# MINKOWSKI SUMS AND SYMMETRIZATIONS

J. BOURGAIN J. LINDENSTRAUSS V.D. MILMAN

Institut des Hautes Etudes Scientifiques
35, route de Chartres
91440 Bures-sur-Yvette, France

Mars 1988

IHES/M/88/14

## MINKOWSKI SUMS AND SYMMETRIZATIONS

**J. Bourgain** IHES, France

**J. Lindenstrauss** Hebrew University

**V.D. Milman** Tel Aviv University

### 0. Introduction

It was found out recently that the empirical distribution method from probability theory has interesting (and perhaps even surprising) consequences in finite-dimensional convexity theory. The relevance of this method to the problem of approximating zonoids by zonotopes with few summands was first pointed out by Schechtman [S]. In [B.L.M] we refined the method and results of Schechtman. In section 6 of [B.L.M] we pointed out that the empirical distribution method gives interesting results if we consider instead of Minkowski sums of segments (i.e., zonotopes) Minkowski sums of more general convex sets. Here we intend to explore this direction in more detail.

In the first section we present results whose proofs are relatively easy modulo known results in the literature. We present those results because of their intrinsic interest from the geometric point of view. There are obviously many variants of the results which can be proved along similar lines. One of the interesting aspects of some of the results of this section, as well as the result of the second section, is that they combine the empirical distribution method with the concentration of measure technique in the spirit of [F.L.M] or more generally [M.S].

In the second section we consider Minkowski symmetrizations (sometimes also called Blaschke symmetrizations) of convex sets. We obtain there a rather sharp estimate on the number of steps needed in order to obtain approximately a ball if we start from an arbitrary convex body and choose the directions of symmetrization randomly and independently. Technically, section 2 is more involved than section 1.

### 1. Minkowski Sums

We start by recalling the basic tool from the empirical distribution method namely the so called Bernstein's inequality. We shall state here three versions of this basic inequality as all three will be used in the sequel. The proofs of all versions are simple and very similar (and of course also classic). We shall present here the proof of the third version (the proofs of (i) and (ii) were recalled in [B.L.M]; a martingale version of (ii) will be proved in the next section). Before stating the result, we recall the definition of Orlicz norms which enter in its statement.

Let $(\Omega, \mu)$ be a probability space and let $\Phi(t)$ be a strictly increasing convex function on $[0, \infty)$ so that $\Phi(0) = 0$ and $\lim_{t\to\infty} \Phi(t) = \infty$. We denote by $L_\Phi(\mu)$ the space of all real valued measurable functions on $\Omega$ so that $\int_\Omega \Phi(|f|/\lambda) d\mu < \infty$ for some $\lambda > 0$ and put

$$\|f\|_{L_\Phi(\mu)} = \inf\{\lambda > 0 \; ; \; \int_\Omega \Phi(|f|/\lambda) d\mu \leq 1 \, .\} \tag{1}$$

We shall be concerned here only with the two functions

$$\psi_1(t) = e^t - 1 \quad , \quad \psi_2(t) = e^{t^2} - 1 \tag{2}$$

(besides of course the functions $t^p$, $1 \leq p < \infty$, which give rise to the usual $L_p$ spaces).

**Proposition 1.** *Let $\{g_j\}_{j=1}^N$ are independent random variables with mean 0 on some probability space $(\Omega, \mu)$.*

*(i) Assume that the $\{g_j\}$ are all bounded and that $\|g_j\|_1 \leq 2$ and $\|g_j\|_\infty \leq A$ for all $j$ and some constant $A$; then, for $0 < \varepsilon < 1$,*

$$\operatorname{Prob}\Big\{\Big|\sum_{j=1}^N g_j\Big| > \varepsilon N\Big\} \leq 2\exp(-\varepsilon^2 N/8A) \, . \tag{3}$$

*(ii) Assume that $\{g_j\}$ belong to $L_{\psi_1}(\mu)$ and that $\|g_j\|_{L_{\psi_1}} \leq A$ for all $j$ and some constant $A$. Then, for $0 < \varepsilon < 4A$,*

$$\operatorname{Prob}\Big\{\Big|\sum_{j=1}^N g_j\Big| > \varepsilon N\Big\} \leq 2\exp(-\varepsilon^2 N/16A^2) \, . \tag{4}$$

*(iii) Assume that the $\{g_j\}$ belong to $L_{\psi_2}(\mu)$ and that $\|g_j\|_{L_{\psi_2}} \leq A$ for all $j$ and some constant $A$. Then, for all $\varepsilon > 0$,*

$$\operatorname{Prob}\Big\{\Big|\sum_{j=1}^N g_j\Big| > \varepsilon N\Big\} \leq 2\exp(-\varepsilon^2 N/8A^2) \, . \tag{5}$$

Note that in contrast to (i) and (ii) we do not have in (iii) any restriction of the size of $\varepsilon$. We will in fact use (iii) below also for large $\varepsilon$.

which matters is that the constant in the right hand side of (6) behaves like $p^{1/2}$ as $p \to \infty$ and this is not hard to verify. Another form of stating (6) and (7) is that

$$(8) \qquad \Big(\int_0^1 \Big|\sum_{i=1}^n r_i(t)u_i\Big|^p dt\Big)^{1/p} \le cp^{1/2}\Big(\sum_{i=1}^n |u_i|^2\Big)^{1/2} \le c_1 p^{1/2} \int_0^1 \Big|\sum_{i=1}^n r_i(t)u_i\Big| dt$$

where $\{r_i\}$ denote the Rademacher functions. By expanding $\exp(x^2)$ into power series we deduce from (8) that for some constant $c$ (we use here and below the same notation for different absolute constants in different formulas)

$$(9) \qquad \Big\|\sum_{i=1}^n r_i(t)u_i\Big\|_{L_{\psi_2}(0,1)} \le c\Big\|\sum_{i=1}^n r_i(t)u_i\Big\|_{L_1(0,1)} .$$

Khintchine's inequality has been generalized in many directions and several of these generalizations will be useful to us below.

Kahane [Ka] and Kwapien [Kw] (cf. also [L.T, p.74] for a proof and [Ta] for a new approach) showed that (8) and (9) remain valid if we replace the scalars $\{u_i\}_{i=1}^n$ by vectors in a Banach space $X$ (the constants being independent of $X$ as well as of the particular choice of $n$ and $\{u_i\}_{i=1}^n$). Thus we have

$$(10) \qquad \Big\|\sum_{i=1}^n r_i(t)u_i\Big\|_{L_{\psi_2}((0,1),X)} \le c\Big\|\sum_{i=1}^n r_i(t)u_i\Big\|_{L_1((0,1),X)} .$$

A second generalization of (9) is Borell's inequality [Bo] (cf. [G.M] for the current point of view) which follows from the Brunn Minkowski inequality. Let $K$ be a convex body symmetric with respect to the origin in $R^n$ with volume 1. Let $y \in R^n$ be a vector of Euclidean norm 1 and let $\varphi(t)$ be the $(n-1)$-dimensional volume of $K \cap \{x; \langle x, y\rangle = t\}$. Then for some absolute constants $c$ and $C$

$$(11) \qquad \varphi(t) \le C\varphi(0)\exp\big(-c\varphi(0)|t|\big)$$

and consequently if $\mu$ is the Lebesgue measure restricted to $K$ then

$$(12) \qquad \|\langle \cdot, y\rangle\|_{L_{\psi_1}(\mu)} \le c_1\|\langle \cdot, y\rangle\|_{L_1(\mu)}$$

In this case we cannot replace in general $L_{\psi_1}$ by $L_{\psi_2}$ (e.g., for the unit ball in $\ell_1^n$, suitably normalized). For $K$ the cube (12) becomes a weak form of Khintchine inequality. A discussion of the relation between Borell's inequality and Kahane's inequality is presented in Appendix II of [M.S].

**Proof of (iii).** Without loss of generality we may assume that $A = 1$ and thus, for $1 \le j \le N$,

$$\int_\Omega \exp(g_j^2)\,d\mu \le 2 .$$

In particular

$$\int_\Omega |g_j|^{2k} d\mu \le 2 \cdot k! , \qquad 1 \le k < \infty,$$

and hence also $\int_\Omega |g_j|^{2k-1} d\mu \le 2 \cdot k!$ for $k \ge 1$. If $\lambda \ge 1$ we deduce that

$$\begin{aligned}\int_\Omega \exp(\lambda g_j)\,d\mu &\le 1 + \sum_{k=2}^{\infty} \int_\Omega |\lambda g_j|^k d\mu / k! \\ &\le 1 + \lambda^2 + 2\sum_{k=2}^{\infty} \lambda^{2k} k! \big((2k!)^{-1} + (2k-1)!^{-1}\big) \le e^{2\lambda^2} .\end{aligned}$$

This estimate holds also for $0 < \lambda \le 1$. Consequently,

$$\begin{aligned}\operatorname{Prob}\Big\{ \sum_{j=1}^{N} g_j > \varepsilon N \Big\} e^{\varepsilon\lambda N} &\le \int_\Omega \exp\big(\lambda \sum_{j=1}^{N} g_j\big)\,d\mu \\ &= \prod_{j=1}^{N} \int_\Omega \exp(\lambda g_j)\,d\mu \le e^{2\lambda^2 N} .\end{aligned}$$

By taking $\lambda = \varepsilon/4$ we deduce (5). □

In order to apply in a significant way Proposition 1 in convexity theory it is useful to have at our disposal distribution inequalities which ensure that for some functions $g$ which are of geometric interest the norm $\|g\|_1$ is comparable to $\|g\|_{L_{\psi_1}}$ or even $\|g\|_{L_{\psi_2}}$. The basic result in this direction is Khintchine's inequality which states the following: There is an absolute constant $c$ so that for every choice of reals $\{u_i\}_{i=1}^{n}$ we have

(6) $$\Big(\text{Average}\Big|\sum_{i=1}^{n} \pm u_i\Big|^p\Big)^{1/p} \le c p^{1/2} \Big(\sum_{i=1}^{n} u_i^2\Big)^{1/2} , \qquad 1 \le p < \infty$$

(7) $$\Big(\sum_{i=1}^{n} u_i^2\Big)^{1/2} \le \sqrt{2}\,\text{Average}\Big|\sum_{i=1}^{n} \pm u_i\Big| ,$$

where the averages are taken over all $2^n$ possible choice of signs. The best constants in (6) and (7) were found in [Sz] and [H] but their exact values are not important to us. The only thing

Another direction in which (9) or (10) can be generalized is by replacing the Rademachers by more general random variables. In particular these inequalities hold if the Rademachers are replaced by independent normalized Gaussian variables [L.S]. Marcus and Pisier [Ma.P] discussed this question in detail and deduced from the Gaussian version of Khintchine's inequality the following. Let $\| \quad \|$ be any norm on $R^n$, let $x \in R^n$ and let $\mu_n$ the normalized Haar measure on the orthogonal group $O^n$. Then for some absolute $c$

$$(13) \qquad \|Ux\|_{L_{\psi_2}(\mu_n)} \leq c\|Ux\|_{L_1(\mu_n)} \, .$$

After presenting these preliminaries we can turn to the main subject of this section, i.e., results concerning Minkowski sums of convex bodies. We start by recalling a result from [B.L.M] (theorem 6.3. there).

**Theorem 2.** *Let $K$ be a compact convex set in $R^n$ and let $0 < \varepsilon < 1/2$. There is a constant $r(K) = r$ and there are orthogonal transformations $\{U_j\}_{j=1}^N$ on $R^n$ with*

$$(14) \qquad N \leq cn\varepsilon^{-2} \log \varepsilon^{-1}$$

*so that*

$$(15) \qquad (1-\varepsilon)rB^n \subset N^{-1} \sum_{j=1}^{N} U_j K \subset (1+\varepsilon)rB^n$$

Here $B^n$ denote the Euclidean ball in $R^n$ and $c$ is, as usual, an absolute constant. Since the method of proof of Theorem 2 serves as a model to several of the theorems below we recall briefly its proof.

**Proof:** Without loss of generality we may assume that $K$ is symmetric with respect to the origin (otherwise replace $K$ by $K + (-K)$) and that it has non-empty interior. By duality (15) is equivalent to

$$(16) \qquad (1-\varepsilon)r|||x||| \leq N^{-1} \sum_{j=1}^{N} \|U_j^* x\|_* \leq (1+\varepsilon)r|||x|||$$

where $|||\cdot|||$ denote the Euclidean norm and $\| \quad \|_*$ is the norm whose unit ball is the polar to $K$.

We choose

$$(17) \qquad r = \int_{S^{n-1}} \|x\|_* d\sigma_n(x) = \int_{O^n} \|Ux\|_* d\mu_n(U)$$

where $\sigma_n$ is the normalized rotation invariant measure on $S^{n-1}$ and $\mu_n$ is the Haar measure on $O^n$. Let $\{x_i\}_{i=1}^m$ be an $\varepsilon$-net with respect to $|||\cdot|||$ on $S^{n-1}$ with $m \le (4/\varepsilon)^n$. For each $i$ we consider the random variables $\{g_{i,j}\}_{j=1}^N$ on $(O^n, \mu_n)$ defined by

$$g_{i,j}(U) = \|Ux_i\|_* - r \qquad 1 \le j \le N$$

where $N$ will be determined shortly. Clearly, for every $i$ and $j$, $g_{i,j}$ has mean 0 and $\|g_{i,j}\|_{L_1(\mu_n)} \le 2r$. It follows from (13) that $\|g_{i,j}\|_{L_{\psi_2}(\mu_n)} \le c_1$ for some absolute constant $c_1$. From (5) we deduce that for every $\varepsilon > 0$ and every $1 \le i \le m$

$$\operatorname{Prob}\left\{(U_1, \ldots, U_N)\ ;\ \left|N^{-1}\sum_{j=1}^N \|U_j x_i\|_* - r\right| < \varepsilon r\right\} \le 2\exp(-\varepsilon^2 N/8c_1^2)\ .$$

Consequently if $2\exp(-\varepsilon^2 N/8c_1^2) \le 1/m$, i.e., if $N$ satisfies (14) there exist $\{U_j\}_{j=1}^N$ so that (16) holds if we take as $x$ any element of the $\varepsilon$-net $\{x_i\}_{i=1}^m$. This implies that (16) holds for every $x$ with $\varepsilon$ replaced by $\Delta = 3\varepsilon/(1-\varepsilon)$ in view of the following simple and well known lemma.

**Lemma 3.** *Let $T$ be a bounded linear map from a Banach space $X$ into a Banach space $Y$. Let $0 < \varepsilon < 1$ and assume that for an $\varepsilon$-net $\mathcal{F}$ of the unit sphere of $X$, $\big|\|Tx\| - \|x\|\big| < \varepsilon$ for every $x \in \mathcal{F}$. Then $\big|\|Tx\| - \|x\|\big| < 3\varepsilon\|x\|/(1-\varepsilon)$ for every $x \in X$.*

**Proof:** Put $\Delta = \sup\left\{\big|\|Tx\| - \|x\|\big|\ ,\ \|x\| = 1\right\}$. For every $x \in X$ with $\|x\| = 1$ there is a $u \in \mathcal{F}$ wo that $\|x - u\| \le \varepsilon$. We have

$$\big|\|Tx\| - \|x\|\big| \le \big|\|Tu\| - \|u\|\big| + \big|\|T(x-u)\| - \|x-u\|\big| + 2\|x-y\|\ .$$

Hence $\Delta \le \varepsilon + \varepsilon\Delta + 2\varepsilon$ or $\Delta \le 3\varepsilon/(1-\varepsilon)$. This concludes the proof of the Lemma as well as of Theorem 2. □

**Remark.** The dependence of $N$ in (14) on $n$ and $\varepsilon$ is close to being optimal in the case $K$ is a segment, especially for large $n$. This question was discussed in some detail in section 6 of [B.L.M]. It is known that if $K$ is a segment one can replace (14) by $N \le c\varepsilon^{-2}n$ ([G]). It is likely that the $|\log\varepsilon|$ factor can be dropped also for general $K$, we did not check this point.

Our next result is of a very similar nature to Theorem 2. It uses instead of the orthogonal group Kahane's theorem. The point in this application is to show that the fact that a relatively small number $N$ is needed in (14) is not a consequence of some deep group theoretical properties of the orthogonal group. A similar phenomenon happens if we consider the simpler and obviously discrete setting of the operation of changing signs.

Let $\| \quad \|$ be a norm in $R^n$ and let $\{e_i\}_{i=1}^n$ denote the unit vector basis. We define a new norm $\|x\|_0$ in $R^n$ by putting for $x = \sum_{i=1}^n \alpha_i e_i$

$$\|x\|_0 = \underset{\pm}{\text{Average}} \, \|\sum_{i=1}^n \pm\alpha_i e_i\| = \int_0^1 \|\sum_{i=1}^n r_i(t)\alpha_i e_i\| dt \,. \tag{18}$$

The unit vectors $\{e_i\}_{i=1}^n$ form a 1-unconditional basis with respect to $\|\cdot\|_0$ (if they are already a 1-unconditional basis of $\| \quad \|$ then of course $\|x\| = \|x\|_0$ for all $x$).

**Proposition 4.** *Let $\|\cdot\|_0$ be defined as in (18). Then for every $\frac{1}{2} > \varepsilon > 0$ there are $N$ choices of $n$-tuples of signs $\{\theta_1^j, \dots, \theta_n^j\}_{j=1}^N$ with $N$ given by (14) so that*

$$(1-\varepsilon)\|x\|_0 \le N^{-1} \sum_{j=1}^N \|\Sigma\theta_i^j \alpha_i e_i\| \le (1-\varepsilon)\|x\|_0 \,, \qquad x = \sum_{i=1}^n \alpha_i e_i \,. \tag{19}$$

**Proof:** Use exactly the same argument as in the proof of Theorem 2, just replace the use of (13) by that of (10). □

We now pass to a situation where $L_{\psi_2}$ estimates are not available and we have to use $L_{\psi_1}$ estimates. Given a convex symmetric body $K$ in $R^n$ normalized so that its volume is 1. We associate to it a zonoid $M(K)$ (called the centroid or moment zonoid of $K$) whose support functional (i.e., the norm determined by its polar) is given by

$$h_{M(K)}(u) = \int_K |\langle x, u\rangle| d\mu(x)$$

where $\mu$ denotes here the usual Lebesgue meaure. The zonoid $M(K)$ should not be confused with another zonoid associated to $K$, namely the projection body of $K$ (which will be discussed in some detail in the paper [B.L] in this volume). In contrast to the projection body which can be an arbitrary zonoid the zonoid $M(K)$ defines a norm which is up to an absolute constant $\tilde{c}$ (independent of $K$ or $n$ Euclidean. This is an immediate consequence of (12) above ((12) also implies of course that $\|\langle\cdot\ y\rangle\|_{L_1(\mu)}$ is equivalent to $\|\langle\cdot\ y\rangle\|_{L_2(\mu)}$). Like any zonoid which is of type 2 it can be (in view of [B.L.M]) approximated up to $\varepsilon$ by a zonotope having $c(\varepsilon)n$ summands. The argument used above combined with (12) gives however a little more specific information.

**Proposition 5.** *Let $K$ be a symmetric convex body with volume 1 in $R^n$. Then for $\varepsilon < 1/2$ and for $N$ satisfying (14), we have with probability $1 - o(\varepsilon)$ that*

$$(1-\varepsilon)M(K) \subset N^{-1}\sum_{j=1}^{N}[-x_j, x_j] \subset (1+\varepsilon)M(K) \tag{20}$$

*if the $\{x_j\}_{j=1}^N$ are chosen randomly (with uniform distribution on $K$) and independently.*

**Proof:** By duality (20) is equivalent to

$$\left|1 - N^{-1}\sum_{j=1}^{N}|\langle u, x_j\rangle|\right| < \varepsilon \tag{20'}$$

whenever $\int_K |\langle u, x\rangle| d\mu = h_{M(K)}(u) = 1$. We consider now an $\varepsilon$-net $\{u_i\}_{i=1}^m$ in the set $\{u\ ;\ h_{M(K)}(u) = 1\}$ and consider the random variables $|\langle u_i, \cdot\rangle| - 1$ on $K$. The same argument as that of the proof of Theorem 2 (using now (4) and (12)) shows that (20') holds with a large probability if the $\{x_j\}_{j=1}^N$ are chosen independently and uniformly distributed on $K$. □

We come back to the theme of Theorem 2 namely the representation of the Euclidean ball as a Minkowski sum. We shall show, by combining the proof of Theorem 2 with some results from [F.L.M], that the estimate on $N$ given in (14) can be improved somewhat if $K$ has an interior point and the Euclidean norm is properly chosen (or equivalently, we first apply a proper affine transformation on $K$). For some $K$ the improvement on the estimate of $N$ can be very substantial.

**Theorem 6.** *Let $K$ be a convex symmetric body in $R^n$ so that the Euclidean ball $B^n$ is the ellipsoid of minimal volume containing $K$. Then (15) holds with $N$ satisfying*

$$N \le cn(\log n)^{-1}\varepsilon^{-2}|\log \varepsilon|\ . \tag{21}$$

*If moreover the space $R^n$ normed by the polar $K^\circ$ of $K$ has cotype $q$ for some $q < \infty$ with cotype constant $C_q(K^\circ)$ then we can replace (21) by*

$$N \le c_1 C_q(K^\circ)^2 n^{1-2/q}\varepsilon^{-2}|\log \varepsilon|\ . \tag{22}$$

**Proof:** By the well known result of F. John [J] we have that $n^{-1/2}B^n \subset K \subset B^n$, i.e.,

$$n^{-1/2}|||x||| \le \|x\|_* \le |||x||| \qquad x \in R^n$$

where $|||\cdot|||$ and $\|\cdot\|_*$ denote the norms whose unit balls are $B^n$ and $K^\circ$ respectively. It was shown in [F.L.M] that for some positive absolute constant $c_2$

$$(23) \qquad r(K) = \int_{S^{n-1}} \|x\|_* d\sigma_n(x) \geq c_2(\log n/n)^{1/2}$$

and that $\|x\|_*$ is strongly concentrated aroung $r(K)$ in the sense that for $m = 0, 1, 2, \ldots,$

$$(24) \qquad \sigma_n\{x \in S^{n-1}, mn^{-1/2} \leq |\|x\|_* - r(K)| \leq (m+1)n^{-1/2}\} \leq 4\exp\big(-(m-c_3)^2/2\big)$$

where $c_3$ is an absolute constant. It follows from (23) and (24) that for $x$ with $|||x||| = 1$

$$\big|\|Ux\| - r(K)\big|_{L_{\psi_2}(O^n,\mu_n)} \leq c_4 n^{-1/2} \leq c_4 r(K)/c_2(\log n)^{1/2} .$$

By the reasoning in the proof of Theorem 2 we get that for every $\varepsilon > 0$ there exist $\{U_j\}_{j=1}^N$ in $O^n$ with $N$ satisfying (21) so that

$$\big|N^{-1}\sum_{j=1}^N \|U_j^* x\|_* - r(K)\big| < \varepsilon r(K)$$

for every $x$ in an $\varepsilon$-net of $S^{n-1}$. An application of Lemma 3 and duality conclude the proof for general $K$. If we have information on $C_q(K^\circ)$ for some $q < \infty$ then as shown in [F.L.M] we can improve (23) to

$$r(K) \geq c_5 C_q(K^\circ)^{-1} n^{1/q-1/2}$$

and this yields the estimate (22) on $N$. □

**Remarks.** 1. The estmates (21) and (22) are the best possible in general, at least as far as the dependence on $n$ is concerned. We shall explain this in the case of (21). Let $K$ be the octahedron in $R^n$ and let $\{T_jK\}_{j=1}^N$ be affine images of $K$ (not necessarily rotations!) such that $\sum_{j=1}^N T_jK$ is up to 2 say an Euclidean ball. Then $N \geq cn/\log n$. Indeed, by passing to the duals the assumption implies that the Banach Mazur distance from $\ell_2^n$ to a subspace of $Z = (\ell_\infty^n \oplus \cdots \oplus \ell_\infty^n)_1$ (with $N$ summands) is at most 2. By counting extreme points in the unit ball of $Z^*$ we get that $Z$ embeds isometrically in $\ell_\infty^{(2n)^N}$. Hence, as observed in [F.L.M], $c_1 \log\big((2n)^N\big) \geq n$ which is the desired result. The same reasoning gives the following corollary of (21) in Banach space terminology. For every $n$ dimensional Banach space $X$, the space $\ell_2^n$ is of Banach Mazur distance $\leq 2$ from a subspace of $(X \oplus X \oplus \cdots \oplus X)_1$ with $cn/\log n$ summands. This fact however, also follows directly from [F.L.M] without applying the empirical distribution method.

2. One should note that in the proof above we had to use **Proposition 1** (iii) for large $\varepsilon$. In the proof of Theorem **2** itself, we used only small $\varepsilon$ and thus we could have used there only $L_{\psi_1}$ estimates and Proposition 1 (ii).

The previous result becomes of special interest if we consider bodies $K$ for which $C_2(K^\circ)$ is bounded (e.g., $K$ the $n$-dimensional cube). In this case we get in (22) an estimate on $N$ which depends only on $\varepsilon$ but not on $n$. A variant of this observation is the following known fact.

**Proposition 7.** *Let $K$ be a convex symmetric body in $R^n$ so that the Euclidean ball $B^n$ is the ellipsoid of minimal volume containing $K$. Then there is an orthogonal transformation $U$ so that*

(25)
$$\lambda^{-1}B^n \subset \tfrac{1}{2}(K+UK) \subset B^n$$

*where*

(26)
$$\lambda \le cC_2(K^\circ)\ln\left(C_2(K^\circ)+1\right)$$

**Proof:** The right hand inclusion in (25) is clear for every choice of $U$ so we have only to check the other inclusion relation. We denote the norm induced by $B^n$, respectively $K^\circ$, by $|||\cdot|||$ and $\|\quad\|_*$. It was proved in [K] in the case of the cube and in [M], [M.P] for a general $K$ that the set of subspaces $Y$ of $R^n$ of dimension $\left[(n+1)/2\right]$ for which

(27)
$$|||y||| \le c_1\|y\|_* C_2(K^\circ)\ln\left(C_2(K^\circ)+1\right) \qquad y\in Y$$

has measure $> \frac{1}{2}$ (with respect to the usual normalized measure on the Grassmanian). Hence there is a subspace $Y\subset R^n$ so that (27) holds for $Y$ as well as its orthogonal complement (with respect to $|||\quad|||$) $Y^\perp$. Choose now $U$ to be the orthogonal map defined by

$$Uy=y\ ,\ y\in Y \quad , \quad Uy=-y\ ,\ y\in Y^\perp\ .$$

Then for $x=y+z$ with $y\in Y$ and $z\in Y^\perp$ we have

(28)
$$\|Ux\|_* + \|x\|_* \ge \|y\|_* + \|z\|_* \ge \left(cC_2(K^\circ)\ln C_2(K^\circ)\right)^{-1}|||x|||$$

and this proves (25). □

**Remark.** It is of some interest to look at (28) also in the following form. By definition of cotype 2 we have for every choice of $\{x_i\}_{i=1}^m$ in $R^n$

$$C_2(K^\circ)\int_0^1 \|\sum_{i=1}^m r_i(t)x_i\|_* dt \geq \left(\sum_{i=1}^m \|x_i\|^2\right)^{1/2} .$$

Inequality (28) shows that also an inequality in the other direction holds (of course not the obvious reverse inequality since this would imply type 2 and thus that $\| \quad \|_*$ is uniformly equivalent to an inner product norm). The inequality we get is that

$$(29)\qquad \int_0^1 \|\sum_{i=1}^m r_i(t)x_i)\|_* dt \leq \tilde{c}C_2(K^\circ)\ln 2C_2(K^\circ)\left(\left(\sum_{i=1}^m \|x_i\|_*^2\right)^{1/2} + \left(\sum_{i=1}^m \|Ux_i\|_*^2\right)^{1/2}\right) .$$

Indeed, we have

$$\int_0^1 \|\sum_{i=1}^m r_i(t)x_i\|_* dt \leq \int_0^1 |||\sum r_i(t)x_i|||dt \leq c_1\left(\sum_{i=1}^m |||x_i|||^2\right)^{1/2}$$

and by (28) we deduce (29).

We now give a result on a very general Minkowski sum. The proof is straightforward generalization of the argument of Schechtman in [S].

**Theorem 8.** *Let $\{K_i\}_{i=1}^m$ be symmetric convex bodies in $R^n$ and let $K = \sum_{i=1}^m K_i$. Then for every $0 < \varepsilon < 1/2$ there is subset $\{i_j\}_{j=1}^N$ of $\{1,\dots,m\}$ with*

$$(30)\qquad N \leq cn^2\varepsilon^{-2}|\log\varepsilon|$$

*and scalars $\{\lambda_j\}_{j=1}^N$ so that*

$$(31)\qquad (1-\varepsilon)K \subset \sum_{j=1}^N \lambda_j K_{i_j} \subset (1+\varepsilon)K .$$

**Proof:** We pass again to the duals. Let $\|\cdot\|_i$ be the norm in $R^n$ whose unit ball is $K_i^\circ$. Then $\|x\| = \sum_{i=1}^m \|x\|_i$ is the norm whose unit ball is $K^\circ$. Let $\{x_k\}_{k=1}^n$ be an Auerbach basis of $R^n$ with respect to $\|\cdot\|$. In other words $\|x_k\| = 1$ for $1 \leq k \leq n$ and for any $x \in R^n$ with $\|x\| \leq 1$ we have $x = \sum_{k=1}^n a_k x_k$ with $|a_k| \leq 1$ for all $k$ Let $\alpha_i = \sum_{k=1}^n \|x_k\|_i$, $1 \leq i \leq m$. Then $\sum_{i=1}^m \alpha_i = \sum_{k=1}^n \|x_k\| = n$ and for every $x$ with $\|x\| = 1$ we have $\|x\|_i \leq \alpha_i$. Let $\nu$ be the probability measure on $\Omega = \{1,2,\dots,m\}$ defined by $\nu(i) = \alpha_i/n$. Let $\{u_t\}_{t=1}^s$ be an $\varepsilon$-net with respect to

$\|\cdot\|$ of the unit sphere in $R^n$ with the same norm and with $s \le (4/\varepsilon)^n$. The functions $\{g_t\}_{t=1}^s$ on $\Omega$ defined by

$$g_t(i) = (n\|u_t\|_i/\alpha_i) - 1 \qquad 1 \le i \le m \ .$$

These are functions of mean 0 with $\|g_t\|_{L_\infty(\nu)} \le n$ and $\|g_t\|_{L_1(\nu)} \le 1$. By applying (3) we can find $\{i_j\}_{j=1}^N$ so that for every $1 \le t \le s$

$$\left| N^{-1} \sum_{j=1}^N n\|u_t\|_{i_j}/\alpha_{i_j} - 1 \right| < \varepsilon \tag{32}$$

provided that $2\log s = 2n|\log 4/\varepsilon| \le \varepsilon^2 N/8n$, i.e., that (30) holds. Now by Lemma 3 it follows from (32) that

$$(1-\Delta)K \subset \frac{1}{N}\sum_{j=1}^N n\alpha_{i_j}^{-1} K_{i_j} \subset (1+\Delta)K$$

where $\Delta = 3\varepsilon/(1-\varepsilon)$. □

It is likely that at least some of the refinements of Schechtman's result which were worked out in [B.L.M] carry over to the setting of Theorem 8. We did not pursue this point.

## 2. Minkowski symmetrizations

Let $K$ be a compact convex set in $R^n$ and let $u$ be any vector in $S^{n-1} = \{u \ ; \ |||u||| = 1\}$, where as before $|||\cdot|||$ denotes the Euclidean norm. We denote by $\pi_u$ the reflection with respect to the hyperplane through 0 orthogonal to $u$, i.e.,

$$\pi_u x = x - 2\langle x, u\rangle u \ . \tag{33}$$

Obviously $\pi_u$ is an orthogonal transformation in $R^n$. The Minkowski symmetrization $K$ with respect to $u$ (or with respect to hyperplane orthogonal to $u$) is defined to be the convex set $\frac{1}{2}(\pi_u K + K)$. This type of symmetrization was apparently used first by Blaschke in his book "Kreis and Kugel" and is therefore also called Blaschke symmetrization. The main result in this section is a proof of the fact that for every $\varepsilon > 0$ we get from $K$ after performing $cn\log n + c(\varepsilon)n$ random Minkowski symmetrizations a body which is (with a large probability) up to $\varepsilon$ an Euclidean ball. By a random Minkowski symmetrization we mean an operation in which $u$ is chosen randomly on $S^{n-1}$ with the usual rotation invariant distribution. If we perform several operations the $u$'s in different steps are chosen independently. After $N$ symmetrizations by vectors $u_1, \ldots, u_N$, the resulting body may be written in the form

$$2^{-N} \sum_{D \subset \{1,\ldots,N\}} \prod_{i \in D} \pi_{u_i}(K)$$

where the sum is taken over all subsets $D \subset \{1, \ldots, N\}$. As in section 1, we want this Minkowski sum to approximate the euclidean ball for a random choice of vectors $u_1, \ldots, u_N$ in the sphere $S^{n-1}$, where $N$ should be taken as small as possible.

This result on Minkowski symmetrizations is in some sense related to the fact discovered by Diaconis and Shashahani (cf. for example [DS]) that the product of $\frac{1}{2} n \log n + cn$ random reflections $\pi_u$ produce a random orthogonal transformation on $R^n$. It does not seem however that this fact can be used for simplifying the proof of our main result. The methods we use here are different from the tools employed in [DS]. In particular we do not use here results from the theory of group representations.

In the proof given here, the methods of the previous section have to be combined with additional techniques, given in Lemmas 10,11,12. We will make use of the concentration phenomena on the orthogonal group (see Lemma 13) besides the concentration on the sphere.

The version of Bernstein's inequality we use in this section involves martingales rather than independent random variables. More specifically we use the following martingale version of Proposition 1(ii).

**Proposition 9.** *Let $(\Omega, \mathcal{B}, \mu)$ be a probability measure space and let $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots \subset \mathcal{B}_N$ be sub-$\sigma$-algebras of $\mathcal{B}$. Let $f$ be a $\mathcal{B}_N$ measurable real valued function and let put $E(f \mid \mathcal{B}_i) = f_i$, $1 \le i \le N$. Assume further that for some constant $A$*

$$E\big(\exp(|f_i - f_{i-1}|/A \mid \mathcal{B}_{i-1}\big) \le 2 \;, \qquad 1 \le i \le N \tag{34}$$

*($\mathcal{B}_0$ is the trivial $\sigma$-field and $f_0 = E(f \mid \mathcal{B}_0) = \int_\Omega f(\omega) d\mu$). Then for $0 < \varepsilon < 4A$*

$$\text{Prob}\,\Big\{\big|f - \int_\Omega f\,d\mu\big| \ge \varepsilon N\Big\} \le 2\exp(-\varepsilon^2 N/16A^2) \;. \tag{35}$$

Note that the assumption (34) is stronger than $\|f_i - f_{i-1}\|_{L_{\psi_1}} \le A$; it is a pointwise condition on the conditional expectation.

**Proof:** There is no loss of generality to assume that $A = 1$ and that $\int_\Omega f\,d\mu = 0$. Put $d_i = f_i - f_{i-1}$ and note that $E(d_i \mid \mathcal{B}_{i-1}) = 0$. Also by (34) $E\big(|d_i|^k \mid \mathcal{B}_{i-1}\big) \le 2 \cdot k!$ for

$1 \le i \le N$ and all $k$. Hence for $0 < \lambda < 1/2$

$$E\big(\exp(\lambda\, d_i) \mid \mathcal{B}_{i-1}\big) \le 1 + 2\sum_{k=2}^{\infty} \lambda^k \le 1 + 4\lambda^2 \ , \qquad 1 \le i \le N \ , \ \omega \in \Omega \ .$$

Consequently for $i \le N$

$$\int_\Omega \exp(\lambda f_i) d\mu = \int_\Omega E\big(\exp(\lambda f_i) \mid \mathcal{B}_{i-1}\big) d\mu$$
$$= \int_\Omega \exp(\lambda f_{i-1}) \cdot E\big(\exp(\lambda d_i) \mid \mathcal{B}_{i-1}\big) d\mu \le (1 + 4\lambda^2) \int_\Omega \exp(\lambda f_{i-1}) d\mu \ .$$

It follows that

$$\mathrm{Prob}\{f > \varepsilon N\} \exp(\varepsilon\lambda N) \le \int_\Omega \exp(\lambda f) d\mu \le (1 + 4\lambda^2)^N \le \exp(4\lambda^2 N) \ .$$

By taking $\lambda = \varepsilon/8$ ($\le 1/2$) we get

$$\mathrm{Prob}\{f_N > \varepsilon N\} \le \exp(-\varepsilon^2 N/16)$$

and this implies (35)

We shall apply Proposition 9 in the proof of

**Lemma 10.** *For $x \in S^{n-1}$ and $u = (u_1, u_2, \ldots, u_N) \in (S^{n-1})^N$ put*

$$\varphi_u(x) = \sup_{y \in S^{n-1}} 2^{-N} \sum_D \Big|\Big\langle \prod_{i \in D} \pi_{u_i} x, y \Big\rangle\Big| \tag{36}$$

*where the sum is taken over all subsets $D$ of $\{1, \ldots, N\}$. Then for fixed $x$, $0 < \delta < \delta_0$, $n > n(\delta)$ and $N = c_1 n |\log \delta|$ we have*

$$\mathrm{Prob}\,\big\{u \ ; \ \varphi_u(x) > \delta\big\} \le \exp(-c_2 \delta^2 n / |\log \delta|) \tag{37}$$

*where $\delta_0, c_1$ and $c_2$ are absolute constants.*

**Proof:** Clearly $\varphi_u(x) = 2^{-N} \| \sum_D \theta_D \prod_{i \in D} \pi_{u_i} x \|$ for suitable signs $\theta_D$ (i.d., $\theta_D = \pm 1$). It follows that

$$\varphi_u(x)^2 \le 4^{-N} \sum_D \sum_{D'} \Big|\Big\langle \prod_{i \in D} \pi_{u_i} x, \prod_{i \in D'} \pi_{u_i} x \Big\rangle\Big| \tag{38}$$

Every summand in the right hand side of (38) is at most 1. The number of summands for which the symmetric difference $D \Delta D'$ has cardinality $\overline{\overline{D \Delta D'}} \le N/4$ is $o(4^N)$ and thus for $n > n(\delta)$

is at most $\delta 4^N/4$. We consider the other summands and take a choice of $D$ and $D'$ so that $\overline{\overline{D\Delta D'}} \geq N/4$. We consider

$$f_{D,D'}(u) = f(u) = \langle \prod_{i\in D} \pi_{u_i} x \ , \ \prod_{i\in D'} \pi_{u_i} x \rangle$$

as a martingale taking as the sub-$\sigma$-algebras $\mathcal{B}_i$ those subsets of $(S^{n-1})^N$ which depend only on the $u_j$ with $j > N - i$. The indices $i$ which contribute a non-zero summand to the presentation of $f$ as a sum of martingale differences are exactly those which belong to $D\Delta D'$. We explain this by considering $i = 1$. If $1 \notin D \cup D'$, then obviously $f$ is already $\mathcal{B}_{N-1}$ measurable. The same is true if $1 \in D \cap D'$ since $\pi_{u_1}$ being orthogonal implies that

$$f(u) = \langle \prod_{i\in D\sim\{1\}} \pi_{u_i} x \ , \ \prod_{i\in D'\sim\{1\}} \pi_{u_i} x \rangle \ .$$

Assume that $1 \in D$ but $1 \notin D'$. Put $v = \prod_{i\in D\sim\{1\}} \pi_{u_i} x$, $w = \prod_{i\in D'} \pi_{u_i} x$. Then

$$f(u) = \langle \pi_{u_1} v, w\rangle = \langle v, w\rangle - 2\langle v, u_1\rangle\langle w, u_1\rangle \ .$$

Put

$$f_{N-1}(u) = E\{f \mid \mathcal{B}_{N-1}\} = \langle v, w\rangle - 2 \int_{S^{n-1}} \langle v, u\rangle\langle w, u\rangle d\sigma_n(u)$$

$$= \left(1 - \frac{2}{n}\right) \langle v, w\rangle \ .$$

Note that

$$f - f_{N-1} = \frac{2}{n}\langle v, w\rangle - 2\langle v, u_1\rangle\langle w, u_1\rangle$$

and that $\|\langle \,\cdot\, v\rangle^2\|_{L_{\psi_1}(S^{n-1})} \approx n^{-1}$. Hence, for some $c_3 > 0$ and all $v$ and $w$

$$E\big(\exp(c_3 n|f - f_{N-1}|) \mid \mathcal{B}_{N-1}\big) \leq 2 \ .$$

The function $f_{N-1}$ is of the same form as $f$ only with the additional factor $\left(1 - \frac{2}{n}\right)$. We can now repeat the same reasoning with $i = 2$ and so on. We get in particular that

$$E(f) = \int_{(S^{n-1})^N} f(d\sigma_n)^N = \left(1 - \frac{2}{n}\right)^{N'} \quad \text{where} \quad N' = \overline{\overline{D\Delta D'}} \ .$$

Since we assume that $N' \geq N/4$ we can, by a suitable choice of $c_1$ (in the definition of $N$), ensure that $|E(f)| < \delta/4$. Since $\delta/2N \leq 4/c_3 n$ for $\delta < \delta_0$ we may apply Proposition 9 and get that for some absolute $c_4 > 0$

$$\begin{aligned}\operatorname{Prob}\{|f| \geq \delta/2\} &\leq \operatorname{Prob}\left\{|f - E(f)| \geq \delta/4\right\} \\ &\leq 2\exp\left(-\delta^2(c_3 n)^2/16 \cdot 4 \cdot N'\right) \leq \exp(-c_4 \delta^2 n/|\log \delta|) \ .\end{aligned}$$

For $D$ and $D'$ with $\overline{\overline{D\Delta D'}} \geq N/4$ put

$$\chi_{D,D'}(u) = \begin{cases} 1 & \text{if } |f_{D,D'}(u)| \geq \delta/2 \\ 0 & \text{otherwise .} \end{cases}$$

Then

$$\begin{aligned} \operatorname{Prob}\left\{u \; ; \; \chi_{D,D'}(u) = 1 \text{ for at least } \delta \cdot 4^N/4 \text{ pairs } D, D'\right\} &\leq 4(4^N\delta)^{-1} E\Big(\sum_D \sum_{D'} \chi_{D,D'}\Big) \leq \\ &\leq 4\delta^{-1}\exp(-c_4\delta^2 n/|\log\delta|) \;. \end{aligned}$$

If $u$ is such that for at most $\delta \cdot 4^N/4$ pairs $D, D'$ with $\overline{\overline{D\Delta D'}} \geq N/4$ we have $|f_{D,D'}(u)| > \delta/2$ then clearly

$$\varphi_u(x) < \delta/4 + \delta/4 + \delta/2 = \delta \;. \qquad \square$$

Our next goal is to strengthen Lemma 10 by showing that $\varphi_u(x)$ is small for most $u$ and all $x$ and not only a single $x$.

**Lemma 11.** *Let $\varphi_u(x)$ be as in (36). Then for $N = c_5\delta^{-2}|\log\delta|^3 n$, $0 < \delta < \delta_0$ and $n > n(\delta)$*

$$\operatorname{Prob}\left\{u \; ; \; \varphi_u(x) > \delta \text{ for some } x \in S^{n-1}\right\} \leq \exp(-c_6 n|\log\delta|) \;. \tag{39}$$

**Proof:** Let $k$ and $N_1$ be integers, let $u \in (S^{n-1})^{N_1}$ and $v \in (S^{n-1})^{kN_1}$. We consider $(u,v)$ as an element of $(S^{n-1})^{(k+1)N_1}$ and there is an obvious meaning to $\varphi_{u,v}(x)$ for $x \in S^{n-1}$. Note that $\varphi_{u,v}(x) \leq \varphi_v(x)$ for all $u, v$ and $x$. We assume that $\varepsilon, \delta$ and $N_1$ are such that for all $x \in S^{n-1}$

$$\operatorname{Prob}\left\{u \in (S^{n-1})^{N_1} \; ; \; \varphi_u(x) \geq \delta/2\right\} \leq \varepsilon \;. \tag{40}$$

Since, for each fixed $v_0$, $\varphi_{u,v_0}(x)$ is bounded from above by the arithmetic average of $2^{kN_1}$ of expressions of the form $\varphi_u(z)$ for suitable vectors $z$ in $S^{n-1}$ we get by the same reasoning as that used at the end of the proof of Lemma 10 that

$$\operatorname{Prob}\left\{u \in (S^{n-1})^{N_1} \; ; \; \varphi_{u,v_0}(x) \geq \delta\right\} \leq 2\varepsilon/\delta \;.$$

Hence, for all $x \in S^{n-1}$,

$$\operatorname{Prob}\left\{(u,v) \; ; \; \varphi_{u,v}(x) \geq \delta\right\} \leq \operatorname{Prob}\left\{v \; ; \; \varphi_v(x) \geq \delta\right\} \cdot 2\varepsilon/\delta \;. \tag{41}$$

By induction on $k$ we deduce that

$$\operatorname{Prob}\left\{v \in (S^{n-1})^{kN_1} \; ; \; \varphi_v(x \;\geq \delta\right\} \leq (2\varepsilon/\delta)^k \;. \tag{42}$$

By Lemma 10, (40) holds with $N_1 = c_1 n|\log\delta|$ and $\varepsilon = \exp(-8c_2\delta^2 n/|\log\delta|)$. Hence by taking $k \sim |\log\delta|^2\delta^{-2}$ in (42) we get that for $N \approx n\delta^{-2}|\log\delta|^3$, $x \in S^{n-1}$, $0 < \delta < \delta_0$ and $n > n(\delta)$

$$\operatorname{Prob}\left\{v \in (S^{n-1})^N \ ; \ \varphi_v(x) \geq \delta\right\} < (\delta/4)^{2n} \ .$$

Let $\{x_\alpha\}_{\alpha\in A}$ be a $\delta$ net in $S^{n-1}$ with $\overline{\overline{A}} \leq (4/\delta)^n$. Then

$$\operatorname{Prob}\left\{v \ ; \varphi_v(x_\alpha) \geq \delta \ \text{ for some } \alpha \in A\right\} < (\delta/4)^n \ . \tag{43}$$

A trivial argument, similar to the proof of Lemma 3, shows that if $\varphi_v(x_\alpha) \leq \delta$ for all $x_\alpha$ then $\varphi_v(x) \leq \delta/(1-\delta) \leq 2\delta$ for all $x \in S^{n-1}$ and this concludes the proof of our assertion. □

**Remark.** It is evident from the proof that if we take a larger value of $N$ we get (39) with a smaller value for the probability. This holds also for the results below, in particular for Theorem 14.

Our next lemma is a simple exhaustion argument.

**Lemma 12.** *Assume that $\{w_\alpha\}_{\alpha\in A}$ is a family of vectors in $S^{n-1}$ so that for some $\delta > 0$*

$$\sup_{y\in S^{n-1}} \overline{\overline{A}}^{-1} \sum_{\alpha\in A} |\langle w_\alpha, y\rangle| \leq \delta \ .$$

*Let $0 < \lambda < 1$ and let $k$ be an integer. Then there are disjoint families $\mathcal{F}_\beta = \{\beta_i\}_{i=1}^k$, $\beta \in B$, of elements of $A$ so that $\overline{\overline{\bigcup_{\beta\in B} \mathcal{F}_\beta}} \geq (1-\lambda)\overline{\overline{A}} - k$ and so that for every $\beta \in B$ there is an orthonormal set of vectors $\{v_{\beta_i}\}_{i=1}^k$ satisfying*

$$|||v_{\beta_i} - w_{\beta_i}||| \leq \delta 4^k/\lambda \ , \qquad 1 \leq i \leq k \ . \tag{44}$$

**Proof:** Assume that we already have families $\mathcal{F}_\beta$ as above and that $\overline{\overline{\bigcup_\beta \mathcal{F}_\beta}} \leq (1-\lambda)\overline{\overline{A}} - k$. Then we can construct by induction on $i$ distinct elements $\{a_i\}_{i=1}^k$ in $A$ which do not belong to the $\mathcal{F}_\beta$ and for which

$$\sum_{j=1}^{i} |\langle w_{a_j}, w_{a_{i+1}}\rangle| \leq i\delta/\lambda \ , \qquad 1 \leq i \leq k-1 \ . \tag{45}$$

Indeed, put $\widetilde{A} = A \sim \left(\{a_j\}_{j=1}^i \cup \bigcup_\beta \mathcal{F}_\beta\right)$ then $\overline{\overline{\widetilde{A}}} \geq \lambda\overline{\overline{A}}$ and by our assumption

$$\sum_{j=1}^{i} \sum_{\alpha\in\widetilde{A}} |\langle w_{a_j}, w_\alpha\rangle| \leq \sum_{j=1}^{i} \sum_{\alpha\in A} |\langle w_{a_j}, w_\alpha\rangle| \leq i\delta\overline{\overline{A}} \ .$$

Hence there is an $\alpha \in \widetilde{A}$ which can serve as $a_{i+1}$ in (45). Let $\{v_{a_i}\}_{i=1}^k$ be the system obtained from $\{w_{a_i}\}_{i=1}^k$ by the Gram Schmidt orthogonalization procedure. Then

$$\begin{aligned}|||v_{a_{i+1}} - w_{a_{i+1}}||| &\le 2\sum_{j=1}^{i} |\langle w_{a_{i+1}}, v_{a_j}\rangle| \le \\ &\le 2\sum_{j=1}^{i} |\langle w_{a_{i+1}}, w_{a_j}\rangle| + 2\sum_{j=1}^{i} |||w_{a_j} - v_{a_j}||| \le \\ &\le 2\delta i/\lambda + 2\sum_{j=1}^{i} |||w_{a_j} - v_{a_j}||| \,.\end{aligned}$$

From this inequality we deduce by induction on $i$ that

$$|||v_{a_i} - w_{a_i}||| \le \delta \cdot 4^i/\lambda\,, \qquad 1 \le i \le k\,.$$

□

The next lemma is a concentration of measure result on the orthogonal group.

**Lemma 13.** *Let $\|\cdot\|$ be a norm on $R^n$ so that $\|x\| \le |||x|||$ for all $x$. Let $\{x_i\}_{i=1}^k$ be orthonormal vectors (with respect to $|||\ \ |||$), and let $\mu_n$ be Haar measure on the orthogonal group $O^n$. Then*

$$\mu_n\{U \in O^n\ ;\ \big|k^{-1}\sum_{i=1}^{k} \|Ux_i\| - \int_{O^n} \|Ux\| d\mu_n(U)\big| \ge \varepsilon\} \le \exp(-c\varepsilon^2 nk)\,. \tag{46}$$

**Proof:** We use the following known fact (cf. [MS], p. 29). Let $F$ be a function from $O^n$ to $\mathbb{R}$ which has Lipschitz constant $\lambda$ if $O^n$ is taken in the Hilbert-Schmidt metric, i.e.,

$$|F(U) - F(V)| \le \lambda |||U - V|||_{\text{H.S.}}$$

Then

$$\mu_n\{U \in O^n\ ;\ \big|F(U) - \int_{O^n} F(U) d\mu_n(U)\big| \ge \varepsilon\} \le \exp(-c\varepsilon^2 n/\lambda^2)\,.$$

We use here the function $F(U) = k^{-1}\sum_{i=1}^{k} \|Ux_i\|$. Note that

$$\begin{aligned}|F(U) - F(V)| &\le k^{-1}\sum_{i=1}^{k} \|(U-V)x_i\| \le \\ &\le k^{-1}\sum_{i=1}^{k} |||(U-V)x_i||| \le k^{-1/2}\big(\sum_{i=1}^{k} |||(U-V)x_i|||^2\big)^{1/2} \le \\ &\le k^{-1/2}|||U-V|||_{\text{H.S.}}\end{aligned}$$

Thus we may take $\lambda = k^{-1/2}$ in the estimate above and this proves (46). □

We are now ready to prove the main result.

**Theorem 14.** *Let $K$ be a symmetric convex body in $\mathbb{R}^n$ and let $\varepsilon > 0$. If $n > n_0(\varepsilon)$ and if we perform $N = cn\log n + c(\varepsilon)n$ random Minkowski symmetrizations on $K$ we obtain with probability $1 - \exp\big(-\widetilde{c}(\varepsilon)n\big)$ a body $\widetilde{K}$ which satisfies*

$$(1-\varepsilon)rB^n \subset \widetilde{K} \subset (1+\varepsilon)rB^n$$

*for a suitable $r = r(K)$ .*

The meaning of random Minkowski symmetrizations was explained in the beginning of this section. As usual $c$ denotes an absolute constant while $c(\varepsilon)$ and $\widetilde{c}(\varepsilon)$ denote positive constants depending only on $\varepsilon$.

As was the case in section 1 it is more convenient to carry out the proof in the dual space because of the simple behaviour of the support function of a Minkowski sum. We first formulate the theorem in the dual form

**Theorem 14′.** *Let $\|\cdot\|$ be a norm in $\mathbb{R}^n$ and let $N = cn\log n + c(\varepsilon)n$ (where $\varepsilon > 0$ and $n > n_0(\varepsilon)$). Then for all $\{u_i\}_{i=1}^N$ in $(S^{n-1})^N$ with the exception of a set of $(\sigma_n)^N$ measure at most $1 - \exp\big(-\widetilde{c}(\varepsilon)n\big)$ we have*

$$(47)\qquad (1-\varepsilon)r|||x||| \le 2^{-N}\sum_D \Big\|\prod_{i\in D}\pi_{u_i}x\Big\| \le (1+\varepsilon)r|||x||| , \qquad x \in X$$

*where the sum is taken over all subsets $D$ of $\{1,\dots,N\}$ and*

$$(48)\qquad r = \int_{S^{n-1}} \|x\|d\sigma_n(x) = \int_{O^n} \|Ux\|d\mu_n(U) .$$

**Proof:** We assume as we may that $\|x\| \le |||x|||$ for all $x$ and that for some vector $e$, $\|e\| = |||e||| = 1$. Then clearly $\|x\| \ge |\langle e, x\rangle|$ and thus

$$(49)\qquad 1 \ge r \ge \int_{S^{n-1}} |\langle e,x\rangle| d\sigma_n(x) \approx 1/\sqrt{n} .$$

We shall perform the operations in steps. We shall show below that the following holds.

**Statement a.** *If $r < 1/8$, then with $N_1 < c_1 n$ we have*

$$(50)\qquad \|x\|_2 = 2^{-N_1}\sum_D \Big\|\prod_{i\in D}\pi_{u_i}x\Big\| \le \tfrac{1}{2}|||x||| , \qquad x \in \mathbb{R}^n$$

*with probability of at least $1 - \exp(-x_2 n)$.*

In proving Statement a, we shall use Lemma 11 for a certain absolute value of $\delta$. Note that because of the rotation invariance of $\sigma_n$ and the fact that the $\pi_u$ are unitary operators we have that $r = \int_{S^{n-1}} \|x\|_2 d\sigma_n(x)$, i.e., $r$ does not change in this operation. If $r < 1/16$ then by applying Statement **a** to the norm $2\|x\|_2$ we obtain a norm $2\|x\|_3$ which satisfies $\|x\|_3 \leq \frac{1}{4}|||x|||$. (Statement **a** uses only the normalizing condition $\|x\| \leq |||x|||$. The requirement on the existence of $e$ with $\|e\| = |||e||| = 1$ is used only to obtain a lower estimate on $r$ in (49). As mentioned, $r$ is fixed throughout the whole procedure.) Note that because of the special form of our operation if follows that $\|x\|_3$ can be obtained directly from $\|x\|$ via the formula.

$$\|x\|_3 = 2^{-2N_1} \sum_D \|\prod_{i\in D} \pi_{u_i} x\|$$

where here of course, the sum is over all subsets $D$ of $\{1, 2, \ldots, 2N_1\}$. In view of (49) we will get, after $\frac{1}{2}\log n$ steps, to a norm for which the ratio between its maximum and average on $S^{n-1}$ is at most 8. So far we have performed $\frac{1}{2}c_1 n \log n$ symmetrizations and the probability that we fail to get a good norm is at most $\frac{1}{2}\log n \exp(-c_2 n) \leq \exp(-c_3 n)$ (for $n \geq n_0$). Thus to conclude the proof of the theorem it suffices to verify besides Statement **a** also

**Statement b.** *If $r \geq 1/8$, then for all $\varepsilon > 0$ and $n \geq n(\varepsilon)$ and with $N_2 = c(\varepsilon)n$*

$$(51) \qquad (1-\varepsilon)r|||x||| \leq 2^{-N_2} \sum_D \|\prod_{i\in D} \pi_{u_i} x\| \leq (1+\varepsilon)r|||x|||\ , \qquad x \in X$$

*with probability of at least $1 - \exp\big(-c_1(\varepsilon)n\big)$.*

**Proof of Statement a.** We shall use as $\delta, \lambda$ and $k$ absolute constants which will be determined shortly. By Lemmas 11 and 12 we get with probability $1 - (\delta/4)^n$ (see (43)) that if $N_1 = c_5\delta^{-2}|\log\delta|^3 n$ we can decompose, for every $x \in S^{n-1}$, the set consisting of the $2^{N_1}$ vectors $\prod_{i\in D} \pi_{u_i} x$ into disjoint families $\{\mathcal{F}_\beta\}_{\beta\in B}$ consisting of $k$ elements each and a remainder of size $\lambda 2^{N_1}$. For each such family $\mathcal{F} = \{w_j\}_{j=1}^k$ there are orthonormal vectors $\{v_j\}_{j=1}^k$ so that $|||w_j - v_j||| \leq \delta \cdot 4^k/\lambda$ for all $j$. By Lemma 13

$$(52) \qquad \mu_n\{U \in O^n\ ;\ k^{-1}\sum_{j=1}^k \|Uv_j\| - r \geq 1/8\} \leq \exp(-c_4 kn)\ .$$

Hence

$$\mu_n\{U \in O^n\ ;\ k^{-1}\sum_{j=1}^k \|Uw_j\| - r \geq 1/8 + \delta\cdot 4^k/\lambda\} \leq \exp(-c_4 kn)\ .$$

Consequently the set of $U \in O^n$ for which

$$k^{-1}\sum_{j=1}^k \|Uw_{\beta_j}\| \leq r + 1/8 + \delta \cdot 4^k/\lambda$$

for at least $(1-\lambda)\overline{\overline{B}}$ of the families $\mathcal{F}_\beta$ have a measure not smaller than $1-\lambda^{-1}\exp(-c_4kn)$ (use the same argument as that at the end of the proof of Lemma 10). We deduce from this that there is a set of $\{u_i\}_{i=1}^{N_1}$ of measure at least $1-(\delta/4)^n$ so that for all $x \in S^{n-1}$ the set of $U \in O^n$ for which

$$2^{-N_1}\sum_D \|U\prod_{i\in D}\pi_{u_i}x\| \le r + \frac{1}{8} + \delta\cdot 4^k/\lambda + \lambda + \lambda \tag{53}$$

has measure $\ge 1-\lambda^{-1}\exp(-c_4kn)$. By Fubini's theorem we deduce that for some $U_0 \in O^n$, (53) holds for all $\{u_i\}_{i=1}^{N_1}$ in a set of measure $\ge 1-(\delta/4)^n-\lambda^{-1}\exp(-c_4kn)$ and all $x$ in a set of measure $\ge 1-\lambda^{-1}\exp(-c_4kn)$. Since

$$U\prod_{i\in D}\pi_{u_i}x = \prod_{i\in D}\pi_{Uu_i}Ux$$

and the map $\{u_i\}_{i=1}^{N_1} \to \{Uu_i\}_{i=1}^{N_1}$ is measure preserving there is no loss of generality to assume that $U_0 =$ identity. We choose $\lambda, \delta$ and $k$ so that $2\lambda = 1/16$, $\delta 4^k/\lambda = 1/16$ and then $\lambda^{-1}\exp(-c_4kn)$ takes the form $\exp(-c_5|\log\delta|n)$ and we are still free to choose $\delta$ as small as we please. For $\delta$ sufficiently small a set on $S^{n-1}$ with measure $\ge 1-\exp(-c_6|\log\delta|n)$ contains a $1/4$ net. Thus for a $1/4$ net on $S^{n-1}$ and a set of $\{u_i\}_{i=1}^{N_1}$ of measure $\ge 1-2\exp(-c_6|\log\delta|n)$ we have

$$2^{-N_1}\sum_D \|\prod_{i\in D}\pi_{u_i}x\| \le 3/8\ . \tag{54}$$

This implies that (54) holds for the same $\{u_i\}_{i=1}^{N_1}$ and all $x \in S^{n-1}$ with $3/8$ replaced by $3/8\cdot\left(1-\frac{1}{4}\right)^{-1} = \frac{1}{2}$ and this proves (50).

**Proof of Statement b.** The proof is almost identical to the proof of Statement **a** only now $k, \delta$ and $\lambda$ will depend on the given $\varepsilon$.

Instead of (52) we will get now

$$\mu_n\{U \in O^n\ ;\ |k^{-1}\sum_{j=1}^{k}\|U_{V_j}\| - r| \ge \varepsilon r/4\} \le \exp(-c_7\varepsilon^2 kn)\ .$$

We shall choose now $k$ and $\lambda$ and $\delta$ so that $2\lambda = \varepsilon/32$ $(\le \varepsilon r/4)$ and $\delta\cdot 4^k/\lambda = \varepsilon/32$. This gives for the measure of the exceptional set (of $\{u_i\}_{i=1}^{N_1}$ or of $x$) an estimate of $\exp\big(-c_8\varepsilon^2|\log(\varepsilon/\delta)|n\big)$. For non exceptional $\{u_i\}_{i=1}^{N_1}$ and $x \in S^{n-1}$ we get that

$$|2^{-N_1}\sum_D \|\prod_{i\in D}\pi_{u_i}x\| - r| \le 3\varepsilon r/4\ . \tag{55}$$

We choose now $\delta(\varepsilon)$ small enough so that by an argument using nets on $S^{n-1}$ we can ensure that (55) holds for all $x \in S^{n-1}$ with $3\varepsilon r/4$ replaced by $\varepsilon r$. $\square$

**Remarks.** 1. The theorem and its proof are valid also for non symmetric convex bodies $K$ which contains $O$ in their interior. We just have to work throughout with norms such that $\|-x\| \neq \|x\|$.

2. The dependence of the $c(\varepsilon)$ (appearing in the statement of Theorem 14) on $\varepsilon$ which we get is $c(\varepsilon) \approx \exp(\alpha\varepsilon^{-2}|\log\varepsilon|)$.

3. The dependence of $N$ on $n$ in Theorem 14 is optimal. Let $e$ be some vector in $\mathbb{R}^n$ with $|||e||| = 1$ and consider the norm $\|x\| = |\langle e, x\rangle| + |||x|||/n^{1/2}$. For this norm $r = \int_{S^{n-1}} \|x\| d\sigma_n(x) \approx n^{-1/2}$. After $k$ random symmetrizations we get from the given norm a norm whose expected value at $x$ is $\geq \left(1 - \frac{2}{n}\right)^k |\langle e, x\rangle| + |||x|||/n^{1/2}$. Thus unless $k \geq \frac{1}{2} n \log n$ we get a norm whose expected value for $e$ is significantly larger than its average (i.e., $r$). If we consider also deterministic symmetrizations then it is no longer clear to us that $cn$ symmetrizations will not suffice always. For the deterministic case it is evident that in general we need at least $n - c\log n$ steps. Indeed, if we symmetrize with respect to $\{u_i\}_{i=1}^k$ we do not change the norm at all of vectors $x$ which are orthogonal to all the $\{u_i\}_{i=1}^k$. For some norms (e.g., the norm in $\ell_\infty^n$) it is known that there are no subspaces of dimension $\geq c\log n$ which are Euclidean up to 2,say.

## References


[B.L] J. Bourgain and J. Lindenstrauss, Projection bodies, This volume.

[B.L.M] J. Bourgain, J. Lindenstrauss and V. Milman, Approximation of zonoids by zonotopes, Preprint I.H.E.S., September 1987, 62pp, to appear in Acta Math.

[Bo] C. Borell, The Brunn-Minkowski inequality in Gauss spaces, Inventiones Math. 30 (1975), 207-216.

[D.S] P. Diaconis and M. Shahshahani, Products of random matrices as they arise in the study of random walks on groups, Contemporary Math. 50 (1986), 183-195.

[F.L.M] T. Figiel, J. Lindenstrauss and V.D. Milman, The dimensions of almost spherical sections of convex bodies, Acta Math. 139 (1977), 53-94.

[G] Y. Gordon, Some inequalities for Gaussian processes and applications, Israel J. Math. 50 (1985), 265-289.

[G.M] M. Gromov and V.D. Milman, Brunn theorem and a concentration of volume of convex bodies, GAFA Seminar Notes, Israel 1983-1984.

[H] U. Haagerup, The best constants in the Khinchine inequality, Studia Math, 70 (1982), 231-283.

[J] F. John, Extremum problems with inequalities as subsidiary conditions, Courant Anniversary Volume, Interscience, New York, 1948, 187-204.

[K] B.S. Kashin, Sections of some finite dimensional sets and classes of smooth functions, Izv. ANSSSR, ser. mat. 41 (1977), 334-351 (Russian).

[Ka] J.P. Kahane, Series of Random Functions, Heath Math. Monographs, Lexington, Mass., Heath & Co., 1968.

[Kw] S. Kwapień, Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients, Studia Math. 44 (1972), 583-595.

[L.T] J. Lindenstrauss and L. Tzafriri, Classical Banach spaces, vol II, Function spaces, Ergebnisse der Math., v. 97, Springer Verlag 1979.

[L.S] H. Landau and L. Shepp, On the supremum of a Gaussian process, Sankhya A32 (1970), 369-378.

[Ma.P] M.B. Marcus and G. Pisier, Random Fourier series with applications to harmonic analysis, Ann. Math. Studies, 101, Princeton 1981.

[M] V.D. Milman, Random subspaces of proportional dimension of finite dimensional normed spaces: Approach through the isoperimetric inequality, Banach spaces, Proc. Missouri Conference 1984, Springer Lecture Notes #1166, 106-115.

[M.P] V.D. Milman and G. Pisier, Banach spaces with a weak cotype 2 property, Israel J. Math., 54 (1986), 139-158.

[M.S] V.D. Milman and G. Schechtman, Asymptotic theory of finite dimensional normed spaces, Springer Lecture Notes #1200 (1986).

[S] G. Schechtman, More on embedding subspaces of $L_p$ in $\ell_r^n$, Compositio Math. 61 (1987), 159-170.

[Sz] S.J. Szarek, On the best constant in the Khinchine inequality, Studia Math. 58 (1976), 197-208.

[Ta] M. Talagrand, An isoperimetric theorem on the cube and the Khinchine-Kahane inequalities, Preprint.