have $\Phi(2x) \leq C\Phi(x)$. The following relates the conditions of sample boundedness and sample continuity. We send the reader, e.g., to [5] for the notion of the separable process.

Theorem 2.9. *Assume that* $\lim_{x\to\infty}\Phi(x)/x = \infty$. *Then the following are equivalent.*

(a) Φ *satisfies the* Δ_2 *condition.*

(b) *For each compact space* (T, d) *that satisfies* $S(T, d, \Phi) < \infty$, *each separable process* $(X_t)_{t\in T}$ *that satisfies* $(*)$ *is sample continuous.*

Proof. (b) $\Rightarrow$ (a). Assume that Φ fails the Δ_2 condition. Then there is an increasing sequence $(a_i)_{i\geq 2}$ with $\Phi(2a_i) \geq 2^{i+1}(\Phi(a_i) + 1)$. Denote by n_i the integer part of $2^{-i}\Phi(2a_i)$, so $n_i \geq 2\Phi(a_i)$. We set $T = \bigcup_{2\leq i\leq\infty} T_i$, where $T_\infty = \{\infty\}$ and $\operatorname{card} T_i = n_i$ if $i < \infty$. For t in T_i, $u \in T$, $u \neq t$, we set $d(t,u) = \max(a_i^{-1}, a_j^{-1})$ (where $a_\infty^{-1} = 0$). We note that (T, d) is ultrametric. Define the probability m on T by $m(\{\infty\}) = \frac{1}{2}$, $m(\{t\}) = 2^{-i}/n_i$ if $t \in T_i$. We note that $a_i^{-1}\Phi^{-1}(2^i n_i) \leq 2$, so if we apply Theorem 4.2 (to be proved later), we see that $S(T, d, \Phi) < \infty$. For each $i \geq 2$, denote by $(A_{i,t})_{t\in T_i}$ a partition of Ω in sets of probability $1/n_i$. Define $X_\infty = 0$, and $X_t = 1_{A_{i,t}}$ if $t \in T_i$. It is straightforward to see that this process satisfies $(*)$, since $2\Phi(a_i) \leq n_i$. But for each ω, $t \to X_t(\omega)$ is not continuous at ∞, because it takes the value 1 in each set T_i. This completes the proof.

(a) $\Rightarrow$ (b). Pisier [12] noted that (b) holds for $\Phi = |\cdot|^p$, $p > 1$. His proof relies on the fact that p-summing operators are p-radonifying. Our argument is of the same nature (and could be used to prove an abstract theorem), but an important difference is that L_Φ need not be reflexive. There seems no way in the proof to avoid the use of more sophisticated tools than in the rest of the paper, but we have tried to give the most elementary proof possible. We will need the following facts, true for any conjugate Young functions Φ, Ψ [we still assume that $\lim_{x\to\infty} \Phi(x)/x = \infty$]. □

LEMMA 2.10. (a) *For any probability space* (Ω, Σ, μ), $L_\Phi = L_\Phi(\mu)$ *is isomorphic to a dual space* (*actually to the dual of the closure of* L^∞ *in* L_Ψ).

(b) *If* Φ *satisfies the* Δ_2 *condition*, L^∞ *is dense in* L_Φ.

LEMMA 2.11. *If* Φ *satisfies the* Δ_2 *condition*, $L_\Phi(\mu)$ *has the Radon–Nikodym property* (R.N.P.) [2].

PROOF. Denote by A the unit ball of L_∞, so A is $\sigma(L_\infty, L_1)$ compact, and $\sigma(L_\Phi, L_\Psi)$ compact. By Lemma 2.10, the span of A is dense in L_Φ, so L_Φ is weakly compactly generated. Since it is isomorphic to a dual space, the result follows from [2].

Since we assume $S(T, d, \Phi) < \infty$, consider the functional θ on the space G of continuous bounded functions on $T \times T \setminus \Delta$ given by Theorem 2.3. For g in G, define $\|g\|_\Phi = \operatorname{Inf}\{\lambda > 0; \theta(\Phi(g/\lambda)) \le 1\}$. Denote by $L_\Phi(\theta)$ the completion of G with respect to $\|\cdot\|_\Phi$. The basic fact is that $L_\Phi(\theta)$ has the R.N.P. To see it, one can introduce the Stone–Cech compactification K of $T \times T \setminus \Delta$, so $C(K)$ can be identified with G and θ can be identified with a Radon probability θ' on K; and $L_\Phi(\theta)$ is isometric to $L_\Phi(\theta')$, so the result follows from Lemma 2.11.

We fix a process $(X_t)_{t\in T}$ that satisfies $(*)$, with basic probability space (Ω, Σ, P). We denote by $\operatorname{Lip}(T)$ the space of Lipschitz functions on T, provided with the seminorm $\|f\|_L = \operatorname{Sup}_{t\ne u}|[f(t) - f(u)]/d(t,u)|$. We can define a vector measure m_1 valued in $\operatorname{Lip}(T)$ by $m_1(A) = \int_A X_t dP$ for A in Σ. [We have $m_1(A) \in \operatorname{Lip}$ by condition $(*)$.] We consider the operator U from $\operatorname{Lip}(T)$ to G given by $U(f)(t,u) = [f(t) - f(u)]/d(t,u)$, so $\|U\| \le 1$. We denote by j the canonical injection from G to $L_\Phi(\theta)$. We denote by Z the closure of the range of $j \circ U$ in $L_\Phi(\theta)$. Consider the vector measure m given by $m(A) = j\circ U(m_1(A))$ for A in Σ. It is valued in Z. Consider now a finite partition $(A_i)_{i\le n}$ of Ω. For $t \ne u$, we have $E(\Phi((X_t - X_u)/d(t,u))) \le 1$, so by convexity of Φ, we have

$$\sum_{i\le n} P(A_i)\Phi\big(E\big(1_{A_i}(X_t - X_u)/d(t,u)P(A_i)\big)\big) \le 1,$$

i.e., $\sum_{i\le n} P(A_i)\Phi(U(m_1(A_i))(t,u)/P(A_i)) \le 1$. As in the proof of Theorem 2.3, (d) $\Rightarrow$ (c), we see that this implies $\sum_{i\le n} P(A_i)\|m(A_i)/P(A_i)\|_\Phi \le 2$, so $\sum_{i\le n}\|m(A_i)\| \le 2$. We have shown that m has bounded variation; the definition of the Radon–Nikodym property implies that there is a measurable map h_1 from Ω to Z such that $\int \|h_1\|_\Phi \, dP < \infty$ and $\int_A h_1 \, dP = m(A)$ for A in Σ.

We fix a point s in T, and we consider the operator V from G to $C(T)$ given by $V(f)(t) = d(s,t)f(t,s)$ for $t \ne s$ and $V(f)(s) = 0$. The basic property of θ is that for f in $\operatorname{Lip}(T)$, we have $\|V(U(f))\|_\infty \le 1$ whenever $\theta(\Phi(U(f))) \le 1$, so

$\|V \circ U(f)\|_\infty \le \|U(f)\|_\Phi$. It follows that V can be extended by a norm one operator $\overline{V}$ from Z to $C(T)$. We set $h = \overline{V} \circ h_1$, so for each ω, $h(\omega) \in C(T)$. Also for each A in Σ, $V \circ m(A) = \int_A h\,dP$, so for each t in T, $V \circ m(A)(t) = \int_A h(\omega)(t)\,dP(\omega)$. By definition of m, we have $V \circ m(A)(t) = \int_A (X_t - X_s)\,dP$. It follows that $X_t(\omega) - X_s(\omega) = h(\omega)(t)$ ω a.s. The exceptional set might depend on t; but it does not if the process (X_t) is separable, so $t \to X_t(\omega)$ is continuous for almost each ω. This concludes the proof. □

3. Necessary conditions. We first prove Theorem 1.3. So we assume $S = S(T, d, \Phi) < \infty$, and we want to produce a measure m on T that satisfies (1.1). Condition (d) of Theorem 2.3 gives us a functional θ on G that we can consider as a finitely additive measure on $T \times T \setminus \Delta$. The obvious approach is to try to construct m from θ. The most obvious choice for m is the average of the "marginals" of θ. It is almost surprising that such a straightforward approach works perfectly.

Given a continuous function h on T, define $h_1, h_2 \in G$ by the formula $h_1(t,u) = h(t)$, $h_2(t,u) = h(u)$. There are unique probability m_1, m_2 on T such that $\theta(h_1) = \int h(t)\,dm_1$, $\theta(h_2) = \int h(t)\,dm_2$, because $h \to \theta(h_1)$ [resp. $h \to \theta(h_2)$] defines a positive linear functional on $C(T)$ (we remind the reader that T is now assumed to be compact). We set $m = (m_1 + m_2)/2$. The following lemma spells out the relationship between θ and m that we will use.

LEMMA 3.1. *Let $(A_i)_{1 \le i \le n}$ be a decreasing sequence of closed subsets of T. Let $(d_i)_{1 \le i \le n+1}$ be positive numbers. Consider a function $g \in G$, $g \ge 0$, such that $g \le d_{n+1}$ and $g(t,u) \le d_i$ whenever $t, u \notin A_i$. Then*

$$\theta(g) \le 2\Big(d_1 + \sum_{i \le n} d_{i+1} m(A_i)\Big).$$

PROOF. For $1 \le i \le n$, consider a continuous function $f_i \ge 0$ on T with $f_i \ge 1_{A_i}$. Define $g_i(t,u) = f_i(t) + f_i(u)$, so

$$\theta(g_i(t,u)) = \int f_i\,dm_1 + \int f_i\,dm_2 = 2\int f_i\,dm.$$

We have $g \le d_1 + \sum_{1 \le i \le n} d_{i+1} g_i$, so $\theta(g) \le 2(d_1 + \sum_{1 \le i \le n} d_{i+1} \int f_i\,dm)$. Taking the infimum over the choices of f_i yields the result. □

We now prove (1.1). We fix x in T; we denote by s the smallest integer for which $2^{-s} < D = D(T)$. For $n \ge s$, we set $a_n = 2^{-n}/m(B(x, 2^{-n}))$. By induction over l, we define $n(1) = s$, and

$$n(l) = \inf\{n > n(l-1);\ a_n > a_{n(l-1)}\}.$$

Now fix $N > 1$. We consider an increasing sequence b_l, for $1 \le l \le N$. Let $h(v)$ be the unique piecewise affine function from $\mathbb{R}^+$ to $\mathbb{R}^+$ that is 0 for $v \ge 2^{-s}$, that is of slope $-b_l$ between $2^{-n(l+1)}$ and $2^{-n(l)}$ and that is constant for $v \le 2^{-n(N+1)}$.

Consider the function f on T given by $f(t) = h(d(x,t))$. We note that

$$\text{(3.1)} \qquad \operatorname*{Inf}_t f(t) = 0, \qquad f(x) \geq \sum_{1 \leq l \leq N} 2^{-n(l)-1} b_l.$$

Since the sequence (b_l) increases, if for some l, $2 \leq l \leq N+1$, we have $d(x,t)$, $d(x,u) \geq 2^{-n(l)}$ then we have

$$|f(t) - f(u)| \leq |d(x,t) - d(x,u)| b_{l-1} \leq d(t,u) b_{l-1}.$$

It follows from Lemma 3.1 that

$$\theta\left(\Phi\left(\frac{f(t)-f(u)}{d(t,u)}\right)\right) \leq 2\sum_{l \leq N} m\big(B(x, 2^{-n(l)})\big)\Phi(b_l).$$

Using (3.1) and condition (d) of Theorem 2.3, we have

$$\text{(3.2)} \qquad 2\sum_{l \leq N} m\big(B(x,2^{-n(l)})\big)\Phi(b_l) \leq 1 \Rightarrow \sum_{l \leq N} 2^{-n(l)-1} b_l \leq S.$$

Let $\lambda > 0$, and make the choice $b_l = \psi(\lambda a_{n(l)})$. The definition of the sequence $(n(l))$ shows that this sequence increases. We note that by (1.1), for $u > 0$ we have

$$\Phi(\psi(u)) \leq \psi(u)\phi(\psi(u)) = u\psi(u),$$

so from (3.2) it follows that

$$2\sum_{l \leq N} \lambda 2^{-n(l)} \psi\big(\lambda a_{n(l)}\big) \leq 1 \Rightarrow \sum_{l \leq N} 2^{-n(l)} \psi\big(\lambda a_{n(l)}\big) \leq 2S.$$

If we choose λ such that equality occurs on the left, we see that $\lambda \geq 1/4S$, and that $\sum_{l \leq N} 2^{-n(l)}\psi(\lambda a_{n(l)}) \leq 2S$. Since ψ is increasing, we have

$$\sum_{l \leq N} 2^{-n(l)} \psi\big(a_{n(l)}/4S\big) \leq 2S;$$

since N is arbitrary, we have $\sum_{l \geq 1} 2^{-n(l)}\psi(a_{n(l)}/4S) \leq 2S$. Since $a_n \leq a_{n(l)}$ for $n(l) \leq n < n(l+1)$, we have

$$\sum_{n(l+1) > n \geq n(l)} 2^{-n}\psi\left(\frac{a_n}{4S}\right) \leq 2^{-n(l)+1}\psi\left(\frac{a_{n(l)}}{4S}\right),$$

so we have

$$\sum_{n \geq s} 2^{-n}\psi\left(\frac{a_n}{4S}\right) \leq 4S$$

and (1.1) follows from the inequality

$$\int_{2^{-n}}^{2^{-n+1}} \psi\left(\frac{\varepsilon}{8S\mu(B(x,\varepsilon))}\right) d\varepsilon \leq 2^{-n}\psi\left(\frac{2^{-n+1}}{8S\mu(B(x,2^{-n}))}\right).$$

Remark. From (2.1), we see that (1.1) implies

$$\forall x \in T, \qquad \int_0^D \frac{m(B(x,\varepsilon))}{\varepsilon}\Psi\left(\frac{\varepsilon}{8Sm(B(x,\varepsilon))}\right) d\varepsilon \leq \frac{1}{2}.$$

This is often more convenient that (1.1) because the integrand is a convex function of m, as is shown by Lemma 2.2. On the other hand, the easily proved inequality $x\psi(x) \le \Psi(2x)$ shows that it has essentially the same strength.

We now present a way to construct majorizing measures, which is based on an idea of Fernique [4]. We denote throughout the paper the set of probability measures on T by $P(T)$. For $\mu \in P(T)$, $a \ge 0$, we set

$$\xi_{\mu,a}(x) = \int_a^D \xi(\mu(B(x,\varepsilon)))\,d\varepsilon,$$

where $D = D(T)$ as usual. It is a simple exercise to show that $\xi_{\mu,a}(x)$ is a continuous function of x if $a > 0$.

PROPOSITION 3.2. *Assume that for some constant $A > 0$ and for each $a > 0$, we have $\int_T \xi_{\mu,a}(x)\,d\mu(x) \le A$ for all $\mu \in P(T)$. Then there exists $m \in P(T)$ such that $\xi_{m,0}(x) \le A$ for each x in T.*

PROOF. Fix $a > 0$ and $B > A$. Consider the set $\mathscr{C}$ of continuous functions f on T for which there is μ in $P(T)$ such that $\xi_{\mu,a}(x) \le f(x)$ for all x. Since ξ is convex by Lemma 2.2, the map $\mu \to \xi_{\mu,a}$ is convex and $\mathscr{C}$ is convex. There must exist $f \in \mathscr{C}$ such that $f \le B$, for otherwise the Hahn–Banach theorem would produce $\nu \in P(T)$ such that $\int f\,d\nu \ge B$ whenever $f \in \mathscr{C}$. This would imply $\int \xi_{\nu,a}\,d\nu \ge B > A$, which contradicts our hypothesis. So, for each n, we can find μ_n in $P(T)$ such that $\xi_{\mu_n,2^{-n}}(x) \le A + 2^{-n}$ for all x in T. Since the map $\mu \to \xi_{\mu,0}(x)$ is lower semicontinuous for each x, any weak $*$ cluster point m of the sequence μ_n satisfies $\xi_{m,0}(x) \le A$ for all x. □

The reader has noted that the convexity of ξ is essential in the above proof; in particular, we could not use the function $\Phi^{-1}(1/x)$ instead of ξ.

We now suppose that T is ultrametric, and we turn to the proof of Theorem 1.4. Let s be the smallest integer with $2^{-s} < D$. We fix μ in $P(T)$; we define the process $(X_t)_{t\in T}$ with basic probability space (T, μ) as follows:

$$X_t = \sum_{n\ge s} 2^{-n-2}\Phi^{-1}\left(\frac{1}{\mu(B(t,2^{-n}))}\right)1_{B(t,2^{-n})}.$$

Since $2^{-s} < D$, we have $\inf_t X_t(v) = 0$ for each v in T. We now prove that this process satisfies $(*)$. Fix t, u in T. Let p be the largest such that $d(t,u) \le 2^{-p}$, so that $d(t,u) > 2^{-p-1}$. By ultrametricity, we have $B(t,2^{-n}) = B(u,2^{-n})$ for all $n \le p$. It follows that

$$\left|\frac{X_t - X_u}{d(t,u)}\right| \le \sum_{n>p} 2^{-n+p-1}\left(\Phi^{-1}\left(\frac{1}{\mu(B(t,2^{-n}))}\right)1_{B(t,2^{-n})} + \Phi^{-1}\left(\frac{1}{\mu(B(u,2^{-n}))}\right)1_{B(u,2^{-n})}\right).$$

Since Φ is convex, we get

$$\Phi\left(\frac{X_t - X_u}{d(t,u)}\right) \le \sum_{n>p} 2^{-n+p-1}\left(\frac{1}{\mu(B(t,2^{-n}))}1_{B(t,2^{-n})} + \frac{1}{\mu(B(u,2^{-n}))}1_{B(u,2^{-n})}\right),$$

so $E\Phi((X_t - X_u)/d(t,u)) \le 1$. Since $\inf X_t = 0$, the definition of S shows that

$$E\sup X_t = \int_T\left(\sum_{n\ge s} 2^{-n-2}\Phi^{-1}\left(\frac{1}{\mu(B(t,2^{-n}))}\right)\right) d\mu(t) \le S.$$

Since $\xi(x) \le \Phi^{-1}(1/x)$, we get $\int_T \xi_{\mu,0}(x)\,d\mu(x) \le 4S$, and Theorem 1.4 follows from Proposition 3.2 and Lemma 2.1.

We now compare the integral conditions of (1.1) and (1.2). We first show (what is not absolutely obvious) that the condition (1.2) is stronger than (1.1) (within a factor of 2). If (1.2) holds, for $0 < \varepsilon \le D$ we have in particular $\varepsilon\Phi^{-1}(1/m(B(x,\varepsilon))) \le 8S$, so using (2.1), we have

$$\frac{1}{m(B(x,\varepsilon))} \le \Phi\left(\frac{8S}{\varepsilon}\right) \le \frac{8S}{\varepsilon}\phi\left(\frac{8S}{\varepsilon}\right),$$

which gives

$$\frac{\varepsilon}{8S}\psi\left(\frac{\varepsilon}{8Sm(B(x,\varepsilon))}\right) \le 1,$$

so we get, using (2.1) again,

$$\Phi\left(\psi\left(\frac{\varepsilon}{8Sm(B(x,\varepsilon))}\right)\right) \le \frac{\varepsilon}{8Sm(B(x,\varepsilon))}\psi\left(\frac{\varepsilon}{8Sm(B(x,\varepsilon))}\right) \le \frac{1}{m(B(x,\varepsilon))},$$

so

$$\psi\left(\frac{\varepsilon}{8Sm(B(x,\varepsilon))}\right) \le \Phi^{-1}\left(\frac{1}{m(B(x,\varepsilon))}\right)$$

and finally

$$\int_0^D \psi\left(\frac{\varepsilon}{8Sm(B(x,\varepsilon))}\right) d\varepsilon \le 8S.$$

In the case where $\Phi(x) = x^p$, $p > 1$, conditions (1.1) and (1.2) are not equivalent; but the following result shows that they are equivalent when Φ grows fast enough.

Proposition 3.3. *Assume that there is $L > 0$ and a sequence $\alpha_n > 0$ with $\sum \alpha_n < \infty$ such that*

$$\forall v \geq \alpha_n, \quad 2^n \phi(2^n v) \leq \Phi(2^n L v). \tag{3.3}$$

Then condition (1.2) *implies that*

$$\forall x \in X, \quad \int_0^D \Phi^{-1}\left(\frac{1}{m(B(x, \varepsilon))}\right) d\varepsilon \leq MS,$$

where M depends only on L and $\sum \alpha_n$.

Proof. Condition (3.3) can be reformulated as follows: There exists a function h from $[0,1]$ to $\mathbb{R}$ with $\int_0^1 h(\varepsilon)/\varepsilon \, d\varepsilon < \infty$, such that $(1/\varepsilon)\phi(v/\varepsilon) \leq \Phi(Lv/\varepsilon)$ whenever $v \geq h(\varepsilon)$. Set $a(\varepsilon) = 1/m(B(x, \varepsilon/8S))$. Then condition (2.1) implies $\int_0^{D/8S} \psi(\varepsilon a(\varepsilon)) \, d\varepsilon \leq \frac{1}{2}$. Let $c(\varepsilon) = \varepsilon\psi(\varepsilon a(\varepsilon))$, so that $\int_0^{D/8S} c(\varepsilon)/\varepsilon \, d\varepsilon \leq \frac{1}{2}$. Set $b(\varepsilon) = \max(c(\varepsilon), h(\varepsilon))$. We have

$$a(\varepsilon) = \frac{1}{\varepsilon}\phi\left(\frac{c(\varepsilon)}{\varepsilon}\right) \leq \frac{1}{\varepsilon}\phi\left(\frac{b(\varepsilon)}{\varepsilon}\right) \leq \Phi\left(\frac{Lb(\varepsilon)}{\varepsilon}\right),$$

so that $\Phi^{-1}(a(\varepsilon)) \leq Lb(\varepsilon)/\varepsilon$ and

$$\int_0^{D/8S} \Phi^{-1}(a(\varepsilon)) \, d\varepsilon \leq L\left(\frac{1}{2} + \int_0^1 \frac{h(\varepsilon)}{\varepsilon} \, d\varepsilon\right),$$

since $D \leq S$ (as is easily seen). This proves the result. □

Condition (3.3) fails when Φ is a power function, but is satisfied when Φ increases fast enough. Elementary computations show that $\Phi(x) = x^{\gamma \log\log(x+e)}$ already satisfies (3.3) for all $\gamma > 0$. These functions increase faster than any power of x, but not so much faster.

We now conclude this section by a brief discussion of sample continuity. The most natural property to consider in that respect is the following (as in Theorem 2.9).

(3.4) Every separable process $(X_t)_{t \in T}$ that satisfies $(*)$ is sample continuous.

Another condition, easier to work with, is:

(3.5) For every $\alpha > 0$, there is $\delta > 0$, such that for every process $(X_t)_{t \in T}$ that satisfies $(*)$, we have $E\{\operatorname{Sup}_{d(t,u) \leq \delta}(X_t - X_u) \leq \alpha$.

It is routine to see that (3.5) implies (3.4). When $\Phi(x) = |x|$ the case of the unit interval with the usual distance, is easily checked to be an example that the converse does not hold (see the discussion after Proposition 2.6); but we do not have examples when $\lim_{x \to \infty} \Phi(x)/x = \infty$.

PROPOSITION 3.4. *There exists a universal constant* κ *with the following property: If for any process* $(X_t)_{t \in T}$ *that satisfies* $(*)$, *we have*

$$E \sup_{d(t,u) \le \delta} (X_t - X_u) \le \alpha,$$

then there exists m *in* $P(T)$ *such that*

$$\forall x \in T, \qquad \int_0^\delta \psi\left(\frac{\varepsilon}{\kappa m(B(x,\varepsilon))}\right) d\varepsilon \le \kappa\alpha.$$

If T *is ultrametric, we can even say that*

$$\forall x \in T, \qquad \int_0^\delta \Phi^{-1}\left(\frac{1}{m(B(x,\varepsilon))}\right) d\varepsilon \le \kappa\alpha.$$

PROOF. The first assertion is proved as Theorem 1.3, but using Theorem 2.4 instead of Theorem 2.3; the second assertion is proved as Theorem 1.4; the details are left to the reader. □

4. Sufficient conditions. To show that our results are extensions of the Kôno–Pisier Theorem 1.2, we first relate covering numbers and majorizing measures.

PROPOSITION 4.1. *There exists* m *in* $P(T)$ *such that*

$$\forall x \in T, \qquad \int_0^D \Phi^{-1}(1/m(B(x,\varepsilon)))\, d\varepsilon \le 4\int_0^D \Phi^{-1}(N(\varepsilon, T, d))\, d\varepsilon.$$

PROOF. For $\varepsilon > 0$ and μ in $P(T)$, consider

$$I(\mu, \varepsilon) = \int_T \Phi^{-1}\left(\frac{1}{\mu(B(x,\varepsilon))}\right) d\mu(x).$$

Let $N = N(\varepsilon/2, T, d)$. There exists a partition of T in sets $(B_i)_{i \le N}$ such that for x in B_i, we have $B_i \subset B(x,\varepsilon)$, so $\mu(B(x,\varepsilon)) \ge \mu(B_i)$; this shows that

$$I(\mu,\varepsilon) \le \sum_{i \le n} \mu(B_i)\Phi^{-1}\left(\frac{1}{\mu(B_i)}\right).$$

We have $\sum_{i \le N} \mu(B_i) = 1$, and the map $v \to v\Phi^{-1}(1/v)$ is concave by Lemma 2.2; so we have

$$I(\mu,\varepsilon) \le N \sum_{i \le N} \frac{1}{N}\mu(B_i)\Phi^{-1}\left(\frac{1}{\mu(B_i)}\right) \le N\left(\frac{1}{N}\Phi^{-1}\left(\frac{1}{N}\right)\right) = \Phi^{-1}\left(\frac{1}{N}\right)$$

and we get

$$\int_T \int_0^D \Phi^{-1}\left(\frac{1}{\mu(B(x,\varepsilon))}\right) d\varepsilon\, d\mu(x) \le 2\int_0^{D/2} \Phi^{-1}(N(\varepsilon, T, d))\, d\varepsilon.$$

The conclusion then follows from Lemma 2.1 and Proposition 3.2. □

The following simple result is a good illustration of the basic idea of majorizing measures.

THEOREM 4.2. *Assume that T is ultrametric. Let $m \in P(T)$, and let*

$$\sup_{x \in T} \int_0^D \Phi^{-1}\left(\frac{1}{m(B(x, \varepsilon))} \right) d\varepsilon = M.$$

Then $S(T, d, \Phi) \leq 48M$.

PROOF. Let s be the largest integer such that $2^{-s} \geq D$. For $n \geq p$, we denote by $\mathscr{B}_n$ the family of balls of T of radius 2^{-n}. By hypothesis, and since $\xi(v) \leq \Phi^{-1}(1/v)$, we get

$$\forall x \in T, \qquad \sum_{n>s} 2^{-n} \xi(m(B(x, 2^{-n}))) \leq 3M \tag{4.1}$$

Since T is ultrametric, if $x \in B \in \mathscr{B}_n$, we have $B(x, 2^n) = B$. (This is the essential point where ultrametricity helps.) Since $\mathscr{B}_n$ is a partition of T, if we integrate (4.1) over T with respect to m, we get

$$\sum_{n>s} 2^{-n} \sum_{B \in \mathscr{B}_n} \frac{2^{-n}}{\Psi^{-1}(1/m(B))} \leq 3M. \tag{4.2}$$

For B in $\mathscr{B}_n$, we fix a point $x(B)$ in B. There is a unique element in $\mathscr{B}_s$, namely T, and we denote by $x(T)$ the corresponding point. For B in $\mathscr{B}_n$, $n > s$, we denote by B' the unique ball of $\mathscr{B}_{n-1}$ that contains B. We denote by δ_t the unit point mass at t; and we define the subprobability ν on $T \times T$ given by

$$\nu = \frac{1}{3M} \sum_{\substack{n>s \\ B \in \mathscr{B}_n}} \frac{2^{-n}}{\Psi^{-1}(1/m(B))} \delta_{x(B)} \otimes \delta_{x(B')}.$$

According to Proposition 2.5, it is enough to show that for any Lipschitz function f on T that satisfies

$$\iint_{T \times T \setminus \Delta} \Phi\left(\frac{(f(t) - f(u))}{d(t, u)} \right) d\nu(t, u) \leq 1, \tag{4.3}$$

we have $\sup_{t,u}(f(t) - f(u)) \leq 24M$. We will show that for each t we have $|f(t) - f(x(T))| \leq 12M$. We fix t, and, to simplify the notation, we set $B_n = B(t, 2^{-n})$ and $x_n = x(B_n)$, for $n \geq s$, so $x_s = x(T)$. It follows from (4.3) and the definition of ν that

$$\sum_{n>s} \frac{2^{-n}}{\Psi^{-1}(1/m(B_n))} \Phi(2^{n-1}(f(x_n) - f(x_{n-1}))) \leq 3M.$$

Since f is continuous, we have $f(t) = \lim_{n \to \infty} f(x_n)$, so we have

$$|f(t) - f(x(T))| \leq \sum_{n>s} |f(x_n) - f(x_{n-1})|.$$

To each term on the right, we apply (2.3) with

$$u = |f(x_n) - f(x_{n-1})|, \qquad a = 2^{-n+1}, \qquad b = \frac{2^{-n+1}}{\Psi^{-1}(1/m(B_n))}.$$

We note that $b\Psi(a/b) = 2^{-n+1}\xi(m(B_n))$, so we have

$$\begin{aligned} |f(t) - f(x(T))| &\le \sum_{n>s} \frac{2^{-n+1}}{\Psi^{-1}(1/m(B_n))} \Phi\big(2^{n-1}(f(x_n) - f(x_{n-1}))\big) \\ &\quad + \sum_{n>s} 2^{-n+1}\xi(m(B_n)) \\ &\le 12M \end{aligned}$$

by (4.1) and (4.2). This completes the proof. □

We now turn to general spaces.

THEOREM 4.3. *Assume there are an m in $P(T)$ and numbers A, B, such that the following conditions hold, where s is the largest integer for which $D \le 8^{-s}$:*

$$\forall x \in T, \quad \int_0^D \Phi^{-1}\left(\frac{1}{m(B(x,\varepsilon))}\right) d\varepsilon \le A. \tag{4.4}$$

(4.5) *For any collection of finite subsets $(J_n)_{n>s}$ of T, such that any two elements of J_n are at distance greater than or equal to $4 \cdot 8^{-n}$, we have*

$$\sum_{\substack{n>s \\ x\in J_n}} \frac{8^{-n}}{\Psi^{-1}(1/m(B(x,8^{-n})))} \le B.$$

Then there exists an ultrametric space (U, δ) such that T is the image of U by a 1-Lipschitz map, and that $S(U, \delta, \Phi) \le K(A + B)$, where K is numerical. In particular, $S(T, d, \Phi) \le K(A + B)$.

Condition (4.5) looks artificial at first sight. In the proof of Theorem 4.2, we have shown that, when T is ultrametric, we can deduce (4.5) from (4.4) by integration of (4.4) with respect to m. This does not work in general spaces, because a small motion of x can create a huge variation of $m(B(x, \varepsilon))$. One way to go around the problem is to assume an integral condition that involves not only the measure of $B(x, \varepsilon)$, but that of balls $B(y, \varepsilon)$ for y close to x (see, e.g., [5]). For example (4.5) follows from an integral condition of the type

$$\forall x \in T, \quad \int_0^D \Phi^{-1}\left(\frac{1}{\operatorname{Inf}\{m(B(y,\varepsilon)); d(x,y) \le \varepsilon\}}\right) d\varepsilon \le A \tag{4.6}$$

by integration with respect to m. Unfortunately, we see no reason why a condition of the type (4.6) would be even close to being necessary. So this type of condition is a bit artificial. We will however prove later on that for a very large class of functions Φ, condition (4.5) follows automatically from (4.4), but the argument is not trivial. We will also give an example that illustrates the kind of very complicated structure that might arise.

PROOF OF THEOREM 4.3. We fix $N \geq s$. For $s \leq n \leq N$, we construct subsets $T_{n,i}$ of T and points $x_{n,i}$ of T. For $n = N$ we proceed by induction on $i \leq 1$, in such a way that, setting $T_{N,0} = \varnothing$,

$$(4.7) \qquad m\big(b(x_{N,i}, 8^{-N})\big) = \sup\left\{ m\big(B(x, 8^{-N})\big);\ x \notin \operatorname{int}\left(\bigcup_{j<i} T_{N,j} \right) \right\},$$

$$(4.8) \quad T_{N,i} = \left\{ y \in T \colon d(y, x_{N,i}) < 4 \cdot 8^{-N}, \forall j < i, d(y, x_{N,j}) \geq 4 \cdot 8^{-N} \right\}.$$

Since T is totally bounded, this construction stops at some index $i(N)$, for which $T = \bigcup_{i \leq i(N)} T_{N,i}$. Assuming now that the sets $T_{n+1,i}$ have been constructed, we construct by induction on $i \geq 1$, sets $T_{n,i}$ and points $x_{n,i}$ of T as follows, setting $T_{n,0} = \varnothing$,

$$(4.9) \qquad m\big(B(x_{n,i}, 8^{-n})\big) = \sup\left\{ m\big(B(x, 8^{-n})\big);\ x \notin \operatorname{int}\left(\bigcup_{j<i} T_{n,j} \right) \right\},$$

$$(4.10) \quad T_{n,i} = \bigcup\left\{ T_{n+1,k} \colon T_{n+1,k} \cap B(x_{n,i}, 4 \cdot 8^{-n}) \neq \varnothing, \forall j < i, T_{n+1,k} \not\subset T_{n,j} \right\}.$$

Since there are only finitely many sets $T_{N,i}$, this construction stops at some index $i(n)$ for which $T = \bigcup_{i \leq i(n)} T_{n,i}$.

We observe that by conditions (4.7) to (4.10), we have $d(x_{n,i}, x_{n,j}) \geq 4 \cdot 8^{-n}$ for $i \neq j$. Also, by decreasing induction on n, one sees that the diameter of each set $T_{n,i}$ is less than or equal to 8^{-n+2}.

Consider a subset U_N of T that contains exactly a point in each set $T_{N,i}$, $i \leq i(N)$. Consider the distance δ_N on U_N given by $\delta_N(x, y) = 8^{-m+2}$, where m is the largest integer less than or equal to N such that x and y belong to the same set $T_{m,i}$ for some $i \leq i(m)$. Since $T_{m,i}$ has a diameter less than or equal to 8^{-m+2}, the canonical injection from (U_N, δ_N) into (T, δ) is 1-Lipschitz.

Our next task is to estimate $S(U_N, \delta_N, \Phi)$. For $s \leq n \leq N$, $i \leq i(n)$, we choose any point $y_{n,i} \in U_N \cap T_{n,i}$. We denote $\bar{y}_{n,i}$ the point $y_{n-1,j}$, where j is the index less than or equal to $i(n-1)$ such that $y_{n,i} \in T_{n-1,j}$. It follows from (4.5) that the measure ν on $U_N \times U_N$ given by

$$\nu = B^{-1} \sum_{\substack{s \leq n \leq N \\ i \leq i(n)}} \frac{8^{-n}}{\Psi^{-1}\big(1/m\big(B(x_{n,i}, 8^{-n})\big)\big)} \delta_{y_{n,i}} \otimes \delta_{\bar{y}_{n,i}}$$

is a subprobability. [Here the balls $B(x, \varepsilon)$ are balls in T.] Consider now a function f on U such that

$$(4.11) \qquad \iint_{U_N \times U_N \setminus \Delta} \Phi\left(\frac{f(t) - f(u)}{\delta_N(t, u)} \right) d\nu(t, u) \leq 1$$

and consider a point $x \in U_N$. For $s \le n \le N$, denote $l(n)$ the index $l \le i(n)$ such that $x \in T_{n,l}$, and set $x_n = y_{n,l(n)}$. Since $\bar{y}_{n,l(n)} = y_{n-1,l(n-1)}$, we see from (4.11) and the definition of ν that

$$(4.12) \qquad \sum_{s<n\le N} \frac{8^{-n}}{\Psi^{-1}\left(1/m\left(B\left(x_{n,l(n)}, 8^{-n}\right)\right)\right)} \Phi\left(\frac{f(x_n) - f(x_{n-1})}{8^{-n+2}}\right) \le B.$$

Also, from (4.7) and (4.9) we see that since $x \notin T_{n,j}$ for $j < l(n)$, we have

$$m\left(B\left(x_{n,l(n)}, 8^{-n}\right)\right) \ge m(B(x, 8^{-n}))$$

and thus, from (4.4), we have

$$(4.13) \qquad \sum_{s<n\le N} 8^{-n}\xi\left(m\left(B\left(x_{n,l(n)}, 8^{-n}\right)\right)\right) \le \sum_{s<n\le N} 8^{-n}\xi(m(B(x, 8^{-n}))) \le 2A.$$

As in the proof of Theorem 4.2, we use (2.3) to deduce from (4.12) and (4.13) that $|f(x_n) - f(x_s)| \le K(A + B)$, where K is a number. Since, obviously, $i(s) = 1$, $x_s = y_{s,1}$ is independent of x; so we have $\sup_{U_N \times U_N} |f(t) - f(u)| \le 2K(A + B)$. And it follows from Proposition 4.2 that $S(U_N, \delta_N, \Phi) \le 2K(A + B)$. It follows from Theorem 1.4 that there exists a probability m_N on U_N such that

$$\forall x \in U_N, \qquad \int_0^D \Phi^{-1}\left(\frac{1}{m_N(B(x, \varepsilon))}\right) d\varepsilon \le 16K(A + B),$$

where the balls refer now to the distance δ_N. This bound is independent of N; each point of T is within distance $4 \cdot 8^{-N}$ of U_N. By an easy compactness argument, which we leave to the reader, we can find an ultrametric space (U, δ), such that T is the image of U by a 1-Lipschitz map, and such that there exists a probability m on U such that

$$\forall x \in U, \qquad \int_0^D \Phi^{-1}\left(\frac{1}{m(B(x, \varepsilon))}\right) d\varepsilon \le 16K(A + B).$$

In particular, Theorem 4.2 shows that $S(U, \delta, \Phi) \le K'(A + B)$, where K' is a number. This completes the proof. □

Remark. It follows from Theorem 1.4 that there is on U a probability measure that satisfies (1.2); but it is not clear how to construct the measure using m.

We now investigate the relationship between (4.4) and (4.5). We recall that we say that Ψ satisfies the Δ_2 condition (with constant C) if for $t \ge 0$ we have $\Psi(2t) \le C\Psi(t)$. Typically, Ψ satisfies the Δ_2 condition if $\Phi(x) = x^p$, $p > 1$, but fails it for $\Phi(x) = x(\log(x + 1))^a$, $a > 0$. To simplify the notation, we set $\eta(t) = t\xi(t) = 1/\Psi^{-1}(1/t)$.

LEMMA 4.4. *If Ψ satisfies the Δ_2 condition with constant C, then for $v, w \le 1$, we have*

$$\xi(v) > 2C^2\xi(w) \Rightarrow \eta(w) > 4\eta(v). \tag{4.14}$$

PROOF. Suppose $\eta(w) \le 4\eta(v)$. Set $a = 1/\eta(v)$, $b = 1/\eta(w)$. Since $a \le 4b$, we have

$$1/v = \Psi(a) \le C^2\Psi(b) = C^2/w.$$

Since Φ^{-1} is concave, we have by Lemma 2.1,

$$\xi(v) \le \Phi^{-1}(1/v) \le C^2\Phi^{-1}(1/w) \le 2C^2\xi(w). \qquad \square$$

REMARK. A condition of the type (4.14) actually implies that Ψ satisfies the Δ_2 condition.

THEOREM 4.5. *Assume that for some $\tau > 0$ we have for $v, w \le 1$,*

$$\xi(v) > 2^{2^\tau}\xi(w) \Rightarrow \eta(w) > 4\eta(v). \tag{4.15}$$

Let $m \in P(T)$, such that for some number A, we have

$$\forall x \in T, \qquad \int_0^D \xi(m(B(x,\varepsilon)))\,d\varepsilon \le A. \tag{4.16}$$

Let s be the largest integer for which $D \le 2^{-s}$. For $n \ge s$, consider a subset J_n of T such that any two different points of J_n are at distance greater than or equal to 2^{-n+2}. Then

$$\sum_{\substack{n \ge s \\ x \in J_n}} 2^{-n}\eta(m(B(x,2^{-n}))) \le 2^{\tau+4}A.$$

PROOF. Suppose, if possible, that for families J_n as above and some $N > s$, we have

$$\sum_{\substack{s \le n \le N \\ x \in J_n}} 2^{-n}\eta(m(B(x,2^{-n}))) > 2^{\tau+4}A. \tag{4.17}$$

By induction over $k \ge 1$, we are going to show that

$$\sum_{\substack{s \le n \le N \\ x \in J_n}} 2^{-n}\eta\big(m\big(B\big(x,2^{-n+1}(1-2^{-k})\big)\big)\big) > 2^{k+\tau+3}A. \tag{4.18}$$

Since $\eta(t) \le \eta(1)$ for $t \le 1$, this is a contradiction for k large. For $k = 1$, (4.18) reduces to (4.17). Assume that (4.18) has been proved for k. For $0 \le q \le 2^\tau$, we define

$$B(x,n,k,q) = B\big(x,2^{-n+1}(1-2^{-k}+q2^{-k-1-\tau})\big),$$

so $B(x,n,k,q) \subset B(x,n,k,q+1)$. For x in J_n, we denote by $q(x,n)$ a number

$0 \le q < 2^\tau$ such that the ratio

$$\beta(x,n) = \frac{\xi(m(B(x,n,k,q+1)))}{\xi(m(B(x,n,k,q)))}$$

is maximum among all possible choices of q. [So $\beta(x,n) \le 1$.] We note that

$$y \in B(x,n,k,q) \Rightarrow B(y, 2^{-n-k-\tau}) \subset B(x,n,k,q+1). \tag{4.19}$$

It follows from (4.16) that

$$\forall y \in T, \qquad \sum_{n \ge s} 2^{-n}\xi(m(B(y,2^{-n}))) \le 2A,$$

so we have

$$\forall y \in T, \qquad \sum_{n \ge s} 2^{-n}\xi\big(m\big(B\big(y, 2^{-n-k-\tau}\big)\big)\big) \le 2^{k+\tau+1}A.$$

It follows that

$$\sum_{n \ge s} \int_T 2^{-n}\xi\big(m\big(B\big(y,2^{-n-k-\tau}\big)\big)\big)\, dm(y) \le 2^{k+\tau+1}A. \tag{4.20}$$

It follows from (4.19) that

$$\begin{aligned}
&\int_{B(x,\,n,\,k,\,q(x,\,n))} 2^{-n}\xi\big(m\big(B\big(y,2^{-n-k-\tau}\big)\big)\big)\, dm(y) \\
&\quad \ge 2^{-n}\mu(B(x,n,k,q(x,n)))\xi\big(m(B(x,n,k,q(x,n)+1))\big) \\
&\quad = 2^{-n}\beta(x,n)\eta\big(B(x,n,k,q(x,n))\big) \\
&\quad \ge 2^{-n}\beta(x,n)\eta\big(B\big(x,2^{-n+1}(1-2^{-k})\big)\big).
\end{aligned}$$

Since the balls $B(x, n, k, q(x, n))$, for x in J_n, are disjoint, we get from (4.20) that

$$\sum_{\substack{s \le n \le N \\ x \in J_n}} 2^{-n}\beta(x,n)\eta\big(B\big(x,2^{-n+1}(1-2^{-k})\big)\big) \le 2^{k+\tau+1}A.$$

It follows that

$$\sum 2^{-n}\eta\big(B\big(x,2^{-n+1}(1-2^{-k})\big)\big) \le 2^{k+\tau+2}A,$$

where the summation is taken over the $x \in J_n$, $s \le n \le N$, for which $\beta(x,n) \ge \frac{1}{2}$. It then follows from (4.18) that

$$\sum 2^{-n}\eta\big(B\big(x,2^{-n+1}(1-2^{-k})\big)\big) \ge 2^{k+\tau+2}A,$$

where the summation is over the $x \in J_n$, $s \le n \le N$, for which $\beta(x,n) < \frac{1}{2}$. But if $\beta(x,n) < \frac{1}{2}$, we have

$$2^{2^\tau}\xi\big(m\big(B\big(x,2^{-n+1}(1-2^{-k-1})\big)\big)\big) \le \xi\big(m\big(B\big(x,2^{-n+1}(1-2^{-k})\big)\big)\big),$$

so by (4.15) we have

$$\eta\left(m\left(B\left(x,2^{-n+1}(1-2^{-k-1})\right)\right)\right) \geq 4\eta\left(m\left(B\left(x,2^{-n+1}(1-2^{-k})\right)\right)\right).$$

So we have

$$\sum_{\substack{x\in J_n\\ s\leq n\leq N}} 2^{-n}\eta\left(m\left(B\left(x,2^{-n+1}(1-2^{-k-1})\right)\right)\right) \geq 4\cdot 2^{k+\tau+2}A = 2^{(k+1)+\tau+3}A.$$

This completes the proof. □

Putting together the previous results, we now have the main result of this section.

THEOREM 4.6. *Assume that Ψ satisfies the Δ_2 condition with constant C. Assume that there is $m \in P(T)$ such that*

$$\forall x\in T,\qquad \int_0^D \Phi^{-1}\left(\frac{1}{m(B(x,\varepsilon))}\right)d\varepsilon \leq A.$$

Then T is the image by a contraction of an ultrametric space (U,δ) such that $S(U,\delta,\Phi) \leq KA(1+\log C)$, where K is numerical. In particular, $S(T,d,\Phi) \leq KA(1+\log C)$.

A version of Theorem 4.6, adapted to condition (3.5), can be stated and proved along the same lines; we leave this to the reader.

We are going next to outline the construction of a metric space T with a genuinely complex structure. This space is well adapted to the function $\Phi(x) = |x|$. It is likely that the principle of the construction could be used to get examples for certain functions Φ with $\lim_{x\to\infty}\Phi(x)/x = \infty$ (but very slowly). What happens for functions Φ of the type $x(\log(1+x))^a$, $a>0$, is entirely open.

EXAMPLE 4.7. There exists a compact metric space (T,d), and m in $P(T)$ such that

$$(4.21)\qquad \forall x\in T,\qquad \int_0^D \frac{1}{m(B(x,\varepsilon))}\,d\varepsilon < \infty,$$

but such that whenever (T,d) is the image of an ultrametric space (U,δ) by a contraction, we have $S(U,\delta,|\cdot|) = \infty$. [It is however possible to show that (4.21) imples that $S(T,d,|\cdot|) < \infty$.]

PROOF. According to Theorem 1.4, if an ultrametric space (U,δ) satisfies $S = S(U,\delta,|\cdot|) < \infty$, there exists μ in $P(U)$ such that for x in T, $\int_0^D 1/\mu(B(x,\varepsilon))\,d\varepsilon \leq 8S$. [In the proof of Theorem 1.4, one has to replace $\xi(x)$ by $1/x$.] Integrating this condition with respect to μ as in the proof of Theorem 4.2 shows that $\int_0^{D(U)} N(\varepsilon,U,\delta)\,d\varepsilon < \infty$. If T is the image of U by a contraction, we then have $\int_0^D N(\varepsilon,T,d)\,d\varepsilon < \infty$. So, it is enough to construct (T,d) and m that satisfies (4.21), but such that $\int_0^D N(\varepsilon,T,d)\,d\varepsilon = \infty$. The main point is to show that given A arbitrarily large, one can find a space (T, d) and m in $P(T)$ such that

$$A \operatorname*{Sup}_{x\in T} \int_0^D \frac{1}{m(B(x,\varepsilon))} d\varepsilon \le \int_0^D N(\varepsilon, T, d)\, d\varepsilon.$$

One then concludes in a routine way by gluing these spaces together. (While now it might not be obvious how to do that, this should be the case after reading the rest of the proof.) We fix $q > 1$. For $p \ge 0$, we consider the space $Z_p = \{l2^{-q-p}; 0 \le l \le 2^q\}$, with the distance δ_p induced by $\mathbb{R}$. We set $a = (\sum_{l=0}^{2^q} 3^{-l})^{-1}$, so $\frac{3}{4} \le a \le 1$. We consider the probability μ_p on Z_p given by $\mu_p(\{l2^{-p-q}\}) = a3^{-l}$. [The idea is that $\mu_p(B(x,\varepsilon))$ will be a very fast increasing function of ε; as seen in the proof of Theorem 4.5, this is at the heart of the matter.] Set $C = C(q) = 3^{2^q}2^{-q}$. We first note that for some universal constant K_1, we have for $0 \le l \le 2^q$,

$$\begin{aligned}2^{-p-q}3^l &\le \sum_{p\le i\le p+q} 2^{-i}/\mu_p\big(B(l2^{-p-q},2^{-i})\big)\\ &\le K_1 2^{-p-q}3^l \le K_1 2^{-p}C,\end{aligned} \tag{4.22}$$

because the term for $i = p + q$ is dominant. It follows that

$$\int_{Z_p} \sum_{p\le i\le p+q} 2^{-i}/\mu_p\big(B(x,2^{-i})\big)\, d\mu_p(x) \le K_1 2^{-p+1}. \tag{4.23}$$

Also, it is clear that

$$\sum_{p\le i\le p+q} 2^{-i}N\big(2^{-i}, Z_p, d_p\big) \ge q2^{-p}. \tag{4.24}$$

To simplify the notation, we make the following convention. For a point x in a finite metric space T, and a measure μ on T, we denote by $S(x,\mu)$ the sum $\sum 2^{-i(l)}/\mu(B(x,2^{-i(l)}))$, where the sequence $i(l)$ is defined by induction by $i(1) = 1$ and

$$i(l+1) = \inf\{i > i(l);\ B(x,2^{-i(l)}) \setminus B(x,2^{-i}) \ne \varnothing\}.$$

(This sequence has a finite length.) We note that

$$\sum_{i\ge 1} 2^{-i}/\mu\big(B(x,2^{-i})\big) \le 2S(x,\mu).$$

By induction over $k \ge 1$ we construct finite metric spaces (Y_k, d_k) of diameter less than or equal to 1, $n(k) \in \mathbb{N}$ and $m_k \in P(Y_k)$ such that the following hold.

$$\text{If } x, y \in Y_k,\ x \ne y, \text{ then } d_k(x,y) \le 2^{-n(k)}. \tag{4.25}$$

$$\forall x \in Y_k, \qquad S(x, m_k) \le 2(1 + K_1)C. \tag{4.26}$$

$$\sum_{1\le i\le n(k)} 2^{-i}N\big(2^{-i}, Y_k, d_k\big) \ge \frac{q}{2K_1}\int_{Y_k} S(x, m_k)\, dm_k(x). \tag{4.27}$$

$$m_k(\{x \in Y_k;\ S(x,m_k) \le C\}) \le (1 - a3^{-2^q})^k. \tag{4.28}$$

The point of the construction is that if k is large enough that $(1 - a3^{-2^q})^k \leq \frac{1}{2}$, we get from (4.26) and (4.28) that

$$\sum_{i \leq n(k)} 2^{-i} N(2^{-i}, Y_k, d_k) \geq qC/4K_1.$$

Since q is arbitrarily large, and since K_1 does not depend on q, when we compare to (4.26), we see that we have the example we seek. We now proceed to the induction. For $k = 1$, we take $(Y_1, d_1) = (Z_0, \delta_0)$, $n(1) = q$. (4.26) follows from (4.22), and (4.27) follows from (4.23) and (4.24); (4.28) follows from (4.22) for $l = 2^q$, since $\mu_0(\{1\}) \geq a3^{-2^q}$. Assume now that $Y_k, d_k, m_k, n(k)$ have been constructed. The space Y_{k+1} is a union of pieces H_x for x in Y_k. If $t \in H_x$, $u \in H_y$, $x \neq y$, we have $d_{k+1}(t, u) = d_k(x, y)$, and also $m_{k+1}(H_x) = m_k(\{x\})$. If $S(x, m_k) \geq C$, we just take $H_x = \{x\}$. If $S(x, m_k) < C$, we consider two cases.

Case 1. If $2^{-n(k)-1} \geq Cm_k(\{x\})$, denote by $p = p(x)$ the largest integer for which $2^{-p} \geq Cm_k(\{x\})$, and set $H_x = Z_p$, $m_{k+1}(A) = m_k(\{x\})\mu_p(A)$ for $A \subset H_x$, $d_{k+1}(t, u) = \delta_p(t, u)$ for $x, u \in H_x$.

Case 2. If $2^{-n(k)-1} \leq Cm_k(\{x\})$, let N be the smallest integer for which $N2^{-n(k)-1} \geq Cm_k(\{x\})$. We take for H_x the union of N copies $(U_i)_{i \leq N}$ of $Z_{n(k)+1}$. If $t \in U_i$, $u \in U_j$, $i < j$, we set $d_{k+1}(t, u) = 2^{-n(k)-1}$; if $t, u \in U_i$, we set $d_{k+1}(t, u) = \delta_{n(k)+1}(t, u)$; if $A \subset U_i$, we set $m_{k+1}(A) = m_k(\{x\})\mu_{n(k)+1}(A)/N$.

Finally, we take $n(k + 1)$ large enough that (4.25) holds. It is straightforward to see that (4.25) to (4.28) hold so we leave the checking to the reader. □

5. The homogeneous case. We first study the case where $T = [0,1]^k$, $k \geq 2$, and the distance is the usual distance, but where Φ is arbitrary.

Theorem 5.1. *There is a constant K, depending only on k, such that the following hold.*

(a) *If $S = S(T, d, \Phi)$, we have $\int_0^1 \varepsilon^{k-1}\Psi(\varepsilon^{1-k}/KS)\, d\varepsilon \leq K$.*

(b) *If for some $A > 0$, we have $\int_0^1 \varepsilon^{k-1}\Psi(\varepsilon^{1-k}/A)\, d\varepsilon \leq B$, then $S(T, d, \Phi) \leq KA(1 + B)$.*

Proof. The Euclidean distance is not well adapted to the study of cubes, so we use instead the distance d given $d(x, y) = \max_{k \leq 1} |x_i - y_i|$ when $x = (x_i)_{i \leq k}$, $y = (y_i)_{i \leq k}$. Since this is equivalent to the usual distance, one checks easily that it is enough to prove the theorem for (T, d). We denote by dx the Lebesgue measure on T.

Proof of (a). By Theorem 1.3, there is m in $P(T)$ such that

$$\forall x \in T, \qquad \int_0^1 \psi\left(\frac{\varepsilon}{8Sm(B(x, \varepsilon))}\right) d\varepsilon \leq 4S.$$

We note that by (2.1), we have $\Psi(u) \le u\psi(u)$, for $u > 0$, so we get

$$(5.1) \qquad \forall x \in T, \qquad \int_0^1 \varepsilon^{-1} m(B(x,\varepsilon)) \Psi\left(\frac{\varepsilon}{8Sm(B(x,\varepsilon))}\right) d\varepsilon \le \frac{1}{2}.$$

Now, by Lemma 2.2, the function $u \to u\Psi(\varepsilon/u)$ is convex. If we integrate (5.1) with respect to dx, and we set $a(\varepsilon) = \int_T m(B(x,\varepsilon))\,dx$, we get

$$\int_0^1 \varepsilon^{-1} a(\varepsilon) \Psi(\varepsilon/8Sa(\varepsilon))\,d\varepsilon \le \tfrac{1}{2}.$$

Now

$$a(\varepsilon) = \int_{\{d(x,y)\le\varepsilon\}} dm(y)\,dx = \int\left(\int_{\{d(x,y)\le\varepsilon\}} dx\right) dm(y)$$
$$\le \int (2\varepsilon)^k\,dm(y) \le (2\varepsilon)^k.$$

Since $u \to u\Psi(\varepsilon/u)$ decreases, the proof is complete. □

Proof of (b). We could produce a measure ν on $T \times T$ that satisfies the conditions of Proposition 2.5; but this would be cumbersome, so we will proceed directly. (The proof actually amounts, with the notation of Theorem 2.3, to using a functional θ on G that does not arise from a probability on $T \times T$.) As in the proof of Theorem 2.3, we can reduce to the case where Ω is finite. In that case, the maps $t \to X_t(\omega)$ are Lipschitz, so they are almost everywhere differentiable; since $E\Phi((X_t - X_u)/d(t,u)) \le 1$ for each t, u, we have $E\Phi(\partial X_t/\partial t_i) \le 1$ whenever $i \le k$, for almost all t. Since Φ is convex, we have

$$E\Phi\left(\frac{1}{k}\sum_{i\le k}\left|\frac{\partial X_t}{\partial t_i}\right|\right) \le 1,$$

so

$$E\left(\int_T \Phi\left(\frac{1}{k}\sum_{i\le k}\left|\frac{\partial X_t}{\partial t_i}\right|\right)\right) dt \le 1.$$

We are now reduced to the following question.

Let f be a Lipschitz function on T. Set $c(t) = \sum_{i\le k}|\partial f/\partial t_i|$. Assuming that $\int\Phi(c(t)/k)\,dt \le C$, we have to prove that $\mathrm{Sup}_{t,u}(f(t) - f(u)) \le KA(C+B)$. This is a Sobolev-type inequality.

Let us fix $t = (t_i)_{i\le k}$ in T. We consider the probability measure λ_t on the boundary H of T such that on each face $\{u_i = 0\} \cap H$ (resp. $\{u_i = 1\} \cap H$) of H, for $i \le k$, λ has density t_i/k [resp. $(1 - t_i)/k$] with respect to the $(k-1)$-dimensional measure of the face. For any function g on T, we have

$$(5.2) \qquad \int_T g(x)\,dx = \int_H\int_0^1 g((1-s)t + sx)s^{k-1}\,ds\,d\lambda_t(x).$$

For $0 \le s \le 1$, we define

$$a(s) = \int_H f((1-s)t + sx)\, d\lambda_t(x),$$

so $a(0) = f(t)$, and $a(1) = \int f\, d\lambda_t$. Set $b(s) = \int_H c((1-s)t + sx)\, d\lambda_t(x)$, and denote by Df the differential of f. Then $a'(s)$ exists a.e. and

$$|a'(s)| = \left| \int_H Df((1-s)t + sx)(x - t)\, d\lambda_t(x) \right|$$
$$\le \int c((1-s)t + sx)\, d\lambda_t(x) = b(s),$$

because the components of $x - t$ are less than or equal to 1. Now we have, by convexity of Φ, Jensen's inequality and (5.2) that

$$\int_0^1 \Phi\left(\frac{b(s)}{k} \right) s^{k-1}\, ds \le \int_H \int_0^1 \Phi\left(\frac{1}{k} c((1-s)t + sx) \right) s^{k-1}\, ds\, d\lambda_t(x)$$
$$= \int_T \Phi\left(\frac{c(x)}{k} \right) dx \le C.$$

Using (2.3), we have

$$b(s) \le kAs^{k-1}\Phi\left(\frac{b(s)}{k} \right) + kAs^{k-1}\Psi\left(\frac{1}{As^{k-1}} \right). \tag{5.3}$$

Since a is Lipschitz, we have

$$\left| f(t) - \int f\, d\lambda_t \right| = |a(0) - a(1)| \le \int_0^1 |a'(s)|\, ds \le \int_0^1 b(s)\, ds,$$

and this is less than or equal to $K(A + C)$.

The theorem will be proved if we show that for t, u in T,

$$\left| \int f\, d\lambda_t - \int f\, d\lambda_u \right| \le K(A + C).$$

Writing $F_i = H \cap \{u_i = 0\}$, $F_i' = H \cap \{u_i = 1\}$ and denoting by μ the $(k-1)$-dimensional measure on F_i, F_i', we have

$$\left| \int f\, d\lambda_t - \int f\, d\lambda_u \right| \le \frac{1}{k} \sum_{i \le k} \left| \int_{F_i} f\, d\mu - \int_{F_i'} f\, d\mu \right|.$$

But we have

$$\left| \int_{F_i} f\, d\mu - \int_{F_i'} f\, d\mu \right| \le \int_T \left| \frac{\partial f}{\partial t_i}(x) \right| dx$$

and thus

$$\left| \int f\, d\lambda_t - \int f\, d\lambda_u \right| \le \int \frac{c(t)}{k}\, dt.$$

Writing, by (2.3),

$$\frac{c(t)}{k} \le A\Phi\left(\frac{c(t)}{k}\right) + A\Psi\left(\frac{1}{A}\right) \le A\Phi\left(\frac{c(t)}{k}\right) + A\int_0^1 s^{k-1}\Psi\left(\frac{1}{As^{k-1}}\right) ds$$

completes the proof. □

REMARKS. 1. It would be interesting to extend Theorem 5.1 to compact subsets of $\mathbb{R}^{k-1}$ with nonempty interior. While the first part of the theorem carries through, it is unclear how to proceed with the second part. One problem in particular is that it is unclear if a process that satisfies $(*)$ and is defined on a subset L of $\mathbb{R}^{k-1}$ can be extended to a cube containing L, while still satisfying $(*)$.

2. The proof has shown that the necessary condition of Theorem 1.3 is sufficient.

We now turn to the extension of Theorem 1.1. We fix a concave function η from $[0,1]$ to $[0,1]$, with $\eta(0) = 0$, so $\eta(x)/x$ decreases on $[0,1]$. We set $d(t, u) = \eta(|t-u|)$. We construct by induction the sequence $(a_n)_{n\ge 1}$ as follows. We set $a_1 = \frac{1}{2}$, and

$$a_{n+1} = \sup\{0 < a \le a_n;\ \eta(a) \le \eta(a_n)/2;\ \eta(a)/a \ge 2[\eta(a_n)/a_n]\}.$$

If the set on the right is empty, we stop the construction. We note that $a_{n+1} \le a_n/2$, and that we have either $\eta(a_{n+1}) = \eta(a_n)/2$ or $\eta(a_{n+1})/a_{n+1} = 2[\eta(a_n)/a_n]$.

THEOREM 5.2. *For some universal constant K, we have*

$$(5.4)\quad K^{-1}\sum \eta(a_n)\Phi^{-1}(1/a_n) \le S([0,1], d, \Phi) \le K\sum \eta(a_n)\Phi^{-1}(1/a_n)$$

(*where the summation involves the a_n actually constructed*).

We first make the link between (5.4) and the integral condition of Theorem 1.2.

PROPOSITION 5.3. *Assume that for some constant $C > 0$, we have, for each $0 < a \le 1$,*

$$\int_0^a \Phi^{-1}\left(\frac{1}{t}\right) dt \le Ca\Phi^{-1}\left(\frac{1}{a}\right), \qquad \int_a^1 \Phi^{-1}\left(\frac{1}{t}\right)\frac{dt}{t} \le C\Phi^{-1}\left(\frac{1}{a}\right).$$

Then for some universal constant K we have

$$\frac{1}{KC}\int_0^{1/2}\frac{\eta(t)}{t}\Phi^{-1}\left(\frac{1}{t}\right) dt \le S([0,1], d, \Phi) \le K\int_0^{1/2}\frac{\eta(t)}{t}\Phi^{-1}\left(\frac{1}{t}\right) dt.$$

PROOF. The right-hand inequality follows from (5.4) and the fact that, since $a_{n+1} \le a_n/2$, we have

$$\int_{a_{n+1}}^{a_n} \frac{n(t)}{t}\Phi^{-1}\left(\frac{1}{t}\right) dt \ge (a_n - a_{n+1})\frac{\eta(a_n)}{a_n}\Phi^{-1}\left(\frac{1}{a_n}\right) \ge \frac{\eta(a_n)}{2}\Phi^{-1}\left(\frac{1}{a_n}\right).$$

For the left-hand inequality we distinguish two cases.

CASE 1. $\eta(a_{n+1}) = \eta(a_n)/2$. Then $\eta(t) \le 2\eta(a_{n+1})$ for $a_{n+1} \le t \le a_n$, so

$$\int_{a_{n+1}}^{a_n} \frac{n(t)}{t}\Phi^{-1}\left(\frac{1}{t}\right) dt \le 2\eta(a_{n+1})\int_{a_{n+1}}^{1} \Phi^{-1}\left(\frac{1}{t}\right)\frac{dt}{t}$$
$$\le 2C\eta(a_{n+1})\Phi^{-1}\left(\frac{1}{a_{n+1}}\right).$$

CASE 2. $\eta(a_{n+1})/a_{n+1} = 2[\eta(a_n)/a_n]$. Then $n(t)/t \le 2[\eta(a_n)/a_n]$ for $a_{n+1} \le t \le a_n$, so

$$\int_{a_{n+1}}^{a_n} \frac{n(t)}{t}\Phi^{-1}\left(\frac{1}{t}\right) dt \le 2\frac{\eta(a_n)}{a_n}\int_{0}^{a_n}\Phi^{-1}\left(\frac{1}{t}\right) dt \le 2C\eta(a_n)\Phi^{-1}\left(\frac{1}{a_n}\right).$$

It follows that

$$\int_{0}^{1/2}\frac{n(t)}{t}\Phi^{-1}\left(\frac{1}{t}\right) dt \le 4C\sum\eta(a_n)\Phi^{-1}\left(\frac{1}{a_n}\right). \qquad \square$$

The proof of Theorem 5.2 is rather instructive; but it will be clearer to discuss the main feature after going through the details. We will assume throughout the proof that the sequence (a_n) is infinite. The other case is similar, and the necessary modifications are left to the reader.

PROOF OF THE LEFT INEQUALITY OF (5.4). Set $h_n = \frac{1}{72}[\eta(a_n)/a_n]\Phi^{-1}(1/a_n)$, so the sequence (h_n) increases. Consider the function f on $\mathbb{R} \setminus \{0\}$, that satisfies $f(-x) = f(x)$, $f(x) = 0$ if $|x| \ge \frac{1}{2}$ and is of slope $-h_n$ between a_{n+1} and a_n. We denote by m Legesgue's measure on $[0,1]$. Consider the process $(X_t)_{t\in[0,1]}$, with basic probability space $([0,1], m)$, given by $X_t(\omega) = f(\theta(\omega - t))$, where $\theta(\omega - t)$ is the unique point in $]-\frac{1}{2}, \frac{1}{2}]$ congruent to $\omega - t$ modulo 1.

Assume that we know that

$$(5.5) \qquad \forall t, 0 < t \le \frac{1}{2}, \qquad \int\Phi\left(\frac{f(t+u) - f(u)}{\eta(t)}\right) du \le 1.$$

Then we see easily that (X_t) satisfies $(*)$. Moreover, $\operatorname{Inf} X_t(w) = 0$ for each t, and $\operatorname{Sup} X_t(w) = \operatorname{Sup} f$. So we have $\operatorname{Sup} f \le S([0,1], d, \Phi)$. Since $a_{n+1} \le a_n/2$, we have $\operatorname{Sup} f \ge \frac{1}{144}\sum_{n\ge 1}\eta(a_n)\Phi^{-1}(1/a_n)$, and this finishes the proof.

We now prove (5.5). Let $p \ge 1$ be such that $a_{p+1} \le t \le a_p$. We distinguish two cases.

CASE 1. We have $\eta(a_{p+1}) = \eta(a_p)/2$. In that case, $\eta(t) \geq \eta(a_{p+1}) \geq \eta(a_p)/2$. We set $g(u) = [f(u) - f(a_p)]^+$, $h(u) = \min(f(u), f(a_p))$. The convexity of Φ shows that

$$\Phi\left(\frac{f(t+u) - f(u)}{\eta(t)}\right) \leq \frac{1}{3}\Phi\left(\frac{3g(t+u)}{\eta(t)}\right) + \frac{1}{3}\Phi\left(\frac{3g(u)}{\eta(t)}\right) + \frac{1}{3}\Phi\left(\frac{3(h(t+u) - h(u))}{\eta(t)}\right),$$

so it is enough to prove that

$$\int \Phi\left(\frac{6g(u)}{\eta(a_p)}\right) du \leq 1, \tag{5.6}$$

$$\int \Phi\left(\frac{6(h(t+u) - h(u))}{\eta(a_p)}\right) du \leq 1. \tag{5.7}$$

We prove (5.6). We have

$$\int \Phi\left(\frac{6g(u)}{\eta(a_p)}\right) du \leq \sum_{n \geq p} a_n \Phi\left(\frac{6(f(a_{n+1}) - f(a_p))}{\eta(a_p)}\right). \tag{5.8}$$

Also,

$$|f(a_{n+1}) - f(a_p)| \leq \sum_{p \leq i \leq n} a_i h_i = \frac{1}{72} \sum_{p \leq i \leq n} \eta(a_i) \Phi^{-1}\left(\frac{1}{a_i}\right),$$

so

$$\frac{6|f(a_{n+1}) - f(a_p)|}{\eta(a_p)} \leq \frac{1}{12} \sum_{p \leq i \leq n} \frac{\eta(a_i)}{\eta(a_p)} \Phi^{-1}\left(\frac{1}{a_i}\right).$$

Since $\eta(a_{n+1}) \leq \eta(a_n)/2$, we have $\sum_{i \geq p} \eta(a_i)/\eta(a_p) \leq 2$, so the convexity of Φ shows that

$$\Phi\left(\frac{6(f(a_{n+1}) - f(a_p))}{\eta(a_p)}\right) \leq \frac{1}{6} \sum_{p \leq i \leq n} \frac{\eta(a_i)}{a_i \eta(a_p)},$$

so from (5.8)

$$\int \Phi\left(\frac{6g(u)}{\eta(a_p)}\right) du \leq \frac{1}{6} \sum_{n \geq p} \sum_{p \leq i \leq n} \frac{a_n \eta(a_i)}{a_i \eta(a_p)}.$$

Since $\eta(a_i) \leq 2^{p-i}\eta(a_p)$ and $a_n \leq 2^{i-n} a_i$, we have

$$\int \Phi\left(\frac{6g(u)}{\eta(a_p)}\right) du \leq \frac{1}{6} \sum_{n \geq p} (n - p + 1) 2^{-n+p} = \frac{1}{6} \sum_{i \geq 0} (i+1) 2^{-i} = 1$$

and this proves (5.6).

We prove (5.7). We observe that for $1 \le i \le p-1$, if the interval $[u, u+t]$ does not meet the interval $[-a_i, a_i]$, we have $|h(t+u) - h(u)| \le th_i$; so this holds outside a set of measure less than or equal to $4a_i$. Also we always have $|h(t+u) - h(u)| \le th_{p-1}$.

So we have

$$\int \Phi\left(\frac{6(h(t+u) - h(u))}{\eta(a_p)}\right) du \le \sum_{i \le p-1} 4a_i \Phi\left(\frac{6th_i}{\eta(a_p)}\right)$$

$$\le \sum_{i \le p-1} 4a_i \Phi\left(\frac{a_p \eta(a_i)}{12 a_i \eta(a_p)} \Phi^{-1}\left(\frac{1}{a_i}\right)\right).$$

Using the convexity of Φ, and since $\eta(a_i)/a_i \le 2^{i-n}[\eta(a_p)/a_p]$, this is $\le \frac{1}{3}$. This proves (5.7).

CASE 2. We have $\eta(a_{p+1})/a_{p+1} = 2[\eta(a_p)/a_p]$. In that case $\eta(t)/t \ge \eta(a_p)/2a_p$. We set

$$g(u) = \left[f(u) - f(a_{p+1})\right]^+, \qquad h(u) = \operatorname{Inf}\left(f(u), f(a_{p+1})\right).$$

The computation is similar to that in Case 1 and is left to the reader. □

PROOF OF THE RIGHT INEQUALITY OF (5.4). We will use Proposition 2.5. Define $a_0 = 0$. For $n \ge 0$, denote by $k(n)$ the largest integer for which $2^{-k(n)} \ge a_n$, so $k(1) = 1$, $k(0) = 0$, $k(n+1) > k(n)$ (since $a_{n+1} \le a_n/2$) and $a_n \ge 2^{-k(n)-1}$. We set

$$J = \{1\} \cup \{n > 2; \eta(a_n) = \eta(a_{n-1})/2\},$$

so for $n \notin J$, we have $\eta(a_n)/a_n = 2\eta(a_{n-1})/a_{n-1}$. For $x \in [0,1]$, $n \ge 0$, we denote by $I_{n,x}$ the unique interval of the type $[i2^{-k(n)}, (i+1)2^{-k(n)}[$ that contains x. For each $n \ge 1$, x in $[0,1]$, let $\nu_{n,x}$ be such that $2^{-k(n-1)}\nu_{n,x}$ is the restriction of Lebesque's measure to $I_{n-1,x}$. Let $\mu_{n,x}$ be such that $2^{-k(n)}\mu_{n,x}$ is the restriction of Lebesgue measure to the unique interval $I'_{n,x}$ of the type $[i2^{-k(n)}, (i+1)2^{-k(n)}[$ immediately to the right of $I_{n,x}$. (If there is no such interval take $\nu_{n,x} = \delta_x$.) We set

$$c_n = \eta(a_n)\xi(a_n) \quad \text{if } n \in J, \qquad c_n = \eta(a_{n-1})\xi(a_{n-1}) \quad \text{if } n \notin J.$$

From Lemma 2.1, and the fact that $\eta(1)\Phi^{-1}(1) \le 2\eta(\frac{1}{2})\Phi^{-1}(\frac{1}{2})$, it follows that $A := \sum_{n \ge 1} c_n \le 3\sum_{n \ge 1} a_n \Phi^{-1}(1/a_n)$. Set $\nu_x = 1/A(\sum_{n \in J} c_n \nu_{n,x} + \sum_{n \notin J} c_n \mu_{n,x})$. Consider the subprobability $\nu = \int_0^1 \delta_x \otimes \nu_x \, dx$. According to Proposition 2.5, it is enough to show that if a Lipschitz function f satisfies

$$\int \Phi\left(\frac{f(t) - f(u)}{d(t,u)}\right) d\nu(t,u) \le 1, \tag{5.9}$$

then for some universal constant K, for all t in $[0,1]$, we have

$$\left|f(x) - \int_0^1 f(t)\,dt\right| \le KA.$$

We fix x. We set $b_n = \int f\,d\nu_{n,x}$, so $f(x) = \lim b_n$ and $b_1 = \int_0^1 f(t)\,dt$. Set $q(n) = 2^{k(n)-k(n-1)}$. For $n \ge 1$, $I_{n-1,x}$ is the union of $q(n)$ intervals of the type $[i2^{-k(n)}, (i+1)2^{-k(n)}[$, that we denote by $J_{n,l,\,0\le l<q(n)}$. Let $\theta_{n,l}$ be the probability such that $2^{-k(n)}\theta_{n,l}$ is the restiction of Lebesgue's measure to $J_{n,l}$. We set $b_{n,l} = \int f\,d\theta_{n,l}$. We note that

$$b_n = q(n)^{-1} \sum_{0\le l<q(n)} b_{n,l}. \tag{5.10}$$

Also, we note that by definition of ν, we have

$$\begin{aligned}\nu \ge {} & \frac{1}{A}\sum_{n\in J} 2^{-k(n)} c_n \nu_{n+1,x} \otimes \nu_{n,x} \\ & + \frac{1}{A}\sum_{n\notin J} 2^{-k(n)} c_n \sum_{0\le l\le q(n)-2} \theta_{n,l} \otimes \theta_{n,l+1}.\end{aligned} \tag{5.11}$$

We note that for t, u in $I_{n-1,x}$, we have $d(t,u) \le \eta(2^{-k(n-1)}) \le 2\eta(a_{n-1})$; and for t in $J_{n,l}$, u in $J_{n,l+1}$, we have $d(t,u) \le \eta(2^{-k(n)+1}) \le 4\eta(a_n)$. So, by (5.9) and the convexity of Φ, we have that

$$\sum_{n\in J} a_n c_n \Phi\left(\frac{b_n - b_{n+1}}{2\eta(a_n)}\right) + \sum_{n\notin J}\ \sum_{0\le l\le q(n)-2} c_n a_n \Phi\left(\frac{b_{n,l} - b_{n,l+1}}{4\eta(a_n)}\right) \le A.$$

We note that $a_{n-1}/2a_n \le q(n) \le 2a_{n-1}/a_n$. If $n \notin J$, we have $1/q(n)\eta(a_n) \ge a_n/2a_{n-1}\eta(a_n) = 1/4\eta(a_{n-1})$. Since Φ is convex, we have

$$\begin{aligned}&\sum_{0\le l\le q(n)-2} \Phi((b_{n,l} - b_{n,l+1})/4\eta(a_n)) \\ &\quad \ge (q(n)-1)\Phi\Big(\sum_{0\le l\le q(n)-2} |b_{n,l} - b_{n,l+1}|/4q(n)\eta(a_n)\Big) \\ &\quad \ge (a_{n-1}/4a_n)\Phi\Big(\sum_{0\le l\le q(n)-2} |b_{n,l} - b_{n,l+1}|/16\eta(a_{n-1})\Big).\end{aligned} \tag{5.12}$$

Now b_{n+1} is one of the $b_{n,l}$, say $b_{n+1} = b_{n,j}$. For $0 \le l < q(n)$, we have

$$|b_{n+1} - b_{n,l}| = |b_{n,j} - b_{n,l}| \le \sum_{0\le l\le q(n)-2} |b_{n,l} - b_{n,l+1}|,$$

so from (5.10) we have

$$|b_{n+1} - b_n| \le \sum_{0\le l\le q(n)-2} |b_{n,l} - b_{n,l+1}|.$$

It now follows from (5.12) that

$$\sum_{n\in J} \frac{\eta(a_n)}{\Psi^{-1}(1/a_n)}\Phi\left(\frac{b_n - b_{n+1}}{2\eta(a_n)}\right) + \sum_{n\notin J} \frac{\eta(a_{n-1})}{\Psi^{-1}(1/a_{n-1})}\Phi\left(\frac{b_n - b_{n+1}}{16\eta(a_{n-1})}\right) \le 4A.$$

It now follows from (2.3), and by a now familiar computation that $\sum_{n\geq 1}|b_n - b_{n-1}| \leq 64A$. This completes the proof. □

The method of the proofs of Theorems 4.2, 5.1 and 5.2 are similar. If f is Lipschitz on T, and $\int_{T\times T}\Phi([f(t)-f(u)]/d(t,u))\,d\nu(t,u) \leq 1$, to bound $\sup_{t,u}(f(t)-f(u))$ one finds bounds for $|f(x) - \int f\,d\lambda|$, where λ is a measure independent of x. This is done by finding a suitable "path" from δ_x to λ. In the ultrametric case, we go from δ_x to λ by big jumps. In the case of Theorem 5.1, one has to make a large number of very small jumps (as becomes more apparent if one writes a "discretized" version of the proof). The proof of Theorem 5.2 presents features of both cases. If $n \in J$, we jump directly from $\nu_{n,x}$ to $\nu_{n-1,x}$, but if $n \notin J$, we make $q(n)$ small jumps in between. The next result we present is definitely a level of complexity above the previous examples, but it is a fascinating fact that beyond the details, the proof is similar to that of Theorem 5.2, in the sense that it mixes the techniques of small jumps and big jumps. It is tempting to believe that these similarities do not arise by chance and that there is some general underlying principle, yet to be formulated.

We denote by U the one-dimensional torus, provided with the usual distance d'. We fix $q \in \mathbb{N}$. We provide U^q with the distance d_1 given by $d_1(z, y) = \max_{i\leq q} d'(x_i, y_i)$ for $x = (x_i)_{i\leq q}$, $y = (y_i)_{i\leq q}$. We denote by H a ultrametric compact group, with distance d_2. We provide $T = U^q \times H$ with the distance d given by $d((x,u),(y,v)) = \max(d_1(x,y), d_2(u,v))$. We propose to compute $S(T, d, |\cdot|^p)$. The complexity of the situation entirely arises from the metric structure of T. (An identical result would hold for $T = [0,1]^q \times H$.) We first note that by Theorem 5.1, if $q \geq 2$, and $S(T, d, |\cdot|^p) < \infty$, we must have $p > q$. It can also be shown that (provided H is infinite) this is also the case for $q = 1$. So we fix $p > q$. We define r by $1/r + q/p^2 = 1$. We denote by $0, e', e$ the units, and by m_1, m_2, m the normalized Haar measures of U^q, H, T, respectively. We set $\mu_n = m_2(B(e', 2^{-n}))$ (where the ball is of course in H). We denote by n_0 the largest integer for which $2^{-n_0} \geq D(H)$.

Theorem 5.4. *There is constant K, depending on p and q only, such that*

$$K^{-1}S \leq \left[1 + \sum_{n\geq n_0} 2^{-n}\mu_n^{-1/p} + \left(\sum_{n\geq n_0}\left(2^{-n(p-q)}/\mu_n\right)^{r/p}\right)^{1/r}\right] \leq KS. \tag{5.13}$$

In the proof, we denote by K a number depending only on p and q, that is not necessarily the same at each line.

Proof of the right inequality. We observe first that $S = S(T, d, |\cdot|^p) \geq S(U^q, d_1, |\cdot|^p)$, so $1 \leq KS$. Also, we have $S \geq S(H, d_2, |\cdot|^p)$. By Theorem 1.4, there is a $\nu \in P(T)$ with $\int_0^{D(H)}(\nu(B(e',\varepsilon)))^{-1/p}\,d\varepsilon \leq KS$. If we integrate with respect to m_2, and we use the convexity of $t \to t^{-1/p}$, we see that

$\int_0^{D(H)}(m_2(B(e',\varepsilon)))^{-1/p}\,d\varepsilon \le KS$, so

$$\sum_{n\ge n_0} 2^{-n}\mu_n^{-1/p} \le KS. \tag{5.14}$$

For $n \ge n_0$, we set $c_n = 2^{-n((p/q)-1)}\mu_n^{-1/p}$. It remains to show that $\sum_{n\ge n_0} c_n^r \le KS^r$. We will use the same principle as before: If f is a continuous function on T that satisfies

$$\forall\, w \in T, \qquad \int\left|\frac{f(z+w)-f(z)}{d(w,e)}\right|^p dm(z) \le 1 \tag{5.15}$$

the process given by $X_t(w) = f_t(w)$ satisfies $(*)$, so we have $\operatorname{Sup}_{t,u}(f(t) - f(u)) \le S$. However, in the present case, the appropriate function f to use is *not* a function of the distance to e.

We fix $N \ge s \ge n_0$. Consider a sequence $(t_k)_{k\ge s}$, with $t_s \le \frac{1}{2}$, and a sequence $(a_n)_{n\ge s}$, $a_n \ge 0$. We denote by h_n the piecewise affine function from $\mathbb{R}^+$ to $\mathbb{R}^+$ such that $h(x) = 0$ if $x \ge t_s$, of slope $-a_k$ between t_{k+1} and t_k for $k \le \min(n, N)$, and that is constant between 0 and $t_{\min(n,N)+1}$, so that $h_n = 0$ for $n < s$. We set $b_k = (t_k - t_{k+1})a_k$. For z in T, $z = (x, u)$, $x \in U^q$, $u \in H$, we define $f(z) = h_n(d_1(0, x))$ when $2^{-n+1} < d_2(e', u) \le 2^{-n}$. We note that $\operatorname{Inf}\{f = 0$, and that $f(e) = \sum_{s\le n\le N} b_n$. To prove (5.15), we will reduce to the cases $w \in \{0\} \times H$ or $w \in U^q \times \{e'\}$.

First step. Set $\theta = 2^{-r(1-1/p)}$, so $\theta < 1$. We prove the existence of a number $\alpha > 0$, depending only on p, q, such that if $t_{k+1} \le \theta t_k$ for $k \ge s$, and if

$$\forall\, i \ge s, \qquad \left(2^i t_i a_i\right)^p \mu_i t_i^q \le \alpha, \tag{5.16}$$

then, whenever $w = (0, u) \in \{0\} \times H$ we have

$$I(u) = \int\left|\frac{f(z+w)-f(z)}{d(w,e)}\right|^p dm(z) \le 2^{-p}. \tag{5.17}$$

We define k by $2^{-k-1} < d_2(e', u) \le 2^{-k}$. Let $z = (x, v)$. Define n and n' by

$$2^{-n-1} < d_2(e', v) \le 2^{-n}, \qquad 2^{-n'-1} < d_2(e', v+u) \le 2^{-n'},$$

so we have $f(z) = h_n(d_1(0, x))$; $f(z + w) = h_{n'}(d_1(0, x))$. We have $f(z) = f(z + w)$ if $n = n'$. By ultrametricity, we have

$$d_2(e', v) = d_2(v, u + v) \le \max(d_2(e', v), d_2(e', u + v)),$$
$$d_2(e', v) \le \max(d_2(e', u + v), d_2(v, u + v)),$$
$$d_2(e', u + v) \le \max(d_2(e', u), d_2(u, u + v)),$$

so we can have $n \ne n'$ only if $n = k$, $n' > k$ or $n' = k$, $n > k$, in which case we have

$$|f(z) - f(z + w)| = |h_\tau(d_2(e', x)) - h_k(d_2(e', x))|$$

for $\tau = \max(n, n')$. It follows that for $n \ge k + 1$, if we have $|f(z) - f(z + w)| >$

$\sum_{k<i<n} b_i$, then $d_1(0, x) \le t_n$ and either $d_2(e', v) \le 2^{-n}$ or $d_2(e', v + u) \le 2^{-n}$. This shows that

$$m\left(\left\{z; |f(z) - f(z + w)| > \sum_{k<i<n} b_i\right\}\right) \le 2^{q+1}\mu_n t_n^q,$$

where if $n = k + 1$, we set $\sum_{k<i<n} b_i = 0$.

For a function g, and an increasing sequence $(d_i)_{i\ge 0}$ with $d_0 = 0$, we have $\int g^p\, dm \le \sum_{i\ge 0} d_{i+1}^p m(\{g > d_i\})$; so we have, using the convexity of $t \to t^p$,

$$\begin{aligned} I(u) &\le \sum_{n>k} 2^{q+1}\mu_n t_n^q \Big(2^{k+1} \sum_{k<i\le n} b_i\Big)^p \\ &\le \sum_{n>k} 2^{q+1+2p}\mu_n t_n^q \Big(\sum_{k<i\le n} 2^{k-i-1} 2^i b_i\Big)^p \\ &\le \sum_{n>k} 2^{q+1+2p}\mu_n t_n^q \Big(\sum_{k<i\le n} 2^{k-i-1}\big(2^i b_i\big)^p\Big). \end{aligned}$$

Since $t_{n+1} \le \theta t_n$, we have $\sum_{n\ge i} \mu_n t_n^q \le [1/(1 - \theta^q)]\mu_i t_i^q$, so we find

$$I(u) \le \frac{1}{1 - \theta^q} 2^{q+1+2p} \sum_{i>k} 2^{k-i}\big(2^i a_i t_i\big)^p \mu_i t_i^q$$

and this proves the claim.

Second step. For $z = (x, v)$, we set

$$M(z) = \limsup_{y\to x} \left|\frac{f((x, v)) - f((y, v))}{d(y, v)}\right|.$$

Then for some number β depending only on p and q, if we have

$$\sum_{s\le n\le N} a_n^p t_n^q \mu_n \le \beta, \tag{5.18}$$

we also have $\int_T M(z)^p\, dm(z) \le 2^{-p}$.

To prove this, we note that we have $|M(z)| \le \max_{i<n} a_i$ unless $d_1(e, x) \le t_n$ and $d_2(e, v) \le 2^{-n}$; so we have $m(\{z; M(z) > \max_{i<n} a_i\}) \le 2^q t_n^q \mu_n$. This easily implies the result.

Third step. We show that (5.16) and (5.18) imply (5.15). We set $w = (y, u)$, so $d(e, w) = \max(d_1(0, y), d_2(e', u))$. Let $w_1 = (y, e')$, $w_2 = (0, u)$. We have

$$|f(z + w) - f(z)| \le |f(z + w) - f(z + w_1)| + |f(z + w_1) - f(z)|.$$

Using the inequality $|a + b|^p \le 2^{p-1}(a^p + b^p)$, we have

$$\int \left|\frac{f(z + w) - f(z)}{d(e, w)}\right|^p dm(z) \le 2^{p-1}[I(u) + I],$$

where $I = \int |[f(z + w_1) - f(z)]/d_1(0, y)|^p \, dm(z)$. Write $z = (x, v)$. Writing $t \to x + ty$ for the natural path from x to $x + y$, we have

$$|f(z + w_1) - f(z)| \le d_1(0, y) \int_0^1 M((x + ty, v)) \, dt.$$

Since $\int_T M(z)^p \, dm(z) \le 2^{-p}$, the convexity of $t \to t^p$, integration and Fubini's theorem show that $I \le 2^{-p}$, and this completes the proof of this step.

Fourth step. For $n \ge n_0$, we recall that $c_n = 2^{n(q/p-1)} \mu_n^{-1/p}$. We note that $c_n \le 2^{1-q/p} c_{n+1}$. We define n_1, n_2, τ as the smallest integers for which, respectively,

$$S 2^{-n_1(p-1)} \mu_{n_1}^{1/p} \le 1, \qquad \beta^{-1/2} 2^{-n_2} \le \tfrac{1}{2}, \tag{5.19}$$
$$\left(\beta^{-1/r} 2^{1/r} 2^{-\tau(p-1)}\right)^{rp} \le \tfrac{1}{2}.$$

By (5.14) we have $c_n \le K 2^{nq/p} S$. It follows that $\sum_{n_0 \le n \le n_2} c_n^r \le K S^r$. If $n < n_1$, we have $\mu_n^{-1/p} \le S 2^{-np+n}$, so $c_n \le S 2^{-np(1-q/p^2)} = S 2^{-np/r}$. This implies that $\sum_{n_0 \le n < n_1} c_n^r \le K S^r$. Also, $\sum_{n_1 \le n \le n_1 + \tau} c_n^r \le K c_{n_1+\tau}^r$. Let $n_4 = \max(n_1 + \tau, n_2)$. We are left to show that $\sum_{n \ge n_4} c_n^r \le K S^r$.

In the preceding steps, we have shown the following. If $t_{n+1} \le \theta t_n$, $t_s \le \frac{1}{2}$, then conditions (5.16) and (5.18) imply that $f(e') = \sum_{s \le n \le N} a_n (t_n - t_{n+1}) \le S$, so $\sum_{s \le n \le N} a_n t_n \le S/(1 - \theta)$. Let $\gamma > 0$. We set $t_n = 2^{-n} \gamma^{-p/q} c_n^{r/p}$ so we always have $t_{n+1} \le \theta t_n$. Algebra shows that $t_n = K 2^{-nr(1-1/p)} \mu_n^{r/p^2}$. We set $a_n = \alpha^{1/p} \gamma c_n^r / t_n$. More algebra shows that (5.15) holds, while (5.18) is equivalent to $\gamma^{p^2/q} \sum_{s \le n \le N} c_n^r \le \beta$, so we choose $\gamma = (\beta \sum_{s \le n \le N} c_n^r)^{-q/p^2}$ and we have

$$t_s = \left[\beta^{-1/r} \left(\sum_{s \le n \le N} c_n^r\right)^{1/r} 2^{-s(p-1)} \mu_s^{1/p}\right]^{r/p}. \tag{5.20}$$

We now know that $t_s \le \frac{1}{2}$ implies $\sum_{s \le n \le N} a_n t_n \le S/(1 - \theta)$, so

$$\left(\sum_{s \le n \le N} c_n^r\right)^{1/r} \le K_1 S$$

for some K_1 depending only on p and q. For $s \ge n_4$, take $N = s$, so $t_s = \beta^{-1/p} 2^{-s} \le \frac{1}{2}$; this shows that $c_s \le K_1 S$. We show that $(\sum_{n \ge n_4} c_n^r)^{1/r} \le K_1 S$ (which finishes the proof). Otherwise let N be the smallest such that $\sum_{n_4 \le n \le N} c_n^r > K_1^r S^r$, and let $s = n_4$. Since $c_N^r \le K_1^r S^r$, we have $\sum_{n_4 \le n \le N} c_n^r \le 2 K_1^r$. It follows that from (5.19), (5.20), the choice of n_1 and the fact that $n_4 \ge n_1$ that $t_s \le \frac{1}{2}$, so that $(\sum_{n_4 \le n \le N} c_n^r)^{1/r} \le K_1 S$. This contradiction finishes the proof. □

Proof of the left inequality. For the same reasons as in Theorem 5.6, we will proceed directly instead of using Proposition 2.5. We define the sequence $n(i)$ by induction as follows. We set $n(1) = n_0$,

$$n(i) = \inf\{n > n(i - 1), B(e', 2^{-n(i-1)}) \ne B(e', 2^{-n})\}.$$

(We assume H to be infinite for definiteness, but the case where H is finite is similar.)

We set $L_i = B(e', 2^{-n(i)}) \setminus B(e', 2^{-n(i+1)})$; we note that $m_2(L_i) \geq \mu_{n(i)}/2$ and that $d(e', u) \leq 2^{-n(i)+1}$ whenever $u \in L_{i-1}$. We denote by λ_i the measure given by $m_2(L_i)\lambda_i(A) = m_2(L_i \cap A)$. We define $\nu_i \in P(T)$ by $\nu_i = \delta_0 \otimes \lambda_i$. We set $A_1 = (\sum_{n \geq n_0} c_n^r)^{1/r}$; $A_2 = \sum_{n \geq n_0} 2^{-n}\mu_n^{-1/p}$. For $i \geq 1$, set $d_i = c_{n(i)}^r/2A_i^r + 2^{-n(i)}\mu_{n(i)}^{-1/p}/2A_2$; hence $\sum_{i \geq 1} d_i \leq 1$. We set $\nu = \sum_{i \geq 1} d_i \nu_i$. Consider a process $(X_t)_{t \in T}$ that satisfies (*). To prove a bound on $\operatorname{Sup}_{t,u}(X_t - X_u)$, we can, as in the proof of Theorem 2.3, assume that Ω is finite; using convolution, we can assume that for ω in r, and u in H, the map $x \to X_{(x,u)}(\omega)$ is differentiable on U^q; we denote by $D(x, u, \omega)(v)$ the value of this differential on the vector v of $\mathbb{R}^q$. Define $\|v\| = \max_{i \leq q} |v_i|$. Using (*) and Fubini's theorem, we see that

$$E\int\int_{u \neq e'} \left| \frac{X_{z+(0,u)} - X_z}{d_2(e', u)} \right|^p dm(z)\, d\nu(u) \leq 1,$$

and that if $\|v\| = 1$, we have

$$E\int |D(x, u, \omega)(v)|^p \, dm((x, u)) \leq 1.$$

For $z = (x, u)$, write $|\nabla X(z)|(\omega) = \operatorname{Sup}_{\|v\|=1} |D(x, u, \omega)(v)|$; we have

$$E\int \left| \frac{1}{q} \nabla X(z) \right|^p dm(z) \leq 1.$$

We have now reduced the problem to the following: Suppose we have a Lipschitz function f on G. We set

$$I_1 = \int\int \left| \frac{f(z + (0, u)) - f(z)}{d_2(u, e')} \right|^p dm(t)\, d\nu(u), \tag{5.21}$$

$$I_2 = \int |\nabla f(z)|^p \, dm(z), \tag{5.22}$$

and then we have $\operatorname{Sup}_{t,u}(f(t) - f(u)) \leq K(I_1^{1/p} + I_2^{1/p})(A_1 + A_2)$. Since these conditions over f are invariant by translation, it is enough to show that $|f(e) - \int f\, dm| \leq K(I_1^{1/p} + I_2^{1/p})(A_1 + A_2)$. For $n \geq n_0$, we define $t_n = \min(\frac{1}{2}, 2^{-n} c_n^{-r/p} A_1^{r/p})$. For $i \geq 1$ and y in U^q, we set

$$f_i(y) = \textstyle\int_H f(y, u)\, d\lambda_i(u).$$

From (5.21), we see that if we set

$$J_i = \frac{1}{m_2(L_{i-1})} \int\int\int_{v \in L_i;\, u \in L_{i-1}} \left| \frac{f((y, v + u)) - f((y, v))}{d_2(u, e')} \right|^p dm_1(y)\, dm_2(u)\, dm_2(v),$$

we have

$$\sum_{i\geq 2}\left(\frac{c^r_{n(i)}}{2A_1^r}+\frac{2^{-n(i)}\mu_{n(i)}^{-1/p}}{2A_2}\right)J_i \leq I_1. \tag{5.23}$$

We note that for $v \in L_i$, λ_{i-1} is invariant under translation by v. So from (5.23) we see that [since $d(u, e') \leq 2^{-n(i)+1}$, $m_2(L_i) \geq \mu_{n(i)}/2$]

$$\sum_{i\geq 2} c^r_{n(i)}\mu_{n(i)}\int\left|\frac{f_i(y)-f_{i-1}(y)^p}{2^{-n(i)+1}}\right|dm_1(y) \leq 4A_1^rI_1. \tag{5.24}$$

From (5.22), we have

$$\sum_{i\geq 1}\mu_{n(i)}\int|\nabla f_i(z)|^p\,dm(z) \leq 2I_2. \tag{5.25}$$

We now denote by α_i (resp. β_i) the average of f_i (resp., f_{i-1}) over the ball (in U^q) of center 0 and radius $t_{n(i)}$. By (5.24), we have

$$\sum_{i\geq 2} 2^{pn(i)}c^r_{n(i)}\mu_{n(i)}t^q_{n(i)}|\alpha_i-\beta_i|^p \leq KA_1^rI_1. \tag{5.26}$$

We define p' by $1/p + 1/p' = 1$. Denote by j the first integer greater than or equal to 2 for which $t_{n(j)} \leq \frac{1}{2}$. Algebra shows that for $i \geq j$, we have

$$\left(2^{n(i)}c^{r/p}_{n(i)}\mu^{1/p}_{n(i)}t^{q/p}_{n(i)}\right)^{-p'} = A_1^{-rq/p^2}c^r_{n(i)},$$

so, by Holder's inequality, we have from (5.26) that

$$\sum_{i\geq j}|\alpha_i-\beta_i| \leq KI_1^{1/p}A_1. \tag{5.27}$$

From (5.23) it follows that

$$\sum_{i\geq 2} 2^{(p-1)n(i)}\mu^{1/p'}_{n(i)}\int|f_i(y)-f_{i-1}(y)|^p\,dm_1(y) \leq KA_2I_1.$$

If $j > 2$, it follows that

$$\sum_{2\leq i<j} 2^{(p-1)n(i)}\mu^{1/p'}_{n(i)}|\alpha_i-\beta_i|^p \leq KA_2I_1,$$

so, by Holder's inequality, we have

$$\sum_{2\leq i\leq j}|\alpha_i-\beta_i| \leq \left(\sum_{i<j}2^{-n(i)}\mu^{-1/p}_{n(i)}\right)^{1/p'}(KA_2I_1)^{1/p} \leq KA_2I_1^{1/p}.$$

For $r \leq \frac{1}{2}$, we denote by $a_i(r)$ [resp., $b_i(r)$] the average of f_i (resp. $|\nabla f_i|$) on the sphere of center 0 and radius r. As in the proof of Theorem 5.1, we have $a'(r) \leq b(r)$. Let $Q_i = \operatorname{Sup}_{r, r'\leq t_i}(a(r) - a(r'))$. Then we have $|\alpha_i - \beta_{i+1}| \leq Q_i$.

We have, using Holder's inequality,

$$Q_i \le \int_0^{t_{n(i)}} |a_i'(r)|\,dr \le \int_0^{t_{n(i)}} b(r)\,dr \le \int_0^{t_{n(i)}} r^{-(q-1)/p}\big(r^{(q-1)/p} b(r)\big)\,dr$$

$$\le K t_{n(i)}^{1-q/p}\left(\int_0^{t_{n(i)}} r^{q-1} b^p(r)\,dr\right)^{1/p}.$$

Now $\int_0^{1/2} r^{q-1} b^p(r)\,dr \le K\int |\nabla f_i(z)|^p\,dm_1(z)$, so we have

$$|\alpha_i - \beta_{i+1}| \le t_{n(i)}^{1-q/p}\mu_{n(i)}^{-1/p}\left(\mu_{n(i)}\int |\nabla f_i(z)|^p\,dm(z)\right)^{1/p}.$$

Algebra shows that

$$t_i^{p'(1-q/p)}\mu_{n(i)}^{-p'/p} = A_1^{(rp'/p)(1-q/p)} c_{n(i)}^r,$$

so we have by Holder's inequality

$$\sum_{i\ge 1} |\alpha_i - \beta_{i+1}| \le K A_1^{(r/p)(1-q/p)} A_1^{r/p'} I_2^{1/p} = K A_1 I_2^{1/p}.$$

We now have

$$\sum_{i\ge 1} |\alpha_i - \alpha_{i+1}| \le \sum_{i\ge 1} |\alpha_{i+1} - \beta_{i+1}| + \sum_{i\ge 1} |\alpha_i - \beta_{i+1}| \le K(A_1 + A_2)\big(I_1^{1/p} + I_2^{1/p}\big).$$

This shows that $|f(e) - \alpha_1| \le K(A_1 + A_2)(I_1^{1/p} + I_2^{1/p})$. If $t_1 \ge \frac{1}{2}$, we have $\alpha_1 = \int_T f\,dm$, and we are done. Otherwise, let $g(x)$ be the average of f on $B(x, t_1) \times H$, so $\alpha_1 = g(e)$. The proof of Theorem 5.1 and (5.22) show that

$$\left|\alpha_1 - \int g(x)\,dm_1(x)\right| \le K\left(\int |\nabla g(x)|^p\,dm_1(x)\right)^{1/p} \le K I_2^{1/p}.$$

Since $\int g(x)\,dm_1(x) = \int_T f\,dm$, this finishes the proof. □

In the three cases we have studied in this section, it is possible, by a slight modification of our arguments, to see that condition (3.5) holds whenever $S < \infty$.

PROBLEM 5.5. Let T be a compact group, d a translation invariant distance on T, Φ an Orlicz function. Assume that $S(T, d, \Phi) < \infty$. Does it follow that for every $\alpha > 0$, there is $\delta > 0$ such that each process which satisfies (*) satisfies $E\,\mathrm{Sup}_{d(t,u)<\delta}(X_t - X_u) \le \alpha$?

A positive answer would be a generalization of the fact that bounded stationary Gaussian processes are continuous. [For a Gaussian process, the distance d on T given by $d(t, u) = \|X_t - X_u\|_2$ satisfies $S(T, d, \Phi_2) < \infty$, where $\Phi_2(x) = (\exp x^2) - 1$.]

6. Processes with values in exponential-type Orlicz spaces. In this section, we no longer consider general Orlicz norms, but only the norms $\|\cdot\|_{\Phi_q}$ associated to the functions $\Phi_q(x) = (\exp|x|^q) - 1$ for $q > 0$. [If $q < 1$, $\Phi_q(x)$ is not convex for $|x| < h_q = ((1-q)/q)^{1/q}$, so Φ_q is not a Young function; it is however convex for $x > h_q$, and this is what really matters; one can for example replace $\Phi_q(n)$ by $\exp(x + h_q)^q - \exp h_q^q$.] We fix $q' > q$. Consider a compact metric space T and μ in $P(T)$.

PROBLEM 6.1. When is it true that all processes $(X_t)_{t\in T}$ that satisfy

$$\forall\, t, u \in T, \qquad \|X_t - X_u\|_{\Phi_q} \le d(t,u)$$

and that are $\mu \otimes P$ measurable have their trajectories a.e. in $L_{\Phi_{q'}}(\mu)$?

PROBLEM 6.2. When does this occur for all μ in $P(T)$?

These problems have been investigated by Marcus and Pisier [11], and by Weber [16]. Problem 6.1 is by far the hardest; it is not clear whether it has a completely satisfactory answer in general spaces, but we will solve it when T is ultrametric. We will solve Problem 6.2 in general.

THEOREM 6.3. *For a compact ultrametric space T, and μ in $P(T)$, the following are equivalent.*

(a) *For each process $(X_t)_{t\in T}$ that satisfies Problem 6.1, and such that $X_t(\omega)$ is $\mu \otimes P$ measurable, the trajectories $t \to X_t(\omega)$ belong to $L_{\Phi_{q'}}(\mu)$ a.e.*

(b) *There is $m \in P(T)$ and a constant A such that*

$$\forall x \in T, \forall n,$$

$$\sum_{k\le N} 2^{-k}\left(\log\frac{1}{m(B(x,2^{-k}))}\right)^{1/q} \le A\left(\log\frac{2}{\mu(B(x,2^{-N}))}\right)^{1/q'}. \tag{6.1}$$

PROOF. To understand better (6.1), we note that $\log[1/m(B(x,2^{-k}))] = 0$ if $2^{-k} \ge D(T)$. Many of the arguments are modifications of arguments ven; we will not repeat them, and we will leave the checking to the reader.

(a) $\Rightarrow$ (b). We fix v in T. For a function f on T, we denote by $M(f)$ the $L_{\Phi_{q'}}(\mu)$ norm of the function $t \to f(t) - f(v)$, so $M(X.)$ is the random variable equal at ω to $M(f)$, where $f(t) = X_t(\omega)$. The first step is to prove, as in Theorem 2.3, that for some number S, every process that satisfies the condition in Problem 6.1 satisfies $E(M(X.)) \le S$. The second step is to find, as in Theorem 2.3, a linear functional θ on the set G of continuous bounded functions on $T \times T \setminus \Delta$ such that whenever f is Lipschitz on T and

$$\theta\left(\Phi_q\left(\frac{f(t)-f(u)}{d(t,u)}\right)\right) \le 1, \tag{6.2}$$

we have $M(f) \le S$. For $x \in T$, denote by $n(x)$ the largest integer such that $2^{-n(x)} \ge d(x,v)$. For a sequence (a_n), consider the function f given by $f(t) = a_n$ if $2^{-n-1} < d(t,x) \le 2^{-n}$, where $a_n = a_N$ if $n \ge N$ and $a_n = 0$ if $n < n(x)$. If (6.2) holds, we have $M(f) \le S$, so, in particular, $\mu(B(x,2^{-N})) \times (\exp(a_N/S)^{q'} - 1) \le 1$, and $a_N \le S(\log(1 + [1/K\mu(B(x,2^{-N}))]))^{1/q'}$. As in the proof of Theorem 1.3, we construct from θ a probability m on T that satisfies for all x,

$$\sum_{n(x)\le n\le N} 2^{-n}\left(\log\frac{1}{m(B(x,2^{-n}))}\right)^{1/q}$$
(6.3)
$$\le KS\left(\log\left(1+\frac{1}{\mu(B(x,2^{-N}))}\right)\right)^{1/q'}.$$

If we replace m by $\frac{1}{2}(m+\delta_v)$, we then get (6.1).

(b) $\Rightarrow$ (a). We denote by s the largest integer for which $2^{-s}\ge D(T)$. For $n\ge s$, we denote by $\mathscr{B}_n$ the family of balls of T of radius 2^{-n}. We fix v in T. To each B in $\mathscr{B}_n$, we associate one of its points $x(B)$, in such a way that $x(B)=v$ if $v\in B$. For $n>s$, we denote by B' the unique element of $\mathscr{B}_{n-1}$ that contains B. Replacing m by $\frac{1}{2}(m+\mu)$, we can suppose that $2m\ge\mu$. We consider the probability ν on $T\times T$ given by

$$\nu=\sum_{\substack{n>s\\ B\in\mathscr{B}_n}} 2^{s-n}m(B)\delta_{x(B)}\otimes\delta_{x(B')}.$$

It is enough to show that if f is a Lipschitz function on $T\times T$ such that $\iint\Phi_q([f(t)-f(u)]/d(t,u))\,d\nu(t,u)\le 1$, we have $M(f)\le K$. (Here and in what follows, K is a constant depending only on A, q, q', s, S, but can vary from line to line.) We note that

$$\sum_{\substack{n>s\\ B\in\mathscr{B}_n}} 2^{-n}m(B)\Phi_q\left(\frac{f(x(B))-f(x(B'))}{2^{-n+1}}\right)\le 2^{-s}. \tag{6.4}$$

To prove that $M(f)\le KA$, it is enough to show that for any Borel subset C of T, we have

$$\int_C |f(t)-f(v)|\,d\mu(t)\le K\mu(C)\left(1+(\log(1/\mu(C)))^{1/q'}\right). \tag{6.5}$$

We define f^n as being constant on each B in $\mathscr{B}_n$, and equal to $f(x(B))$ on B. We have $|f(t)-f(v)|\le\sum_{n>s}|f^n(t)-f^{n-1}(t)|$. For B in $\mathscr{B}_n$, we have

$$\int_{C\cap B}|f^n-f^{n-1}|\,d\mu=\mu(C\cap B)|f(x(B))-f(x(B'))|. \tag{6.6}$$

We note the following inequality, which is more convenient here than Young's inequality. For some $\gamma>0$, and all $a, b>0$, we have

$$ab\le\gamma\left(\exp(2a)^q-1\right)+b(\log^+ b)^{1/q}.$$

This is trivial by distinguishing whether or not $b\le\exp a^q$. So, for $u, a, b>0$, we have

$$u\le\gamma b\Phi_q\left(\frac{2u}{a}\right)+a\left(\log^+\frac{a}{b}\right)^{1/q}.$$

We use this for $u = |f(x(B)) - f(x(B'))|$, $a = 2^{-n+2}$, $b = 2^{-n+2}\mu(C)m(B)/\mu(C \cap B)$ and (6.6) yields

$$\int_{C\cap B} |f^n - f^{n-1}|\, d\mu \le 2^{-n+2}\gamma\mu(C)m(B)\Phi_q\left(\frac{f(x(B)) - f(x(B'))}{2^{-n+1}}\right) + 2^{-n+2}\mu(C \cap B)\left(\log^+\left(\frac{\mu(C\cap B)}{\mu(C)m(B)}\right)\right)^{1/q}.$$

We rewrite this as

$$\int_{C\cap B} |f^n + f^{n-1}|\, d\mu \le \int_C 2^{-n+2}\gamma m(B)\Phi_q\left(\frac{f(x(B)) - f(x(B'))}{2^{-n+1}}\right) d\mu + \int_{C\cap B} 2^{-n+2}\left(\log^+\left(\frac{\mu(C\cap B)}{\mu(C)m(B)}\right)\right)^{1/q}. \tag{6.7}$$

We need careful bounds of the last term. We denote by $\mathscr{H}$ the family of balls B in $\bigcup_{n\ge s}\mathscr{B}_n$ such that $\mu(B) \le \mu^2(C)$, and by N the largest integer such that $2^{-N}(\log[1/\mu(C)])^{1/q-1/q'} \ge 2^{1/q}A$. For $n > N$, we have, since $\mu \le 2m$,

$$2^{-n+2}\left(\log^+\left(\frac{\mu(C\cap B)}{\mu(C)m(B)}\right)\right)^{1/q} \le 2^{N-n+3}2^{-N-1}\log^+\left(\frac{2}{\mu(C)}\right)^{1/q} \le 2^{N-n+3}K\left(1 + \left(\log\frac{1}{\mu(C)}\right)^{1/q'}\right). \tag{6.8}$$

For $n < N$, if $B \in \mathscr{H}$, we have

$$\left(\log\frac{1}{m(B)}\right)^{1/q} \le 2^n A\left(\log\frac{2}{\mu(B)}\right)^{1/q'}, \qquad 2^n A \le 2^{-1/q}\left(\log\frac{2}{\mu(C)}\right)^{1/q-1/q'},$$

so we have

$$\left(\log\frac{1}{m(B)}\right)^{1/q} \le 2^{-1/q}\left(\log\frac{2}{\mu(C)}\right)^{1/q-1/q'}\left(\log\frac{2}{\mu(B)}\right)^{1/q'} \le 2^{-1/q}\left(\log\frac{2}{\mu(B)}\right)^{1/q},$$

since $\mu(C) \ge \mu^2(C) \ge \mu(B)$. This implies $\mu(B) \le 2m(B)^2$. Since $\mu(B) \le \mu(C)^2$, we have $\mu(B) \le 2m(B)\mu(C)$, so $\mu(B \cap C) \le 2m(B)\mu(C)$; it follows that

$$2^{-n+2}\left(\log^+\left(\frac{\mu(B\cap C)}{m(B)\mu(C)}\right)\right)^{1/q} \le K2^{-n}. \tag{6.9}$$

Finally, if $B \notin \mathscr{H}$, $n < N$, we use

$$2^{-n+2}\left(\log^+\left(\frac{\mu(B \cap C)}{m(B)\mu(C)}\right)\right)^{1/q} \leq 2^{-n+2}\left(\log\frac{1}{m(B)}\right)^{1/q}. \tag{6.10}$$

For x in T, if $n(x)$ is the largest such $B(x, 2^{-n}) \notin \mathscr{H}$, we note that by (6.1)

$$\begin{aligned}\sum_{n \leq n(x)} 2^{-n}\left(\log\frac{1}{m(B(x,2^{-n}))}\right)^{1/q} &\leq A\left(\log\frac{2}{\mu(B(x,2^{-n(x)}))}\right)^{1/q'} \\ &\leq A\left(\log\frac{2}{\mu(C)^2}\right)^{1/q'}.\end{aligned} \tag{6.11}$$

Summation of the inequalities (6.7) over all the balls in $\bigcup_{n \geq s} \mathscr{B}_n$, together with (6.4) and (6.8) to (6.11) imply (6.5). This concludes the proof. □

THEOREM 6.4. *For a compact metric space, the following are equivalent.*

(a) *For each process $(X_t)_{t \in T}$ that satisfies the condition in Problem 6.1 and each $\mu \in P(T)$ such that $X_t(\omega)$ is $\mu \otimes P$ measurable, the trajectories $t \to X_t(\omega)$ belong to $L_{\Phi_{q'}}(\mu)$ a.s.*

(b) *The covering numbers $N(\varepsilon) = N(T, d, \varepsilon)$ satisfy*

$$\operatorname*{Sup}_{\varepsilon>0} \varepsilon(\log N(T, d, \varepsilon))^{1/q - 1/q'} < \infty.$$

PROOF. (a) ⇒ (b) is due to Marcus and Pisier [11]; we prove it for completeness. If $M_\mu(X.)$ denotes the $L_{\Phi_{q'}}(\mu)$ norm of $t \to X_t(\omega) - X_v(\omega)$ where v is fixed, one first proves that for some A, we have $EM_\mu(X.) \leq A$ for all μ in $P(T)$. Then consider a family $(x_i)_{i \leq N}$ of points of T such that $d(x_i, x_j) \geq 2\varepsilon$ if $i \neq j$. Let $\mu = (1/N)\sum_{i \leq N} \delta_{x_i}$. For ω in T, set $X_t(\omega) = \max(0, 1 - d(t, \omega)/\varepsilon)$, and take (T, μ) as basic probability space. It is easy to see that the process αX_t, for $\alpha = \varepsilon(\log(N/2))^{1/q}$, satisfies the condition in Problem 6.1. All the points x_i, except at most one, are at distance greater than or equal to ε from v. For such a point x_i, if $\beta = M_\mu(X(x_i))$, we have $(1/N)(\exp(\alpha/\beta)^{q'} - 1) \leq 1$, so $\beta \geq \alpha(\log(N+1))^{-1/q'}$. Since $A \geq EM_\mu(X.) \geq [(N-1)/N]\beta$, we have $2A \geq \varepsilon(\log(N/2))^{1/q}(\log(N+1))^{-1/q'}$, which easily implies (b).

(b) ⇒ (a). Let $N_n = N(2^{-n})$. Denote by s the largest integer with $2^{-s} \geq D(T)$. Let $Q_n = \prod_{s \leq k \leq n} N_k$. We can clearly find an increasing sequence $(\mathscr{H}_n)$ of finite partitions of T, such that card $\mathscr{H}_n \leq Q_n$, each set of $\mathscr{H}_n$ has diameter less than or equal to 2^{-n+1} and is Borel. The partitions $\mathscr{H}_n$ generate the Borel σ-algebra of T. For each $H \in \mathscr{H}_n$, we pick $x(H) \in H$; we set $v = x(T)$. For H in $\mathscr{H}_n$, $n > s$, we denote by H' the unique elment of $\mathscr{H}_{n-1}$ that contains H. We consider the subprobability on $T \times T$ given by

$$\nu = \sum_{\substack{n>s \\ H \in \mathscr{H}_n}} \frac{2^{s-n}}{Q_n} \delta_{x(H)} \otimes \delta_{x(H')}.$$

Let μ be in $P(T)$. The point is as usual to prove that if f is a Lipschitz function on T that satisfies

$$\int_{T\times T}\Phi\left(\frac{f(t)-f(u)}{d(t,u)}\right)d\nu(t,u)\le 1,$$

then $M_\mu(f)\le K$. To that purpose, we prove (6.5). Denote by f^n the function given by $f^n(t)=f^n(x(H))$ if $t\in H\in\mathscr{H}_n$, so again $f(t)-f(v)=\sum_{n>s}(f^n(t)-f^{n-1}(t))$. For $B\in\mathscr{H}_n$, set $m(B)=\frac{1}{2}(\mu(B)+2^{s-n}/Q_n)$. As before, (6.7) holds.

Denote by N the largest integer with $2^{-N}\ge(\log^+[2/\mu(C)])^{1/q'-1/q}$. For $n>N$, we write

$$2^{-n}\left(\log^+\left(\frac{\mu(B\cap C)}{m(B)\mu(C)}\right)\right)^{1/q}\le 2^{-n}\log^+\left(\frac{2}{\mu(C)}\right)^{1/q}$$

$$\le 2^{N-n+1}\left(\log\left(\frac{2}{\mu(C)}\right)\right)^{1/q'}.$$

For $n\le N$, we have, since $\mu(B\cap C)\le\mu(C)$,

$$2^{-n}\left(\log^+\left(\frac{\mu(B\cap C)}{m(B)\mu(C)}\right)\right)^{1/q}\le 2^{-n}\left(\log^+\left(2^{n-s+1}Q_n\right)\right)^{1/q}.$$

A simple computation using (b) shows that

$$\sum_{n\le N}2^{-n}\left(\log\left(2^{n-s+1}Q_n\right)\right)^{1/q}\le K\left(\log\frac{2}{\mu(C)}\right)^{1/q'}$$

and (6.5) follows from summation of the relations (6.7) over all $B\in\bigcup_{n\ge s}\mathscr{H}_n$. The proof is complete. □

REFERENCES

[1] ASSOUAD, P. (1969). Le mouvement brownien sur [0, 1]; applications Φ-sommantes et (Φ − Ψ)-sommantes. *Séminaire Laurent Schwartz 1969–1970*. Ecole Polytechnique, Palaiseau.

[2] DIESTEL, J. and UHL, J. S. (1977). *Vector Measures. Math. Surveys Monogr.* **15**. Amer. Math. Soc., Providence, R.I.

[3] DUDLEY, R. M. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Funct. Anal.* **1** 290–330.

[4] FERNIQUE, X. (1978). Caractérisation de processus à trajectoires majorées ou continues. *Séminaire de Probabilités XII. Lecture Notes in Math.* **649** 691–706. Springer, Berlin.

[5] FERNIQUE, X. (1983). Régularité de fonctions aléatoires non gaussiennes. *Ecole d'Eté de Probabilités de Saint-Flour XI-1981. Lecture Notes in Math.* **976** 1–74. Springer, Berlin.

[6] IBRAGIMOV, I. (1973). Properties of sample functions for stochastic processes and embedding theorems. *Theory Probab. Appl.* **18** 442–453.

[7] IBRAGIMOV, I. (1979). Sur la regularité des fonctions aléatoires. *C. R. Acad. Sci. Paris Sér. A* **289** 545–547.

[8] KLASS, M. J. and HAHN, M. G. (1977). Sample continuity of square integrable process. *Ann. Probab.* **5** 361–377.

[9] KÔNO, N. (1978). Best possibility of an integral test for sample continuity of L_p processes ($p\ge 2$). *Proc. Japan. Acad.* **54** 197–201.

[10] KÔNO, N. (1980). Sample path properties of stochastic processes. *J. Math. Kyoto Univ.* **20** 295–313.

[11] MARCUS, M. and PISIER, G. (1985). Stochastic processes with sample paths in exponential Orlicz spaces. *Probability in Banach Spaces V. Lecture Notes in Math.* **1153** 325–358. Springer, Berlin.

[12] PISIER, G. (1980). Conditions d'entropie assurant la continuité de certains processus. *Séminaire d'Analyse Fonctionnelle 1979–1980. Exposés XIII–XIV.* Ecole Polytechnique, Palaiseau.

[13] PISIER, G. (1980). Some applications of the metric entropy condition to harmonic analysis. *Proc. in Harmonic Analysis (Storrs). Lecture Notes in Math.* **995** 123–154. Springer, Berlin.

[14] TALAGRAND, M. (1987). Regularity of Gaussian processes. *Acta Math.* **159** 99–149.

[15] WEBER, M. (1983). Analyse infinitésimale de fonctions aléatoires. *Ecole d'Eté de Probabilités de Saint-Flour XI-1981. Lecture Notes in Math.* **976** 384–465. Springer, Berlin.

[16] WEBER, M. (1988). Stochastic processes with values in exponential-type Orlicz spaces. *Ann. Probab.* **16** 1365–1371.

[17] ZAANEN, A. C. (1953). *Linear Analysis.* North-Holland, Amsterdam.

EQUIPE D'ANALYSE, TOUR 46
U.A. AU C.N.R.S. N° 754
UNIVERSITÉ PARIS VI
4 PLACE JUSSIEU
75230 PARIS CEDEX 05
FRANCE

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