# Appendix 2

## Liu's new characterization of boundedness of Gaussian processes

In the striking contribution [L25], J. Liu proposes a new information-theoretic characterization of boundedness of Gaussian processes, with in particular a lifting argument reducing to stationary Gaussian processes for which the classical metric entropy numbers produce a full and simple description by the Dudley-Fernique theorem. This appendix briefly describes the principle and the development of this new perspective. The exposition actually carefully follows the note [vH25] by R. van Handel (which provides in addition an improved comparison principle for subGaussian processes). All the notation are taken from Chapter 6 (reproduced below) and van Handel's note. Full details are of course provided by Liu's original contribution [L25].

Roughly speaking, the main results on boundedness of Gaussian processes presented in Chapter 6 express that for a centered Gaussian process $X = (X_t)_{t \in T}$ indexed by some (finite) parameter set $T$ equipped with the intrinsic (pseudo-) metric $d(s,t) = (\mathbb{E}|X_s - X_t|^2)^{1/2}$, $s, t \in T$, the Dudley entropy bound holds true

$$\mathbb{E}\big(\sup_{t \in T} X_t\big) \lesssim \int_0^\infty \big(\log N(T, d; \varepsilon)\big)^{1/2} \, d\varepsilon \tag{1}$$

up to a numerical constant, and turns into an equivalence

$$\mathbb{E}\big(\sup_{t \in T} X_t\big) \sim \int_0^\infty \big(\log N(T, d; \varepsilon)\big)^{1/2} \, d\varepsilon \tag{2}$$

for stationary Gaussian processes as shown by X. Fernique. Short and simple proofs of these results are exposed in Chapter 6.

Since the metric entropy numbers $N(T, d; \varepsilon)$, $\varepsilon > 0$, are not enough to characterize boundedness of arbitrary Gaussian processes, M. Talagrand [Ta87] showed that the more advanced technology of majorizing measures (or families of weights) introduced by X. Fernique may be used to provide a characterization (conjectured by X. Fernique) in the form

$$\mathbb{E}\big(\sup_{t \in T} X_t\big) \sim \inf_m \sup_{t \in T} \int_0^\infty \bigg(\log \frac{1}{m(B(t, \varepsilon))}\bigg)^{1/2} d\varepsilon \;=\; \gamma_2(T, d) \tag{3}$$

up to numerical constants, where the infimum is running over all probability measures $m$ on $T$. In a modern exposition, majorizing measures may be replaced by admissible partitions relying entirely on the metric structure of $(T,d)$ (cf. Appendix 1).

The proof of (3) by M. Talagrand extracts an ultrametric (tree) structure from $(T,d)$ supporting enough information on the boundedness of the Gaussian process (cf. [Ta87,Ta14]). The main new output by J. Liu [L25] is that actually it may be possible to reduce the study of general Gaussian processes to stationary ones for which the easier Dudley-Fernique characterization may be used.

The discussion is developed in the setting of a finite index set $T$. All the conclusions extend, more or less routinely, to countable parameter sets, and more generally to separable processes (cf. [L25]).

If $X = (X_t)_{t\in T}$ is a stochastic process on some probability space $(\Omega,\mathcal{A},\mathbb{P})$ such that $\sup_{t\in T}\mathbb{E}(|X_t|)<\infty$, and $m$ a probability measure on $T$, set

$$\mathcal{F}(X,m) \,=\, \sup \mathbb{E}(X_Z) \tag{4}$$

where the supremum is taken over all random variables $Z$ on $T$ with distribution $m$. Since only depending on the law $P_X$ of $X$ on $\mathbb{R}^T$, it is also denoted $\mathcal{F}(P_X,m)$. This functional was introduced by X. Fernique in the seminal works [F76,F81] to try to understand the expected supremum of random processes, in particular Gaussian processes. Fernique's intuition was that the (or a) majorizing measure (as the infimum in (3)) should be closely connected to the distribution of the maximizer of the Gaussian process, that is the distribution of the random variable $\tau$ such that $X_{\tau(\omega)}(\omega) = \sup_{t\in T}X_t(\omega)$ (assume $T$ is separated by the metric $d$). This intuition appears to be correct as developed in Section 4 of [Ta87].

Given $T$ finite, for $M$ an integer $\geq 1$ let

$$\mathcal{P}_M \,=\, \left\{ \frac{1}{M}\sum_{k=1}^{M}\delta_{t_k}\,;t_1,\ldots,t_M\in T\right\}$$

be the set of empirical probability measures of size $M$ on the points of $T$. For integers $M,N\geq 1$ and $m\in\mathcal{P}_M$, consider next the set

$$\mathbf{T}_N(m) \,=\, \left\{\mathbf{t}=(t_1,\ldots,t_{MN})\in T^{MN}\,;\frac{1}{MN}\sum_{k=1}^{MN}\delta_{t_k}=m\right\}$$

of sequences of $T^{MN}$ in which each $t\in T$ appears exactly $MNm(\{t\})$ times.

Liu's fundamental lifting principle is then described by the following statement ([L25, Lemma 5], [vH25, Proposition 3.1]).

**Theorem.** *Under the preceding notation,*

$$\mathcal{F}(X,m) \,=\, \lim_{N\to\infty}\mathbb{E}\Big(\sup_{\mathbf{t}\in\mathbf{T}_N(m)}\mathbf{X}^{m,N}_{\mathbf{t}}\Big)$$

*where*

$$\mathbf{X}^{m,N}_{\mathbf{t}} \;=\; \frac{1}{MN}\sum_{k=1}^{MN} X^{(k)}_{\mathbf{t}_k}, \quad \mathbf{t} = (t_1, \ldots, t_{MN}) \in \mathbf{T}_N(m),$$

*where $X^{(1)}, X^{(2)}, \ldots$ are independent copies of the process $X$.*

By centering of $X$, the limit in $N$ may be shown to be increasing. The main point of the statement is that, for each $N \geq 1$, the process $(\mathbf{X}^{m,N}_{\mathbf{t}})_{\mathbf{t}\in\mathbf{T}_{N(m)}}$ is stationary. Indeed, the symmetric group $\mathcal{S}_{MN}$ acts on $\mathbf{T}_{N(m)}$ by $\sigma(\mathbf{t}) = (\mathbf{t}_{\sigma(1)}, \ldots, \mathbf{t}_{\sigma(MN)})$ for $\sigma \in \mathcal{S}_{MN}$ and $\mathbf{t} = (t_1, \ldots, t_{MN}) \in \mathbf{T}_{N(m)}$. This action is transitive and since

$$\mathbf{X}^{m,N}_{\sigma(\mathbf{t})} \;=\; \frac{1}{MN}\sum_{k=1}^{MN} X^{(\sigma^{-1}(k))}_{\mathbf{t}_k},$$

the processes $(\mathbf{X}^{m,N}_{\sigma(\mathbf{t})})_{\mathbf{t}\in\mathbf{T}_{N(m)}}$ and $(\mathbf{X}^{m,N}_{\mathbf{t}})_{\mathbf{t}\in\mathbf{T}_{N(m)}}$ have the same distribution for every $\sigma \in \mathcal{S}_{MN}$. The theorem thus reduces the computation of Fernique's functional $\mathcal{F}(X,m)$ for an arbitrary random process to the computation of the expected supremum of a stationary process.

Now

$$\mathbb{E}\Big(\sup_{t\in T} X_t\Big) \;=\; \sup_m \mathcal{F}(X,m),$$

and, by continuity of the map $m \mapsto \mathcal{F}(X,m)$ (which easily follows by arguments similar to the ones put forward for (6) below), the supremum may actually be taken over all probability measures in $\mathcal{P}_M$, $M \geq 1$. Given then a centered Gaussian process $X = (X_t)_{t\in T}$ indexed by $T$ (finite), the theorem provides a full characterization of boundedness of $X$ in terms of the entropy integral for the $N$-extended (Gaussian) processes $(\mathbf{X}^{m,N}_{\mathbf{t}})_{\mathbf{t}\in\mathbf{T}_{N(m)}}$ indexed by $\mathbf{T}_{N(m)}$ (as $N \to \infty$) as formulated by the Dudley-Fernique theorem (2), that is

$$\mathbb{E}\Big(\sup_{\mathbf{t}\in\mathbf{T}_{N(m)}} \mathbf{X}^{m,N}_{\mathbf{t}}\Big) \;\sim\; \int_0^\infty \Big(\log\big(\mathbf{T}_N(m), \mathbf{d}; \varepsilon\big)\Big)^{1/2} d\varepsilon$$

for the intrinsic metric

$$\mathbf{d}(\mathbf{s},\mathbf{t})^2 \;=\; \frac{1}{MN}\sum_{k=1}^{MN} d(\mathbf{s}_k,\mathbf{t}_k)^2, \quad \mathbf{s},\mathbf{t} \in \mathbf{T}_N(m).$$

Therefore

$$\begin{aligned}\mathbb{E}\Big(\sup_{t\in T} X_t\Big) &\sim \sup_{m\in\mathcal{P}_M, M\geq 1} \lim_{N\to\infty} \mathbb{E}\Big(\sup_{\mathbf{t}\in\mathbf{T}_{N(m)}} \mathbf{X}^{m,N}_{\mathbf{t}}\Big) \\ &\sim \sup_{m\in\mathcal{P}_M, M\geq 1} \lim_{N\to\infty} \int_0^\infty \Big(\log\big(\mathbf{T}_N(m), \mathbf{d}; \varepsilon\big)\Big)^{1/2} d\varepsilon \end{aligned} \tag{5}$$

(up to numerical constants).

In his work [L25], J. Liu provides an information-theoretic description of the right-hand side of (5) via coding theory, and from which he establishes its equivalence with the Talagrand

$\gamma_2(T,d)$ functional, briefly described below. But (5) already provides a simple direct metric characterization of boundedness of Gaussian process in terms of the classical metric entropy numbers.

Liu's information-theoretic description is as follows (referring to [L25] for a more careful exposition). For $m$ a probability measure on a metric space $(T,d)$ and $\varepsilon > 0$, set

$$\Pi_\varepsilon(m) \,=\, \Big\{(U,V); \mathcal{L}(U) = \mathcal{L}(V) = m,\, \mathbb{E}\big(d(U,V)^2\big) \le \varepsilon^2\Big\}.$$

Define then

$$i_m(\varepsilon) \,=\, \inf_{(U,V)\in\Pi_\varepsilon(m)} \mathrm{I}(U,V)$$

where $\mathrm{I}(U,V)$ is the mutual information for the law $P_{(U,V)}$ of $(U,V)$ defined as the relative entropy

$$\mathrm{I}(U,V) \,=\, \int_T \log\left(\frac{P_{(U,V)}}{P_U \otimes P_V}\right) dP_{(U,V)}$$

of $P_{(U,V)}$ with respect to the product law $P_U \otimes P_V$ of the marginals. J. Liu then establishes by the main theorem above together with basic coding theory that

$$\mathcal{F}(X,m) \,\sim\, \int_0^\infty \big(i_m(\varepsilon)\big)^{1/2}\, d\varepsilon$$

up to numerical constants. The supremum over $m$ of the right-hand side is then showed to be equivalent to the $\gamma_2(T,d)$ functional, providing thus an alternate proof of the majorizing measure theorem.

The note is now completed by the proof of Liu's lifting theorem copying [vH25].

*Proof of the theorem.* The first step is to observe that the Fernique functional $\mathcal{F}(X,m)$ is Lipschitz with respect to the classical Kantorovich distance $\mathrm{W}_1$ on probability measures $\mu,\nu$ on $\mathbb{R}^T$ given by

$$\mathrm{W}_1(\mu,\nu) \,=\, \inf \mathbb{E}\big(\|U-V\|\big)$$

where the infimum is taken over all couples $(U,V)$ of random vectors in $\mathbb{R}^T$ with $\mathcal{L}(U)=\mu$, $\mathcal{L}(V)=\nu$, and where (for example) $\|u-v\| = \sum_{t\in T}|u_t - v_t|$ for $u=(u_t)_{t\in T}, v=(v_t)_{t\in T} \in \mathbb{R}^T$. Namely, given $m$ a probability measure on $T$, and two processes $X=(X_t)_{t\in T}$ and $Y=(Y_t)_{t\in T}$ on $T$ (finite) such that $\sup_{t\in T}\mathbb{E}(|X_t|) < \infty$ and $\sup_{t\in T}\mathbb{E}(|Y_t|) < \infty$, for any coupling $(U,V)$ with $\mathcal{L}(U) = P_X$, $\mathcal{L}(V) = P_Y$

$$\mathcal{F}(P_X,m) - \mathcal{F}(P_Y,m) \,\le\, \sup_Z \mathbb{E}(U_Z - V_Z) \,\le\, \mathbb{E}\big(\|U-V\|\big)$$

where the supremum runs over all random vectors $Z$ on $T$ with distribution $m$. As a consequence

$$\big|\mathcal{F}(P_X,m) - \mathcal{F}(P_Y,m)\big| \,\le\, \mathrm{W}_1(P_X,P_Y). \tag{6}$$

It is also a classical consequence of the law of large numbers and the metric properties of $\mathrm{W}_1$ (cf. e.g. [Vi03]) that

$$\lim_{N\to\infty} \mathbb{E}\bigg(\mathrm{W}_1\bigg(\frac{1}{N}\sum_{k=1}^N \delta_{X^{(k)}}, P_X\bigg)\bigg) \,=\, 0$$

where it is recalled that $P_X$ is the law of $X$ on $\mathbb{R}^T$ and that $X^{(1)}, X^{(2)}, \dots$ are independent copies of $X$. Together with (6),

$$\mathcal{F}(P_X, m) \;=\; \lim_{N\to\infty} \mathbb{E}\bigg(\mathcal{F}\bigg(\frac{1}{N}\sum_{k=1}^{N} \delta_{X^{(k)}}, m\bigg)\bigg). \tag{7}$$

Provided with these technical preliminaries, the main argument of the proof is as follows. Fix $\mathbf{s}$ in $\mathbf{T}_{N(m)}$. Any coupling of $\frac{1}{MN}\sum_{k=1}^{MN}\delta_{X^{(k)}}$ and $m = \frac{1}{M}\sum_{\ell=1}^{M}\delta_\ell$ can be realized by selecting each pair $(X^{(k)}, \mathbf{s}_\ell)$ with probability $\frac{1}{MN}\Pi_{k\ell}$ where $\Pi$ is an $MN \times MN$ bistochastic matrix. Since, by Birkhoff's theorem, the set $B_{MN}$ of bistochastic matrices is the convex hull of permutation matrices,

$$\begin{aligned}
\mathcal{F}\bigg(\frac{1}{MN}\sum_{k=1}^{MN}\delta_{X^{(k)}}, m\bigg) \;&=\; \sup_{\Pi\in B_{MN}} \frac{1}{MN}\sum_{k,\ell=1}^{MN} \Pi_{k\ell} X^{(k)}_{\mathbf{s}_\ell} \\
&=\; \sup_{\sigma\in\mathcal{S}^{MN}} \frac{1}{MN}\sum_{k=1}^{MN} X^{(k)}_{\sigma(\mathbf{s})_k} \\
&=\; \sup_{\mathbf{t}\in\mathbf{T}_N(m)} \mathbf{X}^{m,N}_{\mathbf{t}}.
\end{aligned}$$

Taking expectations and applying (7) concludes the proof.


[Fe76] X. Fernique. Évaluations de processus gaussiens composés. Probability in Banach spaces. Lect. Notes in Math. 526, 67–83 (1976). Springer

[Fe81] X. Fernique. Évaluations de certaines fonctionnelles associées à des fonctions aléatoires gaussiennes. Probab. Math. Statist. 2, 1–29 (1981).

[L25] J. Liu. Simple and sharp generalization bounds via lifting (2025).

[vH25] R. van Handel. On the subGaussian comparison theorem (2025).

[Ta87] M. Talagrand. Regularity of Gaussian processes. Acta Math. 159, 99–149 (1987).

[Ta21] M. Talagrand. Upper and lower bounds of stochastic processes. Modern methods and classical problems. Ergebnisse der Mathematik und ihrer Grenzgebiete 60. Springer (2014).

[Vi03] C. Villani. Topics in optimal transportation. Graduate Studies in Mathematics 58. American Mathematical Society (2003).

## 6. REGULARITY OF GAUSSIAN PROCESSES

In this chapter, we provide a complete treatment of boundedness and continuity of Gaussian processes via the tool of majorizing measures. After the work of R. M. Dudley, V. Strassen, V. N. Sudakov and X. Fernique on entropy, M. Talagrand [Ta2] gave, in 1987, necessary and sufficient conditions on the covariance structure of a Gaussian process in order that it is almost surely bounded or continuous. These necessary and sufficient conditions are based on the concept of majorizing measure introduced in the early seventies by X. Fernique and C. Preston, and inspired in particular by the "real variable lemma" of A. M. Garsia, E. Rodemich and H. Rumsey Jr. [G-R-R]. Recently, M. Talagrand [Ta7] gave a simple proof of his theorem on necessity of majorizing measures based on the concentration phenomenon for Gaussian measures. We follow this approach here. The aim of this chapter is in fact to demonstrate the actual simplicity of majorizing measures that are usually considered as difficult and obscure.

Let $T$ be a set. A Gaussian random process (or better, random function) $X = (X_t)_{t\in T}$ is a family, indexed by $T$, of random variables on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that the law of each finite family $(X_{t_1}, \ldots, X_{t_n})$, $t_1, \ldots, t_n \in T$, is centered Gaussian on $\mathbb{R}^n$. Throughout this work, Gaussian will always mean centered Gaussian. In particular, the law (the distributions of the finite dimensional marginals) of the process $X$ is uniquely determined by the covariance structure $\mathbb{E}(X_s X_t)$, $s, t \in T$. Our aim will be to characterize almost sure boundedness and continuity (whenever $T$ is a topological space) of the Gaussian process $X$ in terms of an as simple as possible criterion on this covariance structure. Actually, the main point in this study will be the question of boundedness. As we will see indeed, once the appropriate bounds for the supremum of $X$ are obtained, the characterization of continuity easily follows. Due to the integrability properties of norms of Gaussian random vectors or supremum of Gaussian processes (Theorem 4.1), we will avoid, at a first stage, various cumbersome and unessential measurability questions, by considering the supremum functional

$$F(T) = \sup\big\{\mathbb{E}\big(\sup_{t\in U} X_t\big); U \text{ finite in } T\big\}.$$

(If $S \subset T$, we define in the same way $F(S)$.) Thus, $F(T) < \infty$ if and only if $X$ is almost surely bounded in any reasonable sense. In particular, we already see that the main question will reduce to a uniform control of $F(U)$ over the finite subsets $U$ of $T$.

After various preliminary results [Fe1], [De]..., the first main idea in the study of regularity of Gaussian processes is the introduction (in the probabilistic area), by R. M. Dudley, V. Strassen and V. N. Sudakov (cf. [Du1], [Du2], [Su1-4]), of the notion of $\varepsilon$-entropy. The idea consists in connecting the regularity of the Gaussian process $X = (X_t)_{t\in T}$ to the size of the parameter set $T$ for the $L^2$-metric induced by the process itself and given by

$$d(s,t) = \big(\mathbb{E}|X_s - X_t|^2\big)^{1/2}, \quad s,t \in T.$$

Note that this metric is entirely characterized by the covariance structure of the process. It does not necessarily separate points in $T$ but this is of no importance. The size of $T$ is more precisely estimated by the entropy numbers: for every $\varepsilon > 0$, let $N(T,d;\varepsilon)$ denote the minimal number of (open to fix the idea) balls of radius $\varepsilon$ for the metric $d$ that are necessary to cover $T$. The two main results concerning regularity of Gaussian processes under entropy conditions, due to R. M. Dudley [Du1] for the upper bound and V. N. Sudakov [Su3] for the lower bound (cf. [Du2], [Fe4]), are summarized in the following statement.

**Theorem 6.1.** *There are numerical constants $C_1 > 0$ and $C_2 > 0$ such that for all Gaussian processes $X = (X_t)_{t\in T}$,*

$$(6.1) \qquad C_1^{-1} \sup_{\varepsilon>0} \varepsilon\big(\log N(T,d;\varepsilon)\big)^{1/2} \le F(T) \le C_2 \int_0^\infty \big(\log N(T,d;\varepsilon)\big)^{1/2}\, d\varepsilon.$$

As possible numerical values for $C_1$ and $C_2$, one may take $C_1 = 6$ and $C_2 = 42$ (see below). The convergence of the entropy integral is understood for the small values of $\varepsilon$ since it stops at the diameter $D(T) = \sup\{d(s,t); s,t \in T\}$. Actually, if any of the three terms of (6.1) is finite, then $(T,d)$ is totally bounded and in particular $D(T) < \infty$. We will show in more generality below that the process $X = (X_t)_{t\in T}$ actually admits an almost surely continuous version when the entropy integral is finite. Conversely, if $X = (X_t)_{t\in T}$ is continuous, one can show that $\lim_{\varepsilon\to 0} \varepsilon(\log N(T,d;\varepsilon))^{1/2} = 0$ (cf. [Fe4]).

For the matter of comparison with the more refined tool of majorizing measures we will study next, we present a sketch of the proof of Theorem 6.1.

*Proof.* We start with the upper bound. We may and do assume that $T$ is finite (although this is not strictly necessary). Let $q > 1$ (usually an integer). (We will consider $q$ as a power of discretization; a posteriori, its value is completely arbitrary.) Let $n_0$ be the largest integer $n$ in $\mathbb{Z}$ such that $N(T,d;q^{-n}) = 1$. For every $n \ge n_0$, we consider a family of cardinality $N(T,d;q^{-n}) = N(n)$ of balls of radius $q^{-n}$ covering $T$. One may therefore construct a partition $\mathcal{A}_n$ of $T$ of cardinality $N(n)$ on the basis of this covering with sets of diameter less than $2q^{-n}$. In each $A$ of $\mathcal{A}_n$, fix a point of $T$ and denote by $T_n$ the collection of these points. For each $t$ in $T$, denote by $A_n(t)$

the element of $\mathcal{A}_n$ that contains $t$. For every $t$ and every $n$, let then $s_n(t)$ be the element of $T_n$ such that $t \in A_n(s_n(t))$. Note that $d(t, s_n(t)) \leq 2q^{-n}$ for every $t$ and $n \geq n_0$.

The main argument of the proof is the so-called chaining argument (which goes back to A. N. Kolmogorov in his proof of continuity of paths of processes under $\mathrm{L}^p$-control of their increments): for every $t$,

$$X_t = X_{s_0} + \sum_{n>n_0} \big(X_{s_n(t)} - X_{s_{n-1}(t)}\big) \tag{6.2}$$

where $s_0 = s_{n_0}(t)$ may be chosen independent of $t \in T$. Note that

$$d\big(s_n(t), s_{n-1}(t)\big) \leq 2q^{-n} + 2q^{-n+1} = 2(q+1)q^{-n}.$$

Let $c_n = 4(q+1)q^{-n}(\log N(n))^{1/2}$, $n > n_0$. It follows from (6.2) that

$$\begin{aligned} F(T) &= \mathbb{E}\big(\sup_{t\in T} X_t\big) \\ &\leq \sum_{n>n_0} c_n + \mathbb{E}\bigg(\sup_{t\in T} \sum_{n>n_0} \big|X_{s_n(t)} - X_{s_{n-1}(t)}\big| I_{\{|X_{s_n(t)} - X_{s_{n-1}(t)}| > c_n\}}\bigg) \\ &\leq \sum_{n>n_0} c_n + \mathbb{E}\bigg(\sum_{n>n_0} \sum_{(u,v)\in H_n} \big|X_u - X_v\big| I_{\{|X_u - X_v| > c_n\}}\bigg) \end{aligned}$$

where $H_n = \{(u,v) \in T_n \times T_{n-1}; d(u,v) \leq 2(q+1)q^{-n}\}$. If $G$ is a real centered Gaussian variable with variance less than or equal to $\sigma^2$, for every $c > 0$

$$\mathbb{E}\big(|G| I_{\{|G|>c\}}\big) \leq \sigma \mathrm{e}^{-c^2/2\sigma^2}.$$

Hence,

$$\begin{aligned} F(T) &\leq \sum_{n>n_0} c_n + \sum_{n>n_0} \mathrm{Card}(H_n) 2(q+1)q^{-n} \exp\big(-c_n^2/8(q+1)^2 q^{-2n}\big) \\ &\leq \sum_{n>n_0} 4(q+1)q^{-n}\big(\log N(n)\big)^{1/2} + \sum_{n>n_0} 2(q+1)q^{-n} \\ &\leq 7(q+1) \sum_{n>n_0} q^{-n}\big(\log N(n)\big)^{1/2} \end{aligned}$$

where we used that $\mathrm{Card}(H_n) \leq N(n)^2$. Since

$$\begin{aligned} \int_0^\infty \big(\log N(T,d;\varepsilon)\big)^{1/2} d\varepsilon &\geq \sum_{n>n_0} \int_{q^{-n-1}}^{q^{-n}} \big(\log N(T,d;\varepsilon)\big)^{1/2} d\varepsilon \\ &\geq (1-q^{-1}) \sum_{n>n_0} q^{-n}\big(\log N(n)\big)^{1/2}, \end{aligned}$$

the conclusion follows. If $q = 2$, we may take $C_2 = 42$.

The proof of the lower bound relies on a comparison principle known as Slepian's lemma [Sl]. We use it in the following modified form due to V. N. Sudakov, S. Chevet and X. Fernique (cf. [Su1], [Su2], [Fe4], [L-T2]): if $Y = (Y_1, \ldots, Y_n)$ and $Z = (Z_1, \ldots, Z_n)$ are two Gaussian random vectors in $\mathbb{R}^n$ such that $\mathbb{E}|Y_i - Y_j|^2 \le \mathbb{E}|Z_i - Z_j|^2$ for all $i, j$, then

$$\mathbb{E}\big(\max_{1\le i\le n} Y_i\big) \le \mathbb{E}\big(\max_{1\le i\le n} Z_i\big). \tag{6.3}$$

Fix $\varepsilon > 0$ and let $n \le N(T, d; \varepsilon)$. There exist therefore $t_1, \ldots, t_n$ in $T$ such that $d(t_i, t_j) \ge \varepsilon$. Let then $g_1, \ldots, g_n$ be independent standard normal random variables. We have, for every $i, j = 1, \ldots, n$,

$$\mathbb{E}\left|\frac{\varepsilon}{\sqrt{2}}\, g_i - \frac{\varepsilon}{\sqrt{2}}\, g_j\right|^2 = \varepsilon^2 \le d(t_i, t_j) = \mathbb{E}|X_{t_i} - X_{t_j}|^2.$$

Therefore, by (6.3),

$$F(T) \ge \mathbb{E}\big(\max_{1\le i\le n} X_{t_i}\big) \ge \frac{\varepsilon}{\sqrt{2}}\, \mathbb{E}\big(\max_{1\le i\le n} g_i\big).$$

Now, it is classical and easily seen that

$$\mathbb{E}\big(\max_{1\le i\le n} g_i\big) \ge c\,(\log n)^{1/2}$$

for some numerical $c > 0$ (one may choose $c$ such that $\sqrt{2}/c \le 6$). Since $n$ is arbitrary less than or equal to $N(T, d; \varepsilon)$, the conclusion trivially follows. Theorem 6.1 is established. □

As an important remark for further purposes, note that simple proofs of Sudakov's minoration avoiding the rather rigid Slepian's lemma are now available. These are based on duality of entropy numbers [TJ] and are presented in [L-T2]. They allow the investigation of minoration inequalities outside the Gaussian setting (cf. [Ta10], [Ta12]). Note furthermore that we will only use the Sudakov inequality in the proof of the majorizing measure minoration principle (cf. Lemma 6.4).

A simple example of application of Theorem 6.1 is Brownian motion $(W(t))_{0\le t\le 1}$ on $T = [0, 1]$. Since $d(s, t) = \sqrt{|s - t|}$, the entropy numbers $N(T, d; \varepsilon)$ are of the order of $\varepsilon^{-2}$ as $\varepsilon$ goes to zero and the entropy integral is trivially convergent. Together with the proof of continuity presented below in the framework of majorizing measures, Theorem 6.1 is certainly the shortest way to prove boundedness and continuity of the Brownian paths.

In Theorem 6.1, the difference between the upper and lower bounds is rather tight. It however exists. The examples of a standard orthogaussian sequence or of the canonical Gaussian process indexed by an ellipsoid in a Hilbert space (see [Du1], [Du2], [L-T2], [Ta13]) are already instructive. We will see later on that the convergence of Dudley's entropy integral however characterizes $F(T)$ when $T$ has a group structure and the metric $d$ is translation invariant, an important result of X. Fernique [Fe4].

If one tries to imagine what can be used instead of the entropy numbers in order to sharpen the conclusions of Theorem 6.1, one realizes that one feature of entropy is that is attributes an equal weight to each piece of the parameter set $T$. One is then naturally led to the possible following definition. Let, as in the proof of Theorem 6.1, $q$ be (an integer) larger than 1. Let $\mathcal{A} = (\mathcal{A}_n)_{n\in\mathbb{Z}}$ be an increasing sequence (i.e. each $A \in \mathcal{A}_{n+1}$ is contained in some $B \in \mathcal{A}_n$) of finite partitions of $T$ such that the diameter $D(A)$ of each element $A$ of $\mathcal{A}_n$ is less than or equal to $2q^{-n}$. If $t \in T$, denote by $A_n(t)$ the element of $\mathcal{A}_n$ that contains $t$. Now, for each partition $\mathcal{A}_n$, one may consider nonnegative weights $\alpha_n(A)$, $A \in \mathcal{A}_n$, such that $\sum_{A\in\mathcal{A}_n} \alpha_n(A) \le 1$. Set then

$$\Theta_{\mathcal{A},\alpha} = \Theta_{\mathcal{A},\alpha}(T,d) = \sup_{t\in T} \sum_n q^{-n} \left( \log \frac{1}{\alpha_n(A_n(t))} \right)^{1/2}. \tag{6.4}$$

It is worthwhile mentioning that for $2q^{-n} \ge D(T)$, one can take $\mathcal{A}_n = \{T\}$ and $\alpha_n(T) = 1$. Denote by $\Theta(T,d)$ the infimum of the functional $\Theta_{\mathcal{A},\alpha}$ over all possible choices of partitions $(\mathcal{A}_n)_{n\in\mathbb{Z}}$ and weights $\alpha_n(A)$. In this definition, we may take equivalently

$$\Theta_{\mathcal{A},m} = \sup_{t\in T} \sum_n q^{-n} \left( \log \frac{1}{m(A_n(t))} \right)^{1/2}$$

where $m$ is a probability measure on $(T,d)$. Indeed, if $\Theta_{\mathcal{A},\alpha} < \infty$, it is easily seen that $D(T) < \infty$. Let then $n_0$ be the largest integer $n$ in $\mathbb{Z}$ such that $2q^{-n} \le D(T)$. Fix a point in each element of $\mathcal{A}_n$ and denote by $T_n$, $n \ge n_0$, the collection of these points. It is then clear that if $m$ is a (discrete) probability measure such that

$$m \ge (1 - q^{-1}) \sum_{n\ge n_0} q^{-n+n_0} \sum_{t\in T_n} \alpha_n\big(A_n(t)\big)\delta_t,$$

where $\delta_t$ is point mass at $t$, the functional $\Theta_{\mathcal{A},m}$ is of the same order as $\Theta_{\mathcal{A},\alpha}$ (see also below). We need not actually be concerned with these technical details and consider for simplicity the functionals $\Theta_{\mathcal{A},\alpha}$. Furthermore, the number $q > 1$ should be thought as a universal constant.

The condition $\Theta(T,d) < \infty$ is called a majorizing measure condition and the main result of this section is that $C^{-1}\Theta(T,d) \le F(T) \le C\Theta(T,d)$ for some constant $C > 0$ only depending on $q$. In order to fully appreciate this definition, it is worthwhile comparing it to the entropy integral. As we used it in the proof of Theorem 6.1, the entropy integral is equivalent (for any $q$) to the series

$$\sum_{n>n_0} q^{-n} \big(\log N(T,d;q^{-n})\big)^{1/2}.$$

We then construct an associated sequence $(\mathcal{A}_n)_{n\in\mathbb{Z}}$ of increasing partitions of $T$ and weights $\alpha_n(A)$ in the following way. Let $\mathcal{A}_n = \{T\}$ and $\alpha_n(T) = 1$ for every $n \le n_0$. Once $\mathcal{A}_n$ ($n > n_0$) has been constructed, partition each element $A$ of $\mathcal{A}_n$ with a covering of $A$ of cardinality at most $N(A,d;q^{-n-1}) \le N(T,d;q^{-n-1})$ and let $\mathcal{A}_{n+1}$ be the collection of all the subsets of $T$ obtained in this way. To each $A$ in $\mathcal{A}_n$, $n > n_0$, we give the weight

$$\alpha_n(A) = \Big(\prod_{i=n_0+1}^{n} N(T,d;q^{-i})\Big)^{-1}$$

($\alpha(T) = 1$). Clearly $\sum_{A\in\mathcal{A}_n} \alpha_n(A) \leq 1$. Moreover, for each $t$ in $T$,

$$\begin{aligned}\sum_{n>n_0} q^{-n}\Big(\log\frac{1}{\alpha(A_n(t))}\Big)^{1/2} &\leq \sum_{n>n_0}\sum_{i=n_0+1}^{n} q^{-n}\big(\log N(T,d;q^{-i})\big)^{1/2}\\ &\leq (q-1)^{-1}\sum_{i>n_0} q^{-i}\big(\log N(T,d;q^{-i})\big)^{1/2}.\end{aligned}$$

In other words,

$$\Theta(T) \leq C\int_0^\infty \big(\log N(T,d;\varepsilon)\big)^{1/2} d\varepsilon$$

where $C > 0$ only depends on $q > 1$.

It is clear from this construction how entropy numbers give a uniform weight to each subset of $T$ and how the possible refined tool of majorizing measures can allow a better understanding of the metric properties of $T$. (Actually, one has rather to think about entropy numbers as the equal weight that is put on each piece of a partition of the parameter set $T$.) This is what we will investigate now. First however, we would like to briefly comment on the name “majorizing measure” as well as the dependence on $q > 1$ in the definition of the functional $\Theta(T,d)$. Classically, a majorizing measure $m$ on $T$ is a probability measure on the Borel sets of $T$ such that

$$\sup_{t\in T}\int_0^\infty \Big(\log\frac{1}{m(B(t,\varepsilon))}\Big)^{1/2} d\varepsilon < \infty \tag{6.5}$$

where $B(t,\varepsilon)$ is the ball in $T$ with center $t$ and radius $\varepsilon > 0$. As the definition of the entropy integral, a majorizing measure condition only relies on the metric structure of $T$ and the convergence of the integral is for the small values of $\varepsilon$. In order to connect this definition with the preceding one (6.4), let $q > 1$ and let $(\mathcal{A}_n)_{n\in\mathbb{Z}}$ be an increasing sequence of finite partitions of $T$ such that the diameter $D(A)$ of each element $A$ of $\mathcal{A}_n$ is less than or equal to $2q^{-n}$. Let furthermore $m$ be a probability measure on $T$. Note that $A_n(t) \subset B(t, 2q^{-n})$ for every $t$. Therefore

$$\begin{aligned}\int_0^\infty \Big(\log\frac{1}{m(B(t,\varepsilon))}\Big)^{1/2} d\varepsilon &\leq C\sum_n q^{-n}\Big(\log\frac{1}{m(B(t,2q^{-n}))}\Big)^{1/2}\\ &\leq C\sum_n q^{-n}\Big(\log\frac{1}{m(A_n(t))}\Big)^{1/2}\end{aligned}$$

where $C > 0$ only depends on $q$. Since $m$ is a probability measure, we can set $\alpha_n(A) = m(A)$ for every $A$ in $\mathcal{A}_n$ and every $n$. It immediately follows that, for every $q > 1$,

$$\inf_m \sup_{t\in T}\int_0^\infty \Big(\log\frac{1}{m(B(t,\varepsilon))}\Big)^{1/2} d\varepsilon \leq C\Theta(T,d)$$

where $C$ only depends on $q$. One can prove the reverse inequality in the same spirit with the help however of a somewhat technical and actually nontrivial discretization lemma (cf. [L-T2], Proposition 11.10). In particular, the various functionals $\Theta(T,d)$ when $q$ varies are all equivalent. We actually need not really be concerned with these technical details since our aim is to show that $F(T)$ and $\Theta(T,d)$ are of the same order (for some $q>1$). (It will actually follow from the proofs presented below that the functionals $\Theta(T,d)$ are equivalent up to constants depending only on $q \geq q_0$ for some universal $q_0$ large enough.)

Now, we start our investigation of the regularity properties of a Gaussian process $X=(X_t)_{t\in T}$ under majorizing measure conditions. The first part of our study concerns upper bounds and sufficient conditions for boundedness and continuity of $X$. The following theorem is due, in this form and with this proof, to X. Fernique [Fe3], [Fe4]. It follows independently from the work of C. Preston [Pr1], [Pr2].

**Theorem 6.2.** *Let $X=(X_t)_{t\in T}$ be a Gaussian process indexed by a set $T$. Then, for every $q>1$,*

$$F(T) \leq C\Theta(T,d)$$

*where $C>0$ only depends on $q$. If, in addition to $\Theta_{\mathcal{A},\alpha}<\infty$ for some partition $\mathcal{A}$ and weights $\alpha$, one has*

$$\lim_{k\to\infty} \sup_{t\in T} \sum_{n\geq k} q^{-n}\left(\log \frac{1}{\alpha_n(A_n(t))}\right)^{1/2} = 0, \tag{6.6}$$

*then $X$ admits a version with almost all sample paths bounded and uniformly continuous on $(T,d)$.*

*Proof.* It is very similar to the proof of the upper bound in Theorem 6.1. We first establish the inequality $F(T)\leq C\Theta_{\mathcal{A},\alpha}(T,d)$ for any partition $\mathcal{A}$ and any family of weights $\alpha$. We may asssume that $T$ is finite. Let $n_0$ be the largest integer $n$ in $\mathbb{Z}$ such that the diameter $D(T)$ of $T$ is less than or equal to $2q^{-n}$. For every $n\geq n_0$, fix a point in each element of the partition $\mathcal{A}_n$ and denote by $T_n$ the (finite) collection of these points. We may take $T_{n_0}=\{s_0\}$ for some fixed $s_0$ in $T$. For every $t$ in $T$, denote by $s_n(t)$ the element of $T_n$ which belongs to $A_n(t)$. As in (6.2), for every $t$,

$$X_t = X_{s_0} + \sum_{n>n_0} \big(X_{s_n(t)} - X_{s_{n-1}(t)}\big).$$

Since the partitions $\mathcal{A}_n$ are increasing,

$$s_n(t) \in A_{n-1}\big(s_n(t)\big) = A_{n-1}(t), \quad n>n_0.$$

In particular, $d(s_n(t),s_{n-1}(t)) \leq 2q^{-n+1}$. Now, for every $t$ in $T$ and every $n>n_0$, let

$$c_n(t) = 2\sqrt{2}q^{-n+1}\left(\log\frac{1}{\alpha_n(A_n(t))}\right)^{1/2}.$$

With respect to the entropic proof, note here the dependence of $c_n$ on $t$ which is the main feature of the majorizing measure technique. Actually, the partitions $\mathcal{A}$ and weights $\alpha$ are used to bound, in the chaining argument, the "heaviest" portions of the process. We can now write, almost as in the proof of Theorem 6.1,

$$\begin{aligned} &F(T) \\ &\quad \le \sup_{t\in T} \sum_{n>n_0} c_n(t) + \mathbb{E}\Big( \sup_{t\in T} \sum_{n>n_0} \big|X_{s_n(t)} - X_{s_{n-1}(t)}\big| I_{\{|X_{s_n(t)}-X_{s_{n-1}(t)}|>c_n(t)\}}\Big) \\ &\quad \le \sup_{t\in T} \sum_{n>n_0} c_n(t) + \mathbb{E}\Big( \sum_{n>n_0} \sum_{u\in T_n} \big|X_u - X_{s_{n-1}(u)}\big| I_{\{|X_u-X_{s_{n-1}(u)}|>c_n(u)\}}\Big) \\ &\quad \le \sup_{t\in T} \sum_{n>n_0} c_n(t) + \sum_{n>n_0} \sum_{u\in T_n} 2q^{-n+1} \exp\big(-c_n^2(u)/8q^{-2n+2}\big). \end{aligned}$$

Therefore

$$\begin{aligned} F(T) &\le \sup_{t\in T} \sum_{n>n_0} c_n(t) + \sum_{n>n_0} 2q^{-n+1} \sum_{u\in T_n} \alpha_n\big(A_n(u)\big) \\ &\le \sup_{t\in T} \sum_{n>n_0} c_n(t) + 2(q-1)^{-1} q^{-n_0+1}. \end{aligned}$$

Since

$$\Theta_{\mathcal{A},\alpha} \ge (\log 2)^{1/2} q^{-n_0-1},$$

the first claim of Theorem 6.2 follows.

We turn to the sample path continuity. Let $\eta > 0$. For each $k$ $(> n_0)$, set

$$\begin{aligned} V = V_k = \big\{(x,y) \in T_k \times T_k; \exists\, u,v \text{ in } T \text{ such that} \\ d(u,v) \le \eta \text{ and } s_k(u) = x, s_k(v) = y\big\}. \end{aligned}$$

If $(x,y) \in V$, we fix $u_{x,y}$, $v_{x,y}$ in $T$ such that $s_k(u_{x,y}) = x$, $s_k(v_{x,y}) = y$ and $d(u_{x,y}, v_{x,y}) \le \eta$. Now, let $s,t$ in $T$ with $d(s,t) \le \eta$. Set $x = s_k(s)$, $y = s_k(t)$. Clearly $(x,y) \in V$. By the triangle inequality,

$$\begin{aligned} |X_s - X_t| &\le |X_s - X_{s_k(s)}| + |X_{s_k(s)} - X_{u_{x,y}}| + |X_{u_{x,y}} - X_{v_{x,y}}| \\ &\qquad + |X_{v_{x,y}} - X_{s_k(t)}| + |X_{s_k(t)} - X_t| \\ &\le \sup_{(x,y)\in V} |X_{u_{x,y}} - X_{v_{x,y}}| + 4 \sup_{r\in T} |X_r - X_{s_k(r)}|. \end{aligned}$$

Clearly,

$$\mathbb{E}\big( \sup_{(x,y)\in V} |X_{u_{x,y}} - X_{v_{x,y}}|\big) \le \eta \big(\mathrm{Card}(T_k)\big)^2.$$

Now, the chaining argument in the proof of boundedness similarly shows that

$$\mathbb{E}\big(\sup_{t\in T} |X_t - X_{s_k(t)}|\big) \le C \sup_{t\in T} \sum_{n\ge k} q^{-n} \left(\log \frac{1}{\alpha_n(A_n(t))}\right)^{1/2}$$

for some constant $C > 0$ (independent of $k$). Therefore, hypothesis (6.6) and the preceding inequalities ensure that for each $\varepsilon > 0$ there exists $\eta > 0$ such that, for every finite and thus also countable subset $U$ of $T$,

$$\mathbb{E}\Big(\sup_{s,t\in U, d(s,t)\le\eta} |X_s - X_t|\Big) \le \varepsilon.$$

Since $(T,d)$ is totally bounded, there exists $U$ countable and dense in $T$. Then, set $\widetilde{X}_t = X_t$ if $t \in U$ and $\widetilde{X}_t = \lim X_t$ where the limit, in probability or in $\mathrm{L}^1$, is taken for $u \to t$, $u \in U$. Then $(\widetilde{X}_t)_{t\in T}$ is a version of the process $X$ with uniformly continuous sample paths on $(T,d)$. Indeed, let, for each integer $n$, $\eta_n > 0$ be such that

$$\mathbb{E}\Big(\sup_{d(s,t)\le\eta_n} \big|\widetilde{X}_s - \widetilde{X}_t\big|\Big) \le 4^{-n}.$$

Then, if $C_n = \{\sup_{d(s,t)\le\eta_n} |\widetilde{X}_s - \widetilde{X}_t| \ge 2^{-n}\}$, $\sum_n \mathbb{P}(C_n) < \infty$ and the claim follows from the Borel-Cantelli lemma. The proof of Theorem 6.2 is complete. □

We now turn to the theorem of M. Talagrand [Ta2] on necessity of majorizing measures. This result was conjectured by X. Fernique back in 1974. As announced, we follow the simplified proof of the author [Ta7] based on concentration of Gaussian measures. This new proof moreover allows us to get some insight on the weights $\alpha$ of the "minorizing" measure.

**Theorem 6.3.** *There exists a universal value $q_0 \ge 2$ such that for every $q \ge q_0$ and every Gaussian process $X = (X_t)_{t\in T}$ indexed by $T$,*

$$\Theta(T,d) \le CF(T)$$

*where $C > 0$ is a constant only depending on $q$.*

*Proof.* The key step is provided by the following minoration principle based on concentration and Sudakov's inequality. It may actually be considered as a strengthening of the latter.

**Lemma 6.4.** *There exists a numerical constant $0 < c < \frac{1}{2}$ with the following property. If $\varepsilon > 0$ and if $t_1, \dots, t_N$ are points in $T$ such that $d(t_k, t_\ell) \ge \varepsilon$, $k \ne \ell$, $N \ge 2$, and if $B_1, \dots, B_N$ are subsets of $T$ such that $B_k \subset B(t_k, c\varepsilon)$, $k = 1, \dots N$, we have*

$$\mathbb{E}\Big(\max_{1\le k\le N} \sup_{t\in B_k} X_t\Big) \ge c\varepsilon(\log N)^{1/2} + \min_{1\le k\le N} \mathbb{E}\Big(\sup_{t\in B_k} X_t\Big).$$

*Proof.* We may assume that $B_k$ is finite for every $k$. Set $Y_k = \sup_{t\in B_k}(X_t - X_{t_k})$, $k = 1, \dots, N$. Then,

$$\sup_{t\in B_k} X_t = (Y_k - \mathbb{E}Y_k) + \mathbb{E}Y_k + X_{t_k}$$

and thus

$$\max_{1\le k\le N} X_{t_k} \le \max_{1\le k\le N} \sup_{t\in B_k} X_t + \max_{1\le k\le N} |Y_k - \mathbb{E}Y_k| - \min_{1\le k\le N} \mathbb{E}\Big(\sup_{t\in B_k} X_t\Big). \tag{6.7}$$

Integrate both sides of this inequality. By Sudakov's minoration (Theorem 6.1),

$$\mathbb{E}\big(\max_{1\le k\le N} X_{t_k}\big) \ge C_1^{-1}\varepsilon(\log N)^{1/2}.$$

Furthermore, the concentration inequalities, in the form for example of (2.9) or (4.2), (4.3), show that, for every $r \ge 0$, and every $k$,

$$\mathbb{P}\big\{|Y_k - \mathbb{E}Y_k| \ge r\big\} \le 2\mathrm{e}^{-r^2/2c^2\varepsilon^2}.$$

This estimate easily and classically implies that

$$\mathbb{E}\big(\max_{1\le k\le N} |Y_k - \mathbb{E}Y_k|\big) \le C_3 c\varepsilon(\log N)^{1/2}$$

where $C_3 > 0$ is numerical. Indeed, by the integration by parts formula, for every $\delta > 0$,

$$\begin{aligned}\mathbb{E}\big(\max_{1\le k\le N} |Y_k - \mathbb{E}Y_k|\big) &\le \delta + \int_\delta^\infty \mathbb{P}\big\{\max_{1\le k\le N} |Y_k - \mathbb{E}Y_k| \ge r\big\}dr \\ &\le \delta + 2N\int_\delta^\infty \mathrm{e}^{-r^2/2c^2\varepsilon^2}dr\end{aligned}$$

and the conclusion follows by letting $\delta$ be of the order of $c\varepsilon(\log N)^{1/2}$. Hence, coming back to (6.7), we see that if $c > 0$ is such that $\frac{1}{C_1} - cC_3 = c$, the minoration inequality of the lemma holds. The value of $q_0$ in Theorem 6.3 only depends on this choice. (Since we may take $C_1 = 6$ and $C_3 = 20$ (for example), we see that $c = .007$ will work.) Lemma 6.4 is proved. □

We now start the proof of Theorem 6.3 itself and the construction of a partition $\mathcal{A}$ and weights $\alpha$. Assume that $F(T) < \infty$ otherwise there is nothing to prove. In particular, $(T,d)$ is totally bounded. We further assume that $q \ge q_0$ where $q_0 = c^{-1}$ has been determined by Lemma 6.4.

For each $n$ and each subset of $T$ of diameter less than or equal to $2q^{-n}$, we will construct an associated partition in sets of diameter less than or equal to $2q^{-n-1}$. Let thus $S$ be a subset of $T$ with $D(S) \le 2q^{-n}$. We first construct by induction a (finite) sequence $(t_k)_{k\ge 1}$ of points in $S$. $t_1$ is chosen so that $F(S\cap B(t_1, q^{-n-2}))$ is maximal. Assume that $t_1, \dots, t_{k-1}$ have been constructed and set

$$H_k = \bigcup_{\ell<k}\big(S\cap B(t_\ell, q^{-n-1})\big).$$

If $H_k = S$, the construction stops (and it will eventually stop since $(T,d)$ is totally bounded). If not, choose $t_k$ in $S\setminus H_k$ such that $F(B_k)$ is maximal where we set $B_k = (S\setminus H_k)\cap B(t_k, q^{-n-2})$. For every $k$, let

$$A_k = (S\setminus H_k)\cap B(t_k, q^{-n-1}).$$

Clearly $D(A_k) \le 2q^{-n-1}$ and the $A_k$'s define a partition of $S$. One important feature of this construction is that, for every $t$ in $A_k$,

$$F\big(A_k \cap B(t, q^{-n-2})\big) \le F(B_k). \tag{6.8}$$

On the other hand, the minoration lemma 6.4 applied with $\varepsilon = q^{-n-1}$ yields (since $q \ge c^{-1}$), for every $k$,

$$F(S) \ge cq^{-n-1}(\log k)^{1/2} + F(B_k). \tag{6.9}$$

We denote by $\mathcal{A}(S)$ this ordered finite partition $\{A_1, \ldots, A_k, \ldots\}$ of $S$. (6.8) and (6.9) together yield: for every $A_k \in \mathcal{A}(S)$ and every $U \in \mathcal{A}(A_k)$,

$$F(S) \ge cq^{-n-1}(\log k)^{1/2} + F(U). \tag{6.10}$$

We now complete the construction. Let $n_0$ be the largest in $\mathbb{Z}$ with $D(T) \le 2q^{-n_0}$. Set $\mathcal{A}_n = \{T\}$ and $\alpha_n(T) = 1$ for every $n \le n_0$. Suppose that $\mathcal{A}_n$ and $\alpha_n(S)$, $S \in \mathcal{A}_n$, $n > n_0$, have been constructed. We define

$$\mathcal{A}_{n+1} = \bigcup \big\{\mathcal{A}(S); S \in \mathcal{A}_n\big\}.$$

If $U \in \mathcal{A}_{n+1}$, there exists $S \in \mathcal{A}_n$ such that $U = A_k \in \mathcal{A}(S)$. We then set $\alpha_{n+1}(U) = \alpha_n(A)/2k^2$. Let $t$ be fixed in $T$. With this notation, (6.10) means that for all $n \ge n_0$,

$$F\big(A_n(t)\big) \ge c\,2^{-1/2} q^{-n-1}\left(\log \frac{\alpha_n(A_n(t))}{2\alpha_{n+1}(A_{n+1}(t))}\right)^{1/2} + F\big(A_{n+2}(t)\big)$$

where we recall that we denote by $A_n(t)$ the element of $\mathcal{A}_n$ that contains $t$. Summing these inequalities separately on the even and odd integers, we get

$$2F(T) \ge c\,2^{-1/2} \sum_{n > n_0} q^{-n-1}\left(\log \frac{\alpha_n(A_n(t))}{2\alpha_{n+1}(A_{n+1}(t))}\right)^{1/2}$$

and thus

$$c(q-1)^{-1}q^{-n_0} + 2F(T) \ge c\,2^{-1/2}(1 - q^{-1}) \sum_{n > n_0} q^{-n}\left(\log \frac{1}{\alpha_n(A_n(t))}\right)^{1/2}.$$

Since $2q^{-n_0} \le D(T)$, and since

$$\begin{aligned} 2F(T) &= \sup\big\{\mathbb{E}\big(\sup_{s,t \in U} |X_s - X_t|\big); U \text{ finite in } T\big\} \\ &\ge \sup_{s,t \in T} \mathbb{E}|X_s - X_t| = \left(\frac{2}{\sqrt{\pi}}\right)^{1/2} D(T), \end{aligned}$$

it follows that, for some constant $C > 0$ only depending on $q$,

$$CF(T) \geq c \sum_{n > n_0} q^{-n} \left( \log \frac{1}{\alpha_n(A_n(t))} \right)^{1/2}.$$

Theorem 6.3 is therefore established. □

It may be shown that if the Gaussian process $X$ in Theorem 6.3 is almost surely continuous on $(T, d)$, then there is a majorizing measure satisfying (6.6). We refer to [Ta2] or [L-T2] for the details.

Theorem 6.3 thus solved the question of the regularity properties of any Gaussian process. Prior to this result however, X. Fernique showed [Fe4] that the convergence of Dudley's entropy integral was necessary for a stationary Gaussian process to be almost surely bounded or continuous. One can actually easily show (cf. [L-T2]) that, in this case, the entropy integral coincides with a majorizing measure integral with respect to the Haar measure on the underlying parameter set $T$ endowed with a group structure. One may however also provide a direct and transparent proof of the stationary case on the basis of the above minoration principle (Lemma 6.4). We would like to conclude this chapter with a brief sketch of this proof.

Let thus $T$ be a locally compact Abelian group. Let $X = (X_t)_{t \in T}$ be a stationary centered Gaussian process indexed by $T$, in the sense that the $L^2$-metric $d$ induced by $X$ is translation invariant on $T$. As announced, we aim to prove directly that for some numerical constant $C > 0$,

$$\int_0^\infty \big( \log N(T, d; \varepsilon) \big)^{1/2} d\varepsilon \leq CF(T).$$

(cf. [Fe4], [M-P], [L-T2] for more general statements along these lines.) Since $d$ is translation invariant,

$$\mathbb{E}\big( \sup_{s \in B(t,\varepsilon)} X_s \big) \quad \text{and} \quad N\big(B(t, \varepsilon), d; \eta\big), \quad \varepsilon, \eta > 0,$$

are independent of the point $t$. They will therefore be simpler denoted as

$$\mathbb{E}\big( \sup_{s \in B(\varepsilon)} X_s \big) \quad \text{and} \quad N\big(B(\varepsilon), d; \eta\big).$$

Let $n \in \mathbb{Z}$. Choose in a ball $B(q^{-n})$ a maximal family $(t_1, \ldots, t_M)$ under the relations $d(t_k, t_\ell) \geq q^{-n-1}$, $k \neq \ell$. Then the balls $B(t_k, q^{-n-1})$, $1 \leq k \leq M$, cover $B(q^{-n})$ so that $M \geq N(B(q^{-n}), d; q^{-n-1})$. Apply then Lemma 6.4 with $\varepsilon = q^{-n-1}$, $q \geq q_0 = c^{-1}$ and $B_k = B(t_k, q^{-n-2})$. We thus get

$$\mathbb{E}\big( \sup_{t \in B(q^{-n})} X_t \big) \geq c q^{-n-1} \big( \log N(B(q^{-n}), d; q^{-n-1}) \big)^{1/2} + \mathbb{E}\big( \sup_{t \in B(q^{-n-2})} X_t \big).$$

Summing as before these inequalities along the even and the odd integers yields

$$F(T) \geq C^{-1} \sum_n q^{-n} \big( \log N(B(q^{-n}), d; q^{-n-1} \big)^{1/2}.$$

Since

$$N(T,d;q^{-n-1}) \le N(T,d;q^{-n})N\big(B(q^{-n}),d;q^{-n-1}\big),$$

the proof is complete.

To conclude, let us mention the following challenging open problem. Let $x_i$, $i \in \mathbb{N}$, be real valued functions on a set $T$ such that $\sum_i x_i(t)^2 < \infty$ for every $t \in T$. Let furthermore $(\varepsilon_i)_{i\in\mathbb{N}}$ be a sequence of independent symmetric Bernoulli random variables and set, for each $t \in T$, $X_t = \sum_i \varepsilon_i x_i(t)$ which converges almost surely. The question of characterizing those "Bernoulli" processes $(X_t)_{t\in T}$ which are almost surely bounded is almost completely open (cf. [L-T2], [Ta14]). The Gaussian study of this chapter of course corresponds to the choice for $(\varepsilon_i)_{i\in\mathbb{N}}$ of a standard Gaussian sequence.

*Notes for further reading.* On the history of entropy and majorizing measures, one may consult respectively [Du2], [Fe4] and [He], [Fe4], [Ta2], [Ta18]. The first proof of Theorem 6.3 by M. Talagrand [Ta2] was quite different from the proof presented here following [Ta7]. Another proof may be found in [L-T2]. These proofs are based on the fundamental principle, somewhat hidden here, that the size of a metric space with respect to the existence of a majorizing measure can be measured by the size of the well separated subsets it contains (see [Ta10], [Ta12] for more on this principle). More on majorizing measures and minoration of random processes may be found in [L-T2] and in the papers [Ta10], [Ta12], and in the recent survey [Ta18] where in particular new examples of applications are described. It is shown in [L-T2] how the upper bound techniques based on entropy or majorizing measures (Theorems 6.1 and 6.2) can yield deviation inequalities of the type (4.2), which are optimal by Theorem 6.3. Sharp bounds on the tail of the supremum of a Gaussian process can be obtained with these methods (see e.g. [Ta13], [Lif2], [Lif3] and the many references therein). On construction of majorizing measures, see [L-T2], [Ta14], [Ta18]. For the applications of the Dudley-Fernique theorem on stationary Gaussian processes to random Fourier series, see [M-P], [L-T2].