Acta Math., 175 (1995), 273–300

Sections of smooth convex bodies via majorizing measures

by

MICHEL TALAGRAND

C.N.R.S.

Dedicated to Xavier Fernique on his 60th birthday

1. Introduction

A central line of research in convexity theory and local theory of Banach spaces is the problem, given a balanced convex set, to find sections of large dimension that are well behaved. The basic theorem in this direction is Dvoretzky's theorem that asserts that an n-dimensional balanced convex set C has sections of dimension at least $\log n$ that are nearly ellipsoids. This is optimal in general. When more regularity is assumed (in the form of cotype hypothesis on the jauge of C) much larger nearly Euclidean sections can be found, as was demonstrated in the landmark paper [FLM]. In a somewhat different direction but in the same spirit is Milman's theorem [M] asserting the existence of subspaces of quotients of finite-dimensional Banach spaces that are nearly Euclidean and of dimension proportional to the dimension of the space. The nearly Euclidean sections constructed in [FLM] are obtained by a random construction, that provides no information on the "direction" of the section. There are however situations where this information is essential. A typical case arises from harmonic analysis, when one considers a finite family of characters $(\gamma_i)_{i\in I}$ on (say) a compact group, and the space E they generate. In that case, not all the subspaces of E are equally interesting; those that are generated by a subset of the characters $(\gamma_i)_{i\in I}$ are translation invariant and of special interest. The starting point of this research is a theorem of Bourgain that asserts that one can find a subset J of I, with $\operatorname{card} J=(\operatorname{card} I)^{2/p}$, such that on the space generated by the characters $(\gamma_i)_{i\in J}$, the L_p and L_2 norms are equivalent. (The basic measure is of course the normalised Haar measure.) Roughly speaking, what Bourgain proved is the following.

THEOREM 1.1 [B]. *Consider a sequence $(\varphi_i)_{i\leqslant n}$ of functions on $[0,1]$ that is orthogonal in L^2 and satisfies $\|\varphi_i\|_\infty\leqslant 1$ for each $i\leqslant n$. Consider $p>2$. Then, for most of the subsets I of $\{1,...,n\}$ with* $\operatorname{card} I=[n^{2/p}]$, *we have, for all numbers $(\alpha_i)_{i\in I}$,*

$$\Big\|\sum_{i\in I}\alpha_i\varphi_i\Big\|_p\leqslant K(p)\Big(\sum_{i\in I}\alpha_i^2\Big)^{1/2} \tag{1.1}$$

where $K(p)$ depends on p only.

It should be observed that the orthogonality in L^2 of the sequence (φ_i) shows that

$$\Big\|\sum_{i\in I}\alpha_i\varphi_i\Big\|_p\geqslant\Big\|\sum_{i\in I}\alpha_i\varphi_i\Big\|_2\geqslant\Big(\sum_{i\in I}\alpha_i^2\|\varphi_i\|_2^2\Big)^{1/2},$$

so that, when the numbers $\|\varphi_i\|_2$ are bounded below (independently of n) the L_p and L_2 norms are equivalent on the span of $(\varphi_i)_{i\in I}$.

Bourgain's proof of (1.1) is an extraordinary achievement and a masterpiece of technique. It however does not clearly show what is the role of the various hypotheses, in particular the orthogonality and the uniform boundedness of the sequence (φ_i). Moreover, it makes strong use of the special properties of the function $x\to x^p$, and Bourgain has to distinguish the cases $2<p<3$, $3<p<4$, $p>4$. The desire to clarify these intriguing features, and to produce a proof with a more transparent scheme, was at the origin of this paper. As could be expected, the special properties of the function $x\to x^p$ are inessential, and their seeming relevance in Bourgain's proof is an artifact of his approach. It turns out that the essential fact is simply that the L_p norm is 2-smooth (see the definition in (1.2) below). It is however considerably more surprising that the conditions of orthogonality and uniform boundedness of the sequence $(\varphi_i)_{i\leqslant n}$, that seem absolutely essential, play in fact only a very limited role, and that Bourgain's theorem is a simple consequence of a general principle. Let us recall that a norm $\|\cdot\|$ on a Banach space X is θ-smooth ($1<\theta\leqslant 2$) if for all vectors x,y in X, with $\|x\|=1$, $\|y\|\leqslant 1$, we have

$$\|x+y\|+\|x-y\|\leqslant 2+C\|y\|^\theta \tag{1.2}$$

where C is independent of x,y. Our main result is as follows:

THEOREM 1.2. *Consider vectors $(x_i)_{i\leqslant n}$ in a Banach space X, and set*

$$\tau=\sup\Big\{\sum_{i\leqslant n}x^*(x_i)^2 : x^*\in X^*,\ \|x^*\|\leqslant 1\Big\}.$$

Assume that there is another norm $\|\cdot\|_\sim$ on X, larger than $\|\cdot\|$, and such that $\|x_i\|_\sim\leqslant 1$ for each $i\leqslant n$. Assume that

$$\textit{the norm } \|\cdot\|_\sim \textit{ is 2-smooth.} \tag{1.3}$$

Consider $\varepsilon>0$ *and* $m=[n^{1-\varepsilon}/\tau]$. *Then, for most of the subsets* I *of cardinal* m, *we have, for all numbers* $(\alpha_i)_{i\in I}$,

$$\Big\|\sum_{i\in I}\alpha_i x_i\Big\| \leqslant K\Big(\sum_{i\in I}\alpha_i^2\Big)^{1/2} \tag{1.4}$$

where K *depends only on* ε *and on the constant implicit in* (1.3).

At first glance, the relationship between the two norms $\|\cdot\|$ and $\|\cdot\|_\sim$ is curious. A natural situation is where $\|\cdot\|_\sim$ is the original norm, and $\|\cdot\|$ is the new norm such that its dual ball is

$$\Big\{x^*\in X^*: \|x^*\|_\sim \leqslant 1, \sum_{i\leqslant n} x^*(x_i)^2 \leqslant \tau\Big\}.$$

Let us first explain the key point of Theorem 1.2. The definition of τ shows that there is x^* in $X_1^*=\{x^*\in X^*: \|x^*\|\leqslant 1\}$ with $\tau=\sum_{i\leqslant n} x^*(x_i)^2$. For most of the subsets I of $\{1,...,n\}$ of cardinal m, $a_I=\sum_{i\in I} x^*(x_i)^2$ will be of order $m\tau/n$. Now

$$a_I^{1/2} \leqslant \sup\Big\{\Big\|\sum_{i\in I}\alpha_i x_i\Big\| : \sum_{i\in I}\alpha_i^2 \leqslant 1\Big\}.$$

Thus, in order for (1.4) to hold, we must have $a_I\leqslant K^2$ for most I, i.e. $m\leqslant K^2 n/\tau$. The size of τ is thus a natural obstacle to how large I can be in (1.4). The rather unexpected content of Theorem 1.2 is that this is the only obstacle under (1.3), and that within the small loss $n^{-\varepsilon}$ we can achieve the optimal size. In geometrical terms, what (1.4) means is that the intersection of the unit ball of C of $\|\cdot\|$ with most of the subspaces generated by m vectors x_i contains large Euclidean balls (for the Euclidean structure generated by the vectors x_i). It is of interest to note that, in contrast with Dvoretsky's theorem, the Euclidean structure plays no special role here, and that if the norm $\|\cdot\|_\sim$ is simply assumed to be θ-smooth rather than 2-smooth, a suitably modified version of Theorem 1.2 remains true (Theorem 1.3 below).

The significant generality of Theorem 1.2 possibly indicates that an entire line of investigation has remained unexplored. Immediate questions raised by this result are whether (1.3) could be weakened (a natural assumption would be to assume that $\|\cdot\|_\sim$ is of type 2) and under which circumstances inequality (1.4) can be reversed. We have no answers to offer at this point.

Let us now explain the relationship between Theorem 1.2 and Bourgain's theorem. Let us take $X=L_p$, and for $\|\cdot\|=\|\cdot\|_\sim$ the norm of L_p, that is known to be 2-smooth. Taking $x_i=\varphi_i$, where $\|\varphi_i\|_\infty\leqslant 1$, and (φ_i) is orthogonal in L^2, it is simple to see that $\tau\leqslant n^{1-2/p}$ (Lemma 2.2). This is where and only where these two hypotheses really come in. Since, however, we cannot take $\varepsilon=0$ in Theorem 1.2, we cannot directly deduce Bourgain's theorem from Theorem 1.2. But in §2 we will show how to decompose naturally the L_p norm in the sum of two pieces. For one of these, the conclusion follows easily from a beautiful technique of Giné and Zinn. For the other, it follows from Theorem 1.2. This approach actually yields new information, and would allow to extend Bourgain's result to norm much more general than the L_p norm. But doing this would be routine and would make things appear more complicated than what they really are. To make the point that new information is obtained, we will simply show that (1.1) remains true when the L_p norm is replaced by the larger $L_{p,1}$ norm.

Before we discuss the methods and the contents of the paper, let us give a precise formulation of Theorem 1.2.

THEOREM 1.3. *Consider a number $1<\theta\leqslant 2$ and its conjugate $\varrho=\theta/(\theta-1)$. Consider vectors $(x_i)_{i\leqslant n}$ in a Banach space X, and the subset $\mathcal{F}$ of $\mathbf{R}^n$ given by*

$$\mathcal{F}=\{(|x^*(x_i)|^\varrho)_{i\leqslant n}:\|x^*\|\leqslant 1,\,x^*\in X^*\}.$$

Set

$$\tau=\sup\Big\{\sum_{i\leqslant n}f_i:f\in\mathcal{F}\Big\}.$$

Assume that there is another norm $\|\cdot\|_\sim$ on X, larger than $\|\cdot\|$, such that $\|x_i\|_\sim\leqslant 1$ for each $i\leqslant n$. Assume that

$$\text{the norm } \|\cdot\|_\sim \text{ is } \theta\text{-smooth.} \tag{1.5}$$

Consider $\varepsilon>0$, and consider an independent sequence $(\delta_i)_{i\leqslant n}$ of random variables with $\delta_i\in\{0,1\}$, $E\delta_i=\delta=1/(\tau(\log n)^K n^\varepsilon)$. Then

$$E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}\delta_i f_i\leqslant\frac{K}{\varepsilon}. \tag{1.6}$$

In the above (and the rest of the introduction), K is a number that depends only on the constant implicit in (1.5).

Remark. For $n\geqslant(K/\varepsilon)^{K/\varepsilon}$, we have $\delta\geqslant n^{-2\varepsilon}/\tau$.

To relate (1.4) and (1.6), we simply observe that for $x^*\in X_1^*$ we have

$$x^*\Big(\sum_{i\in I}\alpha_i x_i\Big)=\sum_{i\in I}\alpha_i x^*(x_i)\leqslant\Big(\sum_{i\in I}|\alpha_i|^\theta\Big)^{1/\theta}\Big(\sum_{i\in I}|x^*(x_i)|^\varrho\Big)^{1/\varrho}.$$

Taking the supremum over x^*, we get that

$$\Big\|\sum_{i\in I}\alpha_i x_i\Big\|\leqslant\Big(\sum_{i\in I}|\alpha_i|^\theta\Big)^{1/\theta}\Big(\sup_{f\in\mathcal{F}}\sum_{i\in I}f_i\Big)^{1/\varrho},$$

and (1.6) implies that this last term is controlled for $I=\{i\leqslant n:\delta_i=1\}$ for most of the choices of (δ_i).

The point of this formulation of Theorem 1.3 is to bring out its true nature: we have to bound the supremum of a large collection of random variables. Sharp probabilistic methods have been developed to do this. At some point, however, one has to prove a suitable smallness condition on the class $\mathcal{F}$, and this is where the link with the geometry of the situation will come in.

The proof of a statement such as (1.6) must start by a correct understanding of the tails of the random variables $\sum_{i\leqslant n}\delta_i f_i$, or, after recentering, $\sum_{i\leqslant n}(\delta_i-\delta)f_i$. The key point here is that writing $\|f\|_\infty=\sup_{i\leqslant n}|f_i|$, $\|f\|_2^2=\sum_{i\leqslant n}f_i^2$, we have

$$\forall u>0,\quad P\Bigl(\sum_{i\leqslant n}(\delta_i-\delta)f\geqslant u\Bigr)\leqslant \exp\Bigl(-\frac{u}{4\|f\|_\infty}\log\frac{u\|f\|_\infty}{\delta\|f\|_2^2}\Bigr). \tag{1.7}$$

In the range where this inequality (that goes back at least to Prokhorov) will be crucial, the log term will be of order $\log n$ and the inequality will look like

$$P\Bigl(\sum_{i\leqslant n}(\delta_i-\delta)f_i\geqslant u\Bigr)\leqslant 2\exp\Bigl(-\frac{u}{\|f\|_\infty}\log n\Bigr). \tag{1.8}$$

The $\log n$ factor plays a central role. What (1.8) also brings to light is the essential role of the supremum norm. The key steps of the proof are to gain a control of the size of $\mathcal{F}$ with respect to this norm. The most common way to gain such a control is via the growth of the covering numbers $N(\mathcal{F},\|\cdot\|_\infty,\varepsilon)$ where $N(\mathcal{F},\|\cdot\|_\infty,\varepsilon)$ is the smallest number of balls of $\mathbf{R}^n$ for the supremum norm of radius $\leqslant\varepsilon$ needed to cover $\mathcal{F}$. This is indeed essentially how the proof will start and in §3 we will prove the following weak version of (1.6):

$$E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}\delta_i f_i\leqslant\frac{K}{\varepsilon}\log n. \tag{1.9}$$

It is in the nature of the problem that the use of covering numbers does not allow one to go beyond (1.9). To improve upon (1.9) we need the sharper tool of majorizing measures, as a way to measure the size of $\mathcal{F}$ with respect to the supremum norm. Majorizing measures were first invented to provide upper bounds on the supremum of Gaussian processes [F], and later proved to be the correct way to characterise continuity and boundedness of these processes [T1] and of certain natural extensions [T5]. Majorizing measures bring, in principle, geometric information about the sets on which they live. In practice, however, the link with geometry is poorly understood, and is a reason why the construction of majorizing measures remains so difficult. The key point of our success in the present situation is that, in this situation, we have been able to establish a clear link with geometry. This link will allow in §4 to show that under the extra information

$$\sup_{f\in\mathcal{F}}\sum_{i\in I} f_i \leqslant B\log n \tag{1.10}$$

(where B is a parameter), the restriction of $\mathcal{F}$ to I is small in the appropriate majorizing measure sense, the smallness depending of course on the value of B. Once this key estimate is obtained, we consider independent random variables δ'_i valued in $\{0,1\}$ with $E\delta'_i = n^{-\varepsilon}$, and we prove in §5 (through general bounds on certain processes that are of independent interest) that, under (1.10), we have

$$E\sup_{f\in\mathcal{F}}\sum_{i\in I}\delta'_i f_i \leqslant K(1+B)/\varepsilon \tag{1.11}$$

which, when combining with (1.9), yield Theorem 1.3 (with a worse dependence on ε).

The crucial part of the argument can be stated as a result on operators that seems worthy to state in its own right.

THEOREM 1.4. *Consider $1<\theta\leqslant 2$ and a norm one operator T from l_θ^n into a Banach space X. Assume that*

$$\text{the norm of } X \text{ is } \theta\text{-smooth.} \tag{1.12}$$

Consider $0<\delta\leqslant 1$ and independent random variables $(\delta_i)_{i\leqslant n}$ with $\delta_i\in\{0,1\}$, $E\delta_i=\delta$, and denote by $(e_i)_{i\leqslant n}$ the canonical basis of l_θ^n. Denote by Z the norm of the restriction of T to the random subspace of l_θ^n generated by the vectors e_i for which $\delta_i=1$. Then

$$EZ\leqslant \frac{K}{(\log(1/\delta))^{1/\varrho}}\left(1+\sup_{i\leqslant n}\|T(e_i)\|(\log n)^{1/\varrho}\right) \tag{1.13}$$

where ϱ is the conjugate of θ and where K depends only on the constant implicit in (1.12).

Comments. (1) We will show that (1.13) is sharp.

(2) Theorem 1.4 is of special interest when $\|T(e_i)\|\leqslant K(\log n)^{-1/\varrho}$ for all $i\leqslant n$.

A last comment is in order. We have claimed that Theorem 1.3 cannot be proved using only covering numbers. Yet Bourgain did prove Theorem 1.1 using only covering numbers. He however uses in an essential way the fact that, as far as covering numbers are concerned, the slices of a certain ball are genuinely smaller than the ball itself, a fact that can also be seen as the ultimate foundation of our arguments.

Acknowledgement. The paper would not have been written without the insight and the generosity of Professor Gluskin, who suggested to the author that the methods of [T6] could possibly provide a new approach to Bourgain's theorem.

2. Giné and Zinn

In this section we will use tools from probability in Banach spaces to deduce Theorem 1.1 from Theorem 1.3. In order to make the point that Theorem 1.3 improves upon Theorem 1.1, we will prove Theorem 1.1 for the $L_{p,1}$ norm rather than the L_p norm (a fact that apparently cannot be obtained by Bourgain's approach that relies on special properties of the L_p norm). We fix $p>2$, and we denote by q its conjugate exponent. Throughout the section, we denote by $K(p)$ a constant that depends on p only, but may vary at each occurence. We denote by λ the Lebesgue measure on $[0,1]$, and we recall that the $L_{p,1}$ norm is (equivalent to)

$$\|f\|_{p,1}=\int_0^\infty \lambda(\{|f|\geqslant t\})^{1/p}\,dt$$

and its dual the $L_{q,\infty}$ norm is given by

$$\|f\|_{q,\infty}=\sup\{t\lambda(\{|f|\geqslant t\})^{1/q} : t\geqslant 0\}.$$

LEMMA 2.1. *Consider a function h with $\|h\|_{q,\infty}\leqslant 1$.*

$$\textit{If } \|h\|_\infty\leqslant A \textit{ then } \|h\|_2\leqslant K(p)A^{1-q/2}. \tag{2.1}$$

$$\textit{If } \lambda(\{|h|\neq 0\})\leqslant A^{-q} \textit{ then } \|h\|_1\leqslant K(p)A^{1-q}. \tag{2.2}$$

Proof. We have

$$\|h\|_2^2=\int_0^\infty 2t\lambda(\{|h|\geqslant t\})\,dt\leqslant 2\int_0^A t^{1-q}\,dt\leqslant \frac{2}{2-q}A^{2-q}$$

and

$$\|h\|_1=\int_0^\infty \lambda(\{|h|\geqslant t\})\,dt\leqslant A^{1-q}+\int_A^\infty t^{-q}\,dt\leqslant \frac{q}{q-1}A^{1-q}. \qquad \square$$

We consider functions φ_i as in Theorem 1.1.

LEMMA 2.2. *If $h\in L_{q,\infty}$, $\|h\|_{q,\infty}\leqslant 1$, then*

$$\sum_{i\leqslant n}\left(\int h\varphi_i\,d\lambda\right)^2\leqslant K(p)n^{1-2/p}. \tag{2.3}$$

Proof. Write $h'=h1_{\{|h|\leqslant n^{1/q}\}}$, $h''=h1_{\{|h|>n^{1/q}\}}$. It suffices to prove (2.3) when either $h=h'$ or $h=h''$. If $h=h'$, this follows from (2.1) with $A=n^{1/q}$ and the orthogonality of $(\varphi_i)_{i\leqslant n}$ since $2/q-1=1-2/p$. If $h=h''$, we simply observe that

$$\int h''\varphi_i\,d\lambda\leqslant \|h''\|_1\|\varphi\|_\infty\leqslant K(p)n^{1/q-1}$$

by (2.2), taking again $A=n^{1/q}$ and since $\lambda(\{|h|>n^{1/q}\})\leqslant n^{-1}$. □

We now show how to decompose the norm $\|\cdot\|_{p,1}$ in two pieces. Consider a number β such that $1/2q<\beta<1/q$, fixed once for all $\left(\text{e.g. } \beta=\frac{2}{3}q\right)$. We set $A=n^{\beta}$ and

$$\|f\|_s=\sup\left\{\int fh\,d\lambda : \|h\|_{q,\infty}\leqslant 1,\ \|h\|_\infty\leqslant A\right\},$$

$$\|f\|_b=\sup\left\{\int fh\,d\lambda : \|h\|_{q,\infty}\leqslant 1,\ \lambda(\{h\neq 0\})\leqslant A^{-q}\right\},$$

so that clearly

$$\|f\|_{p,1}\leqslant \|f\|_s+\|f\|_b. \tag{2.4}$$

We denote again by $\|\cdot\|_s$ and $\|\cdot\|_b$ the dual norms of $\|\cdot\|_s$ and $\|\cdot\|_b$.

We now show how to use Theorem 1.3.

PROPOSITION 2.3. *Consider i.i.d. random variables $(\delta_i)_{i\leqslant n}$ with $\delta_i\in\{0,1\}$, $E\delta_i=\delta=n^{2/p-1}$. Then*

$$E\sup\left\{\Big\|\sum_{i\leqslant n}\delta_i\alpha_i\varphi_i\Big\|_s : \sum_{i\leqslant n}\alpha_i^2\leqslant 1\right\}\leqslant K(p). \tag{2.5}$$

Proof. Since $1/p+1/q=1$, $\beta<1/q$, we observe that

$$\frac{2}{p}-1+\beta(2-q)<\frac{2}{p}-1+\left(\frac{2}{q}-1\right)=0,$$

so that we can find $q'<q$ (depending on p only) such that

$$\varepsilon=-\left(\frac{2}{p}-1+\beta(2-q')\right)>0.$$

We denote by p' the conjugate exponent of q' so that $p'>1$.

We will apply Theorem 1.3 with $\|\cdot\|=\|\cdot\|_s$, $\|\cdot\|_\sim=\|\cdot\|_{p'}$ (which is 2-smooth by classical results [LiTz]). It follows from (2.1) and the orthogonality of $(\varphi_i)_{i\leqslant n}$ that

$$\tau=\sup\left\{\sum_{i\leqslant n}x^*(\varphi_i)^2 : \|x^*\|_\sim\leqslant 1\right\}\leqslant K(p)A^{2-q'}.$$

Thus

$$\delta\tau\leqslant K(p)n^{2/p-1+\beta(2-q')}=K(p)n^{-\varepsilon}$$

and Theorem 1.3 indeed applies. □

The rest of this section is devoted to the proof that

$$E\sup\left\{\Big\|\sum_{i\leqslant n}\delta_i\alpha_i\varphi_i\Big\|_b : \sum_{i\leqslant n}\alpha_i^2\leqslant 1\right\}\leqslant K(p). \tag{2.6}$$

Combining with (2.4) we will then get

$$E\sup\left\{\left\|\sum_{i\leqslant n}\delta_i\alpha_i\varphi_i\right\|_{p,1}:\sum_{i\leqslant n}\alpha_i^2\leqslant 1\right\}\leqslant K(p), \tag{2.7}$$

an improved version of Bourgain's theorem.

The proof of (2.6) is comparatively easy. It will follow a very beautiful scheme of proof invented by Giné and Zinn in [GZ]. While a posteriori simple, this scheme is extremely efficient, and has proved to be of considerable importance. It was first applied in Banach space theory in [T2], where it was unfortunately not clearly attributed to its authors. The method is also a key ingredient in the papers [BT] (upon seeing [T2]) and [T3].

LEMMA 2.4. *Consider a subset $\mathcal{F}$ of $\mathbf{R}^n$, and set*

$$\tau=\sup\left\{\sum_{i\leqslant n}f_i : f=(f_i)_{i\leqslant n}\in\mathcal{F}\right\}.$$

Then

$$E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}\delta_i f_i\leqslant\delta\tau+E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}(\delta_i-\delta)f_i.$$

Proof. Write

$$\sum_{i\leqslant n}\delta_i f_i\leqslant\delta\left(\sum_{i\leqslant n}f_i\right)+\sum_{i\leqslant n}(\delta_i-\delta)f_i\leqslant\delta\tau+\sum_{i\leqslant n}(\delta_i-\delta)f_i,$$

and take the supremum over f and expectation. □

We now consider an independent sequence of Bernoulli r.v., i.e.

$$P(\varepsilon_i=1)=P(\varepsilon_i=-1)=\tfrac{1}{2},$$

that is independent of all other sequences considered.

LEMMA 2.5. $E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}(\delta_i-\delta)f_i\leqslant 2\sup_{f\in\mathcal{F}}\left|\sum_{i\leqslant n}\delta_i\varepsilon_i f_i\right|$.

Proof. Consider an independent sequence $(\delta_i')_{i\leqslant n}$ distributed like $(\delta_i)_{i\leqslant n}$, and independent of all other sequences. Then

$$E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}(\delta_i-\delta)f_i\leqslant E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}(\delta_i-\delta_i')f_i=E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}\varepsilon_i(\delta_i-\delta_i')f_i$$

by symmetry. Now, by the triangle inequality, this last term is bounded by

$$E\sup_{f\in\mathcal{F}}\left|\sum_{i\leqslant n}\varepsilon_i\delta_i f_i\right|+E\sup_{f\in\mathcal{F}}\left|\sum_{i\leqslant n}\varepsilon_i\delta_i' f_i\right|=2E\sup_{f\in\mathcal{F}}\left|\sum_{i\leqslant n}\varepsilon_i\delta_i f_i\right|.$$ □

To prove (2.6) we have to prove that

$$E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}\delta_i f_i\leqslant K(p)$$

where

$$\mathcal{F}=\{((x^*(\varphi_i)^2)_{i\leqslant n}:\|x^*\|_b\leqslant 1\}.$$

It follows from Lemma 2.2 that

$$\tau=\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}f_i\leqslant K(p)n^{1-2/p}.$$

Thus combining Lemmas 2.4 and 2.5 we are reduced to prove that

$$E\sup_{f\in\mathcal{F}}\left|\sum_{i\leqslant n}\delta_i\varepsilon_i f_i\right|\leqslant K(p). \tag{2.8}$$

To prove this, we will work conditionally on $(\delta_i)_{i\leqslant n}$.

LEMMA 2.6. *For a subset I of $\{1,\dots,n\}$, we have*

$$E\sup_{f\in\mathcal{F}}\left|\sum_{i\in I}\varepsilon_i f_i\right|\leqslant K(p)A^{1-q}E\left\|\sum_{i\in I}\varepsilon_i\varphi_i\right\|_b. \tag{2.9}$$

Proof. Consider the subset $\mathcal{G}$ of $\mathbf{R}^n$ given by

$$\mathcal{G}=\{(x^*(\varphi_i))_{i\leqslant n}:\|x^*\|_b\leqslant 1\}$$

so that

$$E\left\|\sum_{i\in I}\varepsilon_i\varphi_i\right\|_b=E\sup_{f\in\mathcal{G}}\left|\sum_{i\in I}\varepsilon_i f_i\right|.$$

Now, we go from $\mathcal{G}$ to $\mathcal{F}$ by taking the square of each component, and it follows from the comparison theorem for Bernoulli processes [T4, Theorem 2.1] that

$$E\sup_{f\in\mathcal{F}}\left|\sum_{i\in I}\varepsilon_i f_i\right|\leqslant 4\sup_{i\in I}\|\varphi_i\|_b E\left\|\sum_{i\in I}\varepsilon_i\varphi_i\right\|_b.$$

But $\|\varphi_i\|_b\leqslant K(p)A^{1-q}$ by (2.2). □

We now turn to the estimation of $E\left\|\sum_{i\in I}\varepsilon_i\varphi_i\right\|_b$. We recall the norm $\|\cdot\|_{\psi_2}$ given by

$$\|f\|_{\psi_2}=\inf\left\{c>0:\int\exp\left(\frac{f}{c}\right)^2 d\lambda\leqslant 2\right\}.$$

LEMMA 2.7. $E\left\|\sum_{i\in I}\varepsilon_i\varphi_i\right\|_{\psi_2}\leqslant K\sqrt{\operatorname{card} I}$.

Proof. The key is the subgaussian inequality

$$P\left(\left|\sum_{i\in I}\varepsilon_i a_i\right|\geqslant t\right)\leqslant 2\exp\left(-\frac{t^2}{2\sum_{i\in I}a_i^2}\right)$$

for all numbers $(a_i)_{i\in I}$, (see [LT, p. 90]). By Fubini's theorem, this implies that

$$E\int\exp\frac{\left(\sum_{i\in I}\varepsilon_i\varphi_i\right)^2}{3\operatorname{card} I}\,d\lambda\leqslant K.$$

The conclusion now follows from the fact that, since for $u\geqslant 1$ we have $e^{x^2/u}\leqslant 1+e^{x^2}/u$, we have $\|f\|_{\psi_2}\leqslant\int\exp f^2\,d\lambda$. $\square$

LEMMA 2.8. $\|f\|_b\leqslant K(p)A^{1-q}\sqrt{\log A}\,\|f\|_{\psi_2}$.

This amounts to prove by duality that if $\|h\|_{q,1}\leqslant 1$, $\lambda(\{|h|\neq 0\})\leqslant A^{-q}$, the norm of h in the Orlicz space $L\sqrt{\log L}$ is at most $K(p)A^{1-q}\sqrt{\log A}$, an elementary fact. $\square$

Combining Lemmas 2.7 and 2.8, we see that

$$E\left\|\sum_{i\in I}\varepsilon_i\varphi_i\right\|_b\leqslant K(p)A^{1-q}\sqrt{\log A}\sqrt{\operatorname{card} I},$$

so that by Lemma 2.6

$$E\sup_{f\in\mathcal{F}}\left|\sum_{i\in I}\varepsilon_i f_i\right|\leqslant K(p)A^{2(1-q)}\sqrt{\log A}\sqrt{\operatorname{card} I}$$

and thus, since for $I=\{i\leqslant n:\delta_i=1\}$, we have $E\sqrt{\operatorname{card} I}\leqslant\sqrt{E\operatorname{card} I}\leqslant\sqrt{n\delta}$, we get

$$E\sup_{f\in\mathcal{F}}\left|\sum_{i\leqslant n}\delta_i\varepsilon_i f_i\right|\leqslant K(p)A^{2(1-q)}\sqrt{\log A}\sqrt{n\delta}.$$

The right-hand side is

$$K(p)\sqrt{\beta}\,n^{2\beta(1-q)+1/p}\sqrt{\log n}$$

so that (2.6) is proved since $\beta>1/2q$ and hence $2\beta(1-q)+1/p<0$.

The proof of (2.7) is complete.

It should be pointed out that our approach to Theorem 1.1 does bring more information than (2.7) actually shows. If q' is such that for some $2/q<\beta<1/q$ we have $\beta(2-q')<1-2/p$ (so that β can be found whenever $q'>(3p-2)/2(p-1)$) then for most subsets I of $\{1,\dots,n\}$ of cardinal $n^{2/p}$, we have, if $\|h\|_{q',\infty}\leqslant 1$, $\|h\|_\infty\leqslant A$, that

$$\left|\int\left(\sum_{i\in I}\alpha_i\varphi_i\right)h\,d\lambda\right|\leqslant K(p,q')\left(\sum_{i\in I}\alpha_i^2\right)^{1/2},$$

or equivalently, by duality, that if $\sum_{i\in I}\alpha_i^2\leqslant 1$, then $\sum_{i\in I}\alpha_i\varphi_i$ is the sum of a function of $\|\cdot\|_{p',1}$ norm at most $K(p,q')$ and of a function of L_1 norm at most $K(p,q')A^{-1}$.

It would be interesting to know how far one can go in this direction.

3. Within $\log n$

In the rest of the paper we are in the setting of Theorem 1.3. We fix $1<\theta\leqslant 2$, and we assume that the norm $\|\cdot\|_\sim$ is θ-smooth, so that, for some constant C,

$$\forall x,y\in X,\ \|x\|_\sim=1,\ \|y\|_\sim\leqslant 1 \quad\Longrightarrow\quad \|x+y\|_\sim+\|x-y\|_\sim\leqslant 2+C\|y\|_\sim^\theta. \tag{3.1}$$

Throughout the rest of the paper, we make the following convention. We denote by K a universal constant, that may vary at each occurence, and by $K(C)$ a constant depending only on θ, C, and that may vary at each occurence. We will denote by ϱ the conjugate exponent of θ.

We consider the subset $\mathcal{F}$ of $\mathbf{R}^n$ given by

$$\mathcal{F}=\{(|x^*(x_i)|^\varrho)_{i\leqslant n}:\|x^*\|\leqslant 1\},$$

and we keep the notation

$$\tau=\sup_{f\in\mathcal{F}}\sum_{i\leqslant n} f_i.$$

The aim of this section is the following.

THEOREM 3.1. *Consider $\varepsilon>0$ and an i.i.d. sequence $(\delta_i)_{i\leqslant n}$ of $\{0,1\}$ valued r.v., with $E\delta_i=\delta=n^{-\varepsilon}/\tau$.*

Then

$$E\sup_{f\in\mathcal{F}}\sum_{i\leqslant n}\delta_i f_i\leqslant\frac{K(C)}{\varepsilon}\log n.$$

We now consider the norm $\|\cdot\|_\infty$ on X^* given by $\|x^*\|_\infty=\max_{i\leqslant n}|x^*(x_i)|$. The key to Theorem 3.1 is to gain control of the covering numbers $N(X_1^*,\|\cdot\|_\infty,\varepsilon)$. This will be done by duality. We denote by U the balanced convex hull of the vectors $(x_i)_{i\leqslant n}$.

LEMMA 3.2. $\log N(U,\|\cdot\|_2,\varepsilon)\leqslant\dfrac{K(C)}{\varepsilon^\varrho}\log n.$

Proof. It is a great pleasure to reproduce this argument of Maurey. Consider $x\in U$, so that $x=\sum_{i\leqslant n}\alpha_i x_i$ with $\sum_{i\leqslant n}|\alpha_i|\leqslant 1$. Consider the X-valued r.v. Y given by

$$P(Y=(\operatorname{sign}\alpha_i)x_i)=|\alpha_i|,$$
$$P(Y=0)=1-\sum_{i\leqslant n}|\alpha_i|,$$

so that $E(Y)=x$. Since $(X,\|\cdot\|_\sim)$ is θ-smooth it is of type θ, with a type θ constant depending only on C, so that, if $(Y_j)_{j\leqslant k}$ denotes an independent sequence distributed like Y, we have

$$E\left(\left\|k^{-1}\sum_{j\leqslant k}(Y_j-x)\right\|\right)\leqslant K(C)\left(k^{-\theta}\sum_{j\leqslant k}E\|Y_j-x\|_\sim^\theta\right)^{1/\theta}\leqslant K(C)k^{-1/\varrho} \tag{3.2}$$

since $\|Y_j - x\|_{\sim} \leqslant 2$. Thus, if $k \geqslant (K(C)/\varepsilon)^{\varrho}$, the right-hand side of (3.2) is $\leqslant \varepsilon$. In particular there is one realisation of the variables Y_j for which $\|x - k^{-1}\sum_{j\leqslant k} Y_j\|_{\sim} \leqslant \varepsilon$. But there are at most $(n+1)^k$ such realisations, so that

$$N(U, \|\cdot\|_2, \varepsilon) \leqslant (n+1)^k. \qquad \square$$

LEMMA 3.3. $\log N(X_1^*, \|\cdot\|_{\sim}, \varepsilon) \leqslant \dfrac{K(C)}{\varepsilon^{\varrho}} \log n$.

Proof. Since $\|\cdot\|_{\sim} \geqslant \|\cdot\|$, the unit ball B of X^* for the dual norm of $\|\cdot\|_{\sim}$ contains X_1^*, so it suffices to show that

$$\log N(B, \|\cdot\|_{\infty}, \varepsilon) \leqslant K(C) \frac{\log n}{\varepsilon^{\varrho}}. \tag{3.3}$$

Since the norm $\|\cdot\|_{\sim}$ is θ-smooth, the dual norm is uniformly convex [LiTz]. Then (3.3) follows from Lemma 3.2 and Proposition 2 (ii) of [BPST] (combined with iteration). $\square$

For $k \geqslant 1$, and $x^* \in X_1^*$, we define

$$f_{i,k}(x^*) = |x^*(x_i)|^{\varrho} 1_{\{2^{-k} < |x^*(x_i)| \leqslant 2^{-(k-1)}\}}.$$

We note that, in order to prove Theorem 3.1, it suffices to prove that for each k,

$$E \sup_{x^* \in X_1^*} \sum_{i \leqslant n} \delta_i f_{i,k}(x^*) \leqslant \frac{K(C)}{\varepsilon} \tag{3.4}$$

(indeed, $\sum_{i\leqslant n} \delta_i f_{i,k}(x^*) \leqslant 2^{-(k-1)\varrho} n$, so only about $\varrho^{-1} \log n$ values of k matter).

An essential ingredient in the proof of (3.4) is a special case of the Prokhorov–Benett inequality. The inequalities to be found in the literature are more precise than what we need, and the extra precision is confusing. For the convenience of the reader, we prove what we need.

PROPOSITION 3.4. *Consider a r.v. Z with $|Z| \leqslant 1$, $EZ = 0$, $EZ^2 \leqslant \delta$. Consider independent copies $(Z_i)_{i\leqslant n}$ of Z, and a sequence $a = (a_i)_{i\leqslant n}$ of numbers. Then for all $t > 0$, if $\|a\|_{\infty} \leqslant a_{\infty}$, $\|a\|_2^2 \leqslant a_2^2$, we have*

$$P\Big(\sum_{i\leqslant n} Z_i \geqslant t\Big) \leqslant \exp\Big(-\frac{t}{4a_{\infty}} \log \frac{t a_{\infty}}{\delta a_2^2}\Big). \tag{3.5}$$

Proof. We start with the elementary inequality

$$e^x \leqslant 1 + x + \tfrac{1}{2} x^2 e^{|x|}$$

(that is obvious on power series expansion) to obtain, for $\lambda \geqslant 0$,

$$E \exp \lambda Z_i \leqslant 1+\tfrac{1}{2}\lambda^2 \delta e^{\lambda} \leqslant \exp \tfrac{1}{2}\lambda^2 \delta e^{\lambda}.$$

Thus if $Y=\sum_{i\leqslant n} a_i Z_i$, by independence we have

$$E \exp \lambda Y \leqslant \exp \tfrac{1}{2}\lambda^2 a_2^2 \delta e^{\lambda a_\infty} \leqslant \exp \frac{\lambda a_2^2 \delta}{2a_\infty} e^{2\lambda a_\infty}$$

using the inequality $x \leqslant e^x$ for $x=\lambda a_\infty$. Thus for any $\lambda \geqslant 0$,

$$P(Y \geqslant t) \leqslant \exp\left(-\lambda t+\lambda \frac{a_2^2 \delta}{2a_\infty} e^{2\lambda a_\infty}\right),$$

and we take

$$\lambda=\frac{1}{2a_\infty} \log \frac{t a_\infty}{\delta a_2^2}$$

if $\lambda \geqslant 0$ (there is nothing to prove otherwise). □

Comments. (1) When $t\|a\|_\infty/\delta\|a\|_2^2 \leqslant 2$, the bound (3.5) is inefficient (the correct exponent is then $-t^2/(K\delta\|a\|_2^2)$ but we will never use it in that range.

(2) We will use (3.5) only for $Z_i=\delta_i-\delta$.

COROLLARY 3.5. *For $u \geqslant 2\delta \operatorname{card} I$,*

$$P\left(\sum_{i\in I} \delta_i \geqslant u\right) \leqslant \exp\left(-\frac{u}{8} \log \frac{u}{2\delta \operatorname{card} I}\right). \tag{3.6}$$

Proof. Take $Z_i=\delta_i-\delta$; observe that the left-hand side is

$$P\left(\sum_{i\in I}(\delta_i-\delta) \geqslant u-\delta \operatorname{card} I\right) \leqslant P\left(\sum_{i\in I}(\delta_i-\delta) \geqslant \tfrac{1}{2}u\right)$$

and use (3.5) with $a_\infty=1$, $a_2^2=\operatorname{card} I$, $t=\tfrac{1}{2}u$. □

We now turn to the proof of (3.4). We fix k, and given $x^* \in X_1^*$, we set

$$I(x^*)=\{i \leqslant n : |x^*(x_i)| \geqslant 2^{-k-1}\}.$$

Thus

$$2^{-(k+1)\varrho} \operatorname{card} I(x^*) \leqslant \tau$$

and

$$\delta \operatorname{card} I(x^*) \leqslant \delta\tau 2^{(k+1)\varrho} \leqslant 2^{(k+1)\varrho} n^{-\varepsilon},$$

so that by (3.6), for $t \geqslant 1$ we have for any x^* in X^*,

$$P\Big(\sum_{i\in I(x^*)} \delta_i \geqslant 2t2^{(k+1)\varrho}\Big) \leqslant \exp\big(-\tfrac{1}{4}\varepsilon t 2^{(k+1)\varrho}\log n\big). \tag{3.7}$$

We fix $t \geqslant 1$, and we set

$$\alpha = P\Big(\sup_{x^*\in X_1^*} \sum_{i\leqslant n} \delta_i f_{i,k}(x^*) > 2^{1+2\varrho}t\Big). \tag{3.8}$$

Consider the largest integer j_0 such that

$$j_0 \exp\big(-\tfrac{1}{4}\varepsilon t 2^{(k+1)\varrho}\log n\big) \leqslant \alpha. \tag{3.9}$$

The main step is the construction, for $j \leqslant j_0$, of points $x_j^* \in X_1^*$ with the following property:

$$\text{if } l<j, \quad \exists i \leqslant n, \quad |x_j^*(x_i)| > 2^{-k}, \quad |x_l^*(x_i)| \leqslant 2^{-k-1}. \tag{3.10}$$

This shows that $\|x_j^* - x_l^*\|_\infty \geqslant 2^{-k-1}$, so that

$$j_0 \leqslant N(X_1^*, \|\cdot\|_\infty, 2^{-k-1}) \leqslant \exp(K(C)2^{k\varrho}\log n),$$

and thus

$$\alpha \leqslant 2\exp\big(-\big(\tfrac{1}{4}\varepsilon t 2^{(k+1)\varrho} - K(C)2^{k\varrho}\big)\log n\big).$$

Combined with (3.8), this proves (3.4) by a routine computation.

The construction of the vectors (x_j^*) is by induction as follows. Having constructed $x_1^*, \dots, x_j^*$ for $j<j_0$, we combine (3.7) with (3.9) to see that we can find a realisation $(\delta_i)_{i\leqslant n}$ such that

$$\sup_{x^*\in X_1^*} \sum_{i\leqslant n} \delta_i f_{i,k}(x^*) > 2^{1+2\varrho}t, \tag{3.11}$$

while, for $l \leqslant j$,

$$\sum_{i\in I(x_l^*)} \delta_i \leqslant 2t2^{(k+1)\varrho}. \tag{3.12}$$

Consider then x_{j+1}^* such that

$$\sum_{i\leqslant n} \delta_i f_{i,k}(x_{j+1}^*) > 2^{1+2\varrho}t.$$

Since $f_{i,k}(x^*) \leqslant 2^{-\varrho(k-1)}$, and since $f_{i,k}(x^*) \neq 0 \Rightarrow |x^*(x_i)| > 2^{-k}$, we see that

$$\operatorname{card}\{i \leqslant n : \delta_i = 1, |x_{j+1}^*(x_i)| > 2^{-k}\} > 2t2^{\varrho(k+1)}.$$

On the other hand, by (3.12), we have, for $l \leqslant j$,

$$\operatorname{card}\{i \leqslant n : \delta_i = 1, |x_l^*(x_i)| \geqslant 2^{-k-1}\} \leqslant 2t2^{\varrho(k+1)}$$

so that (3.10) is obvious. □

4. Majorizing measures

Let us recall the traditional definition of majorizing measures. Given a metric space (T,d), a number $\alpha\geqslant 1$, and an (atomic) probability measure μ on T, we set

$$\gamma_\alpha(T,d,\mu)=\sup_{t\in T}\int_0^\infty\left(\log\frac{1}{\mu(B(t,\varepsilon))}\right)^{1/\alpha}d\varepsilon \tag{4.1}$$

where $B(t,\varepsilon)$ denotes the ball of radius ε centered at t. It is good to observe that the integral is in fact only between 0 and the diameter of T. We set

$$\gamma_\alpha(T,d)=\inf\gamma_\alpha(T,d,\mu)$$

where the infimum is taken over all probability measures.

The aim of this section is to prove the following.

THEOREM 4.1. *Assume that the norm $\|\cdot\|_\sim$ of the Banach space X satisfies* (3.1). *Consider vectors $(x_i)_{i\leqslant m}$ of X, such that $\|x_i\|_\sim\leqslant 1$ for each $i\leqslant m$. Consider a number A, and the subset T_A of $\mathbf{R}^m$ given by*

$$T_A=\left\{(|x^*(x_i)|^\varrho)_{i\leqslant m}:\|x^*\|_\sim\leqslant 1,\ \sum_{i\leqslant m}|x^*(x_i)|^\varrho\leqslant A\right\}.$$

Consider, on T_A, the distance d_∞ induced by the norm $\|\cdot\|_\infty$. Then

$$\gamma_1(T_A,d_\infty)\leqslant K(C)(A+\log m). \tag{4.2}$$

The most powerful idea about majorizing measures is that the "size" of a metric space with respect to the existence of majorizing measures can be measured by the "size" of the well separated subsets it contains [T1]. Successive elaborations of this idea have led to an abstract principle where the idea of separation is somewhat hidden. We state here the case of the principle we need. This result follows from [T5, §§ 1 and 2].

THEOREM 4.2. *Consider a metric space (T,d) and a number $r\geqslant 8$. Assume that for $u\in T$, $k\in\mathbf{Z}$ we are given a number $\varphi_k(u)\geqslant 0$ with the following properties*:

$$k'\geqslant k \implies \varphi_{k'}(u)\geqslant\varphi_k(u). \tag{4.3}$$

Given $k\in\mathbf{Z}$, $u\in T$, $N\geqslant 1$ points $u_1,\dots,u_N$ of T such that (4.4)

$$\forall l\leqslant N,\quad d(u,u_l)\leqslant r^{-k}, \tag{4.4a}$$

$$\forall l,l',\ 1\leqslant l<l'\leqslant N,\quad d(u_l,u_{l'})\geqslant r^{-k-1}, \tag{4.4b}$$

we have, for some number M,

$$\max_{l\leqslant N}\varphi_{k+2}(u_l)\geqslant\varphi_k(u)+\frac{r^{-k}}{M}\log N. \tag{4.4c}$$

Denote by k_0 the largest integer such that $r^{-k_0}\geqslant\operatorname{diam}T$. Then we can find an increasing sequence of partitions $(\mathcal{A}_k)_{k\geqslant k_0}$ of T, and a probability measure μ on T with the following properties:

$$\textit{The diameter of each set } A\in\mathcal{A}_k \textit{ is at most } 2r^{-k}. \tag{4.5}$$

$$\textit{If } A_k(u) \textit{ denotes the unique element of } \mathcal{A}_k \textit{ that contains } u, \textit{ we have} \tag{4.6}$$

$$\forall u\in T,\quad \sum_{k\geqslant k_0} r^{-k}\log\frac{1}{\mu(A_k(u))}\leqslant K(r)MS \tag{4.6a}$$

where $S=\sup\{\varphi_k(u):k\in\mathbf{Z},\ u\in T\}$ and where $K(r)$ depends on r only.

Comments. (1) A crucial point is the subscript $k+2$ rather than $k+1$ in (4.4c).

(2) The reader will note the main drawback of Theorem 4.2: it does not say how to find the functionals φ_k!

Consider the function h_ϱ on $\mathbf{R}$ given by

$$h_\varrho(x)=(\operatorname{sign}x)|x|^\varrho.$$

Thus h_ϱ increases.

Consider the map h from X_1^* into $\mathbf{R}^m$ given by

$$h(x^*)=(h_\varrho(x^*(x_i)))_{i\leqslant m}.$$

Given a number $A>0$, we will apply Theorem 4.2 to the set

$$\mathcal{F}_A=\left\{h(x^*):x^*\in X_1^*,\ \sum_{i\leqslant m}|h_\varrho(x^*(x_i))|\leqslant A\right\}$$

provided with the distance induced by the norm $\|\cdot\|_\infty$. (Thus $k_0=0$.) The reason why we consider the sequence $(h_\varrho(x^*(x_i)))_{i\leqslant m}$ rather than $(|x^*(x_i)|^\varrho)_{i\leqslant m}$ is the following technical fact. Given $u\in\mathcal{F}_A$, $k\geqslant 0$, define

$$C_k(u)=\{x^*\in X_1^*:h(x^*)\in\mathcal{F}_A,\ \|h(x^*)-u\|_\infty\leqslant 2r^{-k}\}. \tag{4.7}$$

Then

$$C_k(u) \text{ is convex.} \tag{4.8}$$

In (4.7) and the rest of the section, $r=8$. It is however clearer to keep the notation r, and see in due time why $r=8$ works.

We observe that

$$\|v-u\|_\infty \leqslant r^{-k} \implies C_{k+2}(v)\subset C_k(u). \tag{4.9}$$

This follows from the triangle inequality, and the fact that $r^{-k}+2r^{-k-2}\leqslant 2r^{-k}$. The purpose of the factor 2 in (4.7) is to create this condition.

We now define, for $u\in\mathcal{F}_A$,

$$\varphi'_k(u)=\sum_{i\leqslant m}(|u_i|-\min(|u_i|,2r^{-k})),$$
$$\varphi''_k(u)=\inf\{\|x^*\|_\sim : x^*\in C_k(u)\}$$

and

$$\varphi_k(u)=\varphi'_k(u)+(\log m)\varphi''_k(u).$$

It is obvious that (4.3) holds, and that

$$\varphi_k(u)\leqslant A+\log m. \tag{4.10}$$

PROPOSITION 4.3. *The functions φ_k satisfy* (4.4), *where $M=K(C)$.*

Before we start the proof, we state a geometrical lemma.

LEMMA 4.4. *For all x^*,y^* in X^* we have*

$$\|x^*\|_\sim\leqslant 1,\ \|y^*\|_\sim\leqslant 1 \implies \left\|\tfrac{1}{2}(x^*+y^*)\right\|_\sim\leqslant 1-\frac{1}{K(C)}\|x^*-y^*\|_\sim^\varrho. \tag{4.11}$$

Proof. This is the classical fact that the dual of a θ-smooth Banach space is ϱ-convex. See [LiTz].

LEMMA 4.5. *If $s\geqslant 2r^{-k}$, we have*

$$|s-t|\leqslant r^{-k} \implies s-\min(s,2r^{-k})+\tfrac{1}{2}r^{-k}\leqslant t-\min(t,2r^{-k-2}). \tag{4.12}$$

Proof. Since $t\geqslant r^{-k}$, this reduces to

$$s-t\leqslant \tfrac{3}{2}r^{-k}-2r^{-k-2}$$

which holds since $s-t\leqslant r^{-k}$. □

We start the proof of (4.4c). Since, for each $i\leqslant n$, we have $||u_i|-|u_{l,i}||\leqslant r^{-k}$, Lemma 4.5 shows that

$$\varphi'_{k+2}(u_l)\geqslant\varphi'_k(u)+\tfrac{1}{2}r^{-k}\operatorname{card}\{i\leqslant m:|u_i|\geqslant 2r^{-k}\}. \tag{4.13}$$

We now consider a parameter K_1, to be determined later.

Case 1. We have

$$\operatorname{card}\{i\leqslant m:|u_i|\geqslant 2r^{-k}\}\geqslant\frac{\log N}{K_1}.$$

In this case, by (4.13), we have

$$\varphi'_{k+2}(u_l)\geqslant\varphi'_k(u)+\frac{r^{-k}}{2K_1}\log N.$$

By (4.9) we have $C_{k+2}(u_l)\subset C_k(u)$, so that $\varphi''_{k+2}(u_l)\geqslant\varphi''_k(u)$ by definition of φ''_k. Thus (4.4c) holds as soon as $M\geqslant 8K_1$.

Case 2. Case 1 does not occur, that is, if we set

$$I=\{i\leqslant m:|u_i|\geqslant 2r^{-k}\}$$

then

$$\operatorname{card} I\leqslant\frac{\log N}{K_1}. \tag{4.14}$$

The purpose of the functional φ' was to create this condition. The main argument starts now.

Step 1. Consider $t>\max_{l\leqslant N}\varphi''_{k+2}(u_l)$. By definition of φ''_{k+2}, for each $l\leqslant N$ we can find $x_l^*\in C_{k+2}(u_l)$ with $\|x_l^*\|_\sim\leqslant t$. By (4.9) we have $x_l^*\in C_k(u)$. We set $v_l=h(x_l^*)$.

We also note, by (4.4b), and since $r=8$,

$$l\neq l' \implies \|v_l-v_{l'}\|_\infty\geqslant r^{-k-1}-4r^{-k-2}\geqslant\tfrac{1}{2}r^{-k-1}. \tag{4.15}$$

Step 2. We claim that if K_1 has been chosen appropriately, we can find a subset L of $\{1,...,N\}$ such that $\operatorname{card} L\geqslant\sqrt{N}$ and with the following property:

$$\forall l,l'\in L,\ l\neq l' \implies \exists i\leqslant m,\ i\notin I,\ |v_{l,i}-v_{l',i}|\geqslant\tfrac{1}{2}r^{-k-1}. \tag{4.16}$$

To see this, consider the following subset of $\mathbf{R}^m$:

$$B=\{(t_i)_{i\leqslant m}:\forall i\in I,\ |t_i|\leqslant 1\}.$$

By simple volume arguments, the set $u+2r^{-k}B$ can be covered by a family of sets $(W_j)_{j\leqslant N_1}$, such that each set W_j is a translate of $\frac{1}{5}r^{-k-1}B$, and where $N_1\leqslant K^{\operatorname{card} I}$.

Thus if $K_1 \geqslant 2\log K$, we see by (4.14) that $N_1 \leqslant \sqrt{N}$. Thus for a certain choice of j, the set W_j contains at least $\sqrt{N}$ points v_l, and we set $L=\{l\leqslant N: v_l \in W_j\}$. Now, if $l, l' \in W_j$, $l\neq l'$, then for some $i\leqslant m$ we have $|v_{l,i}-v_{l',i}|\geqslant \frac{1}{2}r^{-k-1}$ by (4.15). But $i\notin I$ by definition of W_j, so that (4.16) holds.

Step 3. We observe that for $l\leqslant N$, $i\notin I$, then $|v_{l,i}|\leqslant 4r^{-k}$.

If $l, l'\in L$, (4.16) implies that we can find $i\leqslant m$ such that

$$|h_\varrho(x_l^*(x_i))-h_\varrho(x_{l'}^*(x_i))|\geqslant \tfrac{1}{2}r^{-k-1},$$
$$|h_\varrho(x_l^*(x_i))|\leqslant 4r^{-k}, \quad |h_\varrho(x_{l'}^*(x_i))|\leqslant 4r^{-k}.$$

Thus, by definition of h_ϱ,

$$|x_l^*(x_i)-x_{l'}^*(x_i)|\geqslant \frac{1}{K(C)}r^{-k/\varrho},$$

i.e. $\|x_l^*-x_{l'}^*\|_\sim \geqslant (1/K(C))r^{-k/\varrho}$.

Step 4. We fix l_0 in L, and we set

$$R=\sup_{l\in L}\|x_l^*-x_{l_0}^*\|_\sim.$$

Step 5. The ball centered at $x_{l_0}^*$ of radius R contains $\operatorname{card} L\geqslant\sqrt{N}$ points (namely the points $x_l^*-x_{l_0}^*$ for $l\in L$) that are at mutual distance at least $K(C)^{-1}r^{-k/\varrho}$. Thus Lemma 3.3 implies

$$\log\sqrt{N}\leqslant \log L\leqslant K(C)\log m\left(\frac{r^{-k/\varrho}}{R}\right)^{-\varrho},$$

i.e.

$$R^\varrho \geqslant \frac{1}{K(C)}\cdot\frac{r^{-k}}{\log m}\log N.$$

This means that there exists $l_1\leqslant N$ such that

$$\|x_{l_1}^*-x_{l_0}^*\|^\varrho \geqslant \frac{1}{K(C)}\cdot\frac{r^{-k}}{\log m}\log N. \tag{4.17}$$

We now appeal to (4.11), for $x^*=t^{-1}x_{l_1}^*$, $y^*=t^{-1}x_{l_0}^*$ and we get, using (4.17),

$$\left\|\tfrac{1}{2}(x_{l_1}^*+x_{l_0}^*)\right\|_\sim \leqslant t-\frac{1}{K(C)}t^{1-\varrho}\frac{r^{-k}}{\log m}\log N.$$

Now, since $C_k(u)$ is convex, it contains $\frac{1}{2}(x_{l_1}^*+x_{l_0}^*)$, so that, by definition of φ_k'', we have

$$\varphi_k''(u)\leqslant t-\frac{1}{K(C)}t^{1-\varrho}\frac{r^{-k}}{\log m}\log N.$$

We let $t \to \max_{l\leqslant N} \varphi''_{k+2}(u_l)$, and keep in mind that this number is $\leqslant 1$; we get

$$\max_{l\leqslant N} \varphi''_{k+2}(u_l) \geqslant \varphi''_k(u) + \frac{1}{K(C)} \cdot \frac{r^{-k}}{\log m} \log N.$$

We now observe that

$$|s-t| \leqslant r^{-k} \implies s-\min(s, 2r^{-k}) \leqslant t-\min(t, 2r^{-k-2}).$$

Indeed it suffices to consider the case $s \geqslant 2r^{-k}$, for otherwise the left-hand side is zero; but then more is true by Lemma 4.5. Thus

$$\varphi'_{k+2}(u_l) \geqslant \varphi'_k(u).$$

This completes the proof of (4.4c) provided $M \geqslant K(C)$. □

We now complete the proof of Theorem 4.1. We apply Theorem 4.2 (with $r=8$) to the space $(\mathcal{F}_A, d_\infty)$. This is permitted by Proposition 4.3. It follows from (4.6a) and comparison of the integral (4.1) with the series on the left-hand side of (4.6a) that

$$\gamma_1(\mathcal{F}_A, d_\infty) \leqslant K(C)(A+\log m).$$

Now T_A is the image of $\mathcal{F}_A$ under the map $(t_i)_{i\leqslant m} \to (|t_i|)_{i\leqslant m}$, that is a contraction for d_∞. The definition of γ_1 makes it obvious that

$$\gamma_1(T_A, d_\infty) \leqslant \gamma_1(\mathcal{F}_A, d_\infty).$$

This concludes the proof. □

5. Selector processes

In order to apply Lemma 2.4, we now need useful bounds for the quantity

$$E \sup_{f\in\mathcal{F}} \sum_{i\leqslant n} (\delta_i - \delta) f_i.$$

Since the interest of these bounds goes somewhat beyond the present application we will make the (minimal) extra effort to present a sufficiently general result.

THEOREM 5.1. *Consider a set T provided with two distances d_2, d_∞. Consider a family of centered r.v. $(Z_t)_{t\in T}$ and $\delta>0$. Assume that given any $s,t\in T$, $u>0$, we have*

$$d_2(s,t)\leqslant a_2,\ d_\infty(s,t)\leqslant a_\infty \implies P(Z_s-Z_t\geqslant u)\leqslant \exp\left(-\frac{u}{4a_\infty}\log\frac{ua_\infty}{\delta a_2^2}\right). \tag{5.1}$$

Then for each number $U\geqslant 2$ we have

$$E\sup_{t\in T} Z_t\leqslant K\left(\frac{1}{\log U}\gamma_1(T,d_\infty)+\sqrt{\frac{U\delta}{\log U}}\,\gamma_2(T,d_2)\right) \tag{5.2}$$

where K is universal.

Comment. The most striking choices of U in (5.2) are $U=2$ and $U=\delta^{-1/2}$. In our situation we will make the second choice.

Proof. Consider, for $i=2,\infty$, the largest integer k_i such that r^{-k_i} is larger than the diameter of T for d_i. For simplicity we set $\gamma_1=\gamma_1(T,d_\infty)$, $\gamma_2=\gamma_2(T,d_2)$. First we claim that there is an increasing sequence of partitions $(\mathcal{A}_k)_{k\geqslant k_\infty}$, a probability measure μ_∞ on T such that if $A_k(u)$ denotes the element of $\mathcal{A}_k$ containing u, we have

$$\forall t\in T,\quad \sum_{k\geqslant k_\infty} r^{-k}\log\frac{1}{\mu_\infty(A_k(t))}\leqslant K\gamma_1. \tag{5.3}$$

This fact, which has been known for a long time, is also a consequence of Theorem 4.2 by using the functions

$$\varphi_k(u)=\gamma_1-\gamma_1(B_\infty(u,2r^{-k})) \tag{5.4}$$

that are easily seen to satisfy (4.4). We could also construct a similar sequence of partitions for the distance d_2 (with a term $\sqrt{\log}$ rather than log). But this partition would not be appropriate for our purposes, and a "change of variable" is needed. Tools have been developed to perform this in an efficient manner. We appeal to [T4, Theorem 3.2] (in the case where the functions φ_j are given by $\varphi_j(s,t)=r^{2j}d_2(s,t)$) to see that there is an increasing sequence $(\mathcal{B}_k)_{k\geqslant k_2}$ of partitions of T, and a probability measure μ_2 on T such that, if $d_2(B)$ denotes the diameter of B for d_2, and $B_k(u)$ denotes the unique element of $\mathcal{B}_k$ that contains u, we have

$$\forall t\in T,\quad \sum_{k\geqslant k_2}\left(r^k d_2^2(B_k(t))+r^{-k}\log\frac{1}{\mu_2(B_k(t))}\right)\leqslant K\gamma_2. \tag{5.5}$$

Consider an integer k_1 that will be determined later. For $k\geqslant k_2+k_1$ we write $\mathcal{B}'_k=\mathcal{B}_{k-k_1}$, $B'_k(t)=B_{k-k_1}(t)$.

It follows from (5.5) that

$$\forall t\in T,\qquad \sum_{k\geqslant k_2+k_1} r^{-k}\log\frac{1}{\mu_2(B'_k(t))}\leqslant Kr^{-k_1}\gamma_2, \tag{5.6}$$

$$\forall t\in T,\qquad \sum_{k\geqslant k_2+k_1} r^{k}d_2^2(B'_k(t))\leqslant Kr^{k_1}\gamma_2. \tag{5.7}$$

We set $k_0=\min(k_\infty, k_2+k_1)$. For $k_0\leqslant k<k_\infty$, we set $\mathcal{A}_k=\{T\}$. For $k_0\leqslant k<k_2+k_1$, we set $\mathcal{B}'_k=\{T\}$.

We consider the increasing sequence of partitions $(\mathcal{C}_k)_{k\geqslant k_0}$ where $\mathcal{C}_k$ is generated by $\mathcal{A}_k$ and $\mathcal{B}'_k$.

For $C\in\mathcal{C}_k$, we consider an arbitrary point t_C in C. For $k>k_0$, $t\in T$, we set

$$a_k(t)=\frac{2^{k-k_0}}{\mu_\infty(A_k(t))\mu_2(B'_k(t))},$$

and we observe that, combining (5.3) and (5.6), we have

$$\sum_{k\geqslant k_0} r^{-k}\log a_k(t)\leqslant K(r^{-k_0}+\gamma_1+r^{-k_1}\gamma_2). \tag{5.8}$$

For $C\in\mathcal{C}_k$, $k>k_0$, we set

$$u_C=\frac{4}{\log U}r^{1-k}\log a_k(t_C)+r^{k-1}U\delta d_2^2(B'_{k-1}(t_C)).$$

Consider the quantity

$$S=\sup_{t\in T}\sum_{k>k_0} u_{C_k(t)}.$$

We observe that $B'_k(t_{C_k(t)})=B'_k(t)$ (and similarly for A_k). Thus, combining (5.8) and (5.7) we see that

$$S\leqslant\frac{K}{\log U}(r^{-k_0}+\gamma_1+r^{-k_1}\gamma_2)+Kr^{k_1}U\delta\gamma_2.$$

We see from this formula that it is a good idea to take for k_1 the smallest integer such that $r^{-k_1}\leqslant(U\delta\log U)^{1/2}$. Moreover, since $r^{-k_2}\leqslant K\gamma_2$, $r^{-k_\infty}\leqslant K\gamma_1$, we have

$$r^{-k_0}\leqslant r^{-k_\infty}+r^{-k_2-k_1}\leqslant K(\gamma_1+r^{-k_1}\gamma_2)$$

so that

$$S\leqslant K\left(\frac{\gamma_1}{\log U}+\left(\frac{U\delta}{\log U}\right)^{1/2}\gamma_2\right). \tag{5.9}$$

Consider now $v \geqslant 1$. For $C \in \mathcal{C}_k$, $k > k_0$, we denote by C' the unique element of $\mathcal{C}_{k-1}$ that contains C. We claim that

$$P\big(\sup_{t\in T}(Z_t - Z_{t_A}) \geqslant vS\big) \leqslant \sum_{k>k_0} \sum_{C\in\mathcal{C}_k} P(Z_{t_C} - Z_{t_{C'}} \geqslant vu_C). \tag{5.10}$$

Indeed, if $Z_{t_C} - Z_{t_{C'}} \leqslant vu_C$ for all $k > k_0$, all C in $\mathcal{C}_k$, we have by definition of S that for all C in $\mathcal{C}_k$,

$$Z_{t_C} - Z_{t_T} \leqslant vS$$

so that $\sup_{t\in T}(Z_t - Z_{t_T}) \leqslant vS$. (There the supremum is as usual the essential supremum.)

To evaluate the right-hand side of (5.10), we observe that

$$d_\infty(t_C, t_{C'}) \leqslant r^{-k+1}, \quad d_2(t_C, t_{C'}) \leqslant d_2(B_{k-1}(t_C)).$$

Thus, by (5.1) we have

$$P(Z_{t_C} - Z_{t_{C'}} \geqslant vu_C) \leqslant \exp\left(-\frac{vu_C}{4r^{-k+1}} \log \frac{vu_C r^{-k+1}}{\delta d_2^2(B_{k-1}(t_C))}\right) \leqslant a_k(t_C)^{-v}, \tag{5.11}$$

where we have used the fact that the logarithmic term in (5.11) is at least $\log U$ by the choice of u_C. Since $a_k(t_C) \geqslant 2$, we have $a_k(t_C)^{-v} \leqslant 2^{1-v}(a_k(t_C))^{-1}$. Now

$$a_k(t_C)^{-1} = 2^{-(k-k_0)} \mu_\infty(A_k(t_C)) \mu_2(B_k(t_C))$$

so that the sum of these quantities over all choices of k and C is at most 1. Thus the right-hand side of (5.11) is at most 2^{1-v}. The conclusion follows easily. □

THEOREM 5.2. *Assume that the norm* $\|\cdot\|_\sim$ *satisfies* (1.3), *and consider vectors* $(x_i)_{i\leqslant m}$ *such that* $\|x_i\|_\sim \leqslant 1$. *Consider*

$$A = \sup\left\{\sum_{i\leqslant m} |x^*(x_i)|^\varrho : x^* \in X_1^*\right\}.$$

Consider i.i.d. random variables $(\delta_i)_{i\leqslant m}$ *with* $\delta_i \in \{0,1\}$ *and* $E\delta_i = \delta$. *Then*

$$E \sup_{x^*\in X_1^*} \sum_{i\leqslant m} \delta_i |x^*(x_i)|^\varrho \leqslant \frac{K(C)}{\log(1/\delta)}(A + \log m). \tag{5.12}$$

Proof. Using Lemma 2.4, it suffices to prove that

$$E \sup_{x^*\in X_1^*} \sum_{i\leqslant m} (\delta_i - \delta) |x^*(x_i)|^\varrho \leqslant \frac{K(C)}{\log(1/\delta)}(A + \log m) \tag{5.13}$$

(indeed, $\delta\log(1/\delta)\leqslant K$).

To prove (5.13) we will appeal to Theorem 5.1 with $U=\delta^{-1/2}$,

$$T=\{(|x^*(x_i)|^\varrho)_{i\leqslant m}:x^*\in X_1^*\}$$

and $Z_t=\sum_{i\leqslant m}(\delta_i-\delta)t_i$. We observe that (5.1) holds by Proposition 2.3, and we observe that

$$\gamma_1(T,d_\infty)\leqslant K(C)(A+\log m) \tag{5.14}$$

by Theorem 4.1. We also have that

$$\gamma_2(T,d_2)\leqslant K(C)(A+\log m). \tag{5.15}$$

This is an immediate consequence of the definition of A, (5.14) and [T5, Theorem 1.2].

Now, (5.12) follows from (5.2), (5.14), (5.15). □

Proof of Theorem 1.4. We will use Theorem 5.2 with $\|\cdot\|_\sim=\|\cdot\|$. We set

$$\alpha=\sup_{i\leqslant n}\|T(e_i)\|,$$

and, for $i\leqslant n$, we set $x_i=\alpha^{-1}T(e_i)$, so that $\|x_i\|\leqslant 1$.

Consider a subset I of $\{1,...,n\}$, and denote by $Z(I)$ the norm of the restriction of T to the space generated by the vectors $(e_i)_{i\in I}$. Then we have

$$Z(I)=\alpha\Big(\sup_{x^*\in X_1^*}\sum_{i\in I}|x^*(x_i)|^\varrho\Big)^{1/\varrho}.$$

We also note that, since $\|T\|\leqslant 1$, the number A defined in Theorem 5.2 satisfies $A\leqslant\alpha^{-\varrho}$. Use of (5.12) and Hölder's inequality thus complete the proof. □

THEOREM 5.3. *Under the hypothesis of Theorem* 5.2, *there exists a number* $\alpha=\alpha(C)>0$, *depending on C only, such that*

$$E\sup_{x^*\in X_1^*}\sum_{i\leqslant m}\delta_i|x^*(x_i)|^\varrho\leqslant 2A\delta^\alpha+\log m. \tag{5.16}$$

Proof. We set $\beta=\exp(-2K_0(C))$, where $K_0(C)$ denotes the constant of (5.12). We also set

$$A(\delta)=E\sup_{x^*\in X_1^*}\sum_{i\leqslant m}\delta_i|x^*(x_i)|^\varrho.$$

The key to Theorem 5.3 is the fact that

$$A(\delta\beta)\leqslant\tfrac{1}{2}(A(\delta)+\log m). \tag{5.17}$$

Indeed, once this is proved one sees by induction over k that

$$A(\beta^k) \leqslant 2^{-k}A + \log m. \tag{5.18}$$

If α is such that $\beta^\alpha = \frac{1}{2}$, then (5.18) implies

$$A(\beta^k) \leqslant (\beta^k)^\alpha A + \log m$$

so that for each $\delta < 1$, using the above for the largest k with $\beta^k \geqslant \delta$, we get, since $\beta^k \leqslant \delta/\beta$, that

$$A(\delta) \leqslant \frac{1}{\beta^\alpha}(\delta^\alpha A) + \log m = 2\delta^\alpha A + \log m.$$

To prove (5.17), consider i.i.d. random variables $(\delta_i)_{i\leqslant m}$, $\delta_i \in \{0,1\}$, $E\delta_i = \delta$, and i.i.d. random variables $(\beta_i)_{i\leqslant m}$, $\beta_i \in \{0,1\}$, $E\beta_i = \beta$. We observe that the sequence $(\delta_i\beta_i)_{i\leqslant m}$ is i.i.d., $\delta_i\beta_i \in \{0,1\}$, $E(\delta_i\beta_i) = \delta\beta$, and thus

$$A(\delta\beta) = E \sup_{x^* \in X_1^*} \sum_{i \leqslant m} \delta_i\beta_i |x^*(x_i)|^\varrho.$$

Let us denote by E_δ the expectation given the sequence $(\delta_i)_{i\leqslant m}$. We appeal to Theorem 5.2, for the sequence (β_i) rather than the sequence (δ_i), and the set $\{i \leqslant m : \delta_i = 1\}$ rather than $\{1, ..., m\}$ to get

$$E_\delta \sup_{x^* \in X_1^*} \sum_{i \leqslant m} \delta_i\beta_i |x^*(x_i)|^\varrho \leqslant \frac{K_0(C)}{\log(1/\beta)}(A' + \log m) \tag{5.19}$$

where

$$A' = \sup_{x^* \in X_1^*} \sum_{i \leqslant m} \delta_i |x^*(x_i)|^\varrho.$$

Recalling the value of β and taking expectations in (5.19) finishes the proof. □

Proof of Theorem 1.3. The proof is based on three successive reductions following the scheme of proof of (5.16) (that will not be repeated) and based successively on Theorems 3.1, 5.3 and 5.2. We set again

$$A(\delta) = E \sup_{x^* \in X_1^*} \sum_{i \leqslant n} \delta_i |x^*(x_i)|^\varrho$$

where $E\delta_i = \delta$, $\delta_i \in \{0,1\}$, $(\delta_i)_{i\leqslant n}$ is independent. We first appeal to Theorem 3.1, taking the value of ε there equal to $(\log n)^{-1}$, to get

$$A\left(\frac{1}{e\tau}\right) \leqslant K(C)(\log n)^2.$$

Next, we appeal to Theorem 5.3 to get that for $\delta<1$ we have

$$A\left(\frac{\delta}{e\tau}\right) \leqslant K(C)(\delta^{\alpha}(\log n)^2+\log n),$$

so that

$$A\left(\frac{1}{e\tau(\log n)^{1/\alpha}}\right) \leqslant K(C)\log n.$$

Finally, we appeal to Theorem 5.2 to see that for any $\varepsilon>0$, we have

$$A\left(\frac{1}{e\tau(\log n)^{1/\alpha}n^{\varepsilon}}\right) \leqslant \frac{K(C)}{\varepsilon}.$$

This is the statement of Theorem 1.3. □

To conclude, we describe a simple example borrowed from [BT] that will show how sharp Theorems 3.1 and 5.2 are. We consider two integers $k\leqslant m$ and $n=km$. We divide $\{1,...,n\}$ into m sets $(I_l)_{l\leqslant m}$ of cardinal k. We denote by $(e_l)_{l\leqslant m}$ the canonical basis of l_θ^m, and for $i\in I_l$ we set $x_i=e_l$. It is clear that, with the notations of Theorem 1, we have $\tau=k$. Consider $\delta<1$, and an integer r such that $\delta^r\geqslant 1/m$. Consider i.i.d. random variables $\delta_i\in\{0,1\}$, with $E\delta_i=\delta$. Then, with probability $\geqslant 1-(1-1/m)^m$, there exists $l\leqslant m$ such that $\sum_{i\in I_l}\delta_i\geqslant r$. It follows, with the notations of the proof above, that $A(\delta)\geqslant r/K$, so that

$$A(\delta)\geqslant \frac{\log m}{K\log(1/\delta)}. \tag{5.20}$$

Taking $m=2^k$, we then see that Theorem 5.2 is optimal when $A(=\tau)\leqslant \log n$ (it is not optimal when $A\gg \log n$, by Theorem 5.3). Consider now $\varepsilon>0$ with $n^{-\varepsilon}\leqslant 1/\log n$. When $\delta=n^{-\varepsilon}/(\log n)^K\tau$, then $\log 1/\delta\leqslant K\varepsilon$, so that (5.20) becomes

$$A(\delta)\geqslant \frac{1}{K\varepsilon},$$

and this shows that, in this range, Theorem 1.3 is optimal.

References

[B] BOURGAIN, J., Bounded orthogonal sets and the $\Lambda(p)$-set problem. *Acta Math.*, 162 (1989), 227–246.

[BPST] BOURGAIN, J., PAJOR, A., SZAREK, S. J. & TOMCZAK-JAEGERMAN, N., On the duality problem for entropy numbers of operators, in *Geometric Aspects of Functional Analysis*. Lecture Notes in Math., 1376, pp. 50–63. Springer-Verlag, Berlin–New York, 1989.

[BT] BOURGAIN, J. & TZAFRIRI, L., On a problem of Kadison and Singer. *J. Reine Angew. Math.*, 420 (1991), 1–43.

[F] FERNIQUE, X., Régularité des trajectoires des fonctions aléatoires gaussiennes, in *École d'Eté de Probabilités de Saint-Flour IV*. Lecture Notes in Math., 480, pp. 1–96. Springer-Verlag, Berlin–New York, 1975.

[FLM] FIGIEL, T., LINDENSTRAUSS, J. & MILMAN, V., The dimension of almost spherical sections of convex bodies. *Acta Math.*, 139 (1977), 53–94.

[GZ] GINÉ, E. & ZINN, J., Some limit theorems for empirical processes. *Ann. Probab.*, 12 (1984), 929–989.

[LT] LEDOUX, M. & TALAGRAND, M., *Probability in Banach Spaces.* Springer-Verlag, Berlin, 1991.

[LiTz] LINDENSTRAUSS, J. & TZAFRIRI, L., *Classical Banach Spaces*, Vol. II. Springer-Verlag, Berlin–New York, 1979.

[M] MILMAN, V., Almost Euclidean quotient spaces of subspaces of a finite-dimensional normed space. *Proc. Amer. Math. Soc.*, 94 (1985), 445–449.

[T1] TALAGRAND, M., Regularity of Gaussian processes. *Acta Math.*, 159 (1987), 99–149.

[T2] — Embedding subspaces of L_1 into l_1^N. *Proc. Amer. Math. Soc.*, 108 (1990), 363–369.

[T3] — Cotype and $(q,1)$-summing norm in a Banach space. *Invent. Math.*, 110 (1992), 545–556.

[T4] — Regularity of infinitely divisible processes. *Ann. Probab.*, 21 (1993), 362–432.

[T5] — Supremum of certain canonical processes. *Amer. J. Math.*, 116 (1994), 283–325.

[T6] — Construction of majorizing measures, Bernoulli processes and cotype. *Geom. Funct. Anal.*, 4 (1994), 660–717.

MICHEL TALAGRAND
Equipe d'analyse, Tour 46
Boite 186
Université de Paris VI
4, place Jussieu
75230 Paris Cedex 05
France
mit@ccr.jussieu.fr

Received August 15, 1994
Received in revised form February 1, 1995