Notes for seminar on volume ratio

These are notes for a few seminar talks on volume ratio and related notions delivered in fall 2009 and winter 2010.

Contents

1 Introduction

The *volume ratio* of a convex body K in $\mathbb{R}^n$ is Nov 23

$$\mathrm{vr}(K) = \inf\left\{ \left(\frac{\mathrm{vol}(K)}{\mathrm{vol}(E)}\right)^{1/n} : E \text{ is an ellipsoid and } E \subseteq K \right\} .$$

It is easy to see that this quantity is affine invariant, meaning that for any invertible affine map T we have $\mathrm{vr}(TK) = \mathrm{vr}(K)$. Thus

$$\mathrm{vr}(K) = \mathrm{vr}(K_{\mathrm{John}}) = \left(\frac{\mathrm{vol}(K_{\mathrm{John}})}{\mathrm{vol}(B_2^n)}\right)^{1/n} ,$$

where K_{John} denotes an affine image of K which has B_2^n as its maximum volume ellipsoid, as in John's theorem.

In fact, slightly more is true than affine invariance: we have

$$\mathrm{vr}(K) \le \tilde{d}(K, L)\,\mathrm{vr}(L) . \tag{1}$$

Here $\tilde{d}$ is one of the two natural generalizations of Banach–Mazur distance to convex bodies that are not necessarily symmetric. The usual one is

$$d(K, L) = \inf\{\lambda \colon \lambda > 0 \text{ and } SL \subseteq TK \subseteq \lambda SL \text{ for some invertible affine } S, T\} ,$$

but if we wish to allow the "inner" and "outer" L to be negative homothets, we use instead

$$\tilde{d}(K, L) = \inf\{|\lambda| \colon SL \subseteq TK \subseteq \lambda SL \text{ for some invertible affine } S, T\} .$$

Clearly $\tilde{d}(K, L) \le d(K, L)$; we have equality if one (or both) of K and L is symmetric.[1]

To prove (1), consider any ellipsoid E, any invertible affine maps S, T, and any real number λ such that

$$E \subseteq SL \subseteq TK \subseteq \lambda SL .$$

Then

$$\mathrm{vr}(K) = \mathrm{vr}(TK) \le \left(\frac{\mathrm{vol}(TK)}{\mathrm{vol}(E)}\right)^{1/n} \le \left(\frac{\mathrm{vol}(\lambda SL)}{\mathrm{vol}(E)}\right)^{1/n} = |\lambda| \left(\frac{\mathrm{vol}(SL)}{\mathrm{vol}(E)}\right)^{1/n} .$$

[1] As Sasha pointed out, this statement is slightly trickier than it appears. We'd like to say that if $L \subseteq K \subseteq -\lambda L$ (where $\lambda > 0$) and L is symmetric, then we can just replace $-\lambda L$ by λL. But this argument assumes the centre of homothety (i.e., the origin wrt which we scale L to $-\lambda L$) is the same as the centre of symmetry (i.e., the origin for which $L = -L$), and this is not part of the hypothesis. So we need to be a little more careful. But the idea is easy: if L is symmetric, then any negative homothet of L is also a positive homothet (with the same absolute ratio, but with a different centre), so if K can be sandwiched between L and a negative homothet of some ratio then ipso facto it can be sandwiched between L and a positive homothet of the same ratio.

Taking the infimum over all E such that $E \subseteq SL$ yields

$$\mathrm{vr}(K) \le |\lambda|\,\mathrm{vr}(SL) = |\lambda|\,\mathrm{vr}(L)\,,$$

and then taking the infimum over S, T, λ yields the desired result.

An immediate corollary is that

$$\mathrm{vr}(K) \le \tilde{d}(K, B_2^n)\,\mathrm{vr}(B_2^n) = \tilde{d}(K, B_2^n) \le d(K, B_2^n)\,.$$

From John's theorem we have well-known upper bounds on $d(K, B_2^n)$, yielding

$$\mathrm{vr}(K) \le \begin{cases} \sqrt{n} & \text{if } K \text{ is symmetric,} \\ n & \text{in general.} \end{cases} \tag{2}$$

This estimate for symmetric K is pretty good (as we will see, it's the right order), but the estimate for the general case is quite bad[2]; the correct upper bound is again $c\sqrt{n}$.

2 Ball's precise upper bounds on volume ratio

Ball's Theorem[3] For any convex body K in $\mathbb{R}^n$,

$$\mathrm{vol}(K_{\mathrm{John}}) \le \begin{cases} \mathrm{vol}(B_\infty^n) & \text{if } K \text{ is symmetric,} \\ \mathrm{vol}(\triangle^n) & \text{in general.} \end{cases}$$

Here $\triangle^n$ denotes the regular simplex in John's position; it turns out that

$$\mathrm{vol}(\triangle^n) = \frac{n^{n/2}(n+1)^{(n+1)/2}}{n!} \sim ce^n\,. \tag{3}$$

(The asymptotic statement follows by Stirling's approximation.)

A corollary of Ball's theorem is that the cube B_∞^n has the highest volume ratio of all symmetric convex bodies, and the simplex $\triangle^n$ has the highest volume ratio in general. (Another corollary is the reverse isoperimetric inequality; see [4].)

To prove Ball's theorem we need a few facts about John's position.

John's Theorem[4] Every convex body K in $\mathbb{R}^n$ contains a unique ellipsoid of maximum volume. That ellipsoid is B_2^n if and only if $B_2^n \subseteq K$ and there exist contact points $(u_i)_1^m \subseteq \mathrm{bd}\,K \cap \mathrm{bd}\,B_2^n$ and positive weights $(c_i)_1^m$ such that

[2] As Sasha mentioned, we can get the right order by comparing K to $K \cap -K$: as shown by Stein [17], for any K there exists a choice of origin so that $\mathrm{vol}(K \cap -K) \ge \frac{1}{2^n}\mathrm{vol}(K)$, and it easily follows that $\mathrm{vr}(K) \le 2\,\mathrm{vr}(K \cap -K) \le 2\sqrt{n}$.

[3] The symmetric case appeared first in [1], and the general case in [2]. A complete treatment also appears in [4].

[4] The original paper of John [10] can be somewhat difficult to obtain through the usual methods; I have a copy if anyone is interested. Other proofs can be found in [3] and [8].

(i) $\sum c_i u_i \otimes u_i = \mathrm{Id}$, and

(ii) $\sum c_i u_i = 0$.

In condition (i), which is called John's decomposition of the identity, $x \otimes y$ denotes the map $\mathbb{R}^n \to \mathbb{R}$ given by

$$x \otimes y(t) = \langle x, t\rangle y \; .$$

This map is linear and has matrix yx^T (taking x, y to be column vectors); note that $\mathrm{tr}(x \otimes y) = \langle x, y\rangle$. Unpacking the definitions in condition (i) yields the equivalent statement

$$\forall x \in \mathbb{R}^n\colon \; \sum c_i \langle u_i, x\rangle u_i = x \; .$$

Taking the inner product with x on both sides yields

$$\forall x \in \mathbb{R}^n\colon \; \sum c_i \langle u_i, x\rangle^2 = |x|^2 \; . \tag{4}$$

(In fact this statement is equivalent to condition (i).) This is one respect (among many) in which the vectors u_i behave somewhat like an orthonormal basis.

Since $|u_i| = 1$, we have $\mathrm{tr}(u_i \otimes u_i) = 1$, so taking traces in condition (i) yields

$$\sum c_i = n \; . \tag{5}$$

Also since $|u_i| = 1$, the map $u_i \otimes u_i$ is the orthogonal projection onto span $\{u_i\}$. Since $-u_i$ has the same span, the map $(-u_i) \otimes (-u_i)$ is the same projection. If K is symmetric[5], then $-u_i \in \mathrm{bd}\, K \cap \mathrm{bd}\, B_2^n$, that is, $-u_i$ is also a contact point; thus replacing $c_i u_i \otimes u_i$ with

$$\frac{c_i}{2} u_i \otimes u_i + \frac{c_i}{2}(-u_i) \otimes (-u_i)$$

preserves condition (i) and makes condition (ii) automatic. Thus if K is symmetric then we can ignore condition (ii), or assume it freely, as we wish.

Since $u_i \in \mathrm{bd}\, K$, K has a supporting halfspace at u_i. Since $u_i \in \mathrm{bd}\, B_2^n$ and $B_2^n \subseteq K$, this halfspace also supports B_2^n at u_i; but B_2^n only has one supporting halfspace at u_i, and it is $\{x\colon \langle u_i, x\rangle \le 1\}$. Thus, defining

$$\widetilde{K} = \{x\colon (\forall i\colon \langle u_i, x\rangle \le 1)\} \, , \tag{6}$$

we have $K \subseteq \widetilde{K}$. When K is symmetric, we will instead define

$$\widetilde{K} = \{x\colon (\forall i\colon |\langle u_i, x\rangle| \le 1)\} \, , \tag{7}$$

[5] An important detail: if K is symmetric then its maximum volume ellipsoid has the same centre of symmetry. Indeed, if $K = -K$ and $E \subseteq K$, then $-E \subseteq K$, and so $\mathrm{conv}(E \cup -E) \subseteq K$; if $E \ne -E$ then one can show that E can be stretched (in the direction joining the centres of E and $-E$) to obtain an ellipsoid of larger volume in K. Thus if K is symmetric then $K_{\text{John}} = -K_{\text{John}}$.

taking advantage again of the fact that in this case, u_i and $-u_i$ are both contact points.

The second major tool for Ball's theorem is the following inequality.

Normalized Brascamp–Lieb Inequality[6] If $(u_i)_1^m$ are unit vectors and $(c_i)_1^m$ are positive real numbers such that $\sum c_i u_i \otimes u_i = \mathrm{Id}$, then for any measurable nonnegative functions $(f_i)_1^m$,

$$\int_{\mathbb{R}^n} \prod f_i(\langle u_i, x\rangle)^{c_i}\, dx \le \prod \left(\int_{\mathbb{R}} f_i(t)\, dt \right)^{c_i} .$$

This is another example of how the u_i behave somewhat like an orthonormal basis; indeed, if the u_i are an orthonormal basis then (taking all $c_i = 1$) we have equality.

Now we can prove Ball's theorem. The symmetric case is quick: let $K = -K$, let K be in John's position, with (c_i) and (u_i) as in John's theorem, and $\widetilde{K}$ as in (7). Then

$$\begin{aligned} \mathrm{vol}(K) \le \mathrm{vol}(\widetilde{K}) &= \int_{\mathbb{R}^n} [x \in \widetilde{K}]\, dx = \int_{\mathbb{R}^n} \prod [|\langle u_i, x\rangle| \le 1]^{c_i}\, dx \\ &\le \prod \left(\int_{\mathbb{R}} [|t| \le 1]\, dt \right)^{c_i} = \prod 2^{c_i} = 2^{\sum c_i} = 2^n = \mathrm{vol}(B_\infty^n) , \end{aligned}$$

making use of (5), the fact that $B_\infty^n = [-1, 1]^n$, and the Iverson bracket notation, whereby

$$[P] = \begin{cases} 1 & \text{if } P \text{ is true,} \\ 0 & \text{if } P \text{ is false.} \end{cases}$$

If we try to repeat this proof in the general case, with $\widetilde{K}$ defined by (6) instead of (7), we have $[\langle u_i, x\rangle \le 1]$ instead of $[|\langle u_i, x\rangle| \le 1]$, and at the end we are confronted with an integral over $(-\infty, 1]$ instead of over $[-1, 1]$. To remedy this, we introduce a weight function $w(t)$; we wish to end up with

$$\prod \left(\int_{\mathbb{R}} [t \le 1]\, w(t)\, dt \right)^{c_i} ,$$

where $w(t)$ will be chosen as some function that decays to zero (as $t \to -\infty$) fast enough that this integral is finite. In the previous step, then, we should have

$$\int_{\mathbb{R}^n} \prod [\langle u_i, x\rangle \le 1]^{c_i} w(\langle u_i, x\rangle)^{c_i}\, dx .$$

[6] This name and formulation are due to Ball [1]; the original version of Brascamp and Lieb is in [7]. Barthe ([5], [6]) gives a fairly simple proof of the inequality in this formulation, which is repeated in [4].

In order to introduce these weights at this point, we desire that

$$\prod w(\langle u_i, x\rangle)^{c_i} = 1 ,$$

or equivalently, taking logs,

$$\sum c_i \ln w(\langle u_i, x\rangle) = 0 .$$

Since $\sum c_i u_i = 0$, it is easy to see that we will obtain this condition by taking $w(t) = e^t$. Carrying out this plan yields

$$\operatorname{vol}(K) \le e^n ,$$

which as we see from (3) is asymptotically sharp.

The precise upper bound obtained by Ball requires another idea, which is motivated by the observation that the nicest way to present an n-dimensional simplex is as the convex hull of the $n+1$ standard basis vectors in $\mathbb{R}^{n+1}$, or equivalently, as a section of the positive orthant of $\mathbb{R}^{n+1}$. Likewise, we will replace our body $\widetilde{K}$ with a cone in $\mathbb{R}^{n+1}$: the contact points whose polar is $\widetilde{K}$ will be replaced by a set of normal vectors for the cone; we will construct these normal vectors so that they give a decomposition of the identity in $\mathbb{R}^{n+1}$, and apply the normalized Brascamp–Lieb inequality. (We will also use the idea of introducing an exponential weight to each factor.)

So, define weights $(d_i)_1^m$ and unit vectors $(v_i)_1^m$ in $\mathbb{R}^{n+1}$ by Nov 30

$$d_i = \frac{n+1}{n} c_i \quad \text{and} \quad v_i = \frac{1}{\sqrt{n+1}} \begin{bmatrix} \sqrt{n} u_i \\ -1 \end{bmatrix} \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1} .$$

As desired, these vectors, with these weights, give a decomposition of the identity as in John's theorem:

$$\begin{aligned} \sum d_i v_i v_i^T &= \sum \frac{n+1}{n} c_i \cdot \frac{1}{n+1} \begin{bmatrix} \sqrt{n} u_i \\ -1 \end{bmatrix} \begin{bmatrix} \sqrt{n} u_i^T & -1 \end{bmatrix} \\ &= \sum \frac{c_i}{n} \begin{bmatrix} n u_i u_i^T & \sqrt{n} u_i \\ \sqrt{n} u_i^T & 1 \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0^T & 1 \end{bmatrix} = I_{n+1} . \end{aligned}$$

They do not give a balanced configuration, since

$$\sum d_i v_i = \sum \frac{n+1}{n} c_i \cdot \frac{1}{\sqrt{n+1}} \begin{bmatrix} \sqrt{n} u_i \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -\sqrt{n+1} \end{bmatrix} .$$

Define the cone

$$C = \{ y \in \mathbb{R}^{n+1} : (\forall i\colon \langle v_i, y\rangle \le 0) \} .$$

For $y = (x, r) \in \mathbb{R}^n \times \mathbb{R}$, we have

$$\begin{aligned} y \in C &\iff (\forall i\colon \langle v_i, y\rangle \le 0) \\ &\iff (\forall i\colon \langle \sqrt{n} u_i, x\rangle - r \le 0) \\ &\iff (\forall i\colon \langle u_i, x\rangle \le \tfrac{r}{\sqrt{n}}) \\ &\iff r \ge 0 \text{ and } x \in \tfrac{r}{\sqrt{n}}\widetilde{K} \end{aligned}$$

(The condition $r \ge 0$ is obtained because $\sum c_i u_i = 0$, and so the values $\langle u_i, x\rangle$ cannot all be negative.)

By the normalized Brascamp–Lieb inequality,

$$\int_{\mathbb{R}^{n+1}} \prod [\langle v_i, y\rangle \le 0]\, e^{d_i \langle v_i, y\rangle}\, dy \le \prod \left(\int_{\mathbb{R}} [t \le 0]\, e^t\, dt \right)^{d_i} = 1\,.$$

On the other hand,

$$\begin{aligned} &\int_{\mathbb{R}^{n+1}} \prod [\langle v_i, y\rangle \le 0]\, e^{d_i \langle v_i, y\rangle}\, dy \\ &= \int_{\mathbb{R}^{n+1}} [y \in C]\, e^{\langle \sum d_i v_i, y\rangle}\, dy \\ &= \int_0^\infty \int_{\mathbb{R}^n} [x \in \tfrac{r}{\sqrt{n}}\widetilde{K}]\, e^{-r\sqrt{n+1}}\, dx\, dr \\ &= \int_0^\infty \left(\frac{r}{\sqrt{n}} \right)^n e^{-r\sqrt{n+1}}\, dr\, \mathrm{vol}(\widetilde{K}) \\ &= \int_0^\infty \left(\frac{t}{\sqrt{n(n+1)}} \right)^n e^{-t} \frac{1}{\sqrt{n+1}}\, dt\, \mathrm{vol}(\widetilde{K}) \qquad (t = r\sqrt{n+1}) \\ &= \frac{n!}{n^{n/2}(n+1)^{(n+1)/2}}\, \mathrm{vol}(\widetilde{K})\,. \end{aligned}$$

Therefore

$$\mathrm{vol}(\widetilde{K}) \le \frac{n^{n/2}(n+1)^{(n+1)/2}}{n!} = \mathrm{vol}(\triangle^n)\,,$$

as desired.

3 Low volume ratio yields somewhat Euclidean sections

My treatment of this topic follows [15], chapter 6.[7]

[7] Nicole explained the history: in 1977, Kašin [12] (whose name is rendered “Kashin” in some transliteration systems) showed, for $K = B_1^n$, that there exists a decomposition $\mathbb{R}^n = F \oplus F^\perp$ of the type described in corollary 7; in 1978, Szarek [18] gave a different proof, using the method shown here, but did not explicitly isolate the notion of volume ratio; Szarek and Tomczak-Jaegermann [19] did that in 1980, and stated explicitly the result which I here call “Szarek’s theorem”. Perhaps it should be called the Kašin–Szarek theorem, or the Szarek–Tomczak theorem.

Szarek's Theorem Let K be a convex body in $\mathbb{R}^n$ with $B_2^n \subseteq K$ and

$$\left(\frac{\mathrm{vol}(K)}{\mathrm{vol}(B_2^n)}\right)^{1/n} \leq A\,.$$

Then, with high probability, a random k-dimensional subspace F satisfies

$$\forall x \in F\colon \|x\|_K \leq |x| \leq (4\pi A)^{n/(n-k)}\|x\|_K\,.$$

Here F is random in the sense of the uniform (i.e., rotationally invariant) probability measure on the Grassmannian $G_{n,k}$, and the statement holds with probability at least $1 - \frac{1}{2^n}$.

The inequality $\|x\|_K \leq |x|$ is immediate from the hypothesis that $B_2^n \subseteq K$, so we need only prove the second inequality. Furthermore, by homogeneity it is enough to consider $x \in S^{n-1} \cap F$. The plan of the proof is:

(A) Since the volume of K is small, its radial function is small (i.e., $\|\cdot\|_K$ is large) at an average point of S^{n-1}.

(B) Thus the norm is large at an average point of an average section of S^{n-1}.

(C) Thus the norm is large at most points of most sections of S^{n-1}.

(D) Thus the norm is large at all points of most sections of S^{n-1}.

We need several lemmas, of whose proofs I will say very little.

Lemma 1 If $f\colon \mathbb{R}^n \to \mathbb{R}$ is measurable and nonnegative, then

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\,dx &= \int_{S^{n-1}} \int_0^\infty f(r\theta) r^{n-1}\,dr\,d\theta \\ &= n\,\mathrm{vol}(B_2^n) \int_{S^{n-1}} \int_0^\infty f(r\theta) r^{n-1}\,dr\,d\sigma(\theta)\end{aligned}$$

where in the first line, $d\theta$ indicates Lebesgue measure on S^{n-1}, and in the second, σ is the uniform (i.e., rotationally invariant) probability measure on S^{n-1}. (The factor $n\,\mathrm{vol}(B_2^n) = \mathrm{vol}_{n-1}(S^{n-1})$ is the normalizing factor.)

Lemma 2 For a starshaped body K in $\mathbb{R}^n$,

$$\mathrm{vol}(K) = \mathrm{vol}(B_2^n) \int_{S^{n-1}} \frac{1}{\|\theta\|_K^n}\,d\sigma(\theta)\,.$$

Proof Apply lemma 1 to the characteristic function of K. □

Lemma 2 expresses the volume of K as a kind of average of its norm, which is what we need for part (A).

Lemma 3 For a nonnegative measurable function $f\colon S^{n-1} \to \mathbb{R}$,

$$\int_{S^{n-1}} f(\theta)\, d\sigma(\theta) = \int_{G_{n,k}} \int_{S^{n-1}\cap F} f(\theta)\, d\sigma_F(\theta)\, d\mu(F)\ ,$$

where σ is the uniform probability measure on S^{n-1}, σ_F is the uniform probability measure on $S^{n-1} \cap F$, and μ is the uniform probability measure on $G_{n,k}$, the Grassmannian of k-dimensional subspaces of $\mathbb{R}^n$.

Proof Define $\widehat{\sigma}(A) = \int_{G_{n,k}} \sigma_F(A \cap F)\, d\mu(F)$, and show that $\widehat{\sigma}$ is a rotationally invariant probability measure on S^{n-1}, whence $\widehat{\sigma} = \sigma$. (The integral on the RHS is the integral wrt $\widehat{\sigma}$.) □

Lemma 3 expresses the intuitively obvious equivalence between the notions of "average point on S^{n-1}" and "average point on average section of S^{n-1}", as we need for part (B).

Part (C) needs no lemmas, as to pass from average behaviour to behaviour in most places we need merely invoke Markov's inequality.

Lemma 4 For any $x_0 \in S^{n-1}$ and any $\delta \in [0, \sqrt{2}]$,

$$\sigma(x \in S^{n-1} \colon |x - x_0| \le \delta) \ge \left(\frac{\delta}{\pi}\right)^n\ .$$

(See section 9 for a variant of this lemma.)

In part (D) we wish to deduce the norm is large everywhere on the sphere, knowing only that the norm is large for most points, in the sense that the set of points where the norm is small has small measure. We will deduce that the set of points where the norm is small is small in terms of the metric (which is the role of lemma 4), and then use a Lipschitz condition to deduce that the function cannot be very small anywhere. This argument is encapsulated in the next and final lemma.

Lemma 5 Let $f\colon S^{n-1} \to \mathbb{R}$ be Lipschitz and let $t \in [0, (\sqrt{2}/\pi)^n]$. If $\sigma(f \le r) < t$ then, for every $\theta \in S^{n-1}$, $f(\theta) > r - \pi t^{1/n}$.

Proof Let $\delta = \pi t^{1/n} \in [0, \sqrt{2}]$. The hypothesis and lemma 4 show that $\{f \le r\}$ contains no cap of radius δ. Thus every point $\theta \in S^{n-1}$ is within δ of a point ψ where $f(\psi) > r$, and so $f(\theta) \ge f(\psi) - \delta > r - \pi t^{1/n}$. □

Now we can prove Szarek's theorem. Let K be as described therein. Then Dec 11

$$
\begin{aligned}
A^n &\ge \frac{\mathrm{vol}(K)}{\mathrm{vol}(B_2^n)} && \text{(hypothesis)} \\
&= \int_{S^{n-1}} \frac{1}{\|\theta\|_K^n}\, d\sigma(\theta) && \text{(lemma 2; part (A))} \\
&= \int_{G_{n,k}} \int_{S^{n-1}\cap F} \frac{1}{\|\theta\|_K^n}\, d\sigma_F(\theta)\, d\mu(F) && \text{(lemma 3; part (B))}
\end{aligned}
$$

Let $\Omega_0 = \{F \in G_{n,k} \colon \int_{S^{n-1}\cap F} \frac{1}{\|\theta\|_K^n}\, d\sigma_F(\theta) < (2A)^n\}$. By Markov's inequality,

$$
\mu(\Omega_0) = 1 - \mu(\Omega_0) \ge 1 - \frac{1}{(2A)^n} \int_{G_{n,k}} \int_{S^{n-1}\cap F} \frac{1}{\|\theta\|_K^n}\, d\sigma_F(\theta)\, d\mu(F) \ge 1 - \frac{1}{2^n}\,.
$$

(Ω_0 is the set of "most sections" described in part (C).) Let $F \in \Omega_0$. Then, for $r > 0$ to be chosen later, we have (by Markov's inequality again)

$$
\begin{aligned}
\sigma_F(\theta\colon \|\theta\|_K \le r) &= \sigma_F(\theta\colon \tfrac{1}{\|\theta\|_K^n} \le \tfrac{1}{r^n}) \\
&\le r^n \int_{S^{n-1}\cap F} \frac{1}{\|\theta\|_K^n}\, d\sigma_F(\theta) \\
&\le (2Ar)^n
\end{aligned}
$$

(That completes part (C).) As previously noted, since $B_2^n \subseteq K$ we have $\|\cdot\|_K \le |\cdot|$, and so

$$
\big|\|x\|_K - \|y\|_K\big| \le \|x-y\|_K \vee \|y-x\|_K \le |x-y|\,,
$$

that is, $\|\cdot\|_K$ is Lipschitz. (Note that since K is not assumed symmetric, we cannot use the usual "other" triangle inequality $\big|\|x\| - \|y\|\big| \le \|x-y\|$ here.) So by lemma 5,

$$
\forall\theta \in S^{n-1}\cap F\colon \|\theta\|_K \ge r - \pi(2Ar)^{n/k}\,,
$$

provided that r is later chosen so that

$$
(2Ar)^n \le \left(\frac{\sqrt{2}}{\pi}\right)^k . \tag{8}
$$

(Note that we apply lemma 5 to $S^{n-1}\cap F$, which is essentially S^{k-1}, not to S^{n-1}.) Finally, we choose r so that our estimate $r - \pi(2Ar)^{n/k}$ is positive. The simplest way is to make the second term half of the first, that is,

$$
\pi(2Ar)^{n/k} = \frac{r}{2}\,.
$$

Thus we will take

$$
\left(\frac{r}{2}\right)^{n-k} = \frac{1}{\pi^k(4A)^n}\,.
$$

To check that this value of r satisfies (8), note first that $(2Ar)^n = \left(\frac{r}{2\pi}\right)^k$, so it suffices to check that $\frac{r}{2} \le \sqrt{2}$; and indeed, $\left(\frac{r}{2}\right)^{n-k} = \frac{1}{\pi^k(4A)^n} \le 1$, so in fact $\frac{r}{2} \le 1 \le \sqrt{2}$. Thus we obtain the estimate

$$\|\theta\|_K \ge \frac{r}{2} \ge \frac{1}{(4\pi A)^{n/(n-k)}} = \frac{1}{(4\pi A)^{n/(n-k)}}|\theta| \, .$$

The homogeneity of the norm yields the desired statement, completing the proof of Szarek's theorem.

4 Digression: Convex bodies are spiky

A key maneuver in the proof of Szarek's theorem is to apply Markov's inequality to the formula expressing volume as a kind of average of radius (lemma 2). Writing this idea down separately, we have:

Proposition 6 If K is a star-shaped body in $\mathbb{R}^n$, then

$$\sigma(S^{n-1} \cap rK) \le \frac{\mathrm{vol}(rK)}{\mathrm{vol}(B_2^n)} \, .$$

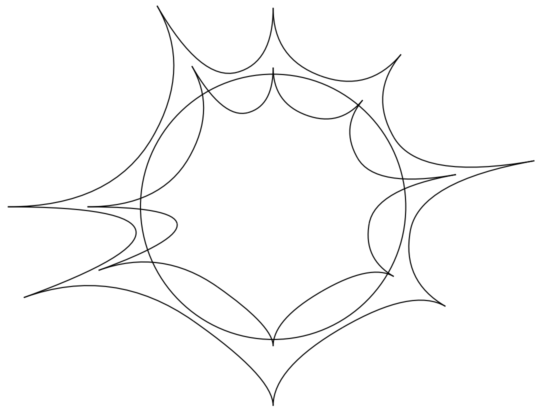
Proof

$$\begin{aligned}\sigma(S^{n-1} \cap rK) &= \sigma(S^{n-1} \cap |r|K) \\ &= \sigma(\theta \in S^{n-1} \colon \|\theta\|_K \le |r|) = \sigma(\theta \in S^{n-1} \colon \tfrac{1}{\|\theta\|_K^n} \ge \tfrac{1}{|r|^n}) \\ &\le |r|^n \int_{S^{n-1}} \frac{1}{\|\theta\|_K^n}\, d\sigma(\theta) = |r|^n \frac{\mathrm{vol}(K)}{\mathrm{vol}(B_2^n)} = \frac{\mathrm{vol}(rK)}{\mathrm{vol}(B_2^n)}\end{aligned}$$

□

Since σ is a probability measure, this upper bound is trivial when $\mathrm{vol}(rK) \ge \mathrm{vol}(B_2^n)$, but when r decreases past this point, this upper bound decreases very rapidly, indeed, like r^n. So if K is even a little bit smaller (in volume) than the Euclidean ball, then it doesn't stick out very much (in terms of area on the sphere); however, as we will see when we compute $\mathrm{vr}(B_1^n)$, a convex body with such volume can stick out quite far (in terms of distance from the origin). This is one reason that high-dimensional convex bodies should be drawn "spiky",[8] even though this makes the picture nonconvex:

[8] Peter pointed out that spiky pictures are good for showing measure phenomena, but, as ever, we should be alert that they do not show metric phenomena well. He also drew my attention to a brief discussion of exactly this kind of picture in [13], which refers to [9] for a similar exponential decay of volume, but using slabs instead of balls.



5 Applications of Szarek's theorem

Corollary 7 Let n be even, and let K be a convex body in $\mathbb{R}^n$ with $B_2^n \subseteq K$ and

$$\left(\frac{\mathrm{vol}(K)}{\mathrm{vol}(B_2^n)}\right)^{1/n} \leq A\,.$$

Then there exists a subspace F of dimension $\frac{n}{2}$ such that

$$\forall x \in F \cup F^\perp\colon \|x\|_K \leq |x| \leq (4\pi A)^2 \|x\|_K\,.$$

Proof Say that a subspace is *good* if the stated inequality holds for vectors in that subspace. By Szarek's theorem, the set of good subspaces F has measure at least $1 - \frac{1}{2^n} > \frac{1}{2}$. And since the map $F \mapsto F^\perp$ preserves the measure on $G_{n,k}$ (indeed, one can check that the measure $\widehat{\mu}(A) = \mu(F^\perp\colon F \in A)$ is rotationally invariant, hence equal to μ), the set of subspaces F such that $F^\perp$ is good also has measure at least $1 - \frac{1}{2^n} > \frac{1}{2}$. Therefore there exists a subspace F such that both F and $F^\perp$ are good. □

For example, $B_2^n \subseteq \sqrt{n} B_1^n$, and we will compute later that Jan 18

$$\left(\frac{\mathrm{vol}(\sqrt{n} B_1^n)}{\mathrm{vol}(B_2^n)}\right)^{1/n} \leq c\,, \tag{9}$$

where c is some constant (meaning in particular that it does not depend on n). Thus, for $n = 2k$, we obtain an orthogonal decomposition $\mathbb{R}^n = F \oplus F^\perp$ with $\dim F = k$ and

$$\forall x \in F \cup F^\perp\colon \|x\|_{\sqrt{n} B_1^n} \leq |x| \leq c' \|x\|_{\sqrt{n} B_1^n}\,,$$

that is,

$$\forall x \in F \cup F^{\perp}: \|x\|_1 \le \sqrt{n}\,|x| \le c'\|x\|_1 .$$

Another nice way to write this is

$$\forall x \in F \cup F^{\perp}: \frac{1}{n}\|x\|_1 \le \frac{1}{\sqrt{n}}|x| \le \frac{c'}{n}\|x\|_1 .$$

This is nice because, writing

$$|||x|||_p = \left(\frac{1}{n}\sum |x_i|^p\right)^{1/p} = \frac{1}{n^{1/p}}\|x\|_p ,$$

it asserts that

$$\forall x \in F \cup F^{\perp}: |||x|||_1 \le |||x|||_2 \le c'|||x|||_1 ,$$

and these norms $|||\cdot|||_p$ are just the usual $L_p(\mu)$ norms if we think of vectors in $\mathbb{R}^n$ as functions $\{1,\dots,n\} \to \mathbb{R}$ and take μ to be the uniform probability measure on $\{1,\dots,n\}$. This motivates the notation L_p^n for these normed spaces (or for their unit balls).

Corollary 8 Let n be even, and let K be a convex body in $\mathbb{R}^n$ with $B_2^n \subseteq K$ and

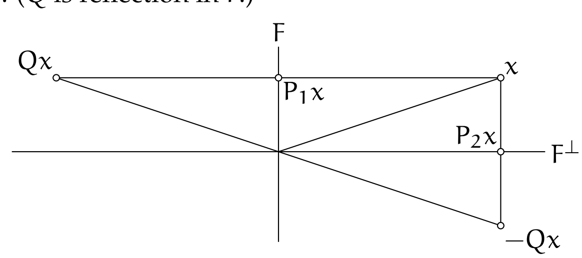
$$\left(\frac{\operatorname{vol}(K)}{\operatorname{vol}(B_2^n)}\right)^{1/n} \le A .$$

Further assume that $K = -K$. Then there exists a symmetric orthogonal map Q such that

$$\forall x \in \mathbb{R}^n: \frac{1}{\sqrt{2}(4\pi A)^2}|x| \le \frac{1}{2}(\|x\|_K + \|Qx\|_K) \le \|x\|_K \vee \|Qx\|_K \le |x| .$$

Proof The second inequality is obvious. The third follows from the facts that $B_2^n \subseteq K$ and Q is (will be) orthogonal.

For the first inequality: let F be as in the previous corollary, let P_1 be the orthogonal projection onto F, let P_2 be the orthogonal projection onto $F^{\perp}$, and let $Q = P_1 - P_2$. (Q is reflection in F.)



We compute, first, that

$$\frac{1}{2}(\|x\|_K + \|Qx\|_K) \geq \frac{1}{2}\|x + Qx\|_K = \|P_1 x\|_K \geq \frac{1}{(4\pi A)^2}|P_1 x| \,.$$

Using the fact that $K = -K$, we have, similarly,

$$\frac{1}{2}(\|x\|_K + \|Qx\|_K) = \frac{1}{2}(\|x\|_K + \|-Qx\|_K) \geq \frac{1}{(4\pi A)^2}|P_2 x| \,.$$

Thus

$$\frac{1}{2}(\|x\|_K + \|Qx\|_K) \geq \frac{1}{(4\pi A)^2}(|P_2 x| \vee |P_1 x|) \geq \frac{1}{\sqrt{2}(4\pi A)^2}|x| \,,$$

as desired. □

In terms of unit balls, this corollary asserts that

$$B_2^n \subseteq K \cap QK \subseteq 2(K \boxplus QK) \subseteq \sqrt{2}(4\pi A)^2 B_2^n \,, \tag{10}$$

where $K \boxplus L$ denotes the body whose norm is the sum of the norms of K and L. (It can be shown that this type of addition is dual to Minkowski addition, in the sense that $(K + L)^\circ = K^\circ \boxplus L^\circ$, and that we have the inclusions

$$K \boxplus L \subseteq K \cap L \subseteq 2(K \boxplus L) \subseteq \frac{1}{2}(K + L) \subseteq \operatorname{conv}(K \cup L) \subseteq K + L \,.$$

The third inclusion is tricky, but the others are straightforward. The last one assumes $0 \in K \cap L$.)

In lecture 4 of [3], Ball proves a version of (10) directly. The method resembles the one used here, but the direct approach yields several advantages: the result is proved for all n (not just even n); we obtain a better constant; the choice of Q is random and with high probability in all of $O(n)$ (instead of just reflections in $\frac{n}{2}$-dimensional subspaces); and we need not assume that K is symmetric. He also deduces Szarek's theorem (for even n) from this statement.

The point of these maneuvers is to show that the phenomena described by Szarek's theorem (a section of the body is somewhat Euclidean) and (10) (combinations of the body and a copy of are somewhat Euclidean) are not merely parallel but, to some extent, mutually deducible. For more on this idea, but focussing on Dvoretzky's theorem, see [14].

The part of (10) that concerns $K \cap QK$ is intuitively reasonable from our previous discussion of "spikiness": since the spikes of K occupy a small area on the sphere, K and QK are not likely to have spikes in the same directions, and so intersecting them will cut off the spikes, leaving something close to B_2^n. Jan 25

6 Volume ratio of B_p^n

As promised earlier, we will now compute the volume of B_p^n (following the method used in [15], page 11), and thereby show how the results of the previous section apply to such balls.

Lemma 9 If $K \subseteq \mathbb{R}^n$ is star-shaped and $0 < p < \infty$, then

$$\operatorname{vol}(K) = \frac{1}{\Gamma(1+\frac{n}{p})} \int_{\mathbb{R}^n} e^{-\|x\|_K^p}\, dx\,.$$

Proof

$$\begin{aligned}
\int_{\mathbb{R}^n} e^{-\|x\|_K^p}\, dx &= \int_{\mathbb{R}^n} \int_{\|x\|_K^p}^{\infty} e^{-t}\, dt\, dx \\
&= \int\int [\|x\|_K^p \le t]\, e^{-t}\, dt\, dx \\
&= \int\int [\|x\|_K \le t^{1/p}]\, e^{-t}\, dt\, dx && \text{(using } p > 0\text{)} \\
&= \int\int [t \ge 0]\, [x \in t^{1/p}K]\, e^{-t}\, dt\, dx && \text{(using } K \text{ star-shaped)} \\
&= \int_0^{\infty} \operatorname{vol}(t^{1/p}K) e^{-t}\, dt \\
&= \int_0^{\infty} t^{n/p} e^{-t}\, dt \operatorname{vol}(K) \\
&= \Gamma(1+\tfrac{n}{p}) \operatorname{vol}(K)
\end{aligned}$$

□

(Lemma 9 appears, with a slightly different proof, as Lemma 7 in [2], but I doubt this is its origin.[9])

Proposition 10 If $0 < p < \infty$ then

$$\operatorname{vol}(B_p^n) = \frac{2^n \Gamma(1+\frac{1}{p})^n}{\Gamma(1+\frac{n}{p})}\,.$$

Proof First we compute that

$$\begin{aligned}
\int_{\mathbb{R}^n} e^{-\|x\|_p^p}\, dx = \int_{\mathbb{R}^n} e^{-\sum |x_i|^p}\, dx &= \int_{\mathbb{R}^n} \prod e^{-|x_i|^p}\, dx \\
&= \left(\int_{\mathbb{R}} e^{-|t|^p}\, dt \right)^n = 2^n \left(\int_0^{\infty} e^{-t^p}\, dt \right)^n,
\end{aligned}$$

[9] Sasha and Nicole suspect this fact has been known for at least a hundred years.

and then the substitution $u = t^p$ yields

$$= 2^n \left(\frac{1}{p} \int_0^\infty u^{\frac{1}{p}-1} e^{-u} \, du \right)^n = 2^n \left(\frac{1}{p} \Gamma\left(\frac{1}{p}\right) \right)^n = 2^n \Gamma\left(1 + \frac{1}{p}\right)^n ,$$

and the lemma yields the claim. □

This method relies on the rather clever lemma 9; it may be reassuring to know that the result for B_p^n can also be obtained by the more obvious method of slicing. Indeed, the cross-section of B_p^n obtained by fixing one coordinate is a scaled copy of B_p^{n-1}; thus

$$\mathrm{vol}(B_p^n) = \int_{-1}^{1} \mathrm{vol}(\text{scaling factor} \cdot B_p^{n-1}) \, dt ,$$

which with a little computation leads to

$$\frac{\mathrm{vol}(B_p^n)}{\mathrm{vol}(B_p^{n-1})} = \frac{2}{p} B\left(\frac{1}{p}, 1 + \frac{n-1}{p}\right) = \cdots = \frac{2\Gamma(1+\frac{1}{p})\Gamma(1+\frac{n-1}{p})}{\Gamma(1+\frac{n}{p})} ,$$

where $B(s,t)$ is the beta integral. Induction on n then yields the desired formula.

Examples

1. $\mathrm{vol}(B_1^n) = \frac{2^n}{n!}$, which can also be obtained by noting that B_1^n consists of 2^n copies (one in each orthant) of a "right-angled simplex", which is a cone whose height is 1 and whose base is the analogous $(n-1)$-dimensional body.

2. $\mathrm{vol}(B_2^n) = \frac{\Gamma(\frac{1}{2})^n}{\Gamma(1+\frac{n}{2})}$. Since $\mathrm{vol}(B_2^2) = \pi$, this yields $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. (Indeed, the method used here to compute $\mathrm{vol}(B_p^n)$ is a generalization of the classical method of computing this value.) Thus

$$\mathrm{vol}(B_2^n) = \frac{\pi^{n/2}}{\Gamma(1+\frac{n}{2})} .$$

3. The case $p = \infty$ is not covered by proposition 10, but if we fix n and let $p \nearrow \infty$, then by the continuity of measure,

$$\mathrm{vol}\left(\bigcup_{p>0} B_p^n \right) = \lim_{n\to\infty} \mathrm{vol}(B_p^n) = 2^n .$$

   Since

$$\mathrm{int}\, B_\infty^n \subseteq \bigcup_{p>0} B_p^n \subseteq B_\infty^n ,$$

   we thus have $\mathrm{vol}(B_\infty^n) = 2^n$. (Of course we already knew this.)[10]

[10] Thanks to Niushan for pointing out that my original argument in this example was wrong.

4. If we fix p and let $n \to \infty$, then by Stirling's approximation,

$$\mathrm{vol}(B_p^n) \sim \frac{2^n \Gamma(1+\frac{1}{p})^n}{\sqrt{\frac{2\pi n}{p}}\left(\frac{n}{pe}\right)^{n/p}},$$

whence

$$\mathrm{vol}(B_p^n)^{1/n} \sim \frac{2\Gamma(1+\frac{1}{p})(pe)^{1/p}}{n^{1/p}} = \frac{c_p}{n^{1/p}}.$$

Thus a natural normalization is $n^{1/p}B_p^n$, which is the unit ball of the L_p^n norm mentioned on page 13.

Now, recall the standard fact that for a probability measure μ, if $0 < p \le q \le \infty$ then $\|\cdot\|_{L_p(\mu)} \le \|\cdot\|_{L_q(\mu)}$). Applying this to the L_p^n norms yields

$$n^{1/q}B_q^n \subseteq n^{1/p}B_p^n. \tag{11}$$

(In fact we only care about $p \ge 1$, because we want convexity.) So if $1 \le p \le 2$, then $n^{1/2}B_2^n \subseteq n^{1/p}B_p^n$, with contact at the vertices of B_∞^n. Those vertices (after scaling) support a decomposition of the identity as in John's theorem, so we conclude that $n^{1/2}B_2^n$ is the maximum volume ellipsoid in $n^{1/p}B_p^n$. Thus

$$\mathrm{vr}(B_p^n) = \left(\frac{\mathrm{vol}(n^{1/p}B_p^n)}{\mathrm{vol}(n^{1/2}B_2^n)}\right)^{1/n} \sim \frac{\Gamma(1+\frac{1}{p})(pe)^{1/p}}{\Gamma(1+\frac{1}{2})(2e)^{1/2}} \qquad \text{if } 1 \le p \le 2,$$

which establishes the promised result for B_1^n (namely (9), on page 12). Again, note that the constant on the right does not depend on the dimension. On the other hand, if $2 \le p \le \infty$, then B_2^n is the maximum volume ellipsoid of B_p^n, and so

$$\mathrm{vr}(B_p^n) = \left(\frac{\mathrm{vol}(B_p^n)}{\mathrm{vol}(B_2^n)}\right)^{1/n} \sim \frac{\Gamma(1+\frac{1}{p})(pe)^{1/p}}{\Gamma(1+\frac{1}{2})(2e)^{1/2}} \cdot \frac{n^{1/2}}{n^{1/p}} \qquad \text{if } 2 \le p \le \infty,$$

which is, alas, an unbounded function of n. Still, if $p < \infty$, this improves the $\sqrt{n}$ estimate we obtained from John's theorem (see (2) on 3).[11]

7 Summary of parameters

K	$\mathrm{vol}(K_{\mathrm{John}})$	$\mathrm{vol}(K_{\mathrm{John}})^{1/n}$	$\mathrm{vr}(K)$	$d(K, B_2^n)$
B_1^n	$2^n n^{n/2}/n!$	$\sim c/\sqrt{n}$	$\sim c$	$\sqrt{n}$
B_2^n	$\pi^{n/2}/\Gamma(1+\frac{n}{2})$	$\sim c/\sqrt{n}$	1	1
B_∞^n	2^n	2	$\sim c/\sqrt{n}$	$\sqrt{n}$
$\triangle^n$	$n^{n/2}(n+1)^{(n+1)/2}/n!$	$\sim e$	$\sim c/\sqrt{n}$	n

[11] As Nicole pointed out, this also gives us an asymptotically sharp lower bound on $d(B_p^n, B_2^n)$. Indeed, for the range $2 \le p \le \infty$, we have $d(B_p^n, B_2^n) \ge \mathrm{vr}(B_p^n) \sim c_p n^{1/2-1/p}$, and $B_2^n \subseteq B_p^n \subseteq n^{1/2-1/p}B_2^n$. For the range $1 \le p \le 2$, take duals.

In this table we can see that, although distance to the Euclidean ball and volume ratio are related (as in (2)), they may nevertheless be quite different: B_1^n and B_∞^n both have extremal distance (for symmetric bodies) but their volume ratios are at the opposite ends of the scale.

The table also shows again that asymmetric bodies can be much further from the ball, but asymmetry does not disturb volume ratio much. (Recall footnote 2, on page 3.)

8 Loose end: Control by ϵ-nets

In the proof of Szarek's theorem, we used the following obvious fact about how controlling a (nice) function on an ϵ-net lets us control it on the whole sphere:

Observation Let Λ be an ϵ-net for S^{n-1} and let $f\colon S^{n-1} \to \mathbb{R}$ be c-Lipschitz. If $\alpha \in \mathbb{R}$ is such that

$$\forall\theta \in \Lambda\colon \alpha \leq f(\theta)\ ,$$

then

$$\forall\theta \in S^{n-1}\colon \alpha - c\epsilon \leq f(\theta)\ .$$

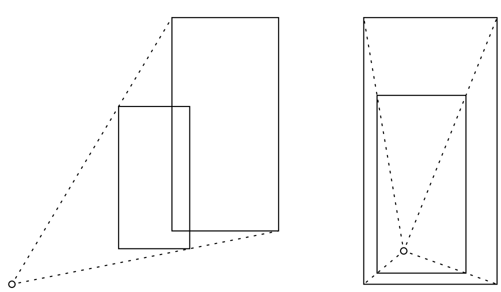
(This observation, with $c = 1$, is implicit in the proof of lemma 5.)

When f is a norm (even an asymmetric one), there is a standard, more powerful, proposition along the same lines. We need the following simple geometric proposition as a lemma.[12] Feb 1

Proposition 11 Let $K \subseteq \mathbb{R}^n$ be closed and convex. Let $A \subseteq \mathbb{R}^n$ be bounded. Let $0 \leq \lambda < 1$. If $A \subseteq (1-\lambda)K + \lambda A$, then $A \subseteq K$.

(We would perhaps like to simply subtract λA from both sides (but this is not possible with Minkowski addition), then factor out the common A on the left (but this is not possible when A is not convex) and then cancel the $(1-\lambda)$.)

For intuition, consider the case that A is a singleton. Then $(1-\lambda)K + \lambda A$ is a smaller homothet of K, in centre A:



[12] Actually, as Sasha pointed out, we can prove proposition 13 more directly and simply (but less geometrically). Such a proof is given in [3], Lemma 9.2.

The proposition states that if A is in the smaller copy then we must be in the situation at right.

Proof We prove the contrapositive. Accordingly, suppose $A \not\subseteq K$. Let $x \in A \setminus K$, and separate x from K by a functional f, so that $f(x) > \sup_K f$. then

$$\sup_{(1-\lambda)K+\lambda A} f = (1-\lambda)\sup_K f + \lambda \sup_A f < \sup_A f$$

with strict inequality because $\lambda < 1$ and $\sup_K f < f(x) \leq \sup_A f < \infty$. Therefore $A \not\subseteq (1-\lambda)K + \lambda A$. □

Corollary 12 If Λ is a ϵ-net for S^{n-1}, where $0 \leq \epsilon < 1$, then

$$\operatorname{cl}\operatorname{conv}\Lambda \subseteq B_2^n \subseteq \frac{1}{1-\epsilon}\operatorname{cl}\operatorname{conv}\Lambda\,.$$

Proof The hypothesis asserts that

$$\Lambda \subseteq S^{n-1} \subseteq \Lambda + \epsilon B_2^n\,.$$

Taking closed convex hulls yields (using the facts that $\operatorname{conv}(A+B) = \operatorname{conv} A + \operatorname{conv} B$ and, for bounded sets, $\operatorname{cl}(A+B) = \operatorname{cl} A + \operatorname{cl} B$)

$$\begin{aligned} \operatorname{cl}\operatorname{conv}\Lambda \subseteq B_2^n &\subseteq \operatorname{cl}\operatorname{conv}\Lambda + \epsilon B_2^n \\ &= (1-\epsilon)\frac{1}{1-\epsilon}\operatorname{cl}\operatorname{conv}\Lambda + \epsilon B_2^n \end{aligned}$$

Applying proposition 11 with $K, A, \lambda := \frac{1}{1-\epsilon}\operatorname{cl}\operatorname{conv}\Lambda, B_2^n, \epsilon$ yields the desired result. □

Proposition 13 Let Λ be an ϵ-net for S^{n-1}, where $0 \leq \epsilon < 1$. Let $K \subseteq \mathbb{R}^n$ be a convex body with $0 \in \operatorname{int} K$. If $\alpha, \beta \in \mathbb{R}$ are such that

$$\forall\theta \in \Lambda\colon \alpha \leq \|\theta\|_K \leq \beta\,,$$

then

$$\forall\theta \in S^{n-1}\colon \alpha - \frac{\epsilon}{1-\epsilon}\beta \leq \|\theta\|_K \leq \frac{1}{1-\epsilon}\beta\,.$$

Proof Since $\|\theta\|_K \leq \beta$ for $\theta \in \Lambda$, we have $\Lambda \subseteq \beta K$, and so $\operatorname{cl}\operatorname{conv}\Lambda \subseteq \beta K$. By corollary 12, $B_2^n \subseteq \frac{1}{1-\epsilon}\beta K$, which yields the desired upper estimate. It also yields that $\|\cdot\|_K$ is $\frac{1}{1-\epsilon}\beta$-Lipschitz, which yields the lower estimate. □

A variant of this proposition is that if

$$\forall\theta \in \Lambda\colon 1-\delta \leq \|\theta\|_K \leq 1+\delta\,,$$

then

$$\forall\theta \in S^{n-1}\colon \frac{1-\delta-2\epsilon}{1-\epsilon} \leq \|\theta\|_K \leq \frac{1+\delta}{1-\epsilon}\,.$$

(This version is proven directly as Lemma 9.2 in [3].)

9 Loose end: Construction of ϵ-nets

In the proof of Szarek's theorem, we used a lower bound on the (relative) measure of a spherical cap, namely lemma 4, which asserted that

$$0 \leq \epsilon \leq \sqrt{2} \implies \sigma(\epsilon\text{-cap}) \geq \left(\frac{\epsilon}{\pi}\right)^n , \tag{12}$$

where σ is, as usual, the uniform probability measure on S^{n-1}. As part of the argument for lemma 5, we deduced that any set smaller than this doesn't contain an ϵ-cap, and so its complement is an ϵ-net. In other words, any big enough subset of the sphere (specifically, any set with measure exceeding $1 - \left(\frac{\epsilon}{\pi}\right)^n$) is an ϵ-net, for $\epsilon \in [0, \sqrt{2}]$.

This construction of an ϵ-net is good when you don't have much control over the set in question (as we didn't, in Szarek's theorem), but if you can choose the set yourself, there is a standard way to construct much smaller ϵ-nets (indeed, finite ones).

First, a sketch of the main ideas: (a) a maximal ϵ-separated set is an ϵ-net, and since (b) any ϵ-separated set yields a packing of $\frac{\epsilon}{2}$-caps in S^{n-1}, we can (c) bound the number of points of an ϵ-separated set by a volumetric argument. On the other hand, (d) *any* ϵ-net yields a cover of S^{n-1} by ϵ-caps, which (e) by a similar volumetric argument yields a lower bound for the size of an ϵ-net. The resulting bounds are:

$$\frac{1}{\sigma(\epsilon\text{-cap})} \leq \text{minimum number of points in } \epsilon\text{-net} \leq \frac{1}{\sigma(\frac{\epsilon}{2}\text{-cap})} .$$

(Since an ϵ-cap is, for small ϵ, essentially the same as ϵB_2^{n-1}, we expect these bounds to be, respectively, something like $\left(\frac{1}{\epsilon}\right)^{n-1}$ and $\left(\frac{2}{\epsilon}\right)^{n-1}$.)

Now, the details behind the sketch above:

(a) Let Λ be a maximal ϵ-separated subset of S^{n-1}. Maximality means that adding any other point of S^{n-1} destroys the ϵ-separation. In other words, every other point of S^{n-1} is within ϵ of some point in Λ.

(b) For Λ to be ϵ-separated means

$$\forall x, y \in \Lambda\colon x \neq y \implies |x - y| \geq \epsilon .$$

Now,

$$\begin{aligned} |x-y| \geq \epsilon &\iff x - y \notin \operatorname{int} \epsilon B_2^n \\ &\iff x - y \notin \operatorname{int} \tfrac{\epsilon}{2} B_2^n - \operatorname{int} \tfrac{\epsilon}{2} B_2^n \qquad (*) \\ &\iff (x + \operatorname{int} \tfrac{\epsilon}{2} B_2^n) \cap (y + \operatorname{int} \tfrac{\epsilon}{2} B_2^n) = \varnothing \end{aligned}$$

Thus Λ is ϵ-separated if and only if its points are the centres of a packing of $\frac{\epsilon}{2}$-caps. (Note that step $(*)$ uses the convexity and symmetry of B_2^n.)

(c) Given a disjoint (or almost disjoint) collection of $\frac{\epsilon}{2}$-caps in S^{n-1}, we can compute

$$1 = \sigma(S^{n-1}) \geq \sigma(\text{union of the } \tfrac{\epsilon}{2}\text{-caps}) = (\text{number of caps})\sigma(\tfrac{\epsilon}{2}\text{-cap}) .$$

(d) For Λ to be an ϵ-net for S^{n-1} means that every point of S^{n-1} is in some ϵ-cap centred at a point of Λ, which means that these caps cover S^{n-1}.

(e) Given a collection of ϵ-caps that cover S^{n-1}, we have

$$1 = \sigma(S^{n-1}) = \sigma(\text{union of the } \epsilon\text{-caps}) \leq (\text{number of caps})\sigma(\epsilon\text{-cap}) .$$

Now, computing $\sigma(\epsilon\text{-cap})$ is somewhat annoying. We can simplify our computations considerably by noting that, by the argument in part (b) above, any ϵ-separated set of points in S^{n-1} yields a collection of almost disjoint copies of $\frac{\epsilon}{2}B_2^n$ with centres on S^{n-1}. Such balls are subsets of $(1+\frac{\epsilon}{2})B_2^n$, so by the same volumetric argument as in part (c), any ϵ-separated set has at most

$$\frac{\operatorname{vol}((1+\frac{\epsilon}{2})B_2^n)}{\operatorname{vol}(\frac{\epsilon}{2}B_2^n)} = \frac{(1+\frac{\epsilon}{2})^n}{(\frac{\epsilon}{2})^n} = \left(1+\frac{2}{\epsilon}\right)^n$$

points. If $\epsilon \in [0,1]$, this is at most $(\frac{3}{\epsilon})^n$, which is for many purposes close enough to optimal.

Note that with such a net Λ we obtain

$$\frac{1}{\sigma(\epsilon\text{-cap})} \leq \operatorname{card}(\Lambda) \leq \left(\frac{3}{\epsilon}\right)^n$$

(where card denotes cardinality) and so

$$0 \leq \epsilon \leq 1 \implies \sigma(\epsilon\text{-cap}) \geq \left(\frac{\epsilon}{3}\right)^n ,$$

which is very close to (12), that is, lemma 4. It is in fact possible to prove Szarek's theorem using this estimate, and we get a slightly better constant.

10 Comparison with Milman's theorem

Let K be a convex body in $\mathbb{R}^n$ with $B_2^n \subseteq K$. Milman's theorem asserts that Feb 8

> For any $\epsilon > 0$, most subspaces F of dimension $\leq c(\epsilon)n(M(K))^2$ satisfy $d(K \cap F, B_2^n \cap F) \leq 1+\epsilon$.

The M parameter is the average of the norm on the sphere:

$$M(K) = \int_{S^{n-1}} \|\theta\|_K \, d\sigma(\theta) ,$$

where σ is, as usual, the uniform probability measure on S^{n-1}. This parameter is not affine invariant, or even linear invariant, although we do have $M(QK) = M(K)$ for orthogonal Q, and $M(bK) = \frac{1}{b}M(K)$ for scalar b.

In analogous language, Szarek's theorem asserts that

> For any k with $1 \le k < n$, most subspaces F of dimension k satisfy
> $$d(K \cap F, B_2^n \cap F) \le \left(4\pi\left(\frac{\mathrm{vol}(K)}{\mathrm{vol}(B_2^n)}\right)^{1/n}\right)^{n/(n-k)} .$$

Comparing these two theorems shows a trade-off between distance and dimension: if you want the section $K \cap F$ to be very close to Euclidean, then you can use Milman's theorem, and the dimension of F might not be large; on the other hand, if you want a high-dimensional section, then you can use Szarek's theorem, and the distance to the Euclidean ball might not be small.

The parameters involved in the two theorems are related:

$$\frac{1}{M(K)} \le \left(\frac{\mathrm{vol}(K)}{\mathrm{vol}(B_2^n)}\right)^{1/n} \le M(K^\circ) . \tag{13}$$

The upper inequality is called Urysohn's inequality; a proof appears in [15], page 6.[13] The lower inequality is easy:

$$\begin{aligned}
\frac{\mathrm{vol}(K)}{\mathrm{vol}(B_2^n)} &= \int_{S^{n-1}} \frac{1}{\|\theta\|_K^n}\, d\sigma(\theta) && \text{(by lemma 2; see page 8)} \\
&= \mathbb{E}(X^{-n}) && \text{(with } X\colon S^{n-1} \to \mathbb{R},\ \theta \mapsto \|\theta\|_K\text{)} \\
&\ge (\mathbb{E}X)^{-n} && \text{(Jensen's inequality: } t \mapsto t^{-n} \text{ is convex)} \\
&= \frac{1}{M(K)^n}
\end{aligned}$$

The lower inequality in (13) already yields good information about nearly-Euclidean sections of B_1^n: recall from section 6 that

$$\mathrm{vr}(B_1^n) = \left(\frac{\mathrm{vol}(\sqrt{n}B_1^n)}{\mathrm{vol}(B_2^n)}\right)^{1/n} \sim c .$$

Therefore

$$\frac{1}{c} \le M(\sqrt{n}B_1^n) \le 1 ,$$

so $M(\sqrt{n}B_1^n)$ is bounded (independently of n). Milman's theorem then yields that $\sqrt{n}B_1^n$ (and hence B_1^n) has $(1+\epsilon)$-Euclidean sections of dimension $c(\epsilon)n$.

[13] As Peter pointed out during the seminar, it also follows immediately by combining the lower inequality with the Blaschke–Santaló inequality.

This approach fails for B^n_∞ because $\mathrm{vr}(B^n_\infty) \sim c\sqrt{n}$, so we only obtain

$$M(B^n_\infty) \geq \frac{c}{\sqrt{n}} ,$$

which yields only that B^n_∞ has $(1+\epsilon)$-Euclidean sections of dimension $c(\epsilon)$. But this is trivial, since if we are content with sections of constant dimension, we can just take one-dimensional sections, which are line segments and therefore exactly Euclidean.[14]

Now, it turns out that, for symmetric K,

$$M(K_{\text{John}}) \geq M(B^n_\infty) \geq c\sqrt{\frac{\log n}{n}} . \tag{14}$$

This yields that every symmetric convex body has $(1+\epsilon)$-Euclidean sections of dimension $c(\epsilon)\log n$, which is (Milman's improvement of) Dvoretzky's theorem.

Traditionally one deduces Dvoretzky's theorem from Milman's theorem not by proving exactly that $M(K_{\text{John}}) \geq M(B^n_\infty)$, but instead by first proving the Dvoretzky–Rogers lemma, which shows that we can trade half the dimensions of our section for some resemblance to the cube (in the new, lower-dimensional section). This approach resembles (14) in spirit.

The last topic of this seminar is to prove the first inequality in (14) by exploiting much of what we have proved already. (The possibility of using this method is the concluding remark of [3].[15])

Proposition 14 If $f\colon \mathbb{R}^n \to \mathbb{R}$ is measurable and positively homogeneous (that is, $f(\lambda x) = \lambda f(x)$ when $\lambda \geq 0$), then

$$\int_{\mathbb{R}^n} f(x)\,d\gamma_n(x) = \frac{\sqrt{2}\,\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \int_{S^{n-1}} f(\theta)\,d\sigma(\theta) ,$$

where γ_n is the standard gaussian probability measure on $\mathbb{R}^n$, which has density $e^{-|x|^2/2}/(2\pi)^{n/2}$, and σ is the uniform probability measure on S^{n-1}.

Proof In brief,

$$\begin{aligned}
\int_{\mathbb{R}^n} f(x)\,d\gamma_n(x) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(\tfrac{x}{|x|})|x|e^{-|x|^2/2}\,dx \\
&= \frac{n\,\mathrm{vol}(B^n_2)}{(2\pi)^{n/2}} \int_{S^{n-1}} \int_0^\infty f(\theta) r^n e^{-r^2/2}\,dr\,d\sigma(\theta) \qquad \text{(by lemma 1)} \\
&= \frac{n\,\mathrm{vol}(B^n_2)}{(2\pi)^{n/2}} \cdot 2^{(n-1)/2}\Gamma(\tfrac{n+1}{2}) \int_{S^{n-1}} f(\theta)\,d\sigma(\theta)
\end{aligned}$$

[14] As Sasha pointed out in seminar, the $c(\epsilon)$ that arises this way is probably less than 1 anyway, so the result is especially useless.

[15] Update: This proof was given explicitly in [16], Proposition 4.11.

Substituting the value of $\mathrm{vol}(B_2^n)$ (see page 16) and simplifying yields the desired result. □

(Incidentally, note that we can compute the normalizing constant $(2\pi)^{n/2}$ for the gaussian measure by invoking lemma 9; see page 15.)

In particular, proposition 14 and Stirling's approximation yield

$$M(K) = \frac{\Gamma(\frac{n}{2})}{\sqrt{2}\,\Gamma(\frac{n+1}{2})} \int_{\mathbb{R}^n} \|x\|_K \, d\gamma_n(x) \sim \frac{1}{\sqrt{n}} \int_{\mathbb{R}^n} \|x\|_K \, d\gamma_n(x) \, .$$

For example, if X is a standard gaussian variable in $\mathbb{R}^n$, then

$$\mathbb{E}|X| = \int_{\mathbb{R}^n} |x| \, d\gamma_n(x) \sim \sqrt{n} M(B_2^n) = \sqrt{n} \, .$$

Proposition 15 If K is a symmetric convex body in $\mathbb{R}^n$, then for all $r \in \mathbb{R}$, $\gamma_n(rK_{\text{John}}) \le \gamma_n(rB_\infty^n)$.

(This proposition resembles the symmetric case of Ball's theorem (see page 3), but for gaussian measure; we use a similar technique to prove it.)

Proof Let $(u_i)_1^m$ and $(c_i)_1^m$ be as in John's theorem. We may assume $r \ge 0$. Then

$$\begin{aligned}
\gamma_n(rK) &\le \gamma_n(\{x \in \mathbb{R}^n : (\forall i : |\langle x, u_i\rangle| \le r)\}) \\
&= \int_{\mathbb{R}^n} [\forall i : |\langle x, u_i\rangle| \le r] \, \frac{e^{-|x|^2/2}}{(2\pi)^{n/2}} \, dx \\
&= \int_{\mathbb{R}^n} [\forall i : |\langle x, u_i\rangle| \le r] \, \frac{e^{-\sum c_i \langle x, u_i\rangle^2/2}}{(2\pi)^{\sum c_i/2}} \, dx \\
&= \int_{\mathbb{R}^n} \prod_{i=1}^m \left([|\langle x, u_i\rangle| \le r] \frac{e^{-\langle x, u_i\rangle^2/2}}{(2\pi)^{1/2}} \right)^{c_i} dx \\
&\le \prod_{i=1}^m \left(\int_{\mathbb{R}} [|t| \le r] \frac{e^{-t^2/2}}{(2\pi)^{1/2}} \, dt \right)^{c_i} \qquad \text{(Brascamp–Lieb; see page 5)} \\
&= \gamma_1([-r, r])^n \\
&= \gamma_n(rB_\infty^n)
\end{aligned}$$

□

Corollary 16 If K is a symmetric convex body in $\mathbb{R}^n$, then $M(K_{\text{John}}) \ge M(B_\infty^n)$.

Proof Indeed, for any symmetric convex body L,

$$\begin{aligned} M(L) &= \frac{\Gamma(\frac{n}{2})}{\sqrt{2}\,\Gamma(\frac{n+1}{2})} \int_{\mathbb{R}^n} \|x\|_L \, d\gamma_n(x) && \text{(by proposition 14)} \\ &= \frac{\Gamma(\frac{n}{2})}{\sqrt{2}\,\Gamma(\frac{n+1}{2})} \int_0^\infty \gamma_n(x\colon \|x\|_L \ge t)\, dt && \text{(distribution formula)} \\ &= \frac{\Gamma(\frac{n}{2})}{\sqrt{2}\,\Gamma(\frac{n+1}{2})} \int_0^\infty (1-\gamma_n(tL))\, dt \end{aligned}$$

Since $\gamma_n(tL)$ is maximal for $L = B^n_\infty$ by proposition 15, it follows that $M(L)$ is minimal for $L = B^n_\infty$. □

References

[1] Keith Ball. Volumes of sections of cubes and related problems. In Joram Lindenstrauss and Vitali Milman, editors, *Israel Seminar on Geometrical Aspects of Functional Analysis (1987/88)*, volume 1376 of *Lecture Notes in Mathematics*, pages 251–260. Springer, 1990. (Cited on pages 3 and 5.)

[2] Keith Ball. Volume ratios and a reverse isoperimetric inequality. *J. London Math. Soc.*, 44:351–359, 1991. doi:10.1112/jlms/s2-44.2.351. (Cited on pages 3 and 15.)

[3] Keith Ball. An elementary introduction to modern convex geometry. In Silvio Levy, editor, *Flavors of Geometry*, volume 31 of *Mathematical Sciences Research Institute Publications*, pages 1–58. Cambridge UP, Cambridge, 1997. http://www.msri.org/publications/books/Book31/files/ball.pdf. (Cited on pages 3, 14, 18, 19, and 23.)

[4] Keith Ball. Convex geometry and functional analysis. In Johnson and Lindenstrauss [11], pages 161–194. (Cited on pages 3 and 5.)

[5] Franck Barthe. Inégalités de Brascamp–Lieb et convexité. *C. R. Acad. Sci. I*, 324:885–888, 1997. In French. (Cited on page 5.)

[6] Franck Barthe. On a reverse form of the Brascamp–Lieb inequality. *Invent. Math.*, 134:335–361, 1998. (Cited on page 5.)

[7] Herm Jan Brascamp and Elliot H. Lieb. Best constants in Young's inequality, its converse, and its generalization to more than three functions. *Adv. Math.*, 20:151–173, 1976. (Cited on page 5.)

[8] Apostolos Giannopoulos and Vitali Milman. Euclidean structure in finite dimensional normed spaces. In Johnson and Lindenstrauss [11], pages 707–779. (Cited on page 3.)

[9] Mikhail Gromov and Vitali Milman. Brunn theorem and a concentration of volume phenomena for symmetric convex bodies. In *Israel Seminar on Geometrical Aspects of Functional Analysis (1983/84)*, pages V.1–V.12. Tel Aviv UP, Tel Aviv, 1984. (Cited on page 11.)

[10] Fritz John. Extremum problems with inequalities as subsidiary conditions. In K. O. Friedrichs, O. E. Neugebauer, and J. J. Stoker, editors, *Studies and Essays Presented to R. Courant on his 60th birthday, January 8, 1948*, pages 187–204. Interscience, New York, 1948. (Cited on page 3.)

[11] William B. Johnson and Joram Lindenstrauss, editors. *Handbook of the Geometry of Banach Spaces*, volume 1. North-Holland, Amsterdam, 2001. (Cited on page 25.)

[12] B. S. Kašin. Diameters of some finite-dimensional sets and classes of smooth functions. *Math. USSR Izv.*, 11(11):317–333, 1977. doi:10.1070/IM1977v011n02ABEH001719. (Cited on page 7.)

[13] Vitali Milman. Surprising geometric phenomena in high dimensional convexity theory. In *Proceedings of ECM2*, volume 2, pages 73–91, Budapest, 1996. Birkhäuser. http://www.math.tau.ac.il/~milman/files/survey4.ps. (Cited on page 11.)

[14] Vitali Milman and Gideon Schechtman. Global versus local asymptotic theories of finite-dimensional normed spaces. *Duke Math. J.*, 90(1):73–93, 1997. (Cited on page 14.)

[15] Gilles Pisier. *The Volume of Convex Bodies and Banach Space Geometry*. Number 94 in Cambridge Tracts in Mathematics. Cambridge UP, 1989. (Cited on pages 7, 15, and 22.)

[16] Gideon Schechtman and Michael Schmuckenschläger. A concentration inequality for harmonic measures on the sphere. In Joram Lindenstrauss and Vitali Milman, editors, *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 1992–94*, number 77 in Operator Theory: Advances and Applications, pages 255–273. Birkhäuser Verlag, 1995. (Cited on page 23.)

[17] S. Stein. The symmetry function in a convex body. *Pac. J. Math.*, 6:145–148, 1956. (Cited on page 3.)

[18] Stanisław J. Szarek. On Kashin's almost Euclidean orthogonal decomposition of ℓ_1^n. *Bull. Acad. Polon. Sci. Sér. Math.*, 26:691–694, 1978. (Cited on page 7.)

[19] Stanisław J. Szarek and Nicole Tomczak-Jaegermann. On nearly Euclidean decompositions of some classes of Banach spaces. *Compos. Math.*, 40(3):367–385, 1980. (Cited on page 7.)