ISRAEL JOURNAL OF MATHEMATICS 79 (1992), 207–224

SUDAKOV-TYPE MINORATION FOR GAUSSIAN CHAOS PROCESSES

BY

MICHEL TALAGRAND*

University of Paris VI, 75230 Paris Cedex 05, France; and The Ohio State University, Columbus, OH 43210, USA

ABSTRACT

Consider a set A of symmetric $n \times n$ matrices $a = (a_{i,j})_{i,j \le n}$. Consider an independent sequence $(g_i)_{i \le n}$ of standard normal random variables, and let $M = E\sup_{a \in A} |\sum_{i,j \le n} a_{i,j} g_i g_j|$. Denote by $N_2(A,\alpha)$ (resp. $N_\epsilon(A,\alpha)$) the smallest number of balls of radius α for the l_2 norm of $\mathbb{R}^{n^2}$ (resp. the operator norm) needed to cover A. Then for a universal constant K we have $\alpha(\log N_2(A,\alpha))^{1/4} \le KM$. This inequality is best possible. We also show that for $\delta \ge 0$, there exists a constant $K(\delta)$ such that $\alpha(\log N_\epsilon(A,\alpha))^{1/(2+\delta)} \le K(\delta)M$.

1. Introduction

Consider an orthogaussian sequence $(g_i)_{i \le n}$. For a subset A of $\mathbb{R}^n$, set $\ell(A) = E\sup_{a \in A} |\sum_{i \le n} a_i g_i|$. The value of $\ell(A)$ in function of the geometry of A is now (in principle) completely elucidated [T1]. An important early result, that is still of considerable use, is as follows.

Remark: Sudakov's minoration. Denote by $N(A,\alpha)$ the smallest number of balls in the ℓ^2 metric of radius α that are needed to cover A. Then, for some universal constant K_0,

$$(1.1) \qquad \alpha(\log N(A,\alpha))^{1/2} \le K_0 \ell(A).$$

∎

* Work partially supported by an N.S.F. grant.
Received January 4, 1991 and in revised form July 3, 1991

Observe that this statement is independent of the dimension n, so is in essence an infinite-dimensional statement.

Consider now a set A of $n \times n$ matrices; set

$$\ell^{(2)}(A) = E \sup_{a \in A} \Big| \sum_{i,j \le n} a_{i,j} g_i g_j \Big|.$$

In contrast with the case of $\ell(A)$, little is known about the relationship between $\ell^{(2)}(A)$ and the geometry of A. The distribution of a Gaussian random variable $\sum_{i \le n} a_i g_i$ depends only on its variance but the structure of a chaos $X = \sum_{i,j \le n} a_{i,j} g_i g_j$ is far more complicated. Consider $\|a\|_2 = (\sum_{i,j \le n} a_{i,j}^2)^{1/2}$, the ℓ^2 norm of a. Consider

$$\|a\|_\epsilon = \sup \left\{ \sum_{i,j \le n} a_{i,j} h_i k_j : \sum_{i \le n} h_i^2 \le 1, \ \sum_{j \le n} k_j^2 \le 1 \right\}.$$

This is the operator norm of a seen as an operator from $\mathbb{R}^n$ to $\mathbb{R}^n$; or, equivalently the norm of a seen as an element of the injective tensor product $\mathbb{R}^n \otimes_\epsilon \mathbb{R}^n$. Observe that, by Cauchy Schwarz inequality, we have $\|a\|_\epsilon \le \|a\|_2$. It is easy to see (see e.g. [L–T] section 3-2 for a proof) that the parameters $\|a\|_2$, $\|a\|_\epsilon$ govern the size of the tails of X. In particular, for some universal constant K, we have

$$(1.2) \qquad \begin{aligned} &P(\|X\| \ge t) \le \exp(-t^2/K\|a\|_2^2) && \text{if } t \le \|a\|_2^2/\|a\|_\epsilon; \\ &P(\|X\| \ge t) \le \exp(-t/K\|a\|_\epsilon) && \text{if } t \ge \|a\|_2^2/\|a\|_\epsilon. \end{aligned}$$

In looking for a generalization of Sudakov's minoration for chaos, it is thus natural to consider both norms $\|\cdot\|_2$ and $\|\cdot\|_\epsilon$.

THEOREM 1.1: *Consider a set A of $n \times n$ symmetric matrices. Denote by $N_2(A, \alpha)$ (resp. $N_\epsilon(A, \alpha)$) the smallest number of ball for $\|\cdot\|_2$ (resp. $\|\cdot\|_\epsilon$) of radius α needed to cover a set $A \subset \mathbb{R}^{n^2}$. Then, for some universal constant K, we have*

$$(1.3) \qquad \alpha(\log N_2(A, \alpha))^{1/4} \le K\ell^{(2)}(A).$$

Moreover, there exists a function $\varphi(x)$ such that $\lim_{x \to \infty} \varphi(x)/x^\delta = 0$ for each $\delta > 0$, and such that

$$(1.4) \qquad (\log N_\epsilon(A, \alpha))^{1/2} \le K \frac{\ell^{(2)}(A)}{\alpha} \varphi\left(\frac{\ell^{(2)}(A)}{\alpha}\right).$$

The reader must have noted the unusual exponent in (1.3); but (1.3) is optimal; a challenging conjecture is to know whether the term in φ can be removed in (1.4). The statement of the theorem is independent of the dimension; it is thus a routine to deduce from this theorem a statement on chaos processes of the form $(\sum_{i,j\geq 1} a_{i,j} g_i g_j)_{a\in A}$ for $A \subset \ell^2(\mathbb{N}^2)$.

It turns out that our proof of Theorem 1.1 makes essential use of methods of local theory of Banach spaces; thus it is natural to state it here in a finite-dimensional setting. Using (1.2), it is shown in [L–T] p. 327 that for some universal constant K, we have

$$\ell^{(2)}(A) \leq K \left(\int_0^\infty (\log N_2(A,\alpha))^{1/2} d\alpha + \int_0^\infty \log N_\epsilon(A,\alpha) d\alpha \right). \tag{1.5}$$

This inequality is the best possible of its type. (Of course one can replace entropy conditions by the corresponding majorizing measure conditions.) The exponents, however are not the same as in (1.3) and (1.4). The reason is that the value of $\ell^{(2)}(A)$ is not determined by the structure of A for the distances induced by $\|\cdot\|_2$ and $\|\cdot\|_\epsilon$. The question of which other simple parameter(s) to use to determine $\ell^{(2)}(A)$, or whether indeed such a parameter exists, is entirely open.

We now give an example showing that (1.3) is optimal. Take $A = \{a$ symmetric; $\|a\|_\epsilon \leq 1\}$. Thus, for $a \in A$,

$$\sum_{i,j\leq n} a_{i,j} g_i g_j \leq \sum_{i\leq n} g_i^2$$

and thus $\ell^{(2)}(A) \leq n$.

PROPOSITION 1.2: *For some constant independent of n, we have*

(i) $\log N_\epsilon\left(A, \frac{1}{2}\right) \geq cn^2$,

(ii) $\log N_2(A, c\sqrt{n}) \geq cn^2$.

Thus for $\alpha = 1/2$, $\alpha(\log N_\epsilon(A,\alpha))^{1/2}$ is of order $\ell^{(2)}(A)$ while for $\alpha = c\sqrt{n}$, $\alpha(\log N_2(A,\alpha))^{1/4}$ is of order $\ell^{(2)}(A)$.

Proof: (i) By volume considerations, $N_\epsilon(A, 1/2) \geq 2^{n(n-1)/2}$.

(ii) We give a simple probabilistic proof. Observe first that we can drop the requirement that A consists of symmetric matrices, by using the map that associates to each $n \times n$ matrix a the $2n \times 2n$ matrix

$$\begin{pmatrix} 0 & a \\ a^t & 0 \end{pmatrix}.$$

Consider an independent doubly indexed Rademacher sequence $(\epsilon_{i,j})_{i,j\leq n}$ (i.e. $P(\epsilon_{i,j}=1)=P(\epsilon_{i,j}=-1)=1/2$). Consider $h,h'\in\mathbb{R}^n$. Then, by the subgaussian inequality $P(|\sum\epsilon_i x_i|\geq t)\leq 2\exp(-t^2/2(\sum x_i^2))$, we see that

$$P\left(|\sum_{i,j\leq n} h_i h'_j \epsilon_{i,j}|\geq t\right)\leq 2\exp\left(-\frac{t^2}{2\|h\|_2^2\|h'\|_2^2}\right). \tag{1.6}$$

LEMMA 1.3: *(See e.g. [P] p. 56.) There exists a subset Z of $\mathbb{R}^n$, such that card $Z\leq 5^n$, $Z\subset 2B$, $B\subset \operatorname{conv} Z$, where B denotes the euclidean ball of $\mathbb{R}^n$, and conv Z the convex hull of Z.*

Thus, since card $Z\leq 5^n$

$$P(\forall h,h'\in Z,\ |\sum_{i,j\leq n} h_i h'_j\epsilon_{i,j}|\leq 16\sqrt{n})\geq\frac{1}{2}.$$

Thus it follows that

$$P(\|(\epsilon_{i,j})\|_\epsilon\leq 16\sqrt{n})\geq\frac{1}{2}.$$

Consider now a family $\eta_{i,j}, 1\leq i,j\leq n$, where $\eta_{i,j}=\pm 1$. Then $(|\epsilon_{i,j}-\eta_{i,j}|^2)_{i,j\leq n}$ is a sequence of independent random variables taking the values 0 and 4 with probability 1/2. Recentering, the subgaussian inequality implies

$$P\left(\sum_{i,j\leq n}|\epsilon_{i,j}-\eta_{i,j}|^2\leq n^2\right)\leq 2\exp -\frac{n^2}{8}.$$

So one needs at least $\frac{1}{4}\exp n^2/8$ balls of radius n for the norm $\|\cdot\|_2$ to cover $\{(\epsilon_{ij})_{i,j\leq n};\|(\epsilon_{ij})\|_\epsilon\leq 16\sqrt{n}\}\subset 16\sqrt{n}A$. Thus at least $\frac{1}{4}\exp n^2/8$ balls of radius $\sqrt{n}/16$ are needed to cover A. ∎

We discuss in more detail the methods and the organization of the paper. Consider another sequence $(g'_i)_{i\leq n}$ of independent $N(0,1)$ variables, independent of $(g_i)_{i\leq n}$. We first "decouple" the chaos and set

$$\bar{\ell}(A)=E\sup_{a\in A}|\sum_{i,j\leq n} a_{i,j}g_i g'_j|.$$

A symmetric bilinear functional $Q(x,y)$ satisfies

$$4Q(x,y)=Q(x+y,x+y)-Q(x-y,x-y).$$

Together with the fact that $(g_i \mp g_i')_{i\le n}$ is distributed like $(\sqrt{2}g_i)_{i\le h}$, this implies that $\bar{\ell}(A) \le \ell^{(2)}(A)$. So it is enough to prove (1.3) and (1.4) when $\ell^{(2)}(A)$ is replaced by $\bar{\ell}(A)$.

It is very tempting to look at the decoupled chaos

$$\sum_{i,j\le n} a_{i,j} g_i(\omega) g_j'(\omega')$$

as

$$\sum_{j\le n}\Big(\sum_{i\le n} a_{i,j} g_i(\omega)\Big) g_j'(\omega')$$

so that, conditionally on ω, this is a Gaussian random variable. Marcus and Pisier have proved a Sudakov-type minoration property for p-stable processes that relies on the fact that these variables are conditionally Gaussian [M–P]. In the p-stable case, an essential ingredient is that the random distance d_ω associated to this conditionally Gaussian process has the property that $P(d_\omega(s,t) < \epsilon)$ decrease very rapidly for $\epsilon \to 0$; this property is also crucial in the work that followed the work of Marcus and Pisier, e.g. [T2]. Unfortunately this property fails in the case of chaos, and another approach is needed. The essential idea is contained in an important result of local theory of Banach spaces due to A. Pajor and N. Tomczack-Jaegermann (following groundbreaking work of V. Milman). This is an improvement of Sudakov's minoration that, roughly speaking, states that the covering provided by (1.1) "needs only about $(\ell(A)/\alpha)^2$ dimensions". The version of this principle that we need is explained and proved in section 2. In section 3, we set up some (elementary) machinery. In Section 4 we prove (1.3). The principle of the proof is to find a set A' of $N \times k$ matrices, when k is of order $(\ell(A)/\alpha)^2, \bar{\ell}(A')$ of order $\ell(A)$, and $N_2(A,\alpha) \le N_2(A',\alpha/2)$. Reiterating the operation reduces to the case of $k \times k$ matrices, where the result follows from trivial volume considerations. In section 5, we prove (1.4). The proof uses a similar principle, but it is more delicate.

2. Random operators

For clarity in the proofs (and to avoid writing many triple sums) it will be useful to use somewhat more abstract notations than done in the introduction. Consider a finite-dimensional Hilbert space H; we denote by γ_H its canonical gaussian measure; if $(e_i)_{i\le n}$ is a basis of H, γ_H is the law of $\sum_{i\le n} g_i e_i$. For $x \in H$, we have

$$\|x\|_2^2 = \int_H \langle x, \omega\rangle^2 d\gamma_H(\omega). \tag{2.1}$$

It will be convenient to denote by ω, ξ, θ (possibly with lower indexes) random variables valued in various finite dimensional Hilbert spaces $G, H, \mathbb{R}^k$. It will always be assumed that when ω(resp. $\xi, \theta, \ldots$) is valued in H, its law is γ_H; and that all of these variables are independent. Thus we will write (2.1) as

$$\|x\|_2^2 = E\langle x, \omega\rangle^2. \tag{2.2}$$

Consider now $k > 0$, given, and an independent sequence $\theta_1, \ldots, \theta_k$ of r.v. distributed like θ. Central to this paper is the consideration of the random operator T^θ from H to $\mathbb{R}^k$ given by

$$T^\theta(x) = \frac{1}{\sqrt{k}}(\langle x, \theta_1\rangle, \ldots, \langle x, \theta_k\rangle).$$

For a subset A of H, we set $\ell(A) = E\sup_{a\in A}|\langle a, \omega\rangle|$. A first important property of T^θ is as follows.

PROPOSITION 2.1: $E(\ell(T^\theta(A))) \leq \ell(A)$.

Proof: Consider a r.v. ξ valued in $\mathbb{R}^k$ and distributed like $\gamma_{\mathbb{R}^k}$. Thus, by definition

$$\ell(T^\theta(A)) = E_\xi \sup_{a\in A}|\langle T^\theta(a), \xi\rangle|.$$

Here (as well as in the rest of the paper) E_ξ denotes expectation with respect to ξ. Using the Fubini theorem we get

$$E(\ell(T^\theta(A))) = E_\xi E_\theta \sup_{a\in A}|\langle T^\theta(a), \xi\rangle|.$$

We denote by $\xi^1, \ldots, \xi^k$ the components of ξ. Then

$$\langle T^\theta(a), \xi\rangle = \frac{1}{\sqrt{k}}\left\langle a, \sum_{i\leq k}\xi^i\theta_i\right\rangle.$$

Since, conditionally on ξ, $\sum_{i\leq k}\xi^i\theta_i$ is distributed like $\|\xi\|\theta$, we have

$$E_\theta \sup_{a\in A}|\langle T^\theta(a), \xi\rangle| = \frac{\|\xi\|}{\sqrt{k}}E_\theta\sup_{a\in A}|\langle a, \theta\rangle| = \frac{\|\xi\|}{\sqrt{k}}\ell(A).$$

The conclusion follows from

$$E\|\xi\| \leq (E\|\xi\|^2)^{1/2} = \left(\sum_{i\leq k} E(\xi^i)^2\right)^{1/2} = \sqrt{k}. \qquad \blacksquare$$

We denote by $K_1, K_2, \ldots$ universal constants. When there is no need to track the constant, we denote it by K (so the value of K may change at each occurrence). The fundamental property of T^θ is expressed in the following result.

THEOREM 2.2: *There exists a universal constant K such that, given a set $A \subset H$, the following event*

$$\forall x, y \in A, \quad \|x - y\| \leq K\left(\|T^\theta(x) - T^\theta(y)\| + \frac{\ell(A)}{\sqrt{k}}\right) \tag{2.3}$$

has probability $\geq 1 - 2e^{-k}$.

If we take $y = 0$, we see that $\|x\|_2 \leq K\ell(A)/\sqrt{k}$ whenever $T^\theta(x) = 0$. We thus recover the theorem of A. Pajor and N. Tomczak-Jaegermann that the sections of A by a random subspace of H of codimension k have a diameter of order $\ell(A)/\sqrt{k}$. It could however hardly be said that Theorem 2.2 is an extension of this result, since the proofs are identical. Our contribution here lies rather in the recognition that Theorem 2.2 is the correct formulation for our purposes.

For completeness we will give the proof of Theorem 2.2. We will follow the approach of [P], Theorem 5.8 (translated in our language).

We start with an elementary lemma.

LEMMA 2.3: *Given $x \in H$, we have, for $u > 0$,*

$$P(\|T^\theta(x)\| \leq u\|x\|) \leq (eu^2)^{k/2}.$$

Proof: By homogeneity we can assume $\|x\| = 1$. Then $\|T^\theta(x)\|^2$ is distributed like $(1/k)\sum_{i=1}^k g_i^2$. For every $\lambda > 0$,

$$E\exp\left(-\lambda\sum_{i=1}^k g_i^2\right) = (E\exp(-\lambda g_1^2))^k = \left(\frac{1}{1+2\lambda}\right)^{k/2}.$$

Therefore,

$$\begin{aligned} P(|T^\theta(x)|^2 < u^2) &= P\left(\exp\left(-\frac{\lambda}{k}\sum_{i=1}^k g_i^2\right) > \exp -\lambda u^2\right) \\ &\leq \left(\frac{1}{1+2\lambda/k}\right)^{k/2} \exp(\lambda u^2). \end{aligned}$$

The conclusion follows by taking $\lambda = k/2u^2$. ∎

Before we turn to the crucial point, let us recall one convenient form of the "concentration of measure" phenomenon for Gaussian measures [P], p. 47. If φ is a semi-norm on H, and $\sigma = \sup_{\|h\|\le 1} \varphi(h)$, then

$$P(|\varphi(\omega) - E\varphi(\omega)| \ge t) \le 2\exp -\frac{t^2}{2\sigma^2}. \tag{2.4}$$

(The use of the best constant in the exponent is essentially irrelevant.)

Lemma 2.4: *For $k \ge 1$, $t > 0$, we have*

$$P\left(\sup_{a\in A} \|T^\theta(a)\| \ge \frac{2\ell(A)}{\sqrt{k}} + t\right) \le 5^k \exp\left(-\frac{kt^2}{8D^2}\right)$$

where $D = \sup\{\|a\|; a \in A\}$.

Proof: By Lemma 1.3, there exists a subset Z of $\mathbf{R}^k$ with card $Z \le 5^k$, consisting of vectors of length ≤ 1, such that

$$\|T^\theta(x)\| \le 2 \sup_{h\in Z} |\langle T^\theta(x), h\rangle|.$$

Thus it suffices to show that for every $h \in Z$, $\|h\| \le 1$, we have

$$P(\sup_{a\in A} |\langle T^\theta(a), h\rangle| > \frac{\ell(A)}{\sqrt{k}} + \frac{u}{2}) \le \exp\left(-\frac{ku^2}{8D^2}\right). \tag{2.5}$$

As observed in the proof of Proposition 2.1, $\sup_{a\in A} |\langle T^\theta(a), h\rangle|$ is distributed like $(1/\sqrt{k}) \sup_{a\in A} |\langle a, \omega\rangle|$. So (2.5) follows from (2.4). ∎

We now conclude the proof of Theorem 2.2. We set $\epsilon = \ell(A)/\sqrt{k}$. By Sudakov's minoration 1.1, we can find a subset S of A of cardinality $\le \exp K_0^2 k$ such that the balls of radius ϵ centered at S cover A. Set $A - A = \{a - b; a, b \in A\}$ and

$$A' = \{x \in A - A; \|x\| \le \epsilon\}$$

so that $\ell(A') \le 2\ell(A)$. If we use Lemma 2.4 for A' instead of A, and for $t = 5\epsilon$, we see that

$$P\left(\sup_{a\in A'} \|T^\theta(a)\| \ge 9\epsilon\right) \le 5^k \exp(-3k) \le e^{-k}.$$

Consider now a number u_0 fixed such that $e^{2K_0^2}(eu_0^2)^{1/2} \leq 1/e$. It follows from Lemma 2.3 that

$$P(\forall s,t \in S, \quad \|T^\theta(s-t)\| \geq u_0\|s-t\|) \geq 1-e^{-k}.$$

It remains to show that (2.3) occurs whenever

$$\forall s,t \in S, \quad \|T^\theta(s-t)\| \geq u_0\|s-t\|, \tag{2.6}$$

$$\sup_{a\in A'} \|T^\theta(a)\| \leq 9\epsilon. \tag{2.7}$$

Indeed consider $x, y \in A$. We can find $s, t \in S$ such that $\|x-s\| \leq \epsilon$, $\|y-t\| \leq \epsilon$. Thus $x-s$, $y-t \in A'$. By (2.7) we have

$$\|T^\theta(x-s)\| \leq 9\epsilon; \|T^\theta(y-t)\| \leq 9\epsilon.$$

On the other hand, by (2.6) we have

$$\|s-t\| \leq \frac{1}{u_0}\|T^\theta(s-t)\|.$$

Thus

$$\begin{aligned}\|x-y\| \leq 2\epsilon + \|s-t\| &\leq 2\epsilon + \frac{1}{u_0}\|T^\theta(s-t)\| \\ &\leq 2\epsilon + \frac{1}{u_0}(18\epsilon + \|T^\theta(x-y)\|).\end{aligned}$$

This completes the proof. ∎

3. Tensor products

Consider two finite-dimensional Hilbert spaces G, H and the space $B(G,H)$ of bilinear forms on $G \times H$. For $a \in B(G,H)$, we set

$$\|a\|_2 = (Ea^2(\omega,\xi))^{1/2}, \tag{3.1}$$

$$\|a\|_\epsilon = \sup\{|a(h',h)| : \|h'\| \leq 1, \|h\| \leq 1\}. \tag{3.2}$$

Thus, if $G = \mathbb{R}^k$, $H = \mathbb{R}^n$, $a = (a_{i,j})_{i\leq k, j\leq n}$, we have $\|a\|_2 = (\sum_{i,j} a_{i,j}^2)^{1/2}$.

Given $x \in G$, we consider the operator V_x (resp. W_x) from $B(G,H)$ (resp. $B(G,\mathbb{R}^k)$) to H(resp. $\mathbb{R}^k$) given by

$$\forall y \in H, \quad \langle V_x(a), y\rangle = a(x,y) \tag{3.3}$$

(resp.

$$\forall y \in \mathbb{R}^k, \quad \langle W_x(a), y\rangle = a(x, y).) \tag{3.4}$$

Given an operator $T : H \to \mathbb{R}^k$, we consider the operator $\overline{T}$ from $B(G, H)$ to $B(G, \mathbb{R}^k)$, given, for $a \in B(G, H)$, $x \in G$, $y \in \mathbb{R}^k$ by

$$\overline{T}(a)(x, y) = a(x, T^t(y)), \tag{3.5}$$

where T^t is the adjoint of T.

LEMMA 3.1: *For $x \in G$ we have*

$$W_x \circ \overline{T} = T \circ V_x. \tag{3.6}$$

Proof: Consider $a \in B(G, H)$, $y \in H$. Then, by definition of W_x, by (3.5), and by the definition of V_x we have successively

$$\begin{aligned}\langle W_x \circ \overline{T}(a), y\rangle &= \overline{T}(a)(x, y)\\ &= a(x, T^t(y))\\ &= \langle V_x(a), T^t(y)\rangle\\ &= \langle T \circ V_x(a), y\rangle.\end{aligned}$$

∎

For $A \subset B(G, H)$, we set

$$\overline{\ell}(A) = E \sup_{a\in A} |a(\omega, \xi)|.$$

LEMMA 3.2: $\overline{\ell}(A) = E\ell(V_\omega(A))$.

Proof: This is just Fubini theorem and the definition of V_ω:

$$\begin{aligned}\overline{\ell}(A) &= E_\omega(E_\xi \sup_{a\in A} |a(\omega, \xi)|)\\ &= E_\omega(E_\xi \sup_{b\in V_\omega(A)} |\langle b, \xi\rangle|)\\ &= E_\omega(\ell(V_\omega(A))).\end{aligned}$$

∎

In a similar way, for $H = \mathbb{R}^n$, we have $\overline{\ell}(A) = E(\ell(W_\omega(A)))$.

LEMMA 3.3: $E\overline{\ell}(\overline{T}^\theta(A)) \leq \overline{\ell}(A)$.

Proof: Using successively Lemma 3.2, Lemma 3.1, the Fubini theorem, Proposition 2.1 and Lemma 3.2 again, we have

$$\begin{aligned} E\overline{\ell}(\overline{T}^\theta(A)) &= E_\theta E_\omega(\ell(W_\omega(\overline{T}^\theta(A)))) \\ &= E_\theta E_\omega(\ell(T^\theta(V_\omega(A)))) \\ &= E_\omega E_\theta(\ell(T^\theta(V_\omega(A)))) \\ &\leq E_\omega(\ell(V_\omega(A))) \\ &= \overline{\ell}(A). \end{aligned}$$ ∎

We finish this section with a few simple facts.

The following is well known, and is a weak version of the integrability property of Gaussian measure.

LEMMA 3.4: *Consider a semi-norm φ on H.*

Suppose that $P(\varphi(\omega) \leq M) \geq 1/2$. Then $(E\varphi^2(\omega))^{1/2} \leq KM$.

LEMMA 3.5:

(a) $P(\|V_\omega(a)\| \geq 3\|a\|_2) \leq 1/9$

(b) *If $P(\|V_\omega(a)\| \leq M) \geq 1/2$, then $\|a\|_2 \leq KM$.*

Proof: (a) Follows from the fact that $Ea^2(\omega,\xi) = \|a\|_2^2$, so that, by the Fubini theorem,

$$\begin{aligned} \|a\|_2^2 = Ea^2(\omega,\xi) &= E_\omega(E_\xi a^2(\omega,\xi)) \\ &= E_\omega(E_\xi\langle V_\omega(a),\xi\rangle^2) \\ &= E_\omega(\|V_\omega(a)\|^2). \end{aligned}$$

(b) We observe that $x \to \|V_x(a)\|$ is a semi-norm on G; so the result follows from Lemma 3.4 and (a).

4. The ℓ_2 norm

The basic fact is as follows.

PROPOSITION 4.1: *There exists a universal constant K with the following property. Consider a subset A of $B(G,H)$. Consider $k \geq 3$ such that $\epsilon/K \geq \overline{\ell}(A)/\sqrt{k}$. Then we can find a subset A' of $B(G,\mathbb{R}^k)$ with the following properties:*

$$\overline{\ell}(A') \leq 2\overline{\ell}(A), \tag{4.1}$$

$$N_2(A,\epsilon) \leq N_2(A',\epsilon/K). \tag{4.2}$$

Proof: Given ω, consider the following event S_ω (that depends on θ only):

$$(4.3) \qquad \forall x, y \in V_\omega(A), \quad \|x-y\| \le K_1 \left(\|T^\theta(x) - T^\theta(y)\| + \frac{\ell(V_\omega(A))}{\sqrt{k}} \right).$$

It follows from Theorem 2.2 that (for a suitable K) we have $P(S_\omega) \ge 1 - 2e^{-k}$. Using the Fubini theorem, we see that if we set

$$R = \{\theta; P(\theta \in S_\omega) \ge 1 - 6e^{-k}\},$$

we have $P(R) \ge 2/3$. On the other hand, from Lemma 3.3, we see that

$$P(\overline{\ell}(\overline{T}^\theta(A)) \le 2\overline{\ell}(A)) \ge 1/2.$$

Thus we can find θ such that $\overline{\ell}(\overline{T}^\theta(A)) \le 2\overline{\ell}(A)$, and $P(S) \ge 1 - 6e^{-k} \ge 3/4$, where $S = \{\omega; \theta \in S_\omega\}$. We fix such a θ, and we set $T = T^\theta$, $A' = \overline{T}^\theta(A)$. We thus have $\overline{\ell}(A') \le 2\overline{\ell}(A)$.

We now prove that, for some universal K_1,

$$(4.4) \qquad \forall a, b \in A, \quad \|a-b\|_2 \le K_1(\|\overline{T}(a) - \overline{T}(b)\|_2 + \overline{\ell}(A)/\sqrt{k}).$$

Indeed, from Lemma 3.5(a), we see that

$$P(\|W_\omega(\overline{T}(a) - \overline{T}(b))\| \le 3\|\overline{T}(a) - \overline{T}(b)\|) \ge 8/9.$$

On the other hand, for $\omega \in S$, and using Lemma 3.1, we have from (4.3) that

$$\begin{aligned}\|V_\omega(a) - V_\omega(b)\| &\le K(\|T(V_\omega(a)) - T(V_\omega(b))\| + \ell(V_\omega(A))/\sqrt{k}) \\ &= K(\|W_\omega(\overline{T}(a) - \overline{T}(b))\| + \ell(V_\omega(A))/\sqrt{k}).\end{aligned}$$

On the other hand, by Lemma 3.2 we have

$$P(\ell(V_\omega(A)) \le 8\overline{\ell}(A)) \ge 7/8.$$

Thus, with probability $\ge 1/2$, we have

$$\|V_\omega(a-b)\| \le K(3\|\overline{T}(a) - \overline{T}(b)\|_2 + 8\overline{\ell}(A)/\sqrt{k}).$$

Together with Lemma 3.5(b), this proves (4.4). Clearly (4.2) follows from (4.4) if $K \ge 2K_1$ (since $\overline{\ell}(A)/\sqrt{k} \le \epsilon/K$). ∎

We are now ready to prove (1.3). Consider $A \subset B(G,H)$. If we apply Proposition 4.1, and then apply it again after exchanging $\mathbb{R}^k$ and G, we see that there exists a universal constant K_2, such that if $\epsilon/K_2 \geq \bar{\ell}(A)/\sqrt{k}$, $k \geq 3$ then we can find a set $A'' \subset B(\mathbb{R}^k \times \mathbb{R}^k)$ such that

$$\bar{\ell}(A'') \leq 4\bar{\ell}(A), \tag{4.5}$$

$$N_2(A,\epsilon) \leq N_2(A'', \epsilon/K_2). \tag{4.6}$$

The ball (for $\|\cdot\|_2$) of radius 8ϵ in $B(\mathbb{R}^k \times \mathbb{R}^k)$ can be covered by at most $\exp(Kk^2)$ balls of radius ϵ/K_2. Thus if one chooses k of order $K_2^2\bar{\ell}(A)^2/\epsilon^2$, we see that

$$N_2(A,\epsilon) \leq \exp\left(\frac{K\bar{\ell}(A)^4}{\epsilon^4}\right) \cdot N_2(A'', 8\epsilon). \tag{4.7}$$

Set $h(\epsilon) = \sup\{N_2(A,\epsilon); \bar{\ell}(A) \leq 1\}$. It follows from (4.7), since $N_2(A'', 8\epsilon) = N_2(A''/4, 2\epsilon)$, that

$$h(\epsilon) \leq (\exp K\epsilon^{-4})h(2\epsilon). \tag{4.8}$$

We now observe that for $a \in A$ we have $E|a(\omega;\xi)| \leq \bar{\ell}(A)$. Since all the moments of a chaos are equivalent (see e.g. [B]) we have $\|a\|_2 = (Ea(\omega,\xi)^2)^{1/2} \leq K_3\bar{\ell}(A)$; thus $h(\epsilon) = 1$ for $\epsilon \geq K_3$. It then follows from (4.8) that $h(\epsilon) \leq \exp K\epsilon^{-4}$.

5. The operator norm

For a set $A \subset B(G,H)$, we set

$$m(A) = E_\xi \sup_{a \in A} \sup_{\|g\| \leq 1} |a(g,\xi)|.$$

Proposition 5.1: *There exists a universal constant K with the following property. Consider $\alpha > 0$, $A \subset B(G,H)$, and k such that $m(A)/\sqrt{k} \leq \alpha/K$. Then we can find $A' \subset B(G,\mathbb{R}^k)$ that satisfies*

$$\bar{\ell}(A') \leq 2\bar{\ell}(A), \tag{5.1}$$

$$N_\epsilon(A,\alpha) \leq N_\epsilon(A', \alpha/2K). \tag{5.2}$$

Proof: Consider the subset B of H given by

$$B = \{V_g(a); a \in A, \|g\| \leq 1\}.$$

Clearly $m(A) = \ell(B)$. Lemma 3.3 and Theorem 2.2 show that we can find θ such that, if we set $A' = \overline{T}^{\theta}(A), T = T^{\theta}$, then (5.1) holds, and moreover

$$\forall u, v \in B, \quad \|u - v\| \leq K\left(\|T(u) - T(v)\| + \frac{m(A)}{\sqrt{k}}\right).$$

Using this for $u = V_g(a)$, $v = V_g(b)$, and using Lemma 3.1 we see that

$$\begin{aligned}\|V_g(a) - V_g(b)\| &\leq K\left(\|W_g(\overline{T}(a)) - W_g(\overline{T}(b))\| + \frac{m(A)}{\sqrt{k}}\right)\\ &\leq K\left(\|\overline{T}(a) - \overline{T}(b)\|_\epsilon + \frac{m(A)}{\sqrt{k}}\right)\end{aligned}$$

so that

$$\|a - b\|_\epsilon \leq K\left(\|\overline{T}(a) - \overline{T}(b)\|_\epsilon + \frac{m(A)}{\sqrt{k}}\right).$$

This obviously finishes the proof. ∎

Consider a semi-norm φ on G. Let $b = \sup_{\|g\|\leq 1} \varphi(g)$. The theorem of Hahn–Banach shows that $\varphi(g) \geq |\langle z, g\rangle|$ for some $z \in G$, $\|z\| = b$. Thus

$$E|\langle z, \omega\rangle| = \sqrt{\frac{2}{\pi}}\|z\| \leq E\varphi(\omega)$$

and thus

$$b \leq \sqrt{\frac{\pi}{2}}E\varphi(\omega) \leq 2E\varphi(\omega). \tag{5.3}$$

Using this for the functional $\varphi(g) = \sup_{a\in A} |a(g, \xi)|$, we get that

$$\sup_{\|g\|\leq 1}\sup_{a\in A} |a(g, \xi)| \leq 2E_\omega \sup_{a\in A} |a(\omega, \xi)|.$$

This shows that $m(A) \leq 2\overline{\ell}(A)$.

Using this observation, as well as two times Proposition 5.1 we get the following:

PROPOSITION 5.2: *There exists a universal constant K with the following property. Consider $\alpha > 0$, $A \subset B(G, H)$, and k such that $\overline{\ell}(A)/\sqrt{k} \leq \alpha/K$. Then we can find $A'' \subset B(\mathbb{R}^k \times \mathbb{R}^k)$ that satisfies*

$$\overline{\ell}(A'') \leq 4\overline{\ell}(A), \tag{5.4}$$

$$N_\epsilon(A, \alpha) \leq N_\epsilon(A'', \alpha/K). \tag{5.5}$$

Observe that we still need of order $(\overline{\ell}(A)/\alpha)^4$ dimensions, so further reductions are necessary. These reductions will be made using Proposition 5.1; but we need to reduce the value of $m(A)$ by breaking A in pieces.

PROPOSITION 5.3: *There exists a universal constant K with the following property. Consider $A \subset B(\mathbb{R}^m \times \mathbb{R}^q)$ and $k \geq 1$. Then we can cover A by translates of sets $(A_i)_{i \leq N}$ such that*

$$\overline{\ell}(A_i) \leq 2\overline{\ell}(A), \tag{5.6}$$

$$E_\xi \sup_{a \in A_i} \sup_{\|g\| \leq 1} |a(g,\xi)| \leq \frac{K\overline{\ell}(A)}{\sqrt{k}}, \tag{5.7}$$

$$N \leq (K\sqrt{qk})^{qk}. \tag{5.8}$$

Proof: In order to avoid introducing new notation, we will replace (5.7) by

$$E_\omega \sup_{a \in A_i} \sup_{\|h\| \leq 1} |a(\omega, h)| \leq \frac{K\overline{\ell}(A)}{\sqrt{k}}. \tag{5.9}$$

As in the proof of Proposition 4.1, with $G = \mathbb{R}^m$, $H = \mathbb{R}^q$ we can find θ such that setting $T = T^\theta$, $A' = \overline{T}(A)$, we have $\overline{\ell}(A') \leq 2\overline{\ell}(A)$, and that with probability $\geq 1/2$, we have, for all $a, b \in A$,

$$\|V_\omega(a) - V_\omega(b)\| \leq K(\|W_\omega(\overline{T}(a)) - W_\omega(\overline{T}(b))\| + \frac{\overline{\ell}(A)}{\sqrt{k}}). \tag{5.10}$$

Consider now a set $C \subset \overline{T}(A)$, and let $D = \overline{T}^{-1}(C) \cap A$. It follows, from (5.10) and (the proof of) Lemma 3.5(b) that

$$E_\omega \sup_{a,b \in D} \sup_{\|h\| \leq 1} |a(\omega,h) - b(\omega,h)|$$

$$\leq K \left(E_\omega \sup_{u,v \in C} \sup_{\|h'\| \leq 1} |u(\omega,h') - v(\omega,h')| + \frac{\overline{\ell}(A)}{\sqrt{k}} \right).$$

Thus we have reduced to the proof of the following statement. If $A' \subset B(\mathbb{R}^m \times \mathbb{R}^k)$, we can cover A' by sets $(A_i)_{i \leq N}$ such that

$$E_\omega(\sup_{u,v \in A_i} \sup_{\|h'\| \leq 1} |u(\omega,h') - v(\omega,h')|) \leq \frac{\overline{\ell}(A')}{\sqrt{k}}$$

when N satisfies

$$N \leq (K\sqrt{mk})^{mk}.$$

We observe that, if $\|h'\| \leq 1$,

$$|u(\omega,h') - v(\omega,h')| \leq \|\omega\| \, \|u - v\|_\epsilon.$$

Since $E_\omega\|\omega\| \le (E\|\omega\|^2)^{1/2} = m^{1/2}$, it suffices to take for the sets A_i balls for $\|\cdot\|_\epsilon$ of radius $\le \bar{\ell}(A')/\sqrt{mk}$. We observe that by (5.4) we have

$$\sup_{a\in A}\|a\|_\epsilon = \sup_{a\in A}\sup_{\|g\|\le 1,\|h\|\le 1}|a(g,h)| \le \frac{\pi}{2}\bar{\ell}(A)$$

so the result follows from the (well known) following lemma.

Lemma 5.4: *Consider a convex balanced set U in $\mathbb{R}^p$. Then U can be covered by at most $(1+2/\alpha)^p$ translates of αU centered on U.*

We now denote

$$N(m,q,\alpha) = \sup\{N_\epsilon(A,\alpha\bar{\ell}(A)); \quad A \subset B(\mathbb{R}^m \times \mathbb{R}^q)\}.$$

Proposition 5.5: *There exists a universal constant K with the following property. If $p \le 1/\alpha^2$, then*

$$N(m,q,\alpha) \le (K\sqrt{qp})^{qp} N\left(m, \frac{K}{p\alpha^2}, \frac{\alpha}{K}\right).$$

Proof: Consider $A \subset B(\mathbb{R}^m \times \mathbb{R}^q)$. We first use Proposition 5.3 for $k = p$. It thus suffices to show that

$$N_\epsilon\left(A, \frac{\alpha}{2}\bar{\ell}(A)\right) \le N(m, K/p\alpha^2, \alpha/K)$$

whenever

$$m(A) = E_\xi \sup_{a\in A}\sup_{\|g\|\le 1}|a(g,\xi)| \le \frac{K\bar{\ell}(A)}{\sqrt{p}}.$$

We can find

$$k \le \frac{K}{p}\alpha^{-2}$$

such that

$$m(A)/\sqrt{k} \le K\bar{\ell}(A)/\sqrt{pk} \le \alpha\bar{\ell}(A)/2K_2$$

where K_2 denotes the constant of Proposition 5.1, so the claim follows from Proposition 5.1. ∎

We now proceed to the proof of (1.4). Consider $p_1 \le \cdots \le p_r \le \alpha^{-2}$. Set $q_1 = m$, $q_s = K^{1+(s-1)(s-2)}p_{s-1}^{-1}\alpha^{-2}$ for $s \le r$. Then, by iteration of Proposition 5.5, we get

$$N(m,m,\alpha) \le \prod_{s\le r}(Kq_s p_s)^{q_s p_s} N\left(m, \frac{K^{1+r(r-1)}}{p_2\alpha^2}, \frac{\alpha}{K^r}\right).$$

We specialize now to the case where p_s is (of order) u^s for some $u > 1$. Then for $s \leq r$, we have $q_s p_s \leq K^{r^2} u\alpha^{-2}$, so that

$$N(m,m,\alpha) \leq (K^{r^2} u\alpha^{-2})^{rK^{r^2} u\alpha^{-2}} N\left(m, \frac{K^{1+r(r-1)}}{u^r\alpha^2}, \frac{\alpha}{K^r}\right).$$

Take now $u = \alpha^{-2/r}$; observe that by Lemma 5.4 we have

$$N\left(m, K^{1+r(r-1)}, \frac{\alpha}{K^r}\right) \leq (K^r\alpha^{-1})^{K^{r^2} m}.$$

So we get (for a new constant K)

$$N(m,m,\alpha) \leq (K\alpha^{-1})^{K^{r^2}(\alpha^{-2-2/r}+m)}.$$

We recall that, by (5.5), we have

$$N(n,n,\alpha) \leq N(m,m,\alpha/K),$$

where $m \leq K\alpha^{-2}$; so, taking r of order $(\log \alpha)^{1/3}$, we get

$$N(n,n,\alpha) \leq \exp(K\alpha^{-2} \exp K(\log \alpha^{-1})^{2/3}).$$

This proves (1.4), where $\varphi(x) = K \exp K(\log x)^{2/3}$.

Remark: (1) This proof is a good illustration of the power of the "iteration method". While we start with a very weak principle (Proposition 5.3), the method yields a reasonably sharp result.

(2) It is possible (and even likely) that a more clever handling of this iteration-type argument would yield a smaller growth of the perburbation term; but in order to get rid entirely of this term, a radically new idea seems to be needed.

References

[B] C. Borell, *On polynomial chaos and integrability*, Prop. Math. Statist. **3** (1984), 191–203.

[M–P] M. Marcus and G. Pisier, *Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes*, Acta Math. **152** (1984), 245–301.

[L–T] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Springer-Verlag, Berlin, 1991.

[P] G. Pisier, *Volume Inequalities in the Geometry of Banach Spaces,* Cambridge University Press, 1989.

[P–T] A. Pajor and N. Tomczak-Jaegermann, *Subspaces of small codimension of finite dimensional Banach spaces,* Proc. Am. Math. Soc. **97** (1986), 637–642.

[T1] M. Talagrand, *Regularity of Gaussian processes,* Acta Math. **155** (1987), 99–149.

[T2] M. Talagrand, *Necessary conditions for sample boundedness of p-stable processes,* Annals Probab. **16** (1988), 1584–1595.