Upper Bounds on the Cross-Sectional Volumes of Cubes and Other Problems

Ben Pooley

March 2012

Contents

1 Preliminaries

1.1 Introduction

In this essay we shall discuss some problems in convex geometry related to (hyper)cubes in $\mathbb{R}^n$. Throughout, we shall denote the centred unit cube in $\mathbb{R}^n$ by $Q_n := [-\frac{1}{2}, \frac{1}{2}]^n$. Of particular interest is the problem of finding bounds on the cross-sectional volumes of cubes. We say "cross-section" to mean the intersection between a body (e.g. Q_n) and an affine subspace of $\mathbb{R}^n$. The simplicity with which this problem can be described makes it quite attractive. It is also puzzling why there are not easier solutions.

The key results discussed will be those proved by Keith Ball in 1986 and 1989, which give upper bounds for the volumes of intersection by hyperplanes (affine subspaces of codimension 1) and for general subspaces respectively. In doing this we also introduce some powerful inequalities which are important in many areas of mathematics, namely the Brünn-Minkowski inequality and the Brascamp-Lieb inequality. We shall conclude this part with a simple conjecture on the maximum cross-sectional volumes.

In the final section, we prove an interesting result about the ubiquity of cross-sections of cubes. Finally we introduce a famous open problem known as the Mahler conjecture, which claims that cubes have minimal Mahler volume.

We aim to make this account readable for Warwick undergraduates in their final year of the MMath course, assuming only a reasonable background in analysis and some geometric intuition.

Before beginning the essay proper we briefly remind the reader of some basic notions and set out some notation for use in the remainder of this account.

1.2 Basic Concepts and Notation

We treat $\mathbb{R}^n$ as an inner product space with the usual product, denoted by $\langle\cdot,\cdot\rangle$ ("$||\cdot||$" will denote the usual ℓ^2 norm). Unless otherwise stated we let $e_1, e_2, \dots, e_n$ denote the standard orthonormal basis of $\mathbb{R}^n$. We denote by $\mathbb{S}^{n-1}$ the unit sphere in $\mathbb{R}^n$. For the orthogonal complement of a set or vector we use the "$\perp$" notation, for example $A^\perp$ or $v^\perp$. The word "subspace" should be understood to mean a linear subspace (containing the origin) unless otherwise specified (e.g. by describing it as affine).

For $A \subseteq \mathbb{R}^n$, "$|A|$" means the ($\dim_H(A)$-dimensional) Hausdorff (Lebesgue) measure of A. If a different dimension is intended we shall write $|A|_k$ for the k-dimensional Hausdorff measure.

We will frequently use the notation of Minkowski sums, which are defined as follows.

Definition 1.1. *Let $A, B \subset \mathbb{R}^n$, the Minkowski sum $A + B$ is defined to be*

$$A + B = \{a + b : a \in A, b \in B\}.$$

When adding a set to a single vector, say A and v respectively, we may write $A + v$ to mean the Minkowski sum $A + \{v\}$.

Scalar multiples of sets will denote dilations and $A - B := A + (-1 \cdot B)$. Moreover, products of intervals and vectors (e.g. $[a, b]v$ or $\mathbb{R}v$) should be understood as line segments, lines or rays, as appropriate.

Topological interior, closure and boundary of a set $A \subseteq \mathbb{R}^n$ will be denoted by $\mathrm{int}(A)$, $\overline{A}$ and ∂A respectively. Open Euclidean balls of radius $r > 0$, centred at $v \in \mathbb{R}^n$ will be denoted by $B_r(v)$ and for a set $A \subseteq \mathbb{R}^n$, $B_r(A) := A + B_r(0)$ means the open r-neighbourhood of A. If a set is called a ball then it should be assumed to be a Euclidean ball unless another norm is indicated. Given a hyperplane $H = v^\perp + tv \subset \mathbb{R}^n$, where $v \in \mathbb{S}^{n-1}$ and $t \in \mathbb{R}$, the (closed) halfspaces induced by H are $H^+ = \{x : \langle x, v\rangle \geq t\}$ and $H^- = \{x : \langle x, v\rangle \leq t\}$.

Now we recall some of the key definitions from convex geometry.

Definition 1.2.

- *A set $A \subseteq \mathbb{R}^n$ is convex if for all $\lambda \in [0, 1]$ and $x, y \in A$ we have $\lambda x + (1 - \lambda)y \in A$.*
- *A convex body is a compact convex set with non-empty interior.*
- *A set A is (centrally) symmetric if $x \in A \Rightarrow -x \in A$.*
- *The convex hull of a set A (denoted* $\mathrm{conv}(A)$*) is the intersection of all convex sets containing A.*
- *If A is convex and $x \in \partial A$, a hyperplane $H \subset \mathbb{R}^n$ is a supporting hyperplane at x if $x \in H$ and $A \subseteq H^+$ or $A \subseteq H^-$.*
- *A body is called star-shaped if it is a union of line segments that all contain the origin.*
- *A convex set is a cone if it contains 0 and is closed under addition.*

2 Cross-Sectional Volumes of Cubes (Hyperplane Case)

2.1 An Exposition of Ball's 1986 Paper

We will now discuss the proof due to Ball that the volume of any $(n-1)$-dimensional cross-section of Q_n is at most $\sqrt{2}$ regardless of n. Moreover if we only consider cross-sections generated by subspaces, the volume of intersection is always at least 1. We will essentially be following a paper by Ball from 1986, namely [9].

To prove the upper bound ($\sqrt{2}$) we first show that it suffices to consider only subspaces. This requires the Brünn-Minkowski inequality for convex sets which is the following (see [18]).

Theorem 2.1. *Let $K, L \subset \mathbb{R}^n$ be convex bodies and $\lambda \in (0,1)$, then*

$$|\lambda K + (1-\lambda)L|^{1/n} \geq \lambda |K|^{1/n} + (1-\lambda)|L|^{1/n}.$$

The Brünn-Minkowski inequality will be discussed in more detail later.

Applying Theorem 2.1 to the problem at hand we see that we can indeed restrict our attention to cross-sections arising from subspaces. In fact the following more general result holds (we generalise a result from [9] so that it is also relevant later).

Lemma 2.2. *Let $d \in \mathbb{Z}_{>0}$ such that $d \leq n$ and let S be an d–dimensional subspace of $\mathbb{R}^n$. Suppose $H = S + u$ is a translation of S by $u \in S^{\perp}$, then $|S \cap Q_n| \geq |H \cap Q_n|$.*

Proof . Let $\widetilde{H} = S - u$, then by symmetry of Q_n and S, we have

$$\widetilde{H} \cap Q_n = (S-u) \cap Q_n = (-S-u) \cap (-Q_n) = -[(S+u) \cap Q_n] = -(H \cap Q_n).$$

Hence $|\widetilde{H} \cap Q_n| = |H \cap Q_n|$. Let P be the orthogonal projection map onto S, then as H and $\widetilde{H}$ are parallel to S, P preserves the volume of subsets of H and $\widetilde{H}$. Now applying Theorem 2.1 with $\lambda = \frac{1}{2}$ we see that

$$\begin{aligned} |S \cap Q_n|^{1/d} &\geq \left| \frac{1}{2} P(H \cap Q_n) + \frac{1}{2} P(\widetilde{H} \cap Q_n) \right|^{1/d} \\ &\geq \frac{1}{2} |P(H \cap Q_n)|^{1/d} + \frac{1}{2} |P(\widetilde{H} \cap Q_n)|^{1/d} \\ &= |H \cap Q_n|^{1/d}, \end{aligned}$$

as required. The first inequality holds because, by convexity of Q_n, $\frac{1}{2}(H \cap Q_n + \widetilde{H} \cap Q_n) \subseteq S \cap Q_n$. Therefore, by linearity of P and the fact that $P|_S = \mathrm{id}_S$ we see that

$$\frac{1}{2}(H \cap Q_n + \widetilde{H} \cap Q_n) = \frac{1}{2} P(H \cap Q_n) + \frac{1}{2} P(\widetilde{H} \cap Q_n).$$

□

The next step in proving the bounds on cross-sectional volume will be to introduce a probability density function (p.d.f.) for a specific random variable so that we can apply some results from probability theory. This approach is not particularly intuitive but looking at the problem in this way allows us to use some powerful tools like the Fourier inversion formula and a handy inequality relating: the supremum of a p.d.f.; the p-norm of the associated random variable and p itself.

The following lemma introduces the (candidate) p.d.f., which we shall call f, and establishes that it is indeed the required density function.

Lemma 2.3. *Let $X_1, X_2, \ldots, X_n$ be independent random variables, each uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$, furthermore let S be a subspace of $\mathbb{R}^n$ and $u = (u_1, \ldots, u_n) \in \mathbb{S}^{n-1}$ a unit normal to S. Then the function $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ given by $f(r) = |(S + ru) \cap Q_n|$ is a p.d.f. for the random variable $X = \sum_{i=1}^{n} u_i X_i$ (where the event space is taken to be the σ-algebra of Borel sets in $\mathbb{R}$).*

Remarks:

- Of course, the idea here is that the random vector $(X_1, \ldots, X_n)$ induces the probability measure on Q_n which agrees with the Lebesgue measure. Therefore, for $\varepsilon > 0$, the probability $P(r-\varepsilon \leq X \leq r+\varepsilon)$ is the volume of Q_n intersected with the ε-neighbourhood of $S + ru$. Hence, taking $\varepsilon \to 0$, it seems natural that f is the p.d.f. of X.
- This lemma summarises part of the introduction of [9]. Note that it is in fact a little stronger, in particular, the proof does not use continuity of f so the result also holds when u is orthogonal to a face of Q_n.

Proof (of Lemma 2.3). Fix $r \in \mathbb{R}$ and $\varepsilon \geq 0$ then

$$\int_{r-\varepsilon}^{r+\varepsilon} f(t)dt = \left| \left\{ x \in Q_n : \left| \sum_{i=1}^{n} u_i x_i - r \right| \leq \varepsilon \right\} \right| = P\left(\left| \sum_{i=1}^{n} u_i X_i - r \right| \leq \varepsilon \right).$$

All we have done here is to rewrite the volume of a certain set using the probability measure induced by the $X_1, \ldots, X_n$ (see remark above).

Of course, we can write any bounded open interval $(a, b) \subset \mathbb{R}$ as $(r - \varepsilon, r + \varepsilon)$ for some $r \in \mathbb{R}$ and $\varepsilon \in [0, \infty)$. Hence for $-\infty < a < b < \infty$, $\int_a^b f(t)dt = P(X \in (a, b))$. Moreover, since Q_n is bounded, this is sufficient to show that the equality still holds if a or b is ∞. Hence by the determination of Borel measures[1] from their values on intervals we have shown that $\int_E f(t)dt = P(X \in E)$ for any Borel set $E \subseteq \mathbb{R}$, as required. $\square$

We will later need f to be continuous at 0. This could be checked directly but we have found the following more general lemma which proves an intuitive property of convex bodies. We also find a further generalisation which will be used in a later section.

Lemma 2.4. *Let $C \subset \mathbb{R}^n$ be a convex body and $H \subset \mathbb{R}^n$ an $(n-1)$–dimensional subspace with unit normal v then the function $g : \mathbb{R} \to \mathbb{R}$ given by $g(x) = |(H + xv) \cap C|$ is continuous at each x such that $H + xv$ contains an interior point of C.*

Proof . Suppose $z \in (H + xv)$ is an interior point of C. Indeed, suppose the open δ-ball $B_\delta(z)$ is contained in C. Now define a conical cap[2] K (see Figure 1) by

$$K := \{\lambda(z + \delta v) + (1 - \lambda)h : h \in (H + xv) \cap C, \lambda \in [0, 1]\}.$$

Consider the family of hyperplanes $H_\varepsilon = H + (x + \varepsilon)v$ parametrised by $\varepsilon > 0$. Define an associated family of conical caps $K_\varepsilon = \{\lambda(z - \delta v) + (1 - \lambda)h : h \in H_\varepsilon \cap C, \lambda \in [0, 1]\}$. Now by convexity of C, we have $K \subset C$ and $K_\varepsilon \subset C$ for all $\varepsilon > 0$. Therefore

$$g(x) = |(H + xv) \cap C| \geq |(H + xv) \cap K_\varepsilon| = \left(\frac{\delta}{\delta + \varepsilon}\right)^{n-1} |H_\varepsilon \cap C| = \left(\frac{\delta}{\delta + \varepsilon}\right)^{n-1} g(x + \varepsilon).$$

Similarly, if $\varepsilon < \delta$, we get the following lower bound on $g(x + \varepsilon)$.

$$g(x + \varepsilon) = |H_\varepsilon \cap C| \geq |H_\varepsilon \cap K| = \left(\frac{\delta - \varepsilon}{\delta}\right)^{n-1} g(x).$$

[1] Some readers might find it useful to think of this argument in terms of the Radon-Nikodym derivatives of measures.

[2] We use the term "conical cap" to mean the convex hull of a convex set with a point added (the "vertex" of the conical cap).

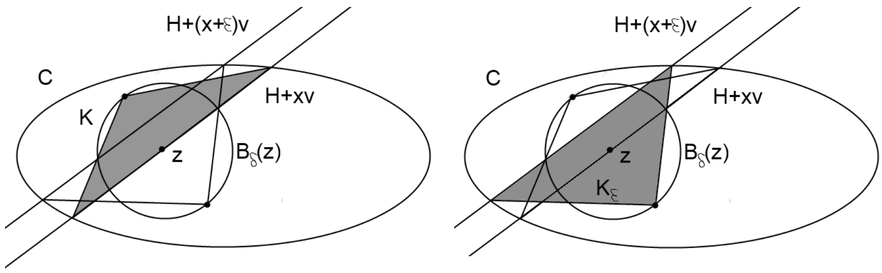

Figure 1: Constructing "conical caps".

Hence for $\varepsilon \in (0, \delta)$ we get

$$\left(\frac{\delta}{\delta - \varepsilon}\right)^{n-1} g(x+\varepsilon) \geq g(x) \geq \left(\frac{\delta}{\delta + \varepsilon}\right)^{n-1} g(x+\varepsilon). \tag{2.5}$$

The argument is easily modified for $\varepsilon \in (-\delta, 0)$ so g is indeed continuous at x. $\square$

Observe that the coefficients in (2.5) are independent of g (only depending on C, x and z) therefore we get the following generalisation.

Corollary 2.6. *If $C \subset \mathbb{R}^n$ is a convex body and $S \subseteq \mathbb{R}^n$ is a d-dimensional subspace then $g : S^\perp \to \mathbb{R}$ given by $g(x) = |(S+x) \cap C|$ is continuous at every x such that $(S+x) \cap \operatorname{int}(C) \neq \emptyset$.*

Proof (Sketch). Fix $\delta > 0$ and $z \in (S + x)$ such that $B_\delta(z) \subset C$. For $y \in S^\perp$ with $||y|| < \delta$, apply the previous proof on the $(d+1)$-dimensional affine subspace containing $S+x$ and $S+x+y$ to obtain the following bounds.

$$\left(\frac{\delta}{\delta - ||y||}\right)^{n-1} g(x+y) \geq g(x) \geq \left(\frac{\delta}{\delta + ||y||}\right)^{n-1} g(x+y).$$

The coefficients depend only on $||y||$ so continuity follows. $\square$

Prior to proving the lower bound on the volume of intersection of Q_n and an $(n-1)$–dimensional subspace we need the aforementioned lemma concerning the p-norm of a random variable, which will enable us to apply Lemma 2.3. This is Lemma 1 from [9]. Note that for a real random variable Y with p.d.f. g, $||Y||_p$ denotes $\left(\int_{-\infty}^{\infty} g(x)|x|^p\right)^{1/p}$ (not to be confused with the L^p norm).

Lemma 2.7. *Let Y be a real random variable with p.d.f. g, then for $p > 0$,*

$$||Y||_p ||g||_\infty \geq \frac{1}{2}(p+1)^{-1/p},$$

where $||g||_\infty$ is of course the essential supremum of g.

Proof . We may assume that Y is symmetric, for if not we can consider the random variable $\bar{Y}$ with p.d.f. $\bar{g}(x) = \frac{1}{2}(g(-x) + g(x))$. Then $||\bar{Y}||_p = ||Y||_p$ and $||\bar{g}||_\infty \leq ||g||_\infty$, so showing $||\bar{Y}||_p ||\bar{g}||_\infty \geq \frac{1}{2}(p+1)^{-1/p}$ would suffice.

Set $G(x) = \int_0^x g(t)dt$ for $x \geq 0$ so $G(0) = 0$ and $G(\infty) = 1/2$ by symmetry. By construction G is absolutely continuous and $G'(x) = g(x)$ for almost all x. We therefore obtain the following.

$$\begin{aligned} 2^{-p} = 2G(\infty)^{p+1} = 2\int_0^\infty (G(x)^{p+1})' dx = 2(p+1)\int_0^\infty g(x)G(x)^p dx \\ \leq 2(p+1)||g||_\infty^p \int_0^\infty g(x)x^p dx. \end{aligned}$$

This last inequality holds because $G(x) \leq x||g||_\infty$ for all $x \geq 0$. The proof is completed by taking the p-th root and replacing $\left(\int_0^\infty g(x)x^p dx\right)^{1/p}$ by $||Y||_p$ □

The lower bound now follows easily, this is Theorem 2 in [9].

Proposition 2.8. *Let $S \subset \mathbb{R}^n$ be a subspace with dimension $(n-1)$ then $|S \cap Q_n| \geq 1$.*

Proof . Fix a unit normal $u = (u_1, \dots, u_n)$ to S and let $X, X_1, \dots, X_n$ and f be as above. Since $|S \cap Q_n| = f(0)$ and f is continuous at 0 (where the maximum is attained), it suffices to show that $||f||_\infty \geq 1$.

To apply Lemma 2.7, consider $||X||_2 = \sqrt{\mathbb{E}(X^2)}$. Since the $X_1, \dots, X_n$ are independent and identically distributed, each with expected value 0, we have $\mathbb{E}(X_i X_j) = 0$ for all $i \neq j$. Moreover,

$$\mathbb{E}(X^2) = \left(\sum_{i=1}^n u_i^2\right)\mathbb{E}(X_1^2) = \mathbb{E}(X_1^2) = \int_{-1/2}^{1/2} x^2 dx = 1/12.$$

Hence $||X||_2 = \frac{1}{2\sqrt{3}}$. Therefore, by Lemma 2.7 (with $p = 2$), $||f||_\infty \geq \frac{1}{2\sqrt{3}||X||_2} = 1$ as required. □

Next we set about proving the upper bound on $f(0)$. To begin with we need a lemma to estimate the L^p norm of the sinc function, defined by

$$\mathrm{sinc}(x) := \begin{cases} \frac{1}{x}\sin(x) & x \neq 0 \\ 1 & x = 0 \end{cases}.$$

The full proof of this lemma can be found in [9] but here we only explain an adaptation of the first part of the proof which leads to a weaker upper bound. The original version (lemma 3 in [9]) is as follows.

Lemma 2.9. *If $p \geq 2$ then*

$$\frac{1}{\pi}\int_{-\infty}^\infty |\mathrm{sinc}(t)|^p dt \leq \sqrt{\frac{2}{p}},$$

with equality if and only if $p = 2$.

However here we only prove the following.

Lemma 2.10. *If $p \geq 2$ then*

$$\frac{1}{\pi}\int_{-\infty}^\infty |\mathrm{sinc}(t)|^p dt \leq \alpha\sqrt{\frac{2}{p}},$$

where

$$\alpha = \begin{cases} \sqrt{\frac{2}{e\log(2)}}\ (\approx 1.030) & \text{if } 2 \leq p < 4 \\ 1 & \text{if } p \geq 4 \end{cases}.$$

Remark: The essential problem with bounding this integral is the "central peak" of the sinc function. Since $\mathrm{sinc}(0) = 1$, its difficult to find a nice bound which decays sufficiently fast as p increases for integrals of the form $\int_{-\varepsilon}^{\varepsilon} |\mathrm{sinc}(x)|^p dx$. In the proof of Lemma 2.9, Ball finds a simple bound when $p \geq 4$ but has to do a lot of work for $2 \leq p < 4$. We exhibit the first part here but apply an interpolation[3] to get the weaker bound for smaller p.

Proof (of Lemma 2.10). To begin with suppose $p \geq 4$. We first show that for $t^2 \leq 36/5$, we have $0 \leq \mathrm{sinc}(t) \leq e^{-t^2/6}$. For this consider the Taylor series expansions (about 0).

$$\mathrm{sinc}(t) = 1 - \frac{t^2}{6} + \frac{t^4}{5!} - \frac{t^6}{7!} + \frac{t^8}{9!} - \cdots < 1 - \frac{t^2}{6} + \frac{t^4}{5!} \quad \text{if } t^2 < \frac{9!}{7!} = 72.$$

Similarly,

$$\begin{aligned} e^{-t^2/6} &= 1 - \frac{t^2}{6} + \frac{t^4}{6^2 \cdot 2!} - \frac{t^6}{6^3 \cdot 3!} + \frac{t^8}{6^4 \cdot 4!} - \frac{t^{10}}{6^5 \cdot 5!} - \cdots \\ &> 1 - \frac{t^2}{6} + \frac{t^4}{6^2 \cdot 2!} - \frac{t^6}{6^3 \cdot 3!} \quad \text{if } t^2 < \frac{6^5 \cdot 5!}{6^4 \cdot 4!} = 30. \end{aligned}$$

Therefore, if $t^2 \leq 36/5 = 6^3 \cdot 3!/180$, we have

$$e^{-t^2/6} - \mathrm{sinc}(t) \geq \frac{t^4}{6^2 \cdot 2!} - \frac{t^6}{6^3 \cdot 3!} - \frac{t^4}{5!} = \frac{t^4}{180} - \frac{t^6}{6^3 \cdot 3!} \geq 0.$$

The fact that $\mathrm{sinc}(t) \geq 0$ for $t^2 \leq 36/5$ follows from the fact that $\sqrt{36/5} < \pi$.

For the sake of brevity, let $c = \sqrt{36/5}$. Now we can use $e^{-t^2/6}$ to estimate the "central part" of the integral in Lemma 2.10 (i.e. the part on $[-c, c]$). We find a cruder estimate on the rest.

$$\begin{aligned} \int_{-\infty}^{\infty} |\mathrm{sinc}(t)|^p dt &\leq \int_{-c}^{c} e^{-pt^2/6} dt + \int_{\mathbb{R}\setminus[-c,c]} |t|^{-p} dt < \int_{-\infty}^{\infty} e^{-pt^2/6} dt + 2\int_{c}^{\infty} t^{-p} dt \\ &= \sqrt{\frac{\pi \cdot 6}{p}} + \frac{2}{(p-1) \cdot c^{p-1}} = \frac{1}{\sqrt{p}} \left(\sqrt{6\pi} + \frac{2\sqrt{p}}{(p-1) \cdot c^{p-1}} \right). \end{aligned}$$

Since $p \geq 4$ we see that $\frac{\sqrt{p}}{p-1} \leq \frac{2}{3}$ and $c^{p-1} \geq c^3 = 216/\sqrt{125}$. Therefore

$$\sqrt{6\pi} + \frac{2\sqrt{p}}{(p-1) \cdot c^{p-1}} \leq \sqrt{6\pi} + \frac{4\sqrt{125}}{3 \cdot 216} < \pi\sqrt{2}.$$

This gives the required bound.

For $2 \leq p < 4$, first notice that we have equality for $p = 2$ (it is a well known fact that $||\mathrm{sinc}||_2 = \sqrt{\pi}$). Moreover, by an application of Hölder's inequality, we deduce that

$$||\mathrm{sinc}||_p^p \leq ||\mathrm{sinc}||_2^{4-p} ||\mathrm{sinc}||_4^{2p-4}.$$

To prove this, one could consider $\int_{\mathbb{R}} |\mathrm{sinc}(x)|^{2\alpha} |\mathrm{sinc}(x)|^{4(1-\alpha)} dx$ for a suitable choice of $\alpha \in [0, 1]$. Now we have already found upper bounds on the terms of the right hand side so we obtain the following.

$$||\mathrm{sinc}||_{L^p}^p \leq \pi^{\frac{4-p}{2}} \left(\frac{\pi}{\sqrt{2}} \right)^{\frac{2p-4}{4}} = \pi \left(\sqrt{2} \right)^{\frac{2-p}{2}} \leq \pi \sqrt{\frac{2}{e \log(2)} \cdot \frac{2}{p}}.$$

The last inequality is the desired result and follows from elementary calculus (finding the maximum of $2^{\frac{2-p}{4}} \sqrt{p/2}$ on the interval $[2, 4]$). □

[3] Thanks to Professor James Robinson for suggesting this trick.

Finally, we will use a generalisation of Hölder's inequality which we state here (see [13], page 67). This can be obtained by applying induction to the standard 2-exponent version. Later on we will discuss a much more general result, namely the Brascamp-Lieb inequality.

Lemma 2.11. *Let $p_i > 1$ for $i = 1, \ldots, n$ be constants such that $\sum_{i=1}^{n} p_i^{-1} = 1$. Then for any functions $g_i \in L^{p_i}(\mathbb{R})$ $(i = 1, 2, \ldots, n)$ we have $g_1 \cdot g_2 \cdot \ldots \cdot g_n \in L^1(\mathbb{R})$ and*

$$||g_1 \cdot \ldots \cdot g_n||_1 \leq \prod_{i=1}^{n} ||g_i||_{p_i}.$$

We are now in a position to prove the upper bound on the volumes of cross-sections of Q_n, this is Theorem 4 in [9] and is the main result of this section.

Theorem 2.12. *If S is an $(n-1)$–dimensional subspace of $\mathbb{R}^n$ then $|S \cap Q_n| \leq \sqrt{2}$.*

Proof . Let $u = (u_1, u_2, \ldots, u_n)$ be a unit normal to S, then by symmetry of the cube, we may assume that $u_i \geq 0$ for $i = 1, 2, \ldots, n$. Moreover, if $u_i = 0$ for some i then $|S \cap Q_n| = |(u_1, \ldots, \hat{u}_i, \ldots, u_n)^{\perp} \cap Q_{n-1}|$ and the problem reduces by one dimension. Thus, by induction, we may assume that $u_i > 0$ for all i.

There are two cases in this proof, the first is a fairly straightforward geometrical argument while the second requires some of the material previously discussed.

Case 1: Suppose that $u_i \geq 1/\sqrt{2}$ for some i, (by symmetry again, we may assume $i = 1$). Consider the cylinder C with "long axis" $e_1 = (1, 0, \ldots, 0)$ defined by $C = \mathbb{R} \times [-\frac{1}{2}, \frac{1}{2}]^{n-1}$. Now clearly $Q_n \subset C$, so $|S \cap Q_n| \leq |S \cap C|$. Moreover $S \cap C$ is the projection in the direction e_1 of $\{0\} \times [-\frac{1}{2}, \frac{1}{2}]^{n-1}$ onto S.

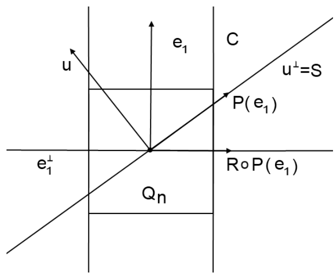

Figure 2: Reducing to the cylinder and projections of e_1.

This projection has the effect of multiplying volumes in $e_1^{\perp}$ by $||P(e_1)||/||R \circ P(e_1)||$ were P and R are orthogonal projections onto S and $e_1^{\perp}$ respectively (see Figure 2). To check this, one might like to consider the 2-plane spanned by e_1 and v. It follows that

$$\begin{aligned}
|S \cap C| &= |Q_{n-1}| \cdot ||P(e_1)||/||R \circ P(e_1)|| \\
&= 1 \cdot ||e_1 - \langle e_1, u\rangle u||/||e_1 - \langle e_1, u\rangle u - e_1(1 - \langle e_1, u\rangle^2)|| \\
&= \frac{1}{u_1\sqrt{1-u_1^2}} \sqrt{(1-u_1^2)^2 + u_1^2(u_2^2 + u_3^2 + \ldots + u_n^2)} \\
&= \frac{1}{u_1\sqrt{1-u_1^2}} \sqrt{(1-u_1^2)^2 + u_1^2(1-u_1^2)} = \frac{\sqrt{1-u_1^2}}{u_1\sqrt{1-u_1^2}} = u_1^{-1}.
\end{aligned}$$

This completes the proof of the first case since we assumed $u_1 \geq 1/\sqrt{2}$.

Case 2: If $u_i < 1/\sqrt{2}$ for all $i = 1, 2, \dots, n$, then consider the random variables $X_1, \dots, X_n$ from Lemma 2.3. For each i, let ϕ_i be the characteristic function of the random variable $u_i X_i$ i.e. $\phi_i(t) = \mathbb{E}(e^{itu_iX_i})$. Each X_i is symmetrically distributed so $\phi_i(t) = \mathbb{E}(e^{-itu_iX_i})$ which is just the Fourier transform of $u_i \mathbb{1}_{[-1/2,1/2]}$. Here $\mathbb{1}_A$ denotes the indicator function (often called the characteristic function) of $A \subseteq \mathbb{R}^n$. Therefore $\phi_i(t) = \operatorname{sinc}(\frac{1}{2}u_i t)$.

Now since the $X_1, \dots, X_n$ are independent (as are the random variables $e^{itu_iX_i}$ for fixed t), the characteristic function ϕ of $X := \sum_{i=1}^n u_i X_i$ is given by

$$\phi(t) = \prod_{i=1}^n \operatorname{sinc}\left(\frac{u_i t}{2}\right).$$

Furthermore, by Lemma 2.3, ϕ is the Fourier Transform of $f(x) = |(S + xu) \cap Q_n|$. Since f is an even function, this is given by $\hat{f}(t) = \int_{-\infty}^{\infty} \cos(xt) f(x)dx$. [4]

Since f is continuous at 0 (by Lemma 2.4), compactly supported and bounded, we may apply the Fourier inversion formula to see that

$$f(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \prod_{i=1}^n \frac{2\sin(u_i t/2)}{u_i t} e^{it\cdot 0} dt = \frac{1}{\pi}\int_{-\infty}^{\infty} \prod_{i=1}^n \operatorname{sinc}(u_i t) dt.$$

Now apply Lemma 2.11 (generalised Hölder's) with $p_i = u_i^{-2}$ (so $\sum_i \frac{1}{p_i} = \sum_i u_i^2 = 1$ and $p_i > 2$) and $g_i(t) = \operatorname{sinc}(u_i t)$ (clearly $\operatorname{sinc} \in L^p(\mathbb{R})$ for all $p > 1$). This shows that

$$f(0) \leq \frac{1}{\pi} \prod_{i=1}^n \left(\int_{-\infty}^{\infty} |\operatorname{sinc}(u_i t)|^{p_i} dt \right)^{1/p_i} = \prod_{i=1}^n \left(\frac{1}{u_i \pi} \int_{-\infty}^{\infty} |\operatorname{sinc}(t)|^{p_i} dt \right)^{1/p_i}.$$

By Lemma 2.9, this gives the bound

$$f(0) \leq \prod_{i=1}^n \left(\frac{\sqrt{2}}{u_i \sqrt{p_i}} \right)^{1/p_i} = \prod_{i=1}^n (\sqrt{2})^{1/p_i} = \sqrt{2}.$$

This is what we set out to prove. □

This completes our exposition of Ball's 1986 paper. We now discuss the Brünn-Minkowski inequality which we previoulsy stated.

2.2 Proof of the Brünn-Minkowski Inequality

Recall the statement of the Brünn-Minkowski inequality (Theorem 2.1). For convex bodies $A, B \subset \mathbb{R}^n$ and $\lambda \in [0, 1]$, we have

$$|\lambda A + (1 - \lambda)B|^{1/n} \geq \lambda |A|^{1/n} + (1 - \lambda)|B|^{1/n}.$$

In some sense this gives a lower bound on the volume of "linear combinations" of pairs of convex bodies, in terms of the volumes of the bodies in question. The proof given here will be based on the proofs in [23] pages 69-71 and [18], however we add several lemmas in order to flesh out the measure-theoretic detail. We begin with a definition (see [23]).

[4] Note that here we have chosen the convention for the Fourier transform given by $\hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-i\xi x} f(x)dx$ as it agrees with the definition of charateristic functions in probability.

Definition 2.13. *A set $e \subset \mathbb{R}^n$ is called elementary if it is the union of finitely many (non-degenerate, closed) axis-parallel cuboids that have pairwise disjoint interiors. Let the collection of all such sets be denoted by E^n.*

We will use the following result from Measure Theory:

Lemma 2.14. *If $A \subset \mathbb{R}^n$ is compact then for any positive ε there exists $e \in E^n$ with $A \subseteq e$ and $|e| < |A| + \varepsilon$. Moreover given any $\delta > 0$ we may assume that $e \subset B_\delta(A)$.*

Proof . From the properties of the product measure λ^n we may choose a countable collection of open axis-parallel cuboids whose union has volume at most $|A| + \varepsilon$ and contains A. By compactness we may assume that this collection is in fact finite. We can then take the closure of each cuboid and subdivide (if necessary) to get an elementary set containing A and with small enough volume.

For the second claim, given any elementary set $e \supseteq A$, subdivide each cuboid into smaller ones with diameter at most $\delta/2$ then we may remove any that do not intersect A to get $e' \supseteq A$ with $e' \subset B_\delta(A)$ and $|e'| \leq |e|$. □

Next we verify that elementary sets can be "cut up" to reduce the number of cuboids[5] and sketch an intuitive result about convex sets.

Lemma 2.15. *If $e \in E^n$ consists of $k \geq 2$ cuboids then there exists an axis-parallel hyperplane H such that $e \cap H^+$ and $e \cap H^-$ both consist of fewer than k cuboids.*

Proof . It obviously suffices to consider the case $k = 2$. Suppose $e = C_1 \cup C_2$ where $C_1 = [a_1, b_1] \times \ldots \times [a_n, b_n]$ and $C_2 = [c_1, d_1] \times \ldots \times [c_n, d_n]$ are cuboids with mutually disjoint interiors. Let

$$P_i = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_i = a_i\},$$

and

$$Q_i = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_i = b_i\},$$

i.e the hyperplanes containig the faces of C_1. Now suppose for contradiction that for each i, neither P_i nor Q_i "separate" C_1 from C_2 i.e. for each of these hyperplanes, one of the corresponding half-spaces intersects the interiors of both cuboids. Then for each i we see that $(a_i, b_i) \cap (c_i, d_i) \neq \emptyset$. Hence $\mathrm{int}(C_1) \cap \mathrm{int}(C_2) \neq \emptyset$. This is a contradiction, thus for some i, either P_i or Q_i has the required property. □

Lemma 2.16. *Let $C \subset \mathbb{R}^n$ be a convex set with interior point x, indeed suppose that $B_\varepsilon(x) \subset C$. Then for any $\delta > 0$, $x + (1 + \delta) \cdot (C - x)$ contains the open neighbourhood $B_{\delta \cdot \varepsilon}(C)$ of C.*

Proof (Sketch). Given $p \in C$ and consider the conical cap $K = \bigcup_{\lambda \in [0,1]} \lambda B_\varepsilon(x) + (1 - \lambda)p$. By convexity $K \subseteq C$. It is easy to check that the conical cap $K' = x + (1 + \delta) \cdot (K - x)$ contains the ball $B_{\delta \cdot \varepsilon}(p)$ and the result follows. □

Finally we need a lemma to relate the volumes of "linear combinations" of convex bodies and approximating elementary sets.

Lemma 2.17. *Let $A, B \subset \mathbb{R}^n$ be convex bodies, then for all $\varepsilon, \delta > 0$ there exist elementary sets $A' \supset A, B' \supset B$ with $|A'| < |A| + \delta$, $|B'| < |B| + \delta$ and $|A' + B'| \leq (1 + \varepsilon)|A + B|$.*

[5]This result is essentially obvious, but finding intuition in higher dimensions is difficult so it is reasonable to check it rigorously.

Proof . Let x be an interior point of $A+B$ and let $\alpha > 0$ such that $B_\alpha(x) \subset A+B$. Now by Lemma 2.14, there exist $A', B' \in E^n$ containing A and B respectively, such that $|B'| < |B| + \delta$, $|A'| < |A| + \delta$. Furthermore, we may assume that $B' \subset B_{\alpha r/2}(B)$ and $A' \subset B_{\alpha r/2}(A)$ where $r = (1+\varepsilon)^{1/n} - 1$. Therefore $A' + B' \subseteq A + B_{\alpha r/2}(0) + B + B_{\alpha r/2}(0) = B_{\alpha r}(A+B)$ since Minkowski sums are commutative.

Now $A+B$ is a convex body, so by Lemma 2.16, we have that $B_{\alpha r}(A+B)$ is a subset of $x + (1+r) \cdot (A+B-x)$ which has volume $(1+r)^n|A+B| = (1+\varepsilon)|A+B|$. Hence $|A'+B'| \le (1+\varepsilon)|A+B|$ as required. □

With these technicalities in place we can now prove the Brünn-Minkowski inequality.

Proof (of Theorem 2.1). We first consider the case when A and B are axis-parallel cuboids with respective side lenths a_i,b_i in the direction of the ith standard-basis vector, $i = 1, 2, \ldots, n$. Observe that $A+B$ is a cuboid with side lengths $a_i + b_i$, hence with volume $\prod_{i=1}^n (a_i + b_i)$. Now by the standard inequality between geometric and arithmetic means we have

$$\left(\prod_{i=1}^n \frac{a_i}{a_i+b_i}\right)^{\frac{1}{n}} + \left(\prod_{i=1}^n \frac{b_i}{a_i+b_i}\right)^{\frac{1}{n}} \le \frac{1}{n}\sum_{i=1}^n \frac{a_i}{a_i+b_i} + \frac{1}{n}\sum_{i=1}^n \frac{b_i}{a_i+b_i} = 1.$$

Thus $|A+B|^{1/n} \ge |A|^{1/n} + |B|^{1/n}$ as required.

Now if A and B are elementary sets consisting of a combined total of $k \ge 2$ cuboids we proceed by induction on k. We have just dealt with the initial case (when $k = 2$) so suppose $k > 2$. We may assume that A contains at least two cuboids. Since A is elementary, by Lemma 2.15 there exists an axis-parallel hyperplane H, such that $A^+ := A \cap H^+$ and $A^- := A \cap H^-$ are both elementary sets with strictly fewer cuboids than A.

Let $\mu = |A^+|/|A|$ then by translating B if necessary, we may assume that $B = B^+ \cup B^-$ where $B^\pm = B \cap H^\pm$ and $|B^+|/|B| = \mu$. Now $A^+ + B^+$ and $A^- + B^-$ are subsets of $A+B$ with disjoint interiors and each made up of at most $k-1$ cuboids. Together with the inductive hypothesis this implies that

$$\begin{aligned} |A+B| &\ge |A^+ + B^+| + |A^- + B^-| \\ &\ge (|A^+|^{1/n} + |B^+|^{1/n})^n + (|A^-|^{1/n} + |B^-|^{1/n})^n \\ &= (\mu + 1 - \mu)(|A|^{1/n} + |B|^{1/n})^n. \end{aligned}$$

This concludes the case for elementary sets. All that remains is to apply some of the measure theory prepared earlier.

Let $A, B \subset \mathbb{R}^n$ be convex bodies and $\varepsilon > 0$. Then by Lemma 2.17 there exist elementary sets A', B' containing A and B respectively such that $|A'+B'|^{1/n} \le [(1+\varepsilon)|A+B|]^{1/n}$. Thus, by the previous case $|A|^{1/n} + |B|^{1/n} \le |A'|^{1/n} + |B'|^{1/n} \le [(1+\varepsilon)|A+B|]^{1/n}$ for any positive ε. This completes the proof (the generalisation to $\lambda A + (1-\lambda)B$ is obvious). □

The Brünn-Minkowski inequality has several generalisations and applications in various areas of mathematics. For a survey of some of these, as well as a disscussion of how it relates to other important inequalites (including Brascamp-Lieb), one might like to access [18] by Gardener.

2.3 An Application to the Busemann-Petty Problem

The upper bound in Theorem 2.12 has a fairly straightforward application to a well-known problem. In 1956 Herbert Busemann and Clinton Petty proposed ten problems in convex geometry (see [6]). The first of these has become known as the Busemann-Petty problem. One statement of the problem is the following.

Problem 2.18. *Let A and B be symmetric convex bodies in $\mathbb{R}^n$. Suppose that for any $(n-1)$–dimensional subspace H we have $|A \cap H| \geq |B \cap H|$, does it follow that $|A| \geq |B|$?*

The answer in general is, surprisingly, negative. In particular if $n \geq 5$, it is not true. However for $n \leq 4$ the answer is indeed affirmative (see [5]).

The bound on the cross-sectional volumes of cubes provides counterexamples for $n \geq 10$. To prove it, we will need the following proposition. Later on we shall flesh out the proof that is sketched in the appendix of [10].

Proposition 2.19. *The volume of the intersection of an $(n-1)$–dimensional subspace of $\mathbb{R}^n$ and the (centred) ball with unit volume in $\mathbb{R}^n$ is strictly increasing (with respect to $n \geq 1$).*

The following theorem is a consequence of this proposition.

Theorem 2.20. *The Busemann-Petty problem has negative solution for $n \geq 10$.*

Proof . Consider the central unit cube Q_n in $\mathbb{R}^n$ and the central ball B_n of unit volume in $\mathbb{R}^n$. As these objects have the same volume, it suffices to prove that the hyperplane (subspace) cross-sectional volumes of B_n are strictly greater than $\sqrt{2}$ (i.e. the maximum cross-sectional volume of Q_n) when $n \geq 10$. Of course, by Proposition 2.19 we only need an estimate in the case $n = 10$.

Denote by v_n the volume of the unit ball in $\mathbb{R}^n$. Recall that v_n is given by

$$v_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}.$$

Here $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$ is the usual gamma function.

Let $a_n = |B_n \cap H|$ for any $(n-1)$–dimensional subspace $H \subset \mathbb{R}^n$, then we have

$$a_n = \frac{v_{n-1}}{v_n^{\frac{n-1}{n}}} = \frac{\pi^{\frac{n-1}{2}}(\Gamma(\frac{n+2}{2}))^{\frac{n-1}{n}}}{\pi^{\frac{n\cdot(n-1)}{2n}}\Gamma(\frac{n+1}{2})} = \frac{(\Gamma(\frac{n+2}{2}))^{\frac{n-1}{n}}}{\Gamma(\frac{n+1}{2})}.$$

Recall that for $n \in \mathbb{Z}_{>0}$, $\Gamma(n) = (n-1)!$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. More generally, $\Gamma(x+1) = x\Gamma(x)$ for $x \in \mathbb{R}_{>0}$. Hence when $n = 10$ we get

$$a_{10} = \frac{\Gamma(6)^{\frac{9}{10}}}{\Gamma(5+\frac{1}{2})} = \frac{5!^{\frac{9}{10}}}{4\frac{1}{2}\cdot 3\frac{1}{2}\cdot 2\frac{1}{2}\cdot 1\frac{1}{2}\cdot \frac{1}{2}\cdot\sqrt{\pi}} > 1.420 > \sqrt{2}.$$

This is what we wanted to show. Furthermore, a similar calculation shows that $a_9 < 1.41 < \sqrt{2}$ so these particular convex bodies do not provide counterexamples in lower dimensions. $\square$

We now return to prove Proposition 2.19 regarding the monotonicity of the volumes of cross-section of balls with unit volume.

One is tempted to try a direct approach, for example by differentiating

$$a_x = \frac{\left(\Gamma\left(\frac{x+2}{2}\right)\right)^{\frac{x-1}{x}}}{\Gamma(\frac{x+1}{2})}.$$

However, doing so does not seem to give the required result, even using some recently proved bounds on Γ and $\Psi = \Gamma'/\Gamma$ (the digamma function) such as those from [16].

Instead we elaborate on the proof sketched by Ball. It is quite elementary but the choice of steps is not obvious so we spend some time on the details. The first step uses the following simple result about integrating convex functions.

Lemma 2.21. *If $f : [a,b] \to \mathbb{R}$ is a C^2 convex function on a bounded interval (i.e. $f''(x) \geq 0$ in (a,b)) then*

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(x)dx \leq \frac{f(a)+f(b)}{2}.$$

Proof . As f is convex, the region above the graph of f, i.e. $U := \{(x,y) \in [a,b]\times\mathbb{R} : y \geq f(x)\}$ is a convex set. Hence U is contained in the upper half-plane generated by any tangent. In particular $U \subseteq \{(x,y) : y \geq f'((a+b)/2)x + C\}$, where $C = f((a+b)/2) - f'((a+b)/2)(a+b)/2$. Thus $f(x) \geq f'((a+b)/2)x + C$ on (a,b), so

$$\int_a^b f(x)dx \geq \frac{b^2-a^2}{2}f'\left(\frac{a+b}{2}\right) + (b-a)f\left(\frac{a+b}{2}\right) - \frac{b^2-a^2}{2}f'\left(\frac{a+b}{2}\right)$$

as required. This proves the left-hand inequality.

For the right hand side, observe that by convexity of U, $f(x) \leq \frac{1}{b-a}[f(a)(b-x)+f(b)(x-a)]$. Integrating gives

$$\begin{aligned}(b-a)\int_a^b f(x)dx &\leq \frac{b^2-a^2}{2}(f(b)-f(a)) + (b-a)(f(a)b - f(b)a)\\ &= \frac{(b-a)^2(f(b)+f(a))}{2}.\end{aligned}$$

This proves the right-hand inequality. ☐

Lemma 2.22. *The function $f : (0,\infty) \to (0,\infty)$ given by $f(x) := \left(\frac{x+1}{x}\right)^{2x+1}$ is decreasing and the function $g : (1,\infty) \to (0,\infty)$ given by $g(x) := \left(\frac{x}{x-1}\right)^{2x-1-1/x}$ is increasing.*

Proof . The derivative of f is given by

$$f'(x) = \left[2\log\left(\frac{x+1}{x}\right) - \frac{2x+1}{x(x+1)}\right]f(x),$$

so it suffices to show that

$$\log\left(1+\frac{1}{x}\right) \leq \frac{2x+1}{2x(x+1)}.$$

Now the function $x \mapsto x^{-1}$ is convex. Therefore by Lemma 2.21 on the interval $(x, x+1)$,

$$\log(x+1) - \log(x) = \int_x^{x+1}\frac{1}{x}dx \leq \frac{1}{2x} + \frac{1}{2(x+1)} = \frac{2x+1}{2x(x+1)}$$

as required.

Similarly, the derivative of g is

$$\begin{aligned}g' &= \left[\left(2+\frac{1}{x^2}\right)\log\left(\frac{x}{x-1}\right) - \left(2x-1-\frac{1}{x}\right)\frac{x-1}{x(x-1)^2}\right]g(x)\\ &= \left[\frac{2x^2+1}{x^2}\log\left(\frac{x}{x-1}\right) - \frac{2x+1}{x^2}\right]g(x).\end{aligned}$$

So we only need to show that

$$\log\left(\frac{x}{x-1}\right) \geq \frac{2x+1}{2x^2+1}.$$

In fact, more is true. If we apply the first inequality of Lemma 2.21 as before on the interval $(x-1,x)$, we get

$$\log(x) - \log(x-1) \geq \frac{1}{x-\frac{1}{2}} = \frac{2}{2x-1} \geq \frac{2x+1}{2x^2+1},$$

as required. □

These give rise to the following result (which is what we actually need).

Corollary 2.23. *If $n \geq 3$ then*

$$\left(\frac{n-1}{n-2}\right)^{\frac{n-3}{2n-5}} \left(\frac{n+1}{n}\right)^{\frac{n}{2n-1}} < \frac{n}{n-1}.$$

Proof . For $n \geq 3$, by Lemma 2.22 we have (by the first part)

$$\left(\frac{n}{n-1}\right)^{\frac{n}{2n+1}} \geq \left(\frac{n-1}{n-2}\right)^{\frac{n}{2n-1}}.$$

Moreover, by the second part

$$\left(\frac{n}{n-1}\right)^{\frac{(2n+1)(n-1)^2(n-3)}{n(2n-1)(n-2)(2n-5)}} \geq \left(\frac{n+1}{n}\right)^{\frac{n-3}{2n-5}}.$$

So the result follows from the easily checked fact that, for $n \geq 3$,

$$\frac{(2n+1)(n-1)^2(n-3)}{n(2n-1)(n-2)(2n-5)} + \frac{n}{2n+1} < 1.$$

□

Now we can prove the proposition.

Proof (of Proposition 2.19). For $n \geq 0$ let $I_n = \int_{-\pi/2}^{\pi/2} \cos^n x dx$ and note that I_n is strictly decreasing and moreover $I_{n+2} = \frac{n+1}{n+2} I_n$. Using the same notation as above, $a_n = v_{n-1} v_n^{-(n-1)/n}$ is the sequence we claim is strictly increasing. It is known that $v_n = I_n v_{n-1}$, so $a_n = v_n^{1/n}/I_n$ and we only need to show that for $n \geq 1$

$$1 < \frac{a_n}{a_{n-1}} = \left[\left(\frac{I_{n-1}}{I_n}\right)^{n-1} \left(\frac{v_n^{\frac{n-1}{n}}}{v_{n-1}}\right)\right]^{\frac{1}{n-1}} = \left[\left(\frac{I_{n-1}}{I_n}\right)^{n-1} \frac{1}{a_n}\right]^{\frac{1}{n-1}}. \tag{2.24}$$

Let $u_n = \left(\frac{I_{n-1}}{I_n}\right)^{n-1}$ then $u_n > 1$ for $n \geq 1$ thus, in particular $u_1 > a_1$ and $u_2 > a_1$. Moreover, from the above observations, one can easily obtain that $a_n = u_n^{\frac{1}{n}} a_{n-1}^{\frac{n-1}{n}}$. It follows that if (u_n) is a strictly increasing sequence and $u_n > a_{n-1}$ then $u_{n+1} > u_n > a_n$. In this case $u_n > a_n$ for all $n \geq 1$ by induction, which would complete the proof as this is exactly (2.24).

Now we show that (u_n) is indeed strictly increasing. Let $x_n = \left(\frac{u_{n+1}}{u_n}\right)^{\frac{1}{2n-1}}$ then

$$x_n = \frac{I_n}{\left(I_{n+1}^n I_{n-1}^{n-1}\right)^{\frac{1}{2n-1}}} = \left(\frac{I_n}{I_{n-1}}\right)\left(\frac{n+1}{n}\right)^{\frac{n}{2n-1}}.$$

Since $1 > \frac{I_n}{I_{n-1}} > \frac{I_n}{I_{n-2}} = \frac{n-1}{n}$ we must have $x_n \to 1$ as $n \to \infty$. Furthermore we shall show that $x_n < x_{n-2}$ for all $n \geq 3$. Hence $x_n > 1$ for $n \geq 1$ so (u_n) is strictly increasing as claimed.

The fact that $x_n < x_{n-2}$ follows directly from Corollary 2.23 with the following observation.

$$\begin{aligned}
\frac{x_{n-2}}{x_n} &= \left(\frac{I_{n-2}\cdot I_{n-1}}{I_{n-3}\cdot I_n}\right)\left(\frac{n-1}{n-2}\right)^{\frac{n-2}{2n-5}}\left(\frac{n}{n+1}\right)^{\frac{n}{2n-1}} \\
&= \left(\frac{n}{n-1}\right)\left(\frac{n-1}{n-2}\right)^{\frac{n-2-(2n-5)}{2n-5}}\left(\frac{n}{n+1}\right)^{\frac{n}{2n-1}} \\
&= \left(\frac{n}{n-1}\right)\left[\left(\frac{n-1}{n-2}\right)^{\frac{n-3}{2n-5}}\left(\frac{n+1}{n}\right)^{\frac{n}{2n-1}}\right]^{-1}.
\end{aligned}$$

This completes the proof. □

3 The Problem for d-Dimensional Subspaces

So far we have only dealt with cross-sections corresponding to hyperplanes, however we may equally think about cross-sections of lower dimension. In this section we exhibit two upper bounds (due to Ball) on the volume of d-dimensional cross-sections of Q_n (for $d \leq n$) which depend only on n and d. In many cases one or other of these bounds is optimal and we present a conjecture on the best upper bound in the remaining cases.

The proof of each bound requires the Brascamp-Lieb inequality, which we first prove.

3.1 The Brascamp-Lieb Inequality

This result is a powerful generalisation of the inequalities of both Hölder and Young. It was originally due to Herm Brascamp and Elliott Lieb (see [7]). We only require a special case (the "geometric" version), but because of the importance of this result we prove it in the original generality (it can also be generalised beyond the original statement).

In 1997 an elegant proof was presented by Barthe in [1]. We shall now commence an exposition and translation of this statement and its proof (with a little simplification). One of interesting things about it is that it relies on some convex geometry, though the result itself is essentially analytic.

Theorem 3.1. *Fix $n \in \mathbb{Z}_{>0}$ also fix an integer $m \geq n$. Let $c_1, c_2, \ldots, c_m \in \mathbb{R}_{>0}$ such that $\sum_{i=1}^m c_i = n$. Suppose $v_1, v_2, \ldots, v_m$ are vectors in $\mathbb{R}^n$ and define*

$$\widetilde{D} = \left\{\frac{\det(\sum_{i=1}^m c_i\gamma_i v_i \otimes v_i)}{\prod_{i=1}^m \gamma_i^{c_i}} : \gamma_i > 0 \text{ for all } i\right\}.$$

We have used the following notation, for $a, b \in \mathbb{R}^n$, $a \otimes b$ is a linear map on $\mathbb{R}^n$ defined by

$a \otimes b(x) = \langle a, x\rangle b$. Now define the following sets:

$$\widetilde{F} = \left\{ \frac{\int_{\mathbb{R}^n} \prod_{i=1}^m (f_i(\langle v_i, x\rangle))^{c_i} dx}{\prod_{i=1}^m ||f_i||_1^{c_i}} : \textit{each } f_i \textit{ is non-negative and integrable} \right\},$$

$$\widetilde{F}_g = \left\{ \frac{\int_{\mathbb{R}^n} \prod_{i=1}^m \exp(-c_i\gamma_i\langle v_i, x\rangle^2) dx}{\prod_{i=1}^m (\int_{\mathbb{R}} \exp(-\gamma_i x^2) dx)^{c_i}} : \gamma_i > 0 \right\},$$

$$\widetilde{E}_g = \left\{ \frac{\int_{\mathbb{R}^n} \sup\left(\left\{\prod_{i=1}^m \exp(-c_i\gamma_i\theta_i^2) : x = \sum_{i=1}^m c_i\theta_i v_i\right\} \cup \{0\}\right) dx}{\prod_{i=1}^m (\int_{\mathbb{R}} \exp(-\gamma_i x^2) dx)^{c_i}} : \gamma_i > 0 \right\}.$$

Furthermore let $F = \sup(\widetilde{F})$, $F_g = \sup(\widetilde{F}_g)$, $E_g = \inf(\widetilde{E}_g)$ and $D = \inf(\widetilde{D})$. Then the conclusion of the theorem is that $F = F_g = \frac{1}{\sqrt{D}}$.

Note that $\widetilde{F}_g$ is the subset of $\widetilde{F}$ corresponding to the restriction that each f_i is a (centered) Gaussian. The "$\ldots \cup \{0\}$" terms appearing in the definitions of $\widetilde{E}_g$ is simply to prevent undesirable behaviour when the $\{v_i\}_{i=1}^m$ does not span $\mathbb{R}^n$, arising from the fact that $\sup \emptyset = -\infty$.

The proof has three parts. We first show that $F_g = \frac{1}{\sqrt{D}}$, then that $E_g \geq D \cdot F$ (it is clear by the remark after the theorem that $F \geq F_g$). Finally we shall show that $E_g = D \cdot F_g$ to complete the proof.

Lemma 3.2. *$F_g = \frac{1}{\sqrt{D}}$. If $D = 0$ then $F_g = \infty$.*

Proof . We show that for every $x \in \widetilde{F}_g$, $\frac{1}{x^2} \in \widetilde{D}$. The same argument also shows the converse (i.e. $x \in \widetilde{D} \Rightarrow \frac{1}{\sqrt{x}} \in \widetilde{F}_g$).

Fix $\gamma_1, \gamma_2, \ldots, \gamma_m > 0$ then the corresponding element of $\widetilde{F}_g$ is

$$y = \frac{\int_{\mathbb{R}^n} \prod_{i=1}^m \exp(-c_i\gamma_i\langle v_i, x\rangle^2) dx}{\prod_{i=1}^m (\int_{\mathbb{R}} \exp(-\gamma_i x^2) dx)^{c_i}}.$$

The denominator is the product of Gaussian integrals and so is easily calculated.

$$\prod_{i=1}^m \left(\int_{\mathbb{R}} \exp(-\gamma_i x^2) dx\right)^{c_i} = \prod_{i=1}^m \left(\frac{\pi}{\gamma_i}\right)^{\frac{c_i}{2}} = \sqrt{\frac{\pi^n}{\prod_{i=1}^m \gamma_i^{c_i}}}.$$

Furthermore the numerator can be calculated by a change of variables (using the fact that $Q(x) := \langle x, \left(\sum_{i=1}^m c_i\gamma_i v_i \otimes v_i\right)(x)\rangle$ is a positive-semidefinite quadratic form and thus has a real matrix-square-root).

$$\int_{\mathbb{R}^n} \prod_{i=1}^m \exp(-c_i\gamma_i\langle v_i, x\rangle^2) dx = \int_{\mathbb{R}^n} \exp(-Q(x)) dx$$
$$= \sqrt{\frac{\pi^n}{\det Q}} \int_{\mathbb{R}^n} \exp(-\pi x^2) dx = \sqrt{\frac{\pi^n}{\det \sum_{i=1}^m c_i\gamma_i v_i \otimes v_i}}.$$

Hence $1/y^2$ is the element in $\widetilde{D}$ corresponding to $\{\gamma_i\}_{i=1}^m$. The result follows easily. □

Lemma 3.3. *If $D \neq 0$, then $E_g \geq F \cdot D \geq F_g \cdot D$*

Proof . As mentioned above, the right-hand inequality is obvious from the definition. We prove the left-hand inequality using a change of variables. Let $f_1, \ldots, f_m, g_1, \ldots, g_m$ be strictly positive, integrable and continuous real valued functions on $\mathbb{R}$. We may make such assumptions about $\{g_i\}_{i=1}^m$ as these will later correspond to Gaussians, for $\{f_i\}_{i=1}^m$ (which later correspond to any non-negative integrable functions) we include an approximation argument at the end of the proof. For each i define T_i by the equation

$$\int_{-\infty}^{T_i(t)} g_i(x)dx = \frac{\int_{\mathbb{R}} g_i}{\int_{\mathbb{R}} f_i} \cdot \int_{-\infty}^{t} f_i(x)dx.$$

Observe that f_i and g_i have anti-derivatives which are increasing and differentiable. Hence each $T_i : \mathbb{R} \to \mathbb{R}$ is well defined and bijective. Moreover, the right-hand side of the above is differentiable and the anti-derivative of g_i has a differentiable inverse on $(0, \infty)$ (by the inverse function theorem). Precomposing by this inverse we see that T_i is differentiable. Indeed we have

$$T_i'(x)g_i(T_i(x)) = \frac{\int_{\mathbb{R}} g_i}{\int_{\mathbb{R}} f_i} f_i(x).$$

Moreover each T_i' is strictly positive since T_i is strictly increasing.

Now the change of variables will be given by $\Theta(x) = \sum_{i=1}^m c_i T_i(\langle x, v_i\rangle)v_i$. Notice that the directional derivatives are given by

$$\frac{\partial}{\partial x_j}\Theta(x) = \sum_{i=1}^m c_i T'(\langle x, v_i\rangle)v_i^j v_i.$$

Where $v_i = (v_i^1, \ldots, v_i^n)$. It follows that the differential is

$$d\Theta(x) = \sum_{i=1}^m c_i T'(\langle x, v_i\rangle)v_i \otimes v_i.$$

Notice that (as in the proof of Lemma 3.2) $d\Theta(x)$ corresponds to a positive-definite quadratic form on $\mathbb{R}^n$ for all x (because $D \neq 0$ so $\{v_i\}_{i=1}^m$ spans $\mathbb{R}^n$). Hence Θ is injective. Using the direct substitution $x = \Theta(y)$, then the change of variables, we get the following inequalities.

$$\begin{aligned}
&\int_{\mathbb{R}^n} \sup_{x=\sum_{i=1}^m c_i\theta_i v_i} \prod_{i=1}^m (g_i(\theta_i))^{c_i} dx \geq \int_{\Theta(\mathbb{R}^n)} \sup_{\Theta(y)=\sum_{i=1}^m c_i\theta_i v_i} \prod_{i=1}^m (g_i(\theta_i))^{c_i} d\Theta(y) \\
&\geq \int_{\mathbb{R}^n} \prod_{i=1}^m [g_i(T_i(\langle y, v_i\rangle))]^{c_i} \det\left(\sum_{i=1}^m c_i T'(\langle y, v_i\rangle)v_i \otimes v_i\right) dy \\
&\geq D\int_{\mathbb{R}^n} \prod_{i=1}^m [g_i(T_i(\langle y, v_i\rangle)) \cdot T_i'(\langle y, v_i\rangle)]^{c_i} dy \\
&= D\prod_{i=1}^m \left(\frac{\int_{\mathbb{R}} g_i}{\int_{\mathbb{R}} f_i}\right)^{c_i} \int_{\mathbb{R}^n} \prod_{i=1}^m [f_i(\langle y, v_i\rangle)]^{c_i} dy.
\end{aligned}$$

To complete the proof, divide by $\prod_{i=1}^m ||g_i||_1^{c_i}$ then take the infimum of the left-hand side with respect to $\{g_i\}_{i=1}^m$ (over all centered Gaussians) and the supremum of the right-hand side with respect to $\{f_i\}_{i=1}^m$ (over positive continuous, integrable functions).

All that remains is to justify the assumption that, for each i, f_i is strictly positive and continuous. We adapt an approach which is used in another paper by Barthe (see [2]).

We show that in calculating F we only need to consider positive and continuous functions. Let $\{f_i\}_{i=1}^m$,$\{v_i\}_{i=1}^m$ and $\{c_i\}_{i=1}^m$ be as they are in the definition of $\widetilde{F}$. It suffices to show that we may assume f_1 to be positive and continuous.

By monotone convergence theorem, we only need to consider functions which are bounded above by centred Gaussians. More precisely if $G(x) = e^{-\pi x^2}$ and f_1 is non-negative and integrable then $\zeta_k(x) := \min(f_1(x), kG(x))$ is increasing (with respect to $k \in \mathbb{Z}_{>0}$) and converges pointwise to f_1 as $k \to \infty$. Hence $||\zeta_k||_1 \to ||f_1||_1$ and moreover

$$\int_{\mathbb{R}^n} \zeta_k(\langle x, v_1 \rangle)^{c_1} \prod_{i=2}^{m} f_i(\langle x, v_i \rangle)^{c_i} dx \to \int_{\mathbb{R}^n} \prod_{i=1}^{m} f_i(\langle x, v_i \rangle)^{c_i} dx.$$

We can now assume that for some Gaussian $\widetilde{G}$, $f_1(x) \leq \widetilde{G}(x)$ for all $x \in \mathbb{R}$. For positivity and continuity, let $G_k(x) = kG(kx)$ and define $\eta_k(x) := \min(f_1 * G_k(x), \widetilde{G}(x))$. Then $\eta_k(x) > 0$ for all x (we may assume that $||f_1||_\infty > 0$) and is continuous (by the properties of convolutions). Furthermore it can be shown that $\lim_{k\to\infty} ||f_1 * G_k - f_1||_1 = 0$ (using Minkowski's integral inequality) and so $\lim_{k\to\infty} ||\eta_k - f_1||_1 = 0$. Hence, passing to a subsequence if necessary, we may assume that η_k converges pointwise to f_1 almost everywhere.

By dominated convergence theorem (we use $\widetilde{G}$ to find the dominating function), we now have

$$\int_{\mathbb{R}^n} \eta_k(\langle x, v_1 \rangle)^{c_1} \prod_{i=2}^{m} f_i(\langle x, v_i \rangle)^{c_i} dx \to \int_{\mathbb{R}^n} \prod_{i=1}^{m} f_i(\langle x, v_i \rangle)^{c_i} dx.$$

This is unless $\widetilde{G}(\langle x, v_1 \rangle)^{c_1} \prod_{i=2}^m f_i(\langle x, v_i \rangle)^{c_i}$ is not integrable, in which case $F = \infty$ anyway.

Hence without loss of generality, we may indeed assume that f_1 is positive and continuous.

□

The final part of Barthe's proof of the Brascamp Lieb inequality uses the following facts and definitions from convex geometry. The first two definitions can be found in [19], (pages 33 and 37 respectively) the third is a quantity discussed in [8].

Definition 3.4. *Let $C \subset \mathbb{R}^n$ be a symmetric convex body then:*

1. *the polar body of C is $C^\circ = \{x \in \mathbb{R}^n : \langle x, y \rangle \leq 1$ for all $y \in C\}$*
2. *the support function of C is the map $h(C) : \mathbb{R}^n \to \mathbb{R}$ given by $h(C)(x) = \sup_{y \in C} \langle x, y \rangle$.*
3. *the Mahler volume (or volume product) of C is defined to be $\mathcal{M}(C) = |C||C^\circ|$.*

Lemma 3.5. *If $C \subset \mathbb{R}^n$ is a symmetric convex body and $T : \mathbb{R}^n \to \mathbb{R}^n$ is a non-singular linear map, then:*

1. *C is uniquely determined by $h(C)$*
2. *$(T(C))^\circ = (T^{-1})^*(C^\circ)$*
3. *$\mathcal{M}(T(C)) = \mathcal{M}(C)$*
4. *If $C = \overline{B}_r(0)$ for $r > 0$ then $C^\circ = \overline{B}_{1/r}(0)$*
5. *C° is a symmetric convex body*
6. *$C^{\circ\circ} = C$ (Duality).*

Proof . **1)** Suppose that C and D are convex bodies and $h(C) = h(D)$. Now for any $y \in \mathbb{S}^{n-1}$, $h(C)(y)$ is the distance from the origin to a supporting hyperplane parallel to $y^{\perp}$. Hence $x \in C$ if and only if $\langle x, y\rangle \leq h(C)(y)$ for all $y \in \mathbb{S}^{n-1}$ (since a convex body is the intersection of all half-spaces which contain it). Similarly, $x \in D$ if and only if $\langle x, y\rangle \leq h(D)(y) = h(C)(y)$ for all $y \in \mathbb{S}^{n-1}$. So $C = D$ as required.
2) $x \in (T(C))^{\circ}$ if and only if $\langle x, T(y)\rangle \leq 1$ for all $y \in C$. An equivalent condition is of course $T^*(x) \in C^{\circ}$ i.e. $x \in (T^{-1})^*(C^{\circ})$.
3) By definition and part 2 we have $\mathcal{M}(TC) = |\det(T)||C||\det(T^{-1})||C^{\circ}| = \mathcal{M}(C)$ as required.
4) By part 2 it suffices to consider $C = \overline{B}_1(0)$. If $x \notin C$ then $\langle x, \frac{x}{||x||}\rangle = ||x|| > 1$ so $x \notin C^{\circ}$. The converse is by Cauchy-Schwarz. So $C^{\circ} = C$ as required.
5) Clearly C° is bounded as C has non-empty interior and hence contains a closed ball, $\overline{B}_{\varepsilon}(0)$ say, so $x \notin \overline{B}_{1/\varepsilon}(0)$ would imply that $\langle x, \varepsilon\frac{x}{||x||}\rangle > 1$ i.e. $x \notin C^{\circ}$. $C^{\circ} = (h(C))^{-1}([0,1])$ and so is closed by continuity of $h(C)$. Symmetry follows directly from the definition. C° has non-empty interior since C is bounded, say $C \subset B_r(0)$ for some $r > 0$ so if $||x|| < \frac{1}{r}$ then $h(C)(x) \leq 1$ i.e. $x \in C^{\circ}$. C° is convex, for if $\lambda \in [0,1]$ and $x, y \in C^{\circ}$ then for all $z \in C$, $\langle (1-\lambda)x + \lambda y, z\rangle \leq 1$ by linearity. Hence $(1-\lambda)x + \lambda y \in C^{\circ}$. So indeed C° is compact, symmetric and convex with non-empty interior i.e. a symmetric convex body as required.
6) It is clear from the definition that $C^{\circ\circ} \supseteq C$. For the converse suppose that $C^{\circ\circ} \supset C$ (i.e. strictly contained), then by taking the polar bodies[6] we would get $C^{\circ} \supset C^{\circ\circ\circ}$. However, just as $C^{\circ\circ} \supseteq C$, we have $C^{\circ\circ\circ} \supseteq C^{\circ}$ which is a contradiction. □

By Lemmas 3.2 and 3.3, to complete the proof of the Brascamp-Lieb inequality we only need to show the following.

Lemma 3.6. *If E_g and F_g are finite then $E_g \cdot F_g = 1$. If $F_g = \infty$ then the $E_g = 0$ and in any case E_g must be finite.*

Proof . Beginning with the last point, E_g is the infimum of a non-empty set of non-negative reals and so cannot be infinite.

We now find alternative expressions for the elements of $\widetilde{F}_g$ and $\widetilde{E}_g$. Let $\gamma_1, \ldots, \gamma_m > 0$ and for $x \in \mathbb{R}^n$ define $N(x) = \sqrt{\sum_{i=1}^m c_i\gamma_i\langle x, v_i\rangle^2}$. If $\{v_i\}_i$ spans $\mathbb{R}^n$ then N is clearly a norm (the triangle inequlity follows from the triangle inequality for the Euclidean norm). In this case the closed unit ball with respect to N is a symmetric convex body (in fact an ellipsoid).

In any case (even if $\{v_i\}_{i=1}^m$ does not span $\mathbb{R}^n$), the triangle inequality holds for N so the set $\mathcal{F} = \{x \in \mathbb{R}^n : N(x) \leq 1\}$ is convex. It is also clearly symmetric, therefore $\mathcal{F}$ is star-shaped. Hence, just as with spheres and balls, we have the relation $|\mathcal{F}| = \frac{|\partial\mathcal{F}|}{n}$ (or both sides are infinite). Therefore, considering the element of $\widetilde{F}_g$ corresponding to the choice of $\gamma_1, \ldots, \gamma_m$, we have

$$\begin{aligned}\frac{\int_{\mathbb{R}^n}\prod_{i=1}^m \exp(-c_i\gamma_i\langle x, v_i\rangle^2)dx}{\prod_{i=1}^m(\int_{\mathbb{R}}\exp(-\gamma_i x^2)dx)^{c_i}} = \frac{\int_{\mathbb{R}^n} e^{-N(x)^2}dx}{\prod_{i=1}^m\sqrt{\pi^{c_i}\gamma_i^{-c_i}}} &= \frac{\int_0^\infty |\partial\mathcal{F}|r^{n-1}e^{-r^2}dr \cdot \prod_{i=1}^m \gamma_i^{\frac{c_i}{2}}}{\pi^{\frac{n}{2}}} \\ &= \frac{|\mathcal{F}|n\int_0^\infty \tilde{r}^{\frac{n-2}{2}}e^{-\tilde{r}}d\tilde{r} \cdot \prod_{i=1}^m \gamma_i^{\frac{c_i}{2}}}{2\pi^{\frac{n}{2}}} \\ &= |\mathcal{F}|\frac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}} \cdot \prod_{i=1}^m \gamma_i^{\frac{c_i}{2}} = \frac{|\mathcal{F}|}{|B_1(0)|} \cdot \prod_{i=1}^m \gamma_i^{\frac{c_i}{2}}.\end{aligned}$$

[6] For symmetric convex bodies A and B, the fact that $A \supset B \Rightarrow A^{\circ} \subset B^{\circ}$ follows from the definiton and convexity.

Note that the introduction of $\tilde{r}$ was just a change of variable.

If $\{v_i\}_i$ does not span $\mathbb{R}^n$ then $|\mathcal{F}| = \infty$ so, by the above calculation, we have $F_g = \infty$. In this case we also have (from the definition) that $E_g = 0$, as required. Hence we may now assume that $\{v_i\}_i^{\perp} = \emptyset$.

Consider the element of $\widetilde{E}_g$ corresponding to the constants $\frac{1}{\gamma_1}, \dots, \frac{1}{\gamma_m}$. Again we define a norm on $\mathbb{R}^n$, this time given by

$$M(x) = \inf \left\{ \sqrt{\sum_{i=1}^m c_i \frac{\theta_i^2}{\gamma_i}} : x = \sum_{i=1}^m c_i \theta_i v_i \right\}.$$

with closed unit ball $\mathcal{E}$ (a symmetric convex body). Then, as above, we have the following.

$$\begin{aligned} \frac{\int_{\mathbb{R}^n} \sup \left\{ \prod_{i=1}^m \exp(-c_i \frac{\theta_i^2}{\gamma_i}) : x = \sum_{i=1}^m c_i \theta_i v_i \right\} dx}{\prod_{i=1}^m \left(\int_{\mathbb{R}} \exp(-\frac{x^2}{\gamma_i}) dx \right)^{c_i}} &= \frac{\int_{\mathbb{R}^n} e^{-M(x)^2} dx}{\prod_{i=1}^m \sqrt{(\pi\gamma_i)^{c_i}}} \\ &= \frac{|\mathcal{E}|}{|B_1(0)|} \cdot \prod_{i=1}^m \gamma_i^{\frac{-c_i}{2}}. \end{aligned}$$

Thus, as we have a 1-to-1 correspondance between $\widetilde{E}_g$ and $\widetilde{F}_g$ arising from the choice of $\gamma_1, \dots, \gamma_m$ it suffices to show that $|\mathcal{F}||\mathcal{E}| = |B_1(0)|^2$ for any such choice.

Here we can apply the previously discussed geometry. In fact we shall see that $\mathcal{E}$ is the polar body of $\mathcal{F}$. This immediately gives the required result by Lemma 3.5 and the fact that $\mathcal{F}$ is an ellipsoid (i.e. the image of $B_1(0)$ under a non-singular linear operator).

Considering the support function of $\mathcal{F}^\circ$, we see that $h(\mathcal{F}^\circ)(x) = N(x)$. This follows from duality. More precisely, if $x \in \mathbb{R}^n \backslash \{0\}$ then $\frac{x}{N(x)} \in \partial\mathcal{F} = \partial\mathcal{F}^{\circ\circ}$ so $h(\mathcal{F}^\circ)(\frac{x}{N(x)}) = 1$ which is enough by homogeneity of $h(\mathcal{F}^\circ)$.

Now by Cauchy-Schwarz applied to $\sum_i \sqrt{c_i\gamma_i}\langle x, v_i\rangle$ and $\sum_i \sqrt{\frac{c_i}{\gamma_i}}\theta_i$, we have that

$$N(x) = \sqrt{\sum_{i=1}^m c_i \gamma_i \langle x, v_i\rangle^2} \geq \sup_{\sum_i \frac{c_i}{\gamma_i}\theta_i^2 \leq 1} \sum_{i=1}^m c_i \theta_i \langle x, v_i \rangle.$$

In particular, setting $\theta_i = \frac{\gamma_i\langle x, v_i\rangle}{N(x)}$ gives equality. Now applying the definition of $\mathcal{E}$, we see that this is exactly $N(x) = \sup_{y\in\mathcal{E}}\langle x, y\rangle = h(\mathcal{E})(x)$. Therefore, since support functions uniquely determine symmetric convex bodies, we have $\mathcal{F}^\circ = \mathcal{E}$. as required. □

We have proved Brascamp-Lieb inequality in its original generality, however we shall only use the following special case (often called the geometric Brascamp-Lieb inequality).

Corollary 3.7. *Fix $m \geq n$ and suppose $\{c_i\}_{i=1}^m \subset \mathbb{R}_{>0}$ and $\{v_i\}_{i=1}^m \subset \mathbb{R}^n$ with*

$$\sum_{i=1}^m c_i = n \tag{3.8}$$

and

$$\sum_{i=1}^m c_i v_i \otimes v_i = I_n. \tag{3.9}$$

Then for any non-negative integrable functions $f_1, f_2, \ldots, f_m$ on $\mathbb{R}$, we have

$$\int_{\mathbb{R}^n} \prod_{i=1}^{m} [f_i(\langle x, v_i \rangle)]^{c_i} dx \leq \prod_{i=1}^{m} \left[\int_{\mathbb{R}} f_i(x) dx \right]^{c_i}.$$

Conditions (3.8) and (3.9) may be refered to as the Fritz John conditions. This corollary follows from the following lemma which is proved in [12], but we shall not prove it here.

Lemma 3.10. *With $\{v_i\}_{i=1}^m$ and $\{c_i\}_{i=1}^m$ as in Corollary 3.7 and D from the statement of the Brascamp-Lieb inequality, we have $D \geq 1$.*

We now switch our attention back to the problem of bounding the cross-sectional volumes of cubes.

3.2 Upper Bounds in the d-Dimensional Case

In this section we discuss two upper bounds on the d-dimensional cross-sectional volumes of $Q_n \subset \mathbb{R}^n$ where d may be less than $n-1$.

The following theorem is the first of these bounds. With the use of the Brascamp-Lieb inequality the proof is quite straightforward. For this result we follow the proof in [11].

Theorem 3.11. *If $S \subseteq \mathbb{R}^n$ is a subspace with dimension $d > 0$ then*

$$|S \cap Q_n| \leq \left(\frac{n}{d}\right)^{\frac{d}{2}},$$

moreover this bound is attained if $d|n$.

Proof . Fix a d-dimensional subspace S, then if P denotes orthogonal projection onto S, we see that $P(e_1), \ldots, P(e_n)$ forms a basis of S. Moreover, for $x \in S$, $\langle x, e_i \rangle = \langle x, P(e_i) \rangle$ for each i. Therefore $S \cap Q_n$ is the set $\{x \in \mathbb{R}^n : |\langle x, P(e_i) \rangle| \leq \frac{1}{2} \text{ for } i = 1, \ldots, n\}$.

Clearly the restriction of P to S is the identity and is also given by

$$P = \sum_{i=1}^{n} e_i \otimes P(e_i) = \sum_{i=1}^{n} P(e_i) \otimes P(e_i).$$

We may assume that for some m with $d \leq m \leq n$ we have $P(e_i) \neq 0 \Leftrightarrow i \leq m$. Then given an isometry $\psi : S \to \mathbb{R}^d$, the vectors $v_i = \frac{\psi(P(e_i))}{||P(e_i)||}$ and constants $c_i = ||P(e_i)||^2 > 0$ $(i = 1, \ldots, m)$ satisfy (3.9). Furthermore,

$$\psi(S \cap Q_n) = \left\{ x \in \mathbb{R}^d : |\langle x, v_i \rangle| \leq \frac{1}{2||P(e_i)||} = \frac{1}{2\sqrt{c_i}} \text{ for } i = 1, \ldots, m \right\}.$$

Notice also that $\{c_i\}_{i=1}^m$ satisfy (3.8), i.e.

$$\sum_{i=1}^{m} c_i = \sum_{i=1}^{n} ||P(e_i)||^2 = d.$$

It is easy to see this by writing each $P(e_i)$ in terms of an orthonormal basis of S and using the fact that $\langle x, e_i \rangle = \langle x, P(e_i) \rangle$ for all $x \in S$.

Therefore, letting $f_i = \mathbb{1}_{[-1/(2\sqrt{c_i}),1/(2\sqrt{c_i})]}$ for each i we conclude, using the Brascamp-Lieb inequality that

$$|S \cap Q_n| = \int_{\mathbb{R}^d} \prod_{i=1}^{m} [f_i(\langle x, v_i \rangle)]^{c_i} dx \leq \prod_{i=1}^{m} \left[\int_{\mathbb{R}} f_i(x) dx \right]^{c_i} = \prod_{i=1}^{m} c_i^{-\frac{c_i}{2}}.$$

We shall shortly sketch a proof that a product of the form $\prod_{i=1}^{n} x_i^{x_i}$ is minimised (subject to each x_i being positive and $\sum_{i=1}^{n} x_i = r$ being constant) when $x_1 = x_2 = \ldots = x_n$. Hence $\prod_{i=1}^{m} c_i^{-\frac{c_i}{2}} \leq \left(\frac{m}{d}\right)^{\frac{d}{2}} \leq \left(\frac{n}{d}\right)^{\frac{d}{2}}$ as required.

We also claimed that if n is an integer multiple of d then this bound is the best possible. Indeed in this case, let $k = \frac{n}{d}$ and consider the subspace $H \subset \mathbb{R}^n$ spanned by

$$\tilde{e}_i := e_{ik+1} + e_{ik+2} + \ldots + e_{(i+1)k},$$

for $i = 0, 1, \ldots, d-1$. Thinking of $\{\tilde{e}_i\}_{i=0}^{d-1}$ as an orthogonal basis for $\mathbb{R}^d$ with $||\tilde{e}_i|| = \sqrt{k}$ it is not difficult to see that $H \cap Q_n$ is isometric to the set $[-\sqrt{k}/2, \sqrt{k}/2]^d$ in $\mathbb{R}^d$. Hence $|H \cap Q_n| = (\sqrt{k})^d = \left(\frac{n}{d}\right)^{\frac{d}{2}}$. □

As mentioned in the previous proof, we sketch a proof of the following lemma.

Lemma 3.12. *If $x_1, \ldots, x_n \in \mathbb{R}_{>0}$ and $\sum_{i=1}^{n} x_i = r$ then $\prod_{i=1}^{n} x_i^{x_i} \geq \left(\frac{r}{n}\right)^r$.*

Proof (Sketch). Let $\Delta \subset \mathbb{R}^n$ be the (open) simplex defined by $\sum_i x_i = r$ (and each x_i positive). Then let $f : \Delta \to \mathbb{R}_{>0}$ be defined by $f(x) = \prod_{i=1}^{n} x_i^{x_i}$. The tangent plane of Δ is spanned by vectors of the form $e_i - e_j$.

Given $x \in \Delta$ we sketch an algorithm to construct a path in Δ, from x to $(\frac{r}{n}, \ldots, \frac{r}{n})$ on which f is always decreasing. We argue by induction. Fix $k \in \mathbb{Z}_{>0}$, $k < n$ and suppose that $x_1 = x_2 = \ldots = x_k$ then we shall construct a path from x to $\tilde{x}$ where $\tilde{x}_1 = \ldots = \tilde{x}_k = \tilde{x}_{k+1}$ and $f(\tilde{x}) \leq f(x)$. Let $\alpha = \frac{1}{k}$ and assume that $x_1 \geq x_{k+1}$ (if not then replace t with $-t$ where appropriate in the following argument). Differentiating f with respect to the direction $v = (-\alpha, \ldots, -\alpha, 1, 0 \ldots, 0)$ (k terms are $-\alpha$) we get

$$\left.\frac{d}{dt}\right|_{t=s} f(x + tv) = \log\left(\frac{x_{k+1} + s}{x_1 - \alpha s}\right) f(x + sv).$$

This is non-positive for $s \leq \tilde{s} := \frac{x_1 - x_{k+1}}{1+\alpha}$. In particular, $f(x) \geq f(x + \tilde{s}v)$ and letting $\tilde{x} = x + \tilde{s}v$, we clearly have $\tilde{x}_1 = \ldots = \tilde{x}_{k+1}$ as required. □

We shall now discuss the second upper bound of this section. This is a theorem from the same paper (i.e. [11]) which gives the best upper bound in the case where $d \geq \frac{n}{2}$. The proof of optimality is again constructive and will be discussed later as part of a more general set of examples.

The proof is essentially a generalisation of the method used in the hyperplane case (Theorem 2.12), so we may only sketch some parts.

Theorem 3.13. *As before, let $S \subset \mathbb{R}^n$ be a d-dimensional subspace ($d > 0$) then*

$$|S \cap Q_n| \leq (\sqrt{2})^{n-d}.$$

Proof (Sketch). Assume for induction that the result holds for any cross-section of Q_{n-1}. As before there are two parts to the proof.

Case 1: We first consider the case where $S^{\perp}$ contains a direction that is sufficiently close to a standard basis direction in $\mathbb{R}^n$. The case $d = n$ is obvious so we also assume that $d \leq n-1$.

Suppose that there exists a unit vector $v \in S^{\perp}$ with some component larger than $\frac{1}{\sqrt{2}}$. Without loss of generality we assume that this is the first component, i.e. $v_1 \geq \frac{1}{\sqrt{2}}$. Denote by C_n the cylinder $\mathbb{R} \times [-\frac{1}{2}, \frac{1}{2}]^{n-1}$ and let P denote orthogonal projection onto $e_1^{\perp}$. Then (as $P\big|_S$ is non-singular) we have, by induction that $P(S \cap C_n) \leq (\sqrt{2})^{n-1-d}$. In other words, $P(S \cap C_n)$ is a subspace of $e_1^{\perp}$ with dimension d.

Let R be the orthogonal projection onto $v^{\perp}$, then $||P \circ R(e_1)||/||R(e_1)|| = v_1$ (see the calculations in the proof of Theorem 2.12). Moreover, since we can find an orthogonal basis of $v^{\perp}$ contained in $\{R(e_1)\} \cup e_1^{\perp}$, we have that for any $x \in v^{\perp}$, $||P(x)||/||x|| \geq v_1$. In particular, by considering an orthogonal basis of $S \subseteq v^{\perp}$ with at least $d-1$ vectors in $e_1^{\perp}$ we see that $|P(S \cap C_n)| \geq |S \cap C_n| v_1$. Hence $|S \cap Q_n| \leq (\sqrt{2})^{n-d}$ as required.

Case 2: Suppose that for any unit vector in $S^{\perp}$, each component is at most $\frac{1}{\sqrt{2}}$ (in absolute value). Let T be the orthogonal projection onto $S^{\perp}$ and define $c_i = ||T(e_i)||^2$ and $w_i = \frac{T(e_i)}{\sqrt{c_i}}$ for $i = 1, \ldots, n$. We may assume (by induction) that for each i, $c_i > 0$, moreover, by the hypothesese of the case, $c_i < \frac{1}{2}$. Notice that as in the proof of Theorem 3.11 we have constructed vectors and constants satisfying the Fritz John conditions (on $S^{\perp}$).

Once again, consider independent random variables $X_1, \ldots, X_n$, each uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$ and let $X = (X_1, \ldots, X_n)$. Observe that the probability density function of $T(X)$ is given by $f : S^{\perp} \to \mathbb{R}_{>0}$ where $f(x) = |(S + x) \cap Q_n|$. By considering the characteristic function of $T(X)$ (i.e. the Fourier transform of f); using the properties of expectaions and applying the Fourier inversion formula, we obtain the following (c.f. the hyperplane case).

$$|S \cap Q_n| = f(0) = \frac{1}{\pi^{n-d}} \int_{S^{\perp}} \prod_{i=1}^{n} \operatorname{sinc}(\sqrt{c_i}\langle w_i, \xi\rangle) d\xi.$$

Thus, taking absolute values and applying the geometric Brascamp-Lieb inequality, we obtain

$$\begin{aligned}
|S \cap Q_n| &\leq \frac{1}{\pi^{n-d}} \prod_{i=1}^{n} \left(\int_{\mathbb{R}} |\operatorname{sinc}(x\sqrt{c_i})|^{\frac{1}{c_i}} dx \right)^{c_i} \\
&= \frac{1}{\pi^{n-d}} \prod_{i=1}^{n} \left(\frac{1}{\sqrt{c_i}} \int_{\mathbb{R}} |\operatorname{sinc}(x)|^{\frac{1}{c_i}} dx \right)^{c_i} \\
&\leq \frac{1}{\pi^{n-d}} \prod_{i=1}^{n} \left(\frac{\pi\sqrt{2c_i}}{\sqrt{c_i}} \right)^{c_i} = (\sqrt{2})^{n-d}.
\end{aligned}$$

The last inequality is just Lemma 2.9, applied n times, with $p = \frac{1}{c_i} \geq 2$. This is what we wanted to show. □

3.3 A Conjecture on the Best Upper Bound in all Cases

We have seen some good upper bounds on the d-dimensional cross-sectional volumes of n-dimensional cubes which are indeed attained in many cases. However, by considering the form of the worst examples when the optimal bounds are known, we can find examples of large cross-sections in every case and conjecture that these are indeed the worst. We first discuss such examples.

Fix $d \le n$ and let r be the remainder from the division of n by d. Also let $k = \frac{n-r}{d}$, then define $v_1, \ldots, v_d \in \mathbb{R}^n$ as follows.

$$\begin{aligned} v_{i+1} &= e_{ki+1} + e_{ki+2} + \ldots + e_{k(i+1)} && \text{for } 0 \le i < d-r \\ v_{d-r+i} &= e_{(d-r)k+(k+1)i+1} + e_{(d-r)k+(k+1)i+2} + \ldots \\ &\quad \ldots + e_{(d-r)k+(k+1)(i+1)} && \text{for } 0 \le i < r \end{aligned}$$

It is easy to check that the $v_1, \ldots, v_d$ are orthogonal and hence span a d-dimensional subspace, call it S, of $\mathbb{R}^n$. Moreover, the intersection $S \cap Q_n$ is the set

$$S \cap Q_n = \left\{ \sum_{i=1}^{d} \alpha_i v_i : -\frac{1}{2} \le \alpha_i \le \frac{1}{2} \text{ for each } i \right\}.$$

Hence

$$|S \cap Q_n| = \prod_{i=1}^{d} |v_i| = \left(\frac{n-r}{d}\right)^{\frac{d-r}{2}} \left(\frac{n-r+d}{d}\right)^{\frac{r}{2}}.$$

One can check that this agrees with the result of Theorem 3.11 when $d \big| n$ ($r = 0$) and with bound in Theorem 3.13 when $2d \ge n$ ($n - r = d$). In particular these examples verify our claims about optimality in Theorem 3.13.

After considering this, the following conjecture seems natural.

Conjecture 3.14. *Let S be a d-dimensional subspace of $\mathbb{R}^n$ and r as above, then*

$$|S \cap Q_n| \le \left(\frac{n-r}{d}\right)^{\frac{d-r}{2}} \left(\frac{n-r+d}{d}\right)^{\frac{r}{2}}.$$

We shall now discuss one of the apparent difficulties in finding a proof for this result. As a first attack on the problem, one might look at the application of the Brascamp-Lieb inequality in Theorem 3.11 as this result looks similar. Furthermore, we may ask whether choosing $c_1, \ldots, c_n$ to satisfy the Fritz John conditions in that proof is too restrictive. After all, the full statement of Brascamp-Lieb gives the following (using the notation of Theorem 3.11).

$$\int_{\mathbb{R}^d} \prod_{i=1}^{n} [f_i(\langle x, v_i \rangle)]^{c_i} dx \le \prod_{i=1}^{n} \left(\int_{\mathbb{R}} f_i \right)^{c_i} \left(\sup \left\{ \frac{\prod_{i=1}^{n} \gamma_i^{c_i}}{\det(\sum_{i=1}^{n} c_i \gamma_i v_i \otimes v_i)} : \gamma_i > 0 \right\} \right)^{\frac{1}{2}}.$$

Here we have only assumed that $\sum_{i=1}^{n} c_i = d$ (and $c_i > 0$ for each i). Now the left-hand side does not depend on the $\{c_i\}_{i=1}^{n}$ as each f_i is an indicator functions. Hence we would like to take the infimum of the right-hand side with respect to $\{c_i\}_{i=1}^{n}$.

Unfortunately this approach will probably not work. Indeed Stefán Valdimarsson has recently proved (see [21]) that 1 is the best "Brascamp-Lieb constant" i.e. the optimal constant in the Brascamp-Lieb[7] inequality. More precisely, for $\{c_i\}_{i=1}^{n}$ and $\{v_i\}_{i=1}^{n} \subset \mathbb{R}^d$ as above, we have

$$\left(\sup \left\{ \frac{\prod_{i=1}^{n} \gamma_i^{c_i}}{\det(\sum_{i=1}^{n} c_i \gamma_i v_i \otimes v_i)} : \gamma_i > 0 \right\} \right)^{\frac{1}{2}} \ge 1,$$

with equality only in the geometric case. Therefore the only hope left for this strategy is if, for a suitable choice of $\{c_i\}_{i=1}^{n}$, $\prod_{i=1}^{n} \left(\int_{\mathbb{R}} f_i\right)^{c_i}$ can be reduced without substantially increasing the Brascamp-Lieb constant.

[7]Valdimarsson actually proves this result for a more general version of the Brascamp-Lieb inequality.

In many cases this is not possible. Indeed, suppose we can find a d-dimensional subspace $S \subset \mathbb{R}^n$ such that $||P(e_i)|| = ||P(e_1)||$ for $2 \leq i \leq n$, where P is the orthogonal projection onto S. In this case, constructing $\{f_i\}_{i=1}^n$ as in Theorem 3.11, we have $f_1 = f_2 = \ldots = f_n$ and therefore

$$\prod_{i=1}^{n} \left(\int_{\mathbb{R}} f_i(x)dx \right)^{c_i} = \left(\int_{\mathbb{R}} f_1(x)dx \right)^{d} = \left(\frac{n}{d} \right)^{\frac{d}{2}}.$$

The final equality holds because as before, the constants $\left(\int_{\mathbb{R}} f_i\right)^{-2}$ satisfy the Fritz John conditions with corresponding vectors $\{P(e_i)/||P(e_i)||\}_{i=1}^n$. It would follow that we could not improve on Theorem 3.11 with this method.

It is easy to check that such subspaces exist when d divides n (and consequently when $n-d$ divides n) by explicit constructions. For example in the case of a hyperplane ($n-d=1$) we can take the subspace $(1,1,\ldots,1)^{\perp}$. However these cases happen to be examples where either Theorem 3.11 or Theorem 3.13 already provide optimal bounds.

It does not seem clear whether such subspaces exist in general, but at least in one case of interest, they can be constructed[8]. Namely for $d = 2$. This follows easily from the fact that the Fritz John conditions[9] are not only necessary but also sufficient for a set of vectors to be the image under projection of an orthogonal set of vectors in some superspace. This follows from a simple result about sets of vectors in Hilbert spaces which was proved for example, by Steinberg (see [20]). To find subspaces with the required properties, it therefore suffices to find vectors with equal lengths that satisfy the Fritz John conditions. This can obviously be done when $d = 2$, for example consider the normal directions of the edges of a regular n-gon.

In summary of this discussion, it seems that if Conjecture 3.14 holds then a proof should take a different approach to the proof of Theorem 3.11 (at least for $d = 2$). However, at present it does not seem clear what this should be.

This concludes our investigation of the cross-sectional volumes of cubes. We now discuss two other interesting problems from convex geometry which involve cubes in some way.

4 Some Other Interesting Problems Related to Cubes

4.1 "Universality" of Cross-Sections

If one were to write a computer programme to generate images of cross-sections of cubes, one would see that a remarkable range of polytopes can be attained (for example see Figures 3 and 4). This is no coincidence. Indeed, we now discuss a result proved in 1967 by Epifanov (see [4]) that any polytope occurs as a cross-section of a cube. The proof begins with the following lemma.

Lemma 4.1. *Fix $n \in \mathbb{Z}_{>0}$ and real numbers $b_{i,j}$ for all $1 \leq i < j \leq n+1$. Then there exist vectors $v_1, \ldots, v_{n+1} \in \mathbb{R}^n$ with $v_i, \ldots, v_n$ linearly independant and $\langle v_i, v_j \rangle = b_{i,j}$ for $i < j$.*

Proof (Sketch). If $n = 1$ this is obvious. Otherwise we argue by induction. Assume we can find such vectors $\{v_i\}_{i=1}^{n+1}$ with respect to an n-dimensional subspace, S, of $\mathbb{R}^{n+1}$. Then, choosing a unit vector $u \in S^{\perp}$ and replacing v_{n+1} with $v_{n+1} + u$, we obtain a basis for $\mathbb{R}^{n+1}$. There is then a unique choice of v_{n+2} that has the prescribed inner product with $v_1, \ldots, v_n$ and $v_{n+1} + u$. $\square$

[8] Thanks to Professor Keith Ball for pointing this out.

[9] When we discuss the Fritz John conditions and only mention the vectors, $\{u_i\}_{i=1}^n$ say, it is meant that $c_i = ||u_i||^2$ and $v_i = u_i/\sqrt{c_i}$ satisfy the (3.8) and (3.9).

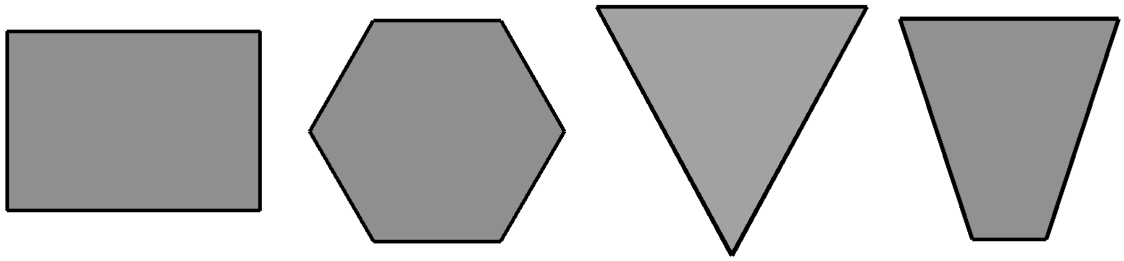

Figure 3: A few hyperplane cross-sections of Q_3 (not to scale) corresponding to the planes $\{x : \langle x, v\rangle = a\}$ where v and a are (from left to right): $v = (1, 1, 0)$, $a = 0$; $v = (1, 1, 1)$, $a = 0$; $v = (1, 1, 1)$, $a = 0.5$ and $v = (1, 1, 0.5)$, $a = 0.5$.

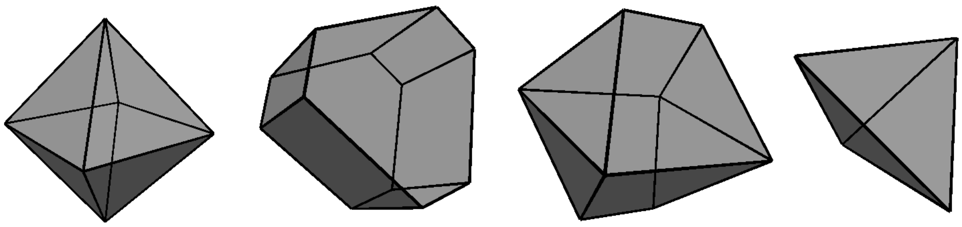

Figure 4: A few hyperplane cross-sections of Q_4 (projected into $\mathbb{R}^3$ and not to scale) corresponding to the hyperplanes $\{x : \langle x, v\rangle = a\}$ where v and a are (from left to right): $v = (1, 1, 1, 1)$, $a = 0$; $v = (1, 1, 1, 0)$, $a = 0.2$; $v = (1, 1, 0.5, 0.5)$, $a = 0$ and $v = (1, 1, 1, 1)$, $a = 1$.

The following result and proof are based on Epifanov's ideas but we have made the argument more explicit. Epifanov mentions cubes but essentially proves something more general, which means that the dimension of the cube constructed is not optimal (see the discussion after the proof). Though we do not improve on this method, we aim to make the construction of the cube clearer.

Proposition 4.2. *Suppose $P \subset \mathbb{R}^n$ is a polytope with non-empty interior and m $(n-1)$–dimensional faces. Then there exists an n-dimensional affine space $S \subset \mathbb{R}^{n+m-1}$ and $r > 0$ such that P is isometric to $S \cap rQ_{n+m-1}$.*

Proof . Let $T \subset \mathbb{R}^{n+m-1}$ be an n-dimensional subspace and let $\psi : \mathbb{R}^n \to T$ be an isometry such that $0 \in \psi(\mathrm{int}(P))$. Also let $\tilde{\nu}_1, \dots, \tilde{\nu}_m \in \mathbb{R}^n$ be unit vectors in the normal directions of the $(n-1)$–dimensional faces of P. Define $\nu_i = \psi(\tilde{\nu}_i) - \psi(0)$ for each $1 \le i \le m$ (i.e. the image of these directions under ψ).

By the lemma, there are vectors $k_1, \dots, k_m \in T^\perp$ such that $\langle k_i, k_j\rangle = -\langle \nu_i, \nu_j\rangle$ for all $1 \le i < j \le m$. In this case one can easily check that the vectors $\{k_i + \nu_i\}_{i=1}^m$ are mutually orthogonal. Moreover $T \cap (k_i + \nu_i)^\perp = T \cap \nu_i^\perp$ is parallel to the face of $\psi(P)$ corresponding to ν_i.

For $1 \le i \le m$ let $h_i = \nu_i + k_i$ and extend to an orthogonal basis $\{h_i\}_{i=1}^{n+m-1}$ of $\mathbb{R}^{n+m-1}$. Without loss of generality assume that there exists M with $m \le M \le n + m - 1$ for which $h_i \in T^\perp$ if and only if $i > M$. Furthermore, for each i and $x \in \mathbb{R}$, define $H_i(x) \subset \mathbb{R}^{n+m-1}$ to be the halfspace, containing 0, induced by the hyperplane $h_i^\perp + xh_i$. Then there exist non-zero constants $a_1, \dots, a_M$ such that $\psi(P) = T \cap (\bigcap_{i=1}^m H_i(a_i)) = T \cap \left(\bigcap_{i=1}^M H_i(a_i)\right)$. We may assume that $(h_i^\perp + a_ih_i) \cap \psi(P) \neq \emptyset$ i.e. each a_i is minimal (in absolute value) with respect to the previous equalities (this only affects the choice of $a_{m+1}, \ldots, a_M$).

Let $r = \max\{\max_{x\in\psi(P)}\langle x, h_i\rangle - \min_{x\in\psi(P)}\langle x, h_i\rangle : 1 \le i \le M\}$, i.e. the maximum width of $\psi(P)$ in any of the directions h_i. We now also see that $\psi(P) \subseteq T \cap \left(\bigcap_{i=1}^{n+m-1} H_i(b_i)\right)$ where $b_i = a_i\left(1 - \frac{r}{|a_i|}\right)$. For those hyperplanes parallel to T, define $a_i = r/2$ and $b_i = -r/2$ for $M < i \le n+m-1$. Then $Q = \bigcap_{i=1}^{n+m-1}(H_i(a_i) \cap H_i(b_i))$ is a cube of side length r with $T \cap Q = \psi(P)$. Thus there is an isometry ϕ of $\mathbb{R}^{n+m-1}$ so that $rQ_{n+m-1} = \phi(Q)$ and if $S := \phi(T)$ we get $\phi \circ \psi(P) = rQ_{n+m-1} \cap S$ as required. $\square$

We now briefly highlight a shortcoming of this result and ponder possible improvements. In the proof we were only really considering cross-sections of cones generated by intersecting half-spaces. Therefore the cube we found will in general be in a higher dimension than necessary. For example, a hexagon can be generated as a 2-dimensional cross-section of a 3-cube. However, the above construction would instead find a hexagonal cross-section of a 7-dimensional cube. The heuristic reason for this is that a cross-section of a cube may intersect faces that are opposite to one another, however in the cone there are no pairs of parallel faces.

This leads us to the question of finding the minimum dimension for a cube with a given polytope as one of its cross-sections. By adapting the above method, it seems likely that we can find an $(n-1+N)$–dimensional cuboid with the correct cross-section, where N is the number of normal directions to faces of the polytope, i.e. the maximum number non-parallel faces. However, this is still not optimal (consider the hexagon again) and moreover, it is not clear whether something similar can be done for cubes.

This concludes our discussion of Epifanov's result. We now arrive at the final part of this essay, in which we introduce a well known open problem.

4.2 The Blaschke-Santaló Inequality and the Mahler Conjecture

In this section we continue the discussion of the Mahler volume, (see Definition 3.4). In particular we consider the bodies that extremise the Mahler volume. We shall prove the Blaschke-Santaló inequality which shows that ellipsoids have maximal Mahler volume and discuss a conjecture on the bodies which minimise this quantity.

Recall that the Mahler volume is invariant under linear transformations. Hence when we say for example, that ellipsoids have maximal Mahler volume, this really does mean that the Mahler volume of any (centred) ellipsoid is maximal in the set of all symmetric convex bodies.

Theorem 4.3. *(Blaschke-Santaló) For any symmetric convex body C,*

$$\mathcal{M}(C) \le \mathcal{M}(B_1(0)).$$

It is also known that equality is only acheived in the case of an ellipsoid. This is discussed in [15], where more general statements for non-symmetric convex bodies are also proved.

For the proof, we will follow the approach of Meyer and Pajor (see [14] or [15]). Prior to this however, we need to discuss an operation known as Steiner symmetrization, the intuition for which is to make a given convex body "more spherical".

To begin with recall that the Hausdorff metric on $\mathcal{K}_n$ (the set of non-empty compact subsets of $\mathbb{R}^n$) is given by

$$d(A, B) = \inf\{\varepsilon > 0 : A \subset B + B_\varepsilon(0) \text{ and } B \subset A + B_\varepsilon(0)\}$$

When we discuss convergence of sets, this is the metric we will use unless stated otherwise.

Definition 4.4. *For a bounded measurable set $S \subseteq \mathbb{R}^n$ and a unit vector $v \in \mathbb{S}^{n-1}$ the Steiner symmetrization* $\mathrm{st}_v(S)$ *is given by*

$$\mathrm{st}_v(S) = \bigcup_{x \in P(S)} \left\{ x + v\left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right] : \lambda = |(x + v\mathbb{R}) \cap S| \right\},$$

where P is the orthogonal projection onto $v^\perp$. In particular suppose S is convex. We can think of S as the union of line-segments in the direction v (one for each $x \in P(S)$) then $\mathrm{st}_v(S)$ is the union of translated copies of those segments such that each is bisected by $v^\perp$.

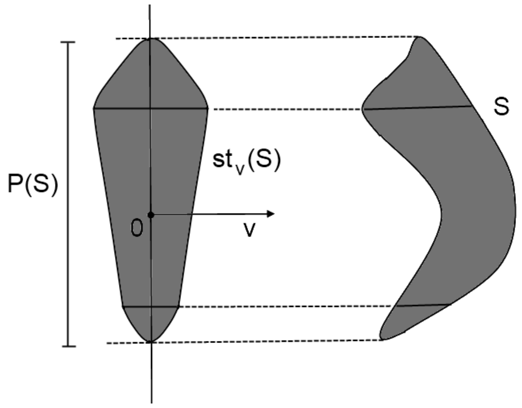

Figure 5: Steiner symmetrization.

We will need the following properties of the Steiner symmetrization (based on pages 169-173 of [17] and [14]).

Lemma 4.5. *Let C, D be symmetric convex bodies in $\mathbb{R}^n$ such that $C \subseteq D$, and let $v \in \mathbb{S}^{n-1}$ then we have the following:*

1. *If C is a ball then* $\mathrm{st}_v(C) = C$
2. $\mathrm{st}_v(C) \subseteq \mathrm{st}_v(D)$
3. $\mathrm{st}_v(C)$ *is a symmetric convex body*
4. $|\mathrm{st}_v(C)| = |C|$
5. $|\mathrm{st}_v(C)^\circ| \geq |C^\circ|$

Proof . Let K be the image of C under the orthogonal projection onto $v^\perp$.
1 & 2) are obvious from the definition.
3) (Sketch) Since C is bounded, (1) and (2) together imply that $\mathrm{st}_v(C)$ is bounded. Let $f : K \to \mathbb{R}$ be given by $x \mapsto |(x + v\mathbb{R}) \cap C|$. Closedness follows from the continuity of f (see Corollary 2.6 and think of $\partial \mathrm{st}_v(C)$ locally as a graph of $\frac{1}{2}f$). For symmetry, observe that K is symmetric and $f(-x) = f(x)$ by symmetry of C.

In two dimensions, with C a trapezium and v parallel to two of the sides of C, $\mathrm{st}_v(C)$ is still a trapezium, hence convex. To deduce convexity in general, let $x, y \in \mathrm{st}_v(C)$, then the convex hull $H = \mathrm{conv}\{[(y + v\mathbb{R}) \cap C] \cup [(x + v\mathbb{R}) \cap C]\} \subset C$ is a trapezium with parallel sides in the direction v. Thus $\mathrm{st}_v(C) \supset \mathrm{st}_v(H)$, the latter being a convex set containing x and y as required (see Figure 6).

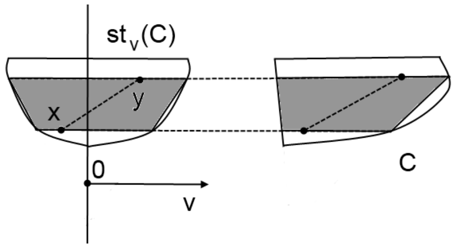

Figure 6: Steiner symmetrization preserves convexity.

Finally, $\mathrm{st}_v(C)$ has non-empty interior since $0 \in \mathrm{int}(C)$ thus $0 \in \mathrm{int}(\mathrm{st}_v(C))$ by (1).
4) Preservation of volume is a consequence of Fubini's Theorem.
5) (from [14]) Assume, without loss of generality, that $v = e_n$ then we can rewrite $\mathrm{st}_v(C)^\circ$ and C° as follows:

$$\begin{aligned} C^\circ = &\left\{(Y, y) \in \mathbb{R}^{n-1} \times \mathbb{R} : \langle Y, X\rangle + xy \leq 1 \text{ for all } X \in K \right. \\ &\left. \text{and } x \in C \cap (X + v\mathbb{R})\right\}, \\ \mathrm{st}_v(C)^\circ = &\left\{(Y, y) \in \mathbb{R}^{n-1} \times \mathbb{R} : \langle Y, X\rangle + y\frac{x_1 - x_2}{2} \leq 1 \text{ for all } X \in K, \right. \\ &\left. x_1, x_2 \in C \cap (X + v\mathbb{R})\right\}. \end{aligned}$$

The second equation follows from the observation that

$$\frac{1}{2}\{x_1 - x_2 : x_1, x_2 \in C \cap (X + v\mathbb{R})\} = \mathrm{st}_v(C) \cap (X + v\mathbb{R}).$$

For $t \in \mathbb{R}$, let H_t denote the hyperplane $v^\perp + tv = \mathbb{R}^{n-1} \times \{t\}$. Then from the equations above, we deduce that, for each $t \in \mathbb{R}$

$$H_t \cap \mathrm{st}_v(C)^\circ \supseteq \frac{1}{2}(H_t \cap C^\circ + H_{-t} \cap C^\circ) = \frac{1}{2}(H_t \cap C^\circ - H_t \cap C^\circ).$$

The equality follows from the symmetry of C°. Now an application of the Brünn-Minkowski inequality shows that the right hand side of the above has volume at least $|H_t \cap C^\circ|$. Therefore we have

$$|\mathrm{st}_v(C)^\circ| = \int_{\mathbb{R}} H_t \cap \mathrm{st}_v(C)^\circ dt \geq \int_{\mathbb{R}} H_t \cap C^\circ dt = |C^\circ|.$$

This is the required result. □

Combining (4) and (5) we see that Steiner symmetrization does not decrease the Mahler volume of a convex body. This is the first of two observations which together give Theorem 4.3 as a corollary. The second is the following (which can be found in [17] pages 172-173).

Lemma 4.6. *Let C be a symmetric convex body in $\mathbb{R}^n$. Then there exists a sequence of vectors $v_1, v_2, \ldots \in \mathbb{R}^n$ such that the sequence of convex bodies $C_k = \mathrm{st}_{v_k} \circ \mathrm{st}_{v_{k-1}} \circ \ldots \circ \mathrm{st}_{v_1}(C)$ converges to a ball $B_r(0)$ (where r is determined by $|C|$).*

This will be proved shortly. We shall also use without proof the following consequence of the Arzelà-Ascoli theorem. For a proof, see [17], pages 85-88.

Lemma 4.7. *Let $C_1, C_2, \ldots$ be a sequence of compact convex sets in $\mathbb{R}^n$, each contained in some ball $B_r(0)$ (i.e. a bounded sequence in some sense). Then the sequence has an accumulation point (with respect to the Hausdorff metric) which is a compact convex set.*

Corollary 4.8. *With $C_1, C_2 \ldots$ as above, suppose additionally that each C_i is a symmetric convex body and $|C_i| = |C_1|$. Then the sequence has an accumulation point which is also a symmetric convex body.*

Proof . Let K be an accumulation point from the lemma. To show symmetry, suppose for contradiction that there is $\varepsilon \in (0,1)$ and $x \in \partial K$ with $\tilde{x} := (\varepsilon - 1)x \in \partial K$. Let $v \in \mathbb{S}^{n-1}$ and $t \in \mathbb{R}$ be such that $v^\perp + tv$ is the supporting hyperplane at $\tilde{x}$. Then there exists $N > 0$ such that $k > N \Rightarrow C_k \subset \{y \in \mathbb{R}^n : \langle y, v\rangle \leq t + \frac{\varepsilon t}{2(1-\varepsilon)}\}$. Now by symmetry of C_k we also have $C_k \subset \{y \in \mathbb{R}^n : \langle y, -v\rangle \leq t + \frac{\varepsilon t}{2(1-\varepsilon)}\}$, but $\langle x, -v\rangle = \langle \frac{\tilde{x}}{1-\varepsilon}, v\rangle = \frac{t}{1-\varepsilon} = t + \frac{\varepsilon t}{1-\varepsilon}$. Hence $x \notin C_k + B_{\frac{\varepsilon t}{2(1-\varepsilon)}}(0)$ which is a contadiction, as C_k was supposed to converge to K.

For non-emptiness of the interior of K we suppose that K contains no ball around the origin. Passing to a subsequence if necessary, we may assume that for any $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{Z}_{>0}$ such that if $i > N_\varepsilon$ then C_i does not contain $B_\varepsilon(0)$. By symmetry and convexity, it would follow that, for $i > N_\varepsilon$, each C_i is contained in some (symmetric) plank of width at most 2ε (i.e. there is a unit vector v such that $|\langle y, v\rangle| < \varepsilon$ for all $y \in C_i$). Moreover, since $\{C_i\}_{i=1}^\infty$ is uniformly bounded, we have $i > N_\varepsilon \Rightarrow |C_i| \leq D\varepsilon$ for some D independent of i and ε. Now ε was arbitrary so we cannot have $|C_i| = |C_1| > 0$ for all i, which is the required contradiction. □

We can now prove Lemma 4.6. The proof of Theorem 4.3 follows directly.

Proof (of Lemma 4.6). For any non-empty compact set $D \in \mathcal{K}_n$, define

$$\rho(D) = \inf\{r > 0 : D \subset B_r(0)\}.$$

Then with C as in the statement, let F be the set of convex bodies which can be obtained from C by applying finitely many Steiner symmetrizations and define $\sigma = \inf\{\rho(D) : D \in F\}$.

Since C is bounded (and by (1) & (2) of Lemma 4.5) we have that any sequences in F are bounded (in the sense of Lemma 4.7). Thus by the corollary, there is a symmetric convex body K and a sequence $C_1, C_2, \ldots \in F$ such that $C_k \to K$ and $\rho(C_k) \to \sigma$. By the definition of the Hausdorff metric it is easy to check that ρ is continuous on $\mathcal{K}_n$ and so $\rho(K) = \sigma$.

It now suffices to show that $K = \overline{B}_\sigma(0)$. By definition of ρ, we have $K \subseteq \overline{B}_\sigma(0)$. Suppose for contradiction that equality does not hold. By convexity of K we must have $K \subset \overline{B}_\sigma(0) \backslash S$ where S is a spherical cap (the intersection of $\overline{B}_\sigma(0)$ and a halfspace). By compactness of $\partial B_\sigma(0)$ there are finitely many unit vectors $v_1, \ldots, v_m$ so that $\partial B_\sigma(0)$ is covered by the images of $S \cap \partial B$ under reflection in the hyperplanes $v_i^\perp$.

From the definition of st_{v_i}, one can check that if $T \subset \partial B_\sigma(0) \backslash K$ and T' is the reflection of T in $v_i^\perp$ then $\mathrm{st}_{v_i}(K) \cap (T \cup T') = \emptyset$ (see Figure 7). It follows that $\mathrm{st}_{v_m} \circ \ldots \circ \mathrm{st}_{v_1}(K) \cap \partial B_\sigma(0) = \emptyset$. Hence $\rho(K) = \sigma$ is not minimal in F, this is the required contradiction. □

Proof (of Theorem 4.3). By Lemma 4.6, given a convex body C, there is a sequence of convex bodies $C = C_1, C_2, C_3, \ldots$, converging to a ball, with $C_{k+1} = \mathrm{st}_{v_k}(C_k)$ for some v_k. It is fairly straightforward to show that the Mahler volume is continuous with respect to the Hausdorff metric (at least when dealling with symmetric convex bodies since neighbourhoods are given by dilations). Therefore $\mathcal{M}(C_k) \to \mathcal{M}(B_\sigma(0)) = \mathcal{M}(B_1(0))$. Furthermore, the sequence $\mathcal{M}(C_k)$ is increasing as a consequence of Lemma 4.5. It follows that $\mathcal{M}(C) \leq \mathcal{M}(B_1(0))$, as required. □

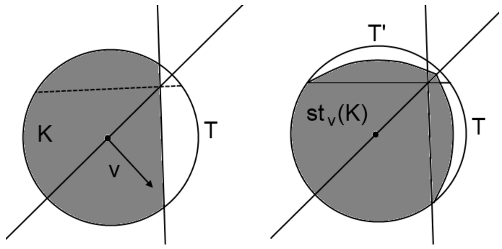

Figure 7: $\mathrm{st}_v(K)$ contains less of the boundary of $B_\sigma(0)$ than K.

We conclude this section with a brief discussion of a problem which is currently open, known as the Mahler conjecture. We state it here in its original form (due to Mahler).

Conjecture 4.9. *(Mahler) For a symmetric convex body $C \subset \mathbb{R}^n$, $\mathcal{M}(C) \geq \mathcal{M}(Q_n)$. That is to say, the minimum Mahler volume of a symmetric convex body is attained by the cube (and therefore by its polar body- known as the cross-polytope).*

The (unit) cross-polytope is the unit ball in the ℓ^1 norm on $\mathbb{R}^n$ and it is easy to check that this is the polar body of the cube (with side length 2).

We shall not discuss attempts to prove this conjecture in detail, but instead point out a few results. For example, it is known that the conjecture holds true in the case of zonoids (i.e. sets that are the limits of finite sums of line-segments). For a proof of this see [24].

One recent paper about this problem (see [3]) shows that cubes are indeed local minimisers of the Mahler volume. Here the word "local" is in the sense of the Banach-Mazur distance, which in the aforementioned paper is defined by

$$\mathrm{d}_{BM}(K, L) = \inf\left\{\frac{b}{a} : aK \subseteq T(L) \subseteq bK \text{ for some } T \in \mathrm{GL}(n)\right\},$$

for symmetric convex bodies K and L. Note that in this form, the Banach-Mazur distance is not a metric but heuristically behaves like the minimum Hausdorff distance between images of the given sets under linear maps. It therefore seems like a natural notion of distance when considering the Mahler volume.

Another recent developement uses a generalisation of the Gauss curvature to show that a convex body with minimum Mahler volume must have a curvature of zero on almost all of its boundary (see [22]). This at least points to the minimisers being polytopes.

References

[1] F.Barthe, *Inégalités de Brascamp-Lieb et convexité*, Comptes Rendus de l'Académie des Sciences **324** (1997), no. 8, 885–888.

[2] ———, *Optimal Young's Inequality and its Converse: a Simple Proof*, Geometric and Functional Analysis **8** (1998), no. 2, 234–242.

[3] F.Nazarov, F.Petrov, D.Ryabogin, and A.Zvavitch, *A Remark on the Mahler Conjecture: Local Minimality of the Unit Cube*, Duke Mathematical Journal **154** (2010), no. 3, 419–430.

[4] G.V.Epifanov, *Universality of the Sections of Cubes*, Mathematical Notes **2** (1967), no. 1, 540–541.

[5] G.Zhang, *A Positive Solution to the Busemann-Petty Problem in* $\mathbb{R}^4$, Annals of Mathematics **149** (1999), 535–543.

[6] H.Busemann and C.M.Petty, *Problems on Convex Bodies*, Mathematica Scandinavica **4** (1956), 88–94.

[7] H.J.Brascamp and E.H.Lieb, *Best Constants in Young's Inequality, Its Converse, and Its Generalization to More than Three Functions*, Advances In Mathematics **20** (1976), 151–173.

[8] J.Bourgain and V.D.Milman, *New volume ratio properties for convex symmetric bodies in* $\mathbb{R}^n$, Inventiones Mathematicae **88** (1987), no. 2, 319–340.

[9] K.Ball, *Cube Slicing in* $\mathbb{R}^n$, Proceedings of the AMS **97** (1986), no. 3, 465–473.

[10] ———, *Some Remarks on the Geometry of Convex Sets*, Geometric Aspects of Functional Analysis **1317** (1988), 224–231.

[11] ———, *Volumes of Sections of Cubes and Related Problems*, Geometric Aspects of Functional Analysis **1376** (1989), 251–260.

[12] ———, *Convex Geometry and Functional Analysis*, Handbook of the Geometry of Banach Spaces, vol. 1, 2001.

[13] Alois Kufner, Oldřich John, and Svatopluk Fučík, *Function Spaces*, 1977.

[14] M.Meyer and A.Pajor, *On Santaló's Inequality*, Geometric Aspects of Functional Analysis **1376** (1989), 261–263.

[15] ———, *On the Blaschke-Santaló Inequality*, Archiv der Mathematik **55** (1990), 82–93.

[16] N.Batir, *Inequalities for the Gamma Function*, Archiv Der Mathematik **91** (2008), no. 6, 554–563.

[17] P.Gruber, *Convex and Discrete Geometry*, Grundlehren der Mathematischen Wissenschaften, vol. 336, Springer, 2007.

[18] R.J.Gardener, *The Brunn-Minkowski Inequality*, Bulletin of the AMS **39** (2002), no. 3, 355–405.

[19] R.Schneider, *Convex Bodies: The Brunn-Minkowski Theory*, Cambridge University Press, 1993.

[20] R.Steinberg, *Note on a Theorem of Hadwiger*, Pacific Journal of Mathematics **6** (1956), no. 4, 775–777.

[21] S.I.Valdimarsson, *Geometric Brascamp-Lieb Has the Optimal Best Constant*, Journal of Geometric Analysis **21** (2011), 1036–1043.

[22] S.Reisner and C.Schütt and E.Werner, *Mahler's Conjecture and Curvature*, International Mathematics Research Notices **2012** (2012), no. 1, 1–16.

[23] Y.D.Burago and V.A.Zalgaller, *Geometric Inequalities*, 1988.

[24] Y.Gordon, M.Meyer, and S.Reisner, *Zonoids With Minimal Volume-Product – A New Proof*, Proceedings of the American Mathematical Society **104** (1988), no. 1, 273–276.